

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE VETERINARIA



## TESIS DOCTORAL

Establecimiento de la Microbiota en Recién Nacidos: distintos Escenarios y Metodologías

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Claudia Saralegui Díez

DIRECTORES

Rosa María del Campo Moreno  
Juan de Dios Caballero Pérez

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE VETERINARIA**



**TESIS DOCTORAL**

Establecimiento de la Microbiota en Recién Nacidos: Distintos Escenarios y Metodologías

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

Claudia Saralegui Díez

DIRECTORES

Rosa María del Campo Moreno  
Juan de Dios Caballero Pérez



**UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE VETERINARIA**

Programa de Doctorado en Veterinaria



**Establecimiento de la Microbiota  
en Recién Nacidos:  
Distintos Escenarios y Metodologías**

MEMORIA PARA OPTAR AL GRADO DE DOCTORA  
PRESENTADA POR

**Claudia Saralegui Díez**

BAJO LA DIRECCIÓN DE

**Rosa María del Campo Moreno  
Juan de Dios Caballero Pérez**

**Madrid, 2023**





## AGRADECIMIENTOS

Para empezar, me gustaría recordar a los profesores y profesoras de Biología y Geología que tuve durante el instituto, los cuales sembraron en mí la curiosidad y la fascinación por todo lo relacionado con la vida, nuestro cuerpo y nuestro planeta. Más adelante, durante la carrera universitaria, me sentí de nuevo motivada al escuchar a los miembros del Departamento de Microbiología que, allá por 2014-2015, me introdujeron en el increíble mundo del mutualismo microbiológico, la simbiosis y todo lo relacionado con la microbiota humana. Son de esos *clicks* de los que no te olvidas en la vida.

En 2016, llego a la guinda del pastel: Rosa. Guiada por mi interés sobre la microbiota y con unas prácticas externas y TFM que realizar, encontré a Rosa del Campo, la cual, desde el minuto uno, me abrió los brazos con una espontaneidad y profesionalidad que dudo mucho vuelva a encontrarme en la vida. Después, Rosa me dio la oportunidad de trabajar en aquello sobre lo que había estudiado, y realizar esta tesis doctoral con un contrato laboral y en un ambiente amable y pedagógico. Algo que, desgraciadamente, no ocurre a menudo. No he conocido a una persona más comprometida con su trabajo y con más ganas de colaborar y compartir que ella. No sólo eso, sino que se ha acabado convirtiendo en una gran amiga y maestra de la vida, levantándome el ánimo y sacando fuerzas cuando no veía nada claro. Estoy infinitamente agradecida por ello y sé que siempre recordaré con cariño esta etapa de tal crecimiento personal y académico.

A Juande, porque en ti siempre he visto un entusiasmo profundo por la Microbiología y un gran respeto por las personas con las que trabajas. Eres un referente para todo aquel que te conozca. Gracias por acceder a codirigir mi tesis junto a Rosa. Tengo al mejor equipo que podría desear.

A todas las personas con las que compartí momentos en el Ramón y Cajal. A Marta Cobo, por enseñarme a dar los primeros pasos en el labo. A Manu, del que he aprendido tanto. A Ana, por convertirte en una auténtica amiga. A Pati, por tan buena compañía y charlas. A Cris, por ser ejemplo de todo y descubrirme a las depredadoras. A Fernando, por haberme dado la oportunidad de trabajar junto a ti y demostrarme que la ciencia es emocionante. A Conchi, por esa última etapa en el labo de infinita ayuda. A todas las personas que han venido a hacer estancias o prácticas al labo: yo también he aprendido mucho intentando enseñar. Al equipo de Concha Gil de Farmacia-UCM, y en especial a Carmen, que me acogieron cuando aún no sabíamos nada de pandemias

mundiales para hacer una estancia en su departamento. En general, a todas las personas del Servicio de Genética, Microbiología y las UCAs que me han ayudado y resuelto dudas o, simplemente, con las que he comido y charlado durante todos esos años. Sola no puedes; con amigas y compañeras: sí.

Gracias también a todo el personal que ha hecho posible la recogida de muestras y, sobre todo, el seguimiento y cuidado de los pacientes. Gracias a las Unidades de Fibrosis Quística de Sevilla, Málaga, el Doce de Octubre y el Ramón y Cajal. Gracias a las UCIN, en especial la de La Paz, que saca adelante a tantas criaturas vulnerables. Gracias a todas las personas que apuestan por la Sanidad y la Investigación Públicas.

A mi red de amistades salvavidas, en especial a Belén (*Sevi*), Auri, Clara S., Mariama, Gabi, Serli y a todas mis ratas peleonas (Ana, Álvaro, Hernán, Marina, Ran, Silvia, Javi, Alex, Irene, Paola). A aquellas con las que descubrí la magia campamental, jugué tantos años de niña y adulta y hasta comencé una asociación de OyTL. A las personas que más me acompañaron y me hicieron reír durante la Universidad: Paula, Paloma, Javi y Juan. Este proceso ha sido largo y la gente va y viene, aunque la tesis se quede. Agradezco también a todas aquellas que estuvieron conmigo, aunque fuera un cachito del camino.

A mi familia. A mi madre y mi padre, por su cariño y apoyo y por la educación que me han dado. A mi hermano Pablo, por haberme descubierto tantas cosas y ser ejemplo. Aunque hayamos estado distanciados físicamente, os he notado siempre cerquita. Os quiero mucho. A toda la familia colmenareña, me encanta vuestro calor y vuestras ganas de hacer piña. A toda la familia Saralegui, en especial a Juan Carlos y Lorenzo. Me habría encantado compartirla con vosotros.

A Chupi. La suerte nunca se olvida. Gracias por tu cariño, tu humor, tu paciencia y tu ternura. El mundo necesita a más como tú. Me siento tremendamente afortunada por tenerte a mi lado.

La vida, en estos años, era eso que pasaba mientras realizaba la tesis. Y ha sido tanto...Me siento agradecida también por los colectivos de los que he podido formar parte, por las aventuras, los viajes, los retos, las formaciones y la terapia que he recibido. La tesis acaba; la vida y la pasión por la Microbiología, siguen.

*A toda la red de personas que me ha acompañado  
en este proceso.*



<b>LISTADO DE ABREVIATURAS</b>	<b>10</b>
<b>ÍNDICE DE FIGURAS</b>	<b>11</b>
<b>ÍNDICE DE TABLAS</b>	<b>15</b>
<b>RESUMEN</b>	<b>17</b>
<b>INTRODUCCIÓN</b>	<b>23</b>
1. MICROBIOMA HUMANO .....	25
1.1. Definiciones y conceptos básicos.....	25
1.2. Funciones de la microbiota .....	27
1.3. Relación con la salud y la enfermedad.....	29
2. HERRAMIENTAS DE LA INVESTIGACIÓN TRASLACIONAL EN MICROBIOLOGÍA.....	32
2.1. Técnicas de estudio de brotes infecciosos.....	32
2.1.1. Técnicas convencionales	32
2.1.2. Secuenciación de genoma completo	33
2.2. Técnicas de estudio del microbioma.....	33
2.2.1. Análisis metagenómico	33
2.2.1.1. <i>Shotgun sequencing</i>	35
2.2.1.2. Secuenciación de amplicones	36
2.2.2. Culturómica	37
2.2.3. Análisis metatranscriptómico	38
2.2.4. Análisis metaproteómico	38
2.2.5. Análisis metabolómico	39
3. ETAPAS DETERMINANTES EN LA ADQUISICIÓN DEL MICROBIOMA INTESTINAL.....	40
3.1. Embarazo .....	41
3.2. Parto.....	42
3.3. Lactancia y etapa infantil .....	42
3.4. Perspectivas futuras .....	44
4. PATRONES DIFERENCIALES EN EL MICROBIOMA NEONATAL .....	45
4.1. Prematuridad.....	45
4.1.1. Definición, factores de riesgo y de asociación	45
4.1.2. Microbioma y morbilidad en recién nacidos prematuros	46
4.2. Fibrosis quística .....	50
4.2.1. Fisiopatología, sintomatología y tratamientos	50

4.2.2.	Primeros años de vida con FQ	53
4.2.3.	Microbioma respiratorio e intestinal en la FQ	54
<b>5.</b>	<b>INTERVENCIÓN TERAPÉUTICA BASADA EN LAS DINÁMICAS ECOLÓGICAS DEL MICROBIOMA.....</b>	<b>58</b>
5.1.	Transferencia de microbiota fecal (TMF).....	58
5.2.	Terapia con bacteriófagos.....	59
5.2.1.	Papel en la función del microbioma	59
5.2.2.	Aplicaciones clínicas	59
5.3.	Bacterias depredadoras.....	60
5.3.1.	Descubrimiento y características	60
5.3.2.	Aplicaciones industriales y agrícolas	61
5.3.3.	Aplicaciones clínicas	62
	<b>HIPÓTESIS Y OBJETIVOS</b>	<b>65</b>
1.	OBJETIVOS GENERALES.....	68
2.	OBJETIVOS ESPECÍFICOS.....	69
2.1.	Técnicas de estudio del microbioma:.....	69
2.2.	Población pediátrica con FQ:.....	69
2.3.	Población neonatal prematura ingresada en UCIN:.....	70
2.4.	Colecciones de cepas patógenas aisladas de estos pacientes y de otras cohortes:.....	70
	<b>CAPÍTULO 1. ABORDAJE DEL ESTABLECIMIENTO DE LA MICROBIOTA INTESTINAL EN PACIENTES CON FIBROSIS QUÍSTICA MEDIANTE DIFERENTES TÉCNICAS ÓMICAS</b>	<b>71</b>
I.	ANTECEDENTES.....	73
II.	JUSTIFICACIÓN Y OBJETIVOS.....	75
III.	MATERIALES Y MÉTODOS.....	77
1.	Diseño metodológico.....	77
2.	Participantes, recogida de muestras e historiales médicos.....	77
3.	Extracción de ADN.....	80
4.	Secuenciación de amplicones de 16S ARNr.....	80
4.1.	Análisis bioinformático	81
4.2.	Procesamiento de lecturas	81
4.3.	Base de datos para la asignación taxonómica	81
5.	Análisis Metaproteómico.....	82
5.1.	Obtención de la fracción proteica bacteriana	82
5.2.	Espectrometría de masas en tándem (MS/MS)	82

5.3.	Análisis de los datos	83
6.	Acuerdos de asignación taxonómica y cuantificación microbiana.....	83
IV.	RESULTADOS .....	85
1.	Inclusión de muestras.....	85
2.	Concordancia estadística de los resultados de la secuenciación de amplicones y la metaproteómica .....	85
3.	Estudio funcional de la microbiota de esta cohorte mediante metaproteómica .....	93
V.	DISCUSIÓN .....	98
VI.	DIVULGACIÓN .....	102
	<b>CAPÍTULO 2. MONITORIZACIÓN DE LA IMPLANTACIÓN DE LA MICROBIOTA INTESTINAL Y NASOFARÍNGEA EN RECIÉN NACIDOS CON FIBROSIS QUÍSTICA</b>	<b>103</b>
I.	ANTECEDENTES .....	105
II.	JUSTIFICACIÓN Y OBJETIVOS .....	108
III.	MATERIALES Y MÉTODOS .....	110
1.	Diseño metodológico .....	110
2.	Extracción de ADN.....	110
3.	Secuenciación de amplicones de 16S ARNr' .....	111
3.1.	Análisis bioinformático	111
3.2.	Procesamiento de lecturas	112
3.3.	Base de datos para la asignación taxonómica	112
4.	Cultivo de muestras, tipificación mediante electroforesis en gel de campo pulsado (PFGE) y estadística de la culturómica .....	112
5.	Identificación de especies de <i>Streptococcus</i> .....	113
6.	Tipado genético de <i>S. aureus</i> y <i>P. aeruginosa</i> mediante MLST.....	114
7.	Análisis estadísticos .....	114
IV.	RESULTADOS .....	117
1.	Características de la cohorte .....	117
2.	Evolución sintomática, tratamientos y principales colonizaciones patogénicas .....	117
3.	Cultivo y tipado de las principales especies bacterianas .....	120
4.	Establecimiento de la microbiota intestinal y respiratoria analizada mediante secuenciación	128
4.1.	Evolución de los principales géneros bacterianos y la alfa diversidad de la microbiota intestinal y respiratoria	131
4.2.	Concordancia entre los análisis dependientes de cultivo y la secuenciación	139

V.	DISCUSIÓN .....	142
VI.	DIVULGACIÓN .....	146
<b>CAPÍTULO 3. INFLUENCIA DE UN BROTE DE <i>Serratia marcescens</i> EN EL ESTABLECIMIENTO DE LA MICROBIOTA INTESTINAL DE RECIÉN NACIDOS PREMATUROS Y ANÁLISIS FILOGENÉTICO DE AISLADOS NOSOCOMIALES CONTEMPORÁNEOS E HISTÓRICOS DE DICHA ESPECIE</b>		<b>147</b>
I.	ANTECEDENTES .....	149
II.	JUSTIFICACIÓN Y OBJETIVOS .....	151
III.	MATERIALES Y MÉTODOS .....	153
1.	Diseño metodológico .....	153
2.	Participantes, recogida de muestras e inclusión de aislados .....	153
3.	Extracción de ADN.....	154
4.	Secuenciación de amplicones del 16S ARNr.....	154
4.1.	Análisis bioinformático	155
4.2.	Procesamiento de las lecturas	155
4.3.	Asignación taxonómica	155
4.4.	Cálculos de diversidad y pruebas estadísticas	155
4.5.	Herramientas para detectar biomarcadores característicos	156
5.	Cultivo de muestras, tipificación mediante electroforesis en gel de campo pulsado (PFGE) y estadística de la culturómica .....	156
6.	Secuenciación de genoma completo (WGS), ensamblaje y anotación.....	157
7.	Análisis filogenético de cepas de estudio y del repositorio global de <i>S. marcescens</i> mediante cgMLST y AcCNET.....	158
8.	Análisis de susceptibilidad antibiótica .....	158
IV.	RESULTADOS .....	160
1.	Características de los neonatos prematuros.....	160
2.	Establecimiento de la microbiota intestinal mediante secuenciación de amplicones.....	161
3.	Caracterización de los aislados cultivables del grupo del brote .....	166
4.	Estudio extendido de las cepas causantes del brote contemporáneo y el histórico. ....	169
V.	DISCUSIÓN .....	178
VI.	DIVULGACIÓN .....	186
<b>CAPÍTULO 4. BACTERIAS DEPREDADORAS COMO CONTROL DE <i>Pseudomonas aeruginosa</i> y <i>Serratia marcescens</i></b>		<b>187</b>

I.	ANTECEDENTES .....	189
II.	JUSTIFICACIÓN Y OBJETIVOS .....	192
III.	MATERIALES Y MÉTODOS .....	194
	1. Cepas y condiciones de crecimiento .....	194
	2. Experimentos de depredación .....	195
	3. Análisis estadísticos .....	196
IV.	RESULTADOS .....	197
	1. La fuente clínica de la presa influye en la actividad de depredación <i>de B. bacteriovorus</i> 109J 197	
	2. Influencia de la variación genética de la depredadora .....	197
V.	DISCUSIÓN .....	204
VI.	DIVULGACIÓN .....	208
	<b>CONCLUSIONES</b>	<b>209</b>
	<b>BIBLIOGRAFÍA</b>	<b>213</b>
	<b>ANEXOS</b>	<b>247</b>

## LISTADO DE ABREVIATURAS

- **16S ARNr**: gen del ADN ribosomal que codifica la subunidad 16S del ribosoma.
- **AGCC**: ácidos grasos de cadena corta.
- **ASV**: variantes de secuencia de amplicón (*amplicon sequence variant*).
- **ATP**: adenosín trifosfato.
- **BALO**: *Bdellovibrio* y organismos similares (*Bdellovibrio and like organisms*).
- **BHI**: infusión cerebro-corazón (*brain-heart infusion*).
- **COG**: grupo de genes ortólogos (*cluster of orthologous genes*).
- **DGGE**: electroforesis en gradiente desnaturizante (*denaturing gradient gel electrophoresis*).
- **DO**: densidad óptica.
- **EPOC**: enfermedad pulmonar obstructiva crónica.
- **FQ**: fibrosis quística.
- **GC**: guanina-citosina.
- **I-PR**: rango de depredación intraespecífica (*intraspecies predation range*).
- **IS**: secuencias de inserción.
- **LEfSe**: grado de impacto del análisis discriminante lineal (*linear discriminant analysis Effect Size*).
- **LME**: lactancia materna exclusiva.
- **LN**: lavados nasofaríngeos.
- **mEq**: mili-equivalentes.
- **MLST**: *multilocus sequence typing*.
- **MP**: metaproteómica.
- **NCBI**: centro nacional de información biotecnológica (*National Center for Biotechnology Information*).
- **PBS**: tampón fosfato salino (*phosphate buffered saline*).
- **PE**: eficiencia de depredación (*predation efficiency*).
- **PFGE**: electroforesis en gel de campo pulsado (*Pulsed-Field Gel Electrophoresis*).
- **PHA**: polihidroxialcanoato.
- **RDP**: Proyecto de base de datos ribosomal (*Ribosomal Database Project*).
- **RQ**: rango intercuartílico.
- **SA**: secuenciación de amplicones.
- **SDS**: dodecilsulfato sódico.
- **TMF**: transferencia de microbiota fecal.
- **UCIN**: unidad de cuidados intensivos neonatales.
- **WGS**: secuenciación de genoma completo (*whole genome sequencing*).

## ÍNDICE DE FIGURAS

- **Figura 1.** Dinámicas ecológicas de la microbiota intestinal. (p. 26)
- **Figura 2.** Vías de comunicación entre la microbiota y el cerebro. (p. 28)
- **Figura 3.** Conexiones en el eje bidireccional intestino-pulmón. (p. 30)
- **Figura 4.** Métodos de estudio del microbioma humano. (p. 35)
- **Figura 5.** Factores que influyen en la composición del microbioma neonatal y bacterias asociadas a las diferentes variables. (p. 41)
- **Figura 6.** Causas de muerte en menores de cinco años en todo el mundo (2016). (p. 46)
- **Figura 7.** Mecanismos fisiopatológicos en las vías respiratorias de la presencia de un gen codificante de CFTR defectuoso. (p. 51)
- **Figura 8.** Representación del modelo de adaptación de clímax y ataque en FQ. (p. 56)
- **Figura 9.** Ciclo de vida bifásico de *B. bacteriovorus*. (p. 61)
- **Figura 10.** Imágenes de microscopía de contraste mostrando células de *B. bacteriovorus* (p. **¡Error! Marcador no definido.**)
- **Figura 11.** Flujo de trabajo y principales resultados obtenidos con ambas metodologías. (p. 86)
- **Figura 12.** Filos, clases y géneros bacterianos identificados exclusivamente por cada una de las metodologías. (p. 88)
- **Figura 13.** Abundancias relativas de los principales taxones detectadas por ambas metodologías. (p. 91)
- **Figura 14.** Cambio composicional de los géneros mayoritarios en cada una de las metodologías. (p. 92)
- **Figura 15.** Evolución de la alfa diversidad (medida a través del índice de Shannon-Weaver y de la riqueza bacteriana) detectada por un método y otro. (p. 93)
- **Figura 16.** Análisis funcional de las proteínas bacterianas de esta cohorte, comparando el grupo de muestras iniciales y el de muestras finales (FQ temprana). (p. 95)

- **Figura 17.** Proteínas humanas encontradas en interacción con el microbioma intestinal de la cohorte. (p. 97)
- **Figura 18.** El eje intestino-cerebro en FQ. (p. 106)
- **Figura 19.** Dendrogramas de los pulsotipos obtenidos por PFGE de las especies *E. faecalis* y *E. faecium*. (p. 122)
- **Figura 20.** Dendrograma de los pulsotipos obtenidos por PFGE de la especie *S. aureus*. (p. 123)
- **Figura 21.** Dendrogramas de los pulsotipos obtenidos por PFGE de las especies Gram negativas *E. coli*, *K. pneumoniae*, *K. oxytoca* y *S. marcescens*. (p. 124)
- **Figura 22.** Representantes clonales de los aislados obtenidos de muestras fecales y nasofaríngeas de la cohorte de FQ. (p. 125)
- **Figura 23.** Detección de clones de las especies Gram positivas identificadas en las muestras fecales y en lavados nasofaríngeos (LN) de los 0 a 12 meses. (p. 126)
- **Figura 24.** Detección de clones de las especies Gram negativas identificadas en las muestras fecales y en lavados nasofaríngeos (LN) de los 0 a 12 meses. (p. 127)
- **Figura 25.** Detección de clones de las especies recuperadas de heces y lavados nasofaríngeos (LN) entre los 12 y los 24 meses. (p. 128)
- **Figura 26.** Evolución de los principales géneros bacterianos en muestras de H y LN en todos los pacientes. (p. 129)
- **Figura 27.** Core bacteriano de heces y LN y patrones de cocolonización en la cohorte de FQ. (p. 130)
- **Figura 28.** Prevalencia y abundancia de los principales géneros bacterianos de la cohorte. (p. 131)
- **Figura 29.** Correlación de los principales géneros con la edad en la cohorte de FQ. (p. 132)
- **Figura 30.** Evolución a lo largo de 2 años de la microbiota mayoritaria en heces. (p. 133)
- **Figura 31.** Evolución a lo largo de 2 años de la microbiota mayoritaria en muestras LN. (p. 134)

- **Figura 32.** Índices de alfa diversidad a lo largo de los dos primeros años de vida en la cohorte de FQ. (p. 135)
- **Figura 33.** Asociaciones positivas y negativas de los principales géneros bacterianos con la edad en la cohorte de FQ, en función de distintas variables. (p. 137)
- **Figura 34.** Cambios composicionales significativos a lo largo del tiempo de los principales géneros bacterianos (N=23) en función de distintas variables en las heces y los LN de la cohorte de FQ. (p. 138)
- **Figura 35.** Prevalencia en la detección de las bacterias cultivables y patógenas clínicas por secuenciación, cultivo rutinario y cultivo de investigación. (p. 140)
- **Figura 36.** Diferencias entre la abundancia relativa de las bacterias cultivables en muestras con cultivo positivo y con cultivo negativo. (p. 141)
- **Figura 37.** Comparación de la diversidad alfa entre el grupo control y brote. (p. 161)
- **Figura 38.** Diferencias entre el grupo brote y control del hospital La Paz. (p. 163)
- **Figura 39.** Diferencias entre los nacidos por cesárea y por parto vaginal. (p. 164)
- **Figura 40.** Principales grupos bacterianos y sus abundancias relativas en todos los individuos y en el grupo control y brote del Hospital La Paz. (p. 165)
- **Figura 41.** Dendrogramas de las especies cultivadas. (p. 168)
- **Figura 42.** Esquema de clones detectados en cada paciente del grupo brote en los distintos tiempos de recogida de muestras. (p. 169)
- **Figura 43.** PFGE conjunto de los aislados de *S. marcescens* encontrados en 2016 (colección contemporánea) y los conservados desde 1969 (colección histórica). (p. 170)
- **Figura 44.** IS detectadas en los 8 aislados de *S. marcescens* estudiados y su frecuencia en los genomas anotados. (p. 171)
- **Figura 45.** Resultado de PlacnetW que refleja las secuencias de plásmidos detectadas. (p. 172)
- **Figura 46.** Análisis filogenético de 452 genomas de *S. marcescens*. (p. 177)

- **Figura 47.** Curvas de depredación de bd109J en la colección *S. marcescens* y células viables antes y después del experimento. (p. 199)
- **Figura 48.** Curvas de depredación de bd109J en la colección *P. aeruginosa* FQ y células viables antes y después del experimento. (p. 199)
- **Figura 49.** Curvas de depredación de bd109J en la colección *P. aeruginosa* BACT y células viables antes y después del experimento. (p. 200)
- **Figura 50.** Mapa de calor con la asociación entre el patrón de resistencia a antibióticos, el origen de cada aislado y el valor de PE. (p. 201)
- **Figura 51.** Susceptibilidad a la depredación de las colecciones de presas clínicas. (p. 202)
- **Figura 52.** Curvas de depredación de bd2637 en la colección *P. aeruginosa* BACT y células viables antes y después del experimento. (p. 203)

## ÍNDICE DE TABLAS

- **Tabla 1.** Limitación del esfuerzo terapéutico y causas de fallecimiento neonatal en función de la edad gestacional en España (2018). % (RQ). (p. 47)
- **Tabla 2.** Morbilidad infecciosa neonatal en 2018. % (RQ). (p. 48)
- **Tabla 3.** Datos demográficos y clínicos de los sujetos con FQ incluidos en el estudio comparado. (p. 79)
- **Tabla 4.** Discrepancias en la nomenclatura o clasificación taxonómica a diferentes rangos (filo/clase/orden/familia/género) por secuenciación de amplicones y metaproteómica . (p. 87)
- **Tabla 5.** Tabla de contingencia para la prueba McNemar. (p. 90)
- **Tabla 6.** Proteínas humanas enriquecidas en cada grupo de la cohorte. (p. 96)
- **Tabla 7.** Condiciones y cebadores para las PCR MLST de *S. aureus*. (p.115)
- **Tabla 8.** Condiciones y cebadores para las PCR MLST de *P. aeruginosa*. (p. 116)
- **Tabla 9.** Características de la cohorte de FQ (p. 119)
- **Tabla 10.** Variables clínicas recogidas a lo largo de los dos primeros años en función del genotipo FQ. (p. 120)
- **Tabla 11.** Características clínicas y demográficas de las cohortes de neonatos del Hospital La Paz. (p. 160)
- **Tabla 12.** Abundancia relativa del género *Serratia* y aislados de la especie *S. marcescens* en el grupo brote. (p. 166)
- **Tabla 13.** Resistoma de los 8 aislados de *S. marcescens* del estudio y principales clases de antibióticos que pueden ser objetivos potenciales de los productos que codifican los genes de resistencia. (p. 174)
- **Tabla 14.** Concentraciones mínimas inhibitorias (CMI) observadas (mg/L) o diámetro de la zona de inhibición (ZD) (mm) de cada antibiótico para las 8 cepas de *S. marcescens* estudiadas. (p. 176)
- **Tabla 15.** Mutaciones cromosómicas acumuladas durante la delección génica. (p. 195)



# **RESUMEN**



**RESUMEN**

La adquisición y el establecimiento del microbioma después del nacimiento puede condicionar la salud del resto de nuestra vida. La predisposición a ciertas patologías o el agravamiento de enfermedades de base está influenciado por las bacterias y otros microorganismos que habitan en nuestro cuerpo. Conocer las dinámicas que ocurren en este complejo ecosistema es necesario para identificar terapias alternativas al uso de antibióticos bajo una perspectiva ecológica. Además, existen múltiples herramientas para abordar la complejidad de la microbiota intestinal y es necesario identificar las fortalezas y debilidades de cada aproximación, así como unificar las bases de datos y la clasificación taxonómica. Estas herramientas permiten, además, reformular conceptos como los brotes o la virulencia potencial de una bacteria, ya que exploran todo el contenido genético, proteico o metabólico de los microorganismos en cuestión. Esta tesis tiene como principal objetivo estudiar la implantación de la microbiota de recién nacidos en diferentes circunstancias con especial vulnerabilidad, y determinar la capacidad de las bacterias depredadoras de erradicar a patógenos clínicos.

En el **primer capítulo**, se evalúa la concordancia entre dos técnicas comunes para el estudio del microbioma intestinal, la secuenciación de amplicones y la metaproteómica, aplicadas en el seguimiento longitudinal de una cohorte de recién nacidos diagnosticados con fibrosis quística durante su primer año de vida. Los resultados muestran que la identificación taxonómica es similar, pero se detectaron importantes diferencias en las proporciones estimadas de ciertas bacterias. En particular, la secuenciación de amplicones detecta menor abundancia de *Bifidobacterium*, pero sobreestima ciertos géneros de los filos *Firmicutes* y *Proteobacteria*, lo que se asocia con los sesgos y limitaciones de ambas metodologías. Además, se identificó a *Ruminococcus gnavus* como bacteria anaerobia metabólicamente activa en el intestino de estos recién nacidos directamente relacionada con un estado de inflamación humana.

En el **segundo capítulo**, se detalla la monitorización de esta cohorte de niños con FQ ampliándola con casos de nuevo diagnóstico, mediante un estudio longitudinal de su microbiota intestinal y nasofaríngea hasta los dos primeros años de vida. Se evaluaron las tendencias de los principales grupos bacterianos mediante secuenciación de amplicones, con especial atención en aquellos patógenos clásicos que causan primoinfecciones durante este periodo vital, y también en bacterias anaerobias conocidas por su implicación en el curso de la FQ. Los resultados mostraron algunas discrepancias entre la

secuenciación de amplicones y las técnicas clásicas de cultivo. Se observó un núcleo compartido de géneros bacterianos entre ambos nichos, incluido a nivel de clones de aislados. Destacó una alta dominancia de *Streptococcus* en las muestras respiratorias y de *Bifidobacterium* en las intestinales. Cada persona presenta clones específicos, aunque en ocasiones se encontraron en más de un paciente.

En el **tercer capítulo**, se compara la implantación de la microbiota intestinal de pacientes prematuros ingresados en una UCIN bajo el contexto de un brote nosocomial causado por *Serratia marcescens* con la de otra cohorte de prematuros ingresados en un periodo sin brote. Se realizó un estudio de casos y controles en la primera y tercera semana de vida, demostrando que la colonización intestinal por *S. marcescens* impacta al resto de la microbiota de estos neonatos. Posteriormente, se realizó un análisis filogenético y fenotípico extenso comparando los aislados del brote contemporáneo con aislados históricos del mismo hospital obtenidos 47 años antes durante un brote. Todos estos aislados se compararon con el resto de las cepas clínicas de *S. marcescens* depositadas en repositorios públicos, determinando su relación filogenética, su resistoma y la presencia de fagos y secuencias de inserción.

En el **cuarto capítulo**, se evaluó la eficiencia de depredación y la tasa de depredación intraespecífica de *Bdellovibrio bacteriovorus* 109J, una bacteria depredadora que se ha propuesto como agente de biocontrol para la erradicación de patógenos Gram negativos. Para ello, se seleccionaron colecciones de *Pseudomonas aeruginosa* procedente de pacientes con bacteriemia y FQ y una colección de *S. marcescens* aisladas de pacientes adultos con bacteriemia y de los neonatos estudiados en el capítulo anterior. *B. bacteriovorus* mostró la mayor eficiencia de depredación en *S. marcescens*, después en la colección de *P. aeruginosa* de FQ, siendo los aislados de *P. aeruginosa* de bacteriemias los menos depredados. Como prueba de concepto, se escogió una depredadora mutante en un único gen, *B. bacteriovorus* bd2637, para observar si los patrones de depredación eran dependientes de la relación específica depredadora-presa, demostrando esta especificidad en los aislados seleccionados. Los resultados muestran cómo es necesario estudiar la susceptibilidad presa y depredadora-específica, tal y como se hace con la susceptibilidad antibiótica clásica, para continuar con las evidencias prometedoras de que estas bacterias pueden servir para controlar infecciones.

## SUMMARY

The acquisition and establishment of the microbiome after birth can affect the state of health for the rest of our lives. The predisposition to certain pathologies or the aggravation of basic diseases is influenced by the bacteria and other microorganisms that inhabit our body. Knowing the dynamics that occur in this complex ecosystem is necessary to identify alternative therapies for antibiotic usage with an ecological perspective. In addition, there are multiple tools to address the complexity of the intestinal microbiota and it is necessary to identify the strengths and weaknesses of each approach, as well as to unify databases and taxonomic classification guidelines. These tools also allow the reformulation of concepts such as outbreaks or the potential virulence of a bacterium, since they explore the entire genetic, protein or metabolic content of the microorganisms in question. The main objective of this thesis is to study the establishment of the newborn microbiota in different circumstances with special vulnerability, and to determine the capacity of predatory bacteria to eradicate clinical pathogens.

The **first chapter** evaluates the concordance between two common techniques for the study of the intestinal microbiome, amplicon sequencing and metaproteomics, applied in the longitudinal follow-up of a cohort of newborns diagnosed with cystic fibrosis during their first year of life. The results show that taxonomic identification is similar, but important differences in the estimated proportions of certain bacteria were identified. In particular, amplicon sequencing detects lower abundance of *Bifidobacterium* but overestimates certain genera of *Firmicutes* and *Proteobacteria* phyla, which is associated with the biases and limitations of both methodologies. In addition, *Ruminococcus gnavus* was identified as a metabolically active anaerobic bacterium in the intestine of these newborns directly related to a condition of human inflammation.

The **second chapter** details the monitoring of this cohort of CF children by prolonging it with newly diagnosed cases, through a longitudinal study of their intestinal and nasopharyngeal microbiota up to the first two years of life. The trends of the main bacterial groups were evaluated by amplicon sequencing, with special attention to those classical pathogens that cause primoinfections during this vital period and, also, to anaerobic bacteria known to be involved during the course of CF. The results showed some discrepancies between amplicon sequencing and classical culture techniques. A shared core of bacterial genera was observed between both niches, including at the level of isolated clones. A high dominance of *Streptococcus* in the respiratory samples and of

*Bifidobacterium* in the intestinal samples stood out. Each individual had specific clones, although sometimes they were found in more than one patient.

In the **third chapter** the establishment of the intestinal microbiota of preterm patients admitted to a NICU in the context of a nosocomial outbreak caused by *Serratia marcescens* with that of another cohort of preterm patients admitted in a period without an outbreak were compared. A case-control study was performed in the first and third weeks of life, showing that intestinal colonization by *S. marcescens* impacts the rest of the microbiota of these neonates. Subsequently, an extensive phylogenetic and phenotypic analysis was performed comparing the contemporary outbreak isolates with historical isolates from the same hospital obtained 47 years earlier during an outbreak. These isolates were compared with the rest of the clinical isolates of *S. marcescens* deposited in public repositories, determining their phylogenetic relationship, resistome, and the presence of phage and insertion sequences.

In the **fourth chapter** the predation efficiency and intraspecific predation rate of *Bdellovibrio bacteriovorus* 109J, a predatory bacterium that has been proposed as a biocontrol agent for the eradication of Gram-negative pathogens, was evaluated. For this purpose, collections of *Pseudomonas aeruginosa* from patients with bacteremia and CF and a collection of *S. marcescens* isolated from adult patients with bacteremia and from the neonates studied in the previous chapter were selected. *B. bacteriovorus* showed the highest predation efficiency on *S. marcescens*, followed by the CF *P. aeruginosa* collection, with *P. aeruginosa* isolates from bacteremia being the least depredated. As a proof of concept, a single gene predator mutant, *B. bacteriovorus* bd2637, was chosen to observe whether predation patterns were dependent on the specific predator-prey relationship, demonstrating this specificity in our isolates. The results show how it is necessary to study this prey-predator relationship, as it is done with classical antibiotic susceptibility, in order to continue with the promising evidence of predatory bacteria being able to control infections.

# **INTRODUCCIÓN**



## INTRODUCCIÓN

### 1. MICROBIOMA HUMANO

#### 1.1. Definiciones y conceptos básicos

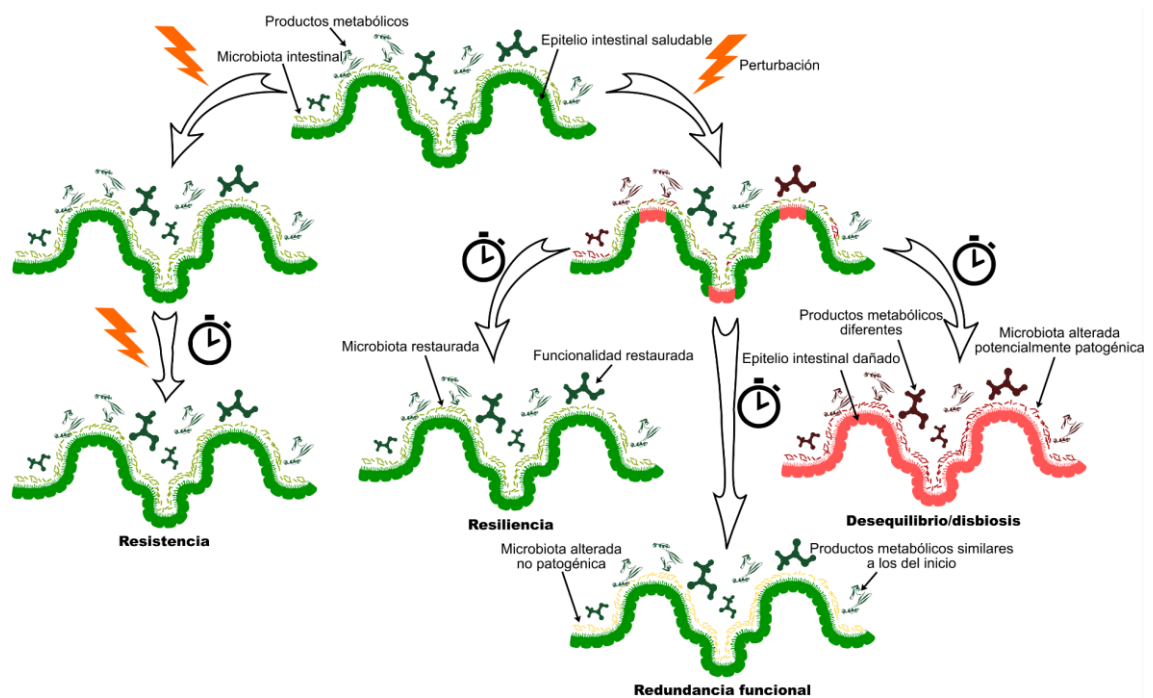
La microbiota humana es el conjunto de microorganismos que habitan en nuestro cuerpo recubriendo epitelios y mucosas, ocupando la interfaz con el exterior, y que establecen con el organismo relaciones simbióticas (de mutuo beneficio), comensales (con un beneficio neutro) o patogénicas (con un perjuicio hacia el hospedador) [1]. Estas poblaciones pueden ser estables (autóctonas) o transitorias (alóctonas). El concepto de microbioma, por otro lado, engloba a los microorganismos del hábitat, su contenido genético y las características ambientales. A veces se usan indistintamente las palabras microbiota y microbioma. Además, tenemos el término metagenoma, que es el conjunto de genomas y genes de los miembros de la microbiota; el metatranscriptoma, o conjunto de transcritos, que revela información sobre la expresión de los genes; y el metaproteoma o contenido proteico de la microbiota, que informa sobre los genes que acaban por traducirse a proteínas. Finalmente, el metaboloma es el conjunto de rutas metabólicas de la microbiota y sus metabolitos [2]. Tal es la estrecha relación que parece darse entre huésped y microbiota a lo largo de la evolución de las especies, que también se habla del término holobionte para referirse al hospedador animal y a todas las comunidades microbianas que habitan en él [3].

A menudo, los estudios sobre el microbioma se centran únicamente en el dominio *Bacteria*. Sin embargo, deberíamos tener en cuenta que las arqueas, los hongos, los helmintos y otros animales microscópicos, e incluso, salvando las definiciones de lo que son los seres vivos, los virus, son actores relevantes del microbioma humano. Por ello, se comienza a hablar de términos más específicos como arqueoma, micobioma o viroma para referirse a estas comunidades [4]–[6].

Dependiendo del lugar anatómico, la microbiota humana está compuesta por comunidades distintas. La más estudiada es la del tracto gastrointestinal, por su gran densidad de microorganismos (se estima que del orden de  $10^{14}$  células bacterianas por cada gramo de tejido) y el fácil acceso a través de muestras fecales o biopsias, que a su vez se compone de varios ecosistemas anatómicamente diferenciados. El colon alberga cerca del 70% de las bacterias que se encuentran en todo el cuerpo humano [7], [8].

La composición de la microbiota intestinal varía en función de numerosos factores que se explicarán más adelante. En general, se consideran miembros comunes las bacterias de los filos *Firmicutes* (actualmente renombrado *Bacillota*), *Bacteroidota* y, en menor proporción, *Proteobacteria* (actualmente renombrado *Pseudomonadota*), *Actinobacteriota*, *Fusobacteriota* y *Verrucomicrobia*<sup>1</sup> [9].

Una característica deseable de la microbiota en condiciones de salud es que sea capaz de mantenerse estable frente a los cambios y perturbaciones producidos por la dieta, los niveles de estrés o el contacto con patógenos exógenos. Se considera estable a una microbiota que, tras una perturbación, no sufre cambios en su composición (tiene resistencia) o bien, si los sufre, que es capaz de restaurar la composición original (tiene resiliencia), o de recuperar sus funciones originales o fenotipo (tiene redundancia funcional) (Figura 1) [10].



**Figura 1. Dinámicas ecológicas de la microbiota intestinal.**

<sup>1</sup> A lo largo de esta tesis se va a nombrar a los principales grupos taxonómicos de acuerdo con la designación más común y familiar (*Firmicutes*, *Proteobacteria*, *Actinobacteriota* y *Bacteroidota*), en lugar de las últimas actualizaciones, para evitar confusión.

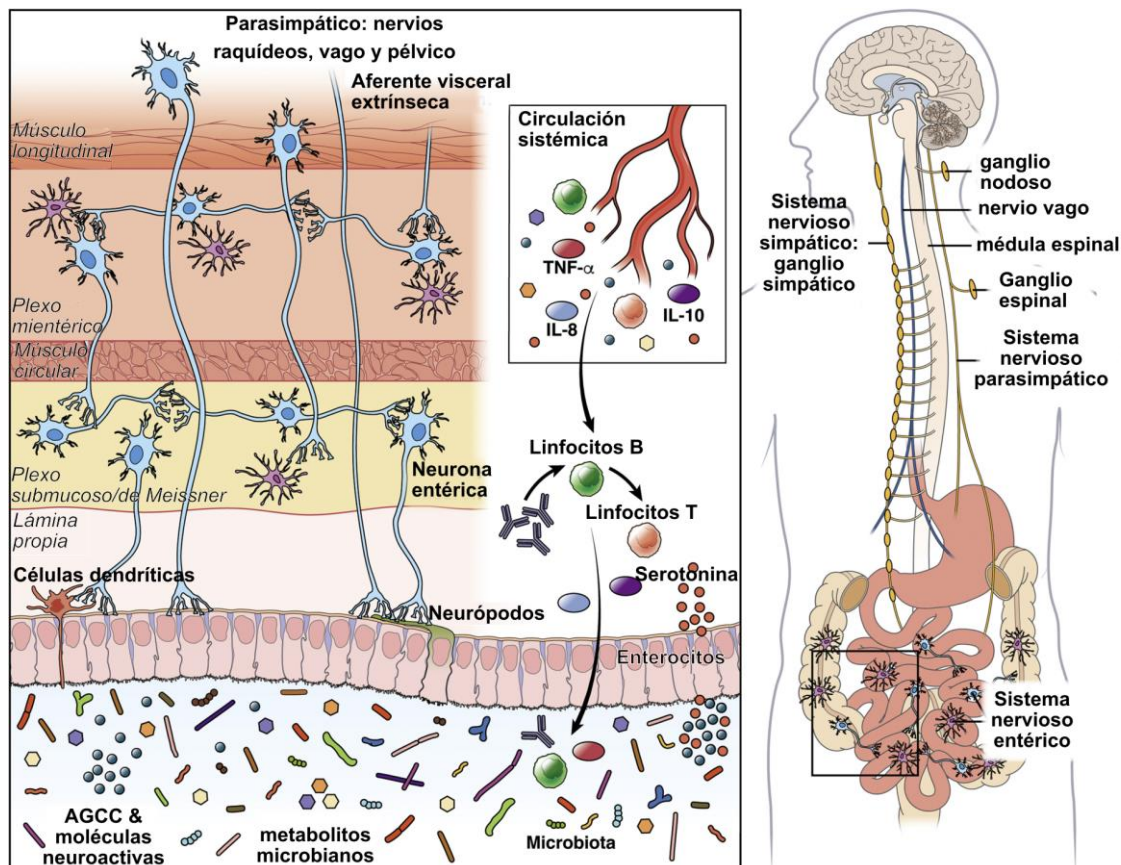
## 1.2. Funciones de la microbiota

La microbiota ejerce numerosas funciones. Por un lado, y en condiciones normales, ayuda al mantenimiento de la integridad de la barrera intestinal, induce a las células de los tejidos linfoides a secretar sustancias antiinflamatorias (IL-10, IL-22 y TGF-beta1) y favorece las uniones estrechas entre las células del epitelio intestinal. Ante la irrupción de patógenos, la microbiota induce al sistema inmunitario para la producción de citocinas proinflamatorias (TNF-alfa, IL-6 o IL-1beta), a la vez que dispone de una batería de sustancias con capacidad antimicrobiana para competir con microorganismos transitorios [11].

Tiene un papel muy relevante en nuestro metabolismo, ya que permite la digestión y asimilación de nutrientes de los cuales no podríamos disponer de otra forma. Las bacterias, especialmente las pertenecientes a los filos *Firmicutes* y *Bacteroidota*, convierten los carbohidratos complejos como la celulosa, hemicelulosa, el almidón resistente, la pectina, la lignina y los oligosacáridos, en ácidos grasos de cadena corta (AGCC), como el butirato, el acetato o el propionato [12]. Estos AGCC tienen numerosas implicaciones en el metabolismo de lípidos, la regulación del apetito y sirven como fuente de energía a los colonocitos, entre otras funciones [13]. Además, las bacterias sintetizan vitaminas del grupo B (tiamina o B1, riboflavina o B2, niacina o B3, biotina o B7, cobalamina o B12) y la vitamina K [14], que facilitan la absorción de calcio y de otros minerales y regulan procesos de remodelación ósea [15]. También transforman ácidos biliares primarios a formas secundarias, y producen, a partir de aminoácidos aromáticos de la dieta, aminas aromáticas [16]. Por último, también participan en la asimilación de drogas y de otros compuestos [17].

Los AGCC, y los productos derivados de los aminoácidos aromáticos y los ácidos biliares anteriormente nombrados no sólo son una fuente de energía, sino que influyen en la maduración de la microglía y en otros aspectos del sistema nervioso central [18]. Estos metabolitos tienen la capacidad de llegar a la barrera hematoencefálica y favorecen la formación de uniones estrechas que impiden la entrada de sustancias no deseadas. Además, inducen la producción de leptina y de un péptido similar al glucagón tipo 1 en las células endocrinas entéricas. Por otro lado, se ha comprobado que miembros típicos de la microbiota como *Escherichia coli* o especies de *Lactobacillus* son capaces de producir neurotransmisores como la dopamina, serotonina, noradrenalina, GABA, acetilcolina o histamina [19]. La compleja y estrecha conexión entre la microbiota

intestinal y sus productos metabólicos con el sistema nervioso es posible gracias al eje microbiota-intestino-cerebro, el cual comunica de manera bidireccional al sistema nervioso entérico con el sistema nervioso central. La mayoría de estas conexiones ocurren a través de la aferencia del nervio vago (Figura 2) [20].



**Figura 2. Vías de comunicación entre la microbiota y el cerebro.** La microbiota interviene en el eje intestino-cerebro a través de la circulación sanguínea por múltiples vías, como la endocrina, la inmunológica o la neuronal. Los neurópodos presentes en las células endocrinas entéricas transducen señales sensoriales desde el medio intestinal al cerebro mediante la formación de conexiones similares a las sinapsis con las neuronas de los nervios aferentes como el nervio vago. Traducción de Margolis, K. G., Cryan, J. F., & Mayer, E. A. (2021). The Microbiota-Gut-Brain Axis: From Motility to Mood. *Gastroenterology*, 160(5), 1486–1501. Doi: 10.1053/j.gastro.2020.10.066 (licencia libre).

En otras localizaciones anatómicas, la microbiota también tiene un papel clave en la homeóstasis. Las bacterias que colonizan el tracto respiratorio secretan sustancias que inhiben la formación de biopelículas por otras bacterias patógenas [21]. A lo largo del tracto urogenital femenino, la microbiota colabora con las células inmunitarias en el

mantenimiento de un pH ácido que mantiene a raya a patógenos que puedan causar infecciones de transmisión sexual [22]. En la piel, previene la invasión de patógenos a través del órgano más expuesto del cuerpo, e interviene en la termorregulación y en la síntesis de vitamina D [23]. Sin embargo, la microbiota también se presenta como un reservorio de potenciales patógenos que se transmiten con facilidad, evento especialmente preocupante en el entorno hospitalario [24].

Finalmente, se ha comprobado que ha existido una coevolución entre la microbiota intestinal y los homínidos, lo que implica que ha habido una transmisión intraespecífica a lo largo de generaciones, principalmente por la línea maternofilial, con posterior transmisión horizontal por parte de otros familiares y del entorno. Esto explica la existencia tanto de una microbiota compartida por los miembros de cada especie, como de una microbiota que se diversifica según características demográficas o culturales [25].

### 1.3. Relación con la salud y la enfermedad

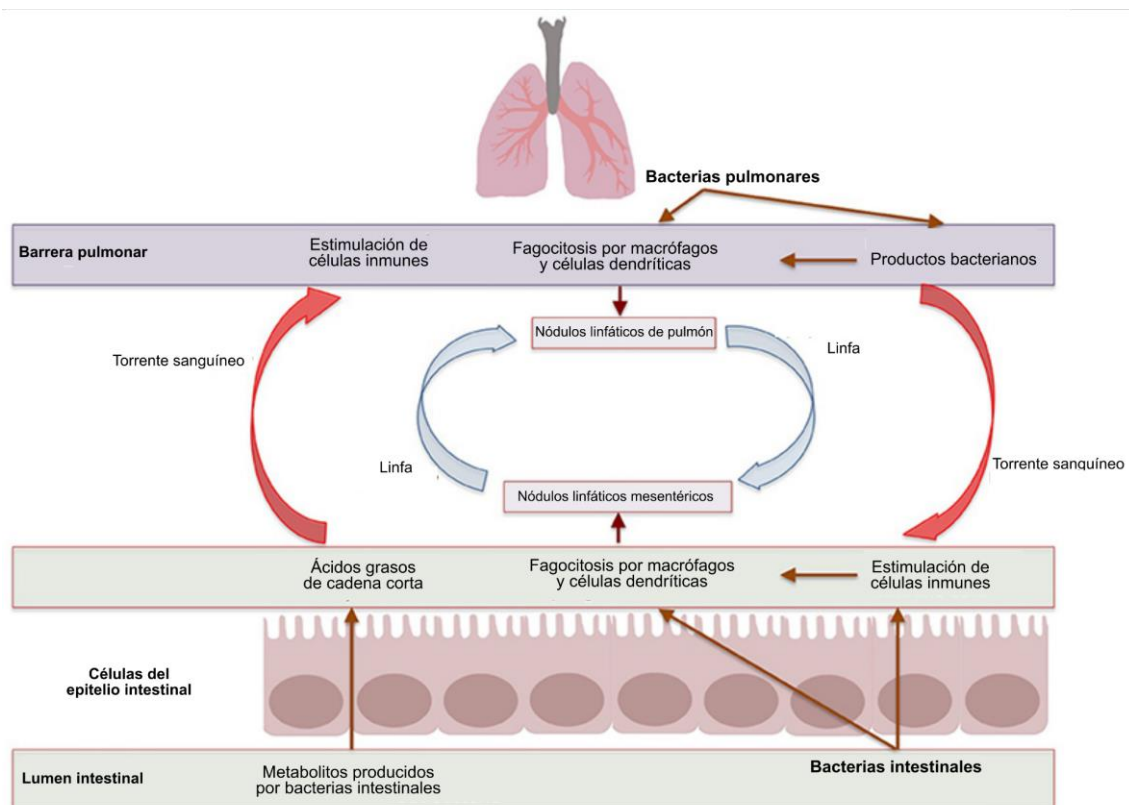
Se ha observado que existe una asociación entre determinados trastornos o enfermedades y un estado de desequilibrio de la microbiota intestinal respecto a lo común en un determinado individuo. Hay asociaciones con efecto protector en *Bifidobacterium* spp., *Akkermansia muciniphila* o *Faecalibacterium prausnitzii* [2], mientras que *Fusobacterium nucleatum* o *Streptococcus gallolyticus* subsp. *gallolyticus*, comunes en la microbiota oral e intestinal, se relacionan con el cáncer de colon [26] o, en el caso de *Yersinia pseudotuberculosis*, con la enfermedad inflamatoria intestinal [27].

Por el momento no disponemos de un consenso que defina los límites de una microbiota sana, aunque hay algunas certezas. Por un lado, se ha asociado la pérdida de diversidad en la microbiota intestinal con la enfermedad inflamatoria intestinal, la otitis media aguda, la infección urinaria o la infección por *Clostridioides difficile* [28], [29]. Sin embargo, en el caso de la microbiota vaginal ocurre lo contrario: el aumento de diversidad bacteriana y la disminución de las especies de *Lactobacillus* se asocian a enfermedades como la vaginosis, la inflamación y el riesgo de adquirir infecciones de transmisión sexual [22].

Para elucidar el papel de la microbiota intestinal en la salud se han utilizado ratones libres de gérmenes o modelos tratados con antibióticos para provocar una disrupción de las comunidades microbianas. Estos estudios han mostrado que, en ausencia de microbiota o en presencia de una microbiota alterada, hay una disminución

de factores que promueven la neurogénesis, falta de sociabilidad, depresión y ansiedad [30]. Asimismo, existen numerosos estudios que apuntan a una relación entre una microbiota alterada y numerosos trastornos neurofisiológicos como la esclerosis múltiple, el Alzheimer y el Parkinson, así como con trastornos de ansiedad o del espectro autista [18].

Se ha propuesto que, al igual que existe el concepto del eje intestino-cerebro, existe el eje intestino-pulmón (Figura 3). Varios estudios han informado de que un desequilibrio en la composición de la microbiota intestinal se asocia a enfermedades respiratorias como el asma, la fibrosis quística (FQ), la neumonía o la gripe [21]. Este eje actúa bidireccionalmente, ya que otra investigación demostró que la estimulación con endotoxina o lipopolisacárido (LPS) en los pulmones de roedores provocaba un aumento en el número de bacterias intestinales [31]. Además, se sospecha una estrecha relación entre la inflamación, la disrupción de la barrera epitelial y los niveles altos de endotoxina en el torrente sanguíneo [32].



**Figura 3. Conexiones en el eje bidireccional intestino-pulmón.** Traducción de S. Anand y S. S. Mande, “Diet, Microbiota and Gut-Lung Connection”, *Front Microbiol*, vol. 9, Sep. 2018. Doi: 10.3389/fmicb.2018.02147 (licencia libre).

Diversos trabajos sugieren que la microbiota podría ser el eslabón perdido que ayude a comprender el mecanismo patogénico de enfermedades autoinmunitarias como la artritis reumatoide, el lupus eritematoso sistémico o la espondiloartritis, debido a su estrecha relación con la maduración del sistema inmunitario innato y con la producción de citocinas [33]. La hipótesis de la higiene propone que el aumento de medidas sanitarias está asociado con una disminución en la incidencia de infecciones en la infancia temprana, pero también con un aumento de las alergias y de otras enfermedades de carácter autoinmunitario. Se ha demostrado que el contacto con bacterias como *E. coli* y, en particular, con su LPS, contribuye a la educación del sistema inmunitario, y se ha asociado a una menor incidencia de alergias, en comparación con el subtipo de LPS de especies de *Bacteroides*, que inhibe la actividad inmunoestimulante del anterior [34].

Todos estos conocimientos se han recopilado gracias al desarrollo de la investigación traslacional, la cual integra la investigación básica y clínica con la medicina asistencial y la monitorización epidemiológica de las enfermedades, en este caso, relacionadas con el microbioma. El objetivo de este enfoque holístico es poner los últimos avances en tecnología al servicio de la vigilancia en problemas de salud para mejorar el abordaje diagnóstico, terapéutico, preventivo y pronóstico.

## 2. HERRAMIENTAS DE LA INVESTIGACIÓN TRASLACIONAL EN MICROBIOLOGÍA

La investigación traslacional es aquella que se origina desde el laboratorio para llegar a aplicarse en la medicina asistencial. En el área de la Microbiología y las Enfermedades Infecciosas, esta visión ha sido relevante, sobre todo, en el seguimiento y caracterización de los brotes y en el estudio del microbioma con el desarrollo de las técnicas -ómicas [35]. Los institutos de investigación biomédica trabajan en sinergia con los servicios de los hospitales públicos de referencia, como es el caso de nuestro instituto IRYCIS de investigaciones sanitarias asociado al Hospital Universitario Ramón y Cajal de Madrid.

### 2.1. Técnicas de estudio de brotes infecciosos

El aumento del número de casos esperados asociados a una infección en un lugar concreto, lo que se conoce como brote infeccioso, es una amenaza frecuente para las poblaciones más vulnerables de todo el mundo. En los hospitales y, en particular, en las Unidades de Cuidados Intensivos (UCI), los pacientes son particularmente susceptibles de sufrir este tipo de eventos, que se agravan si están causados por bacterias con multirresistencia a los antibióticos que sobreviven en este ambiente, asociadas a las superficies y el instrumental (catéteres, tubos, drenajes) [36].

#### 2.1.1. Técnicas convencionales

La mayor parte de los Servicios de Microbiología caracterizan fenotípicamente a las bacterias causantes de infecciones mediante su identificación por espectrometría de masas MALDI-TOF, pruebas bioquímicas o serotipado, realizando además un estudio de su sensibilidad a los antibióticos. Muchos de estos métodos siguen siendo el *gold-standard*, aunque consumen tiempo y recursos y carecen de resolución taxonómica [37].

Para monitorizar los brotes infecciosos, tradicionalmente y aún hoy en día, se emplea la electroforesis en campo pulsado (PFGE), la cual permite detectar patrones similares existentes en diferentes aislados [38]. Estos patrones pueden agruparse usando diferentes algoritmos y se pueden construir árboles o dendrogramas que establezcan la relación entre aislados. Estos protocolos tardan días en realizarse, y no informan del contenido genético de las bacterias, sino de su similitud [39].

Otras técnicas moleculares comunes se basan en la secuenciación de un sólo gen o de varios que suelen ser constitutivos, como es el caso de *Multi-Locus Sequence Typing*

(MLST), cuyas variantes de secuencia permiten discriminar entre aislados de la misma especie. El inconveniente es que esta aproximación no está implementada para todas las especies [40].

### **2.1.2. Secuenciación de genoma completo**

Mediante la secuenciación de genoma completo, se obtiene todo el contenido genético de un organismo a partir del ADN extraído de un aislado puro. En este sentido, es una técnica realmente dependiente de cultivo. Algunas herramientas, además, permiten circularizarlo para obtener la disposición de todos los genes. Con ello, puede explorarse el resistoma o batería de genes de resistencia, el contenido en fagos, los elementos genéticos móviles o las secuencias de inserción y, así, indagar en la capacidad de transmisión entre pacientes y el entorno. Secuenciando el genoma completo de aislados en todo el mundo se pueden estudiar las relaciones filogenéticas dentro de una misma especie e incluso establecer conexiones evolutivas. En países como los Estados Unidos se ha integrado en la rutina de Microbiología Clínica para el estudio de brotes infecciosos. Sin embargo, su implementación se encuentra limitada por el coste, la formación requerida y la necesidad de comunicar rápido y eficazmente los resultados, entre otros [36].

## **2.2. Técnicas de estudio del microbioma**

Para comprender la influencia de la microbiota en la salud humana y en el desarrollo de patologías, es necesario investigar esta relación en las distintas capas de complejidad de nuestra fisiología; es decir, al nivel del genoma, del proteoma, del transcriptoma o del metaboloma (Figura 4). Dependiendo de las cuestiones a responder: *¿qué comunidades hay? ¿qué componente genético tienen? ¿qué funciones realizan? ¿qué proteínas o metabolitos expresan?*, será más racional escoger una herramienta u otra. La perspectiva futura pasa por abrazar el enfoque multiómico, en el cual se integran dos o más análisis de los que se van a exponer a continuación, aunque sigue siendo un área en desarrollo y con unos requerimientos analíticos exigentes. El Proyecto del Microbioma Humano Integrativo (iHMP, <https://hmpdacc.org/ihmp/>) es un ejemplo de lo prometedor que puede ser este abordaje en el futuro [41], [42].

### **2.2.1. Análisis metagenómico**

El material genético codifica la información, tanto de los genes reguladores para procesos biológicos, como de los genes codificantes de proteínas para los productos que las

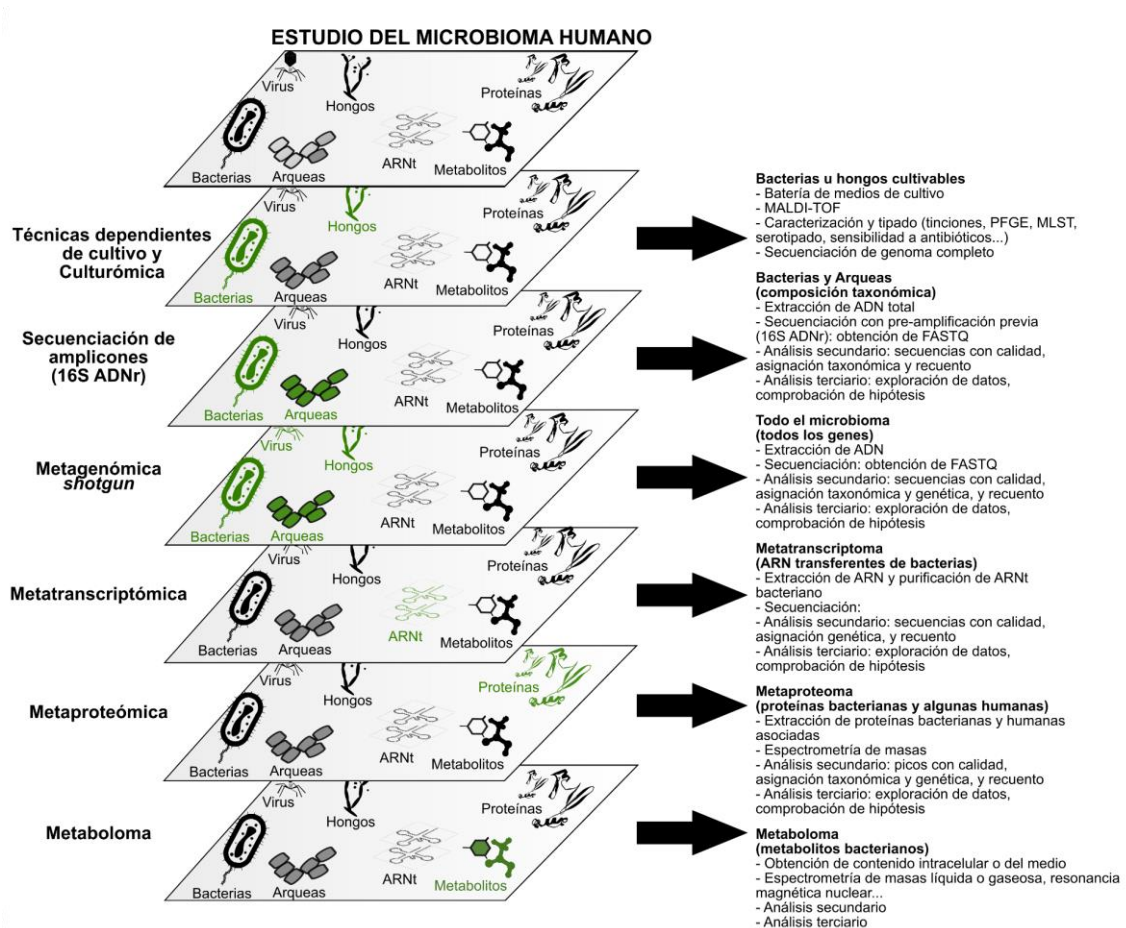
bacterias potencialmente pueden expresar. Además, existen genes que se conservan a lo largo de la evolución debido a su implicación en procesos vitales imprescindibles (los genes constitutivos) y que, por tanto, funcionan como una huella filogenética que permite estudiar relaciones de parentesco entre los organismos. La genómica estudia el contenido de todos ellos, a través de plataformas como Ion Torrent/Ion Proton/Ion Proton, Illumina (MiSeq, NextSeq) y Oxford Nanopore [43].

El flujo de trabajo en la etapa primaria consiste en:

- 1) Aislar el material genético con la pureza y calidad óptimas;
- 2) Generar fragmentos de ADN y
- 3) Preparar librerías genómicas (con secuencias tipo código de barras o etiqueta que identifican el origen de cada fragmento).

Al secuenciar fragmentos de ADN pequeños se pueden obtener errores. Esto se solventa con una buena profundidad de secuenciación, con lo que un nucleótido se secuencia varias veces y el consenso de todas las lecturas de secuenciación que cubran esa región sirve para encontrar la correcta [44]. En la etapa secundaria, se procesan los datos obtenidos en la secuenciación (normalmente archivos FASTQ) hasta la obtención de hojas de datos con la identificación de los genes y su recuento. Finalmente, en una etapa terciaria del análisis se realiza la exploración estadística de los datos.

Las fortalezas de la metagenómica son su bajo coste y su alta implementación en los estudios de microbioma. No obstante, aún son necesarios criterios de estandarización para la extracción de ADN, la elección de los cebadores (en caso de secuenciación de amplicones) y, sobre todo, las bases de datos unificadas. Como limitaciones, destacan su incapacidad para discriminar si el contenido encontrado pertenece a bacterias viables o muertas, la falta de acceso a las bacterias vivas para profundizar en su caracterización fenotípica, el sesgo de profundidad que hace que las especies minoritarias estén infrarrepresentadas, y la naturaleza composicional de los datos metagenómicos; es decir, el hecho de que el recuento de un taxón refleja su abundancia con relación al resto de taxones, y no su recuento absoluto en la muestra de origen [45].



**Figura 4. Métodos de estudio del microbioma humano.**

### 2.2.1.1. *Shotgun sequencing*

Este análisis consiste en secuenciar el ADN total que hay en una muestra, obteniendo no sólo las lecturas de bacterias, sino también de arqueas, hongos, virus, fagos, ADN del huésped, ADN mitocondrial y de plastos, así como elementos extracromosómicos [46]. Esta técnica se ha utilizado ampliamente en el Proyecto del Microbioma Humano y Metagenómica del Tracto Gastrointestinal Humano [47], [48]. Los catálogos de genes obtenidos pueden compararse en bases de datos como KEGG (Enciclopedia Kyoto de Genes y Genomas, <https://www.kegg.jp/kegg/genome/>), la cual, además, relaciona los genes con sus productos codificados y con los procesos biológicos y rutas metabólicas en las que participan [43].

La etapa secundaria en *shotgun* consiste en filtrar las secuencias de baja calidad, eliminar las secuencias duplicadas, contaminantes y las etiquetas, ensamblar las secuencias y agruparlas para reconstruir genomas y zonas de interés, asignar la taxonomía

en base al ADN ribosomal (ADNr), predecir transcritos y anotar funcionalmente a las proteínas predichas por los genes secuenciados. El ensamblaje puede hacerse *de novo* o comparando las lecturas con bases de datos de referencia de especies ya identificadas [46]. Entre las fortalezas del análisis *shotgun*, se encuentra la posibilidad de estudiar una gran diversidad de genes y organismos bajo una misma condición, la identificación a nivel de especie (e incluso de cepa) de las bacterias que dominan el ecosistema y la anotación funcional [49]. Su principal limitación radica en requerir un análisis bioinformático extenso y cualificado. También hay ocasiones en las que la contaminación con el ADN humano dificulta el procesamiento de las muestras, como es el caso de los estudios sobre el microbioma pulmonar en pacientes con FQ [50].

### 2.2.1.2. Secuenciación de amplicones

El objetivo de esta técnica es identificar los grupos microbianos presentes en una comunidad. El fundamento de este análisis se basa en la conservación de genes ribosomales a lo largo de la evolución y en la existencia, dentro de la secuencia de esos genes, de regiones hipervariables que son una marca única de cada especie o grupo de organismos con relación filogenética. En las bacterias, el gen conservado universal es el 16S del ADNr, cuyo transcrito forma parte de la subunidad 30S del ribosoma bacteriano. Durante la etapa primaria, en el paso previo de la preparación de la librería, se debe realizar una amplificación PCR de las regiones hipervariables deseadas, que normalmente son las regiones V1-V2, V3-V4, o V4-V5. Existen numerosos cebadores con diferentes dianas y de su elección dependerá el análisis posterior.

En la etapa secundaria se eliminan las secuencias de baja calidad o de pequeño tamaño y las secuencias de etiqueta, se ensamblan los fragmentos (en caso de haber hecho una aproximación *paired-end*), se encuentran secuencias representativas (ASV) y se recuentan, se asigna taxonómicamente cada ASV a un grupo taxonómico cuya máxima concreción es el nivel de género (a menudo varios ASV se asignan al mismo, o bien quedan ciertos ASV sin asignar y deben eliminarse) y, finalmente, se generan tablas de recuento de ASV/taxones por muestra biológica y árboles filogenéticos. Para la asignación taxonómica se han usado tradicionalmente clasificadores de secuencias como SILVA [51], RDP [52] o *Greengenes* [53]; pero, debido a los resultados variables, ha habido intentos de reconstruir bases de datos unificadas y optimizadas para que la asignación taxonómica sea lo más precisa posible [54], [55].

En la etapa terciaria de un estudio con secuenciación de amplicones, se realiza el cálculo de índices de diversidad alfa y beta y, según el objeto de estudio, análisis de abundancia diferencial o de descubrimiento de biomarcadores, entre otras exploraciones estadísticas. Se emplean herramientas como LEfSe (*Linear discriminant analysis Effect Size*), que determina los rasgos (que pueden ser organismos, clados, ASV, genes o funciones) con mayor probabilidad de explicar las diferencias entre grupos de muestras, acoplando pruebas estándar de significación estadística con pruebas adicionales que evalúan la consistencia biológica y la relevancia del efecto observado [56]. Se emplean páginas web para realizar este tipo de análisis de datos en línea con sus servidores, como es el caso de la plataforma de [Galaxy Huttenhower](#). Por último, existen multitud de paquetes de R para realizar todo tipo de análisis como *phyloseq*, *MetagenomeSeq*, o *microbiome*. Los equipos con experiencia y conocimientos bioinformáticos realizan todo el flujo de trabajo desde la etapa secundaria hasta la obtención de resultados estadísticos y gráficos empleando herramientas de código abierto como Qiime [57].

Entre las limitaciones específicas de la secuenciación de amplicones se encuentran la baja resolución para identificar especies bacterianas y cepas particulares, junto a la falta de consistencia a la hora de elegir los cebadores [58].

### 2.2.2. Culturómica

Curiosamente, el avance en el empleo de las técnicas de secuenciación para el estudio del microbioma humano supuso un renacer de las técnicas dependientes de cultivo. En 2012 se crea el término culturómica para hacer referencia a la implantación masiva de diferentes condiciones de cultivo microbiológico de alto rendimiento, junto con la identificación por MALDI-TOF, y acoplado a la secuenciación de amplicones. Los objetivos de la culturómica son, por un lado, llegar a la resolución de cepa bacteriana y estudiar las características bioquímicas en casos donde discriminar entre cepas o clones sea relevante para la condición o patología y, por otro, indagar en la “materia oscura”, aquella porción de secuencias que siguen sin asignarse a ningún taxón por medio de la secuenciación [59]. Entre sus limitaciones, se encuentra la dificultad para recuperar bacterias anaerobias estrictas, la optimización de los medios de cultivo y los tiempos de incubación extensos [60].

### 2.2.3. Análisis metatranscriptómico

Mediante la secuenciación de ARN mensajero (ARNm) bacteriano se pueden detectar y cuantificar los transcritos sintetizados bajo una determinada condición biológica y conocer el perfil de expresión génica real, a diferencia de la metagenómica, que solo obtiene el perfil potencial. Este abordaje permite conocer la interacción de la microbiota con su hospedador. Una desventaja de esta técnica es que en el ARNm bacteriano falta la cola de poli-A típica del ARNm eucariótico, y esto hace que su aislamiento del resto de ARN, mediante codificación de su ADN complementario (ADNc), sea complicado, teniendo en cuenta que sólo representa el 5% del ARN bacteriano total. Además, el ARNm es muy inestable y faltan métodos estandarizados para su extracción y almacenamiento, así como para el análisis de datos [61]. Por último, esta técnica no informa de las proteínas que realmente estarían traduciendo bajo esa condición. Normalmente, se propone combinar la metatranscriptómica con otros análisis -ómicos, pero esto aumenta considerablemente el presupuesto y el equipo requerido [62].

### 2.2.4. Análisis metaproteómico

Las proteínas son consideradas la unidad funcional de los organismos. En metaproteómica el flujo de trabajo comienza con la extracción, aislamiento y digestión de las proteínas de una muestra para obtener péptidos, que se separan y analizan mediante cromatografía líquida acoplada a espectrometría de masas en tándem (LC-MS/MS). Los espectros se comparan con bases de datos que contienen la información de muchas proteínas y se les asocia una identidad proteica. Mediante metaproteómica, se puede tomar el camino inverso respecto a la metagenómica e inferir qué composición bacteriana hay en una muestra en función de la composición peptídica. El fundamento de este análisis taxonómico es que las proteínas y los péptidos tienen firmas particulares de especies o grupos bacterianos de mayor rango y, gracias a herramientas bioinformáticas como MetaLab, se puede obtener esta información [63]. A diferencia de la metagenómica, la composición taxonómica obtenida por metaproteómica indica qué organismos sintetizan proteínas; es decir, tienen o han tenido el metabolismo activado. No obstante, la cobertura de la técnica depende del número de microorganismos. Las comunidades más complejas tienen menos porcentaje de proteoma por organismo cubierto, siendo las especies más dominantes las que tienen mayor proporción. Se estima que, aproximadamente, entre el 50 y el 80% de las proteínas estimadas se detectan por organismo y experimento, ya que no todo el proteoma es expresado bajo una misma

condición y el tamaño del genoma varía de unos organismos a otros [64]. Los principales retos de la metaproteómica son la complejidad de la preparación de las muestras, el riesgo de contaminación y el análisis de los datos, debido a la formación requerida, a la falta de bases de datos de secuencias de proteínas y a su anotación, muchas veces ambigua [61].

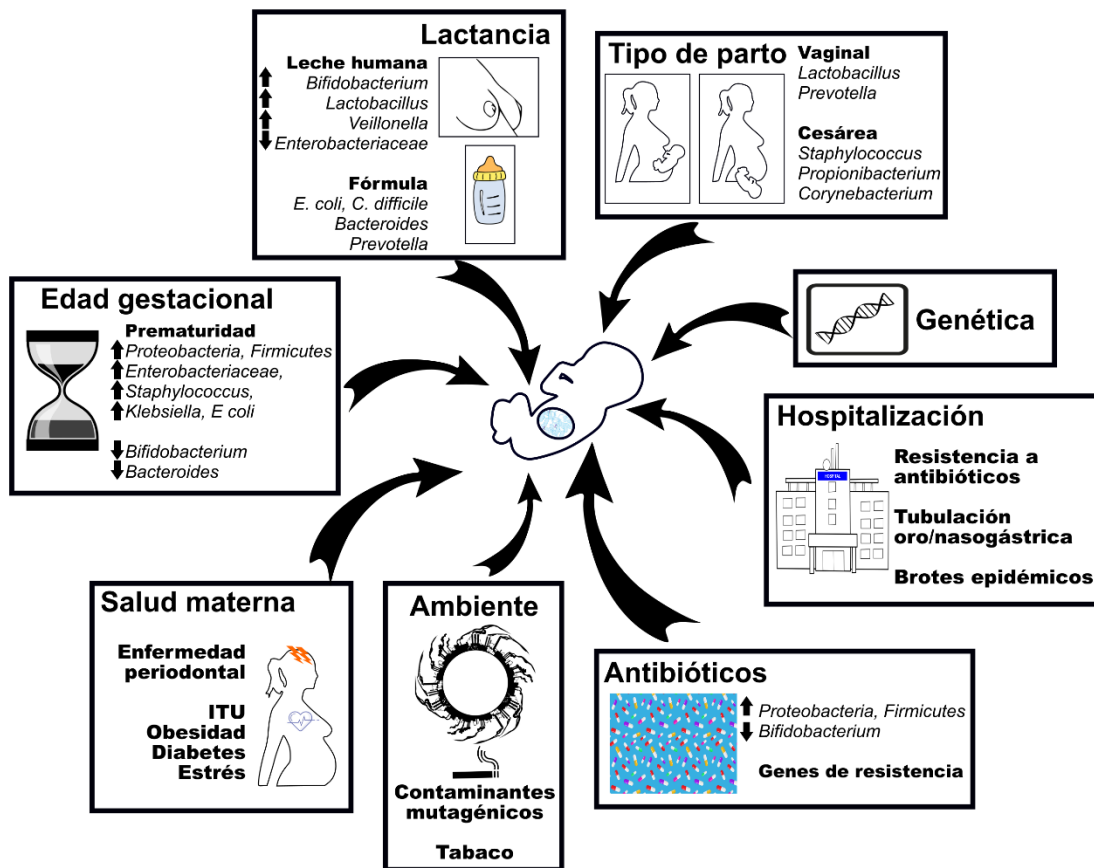
### **2.2.5. Análisis metabolómico**

Los metabolitos son los componentes básicos de la función celular y reflejan las reacciones químicas de las enzimas. Mediante la metabolómica, se estudia también el perfil fenotípico de la microbiota. Permite identificar nuevos biomarcadores o dianas de medicamentos o qué moléculas de las comunidades microbianas están implicadas en un determinado fenómeno biológico. Las técnicas utilizadas en metabolómica se comparten en parte con la metaproteómica, como la cromatografía de gas o líquida acoplada a la espectrometría de masas en tándem (LC o GC MS/MS). También se emplea la resonancia magnética nuclear. Existen dos aproximaciones principales: la metabolómica dirigida y la no dirigida. Esta última es la que se emplea en los estudios exploratorios y no requiere de datos ya existentes que concreten la ruta metabólica a estudiar [65]. Entre sus limitaciones, se encuentran las restricciones de cada técnica analítica, ya que la GC-MS se limita a detectar compuestos volátiles y la LC-MS compuestos orgánicos con carga, así como la relativa dependencia con otras -ómicas para complementar los hallazgos [66].

### 3. ETAPAS DETERMINANTES EN LA ADQUISICIÓN DEL MICROBIOMA INTESTINAL

En general, la microbiota es particular de cada individuo, incluso al nivel de especies bacterianas [67]. Durante su establecimiento, sufre fluctuaciones y adquiere complejidad en su diversidad, con dominancia de géneros como *Faecalibacterium*, *Dialister*, *Roseburia*, *Ruminococcus* y *Bifidobacterium*. Alrededor de los tres años de vida, las variaciones interindividuales disminuyen y se alcanza una composición estable y parecida a la microbiota de adultos [68]. La etapa neonatal e infantil supone una ventana crítica en el desarrollo a nivel metabólico, inmunitario y neurológico, principalmente. En esta ventana, se adquiere plasticidad en los niveles de expresión génica en respuesta a los cambios y perturbaciones del ambiente mediante mecanismos epigenéticos y se apunta a que la adquisición y maduración de la microbiota juega un papel importante en este proceso [69].

La teoría más aceptada sobre la primera colonización de nuestro organismo sostiene que ocurre durante el parto y, sobre todo, en los primeros momentos después del nacimiento [70]. Se han detectado microorganismos, aunque en pequeña abundancia, en el meconio, la primera deposición del neonato que suele ocurrir en las primeras 48 h [71]. Las sucesivas deposiciones reflejan que el recién nacido es colonizado por bacterias cuyo parecido más próximo parece ser la microbiota intestinal de la propia madre [72]. Esto apunta a una transmisión vertical madre-hijo, sobre todo, gracias a la leche materna a través de la ruta enteromamaria, una importante fuente de bacterias, virus y componentes del sistema inmunitario. Posteriormente, bacterias de la familia *Clostridiaceae*, con capacidad para formar esporas y resistir las condiciones ambientales, se transfieren del entorno y otros familiares al lactante. A continuación, se exponen las etapas vitales que determinan la composición de la microbiota neonatal e infantil, mencionando algunos de los factores más importantes (Figura 5).



**Figura 5. Factores que influyen en la composición del microbioma neonatal y bacterias asociadas a las diferentes variables. ITU: infección del tracto urogenital.**

### 3.1. Embarazo

Se ha demostrado que la ganancia de peso, la dieta, el consumo de antibióticos, el consumo de tabaco y la composición de la microbiota de la madre (en especial, la intestinal y la vaginal) durante el embarazo influyen en el perfil de la microbiota del recién nacido en mayor grado que otras variables ampliamente estudiadas como el tipo de parto [73], [74]. La diabetes gestacional y la obesidad materna se han asociado al riesgo incrementado de enfermedades metabólicas en el recién nacido, como diabetes tipo 2 o el hígado graso no alcohólico, ya que se han visto en estos casos niveles de citocinas neonatales incrementados, lo que apunta a una desregulación del sistema inmunitario innato mediada por la microbiota. Estos recién nacidos presentan niveles bajos de *Lactobacillus* y altos de la familia *Lachnospiraceae*, un conocido biomarcador de control glucémico alterado en adultos con diabetes tipo 2 [75].

### 3.2. Parto

El retraso madurativo de los embarazos prematuros va ligado, con frecuencia, a un eje intestino-cerebro alterado, un desequilibrio en la microbiota con poca diversidad alfa, niveles altos de *Staphylococcus* y bajos de *Bifidobacterium*, y diversas morbilidades que se expondrán en otros apartados (ver 4.1) [76]. Los recién nacidos por cesárea presentan una microbiota intestinal diferente a la de los nacidos por parto vaginal. *Propionibacterium* y *Streptococcus*, característicos de la microbiota cutánea de la madre, son los primeros gérmenes en aparecer, mientras que los nacidos por parto vaginal se colonizaban con *Lactobacillus* procedentes de la vagina y representantes de la microbiota intestinal materna. Sin embargo, con frecuencia no se tenía en cuenta el fundamento de la intervención de la cesárea (elegida o de emergencia), lo que parece ser un factor diferencial, ni tampoco otros factores asociados como el uso de antibióticos intraparto, considerándose erróneamente al parto por cesárea como principal responsable de una microbiota aberrante en estos recién nacidos. Las diferencias entre los dos tipos de parto no parecen prolongarse en el tiempo más allá del primer año [77], [78].

Sólo determinadas bacterias pueden adaptarse al ambiente hospitalario y a los utensilios y las superficies de las UCI neonatales (UCIN). Generalmente, resisten la presión antibiótica y son portadoras de genes de resistencia, como *Klebsiella*, *Staphylococcus* y miembros de la familia *Enterobacteriaceae*, que pueden acabar colonizando el intestino de los recién nacidos [79], [80]. Otras bacterias del entorno, como *Gemella*, *Halomonas* o *Acinetobacter*, se consideran ambientales y comensales típicas de estos entornos [81].

### 3.3. Lactancia y etapa infantil

La leche materna contiene una microbiota particular que es transferida, junto con microbiota procedente de la piel y de la areolas mamarias, durante la lactancia. El origen de la microbiota de la leche no se conoce en profundidad, pero se apunta a que existe una retroalimentación entre la microbiota de la piel de madre y lactante, así como una vía de traslocación en la madre intestino-mamaria, también conocida como ruta enteromamaria, anteriormente nombrada. Esto explica la presencia de bacterias anaerobias obligadas como *Bifidobacterium* y también de comensales de la piel como *Streptococcus* y *Staphylococcus* en la leche [82]. Por otro lado, en torno al 20% de los hidratos de carbono de la leche son oligosacáridos no digeribles por las enzimas humanas que sirven de

sustrato energético para las bacterias intestinales, que los fermentan y producen AGCC. La composición de la leche materna varía entre las mujeres y depende del tiempo, la dieta y el genotipo. Además de proporcionar estos sustratos, la leche materna es inmunológicamente activa, ya que contiene inmunoglobulinas secretadas (IgA) y compuestos antimicrobianos que pueden guiar el desarrollo de la microbiota intestinal y el sistema inmunológico del lactante. En consecuencia, la microbiota intestinal de los lactantes alimentados exclusivamente con leche materna difiere de los otros, tanto en su composición como en su función [45]. Actualmente, se aconseja llevar una lactancia materna exclusiva para beneficiarse de sus múltiples propiedades hasta los seis meses de vida, seguida de la introducción de alimentos sólidos y fórmula de continuación hasta los dos años [83]. No obstante, también existen fórmulas suplementadas con prebióticos (compuestos que sirven de sustrato energético para las bacterias mutualistas y comensales) y probióticos (bacterias beneficiosas sin efecto patogénico como cepas comerciales de *Bifidobacterium* y *Lactobacillus*), así como bancos de leche en hospitales públicos para el acceso universal a una lactancia de calidad [84].

El genotipo humano también es un factor que influye en la composición de la microbiota. Los genes maternos determinan la estructura de los oligosacáridos de la leche, y esto tiene un impacto en la forma en que la microbiota infantil responde a la lactancia [45]. Otros genes, como los de la persistencia de la lactasa (que permiten que se produzca esta enzima más allá del periodo de lactancia materna) y los que se encargan de la fucosilación de la mucina y los oligosacáridos de la leche, afectan a la composición de los glicanos del intestino y, por tanto, también influyen en la composición de la microbiota intestinal. Se ha asociado el genotipo de la inmunidad innata del huésped, la composición de la microbiota y la gravedad de la enfermedad inflamatoria intestinal en pacientes pediátricos [85]. La FQ, como se desarrollará más adelante (ver 4.2), también tiene una microbiota alterada desde la etapa pediátrica, lo que acaba favoreciendo la colonización por distintos patógenos [86].

La ingesta de antibióticos a edades tempranas afecta tanto a la composición de la microbiota comensal como a la capacidad del sistema inmunitario de producir sustancias antimicrobianas e inmunoglobulinas. Además, una posología inadecuada puede favorecer la aparición de mecanismos de resistencia [69]. En el manejo de la FQ, a menudo se administran antibióticos a largo plazo sin evidencia de una infección pulmonar crónica y sin estar ligado a una mejoría de la enfermedad, lo que puede resultar perjudicial para el

microbioma y la salud a largo plazo [87].

Finalmente, se ha comprobado que ciertos factores ambientales como la calidad del aire, del agua y de los alimentos, así como el empleo de cosméticos y productos de higiene y sanitarios tienen un impacto en la adquisición y el establecimiento de la microbiota [87]. Los factores socioculturales o materiales, como el hecho de vivir en un país de bajos ingresos, afectan al desarrollo biológico y psicológico infantil, lo que puede impactar en su microbiota [88]. De modo similar, se ha postulado una mayor vulnerabilidad de la microbiota en países industrializados, donde los alimentos procesados y la excesiva higienización han favorecido una pérdida de diversidad y resiliencia microbiana [89].

### **3.4. Perspectivas futuras**

Hace unos años, se pensaba que las alteraciones que sufre la microbiota intestinal en el periodo neonatal podían programar la composición del microbioma de por vida [90]. No obstante, una revisión reciente cuestiona esta teoría debido a la falta de certezas, puesto que todavía no contamos con muchas cohortes humanas grandes seguidas por décadas, existiendo disparidad de resultados entre las investigaciones [91]. Otros estudios apuntan a que las principales fuerzas que impulsan los cambios en la microbiota en la etapa adulta son los movimientos migratorios y la zona de residencia, teniendo en cuenta el efecto de la dieta [92], [93]. De todas formas, no se cuestiona que los cambios inducidos en la etapa neonatal afectan a la capacidad de la microbiota para interactuar con los tejidos y órganos durante esta ventana crítica del desarrollo, ya que existen numerosas asociaciones entre una microbiota neonatal alterada y diversas patologías, tanto en esa etapa como en la adulta [69].

## 4. PATRONES DIFERENCIALES EN EL MICROBIOMA NEONATAL

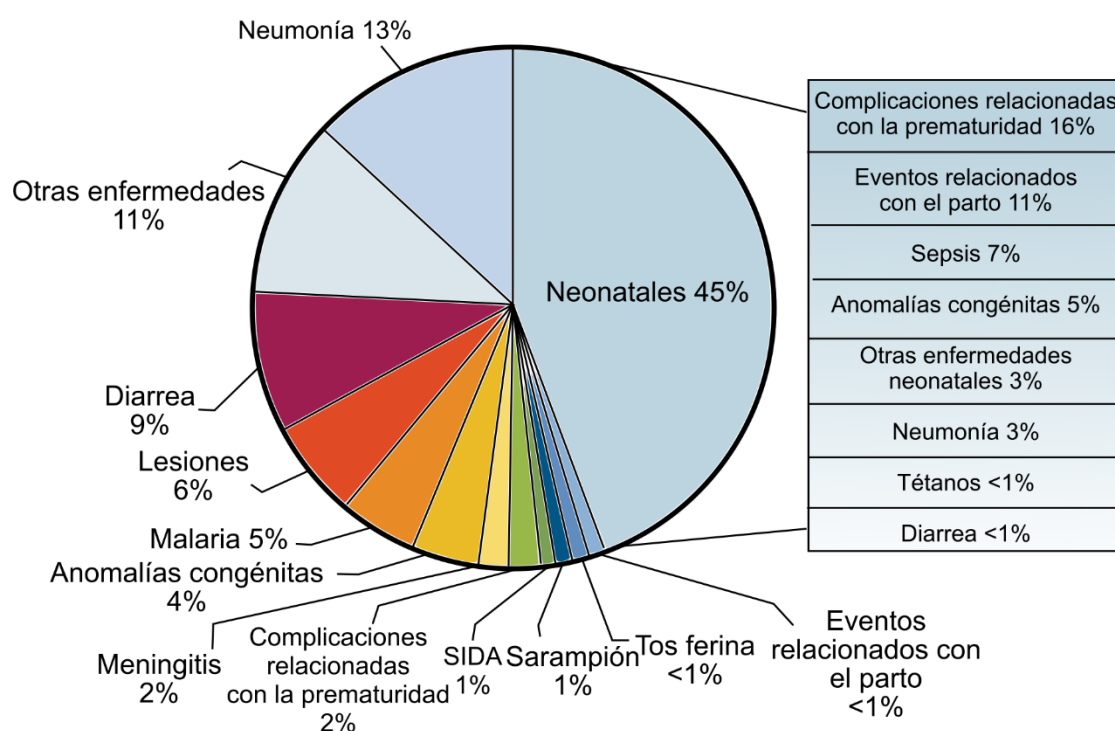
### 4.1. Prematuridad

#### 4.1.1. Definición, factores de riesgo y de asociación

El parto puede clasificarse en función de la edad gestacional en parto a término (de 37 a 42 semanas) o prematuro, que a su vez puede dividirse en extremadamente prematuro (<28 semanas), muy prematuro (<32) y moderadamente prematuro (entre 32 y 37 semanas). El parto prematuro se diferencia en espontáneo (parto espontáneo o parto que sucede a una ruptura de membranas prematura) e inducido debido a factores de riesgo en la madre o el feto, o por otras razones no médicas. El parto prematuro espontáneo es un proceso multifactorial, cuyos factores asociados son: tener antecedentes de parto prematuro, edad materna temprana o avanzada, poco tiempo entre embarazos, bajo índice de masa corporal, embarazos múltiples, antecedentes de enfermedades infecciosas, estrés, consumo de alcohol y tabaco y enfermedades periodontales, entre otros. En cuanto al parto prematuro inducido, algunas de sus causas son: la preeclampsia severa (hipertensión asociada a proteinuria), la ruptura de la placenta, la ruptura uterina, la colestasis (presencia de bilis en el tracto intestinal), el estrés fetal o la restricción del crecimiento fetal [94]. La corioamnionitis (inflamación de las membranas coriónicas y del saco amniótico) y la fusinitis (inflamación del cordón umbilical) están asociadas a los partos extremadamente prematuros. Se postula que las bacterias que suelen causar esta inflamación ascienden desde el tracto urogenital inferior, llegan por vía hematógica, o bien se adquieren durante alguna toma de muestras. Pueden acceder durante la ruptura de membranas y también atravesar las membranas intactas [95]. Asimismo, la obesidad y la diabetes materna se han registrado como dos importantes factores que contribuyen a la prematuridad en los partos a nivel global. Se ha observado que los varones tienen mayor riesgo de prematuridad [96].

Las complicaciones asociadas a la prematuridad constituyen la primera causa de muerte neonatal y la segunda causa conocida de muerte infantil en población menor de cinco años en todo el mundo (Figura 6) [97]. Cada año nacen 15 millones (11% del total) de bebés prematuros [98]. En España, se estima que la incidencia de prematuridad es del 8,3%, y la mortalidad en ese grupo tiene una tasa del 2,8%. A corto plazo, las complicaciones son generalmente respiratorias o relacionadas con la alimentación [99]. Con frecuencia, ingresan en UCIN por largos periodos de tiempo, estando el ingreso y el tiempo de estancia inversamente correlacionados con la edad gestacional. Allí, suelen

recibir tratamiento antibiótico, respiración artificial asistida, alimentación parenteral o fórmulas enriquecidas [70]. Alrededor del 50% de los neonatos prematuros nacen por cesárea, y tienen menores tasas lactancia materna respecto a los nacidos a término (47% vs. 60-70%) [99], y alrededor del 25% sufren problemas en el desarrollo neurológico [100]. A medio y largo plazo, la prematuridad está asociada a mayores tasas de enfermedades crónicas como la diabetes o la hipertensión [98]. El impacto de la prematuridad alcanza la edad juvenil, ya que se ha observado una mayor probabilidad de morir entre los 18 y 36 años en adultos nacidos prematuramente en comparación con los nacidos a término.



**Figura 6. Causas de muerte en menores de cinco años en todo el mundo (2016).** Traducido de Liu L, et al. Levels and Causes of Mortality under Age Five Years. In: Black RE, Laxminarayan R, Temmerman M, Walker N, editors. Reproductive, Maternal, Newborn, and Child Health: Disease Control Priorities, Third Edition (Volume 2). Washington (DC): The International Bank for Reconstruction and Development / The World Bank; 2016 Apr 5. Chapter 4. PMID: 27227226. Licencia libre.

#### 4.1.2. Microbioma y morbilidad en recién nacidos prematuros

Los factores desencadenantes del parto prematuro influyen en la adquisición de la microbiota neonatal y ésta, a su vez, se ha relacionado con las complicaciones anteriormente nombradas. Por un lado, las infecciones que ocurren durante el desarrollo fetal y la última etapa del embarazo prematuro, aunque también pueden darse en

embarazos a término, tienen una respuesta inflamatoria, una carga bacteriana y una durabilidad mayores. Esto puede desencadenar finalmente una sepsis; es decir, una infección bacteriana que alcanza el torrente sanguíneo y desencadena una respuesta inmunitaria desmesurada, cuya incidencia es de 2,27 a 5,14/1.000 neonatos prematuros. La sepsis es una de las principales causas de fallecimiento neonatal y está asociada a la prematuridad (Tabla 1) [95].

**Tabla 1. Limitación del esfuerzo terapéutico y causas de fallecimiento neonatal en función de la edad gestacional en España (2018). % (RQ).**  
Fuente: Sociedad Española de Neonatología. Informe anual SEN1500 2018.

	Edad gestacional al nacimiento (semanas)					
	Total N=286	<24 N=32	24-25 N=124	26-27 N=74	28-29 N=27	≥30 N=29
<b>Limitación del esfuerzo terapéutico</b>	45 (25-67)	63 (0-100)	46 (0-81)	36 (0-83)	42 (0-100)	45 (0-100)
<b>Causas de fallecimiento</b>						
Respiratoria	35 (0-50)	30 (0-100)	36 (0-56)	47 (0-100)	27 (0-46)	10 (0-0)
Sepsis	22 (0-34)	20 (0-38)	24 (0-45)	18 (0-33)	27 (0-63)	24 (0-50)
Neurológica	18 (0-27)	23 (0-50)	18 (0-14)	14 (0-6)	15 (0-38)	24 (0-42)
Malformación	3 (0-0)	3 (0-0)	1 (0-0)	3 (0-0)	-	17 (0-33)
Otras	19 (0-37)	23 (0-8)	19 (0-33)	14 (0-33)	27 (0-50)	24 (0-42)
Desconocida	3 (0-0)	-	3 (0-0)	6 (0-0)	4 (0-0)	-

La sepsis neonatal puede clasificarse en sepsis de inicio temprano (SIT), si ocurre en los tres primeros días, o sepsis de aparición tardía (SAT), si ocurre a los 7-10 días. A las madres se les suele administrar antibioterapia empírica profiláctica para prevenir la SIT, lo que inevitablemente desestabiliza la microbiota neonatal y predispone a la SAT. La adquisición de estreptococos del grupo B (EGB) y de *E. coli* durante el proceso del

parto se ha identificado como la principal causa de SIT en bebés prematuros (Tabla 2) [101]. Mediante técnicas de secuenciación se ha identificado que, justo antes de que ocurra una SAT, hay un aumento brusco de *Proteobacteria* y menores proporciones de *Bifidobacterium*, siendo las especies mayoritarias tanto Gram positivas (*Staphylococcus*, *Streptococcus* y *Enterococcus*) como Gram negativas (*E. coli*, *Klebsiella* y *Serratia*) [101]. Se ha comprobado el origen intestinal de las bacterias causantes de la SAT [102].

**Tabla 2. Morbilidad infecciosa neonatal en 2018. % (RQ).** Fuente: Sociedad Española de Neonatología. Informe anual SEN1500 2018.

	Edad gestacional al nacimiento (semanas)					
	Total N=2168	<24	24-25	26-27	28-29	≥30
Total de no fallecidos en sala de partos	N=2279	N=40	N=276	N=395	N=541	N=1027
<b>Infección neonatal precoz</b>	4,1 (0-6,9)	10,8 (0-0)	10,0 (0-17,6)	4,6 (0-0)	4,6 (0-8,4)	1,9 (0-0)
Total de ingresados más de 72 h	N=2168	N=25	N=236	N=370	N=527	N=1010
<b>Infección neonatal tardía</b>	21,9 (10,7-25)	39,1 (0-100)	52,3 (31,2-72,3)	38,9 (11,8-60,6)	17,7 (0-25)	10,5 (0-13,8)
Estafilococo coagulasa negativo	15,8 (5,5-25)	12,5 (0-0)	34,1 (0-50)	28,0 (0-50)	13,0 (0-24)	8,2 (0-10,8)
Hongos	1,3 (0-1,6)	4 (0-0)	4,4 (0-0)	2,9 (0-0)	0,8 (0-0)	0,1 (0-0)
Por otros patógenos	10,9 (4,2-13,8)	34,8 (0-62,5)	29,8 (0-50)	18,8 (0-34,3)	7,6 (0-12,7)	4,8 (0-6,4)
Más de un patógeno	5,4 (0-7,9)	13,0 (0-6,3)	15,4 (0-33,3)	10,1 (0-14,3)	3,0 (0-0,8)	2,4 (0-2,2)

El entorno hospitalario de las UCIN también impacta en el microbioma y metaboloma de los prematuros ingresados en su primer año de vida [103]. Los microorganismos colonizantes se han encontrado primero en las superficies abióticas, siendo los lavabos, las incubadoras y los tubos nasogástricos los reservorios de

propagación más comunes [79].

Se estima que alrededor del 38% de los brotes infecciosos en UCIN son causados por bacterias ambientales, siendo las más frecuentes *Klebsiella pneumoniae*, *Serratia marcescens*, *Staphylococcus aureus*, *E. coli* y el complejo de *Burkholderia cepacia*. El foco de origen suele ser algún neonato transferido de otra unidad o centro hospitalario, o procedente del entorno comunitario [79], [104]. La transmisión ocurre a través del material quirúrgico, el personal médico y familiares, el agua y la leche contaminadas. Las bacterias multirresistentes a los antibióticos son un problema importante en las UCIN, siendo la producción de beta lactamasas de espectro extendido (ESBL) el mecanismo de resistencia más común [105]. La colonización/infección por estos patógenos tendrá repercusiones en toda la microbiota en su conjunto y en el sistema inmunitario inmaduro, además de que la administración de antibióticos y la ausencia de bacterias comensales protectoras retrasará y dificultará la erradicación de estos patógenos hospitalarios.

Los prematuros tienen un patrón de colonización diferente al de los nacidos a término. Habitualmente se detectan bacterias en el meconio, pero con una baja diversidad [106]. En general, existen grandes fluctuaciones tempranas y un retraso en la colonización de bacterias anaerobias obligadas [107]. Habitualmente, se han detectado proporciones aumentadas de *Bacilli* (*Enterococcus*) y *Gammaproteobacteria* (*Enterobacter*, *Klebsiella* y *E. coli*), seguidas de una colonización tardía y a veces pobre de *Clostridia*, *Lactobacillus* y *Bifidobacterium*. Los filos *Bacteroidota* y *Actinobacteriota* son minoritarios, el filo *Firmicutes* disminuye con el tiempo, mientras que el filo *Proteobacteria* aumenta [108]. Estos patrones de colonización, así como la baja diversidad, se han asociado a numerosas enfermedades, entre ellas la enterocolitis necrosante (ECN), que afecta al 5-10% de los prematuros. Finalmente, su microbiota alcanza cierta estabilidad a las 30-36 semanas de edad corregida<sup>1</sup> y parece que, al cabo del tiempo, se asemeja a la de los nacidos a término [101], [109].

---

<sup>1</sup> Edad corregida: edad del recién nacido si hubiera nacido a las 40 semanas de vida. Siempre es menor que la edad desde el nacimiento. Se aconseja emplearla sólo hasta los dos años de edad corregida.

## 4.2. Fibrosis quística

La FQ es una enfermedad genética causada por la presencia de diversas mutaciones en el gen codificante del regulador de la conductancia transmembrana de la fibrosis quística (CFTR), situado en el cromosoma 7 (7q31.2) y conformado por 27 exones. Se trata de una enfermedad autosómica de herencia recesiva mendeliana, de tal modo que es necesario que ambos progenitores sean portadores de un alelo del gen CFTR defectuoso. Dicho gen codifica un canal transportador de electrolitos, compuesto por 1.480 aminoácidos [110]. Actualmente, se conocen más de 2.000 mutaciones que afectan a este gen, aunque no todas son causantes de enfermedad, siendo la mutación F508del (delección de la fenilalanina de la posición 508) la más frecuente en el haplotipo caucásico. En España, más del 51% de los alelos mutantes en 2016 eran F508del (el 75% de los pacientes portaban esta mutación en homo o heterocigosis). Otras variantes frecuentes en nuestro país son G542X (6% de los alelos), R334W y N1303K (2,7%). Se han clasificado en 7 categorías en base al efecto que producen en la funcionalidad del canal transmembrana. Estas clases están relacionadas con la severidad y el pronóstico de la FQ, considerando a las clases I, II, III y VII como de alto riesgo y a las clases IV, V, y VI de bajo riesgo [111], [112].

Se trata de una enfermedad poco frecuente que afecta a 1/13.500 nacimientos en todo el mundo siendo, a su vez, la enfermedad genética autosómica recesiva más común en la población caucásica (1/3.500) [113]. Comenzó siendo una enfermedad letal para la infancia, pero con el manejo de la malnutrición a través de la suplementación de enzimas pancreáticas, la esperanza de vida de los pacientes ha mejorado considerablemente. El número de adultos con FQ en un país es un indicador de calidad del manejo de la enfermedad. Según el registro de la Sociedad Europea de la Fibrosis Quística, en 2020 en España se registraron 2.505 pacientes, siendo el 55% mayores de 18 años y cerca del 53% varones [114]. En países de ingresos altos, la mediana de la esperanza de vida está en torno a los 40 años [115], [116].

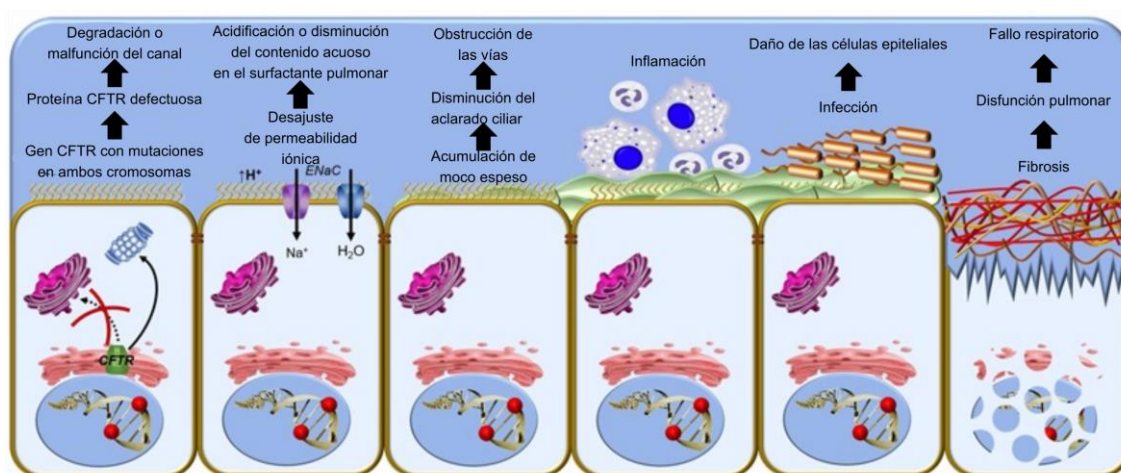
### 4.2.1. Fisiopatología, sintomatología y tratamientos

El canal CFTR es una proteína perteneciente a la superfamilia de los transportadores *cassette* ABC de unión a ATP que se encuentra en la membrana apical de todos los epitelios secretores del cuerpo, donde su principal función es regular la secreción de los iones cloruro. Por otro lado, posee un efecto inhibitorio sobre el canal apical encargado

de absorber iones sodio (ENaC) hacia el interior celular, y también regula el transporte de iones bicarbonato, clave para mantener la homeostasis (pH en torno a 7,1) del surfactante pulmonar y la viscosidad de otras secreciones [117].

En las glándulas sudoríparas, el canal CFTR se expresa, tanto en la membrana apical como en la basolateral, y se encarga de secretar iones cloruro durante la formación de sudor primario y de reabsorberlos en la formación del sudor final. Con un canal CFTR defectuoso, los iones cloruro no salen de las células por la membrana basolateral y tampoco se reabsorben, lo que genera que los iones sodio también queden retenidos en el lumen y se produzca un sudor con alto contenido en sales (>60 mEq/l) [117].

En el entorno pulmonar de FQ se da una retroalimentación entre la obstrucción por un mal aclaramiento mucociliar, la inflamación crónica y la colonización por parte de potenciales patógenos (colonización patogénica). El pH reducido inhibe a proteínas con capacidad antimicrobiana que mantienen la hidratación del moco [118], [119]. Además, la actividad de los neutrófilos se encuentra aumentada y esto, junto con la presencia de bacterias patógenas, aumenta la producción de especies reactivas del oxígeno, generando estrés oxidativo, daño tisular e inflamación. Cabe destacar que el canal CFTR también se expresa en neutrófilos y macrófagos, por lo que, a pesar de que el ambiente pulmonar esté inflamado y haya gran liberación de citocinas y quimiocinas, la inmunidad celular es defectuosa y no logra eliminar a los patógenos ni las consecuencias de la inflamación (Figura 7) [120].



**Figura 7. Mecanismos fisiopatológicos en las vías respiratorias de la presencia de un gen codificante de CFTR defectuoso.** Imagen traducida de Lopes-Pacheco M. (2016). CFTR Modulators: Shedding Light on Precision Medicine for Cystic Fibrosis. *Frontiers in Pharmacology*, 7, 275. Doi: 10.3389/fphar.2016.00275. Licencia Creative Commons.

Esto se traduce en una enfermedad multisistémica que afecta a las glándulas serosas sudoríparas y salivares, y a las glándulas y epitelios del aparato digestivo (páncreas, capa mucosa, hígado y vesícula biliar), reproductor y respiratorio. Se genera un proceso obstructivo e inflamatorio crónico que implica la existencia de diversas patologías y sintomatologías, como diarrea crónica, esteatorrea, enfermedad pulmonar obstructiva crónica (EPOC), insuficiencia pancreática exocrina, infertilidad, y pérdida de sal por el sudor o deshidratación hiponatrémica [117].

La principal causa de morbilidad y mortalidad está relacionada con la afectación pulmonar, en la cual la obstrucción bronquial, la infección y la inflamación conducen a la bronquiectasia y la lesión pulmonar progresiva [121]. Las exacerbaciones pulmonares están asociadas a mayores tasas de mortalidad y al empeoramiento de la función pulmonar. Los patógenos responsables suelen ser virus como el virus respiratorio sincitial o el de la gripe, y bacterias, siendo las principales *Pseudomonas aeruginosa*, *S. aureus*, el complejo de *B. cepacia*, *Achromobacter xylosoxidans*, *Stenotrophomonas maltophilia* y las micobacterias no tuberculosas. En pacientes jóvenes, la prevalencia de colonización/infección por *S. aureus* y *Haemophilus influenzae* es mayor. En general, las infecciones son polimicrobianas y acaban cronificándose, con patógenos cada vez más adaptados y con diversos mecanismos de resistencia antibiótica. El patrón de infección por *P. aeruginosa* suele tener tres fases: primoinfección, infección intermitente e infección crónica [122], [123]. El aislamiento de microorganismos patógenos requiere una amplia batería de medios, condiciones de cultivo y tiempos de incubación largos. Suelen aislarse a partir de muestras nasofaríngeas (lavado con suero salino) durante el primer año de vida, y también de muestra orofaríngea (con hisopos de garganta), lavado broncoalveolar (requiere anestesia) o esputo inducido en población de mayor edad [124].

El Registro de Pacientes de la Sociedad Europea de Fibrosis Quística (ECFSPR) apunta que la incidencia y prevalencia de *P. aeruginosa*, y el complejo *B. cepacia* en todos los grupos de edad disminuyó de 2011 a 2016, mientras que aumentó la de *S. aureus* [123]. En España en 2016, el 13% de los niños y el 38% de los adultos tenían colonización bronquial crónica por *P. aeruginosa* [116]. Existen muchos esfuerzos por identificar biomarcadores que anticipen estos episodios de recrudescimiento de la enfermedad y la colonización patogénica por estas bacterias. La terapia antibiótica en FQ es constante a lo largo de su vida, mientras que los nuevos tratamientos con los moduladores CFTR van destinados a corregir el problema de fondo. Actualmente, se han aprobado cuatro

medicamentos: Kalydeco® (ivacaftor), Orkambi® (lumacaftor e ivacaftor), Kaftrio® (ivacaftor, tezacaftor y elexacaftor) y Symdeko® (tezacaftor e ivacaftor) [125].

#### 4.2.2. Primeros años de vida con FQ

Desde el nacimiento, muchos pacientes con FQ presentan obstrucciones pancreáticas [126] e intestinales [127]. En el tracto intestinal, el déficit de secreción de bicarbonato al lumen puede provocar, en los recién nacidos, la formación de un meconio espeso, cargado de proteínas y pegado a la parte distal del íleo, lo que se conoce como íleo meconial. Es el síntoma más temprano en neonatos con FQ y comienza a desarrollarse *in utero*. Esta manifestación está asociada a las mutaciones de clase I-III del CFTR, y ocurre en, aproximadamente, el 20% de los recién nacidos con FQ [127]. En el hígado, también se generan secreciones biliares hiperviscosas, colestasia y cirrosis [128]. Si bien durante la infancia ya ocurren exacerbaciones respiratorias, la sintomatología pulmonar grave suele debutar en la adolescencia, con un rápido decaimiento de la función pulmonar y bronquiectasias precedidas por infecciones, muchas veces asintomáticas [129]. Se estima que en torno al 30% de los recién nacidos se colonizan de *S. aureus* en los primeros seis meses de vida. La primoinfección por *P. aeruginosa* suele ocurrir entre los tres meses y los dos años [130]. Las exacerbaciones en los primeros dos años se asocian a una caída en el volumen espiratorio forzado en el primer segundo (VEF<sub>1</sub>) y a una ganancia de peso insuficiente a los cinco años. El seguimiento de estos eventos en los primeros años con FQ es complejo, ya que niños sin FQ suelen experimentar también infecciones virales en las vías respiratorias altas [124].

La estrategia del cribado neonatal se basa en la medición de la inmunotripsina reactiva en muestra de sangre seca (normalmente, de una punción en el talón), la cual refleja los niveles de tripsinógeno sérico, que son altos en caso de que exista obstrucción ductal del páncreas. Asimismo, se miden los valores de cloro en sudor y se confirma la mutación/es de CFTR [131]. En España, el cribado neonatal está ampliamente instaurado, con un 91% de los pacientes menores de 5 años diagnosticados por cribado [116], lo que ha mejorado sustancialmente el abordaje temprano de la enfermedad. Si el cribado es positivo, se deriva al recién nacido a la Unidad de FQ de referencia para confirmar el diagnóstico, y se realizan visitas periódicas y pruebas complementarias, primero mensualmente y, a partir del sexto mes, cada 2-3 meses, dependiendo de la evolución. El abordaje en estas Unidades es multidisciplinar, e incluye a las áreas de Nutrición, Neumología, Digestivo, Psicología, y Fisioterapia de los hospitales de referencia [132].

La Fundación de FQ de Estados Unidos publica las guías consenso de diagnóstico (<https://www.cff.org/>), y la web CFTR2 registra variantes del gen CFTR y su relación con el grado de enfermedad (<https://cftr2.org/>). La prueba estandarizada para confirmar el diagnóstico de FQ se basa en los valores de cloro en sudor [131], y para entrar en el Registro Nacional español, se debe cumplir alguno de estos supuestos:

- a) Dos pruebas del sudor positivas;
- b) Una prueba del sudor positiva y dos mutaciones identificadas;
- c) Si no se cumple ninguna de las anteriores, al menos dos de estas condiciones:
  - i) Presencia de dos mutaciones causantes de enfermedad en base a la web CFTR2;
  - ii) Diferencia de potencial transepitelial nasal consistente con el diagnóstico de FQ;
  - iii) Presencia de rasgos clínicos típicos de FQ [116].

El cribado neonatal ha permitido reclutar a una mayor cantidad de pacientes para estudiar la implantación del microbioma, y durante una ventana de oportunidad en la cual la afectación respiratoria aún no es tan grave. Esto permite anticipar las intervenciones terapéuticas y arrojar luz sobre las interacciones entre la microbiota y la fisiología del hospedador.

### **4.2.3. Microbioma respiratorio e intestinal en la FQ**

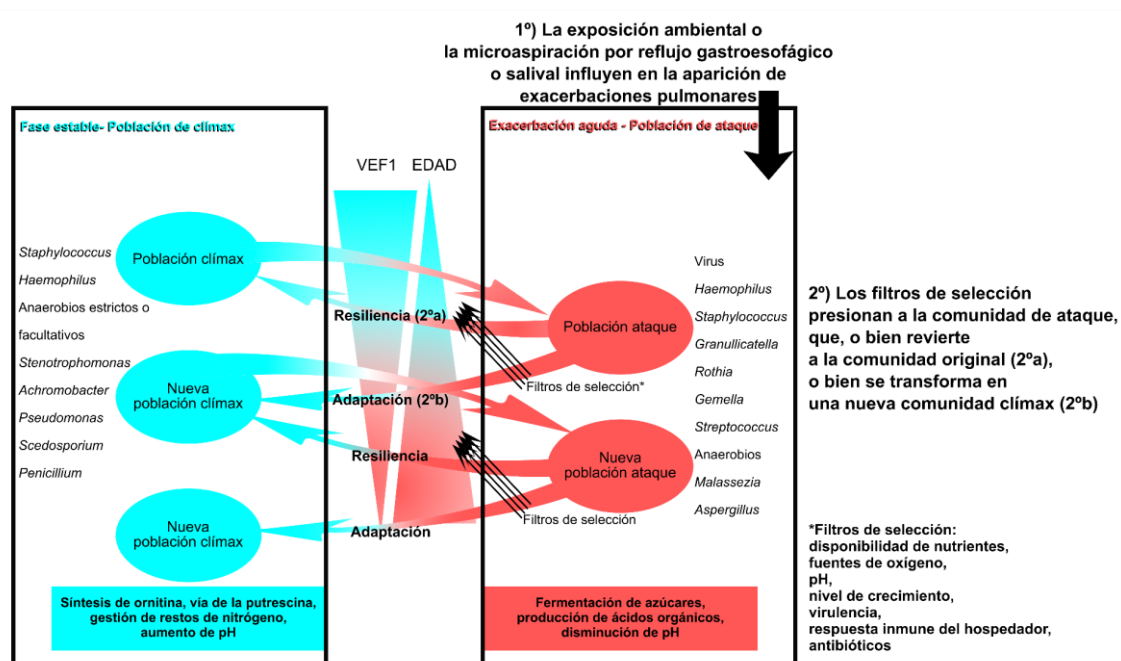
La enfermedad pulmonar en FQ progresa silenciosamente, primero con colonización patogénica, seguida de inflamación crónica, aparición de bronquiectasias y pérdida progresiva de la función pulmonar [133]. Ya en niños con 4 semanas de vida, se ha detectado inflamación en la vía aérea [134]. Las muestras nasofaríngeas suelen ser las empleadas para estudiar la microbiota de las vías respiratorias en pacientes menores de un año, aunque es cierto que se han encontrado diferencias cuando se ha comparado con la composición obtenida de esputo inducido, y puede que no reflejen las comunidades presentes en las vías respiratorias bajas [135].

En base a los estudios de metagenómica, se ha visto que la diversidad alfa en microbiota respiratoria es relativamente alta en recién nacidos y va decayendo con la edad. Se ha propuesto que, en adultos jóvenes, una baja diversidad en la microbiota de las vías respiratorias pueda ser un biomarcador de riesgo para la enfermedad [136]. Las

bacterias más abundantes en las vías respiratorias de lactantes con FQ son anaerobias; *Streptococcus* y *Haemophilus* compiten en ambientes de bajo pH, como es el caso del pulmón de la FQ, detectando otros como *Burkholderia*, *Prevotella*, *Veillonella*, *Gemella*, *Neisseria* y *Methylobacterium*. Por otro lado, los patógenos clásicos como *Pseudomonas*, *Stenotrophomonas* y *Staphylococcus* inicialmente están en bajas proporciones [86]. Los anaerobios tienen una función importante y, aunque por mucho tiempo se consideró que procedían de la contaminación oral de la muestra, existen evidencias de que esto no es así. Por último, a menudo se ha observado que no existe correlación entre la abundancia de *P. aeruginosa* detectada por metagenómica y el cultivo positivo [135].

La modificación ecológica de la microbiota podría cambiar el manejo terapéutico de esta enfermedad. En 2012 se propuso, en base a los estudios de redes ecológicas y de interacción, el modelo de ataque y clímax (Figura 8). Este modelo se basa en que el entorno pulmonar y su microbioma pueden actuar como un ecosistema con diferentes fases. Si recibe una perturbación breve, puede volver a ese estado posteriormente, pero si la perturbación se prolonga, el ecosistema cambia a otro modo de estabilidad diferente, con otra comunidad de microorganismos adaptada al nuevo estado. La ecología restaurativa intenta identificar la perturbación inicial, y generar nuevas alteraciones que restauren su estado. En el caso de la FQ, existiría una fase de ataque en la que existe una comunidad funcional microbiana asociada a las exacerbaciones pulmonares (con genes de virulencia que desencadenan la respuesta inmunitaria innata y la afectación de las vías respiratorias), y otra fase de clímax típica de los periodos clínicamente estables; es decir, de la fase crónica de la enfermedad (colonización permanente, respuestas inmunitarias adaptativas y baja susceptibilidad al tratamiento antibiótico). Esto se ha conceptualizado gracias a la aportación de las técnicas -ómicas al conocimiento sobre la funcionalidad de las comunidades [137]. En base a diferentes hallazgos, se ha propuesto que, en la fase de ataque, se encuentran anaerobios como *Veillonella*, *Lactobacillus*, *Prevotella*, *Streptococcus*, *Rothia* y *Granulicatella*, cuya actividad fermentativa de azúcares genera ácidos y contribuye a disminuir el pH del medio. Se ha comprobado en co-cultivos, que estos anaerobios producen aminoácidos y AGCC a partir de las mucinas que estimulan el crecimiento de *P. aeruginosa* y la producción de sus factores de virulencia. En la fase de clímax y con presión antibiótica, los anaerobios desaparecen y domina *P. aeruginosa* con fenotipo mucoide, la cual prefiere los compuestos nitrogenados (aminoácidos) como fuente de carbono. El producto del metabolismo de las comunidades clímax genera

amonio, lo que contribuye a aumentar el pH. Las cepas de *S. aureus*, a pesar de mostrar relaciones antagónicas de crecimiento con *P. aeruginosa* en cultivo, también se han asociado a la fase clímax. No se descarta que estas comunidades convivan en el mismo momento en nichos compartimentados [138], [139]. Por último, hay que destacar que la secuenciación *shotgun* ha permitido identificar diferentes linajes genéticos de *P. aeruginosa* en un mismo paciente a lo largo del tiempo, lo que podría estar detrás de esos cambios de estado en pacientes de enfermedad avanzada [140]. Es necesario seguir investigando la relación que tienen las bacterias anaerobias con el desencadenamiento de exacerbaciones pulmonares, la inflamación y los patógenos clásicos en recién nacidos con FQ.



**Figura 8. Representación del modelo de adaptación de clímax y ataque en FQ.** Modificado de Soret, P., et al. Respiratory mycobiome and suggestion of inter-kingdom network during acute pulmonary exacerbation in cystic fibrosis. *Sci Rep* 10, 3589 (2020). Doi: 10.1038/s41598-020-60015-4. Licencia libre.

Otros miembros de la microbiota respiratoria de la FQ de los que no se conoce mucho son los hongos, entre ellos *Aspergillus* o *Candida*, los virus bacteriófagos y las bacterias depredadoras (ver apartado 5). Estos dos últimos se consideran agentes de control ecosistémico *top-down*, implicados en la eliminación de poblaciones bacterianas y en la transmisión y reserva, en el caso de los fagos, de genes de resistencia. Se han encontrado en el tejido pulmonar y las secreciones de FQ [141], [142].

La microbiota intestinal y la respiratoria de lactantes con FQ evolucionan de manera similar a lo largo del tiempo, existiendo un núcleo común formado por *Veillonella*, *Streptococcus*, *Bifidobacterium* y *Bacteroides*. Además, hay bacterias que primero se enriquecen en intestino y aparecen más tarde en pulmón. La diversidad aumenta con el tiempo, y la lactancia materna impacta en la microbiota respiratoria. En adultos, la administración de probióticos orales en FQ se ha asociado a una disminución de las exacerbaciones [143] y la disminución de *Parabacteroides* en intestino se ha identificado como predictor de la colonización respiratoria por *P. aeruginosa*. En general, se han detectado proteínas bacterianas y humanas asociadas a la inflamación (calprotectina y lipocalina-2, entre otras) que podrían usarse como biomarcadores. para monitorizar la salud intestinal de pacientes con FQ [144].

## 5. INTERVENCIÓN TERAPÉUTICA BASADA EN LAS DINÁMICAS ECOLÓGICAS DEL MICROBIOMA

Los Postulados de Koch contribuyeron significativamente al entendimiento de la patogenicidad bacteriana. Sin embargo, con el desarrollo de las tecnologías -ómicas, es necesario revisar estos postulados respecto de los conocimientos actuales de microbiota y salud global [145]. Se ha demostrado la resistencia de la microbiota a la colonización por patógenos exógenos activando el sistema inmunitario [146] aunque, del mismo modo, la microbiota puede facilitar el proceso de infección [147]. Por tanto, el curso de la enfermedad estará influenciada por factores del patógeno, el hospedador, la microbiota y el ambiente.

La resistencia a antibióticos es un problema global y se hace urgente encontrar alternativas que acaben con las bacterias multirresistentes y con la transferencia de estos genes de unas a otras a través de diversos mecanismos. La visión holística de que el microbioma puede actuar como un órgano más del cuerpo abre caminos para intervenir en su composición y funcionalidad en la lucha contra las infecciones. Además, reconocer los microorganismos que forman parte de la microbiota y sus mecanismos de colaboración y competencia puede dar pistas para adquirir una visión ecológica, en la que se empleen componentes metabólicamente activos para combatir otros y eliminar el círculo vicioso de las resistencias.

### 5.1. Transferencia de microbiota fecal (TMF)

Este término hace referencia a la administración de una suspensión de microbiota fecal procedente de una persona saludable a un receptor con la intención de tratar una patología, vía gastroscopia, colonoscopia, o mediante cápsulas orales [148]. Actualmente, su única indicación es la infección recurrente por *Clostridioides difficile* (ICDr), una de las principales causas de diarrea comunitaria y nosocomial, con una elevada morbilidad y mortalidad en pacientes de edad avanzada [149], pero se ha utilizado en otras enfermedades intestinales con resultados variables [150][151].

Desde un punto de vista ecológico, el éxito de esta práctica ocurre cuando un ecosistema equilibrado del donante coloniza el tracto gastrointestinal del receptor, por competencia de nicho y otras ventajas ecológicas. Sin embargo, el éxito clínico no se logra siempre [152]. Se ha observado mejora clínica en pacientes con ICDr que recibieron filtrado de heces sin contenido bacteriano, lo que apunta a que el beneficio puede deberse

a la transferencia de metabolitos, agentes antimicrobianos, restos de ADN o fagos [153], [152]. En contraposición a la hipótesis de la existencia de “súper-donantes”, se piensa que factores como la redundancia funcional de la microbiota, la compatibilidad donante-receptor (como, por ejemplo, la tolerancia inmunológica del receptor a la materia fecal del donante), la microbiota preexistente a la recepción de TMF, las diferencias genéticas entre donante y receptor, el tratamiento antibiótico y la dieta del receptor, tienen más importancia en relación con el éxito de la intervención, y deben evaluarse y tenerse en cuenta antes de extraer conclusiones [154].

## **5.2. Terapia con bacteriófagos**

### **5.2.1. Papel en la función del microbioma**

Los virus bacteriófagos que infectan bacterias son unos habitantes más de la microbiota intestinal de los mamíferos, y forman el grupo más mayoritario dentro del viroma: el fagoma [155], [156]. Si bien las interacciones entre bacterias y virus suelen seguir las dinámicas de presa y depredadora con interacciones líticas (el virus rompe y mata a la bacteria), se ha comprobado que la mayoría del fagoma humano está compuesto por bacteriófagos lisogénicos o profagos; es decir, la fase viral en la que el material genético del virus se encuentra integrado en el genoma bacteriano [157]. Se desconocen las dinámicas que activan la fase lítica. Por otro lado, se ha comprobado su función en el proceso de adaptación del bacterioma al estrés o los antibióticos, ya que parecen actuar como reservorio de genes de resistencia a antibióticos [158]. En general, los fagos se consideran moduladores de las poblaciones microbianas [156].

### **5.2.2. Aplicaciones clínicas**

Su uso como agente antibacteriano se popularizó el siglo pasado, pero con el desarrollo de los antibióticos y los desafíos que planteaba el trabajo con fagos, la investigación cesó. Recientemente, han resurgido con fuerza [159] y, por ejemplo, se han empleado para erradicar a patógenos con genes de resistencia a carbapenems como *K. pneumoniae* en modelos de ratón [160]. También se ha observado que la terapia con fagos es relativamente exitosa para acabar con la infección crónica de una cepa multirresistente de *P. aeruginosa* procedente de paciente con FQ en un modelo de ratón, y que su uso profiláctico puede prevenir la infección crónica [161].

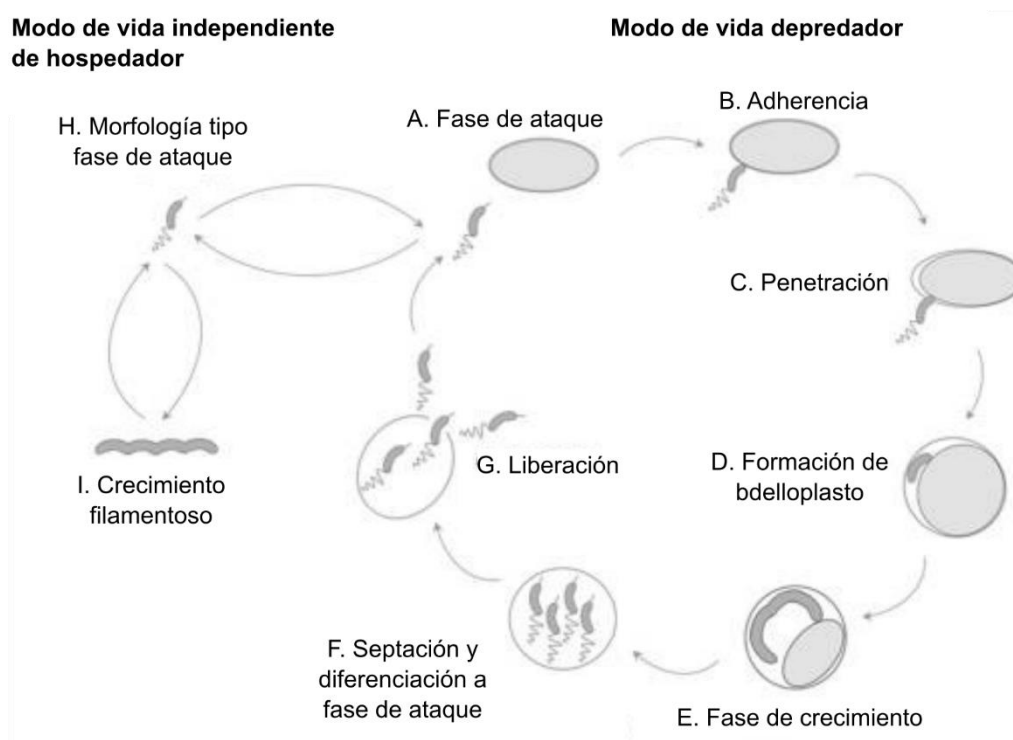
No obstante, existen algunas limitaciones, y los ensayos aún se encuentran en fase preclínica. Además, los resultados en cocultivos no se reflejan en algunos modelos *in*

*vivo*, sobre todo en infecciones intestinales, debido a la dificultad de que los fagos alcancen la diana en suficiente número [162]. Otros desafíos incluyen la posibilidad que las bacterias desarrollen resistencia a la infección, la interacción con el sistema inmunológico del hospedador (que puede reaccionar y eliminar el fago circulante, o generar una respuesta inmune contraproducente), o la expresión de componentes del fago en células eucariotas humanas [163].

### 5.3. Bacterias depredadoras

#### 5.3.1. Descubrimiento y características

Las bacterias depredadoras se caracterizan por tener un tamaño reducido y porque su ciclo vital pasa por depredar a otras bacterias (presas), de una manera similar a los fagos. Su contenido genético codifica, sobre todo, enzimas líticas destinadas a romper la pared celular y la membrana de sus presas. Existe un amplio rango de bacterias susceptibles de ser depredadas [164]. La especie más conocida de bacterias depredadoras es *Bdellovibrio bacteriovorus*, una bacteria Gram negativa que depreda a otras bacterias Gram negativas y pertenece al filo recientemente renombrado como *Bdellovibrionota*. Se descubrió en 1963, mientras un grupo de investigadores trabajaba con bacteriófagos del suelo, y se ha encontrado en numerosos ecosistemas acuáticos y terrestres desde entonces [165]. Tiene forma vibrioide y un flagelo que le aporta motilidad. Es una bacteria depredadora obligada con un ciclo de vida bifásico, con una fase libre o de ataque en la que busca a una presa cercana, y otra fase de crecimiento dependiente de presa. En la segunda fase, se adhiere a la superficie, penetra en el periplasma y aprovecha los recursos de la presa para sintetizar sus componentes. Una vez ha alcanzado cierto número de duplicaciones, rompe a la bacteria por completo, la abandona y entra de nuevo en la fase libre (Figura 9). Las bacterias de géneros similares a *Bdellovibrio*, como *Peridibacter* y *Bacteriovorax*, se engloban en el término BALOs (*Bdellovibrio and like-organisms*). *Micavibrio aeruginosavorus*, una BALO perteneciente al filo *Proteobacteria*, se ha encontrado en otros ecosistemas depredando a *S. maltophilia* y a *P. aeruginosa*. Esta depredadora, a diferencia de *B. bacteriovorus*, depreda de manera epibiótica, adhiriéndose a la presa y consumiendo sus recursos desde fuera para crecer y multiplicarse por división binaria [166].



**Figura 9. Ciclo de vida bifásico de *B. bacteriovorus*.** Traducido de Rotem, O., et al. (2015). Cell-cycle progress in obligate predatory bacteria is dependent upon sequential sensing of prey recognition and prey quality cues. *Proceedings of the National Academy of Sciences of the United States of America*, 112(44), E6028–E6037. Doi: 10.1073/pnas.1515749112. Licencia libre.

La mayoría de las bacterias depredadoras conocidas se han aislado del ambiente [167]–[169]. En los últimos años, y gracias a las tecnologías de secuenciación, se han encontrado evidencias de su material genético en el intestino de mamíferos [170] y en la microbiota humana, como en el pulmón de pacientes con FQ [142]. Sin embargo, aún no se han recuperado viables.

### 5.3.2. Aplicaciones industriales y agrícolas

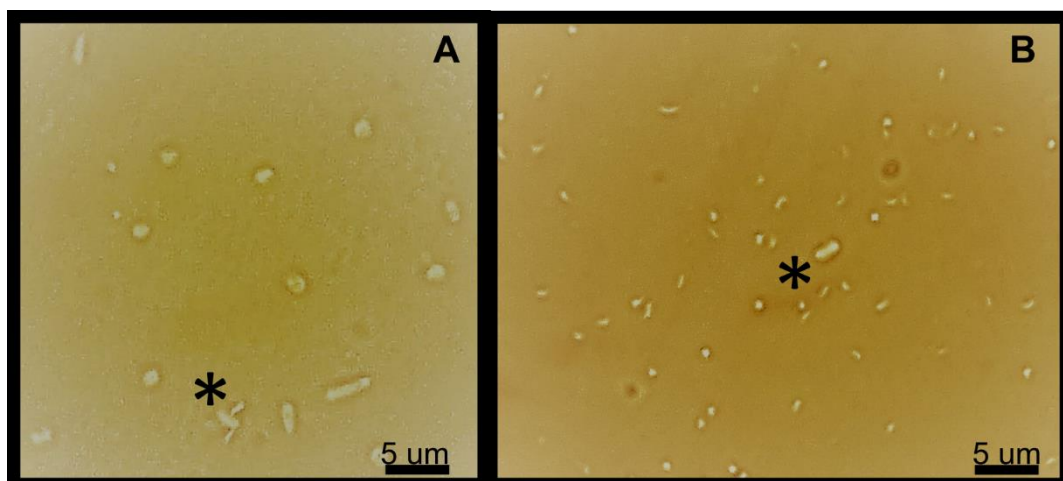
Debido a su gran diversidad enzimática, las bacterias depredadoras se han empleado en biotecnología para numerosas estrategias. *B. bacteriovorus* depreda típicamente a cepas de la especie *Pseudomonas putida*, las cuales poseen la capacidad de sintetizar polihidroxicanoatos (PHA), unos plásticos biodegradables con los que se crean polímeros para la industria. Se han mejorado diversas cepas de *B. bacteriovorus* para perforar a *P. putida* y extraer los plásticos con alto rendimiento [171].

*B. bacteriovorus* elimina fitopatógenos, *Salmonella* en granjas avícolas y erradica la contaminación de *V. cholerae* en piscifactorías. Por otro lado, se sabe que las plantas

de aguas residuales son un reservorio de genes de resistencia a antibióticos, los cuales suelen estar en la forma de ADN libre o en bacteriófagos. Se ha demostrado que *B. bacteriovorus* puede contener los fenómenos de transferencia horizontal de genes de resistencia mediante la eliminación, no sólo de las bacterias recombinantes que adquieren estos genes, sino también de los genes libres, gracias a que liberan gran cantidad de proteasas y nucleasas al exterior. Por último, también son capaces de degradar biopelículas en las maquinarias y en el instrumental quirúrgico, un gran problema, tanto en la industria alimentaria como en la práctica médica [172].

### 5.3.3. Aplicaciones clínicas

El fundamento del empleo de bacterias depredadoras como antimicrobianos vivos es el mismo que el de la terapia con fagos. Existen numerosos estudios *in vitro* realizados en cultivos bacterianos, suero humano y en biopelículas, que demuestran su eficacia para eliminar a una gran variedad de patógenos resistentes a antibióticos, incluidos los del grupo ESKAPE (*Enterococcus faecium*, *S. aureus*, *K. pneumoniae*, *Acinetobacter baumannii*, *P. aeruginosa*, y especies de *Enterobacter*) (Figura 10) [173]. Algunos ensayos en animales también han sido exitosos [174], [175]. No obstante, no se ha explorado en la especificidad de la depredación en función de las condiciones ambientales, así como la posibilidad de ampliar el rango y la eficacia de depredación modificando componentes de la depredadora, tal y como se ha hecho en aplicaciones industriales.



**Figura 10. Imágenes de microscopía de contraste mostrando células de *B. bacteriovorus*. A: depredando a un espécimen de *S. marcescens* (asterisco). B: una presa rodeada de depredadoras (asterisco).**

El empleo de BALOs tiene una serie de ventajas; las bacterias depredadoras no favorecen la transferencia horizontal de genes entre bacterias potencialmente patógenas, como es el caso de los fagos y, además, no dependen del azar de encontrar a una presa y unirse a un receptor, sino que, al tratarse de organismos metabólicamente activos, buscan a su presa [176]. Por otro lado, su producción a escala no depende del empleo de bacterias patógenas (como podría ser el caso de los fagos), si no que pueden emplearse bacterias inocuas. El LPS de *Bdellovibrio* no tiene el grupo fosfato cargado negativamente, lo que les impide unirse con tanta afinidad a los receptores de las células inmunes eucariotas, generando una respuesta inflamatoria mínima. Además, no presentan ninguna actividad sobre las células eucariotas [177]. Si bien se ha planteado la posibilidad de que se genere resistencia a los BALOs, no parece probable, puesto que la depredación ocurre de manera rápida y no se conoce hasta el momento un receptor específico de la presa que pueda ver su expresión disminuida o eliminada. En todo caso, al tratarse de “antibióticos vivos”, la generación de una resistencia en la presa generaría una presión selectiva en la depredadora para sobrellevarla, en línea con las dinámicas evolutivas [178].

Para la incorporación de BALOs a la práctica clínica, es necesario evaluar las posibles limitaciones y desafíos que conlleva. La estabilidad de estas bacterias, la correcta eliminación de las presas en cocultivo durante la fase de producción, la viabilidad y el alcance del lugar anatómico deseado son algunas de ellas. Además, aún se desconocen los mecanismos de interacción y de afinidad por determinadas bacterias a niveles concretos de cepa, y escasean las investigaciones en grandes colecciones clínicas [179]. Como se ha indicado en otros capítulos, las interacciones entre los componentes de la microbiota dependen, en muchos casos, de características específicas de cepa, y aún queda por determinar qué factores determinan el comportamiento polimórfico de la depredación [166], [180].



**HIPÓTESIS  
Y  
OBJETIVOS**



## HIPÓTESIS

La adquisición y el establecimiento del microbioma condiciona la salud del resto de nuestra vida e influye en la predisposición a patologías o el agravamiento de enfermedades de base. Conocer las dinámicas poblacionales de este ecosistema ayudará a identificar terapias alternativas a los antibióticos bajo una perspectiva ecológica. El estudio de la microbiota se puede abordar utilizando múltiples herramientas, pero es necesario identificar las fortalezas y debilidades de cada una. Estas tecnologías también han reformulado y ampliado los conceptos de brote, virulencia y la contribución de la microbiota comensal al desarrollo o no de enfermedades.

En esta tesis doctoral se profundiza en la ventana de oportunidad que ofrecen los primeros meses de vida para estudiar las complejidades del microbioma y su adquisición, así como la evaluación de diferentes técnicas para su análisis. La hipótesis principal de esta tesis es que la composición de la microbiota en recién nacidos varía en función de la técnica de estudio empleada, así como del contexto epidemiológico y la predisposición genética a ciertas enfermedades. Asimismo, se cuestiona el nivel de resolución de las diferentes técnicas que se emplean para el estudio de la microbiota, ya que el empleo de una u otra técnica de estudio deberá seleccionarse en base a qué preguntas queremos responder. Por último, en este trabajo se evalúa el empleo de bacterias depredadoras como recursos basados en la ecología microbiana para erradicar bacterias patógenas clásicas, utilizando estrategias de modulación de la microbiota ecológicas y alternativas a los antibióticos o a la transferencia de microbiota fecal.

### 1. OBJETIVOS GENERALES

- Determinar las diferencias y concordancias en la caracterización de la microbiota intestinal por secuenciación de amplicones y por metaproteómica, y las fortalezas de combinarlas en el estudio de la microbiota y su relación con el huésped en una población neonatal con FQ.
- Detectar la sucesión de clones bacterianos de la microbiota intestinal y respiratoria en lactantes con FQ durante dos años, relacionándola con la composición obtenida mediante secuenciación de amplicones, la colonización patogénica y las exacerbaciones clínicas.
- Comparar el establecimiento de la microbiota intestinal de dos cohortes neonatales ingresadas en UCIN en contextos epidemiológicos distintos.
- Comparar el contenido genético de aislados de *S. marcescens* responsables de brotes nosocomiales contemporáneos e históricos y actualizar la filogenia de la especie.
- Evaluar la capacidad depredadora de *B. bacteriovorus* y su mutante bd2637 con patógenos clínicos procedentes de las anteriores poblaciones pediátricas y adultas.

## 2. OBJETIVOS ESPECÍFICOS

### 2.1. Técnicas de estudio del microbioma:

- Comparar la asignación taxonómica y la detección de bacterias realizada mediante secuenciación de amplicones con la de metaproteómica en muestras fecales.
- Evaluar las diferencias cuantitativas en la abundancia relativa de los principales grupos bacterianos y la diversidad total calculada según la secuenciación de amplicones con la inferida mediante metaproteómica en las heces de una población infantil con FQ.
- Determinar el grado de concordancia entre la detección de patógenos de FQ realizada mediante secuenciación de amplicones en muestras fecales y respiratorias con el cultivo de muestras respiratorias positivo y clínicamente relevante.

### 2.2. Población pediátrica con FQ:

- Analizar las proteínas humanas asociadas a su microbiota fecal.
- Identificar los géneros bacterianos con cambios composicionales significativos a lo largo de los primeros años en muestras fecales y respiratorias.
- Establecer el núcleo compartido de bacterias en muestras fecales y respiratorias y estudiar los patrones de colonización conjunta.
- Estudiar las relaciones entre los cambios composicionales de las principales bacterias fecales y respiratorias con las variables clínicas recogidas durante los dos primeros años.
- Identificar y tipar las principales bacterias recuperables de heces y lavados nasofaríngeos y monitorizar su prevalencia a lo largo del tiempo, atendiendo a los patrones de similitud entre aislados.

**2.3. Población neonatal prematura ingresada en UCIN:**

- Identificar los géneros bacterianos con cambios composicionales significativos en meconio y heces de la tercera semana de vida, atendiendo a diferentes variables clínicas.
- Identificar y tipar las principales bacterias recuperables de meconio y heces y monitorizar su prevalencia a lo largo del tiempo, atendiendo a los patrones de similitud entre aislados, en una población ingresada en UCIN durante un brote nosocomial de *S. marcescens*.
- Comparar la composición de la microbiota fecal y su diversidad de un grupo ingresado durante un brote nosocomial con otro ingresado en otro momento en la misma UCIN sin brote.

**2.4. Colecciones de cepas patógenas aisladas de estos pacientes y de otras cohortes:**

- Determinar el pangenoma, el genoma núcleo y el accesorio, así como el resistoma y el contenido de fagos y secuencias de inserción de los aislados de *S. marcescens* responsables del brote nosocomial anteriormente nombrado y de otro muy similar ocurrido 47 años antes.
- Comparar el resistoma de esos aislados con la resistencia fenotípica a los principales antibióticos de elección.
- Construir un árbol filogenético a partir del cgMLST de estos aislados y todos los de esta especie depositados en bases de datos públicas.
- Analizar el rango de deprecación intraespecífico (I-PR), así como la eficiencia de deprecación (PE) de *B. bacteriovorus* 109J sobre colecciones de aislados clínicos de *S. marcescens* (de bacteriemias pediátricas y adultas) y *P. aeruginosa* (de FQ y de bacteriemias).
- Estudiar las diferencias significativas en el rango de deprecación intraespecífico y la eficiencia de deprecación entre *B. bacteriovorus* 109J y el mutante de un solo gen *B. bacteriovorus* bd2637 sobre la colección de *P. aeruginosa* aislada de bacteriemias.

# **CAPÍTULO 1**

## **ABORDAJE DEL ESTABLECIMIENTO DE LA MICROBIOTA INTESTINAL EN PACIENTES CON FIBROSIS QUÍSTICA MEDIANTE DIFERENTES TÉCNICAS ÓMICAS**

## *Metaproteómica y secuenciación de amplicones...*

*¿Detectan los mismos **grupos** bacterianos?*

*¿Son detectados en las mismas **proporciones**?*

*¿Existen **sesgos** que puedan ser explicados por el uso de cada técnica?*

*¿Puede ser **complementaria** la información obtenida por ambas?*

*¿Pueden detectar por igual los **cambios composicionales** en la microbiota intestinal de lactantes con FQ?*

*¿Son igual de sensibles a la hora de detectar **biomarcadores** asociados al paso del tiempo en esta cohorte?*

## I. ANTECEDENTES

El establecimiento de la microbiota al nacer condiciona en gran medida la salud del resto de nuestra vida [181] pero, en FQ, la colonización bacteriana patogénica de las vías respiratorias determina su supervivencia final [182]. La esperanza de vida de esta enfermedad ha aumentado considerablemente en los últimos años, apareciendo nuevas comorbilidades, principalmente del tracto digestivo [183]. La actividad metabólica de la microbiota intestinal tiene un impacto directo en el estado nutricional y en FQ ejerce un efecto importante en la enfermedad pulmonar infecciosa, estando la correcta nutrición inversamente relacionada con el número y gravedad de las exacerbaciones respiratorias [184].

La microbiota intestinal en FQ está condicionada por los siguientes factores: la disfunción del regulador transmembrana CFTR, que crea una capa de moco deshidratada y espesa, el entorno ácido causado por la secreción alterada de bicarbonato, la mala digestión, la malabsorción, la esteatorrea debida a la insuficiencia pancreática, y la elevada ingesta de antibióticos para tratar las infecciones bacterianas durante las exacerbaciones pulmonares [185], [186]. Numerosas publicaciones han señalado las diferencias significativas en la composición bacteriana y en la diversidad alfa de la microbiota de FQ en comparación con la de lactantes sanos [187]–[189].

Actualmente, la secuenciación de amplicones del gen codificante del 16S ARNr es la técnica más amplia, rápida y rentable para caracterizar y estudiar a nivel composicional comunidades bacterianas complejas como la microbiota intestinal, si bien no discrimina entre bacterias con actividad metabólica o muertas [190]. La principal limitación de esta técnica es la afinidad variable de los cebadores empleados, y así se ha notificado una sobrerrepresentación de ciertos géneros bacterianos, mientras que otros, principalmente el género *Bifidobacterium*, están sistemáticamente infrarrepresentados [191]. Incluso con cebadores optimizados, la secuenciación de amplicones suele proporcionar una asignación taxonómica deficiente a nivel de género y especie debido a la corta longitud de estas regiones. Por otro lado, la metaproteómica aplica la espectrometría de masas de alta resolución para la detección y cuantificación de las proteínas presentes en una muestra. Al ser las proteínas la unidad funcional básica de la mayoría de los procesos biológicos, la metaproteómica aporta, como enfoque novedoso respecto a la secuenciación de amplicones, la caracterización funcional dinámica de la microbiota. Asimismo, a través de la metaproteómica, se puede estudiar la microbiota a

nivel composicional, ya que existen patrones proteicos característicos de cada grupo bacteriano, y esto se observa a distintos niveles taxonómicos. La reciente introducción de herramientas bioinformáticas actualizadas, como MetaLab, ha mejorado considerablemente el análisis de los datos [63]; sin embargo, para las especies menos representadas, el análisis sigue siendo impreciso. Los datos disponibles sugieren que sólo se identifica entre el 10 y el 20% del contenido proteico total de una muestra clínica, lo que corresponde a un tercio de toda la microbiota [35]. Recientemente, se han sugerido estrategias adicionales para solventar los problemas de sensibilidad de los instrumentos de análisis y poder así acceder a las proteínas menos abundantes con importancia biológica, como realizar pasos de pre-fraccionamiento de las muestras [192]. Los puntos fuertes de la metaproteómica incluyen una alta resolución en la identificación a nivel de especie bacteriana, la visión general de la actividad funcional y metabólica de los microorganismos y, sobre todo, la detección de las proteínas del hospedador, lo que permite ampliar el conocimiento de las interacciones entre el huésped y el microbioma [193], [194], algo que resulta especialmente interesante en la FQ [144]. No obstante, la metaproteómica también tiene importantes limitaciones [195], siendo necesaria la estandarización de la técnica aplicada a estudios de microbiota fecal humana para que éstos sean reproducibles, así como la integración con las otras plataformas -ómicas para una caracterización profunda.

La combinación de las técnicas de secuenciación de amplicones y de metaproteómica proporciona una caracterización más amplia de las comunidades microbianas y de su interacción con el huésped [196]–[198], y la integración de los resultados podría contribuir a realizar predicciones precisas [199]. Sin embargo, pocos estudios han incluido un análisis estadístico para comparar los resultados de ambas metodologías, especialmente de muestras difíciles de obtener como las de neonatos, caracterizadas por presentar una cantidad limitada de material biológico y una carga bacteriana potencialmente baja.

## II. JUSTIFICACIÓN Y OBJETIVOS

Si bien es cierto que en algunas ocasiones se realiza un abordaje meta-ómico para el estudio de la microbiota, la norma general es inclinarse por una técnica concreta, ya sea la secuenciación masiva, la secuenciación de amplicones o la metaproteómica. Cada una de estas aproximaciones responde a preguntas diferentes, ya sea la composición de la microbiota o su funcionalidad. Si nos centramos en el análisis composicional, el cual puede abordarse tanto a través de la metaproteómica como de la metagenómica, vemos que existen discordancias entre ambas metodologías. Es necesario determinar las diferencias y similitudes de los resultados de ambas técnicas, así como las fortalezas de combinar estas aproximaciones meta-ómicas en el estudio de la microbiota de muestras complejas. El empleo de metodologías que nacen desde distintos planteamientos puede ser de especial interés para profundizar en las particularidades de personas que nacen con una predisposición genética al desarrollo de infecciones, como es el caso de la FQ, antes de que aparezcan los síntomas de la enfermedad.

Los **OBJETIVOS** de este capítulo fueron:

1. Realizar un análisis estadístico robusto para determinar la concordancia entre la secuenciación de amplicones y la metaproteómica en la determinación de la composición de la microbiota intestinal con una pequeña cohorte representativa de lactantes con FQ.
2. Identificar diferencias en la determinación cualitativa y cuantitativa de grupos bacterianos y relacionarlas con los sesgos derivados del fundamento de cada metodología.
3. Evaluar la sensibilidad de ambas técnicas en la detección de diferencias en la composición de la microbiota intestinal en una cohorte de lactantes con FQ a lo largo del primer año de vida después del diagnóstico.
4. Estudiar las proteínas humanas que interactúan con la microbiota intestinal presentes en la cohorte de lactantes con FQ en el primer mes de vida y tras un año después de su diagnóstico.
5. Identificar las proteínas bacterianas presentes en la cohorte de lactantes con FQ en el primer mes de vida y tras un año después de su diagnóstico, así como su categoría funcional.

### III. MATERIALES Y MÉTODOS

#### 1. Diseño metodológico

Para este capítulo, el diseño se dividió en tres apartados:

Para los objetivos 1 y 2, se planteó un estudio de casos y controles pareado, en el que se compararon los resultados obtenidos por ambas metodologías, tomando como controles los resultados obtenidos por secuenciación de amplicones. La variable de estudio fue la composición de la microbiota intestinal, tanto a nivel cualitativo (presencia/ausencia de grupos bacterianos) como cuantitativo (abundancia relativa de esos grupos bacterianos).

En segundo lugar, y para llevar a cabo el objetivo 3, se planteó un estudio longitudinal para cada metodología, en paralelo. Las variables de estudio fueron las distintas abundancias relativas de los miembros mayoritarios de la microbiota intestinal y las métricas de diversidad alfa (índice Shannon-Weaver y riqueza bacteriana), evaluadas a lo largo del tiempo (de 1 a 12 meses).

En último lugar, y para llevar a cabo los objetivos 4 y 5, se planteó también un estudio de casos y autocontroles, exclusivamente con la metodología de la metaproteómica, en el que se compararon las proteínas bacterianas y humanas obtenidas de las muestras más longevas de cada lactante (casos, llamado grupo de “FQ temprana”) con la muestra correspondiente a un mes de vida (autocontroles, llamado grupo de muestras “inicial”).

#### 2. Participantes, recogida de muestras e historiales médicos

Los sujetos incluidos en este estudio tenían FQ diagnosticada mediante cribado neonatal, prueba de cloruro en sudor y secuenciación de mutaciones, y fueron reclutados durante su primer mes de vida, entre enero de 2018 y enero de 2019. Los lactantes fueron atendidos en 3 hospitales españoles de referencia en FQ, geográficamente distantes. Los sujetos fueron excluidos si habían nacido después de esa fecha, si tenían menos de 3 muestras recogidas o si faltaban datos clínicos importantes. Finalmente, se realizó el seguimiento de ocho lactantes, y los datos relevantes, incluyendo los hábitos alimentarios, el modo de parto y las mutaciones de FQ de los mismos se muestran en la **Tabla 3**. Los datos relacionados con el uso de antibióticos y otros tratamientos se resumen en **Anexos**. La aprobación ética del estudio fue concedida por el Comité de Ética Ramón y Cajal en 2017 (**Anexos**), y todos los padres de los recién nacidos firmaron el consentimiento

informado (**Anexos**). Las muestras fecales se recogieron secuencialmente del pañal de cada lactante durante el primer año de vida en las revisiones médicas rutinarias. Las muestras se almacenaron a -80°C separadas en 2 alícuotas para la posterior secuenciación de amplicones y el análisis metaproteómico.

En este estudio se siguieron las directrices para la presentación de informes sobre el microbioma de *Strengthening The Organization and Reporting of Microbiome Studies* (STORMS) y se puede consultar en **Anexos** [200].

**Tabla 3. Datos demográficos y clínicos de los sujetos con FQ incluidos en el estudio comparado.**

Sujeto	Sexo al nacer	Insuficiencia pancreática	Tipo de parto	Lactancia	Mutación CFTR <sup>1</sup>	Meses de recogida de muestras (nº de muestra)											
						1	2	3	4	5	6	7	8	9	10	11	12
1	H	Sí	Vaginal	Fórmula	G542X/R1162X	1	2						3				
2	V	Sí	Cesárea	Fórmula	F508del/ Y1092X	4				5		6					
3	H	Sí	Vaginal	LME <sup>2</sup>	F508del/ F508del	7			8						9		10
4	V	No	Vaginal	LME	F508del/ 3272-26A->G	11		12			13		14				
5	V	Sí	Vaginal	LME	F508del/ F508del	15			16				17				
6	V	Sí	Vaginal	Fórmula	F508del/N1303K	18					19			20			
7	H	Sí	Vaginal	Fórmula	F508del/ F508del	21			22	23		24					
8	H	Sí	Vaginal	Fórmula	F508del/ R1162X	25		26		27							28

<sup>1</sup> Todos los alelos, a excepción de 3272-26A>G son clasificados como de alto riesgo y, por tanto, todos los lactantes poseían un genotipo de alto riesgo de FQ a excepción del número 4.

<sup>2</sup> LME: lactancia materna exclusiva

### 3. Extracción de ADN

Las muestras de heces se descongelaron lentamente, primero a -20°C 24 h y luego a 4°C otras 24 h, se suspendieron en 500 µL de agua MiliQ y se procedió con el protocolo del kit de extracción QiaAMP (QiaAMP, Alemania) siguiendo las recomendaciones del fabricante. Se suspendieron 0,3-0,5 g. en 1 mL de agua MiliQ y se centrifugaron 5 min a 1.000 rpm para tomar 200 µL de sobrenadante y transferirlo a tubos eppendorf. Se añadieron 100 µL de BT1 y 20 µL de proteinasa K (20 mg/mL). Esto se llevó a un baño de 56°C durante 45 min para la disolución completa. Se añadieron 200 µL de tampón BB3, se homogeneizó y se llevó a un baño de 70°C durante 10 min. Se añadieron 200 µL de etanol 100%, se vorteoó y se centrifugó 3 min a 8.000 rpm. Se transfirió el volumen a una columna de extracción y se llevó a la centrífuga 1 minuto a 8.000 rpm. Se desechó el tubo de recogida, se añadió el reactivo de lavado y se centrifugó de nuevo 1 minuto a 8.000 rpm. Se desechó nuevamente el tubo de recogida y se llevó la columna de extracción con el filtro a otro. Se añadió reactivo de lavado de nuevo y se centrifugó 3 min a 13.000 rpm. Se desechó el tubo de recogida y se llevó la columna de extracción a tubos eppendorf. Se añadieron 50 µL de agua MiliQ y se centrifugó 1 minuto a 13.000 rpm y, posteriormente, se añadieron 30 µL para acabar de eluir el volumen y desechar el filtro. El ADN bacteriano se conservó en tubos eppendorf a 4°C hasta el momento de su utilización. Se determinó la concentración y la calidad de la extracción de ADN en Qubit (*Thermo Fisher Scientific*, MA, E.E.U.U.).

### 4. Secuenciación de amplicones de 16S ARNr

Se realizó la secuenciación con la tecnología Mi-Seq 2x300 bp de extremos pareados (Illumina) de las regiones hipervariables V3 y V4 del 16S ARNr [201] en la Unidad Central de Apoyo a la Genómica Traslacional del Instituto de Investigación Sanitaria Ramón y Cajal (UCA G-T IRYCIS). Los cebadores empleados para la amplificación fueron:

Cebador directo F341:

5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+CCTACGGGNGGCWGC  
AG.

Cebador inverso R805:

5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACA+GGACTACHVGGGTATC  
TAATCC.

#### 4.1. Análisis bioinformático

El procesamiento de las secuencias y la asignación taxonómica derivada de la secuenciación de amplicones del 16S ADN<sub>r</sub> fue realizada con el software QIIME2 (versión 2021.11), mientras que las métricas de diversidad, el análisis estadístico y la exploración de los datos fue realizada con diversos paquetes del software RStudio (versión R 4.2.1.).

#### 4.2. Procesamiento de lecturas

Primeramente, se evaluó la calidad de las lecturas de los archivos FASTQ en la plataforma QIIME2 v.201.11 [202] con el complemento q2-demux. Se ajustaron los parámetros de filtrado y se prosiguió con el complemento DADA2, que realiza un filtrado de calidad de las secuencias, la unión de lecturas directas e inversas, la eliminación de quimeras y la construcción de tablas con el recuento de ASV por muestra. Asimismo, se descartaron los ASV con muy baja frecuencia (<7 lecturas).

#### 4.3. Base de datos para la asignación taxonómica

A partir de la base de datos SILVA 138.1 (octubre 2019), la cual contiene secuencias completas del gen 16S ADN<sub>r</sub>, se obtuvo una base de datos entrenada específicamente para estos análisis, siguiendo las recomendaciones y directrices de diversas publicaciones [203]–[205]. Se eliminaron secuencias replicadas y redundantes, se extrajeron las secuencias correspondientes únicamente a las regiones V3 y V4 del gen, se eliminaron secuencias no bacterianas que pudieran quedar y se realizó la asignación con el clasificador bayesiano *classify-sklearn naive Bayes* mediante el complemento q2-feature-classifier de QIIME2 y el paquete RESCRIPt. Las tablas de abundancia relativa y de contingencia de los ASV incluyeron todos los taxones identificados. Se eliminaron las secuencias no bacterianas o sin asignación taxonómica, así como todas las muestras con menos de 1.000 lecturas.

Las secuencias de nucleótidos se depositaron en el repositorio *Sequence Read Archive* (SRA) del *National Center for Biotechnology Information* (NCBI), *BioProject* ID 719717 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA719717>).

## 5. Análisis Metaproteómico

### 5.1. Obtención de la fracción proteica bacteriana

Las muestras se prepararon mediante centrifugación diferencial y el *pellet* microbiano se procesó por sonicación según un protocolo previamente publicado. En resumen, se suspendieron 0,1-0,3 g de heces en 10 mL de solución salina tamponada con fosfato (PBS) y se mezclaron en un rotador de tubos durante 45 min a 4 °C. Las muestras se centrifugaron a 500 g durante 5 min y el sobrenadante se recogió en un tubo de 50 mL. Se añadieron de nuevo 10 mL de PBS, repitiendo el proceso dos veces y los 3 sobrenadantes (~30 mL) se centrifugaron a 11.000 g. Los gránulos microbianos se suspendieron en 500 µL de tampón de lisis (4% de dodecil sulfato sódico [SDS], 50 mM de Tris-HCl pH 8,0) y se calentaron durante 10 min a 95 °C. Se realizaron cuatro ciclos de sonicación (30 s con un intervalo de 1 min en hielo), con una amplitud del 40%. A continuación, se añadieron perlas de sílice (0,3 g) a cada muestra y se realizaron 5 rondas de batido de perlas (30 segundos con un intervalo de 5 min en hielo) a una velocidad de 6,5 ms<sup>-1</sup>. Se realizó una centrifugación de 14.000 g para eliminar las perlas y los restos celulares. Para eliminar el SDS, las proteínas se precipitaron con metanol/cloroformo y se suspendieron en urea 8 M para la digestión con tripsina en solución. Por último, los péptidos se cuantificaron en un fluorímetro Qubit (Thermo Scientific) y se cargó 1 µg de péptidos para el análisis por cromatografía de fase inversa-nano-líquida con ionización por electrospray en tándem de masas en un sistema EASY-nLC 1000 (Proxeon) acoplado a un espectrómetro de masas Q-Exactive HF (Thermo Scientific Inc.). Los péptidos se cargaron en línea en una columna Acclaim PepMap 100 Trapping (75 µm de diámetro interior × 20 mm, resina C18 de 3 µm y 100 Å de poro; Thermo Scientific) utilizando el tampón A (ácido fórmico al 0,1%) y luego se separaron en una columna analítica de resina C18 (75 µm de diámetro interior × 500 mm, 2 µm y 100 Å de poro; Thermo Scientific Easy Spray Column). Para separar los péptidos se realizó un gradiente de 240 min del 2% al 40% de tampón B (ácido fórmico al 0,1% en acetonitrilo al 100%) en tampón A.

### 5.2. Espectrometría de masas en tándem (MS/MS)

Los datos se obtuvieron mediante adquisición dependiente de datos en modo positivo. De cada barrido MS (entre 350 y 2000 Da), se seleccionaron los 15 precursores más intensos (carga entre 2+ y 5+) por su fragmentación de disociación de alta energía de colisión, con una exclusión dinámica de 10 segundos y una energía de colisión normalizada de 20, y

se adquirieron los correspondientes espectros MS/MS. Los péptidos se eluyeron utilizando un gradiente de 240 min. Los datos brutos de la proteómica MS fueron enviados al Consorcio ProteomeXchange (<http://www.proteomexchange.org>) a través del repositorio asociado Proteomics Identifications Database con el identificador de base de datos PXD029284 [206].

### 5.3. Análisis de los datos

Para el procesamiento de los datos se empleó el software MetaLab, que proporciona una base de datos del microbioma intestinal humano [63], [207]. También se utilizó una base de datos humana descargada de Uniprot DB (<http://www.uniprot.org>) [208], restringida a la taxonomía humana (descargada el 18/02/2020 con 74.451 secuencias), para identificar proteínas humanas. Para la identificación de péptidos y proteínas, la tasa de descubrimiento falso (FDR) Benjamini-Hochberg se fijó en 0,01. Sólo se consideraron los taxones identificados con al menos 2 péptidos y se filtraron manualmente para eliminar los péptidos humanos. La suma de las intensidades de todos los péptidos distintivos asignados a un taxón se utilizó como la abundancia relativa de ese taxón. También se realizó un análisis taxón-función, utilizando la información taxonómica del análisis de enriquecimiento de la plataforma iMetaLab (<http://shiny.imetalab.ca/>) [209].

Para el análisis estadístico funcional de las proteínas microbianas y humanas, se utilizó GraphPad Prism (versión 9.1.1.225). Se realizó un análisis con el programa STRING (versión 11.5), *Search Tool for the Retrieval of Interacting Genes/Proteins*.

### 6. Acuerdos de asignación taxonómica y cuantificación microbiana

Los datos obtenidos de ambas metodologías se compararon, seleccionando sólo los taxones identificados por ambos métodos (taxones compartidos) y resolviendo las incongruencias de anotación. En primer lugar, evaluamos la variable dependiente dicotómica, "presencia de taxones", mediante una prueba de McNemar con corrección de continuidad a nivel de filo, clase y género bacteriano, considerando aquellos taxones con al menos un ASV asignado y al menos 2 péptidos asignados en metaproteómica. La significación se ajustó a  $p < 0,05$ . En segundo lugar, se normalizó la abundancia relativa de cada filo, clase o género por muestra, expresada como porcentaje y se realizó un análisis diferencial de abundancia con aquellos taxones detectados en más del 50% de las muestras por alguna de las dos metodologías. La abundancia composicional se calculó relativa al número total de lecturas de cada bacteria y a la suma total de intensidad LFQ

para secuenciación de amplicones y metaproteómica, respectivamente. Las métricas de diversidad alfa (índice Shannon Weaver y riqueza) se calcularon con el comando *estimate\_richness* del paquete *phyloseq* de RStudio (versión 1.4.1717) a nivel de género bacteriano [210]. El paquete *ggpurb* se empleó para la representación de los datos pareados y *GraphPad Prism* (versión (9.1.1.225) para el análisis estadístico funcional de proteínas humanas y bacterianas. Se realizó una prueba no paramétrica de suma de rangos de Wilcoxon para comparar los rangos medios de abundancia relativa de cada grupo bacteriano y cada métrica de diversidad alfa, pareando por muestras. Los valores de significación se ajustaron con FDR de 0,1.

La monitorización de los cambios en la composición de la microbiota a lo largo del primer año de vida se realizó con el paquete *splinctomeR* (versión 0.1.0.) en RStudio. Se evaluaron sólo los géneros mayoritarios (n=16). Se comparó, en primer lugar, la tendencia de todas las bacterias a lo largo de la serie temporal entre un método y otro con la función *permuspliner* (999 permutaciones), para ver si existían diferencias en general. Asimismo, cada taxón fue analizado individualmente para detectar si había diferencias en algún momento concreto en sus fluctuaciones con la función *slidingspliner*, la cual permite identificar en qué punto temporal hay cambios significativos ( $p < 0,05$ ). Finalmente, se investigó el efecto del tiempo en estas bacterias con cada método, por separado, con la función *trendyspliner*. Lo mismo se aplicó para las métricas de diversidad alfa. La significación estadística se ajustó a FDR=0,1.

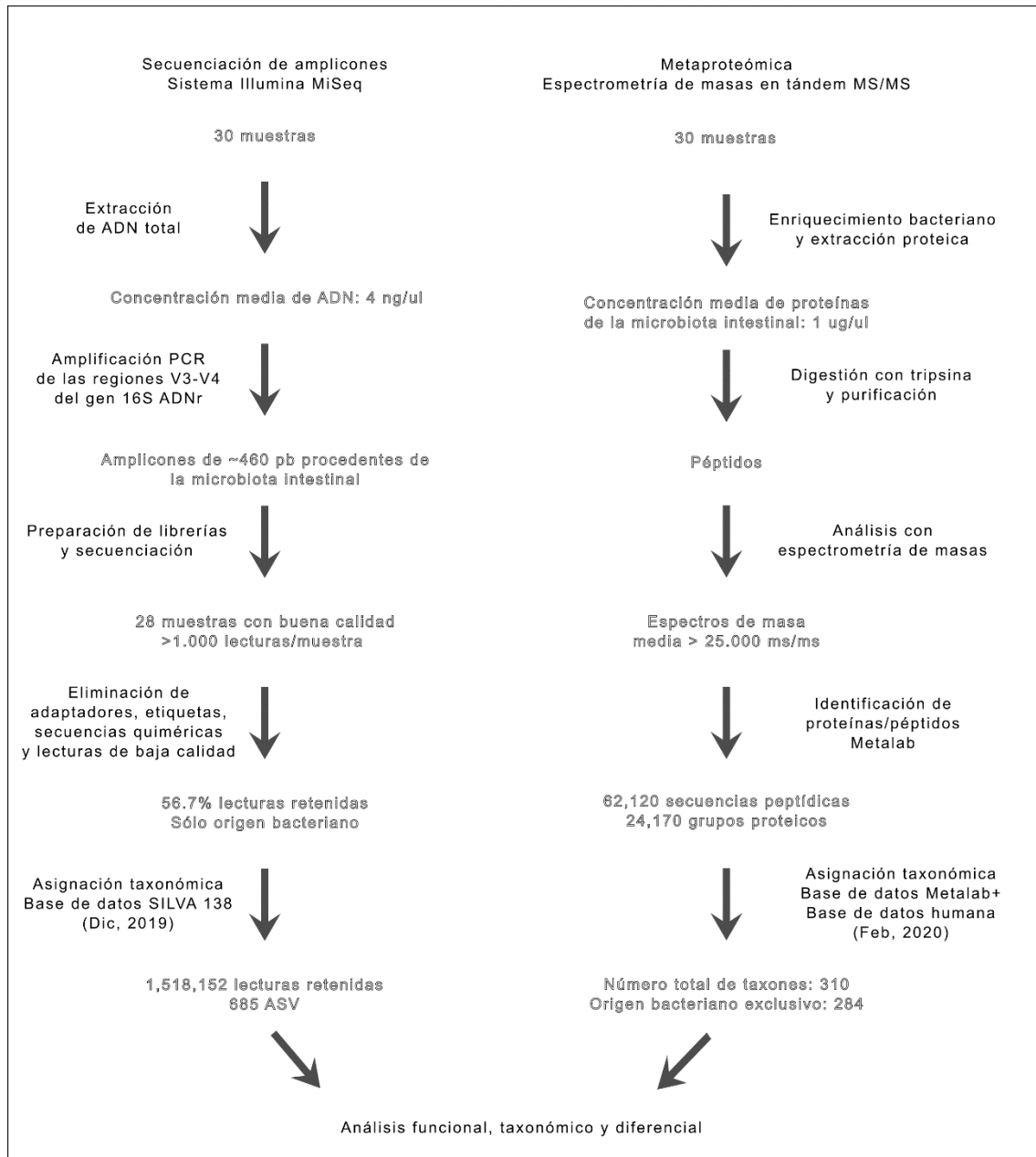
## IV. RESULTADOS

### 1. Inclusión de muestras

Tras realizar la selección de niños en base al número de muestras disponibles, se incluyeron 8 recién nacidos con FQ para este estudio comparativo, cada uno de los cuales aportó 3 o 4 muestras fecales a lo largo de su primer año de vida. En total, se analizaron 56 muestras fecales en el apartado de casos (n=28) y controles (n=28) para comparar las dos metodologías y 16 muestras en el apartado de casos y autocontroles (8 muestras iniciales y 8 muestras de FQ temprana).

### 2. Concordancia estadística de los resultados de la secuenciación de amplicones y la metaproteómica

Globalmente, se identificaron 685 ASV mediante secuenciación de amplicones, correspondientes a 9 filos, 13 clases, 130 géneros, y 310 taxones mediante metaproteómica, de los cuales 286 fueron bacterianos, correspondientes a 8 filos, 17 clases y 65 géneros. Cabe destacar que los resultados de metaproteómica se encuentran, desde el inicio del análisis, ya colapsados a los diferentes rangos taxonómicos, mientras que en la secuenciación de amplicones es posible analizar los datos colapsados o sin colapsar (en ASV) (**Figura 11**).



**Figura 11. Flujo de trabajo y principales resultados obtenidos con ambas metodologías.**

Para realizar el análisis comparativo, primero se analizó la nomenclatura y clasificación que habían llevado a cabo ambas metodologías con sus respectivas bases de datos de referencia, en los niveles taxonómicos de filo, clase, orden, familia y género. Se observaron algunas discrepancias en algunos grupos de bacterias, como la inclusión del filo *Tenericutes* de la metaproteómica en el filo *Firmicutes* según la secuenciación de amplicones y la reciente reclasificación taxonómica de las clases del filo *Proteobacteria*. En general, estas discrepancias se encontraron en los rangos taxonómicos más altos

(Tabla 4)<sup>1</sup>. Las bacterias detectadas exclusivamente por una u otra metodología no representaban más del 10% del total de la composición relativa en ninguno de los rangos especificados, siendo el nivel de género el de mayores diferencias (Figura 12).

**Tabla 4. Discrepancias en la nomenclatura o clasificación taxonómica a diferentes rangos (filo/clase/orden/familia/género) por secuenciación de amplicones y metaproteómica.**

<b>Secuenciación de amplicones</b>	<b>Metaproteómica</b>
<i>Campylobacterota/Campylobacteria/Campylobacterales</i>	<i>Proteobacteria<sup>2</sup>/Epsilonproteobacteria/Campylobacterales</i>
<i>Firmicutes/Bacilli/Mycoplasmatales</i>	<i>Tenericutes/Mollicutes/Mycoplasmatales</i>
<i>Actinobacteriota</i>	<i>Actinobacteria</i>
<i>Bacteroidota</i>	<i>Bacteroidetes</i>
<i>.../Bacilli/Erysipelotrichales</i>	<i>.../Erysipelotrichia/Erysipelotrichales</i>
<i>.../Gammaproteobacteria/Burkholderiales</i>	<i>.../Betaproteobacteria/Burkholderiales</i>
<i>Desulfobacterota/Desulfovibrionia/Desulfovibrionales</i>	<i>Proteobacteria/Deltaproteobacteria/Desulfovibrionales</i>
<i>.../Clostridia/Peptostreptococcales-Tissierellales</i>	<i>.../Tissierellia/ Tissierellales</i>
<i>.../.../.../Escherichia-Shigella</i>	<i>.../.../.../Escherichia</i>
<i>.../.../Lachnospirales/Lachnospiraceae/Ruminococcus_gnavus_group</i>	<i>.../.../Clostridiales/</i>
<i>.../.../Lachnospirales/Lachnospiraceae/Ruminococcus_torques_group</i>	<i>Ruminococcaceae/Ruminococcus</i>

<sup>1</sup> Este análisis se realizó en los meses de octubre a diciembre de 2019 teniendo en cuenta lo que por entonces establecía la base de datos de referencia SILVA 138.1.

<sup>2</sup> La asignación taxonómica fue hecha en 2019 antes de los últimos cambios nominales en los filos bacterianos. *Actinobacteriota* es ahora *Actinomycetota*, *Firmicutes* es *Bacillota* y *Proteobacteria* es ahora *Pseudomonadota*.

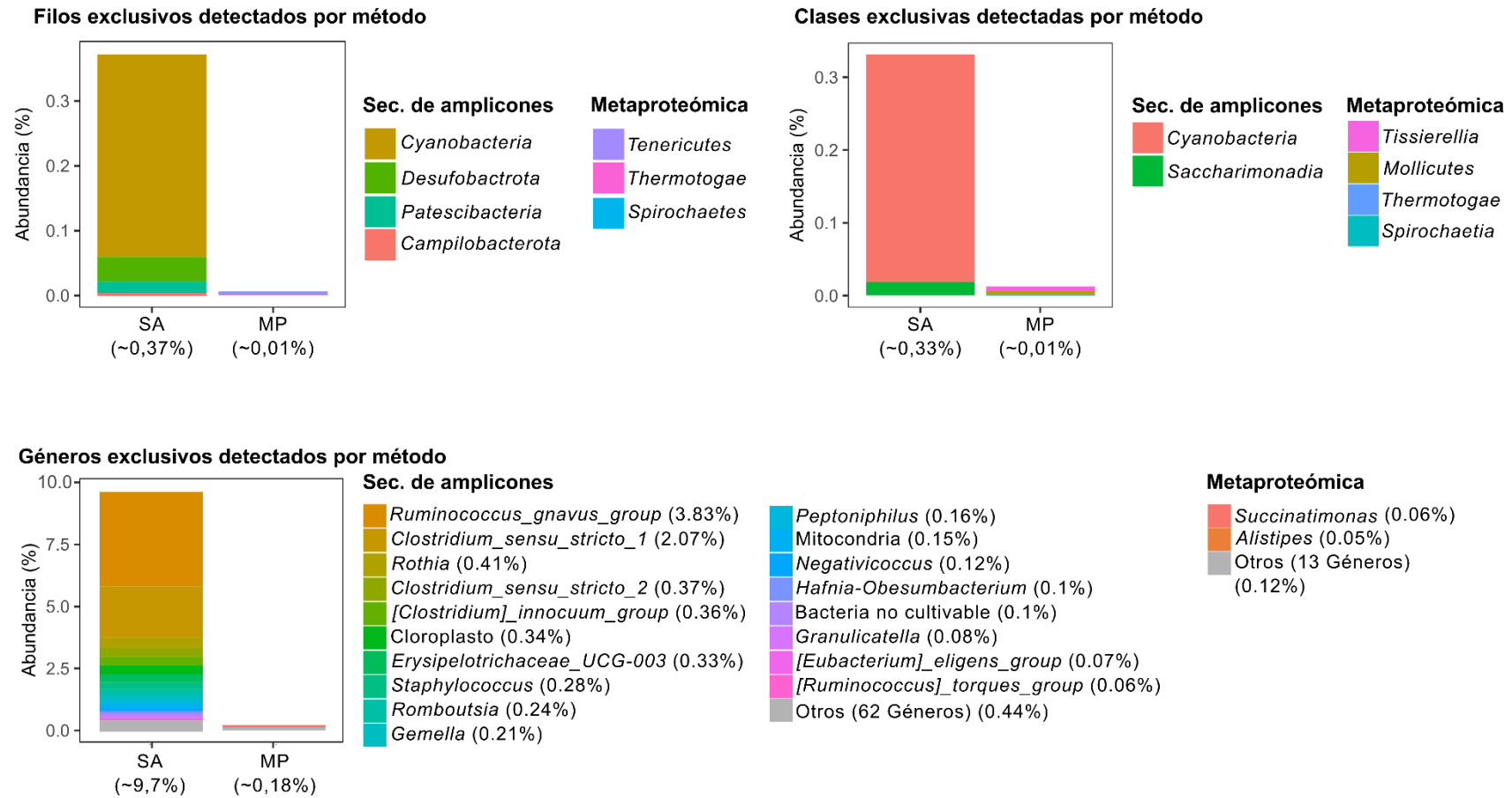


Figura 12. Filos, clases y géneros bacterianos identificados exclusivamente por cada una de las metodologías. SA: secuenciación de amplicones. MP: metaproteómica.

Tras resolver las discrepancias, se analizaron los taxones detectados por ambas metodologías. En primer lugar, se evaluó la presencia o ausencia mediante la prueba McNemar pareada, la cual reveló diferencias, sobre todo, a nivel de género (**Tabla 5**). De igual modo, se realizó el análisis composicional de los taxones mayoritarios. La prueba de los rangos con signo de Wilcoxon no paramétrica mostró diferencias significativas entre las dos técnicas en la abundancia relativa del filo *Actinobacteria*, las clases *Actinobacteria* y *Bacilli*, y los géneros *Blautia*, *Haemophilus*, *Escherichia*, *Enterococcus*, *Streptococcus* y *Bifidobacterium* (**Figura 13**).

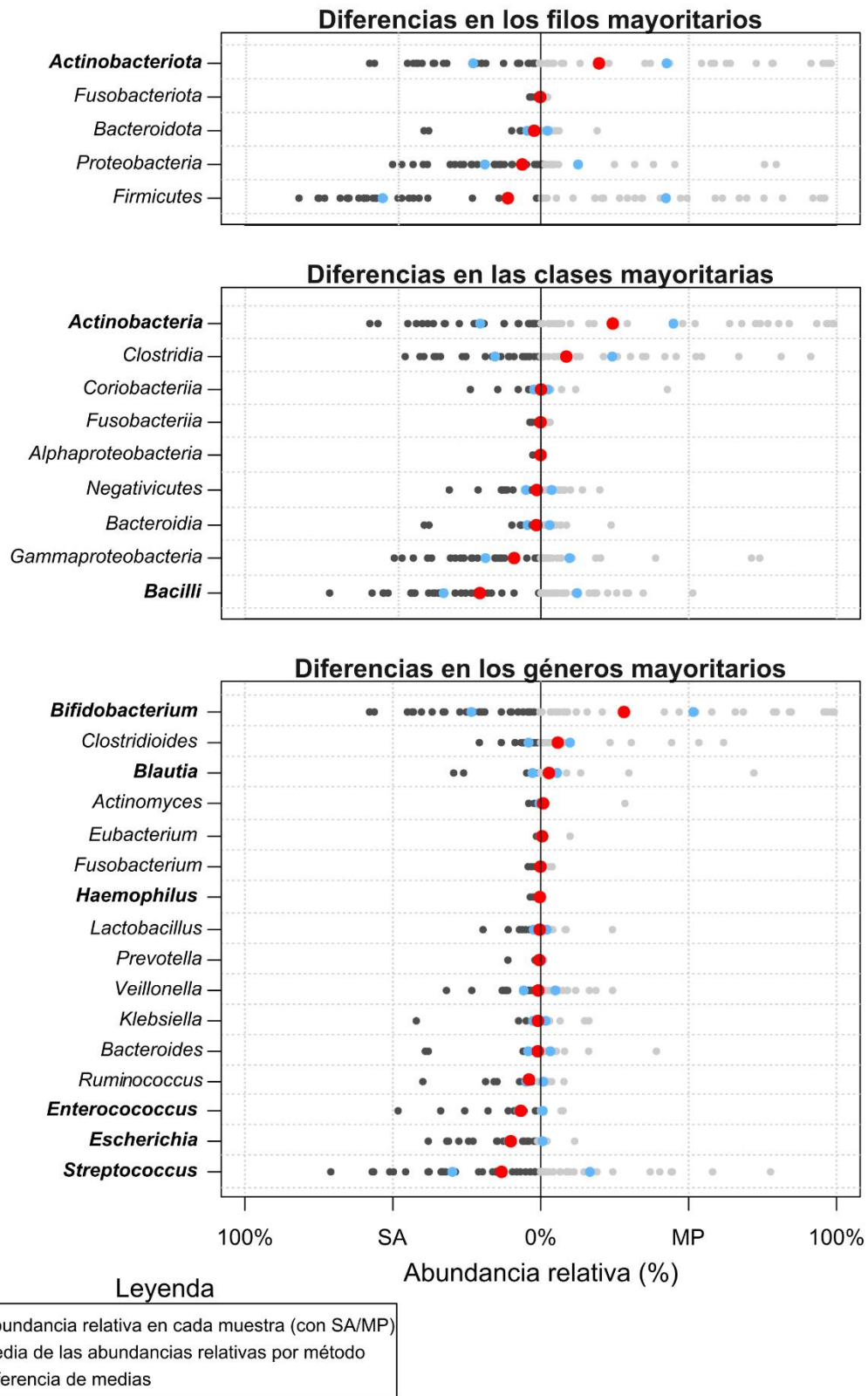
Para tener en cuenta el efecto de cada sujeto a lo largo del tiempo, se evaluó el cambio composicional de los principales géneros bacterianos a lo largo del primer año de vida. Ambas metodologías mostraron tendencias similares para la mayoría (**Figura 14A**), exceptuando *Bifidobacterium*, *Escherichia*, *Enterococcus* y *Ruminococcus*, los cuales casi alcanzan la significancia estadística ( $p < 0,05$  pero  $FDR = 0,13$ ) (**Figuras 14C, 14D, 14E**). De acuerdo con la secuenciación de amplicones, los únicos géneros que cambiaron composicionalmente a lo largo del tiempo fueron *Ruminococcus* ( $p = 0,022$ ) y *Eubacterium* ( $p = 0,007$ ) (**Figura 14B**, panel izquierdo). Según el análisis metaproteómico, los géneros que cambian a lo largo del tiempo en esta cohorte fueron *Fusobacterium* ( $p = 0,02$ ), *Haemophilus* ( $p = 0,029$ ) y *Veillonella* ( $p = 0,03$ ) (**Figura 14B**, panel derecho). Cabe destacar, no obstante, que la significancia estadística puede verse afectada por el pequeño tamaño muestral y, en muchas ocasiones, las diferencias se obtienen en intervalos temporales concretos o en los casos donde las abundancias relativas son más pequeñas.

Los índices de alfa diversidad fueron significativamente distintos entre las dos metodologías (prueba pareada de Wilcoxon  $p < 0,05$ ). Cuando se tuvo en cuenta la serie temporal de muestras repetidas, los métodos siguieron mostrando diferencias, siendo la secuenciación de amplicones la única capaz de detectar una tendencia creciente en el índice Shannon-Weaver (**Figura 15**). Aun así, es importante mencionar que no es del todo adecuado calcular métricas como el índice Shannon con los datos de metaproteómica de los que disponíamos ya que los comandos del paquete de RStudio *phyloseq* advertían de que no había *singletons* en esas bases de datos. Calcular este tipo de índices a partir de intensidades de espectros de masa derivados de la metaproteómica no es una práctica común y debe tenerse en cuenta para futuros trabajos. Las técnicas metagenómicas como la secuenciación de amplicones siguen siendo la mejor y más común herramienta para

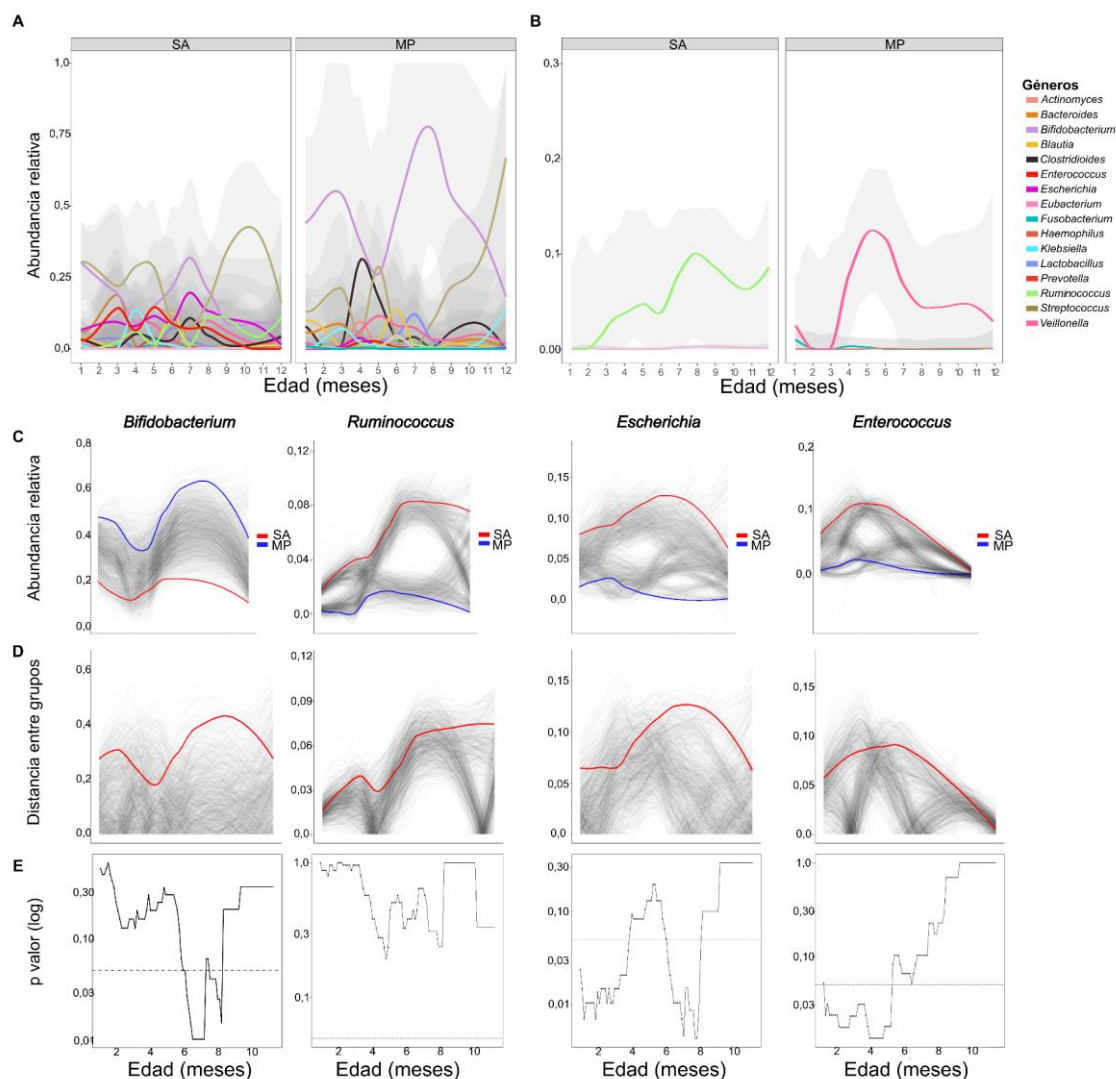
calcular este tipo de métricas.

**Tabla 5. Tabla de contingencia para la prueba Mcnemar. Sólo se muestran los grupos con diferencias significativas ( $p < 0,05$ ).**

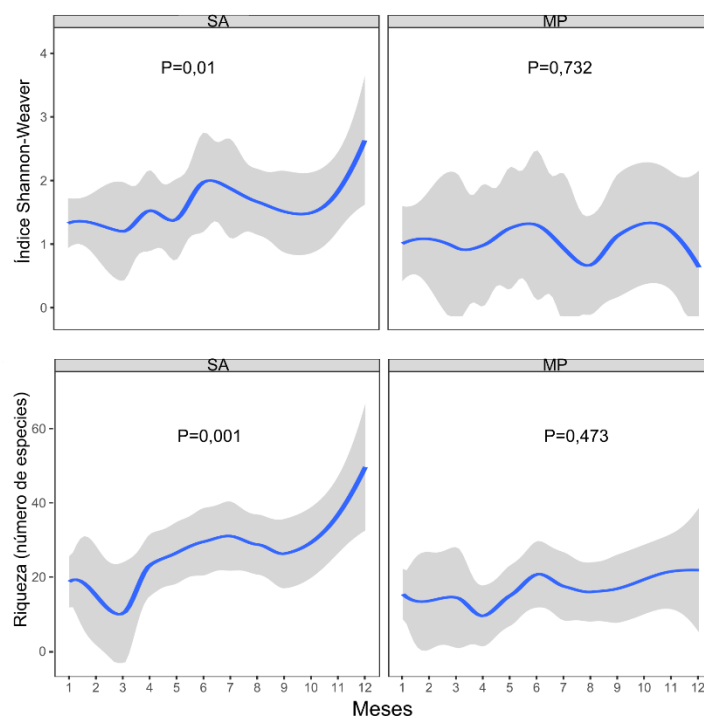
Taxón	Nº de muestras en las que la bacteria se detecta por				p valor
	Secuenciación de amplicones	Metaproteómica	Ambas	Ninguna	
<b>Clase</b>					
<i>Alphaproteobacteria</i>	8	1	9	10	0,02
<b>Género</b>					
<i>Eggerthella</i>	12	0	1	15	0,001
<i>Anaerococcus</i>	11	0	1	16	0,002
<i>Haemophilus</i>	15	2	5	6	0,003
<i>Coprococcus</i>	1	10	1	16	0,007
<i>Lachnoclostridium</i>	0	9	2	17	0,007
<i>Roseburia</i>	0	9	2	17	0,007
<i>Eubacterium</i>	1	14	6	7	0,01
<i>Enterobacter</i>	10	1	3	14	0,01
<i>Enterococcus</i>	10	1	14	3	0,01
<i>Escherichia</i>	11	1	14	2	0,01
<i>Actinomyces</i>	14	4	6	4	0,03
<i>Collinsella</i>	8	1	5	14	0,04



**Figura 13. Abundancias relativas de los principales taxones detectadas por ambas metodologías.** SA: secuenciación de amplicones. MP: metaproteómica. Taxones en **negrita**: aquellos para los que la prueba de Wilcoxon pareada fue significativa ( $p < 0,05$ ).



**Figura 14. Cambio composicional de los géneros mayoritarios en cada una de las metodologías. A:** evolución de la abundancia relativa de los principales géneros bacterianos (n=16). **B:** evolución de los géneros bacterianos con un cambio significativo a lo largo del tiempo detectado por SA (*Ruminococcus* y *Eubacterium*) y por MP (*Fusobacterium*, *Haemophilus* y *Veillonella*). **C:** gráficos obtenidos por la función *permuspliner* (999 permutaciones) donde se muestra la evolución temporal de la abundancia relativa en aquellos géneros con diferencias significativas entre ambas metodologías. **D:** gráficos de distancias entre ambas metodologías. La diferencia (línea roja continua) no es significativa si no se encuentra por encima del 95% de los valores permutados (gris traslúcido). **E:** gráficos obtenidos por *slidingspliner* donde se muestra el valor *p* en cada intervalo especificado (mostrado con 100 intervalos por defecto). La línea de puntos indica  $p = 0,05$ . En ciertos intervalos de tiempo, las diferencias llegan a ser significativas. SA: secuenciación de amplicones. MP: metaproteómica.



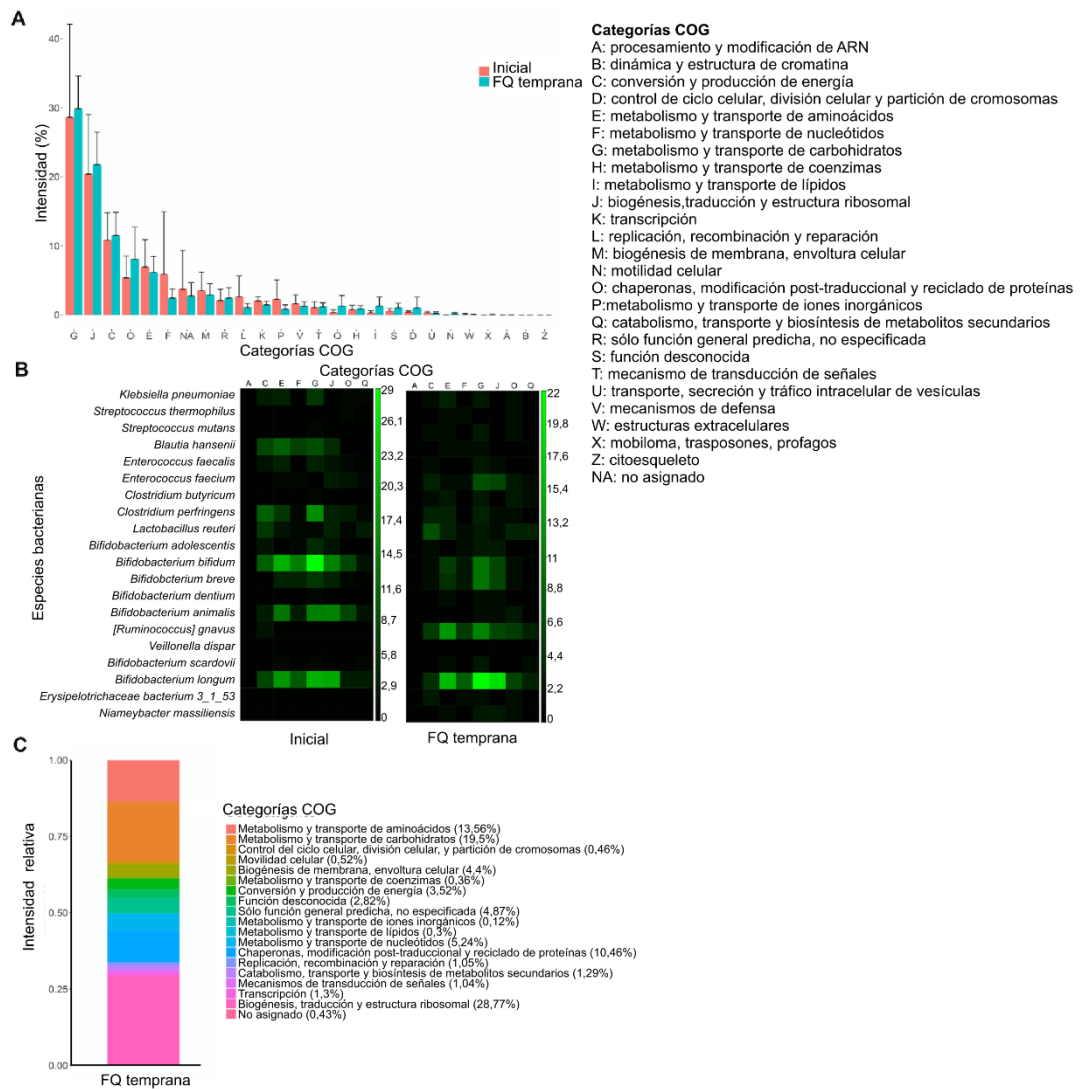
**Figura 15. Evolución de la alfa diversidad (medida a través del índice de Shannon-Weaver y de la riqueza bacteriana) detectada por un método y otro.**

### 3. Estudio funcional de la microbiota de esta cohorte mediante metaproteómica

Se utilizó el programa MetaLab para investigar la función de las proteínas de la microbiota en esta cohorte usando la suma de las intensidades de cada proteína para un análisis cuantitativo. Al comparar el número y la intensidad de la fracción proteica de las muestras iniciales (recién nacidos) con las muestras del final del estudio (FQ temprana), se obtuvieron cantidades muy similares: 5.718 proteínas y  $4,8e+12$  de intensidad al inicio y 7.209 proteínas y  $4,7e+12$  de intensidad al final. Se logró asignar una categoría funcional de grupos ortólogos (COG, por sus siglas en inglés) al 97% de las proteínas analizadas con MetaLab. Se detectaron 14 COGs (6 en el grupo de recién nacidos y 8 en el grupo de FQ temprana) con relevancia estadística, previa normalización de las intensidades proteicas (se dividió la intensidad asignada a cada función entre la suma total de intensidades en cada muestra). A pesar de no haber obtenido significación estadística en el análisis diferencial, probablemente debido al reducido tamaño muestral de este estudio, las categorías funcionales con mayores intensidades fueron “producción y conversión de energía (C), “transporte y metabolismo de carbohidratos” (G), y “traducción, estructura ribosomal y biogénesis” (J) (**Figura 16A**).

La correlación de cada bacteria con su función fue realizada, asimismo, con iMetaLab, empleando la herramienta iMetaShiny. Las proteínas asociadas con *Bifidobacterium* fueron de las más abundantes, tanto en el grupo inicial como en el de FQ temprana. En este último, destacó el enriquecimiento de las vías funcionales asignadas a *Ruminococcus gnavus*, confirmando lo que venía señalando el análisis taxonómico composicional anterior (**Figura 16B**). La cantidad de proteínas asociadas a *R. gnavus* aumentó significativamente entre los dos momentos temporales (de 2 a 69 grupos de proteínas detectados), estando la mayoría de ellas relacionadas con procesos de traducción (**Figura 16C**). Por otro lado, proteínas asociadas a las especies *Clostridium perfringens* y *Blautia hansenii* sufrieron una disminución en las muestras de FQ temprana.

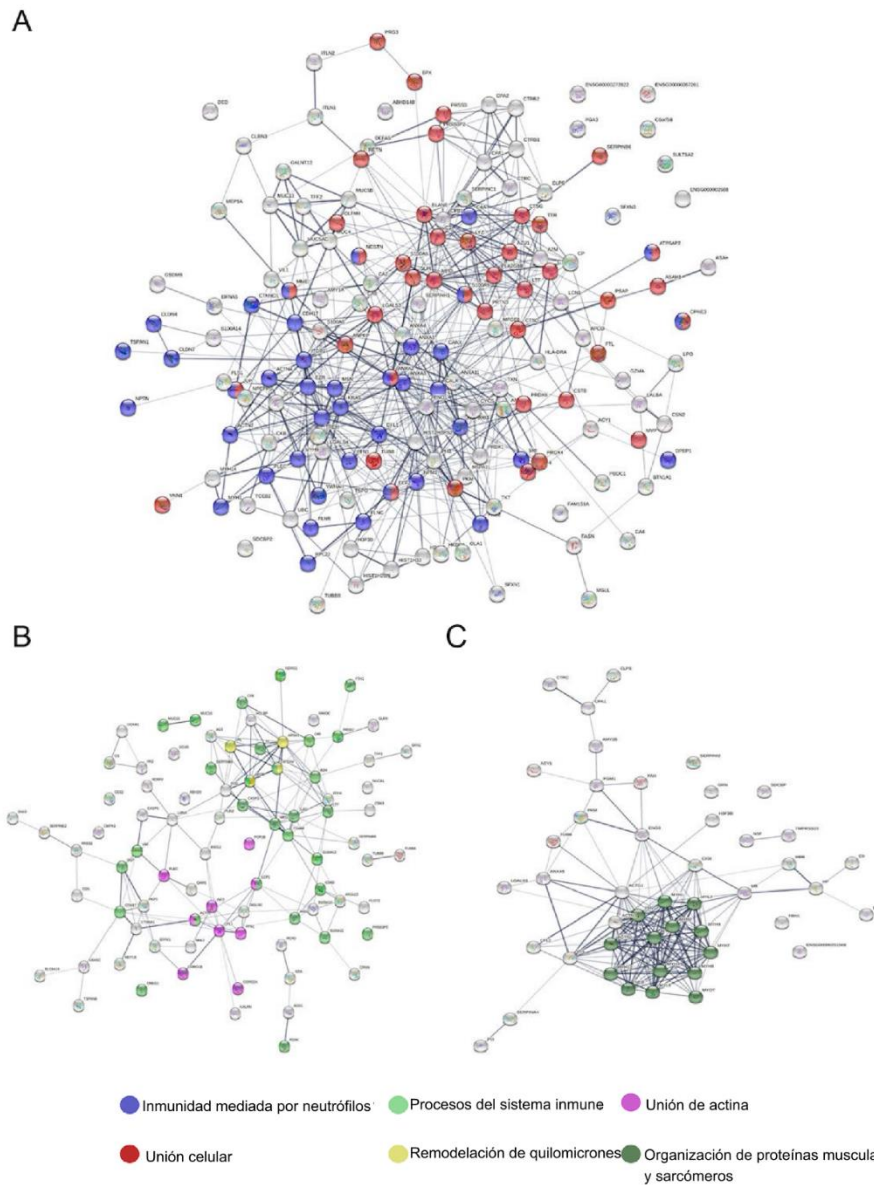
Por último, la metaproteómica permitió analizar la fracción de proteínas humanas asociadas al material analizado. Se detectaron 293 proteínas humanas, que fueron etiquetadas en tres grupos: grupo inicial (85 proteínas exclusivas), FQ temprana (43 proteínas exclusivas) y grupo de proteínas compartidas (n=165). Sólo se detectó un enriquecimiento significativo en 6 de estas últimas (**Tabla 6**). Se llevó a cabo un análisis STRING (versión STRING 11.5) con los tres grupos de proteínas, y se observó una interacción proteína-proteína muy llamativa en el grupo de proteínas compartidas (**Figura 17**). Estas proteínas estaban relacionadas con la inmunidad mediada por neutrófilos (42 proteínas,  $p=1,87e-25$ ), como la calprotectina (S100-A9 y S100-A8), la elastasa neutrófila (ELANE), la catepsina G y la mieloperoxidasa, o relacionadas con la barrera intestinal (40 proteínas,  $p=1,24e-5$ ), como las mucinas (MUC2, MUC4, MUC13), y las proteínas presentes en uniones celulares (**Figura 17A**). En las muestras iniciales, se encontró un número significativamente mayor de proteínas asociadas con el ensamblaje de quilomicrones y el sistema inmunitario en comparación con el grupo de FQ temprana, donde predominaban las proteínas de músculo (**Figura 17B y 17C**).



**Figura 16. Análisis funcional de las proteínas bacterianas de esta cohorte, comparando el grupo de muestras iniciales y el de muestras finales (FQ temprana).** **A:** intensidad relativa de cada una de las categorías funcionales COG en cada grupo de muestras. **B:** correlación entre las categorías funcionales más enriquecidas y las especies bacterianas a las que se les asignan. El gradiente de color indica mayor intensidad relativa (verde claro). **C:** proporciones de las categorías funcionales COG de las proteínas asignadas a la bacteria *R. gnavus*.

**Tabla 6. Proteínas humanas enriquecidas en cada grupo de la cohorte.**

<b>Mayor expresión en grupo INICIAL</b>	<b>Mayor expresión en grupo FQ TEMPRANA</b>	<b><i>P</i> valor</b>
Lactotransferrina		<0,0001
Mieloperoxidasa		<0,0001
	Actina citoplasmática 2	<0,0001
	Quimotripsina-C	<0,0001
	Carboxipeptidasa A1	0,0409
	Miosina-1	<0,0001



**Figura 17. Proteínas humanas encontradas en interacción con el microbioma intestinal de la cohorte. A:** grupo de proteínas exclusivas de las muestras iniciales. **B:** grupo de proteínas exclusivas de las muestras de FQ temprana. **C:** proteínas compartidas por ambos grupos.

## V. DISCUSIÓN

Recientemente, se ha publicado una nueva clasificación de mutaciones genéticas del gen CFTR [112], diferenciando entre mutaciones de alto riesgo y de bajo riesgo, de acuerdo con las manifestaciones clínicas asociadas. Siete de ocho de los niños incluidos en este estudio portaban mutaciones de alto riesgo, por lo que se ha considerado que la población es homogénea. El entorno de la FQ condiciona fuertemente la composición de la microbiota intestinal [211], contribuyendo asimismo a su mal funcionamiento y estabilidad, y estos cambios pueden detectarse ya a las 6 semanas de vida [189]. A pesar de que la secuenciación de amplicones es la técnica más utilizada para determinar la composición de la microbiota intestinal, existe una demanda creciente de técnicas que determinen la funcionalidad global del ecosistema, como la metaproteómica. En este capítulo, se pretendió supervisar el establecimiento de la microbiota intestinal de la FQ durante el primer año de vida, comparando la secuenciación de amplicones y la metaproteómica, e incorporando un análisis estadístico robusto para evaluar la concordancia de ambas técnicas en la asignación taxonómica y la detección de la abundancia bacteriana.

Las discordancias taxonómicas detectadas correspondieron en su mayoría a grupos parafiléticos, que son aquellos que están siendo reclasificados constantemente en base a los últimos avances en filogenia. En este sentido, SILVA 138 es una base de datos más precisa con la que establecer relaciones filogenéticas, ya que las bases de datos de proteínas se actualizan con menos frecuencia.

Curiosamente, la metaproteómica detectó proteínas atribuidas a *Streptococcus thermophilus*, mientras que al utilizar la secuenciación de amplicones, las proteínas parecían corresponder a *Streptococcus salivarius*. Se ha informado previamente de la escasa resolución a la hora de identificar especies de estreptococos secuenciando las regiones V3-V4 del gen 16S ARNr [212]. Otra limitación de la secuenciación de amplicones es la detección de la esporobiota [213], siendo la mayoría de las especies del filo *Firmicutes* endosporulantes. Las esporas resisten las técnicas de extracción de ADN, y sus genomas son a menudo más grandes y al ser secuenciados, se producen menos lecturas por gen por taxón, por lo que en general están infrarrepresentadas en la secuenciación de amplicones. Otros filos bacterianos, como *Actinobacteria* (en particular *Bifidobacterium*), tienen genomas con un alto contenido de GC, y esto también contribuye a su subestimación por las técnicas de secuenciación.

La metaproteómica ha fracasado en la detección de ciertos grupos bacterianos que suelen encontrarse en la microbiota intestinal humana, debido a su resistencia a la lisis celular, especialmente las bacterias Gram positivas [214]. Además, ciertos componentes microbianos y detergentes utilizados para la lisis celular pueden afectar a la digestión postenzimática [215]. En el protocolo de trabajo empleado, se han incluido pasos optimizados, como centrifugaciones en serie para el pre-enriquecimiento microbiano, sonicación con perlas y lisis con tampones con SDS de base, para evitar estas limitaciones. Mientras que este método mejora el número de identificaciones de péptidos y proteínas y la abundancia relativa de *Actinobacteria*, otros protocolos han demostrado una mayor identificación de proteínas de otros grupos taxonómicos como *Proteobacteria*, pero con menor identificación de péptidos y proteínas [216]. El análisis metaproteómico presenta varios retos en el área de la bioinformática, siendo los más comunes la correcta identificación de péptidos, la asignación taxonómica y la calidad de la anotación funcional [215]. Se estima que la metaproteómica sólo puede captar el 10-20% de las proteínas expresadas en los microbiomas intestinales humanos. Los espectros de espectrometría de masas también pueden llegar a saturarse con las proteínas altamente abundantes de las especies dominantes [35].

El análisis estadístico de los resultados reflejó que ambas técnicas realizan la asignación taxonómica bacteriana de manera muy similar, aunque en términos cuantitativos se encontraron algunas diferencias [217]. La secuenciación de amplicones estima una menor abundancia relativa de *Bifidobacterium* en comparación con otras técnicas y, del mismo modo, también se ha reportado una mayor abundancia relativa de algunos miembros de *Firmicutes* y *Proteobacteria* [218].

El pequeño tamaño muestral de nuestra cohorte limita la comparación con otros estudios. Los índices de alfa diversidad de este estudio de lactantes con FQ no alcanzaron las cifras observadas en otras cohortes de FQ ni en la cohorte de control por Antosca *et al.* [189], aunque otros autores han notificado valores inferiores a los de los controles sanos [219]. De este trabajo, se concluye que no es posible comparar métricas de diversidad alfa entre secuenciación de amplicones y metaproteómica, aunque sí podría hacerse en casos donde se emplearan las mismas técnicas de metaproteómica entre sí. En secuenciación, varios ASV (cada uno con un determinado número de lecturas, del orden de 1 a  $10^4$ ) pueden corresponder al mismo taxón bacteriano. Por otro lado, en metaproteómica, se evalúa la suma de intensidades (del orden de  $10^6$  a  $10^{11}$ ) de todos los

péptidos correspondientes al mismo taxón para calcular su abundancia relativa en una muestra. Esto hace que el cálculo de métricas como el índice de Shannon-Weaver, que tienen en cuenta la proporción de cada bacteria, y su comparación entre técnicas, no sea del todo fiable, mientras que la riqueza (número de especies diferentes) sí podría emplearse. Lo que sí se pudo corroborar, a través de la secuenciación de amplicones, es el discreto aumento en la diversidad alfa de las muestras de FQ temprana.

La proteómica fecal en adultos con FQ ha señalado el predominio de *R. gnavus*, *Enterobacteriaceae* y *Clostridia*, junto con una disminución de productores de butirato, como *Faecalibacterium prausnitzii* [144]. En este capítulo, se detectaron cambios de composición similares en una cohorte muy reducida de 8 lactantes a los que se les hizo un seguimiento durante 12 meses justo después de ser diagnosticados de FQ. *R. gnavus* se ha asociado a varias enfermedades gastrointestinales humanas y es una bacteria mucolítica capaz de degradar la mucina colónica humana MUC2, por tanto, tiene una estrecha relación con la capa de moco intestinal [220], [221]. Se ha sugerido que la bacteria puede penetrar en la capa de moco y provocar una activación del sistema inmunitario y respuestas inflamatorias [222], [223]. Asimismo, esta bacteria también se ha relacionado con enfermedades alérgicas y autoinmunes y actualmente se está cuestionando su grado de implicación en ellas [220]-[224].

No se observaron cambios funcionales drásticos en las proteínas bacterianas al comparar las muestras iniciales y las de FQ temprana. Las proteínas más abundantes en las muestras iniciales fueron la formiltetrahidrofolato sintetasa y la piruvato-formato liasa (datos no mostrados), previamente relacionadas con la regulación de las biopelículas [225]. En las muestras de FQ temprana, el enriquecimiento más relevante fue el de la chaperonina GroEL, una proteína de respuesta al estrés cuyo papel principal es el plegado de proteínas, que recientemente se ha observado que tiene funciones en la virulencia y la patogénesis, como la adherencia bacteriana y la invasión inmunitaria [226].

Una de las ventajas más interesantes de la metaproteómica es la detección de proteínas de origen humano. En estas muestras, se detectaron varias proteínas asociadas al mantenimiento del epitelio intestinal y a la respuesta inmunitaria. La lactotransferrina, proteína de secuestro de hierro abundante en la leche materna humana, y la mieloperoxidasa, una enzima microbicida regulada por inflamación presente en los granulocitos de los neutrófilos, estaban significativamente enriquecidas en las muestras iniciales [227], [228]. Todos estos resultados apuntan a un establecimiento alterado de la

microbiota intestinal en los lactantes con FQ, con un predominio de *R. gnavus* dentro de un estado inflamatorio humano.

La principal limitación del estudio es su reducido tamaño muestral. Debido a esto, se incluyeron *p* valores sin ajustar o ligeramente elevados para informar de las diferencias encontradas a nivel de filo, clase y género bacteriano, así como en proteínas bacterianas y humanas, que mostraban tendencias diferenciales, pero sin alcanzar la significación estadística. Los grupos bacterianos y proteicos señalados podrían servir para futuros estudios en donde se realicen caracterizaciones metabólicas en FQ más amplias y con mayor N.

Este estudio proporciona un análisis comparativo ampliado con un sólido apoyo estadístico que podría optimizar el uso de ambos enfoques para la investigación de la microbiota intestinal. La metaproteómica proporciona información sobre la composición y la funcionalidad, así como datos sobre las interacciones entre el huésped y el microbioma. Su punto fuerte es la identificación y cuantificación de *Actinobacteria* y ciertas clases de *Firmicutes*, mientras que su limitación es la falta de precisión en el cálculo de las métricas de diversidad alfa. Teniendo en cuenta todos los resultados, ambas técnicas detectaron una microbiota aberrante en los lactantes con FQ durante su primer año de vida, dominada por el enriquecimiento de *R. gnavus* en un entorno inflamatorio humano.

## VI. DIVULGACIÓN

Los resultados de este capítulo se han publicado en la revista *Microbiology Spectrum* bajo el título: *Statistical Evaluation of Metaproteomics and 16S rRNA Amplicon Sequencing Techniques for Study of Gut Microbiota Establishment in Infants with Cystic Fibrosis* (DOI: [10.1128/spectrum.01466-22](https://doi.org/10.1128/spectrum.01466-22)) y se incluye en **Anexos**.

## **CAPÍTULO 2**

# **MONITORIZACIÓN DE LA IMPLANTACIÓN DE LA MICROBIOTA INTESTINAL Y NASOFARÍNGEA EN RECIÉN NACIDOS CON FIBROSIS QUÍSTICA**

### *Los niños recién nacidos con FQ...*

*¿Tienen más **exacerbaciones** si tienen **genotipo de alto riesgo**?  
¿Tienen más **ingresos**? ¿Toman más **antibióticos**?*

*¿Existe asociación entre la **microbiota respiratoria e intestinal** y la **progresión de la FQ**?*

*¿Está relacionada la **composición de la microbiota intestinal y respiratoria** con la **primocolonización de S. aureus**?  
¿Y con la de **P. aeruginosa**?*

*¿Evoluciona de manera diferente la **microbiota intestinal y nasofaríngea** de los recién nacidos con FQ? ¿Existen **bacterias compartidas** en estos nichos? ¿Cuáles son los **patrones de colonización bacteriana**?*

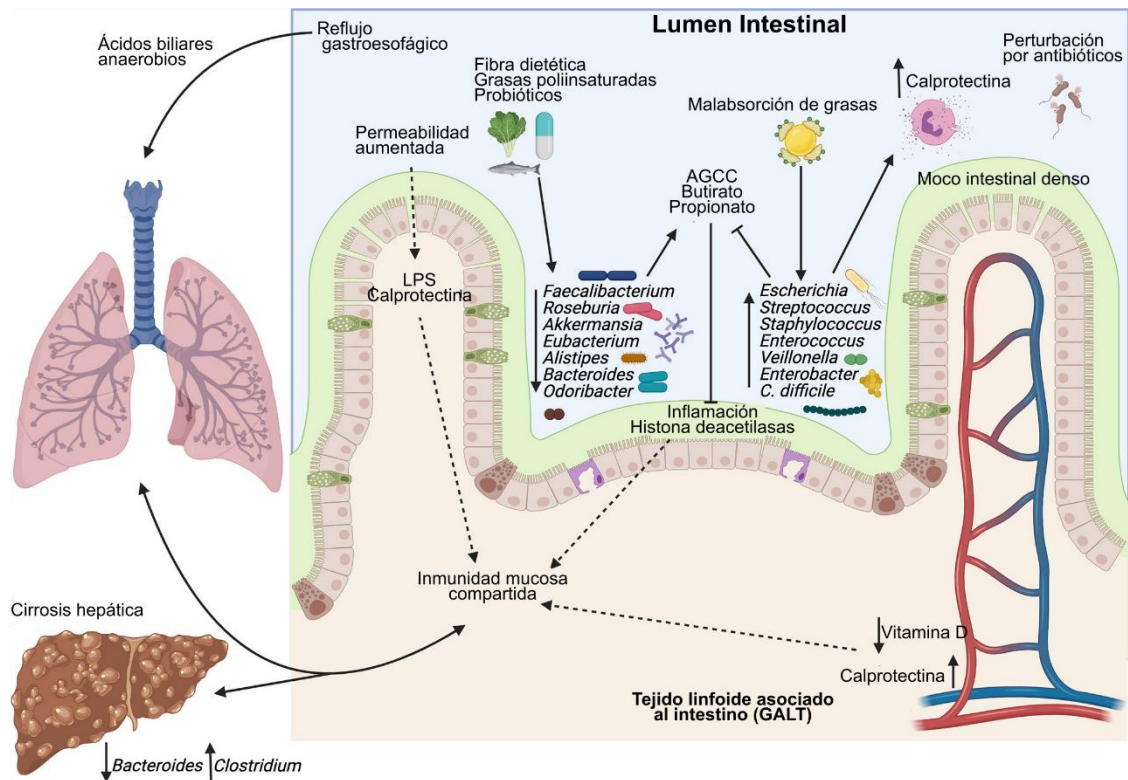
## I. ANTECEDENTES

Los estudios microbiológicos en la FQ se han centrado en la vía aérea, ya que es donde ocurren las mayores complicaciones de la enfermedad. El defecto en el canal CFTR también afecta al tracto digestivo a través de los ácidos biliares, a la secreción pancreática y la secreción de bicarbonato, todo lo cual conduce a un pH intestinal más ácido con una malabsorción de grasa que se elimina posteriormente en las heces y se denomina esteatorrea [229]–[231].

Los antibióticos que se administran para controlar la colonización pulmonar también afectan a la microbiota intestinal, que está a su vez condicionada por la modificación de la dieta, la sustitución de enzimas pancreáticas y las terapias con moduladores del canal CFTR [232]–[234]. El incremento de la esperanza de vida de los sujetos con FQ conlleva la aparición de complicaciones intestinales en etapas posteriores de la vida, como tasas más altas de enfermedad intestinal inflamatoria (EII) y de cáncer de colon [235]–[238].

Actualmente, existe una necesidad de conocer el eje intestino-pulmón en la FQ para implementar medidas terapéuticas que eviten el daño en ambos compartimentos. Este eje se define como la capacidad de la microbiota intestinal para influir en el curso o el resultado de la enfermedad pulmonar subyacente, y viceversa. Hay pruebas sólidas de la existencia de este eje en el asma y en otras enfermedades respiratorias, como la EPOC y la FQ [239], [240]. Se ha documentado la transmisión directa de microorganismos entre ambos compartimentos, pero también existe una comunicación metabólica de la que aún hay pocas evidencias [241].

En niños con FQ de 6 a 12 meses de vida, se ha demostrado que la microbiota intestinal de los que presentan exacerbaciones respiratorias es significativamente diferente de la de los niños sin exacerbación (**Figura 18**) [189], [242]. Además, la diversidad alfa intestinal correlacionó positivamente con un mayor tiempo hasta la exacerbación inicial en los lactantes [185], [242]. Por último, las alteraciones de géneros intestinales también pueden predecir la patología de las vías respiratorias, ya que *Parabacteroides* se redujo significativamente en el intestino antes de la colonización inicial de *Pseudomonas* en una cohorte de lactantes con FQ [242]. Otra de las cuestiones que aún no se han resuelto es el hecho de si existen alguna relación entre el genotipo de FQ y la composición de la microbiota [112].



**Figura 18. El eje intestino-cerebro en FQ.** Resumen de las alteraciones más comunes que se encuentran en el tracto gastrointestinal de las personas con FQ y cómo estos cambios pueden afectar a hígado y pulmón. La FQ está asociada a una disminución en bacterias productoras de AGCC (como *Faecalibacterium*, *Roseburia*, *Akkermansia*, *Eubacterium*, *Alistipes* y *Odoribacter*) y un aumento de las bacterias que degradan AGCC y son potencialmente patógenas (*Escherichia*, *Streptococcus*, *Staphylococcus*, *Enterococcus*, *Veillonella*, *Enterobacter* y *Clostridium difficile*). Las alteraciones microbianas son impulsadas, en parte, por el alto contenido de grasas en la materia fecal, el frecuente tratamiento antibiótico, y la densidad del moco intestinal. Estas alteraciones favorecen un ambiente inflamatorio que afecta de manera sistémica y al pulmón, debido al sistema inmune compartido de las mucosas y la transmisión directa de los contenidos intestinales hacia el tracto respiratorio. La cirrosis hepática está asociada a un descenso de *Bacteroides* y un aumento de *Clostridium*, y puede estar relacionada también con la inmunidad de las mucosas. Traducción de Price CE, O'Toole GA. The Gut-Lung Axis in Cystic Fibrosis. *J Bacteriol.* 2021 Sep 23;203(20):e0031121. doi: 10.1128/JB.00311-21. Licencia libre.

Clásicamente, se considera que la colonización patógena respiratoria comienza con la adquisición de *S. aureus* y, posteriormente, se introduce *P. aeruginosa*. Sin embargo, gracias a las técnicas de secuenciación, se ha podido descubrir que la microbiota es mucho más compleja y que las especies pueden ir alternando, a pesar de no detectarse en los cultivos [243]. La disparidad entre los resultados de los cultivos tradicionales y la secuenciación de amplicones no ha sido suficientemente estudiada, pero un problema añadido es la adecuación de los cultivos a las muestras de FQ. Los microorganismos de

la FQ suelen requerir de tiempos más prolongados de incubación, y es por eso por lo que se ha comenzado a contemplar la secuenciación en los procedimientos de cultivo de muestras respiratorias [244]. El tipado de estos aislados mediante herramientas como el PFGE o el MLST aporta también una perspectiva poblacional que nos permite inferir si varios pacientes comparten los mismos linajes o si, por el contrario, los clones que colonizan a los sujetos con FQ son diferentes de los que infectan a otros individuos con otras patologías.

## II. JUSTIFICACIÓN Y OBJETIVOS

El cribado neonatal de FQ ha servido para anticipar el manejo terapéutico y el seguimiento de estas personas, mejorando sustancialmente su calidad de vida. Ahora disponemos de una cantidad mayor de pacientes más jóvenes para estudiar las dinámicas del microbioma y las interacciones entre la microbiota y la fisiología del hospedador con FQ en etapa de desarrollo.

En los últimos años, se ha puesto especial hincapié en la compleja relación entre las bacterias anaerobias con el desencadenamiento de las exacerbaciones pulmonares, la inflamación y la colonización patogénica en FQ. Aún está por determinar si los resultados observados en población adulta pueden extrapolarse a recién nacidos, o si existen dinámicas de interacción diferentes.

Si bien en el capítulo anterior se hizo un abordaje con distintas metodologías de la microbiota intestinal de estas personas, no se exploró la microbiota respiratoria y la relación existente entre ellas. También quedaron fuera del estudio el cultivo microbiológico y el tipado de las bacterias mayoritarias, su prevalencia y su intercambio intra e interindividual. Incorporando estas aproximaciones, estaremos ampliando la mirada sobre una condición de base genética con repercusiones a nivel de toda la fisiología, que también podrían servir para seleccionar posibles biomarcadores de progreso o recrudescimiento de la FQ.

Los **OBJETIVOS** de este capítulo fueron:

1. Identificar y caracterizar mediante electroforesis en campo pulsado la sucesión de clones bacterianos de la microbiota intestinal y respiratoria en lactantes con FQ durante 2 años, relacionándola con la composición obtenida mediante secuenciación de amplicones, la colonización patogénica y las exacerbaciones clínicas.
2. Identificar los géneros bacterianos con cambios composicionales significativos a lo largo de los primeros años con FQ en muestras fecales y respiratorias, y su relación con las variables clínicas.
3. Establecer el núcleo común o *core* de bacterias en muestras fecales y respiratorias.

### III. MATERIALES Y MÉTODOS

#### 1. Diseño metodológico

Se realizó un estudio descriptivo, longitudinal y retrospectivo en el que se reclutaron recién nacidos diagnosticados con FQ (mediante cribado neonatal, prueba de cloruro sódico en sudor y posterior secuenciación de mutaciones), en cuatro hospitales españoles de referencia (Hospital Universitario Ramón y Cajal de Madrid, Hospital Universitario 12 de Octubre de Madrid, Hospital Regional Universitario de Málaga y Hospital Universitario Virgen del Rocío de Sevilla), entre enero de 2018 y septiembre de 2019. Tras la firma del consentimiento informado (**Anexos**) por parte de los padres, se recogieron muestras de heces del pañal y muestras de lavados nasofaríngeos en las visitas rutinarias al hospital para el seguimiento de la FQ. Las muestras se almacenaron a -80°C en diferentes alícuotas de ~0,3 g (heces) o 1 mL (lavado nasofaríngeo, LN) después de la recogida hasta su procesamiento. Los criterios de inclusión fueron: tener un diagnóstico de FQ en su primer mes de vida, tener identificada la mutación causante de la condición, y tener la información demográfica y clínica relevante de cada sujeto recopilada hasta sus dos años de vida.

Las variables recopiladas sistemáticamente hasta los dos años de edad fueron: tipo de mutación de las dos copias del gen CFTR, tipo de parto, edad gestacional, sexo al nacer, insuficiencia pancreática, tipo de lactancia, mes de introducción a la dieta sólida, fecha de las exacerbaciones respiratorias y de otro tipo de sintomatología, número de ingresos hospitalarios relacionados con la exacerbación respiratoria así como debidos a otras causas, tipo, duración y vía del tratamiento antibiótico, ingesta de probióticos y, por último, momento, duración y tipo de colonización por los patógenos clásicos de la FQ, en especial, *S. aureus* y *P. aeruginosa*.

#### 2. Extracción de ADN

Las muestras se descongelaron lentamente, primero a -20°C durante 24 h y luego a 4 °C otras 24 h. En el caso de las heces, se suspendieron 0,5 g en 1 mL de agua MiliQ y se centrifugaron 5 min a 1.000 rpm para tomar 200 µL de sobrenadante y transferirlo a tubos eppendorf. En el caso de los LN, se partió de 1 mL de suspensión y se incubó previamente con 10% de lisozima (10 mg/mL) y de lisostafina (10 mg/mL) 30 min a 37 °C. Se prosiguió con el protocolo del kit de extracción QiaAMP (QiaAMP, Alemania). Se añadieron 100 µL de BT1 y 20 µL de proteinasa K (20 mg/mL). Esto se llevó a un baño de 56 °C durante 45 min para la disolución completa. Se añadieron 200 µL de tampón

BB3, se homogeneizó y se llevó a un baño de 70 °C durante 10 min. Se añadieron 200 µL de etanol 100%, se homogeneizó y se centrifugó 3 min a 8.000 rpm. Se transfirió el volumen a una columna de extracción y se llevó a la centrífuga 1 min a 8.000 rpm. Se desechó el tubo de recogida, se añadió el reactivo de lavado y se centrifugó de nuevo 1 min 8.000 rpm. Se desechó nuevamente el tubo de recogida y se llevó la columna de extracción con el filtro a otro. Se añadió reactivo de lavado de nuevo y se centrifugó 3 min a 13.000 rpm. Se desechó el tubo de recogida y se llevó la columna de extracción a tubos eppendorf. Se añadieron 50 µL de agua MiliQ y se centrifugó 1 min a 13.000 rpm y, posteriormente, se añadieron 30 µL para acabar de eluir el volumen y desechar el filtro. El ADN bacteriano se conservó en tubos eppendorf a 4 °C hasta el momento de su utilización. Se determinó la concentración y la calidad de la extracción de ADN en Qubit (*Thermo Fisher Scientific*, MA, E.E.U.U.).

Para la extracción a partir de colonias, se suspendió una colonia en 200 µL de BT1 y se prosiguió con el protocolo anterior.

### 3. Secuenciación de amplicones de 16S ARNr´

Se realizó la secuenciación Mi-Seq 2x300 bp paired-end (Illumina) del ARNr 16S de las regiones V3 y V4 [201] en la UCA G-T IRYCIS. Los cebadores empleados para la amplificación fueron:

Cebador directo F341:

5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+CCTACGGGNGGCWGC  
AG.

Cebador inverso R805:

5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACA+GGACTACHVGGGTATC  
TAATCC.

#### 3.1. Análisis bioinformático

El procesamiento de las secuencias y la asignación taxonómica derivada de la secuenciación de amplicones del 16S ADNr fue realizada con el software QIIME2 (versión 2021.11) [202], mientras que las métricas de diversidad, el análisis estadístico y la exploración de los datos fue realizada con diversos paquetes del *software* RStudio.

### 3.2. Procesamiento de lecturas

Primeramente, se evaluó la calidad de las lecturas de los archivos FASTQ en la plataforma QIIME2 con el complemento q2-demux. Se ajustaron los parámetros de filtrado y se prosiguió con el complemento DADA2, que realiza un filtrado de calidad de las secuencias, la unión de lecturas directas (*forward*) e inversas (*reverse*), la eliminación de quimeras y la construcción de tablas con el recuento de ASV por muestra. Asimismo, se descartaron las ASV con muy baja frecuencia (<7 lecturas).

### 3.3. Base de datos para la asignación taxonómica

Las secuencias de nucleótidos se depositaron en el repositorio SRA NCBI, *BioProject* ID 895693 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA895693>). A partir de la base de datos SILVA 138 (Febrero 2022), la cual contiene secuencias completas del gen ADNr 16S, se obtuvo una base de datos entrenada específicamente para estos análisis, siguiendo las recomendaciones y directrices de diversas publicaciones [203]–[205].

### 4. Cultivo de muestras, tipificación mediante electroforesis en gel de campo pulsado (PFGE) y estadística de la culturómica

Las heces y los LN se descongelaron lentamente a -20 °C durante 24 h y a 4 °C durante otras 24 h, para evitar la muerte bacteriana. Se inocularon porciones de entre 0,3-0,5 g de cada muestra en caldo *brain heart infusion* (BHI) (Difco, Detroit, Michigan) y se incubaron a 37 °C durante 24 h a modo de pre-enriquecimiento bacteriano. Las bacterias cultivables se aislaron en medios de agar selectivos y no selectivos del tubo BHI (sembrando distintas diluciones), incluyendo placas de agar de *M-Enterococcus*; agar chocolate; De Man, Rogosa y Sharpe (MRS); Manitol-Salt; McConkey; y agar Columbia con un 5% de sangre de carnero. Los medios de cultivo se compraron a Difco y las placas se incubaron a 37 °C durante 24-72 h, incluyendo 5% de CO<sub>2</sub> para las placas de agar sangre y condiciones de anaerobiosis para las placas MRS.

La identificación de las colonias se llevó a cabo mediante espectrometría MALDI-TOF (Bruker, Alemania), y todos los aislados se conservaron a -80°C en leche semidesnatada esterilizada. Los aislados se tiparon genéticamente mediante PFGE utilizando los protocolos estandarizados y las enzimas de restricción de acuerdo con la especie bacteriana (*SmaI* y *XmaI* para especies de *S. aureus*, *SmaI* para *Enterococcus faecalis* y *E. faecium*, *XbaI* para *E. coli* y, finalmente, *SpeI* para *Klebsiella oxytoca*, *K. pneumoniae* y *S. marcescens*). El análisis del patrón de bandas se realizó con el programa

Phoretix 5.0 (TotalLab, NewCastle upon Tyne, Reino Unido), y la representación de los resultados se hizo a partir de los coeficientes Dice. Se consideró que dos clones eran idénticos cuando el coeficiente era  $<0,8$ . Si estaba próximo al punto de corte, se establecieron sub-clones mediante letras.

Se calculó la frecuencia de cada especie y de cada clon ( $\alpha_h$ ) con la siguiente fórmula, asumiendo que la probabilidad de encontrar cada especie o clon en una muestra fecal o de LN es siempre la misma (1), independientemente del periodo de tiempo que represente.

$$\left( \sum_{k=1}^P \sum_{i=1}^{n_k} \delta_{ik}^h * \frac{1}{n_k} \right) * \frac{1}{P} = \alpha_h$$

Siendo:

- $\alpha_h$ : probabilidad de la presencia del clon o especie  $h$  en  $\rho$ . De 0 a 1.
- P: total de neonatos con FQ
- $n_k$ : número de muestras (heces o LN) de neonato  $k$
- $\delta_{ik}^h$ : 1 si el clon o especie  $h$  está presente en muestra  $i$  de neonato  $k$ ; 0 si ausente

### 5. Identificación de especies de *Streptococcus*

Se amplificó el gen conservado *rpoB* (860 pb) de acuerdo con la bibliografía [245]. Se emplearon los siguientes cebadores:

Directo- AARYTIGGMCCTGAAGAAAT.

Inverso- TGIARTTTRTCATCAACCATGTG.

Se llevaron a cabo 35 ciclos (95 °C 2' – 94 °C 30'' – 52 °C 30'' – 72 °C 1') con 5 min de elongación final a 72 °C. Los amplicones de la PCR se purificaron usando Illustra™ ExoProStar 1-Step™ (GE Healthcare Life Sciences, Little Chalfont, Reino Unido) y secuenciados con la tecnología Sanger por Macrogen Inc. (Gasan-dong, Geumchun-gu, Seúl, Corea). Las secuencias fueron ensambladas y editadas manualmente usando el *software* BioEdit. Por último, se construyeron secuencias consenso de cada una de las especies con el complemento MUSCLE del programa MEGAX y se construyó un árbol filogenético con los representantes.

### 6. Tipado genético de *S. aureus* y *P. aeruginosa* mediante MLST

Para cada uno de los aislados identificados como *S. aureus* o *P. aeruginosa* por MALDI-TOF se amplificaron las regiones correspondientes a los diferentes genes constitutivos, con los cebadores y las condiciones incluidas en las **Tablas 7 y 8**, siguiendo las directrices de la página oficial de MLST [246].

### 7. Análisis estadísticos

Se evaluaron las variables clínicas en función del perfil genotípico de FQ con las pruebas de Fisher y de Mann-Whitney con el valor de significancia estadística  $p < 0,05$ .

A partir de las tablas de recuento de ASV crudas obtenidas de la secuenciación de amplicones, se normalizaron las muestras mediante submuestreo clasificado o SRS por sus siglas en inglés [247] y, posteriormente, se calcularon diferentes índices de diversidad alfa (como métrica cuantitativa se empleó Shannon-Weaver y como métrica con información filogenética, PD) con los paquetes *phyloseq* y *picante* de RStudio. Luego, se trabajó con las tablas de recuento aglomeradas a nivel de género y se transformaron en porcentajes. Para representar gráficas se emplearon paquetes como *ggplot2* o *splinectomeR*, el cual también permitió evaluar cambios composicionales significativos a lo largo del tiempo con las funciones *trendyspliner*, *permuspliner* y *slidingspliner*.

La asociación entre la abundancia relativa de los principales géneros bacterianos y la variable cuantitativa edad en meses se calculó con la correlación de Spearman, así como la asociación entre los principales géneros detectados en LN y heces. Se calculó y representó con el paquete *phylosmith* y *ggplot2* de RStudio.

La diferencia de prevalencia de las especies patógenas detectadas por las técnicas dependientes de cultivo y por secuenciación se evaluó con la prueba no paramétrica para muestras relacionadas de McNemar. La asociación entre la detección por las técnicas dependientes de cultivo y su abundancia relativa obtenida por secuenciación se calculó con la prueba no paramétrica de Mann-Whitney. El valor de significación estadística se fijó en  $p < 0,05$ .

**Tabla 7. Condiciones y cebadores para las PCR MLST de *S. aureus*. D: directo. I: inverso.**

Gen	Programa PCR	Ciclos	Cebadores (5'→3')
<i>arc</i> (456 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 55 °C 30 s. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: TTGATTCACCAGCGCGTATTGTC
			I: AGGTATCTGCTTCAATCAGCG
<i>aro</i> (456 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 55 °C 30 s. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: ATCGGAAATCCTATTTACATTC
			I: GGTGTGTATTAATAACGATATC
<i>glp</i> (465 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 55 °C 30 s. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: CTAGGAACTGCAATCTTAATCC
			I: TGGTAAAATCGCATGTCCAATTC
<i>gmk</i> (429 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 55 °C 30 s. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: ATCGTTTTATCGGGACCATC
			I: TCATTA ACTACAACGTAATCGTA
<i>pta</i> (474 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 55 °C 30 s. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: GTTAAAATCGTATTACCTGAAGG
			I: GACCCTTTTGTGAAAAGCTTAA
<i>tpi</i> (402 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 56 °C 30 s. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: TCGTTCATTCTGAACGTCGTGAA
			I: TTTGCACCTTCTAACAATTGTAC
<i>yqi</i> (516 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 56 °C 30 s. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: CAGCATA CAGGACACCTATTGGC
			I: CGTTGAGGAATCGATACTGGAAC

**Tabla 8. Condiciones y cebadores para las PCR MLST de *P. aeruginosa* D: directo. I: inverso.**

Gen	Programa PCR	Ciclos	Cebadores (5'→3')
<i>acsA</i> (390 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 60 °C 1 min. 4. 72 °C 30 s. 5. 72 °C 10 min.	30	D: GCCACACCTACATCGTCTAT
			I: AGGTTGCCGAGGTTGTCCAC
<i>aroE</i> (498 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 55 °C 1 min. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: ATGTCACCGTGCCGTTCAAG
			I: TGAAGGCAGTCGGTTCCTTG
<i>guaA</i> (373 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 60 °C 1 min. 4. 72 °C 30 s. 5. 72 °C 10 min.	30	D: AGGTCGGTTCCTCCAAGGTC
			I: GACGTTGTGGTGCGACTTGA
<i>mutL</i> (442 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 55 °C 1 min. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: AGAAGACCGAGTTCGACCAT
			I: GGTGCCATAGAGGAAGTCAT
<i>nuoD</i> (366 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 60 °C 1 min. 4. 72 °C 30 s. 5. 72 °C 10 min.	30	D: ACGCGGAGAACGAGGACTAC
			I: TGGCGGTTCGGTGAAGGTGAA
<i>ppsA</i> (370 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 60 °C 1 min. 4. 72 °C 30 s. 5. 72 °C 10 min.	30	D: GGTGACGACGGCAAGCTGTA
			I: GTATCGCCTTCGGCACAGGA
<i>trpE</i> (443 pb)	1. 94 °C 5 min. 2. 94 °C 30 s. 3. 55 °C 1 min. 4. 72 °C 30 s. 5. 72 °C 10 min.	35	D: TTCAACTTCGGCGACTTCCA
			I: GGTGTCCATGTTGCCGTTCC

## IV. RESULTADOS

### 1. Características de la cohorte

Se recogieron 205 muestras (105 heces y 100 lavados nasofaríngeos, LN) pertenecientes a 24 niños nacidos entre enero de 2018 y septiembre de 2019 en cuatro centros hospitalarios de referencia de FQ, recogidas entre el primer mes de vida y los 24 meses. Se descartaron 53 muestras y 10 niños por falta de datos demográficos, clínicos, y/o por tener muestras repetidas insuficientes, así como cuatro muestras que, aunque se procesaron, tuvieron escaso rendimiento en la extracción de ADN y baja cantidad de materia biológica que impidió tener alícuota para su cultivo. Finalmente, fueron 14 los niños reclutados de tres hospitales con un total de 148 muestras (73 heces y 75 LN) para ser procesadas mediante secuenciación, culturómica, o ambas. La mediana de muestras de heces y LN por paciente fueron 5 (0-9).

En la **Tabla 9** se encuentran las características basales clínicas y demográficas de la cohorte estudiada. Las combinaciones de alelos mutantes del gen CFTR fueron: F508del/F508del (6 niños), F508del/3272-26A>G (1), F508del/N1303K (1), F508del/R1162X (1), F508del/Y1092X (1), G542X/R1162X (1), G673X/N1303K (1), L206W/R258G (1) y R334W/F508del (1). La mediana de la edad de introducción a los sólidos fueron los 7 meses (3-10).

### 2. Evolución sintomática, tratamientos y principales colonizaciones patogénicas

Toda la información relativa a las variables clínicas, posología de tratamientos e ingresos hospitalarios se encuentra en **Anexos**. En total, se detectaron 12 patógenos cultivables en los cultivos rutinarios de las consultas: *Chryseobacterium indologenes* (2 pacientes), *E. coli* (2), *H. influenzae* (3), *Haemophilus. parainfluenzae* (1), *K. oxytoca* (1), *K. pneumoniae* (1), *Moraxella catarrhalis* (2), *P. aeruginosa* (9), *S. marcescens* (3), *S. aureus* (10), y *S. maltophilia* (1). La primocolonización de *P. aeruginosa* comenzó a los (mediana de meses (RQ)) 9 (9-20) meses, mientras que la de *S. aureus* a los 6,5 (2-16) meses.

Doce niños tuvieron, al menos, una exacerbación respiratoria, entendida como evento de dificultad respiratoria, acompañado de febrícula, bronquiolitis, tos y secreciones purulentas, y siete de ellos tuvieron que ser ingresados en alguna ocasión a consecuencia. El motivo del ingreso por exacerbación respiratoria estaba asociado a infecciones de *S. aureus*, *C. indologenes*, *H. influenzae* y *P. aeruginosa*. Otros ingresos estuvieron relacionados con fallo de medro (paciente #2, #10 y #11), fiebre y otalgia (#3),

íleo meconial (#11, #14), deshidratación (#6 y #14) y sepsis causada por *S. marcescens* (#11). La mediana de días de ingreso fue de 5 días (rango 1-31 días).

Se empleó la vía intravenosa para administrar cefotaxima, cefuroxima y tobramicina en tres niños debido a infecciones por *C. indologenes* y *P. aeruginosa*. Se administraron ciclos de tobramicina y colismetato *on-off* por vía inhalada a tres pacientes que presentaron colonización intermitente de *P. aeruginosa*. En otras ocasiones, esta colonización se trató con tobramicina inhalada y ciprofloxacina oral, o con cefaclor oral y colismetato inhalado. Las infecciones por *S. aureus* fueron tratadas, normalmente, con amoxicilina-clavulánico oral, mientras que en una ocasión en la que se detectó *S. aureus* resistente a la meticilina (SARM), se administró rifampicina con trimetoprim sulfametoxazol.

No se encontraron diferencias significativas entre los individuos de genotipo de alto riesgo y los de bajo riesgo en cuanto al número de exacerbaciones respiratorias o de otra sintomatología, el número de ingresos por motivos respiratorios u otros, ni del número de meses con tratamiento antibiótico (**Tabla 10**). Estos resultados deben interpretarse con precaución debido al reducido número de individuos de bajo riesgo incluidos en esta cohorte. (ver Discusión).

**Tabla 9. Características de la cohorte de FQ.**

Características basales	Cohorte													
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14
<b>Hospital</b>	1	2	2	2	2	2	3	3	2	2	2	2	2	2
<b>Mutación</b>	G542X/ R1162X	F508del/ Y1092X	F508del/ F508del	F508del/ 3272- 26A>G	F508del/ F508del	F508del/ N1303K	F508del/ F508del	F508del/ R1162X	F508del/ F508del	G673X/ N1303K	L206W/ R258G	F508del/ F508del	F508del/ F508del	R334W/ F508del
F508del en homocigosis: 43%														
<b>Genotipo</b>	AR	AR	AR	BR	AR	AR	AR	AR	AR	AR	BR	AR	AR	BR
Alto riesgo: 79%														
<b>Sexo al nacer</b>	M	V	M	V	V	V	M	V	M	M	M	V	V	V
Varón: 57%														
<b>Edad gestacional</b>	38	41	37	40	39	40	41	38	37	38	33	39	41	38
Prematuridad: 7%														
<b>Tipo de parto</b>	Vaginal	Cesárea	Vaginal	Vaginal	Vaginal	Vaginal	Vaginal	Vaginal	Vaginal	Cesárea	Cesárea	Vaginal	Cesárea	Vaginal
Parto vaginal: 71%														
<b>Tipo de lactancia</b>	Fórmula	Fórmula	LME	LME	LME	Fórmula	Fórmula	Fórmula	Mix	Mix	Mix	LME	Fórmula	Fórmula
Lactancia materna exclusiva (LME): 29%														
<b>Íleo meconial</b>	No	No	No	No	No	No	No	No	No	No	Sí	No	No	Sí
Íleo meconial: 14%														
<b>Insuficiencia pancreática</b>	Sí	Sí	Sí	No	Sí	Sí	Sí	Sí	Sí	Sí	No	Sí	Sí	Sí
Insuficiencia pancreática: 86%														
<b>Heces (n=73)</b>	4	5	6	9	9	6	4	5	7	6	3	4	1	3
Mediana del número de muestras de heces/caso: 5 (1-9)														
<b>LN (n=75)</b>	4	7	7	7	7	6	5	3	7	5	2	4	0	3
Mediana del número de muestras de LN/caso: 5 (0-7)														

**Tabla 10. Variables clínicas recogidas a lo largo de los dos primeros años en función del genotipo FQ.**

Variables clínicas <sup>1</sup>	Genotipo Alto riesgo (N=11)	Genotipo Bajo riesgo (N=3)	p valor
<b>Exacerbación respiratoria</b>	9 (82%)	2 (67%)	0,17
Nº (en dos años)	3 (2-5)	1 (0-2)	0,24
<b>Exacerbación (otra)</b>	4 (36%)	2 (67%)	0,54
Nº (en dos años)	0 (0-1)	2 (1-3)	0,19
<b>Ingresos respiratorios</b>	5 (45%)	2 (67%)	1
Nº (en dos años)	0 (0-1)	1 (0-1)	0,73
<b>Ingresos (otro)</b>	3 (27%)	2 (67%)	0,5
Nº (en dos años)	0 (0-1)	2 (1-3)	0,12
<b>Antibióticos (meses)</b>	5 (3-9,5)	12 (6,5-12,5)	0,64
<b>Primocolonización de <i>P. aeruginosa</i></b>	7 (64%)	2 (67%)	1
Edad (meses)	19 (9-22)	4,5 (3-6)	0,06
<b>Primocolonización de <i>S. aureus</i><sup>1</sup></b>	8 (73%)	2 (67%)	0,77
Edad (meses)	4,5 (1,5-13)	14 (6-22)	0,36

### 3. Cultivo y tipado de las principales especies bacterianas

De las 73 heces sembradas, se recuperaron y tiparon microorganismos de las especies bacterianas descritas en el apartado de Materiales y Métodos en 69 (95%) (**Figuras 19-22**). Las especies más frecuentes fueron *E. faecalis* (EFL) ( $\alpha_h=0,61$ ), *E. coli* (EC) ( $\alpha_h=0,56$ ) y *E. faecium* (EFC) ( $\alpha_h=0,40$ ), mientras que las más minoritarias fueron *S. aureus* (SA) ( $\alpha_h=0,02$ ) y las especies de *Streptococcus* ( $\alpha_h=0,08$ ). Sólo se recuperó un

<sup>1</sup> Los datos sobre primocolonización se expresan en prevalencia (número de casos (% dentro del grupo)) y se analizaron con la prueba de Fisher. El resto, se evaluó con la prueba de Mann Whitney y se expresan en mediana (RQ).

aislado de *S. marcescens* (#11) a partir de un niño que sufrió sepsis por este patógeno durante su estancia hospitalaria en el primer mes. En cuanto a los clones, los más frecuentes en la cohorte fueron EC-19 ( $\alpha_h=0,11$ ), EFL-7, y EFC-21 ( $\alpha_h=0,07$ ), y SA-4 ( $\alpha_h=0,05$ ). De las especies de *Streptococcus*, el más frecuente fue *S. lutetiensis* ( $\alpha_h=0,10$ ), seguido de *S. gallolyticus* sbsp. *pasteurianus* ( $\alpha_h=0,04$ ) y *S. gallolyticus* sbsp. *gallolyticus* ( $\alpha_h=0,03$ ). En 15 ocasiones se observó que el mismo clon se detectaba en dos o más individuos diferentes (en su mayoría *E. coli*), generalmente del mismo hospital regional, salvo el caso anecdótico del *S. aureus* SA-7 (ST398), que fue detectado en dos niños de hospitales y regiones diferentes. Destacaron algunos casos (#4, #5) en los que más de dos clones diferentes se sucedían y/o alternaban a lo largo del tiempo, aunque pertenecían a pulsotipos muy relacionados (**Figuras 23-25**).

De los 75 LN cultivados, se recuperaron y tiparon microorganismos cultivables en 55 (73%) de ellos. Las especies más recuperadas fueron las de *Streptococcus* ( $\alpha_h=0,78$ ), en particular *S. mitis* ( $\alpha_h=0,24$ ), *S. vestibularis* ( $\alpha_h=0,23$ ) y *S. salivarius* ( $\alpha_h=0,22$ ). Después, las especies más frecuentes fueron *S. aureus* ( $\alpha_h=0,41$ ), *E. faecalis* ( $\alpha_h=0,41$ ) y *K. pneumoniae* ( $\alpha_h=0,16$ ), mientras que las minoritarias fueron *E. faecium* ( $\alpha_h=0$ ) y *K. oxytoca* ( $\alpha_h=0,02$ ). Se volvió a recuperar un aislado de *S. marcescens* del neonato #11, y también se encontró en el LN de otro neonato, el #1, a los 2 y 8 meses, el cual sufrió una exacerbación respiratoria y un ingreso por la colonización patogénica de esta bacteria. En cuanto a los clones, los más frecuentes en la cohorte fueron KP-12B ( $\alpha_h=0,15$ ), SA-1B ( $\alpha_h=0,14$ ) y EFL-13 ( $\alpha_h=0,08$ ). El mismo clon se detectó en dos o más individuos diferentes hasta en 6 ocasiones (generalmente clones de *S. aureus*), siempre del mismo hospital regional (**Figuras 23-25**).

La edad de colonización por *S. aureus* detectada en los cultivos fue ligeramente inferior a la informada por consulta, de 1 (1-7) mes ( $p<0,05$ ). Estos aislados de *S. aureus* se correspondieron con los siguientes complejos clonales (CC): ST8-CC8, ST26, ST22-CC22, ST509, ST30-CC30, ST4849-CC15 y ST398 (**Figura 22B**). Sólo se recuperó un aislado de *P. aeruginosa* (ST3446) obtenido a partir de un LN, (caso #4, 6 meses), y un aislado de heces (#5, 3 meses), que no sobrevivió y no pudo tiparse.

Otras bacterias recuperadas de los caldos de cultivo, pero no tipadas, pertenecían en su mayoría a la especie *Staphylococcus epidermidis*, a especies de *Enterococcus*, de *Clostridium*, de *Bifidobacterium* o especies de *Lactobacillus*.

CAPÍTULO II.  
RESULTADOS

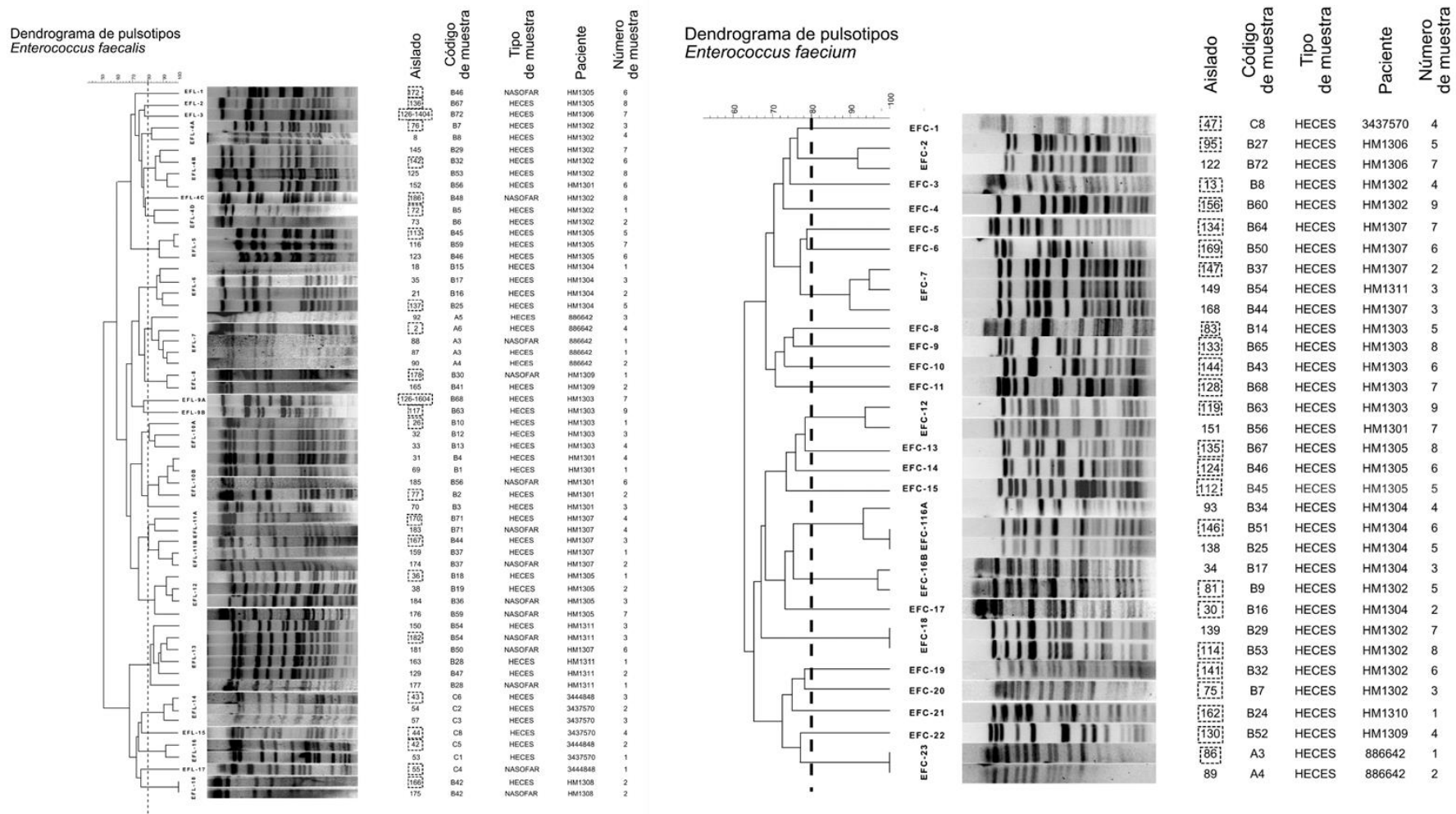


Figura 19. Dendrogramas de los pulsotipos obtenidos por PFGE de las especies *E. faecalis* y *E. faecium*. Recuadros: aislados elegidos como representantes de esos pulsotipos.

Dendrograma de pulsotipos  
*Staphylococcus aureus*

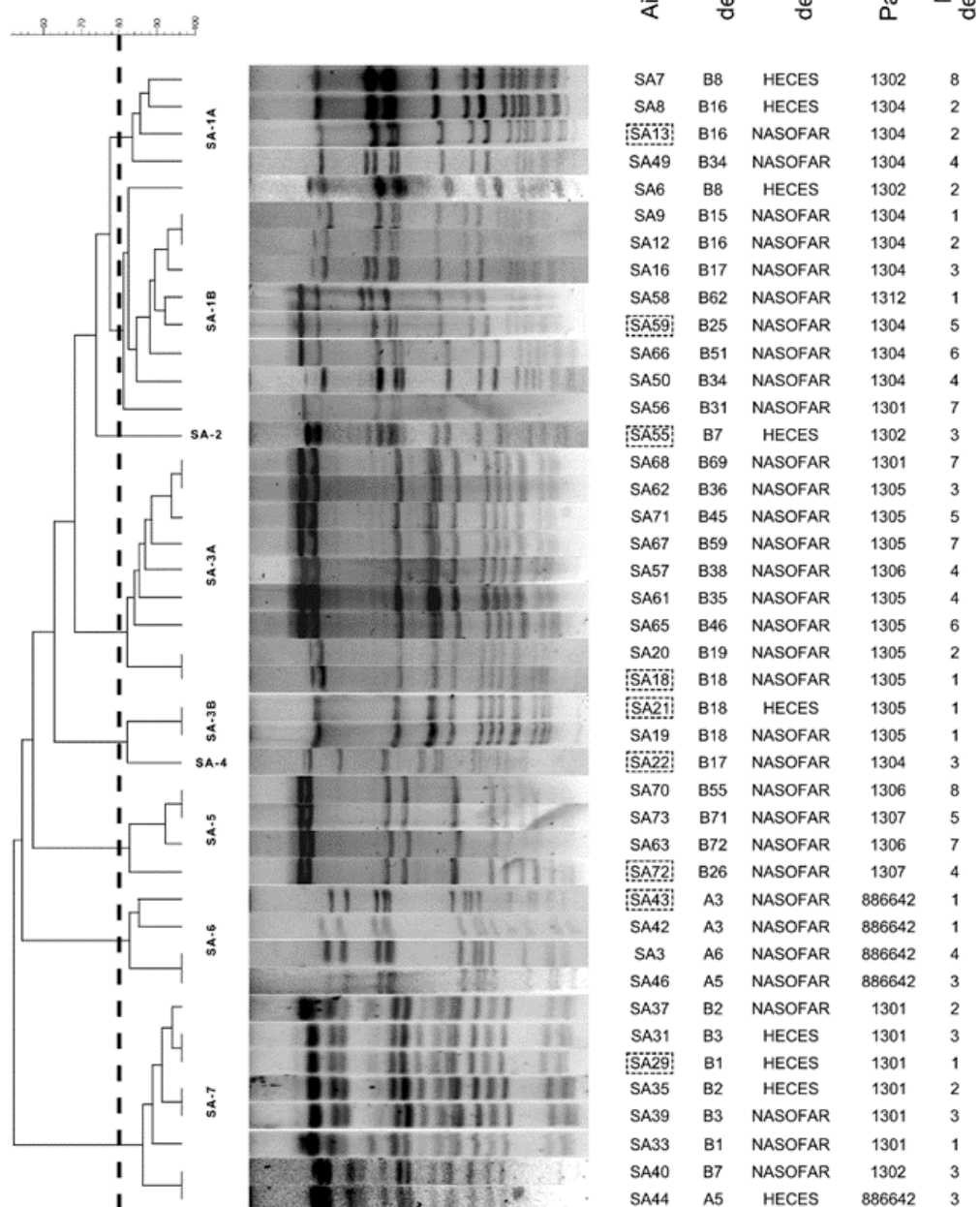


Figura 20. Dendrograma de los pulsotipos obtenidos por PFGE de la especie *S. aureus*. Recuadros: aislados elegidos como representantes de esos pulsotipos.

CAPÍTULO II.  
RESULTADOS

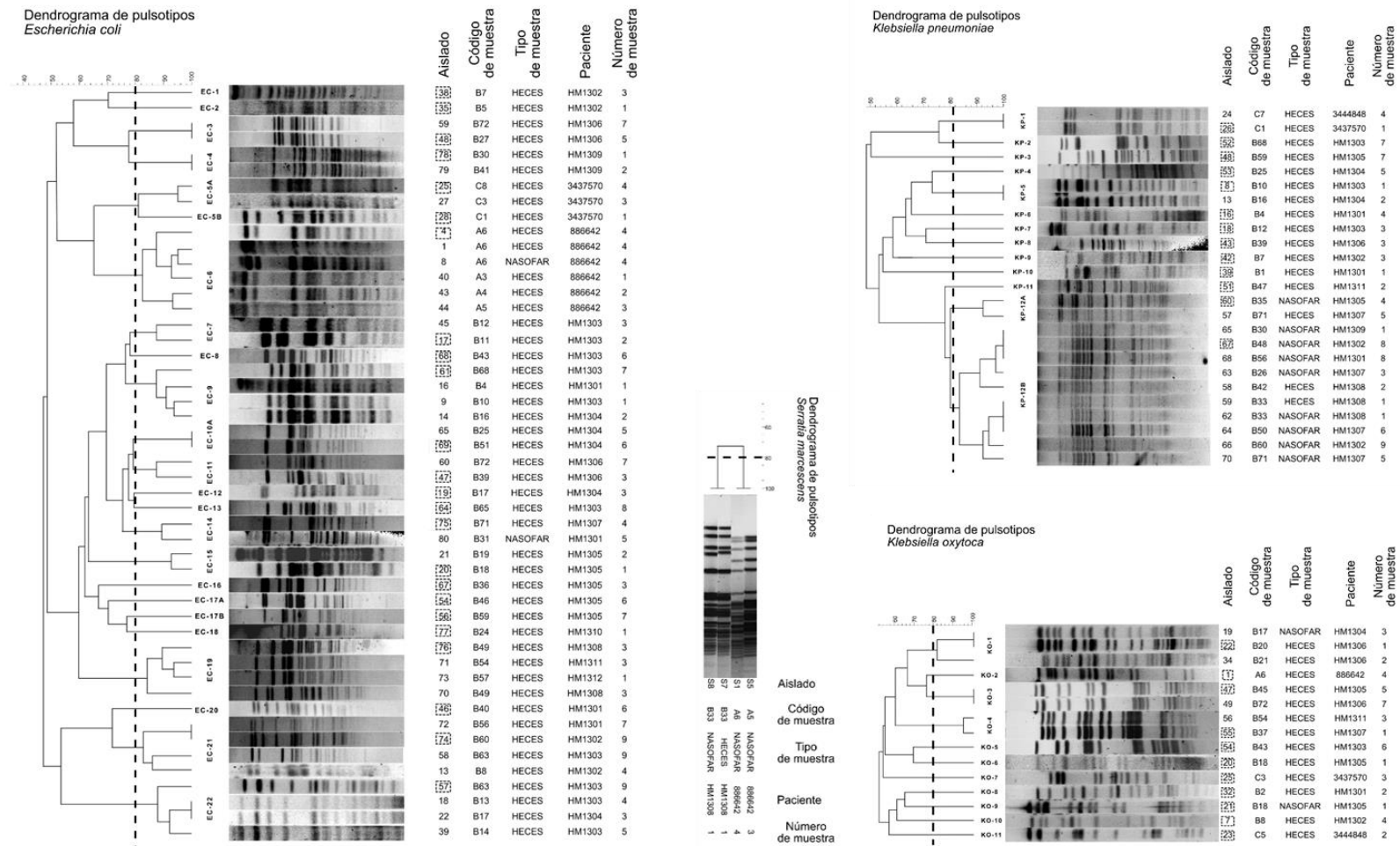
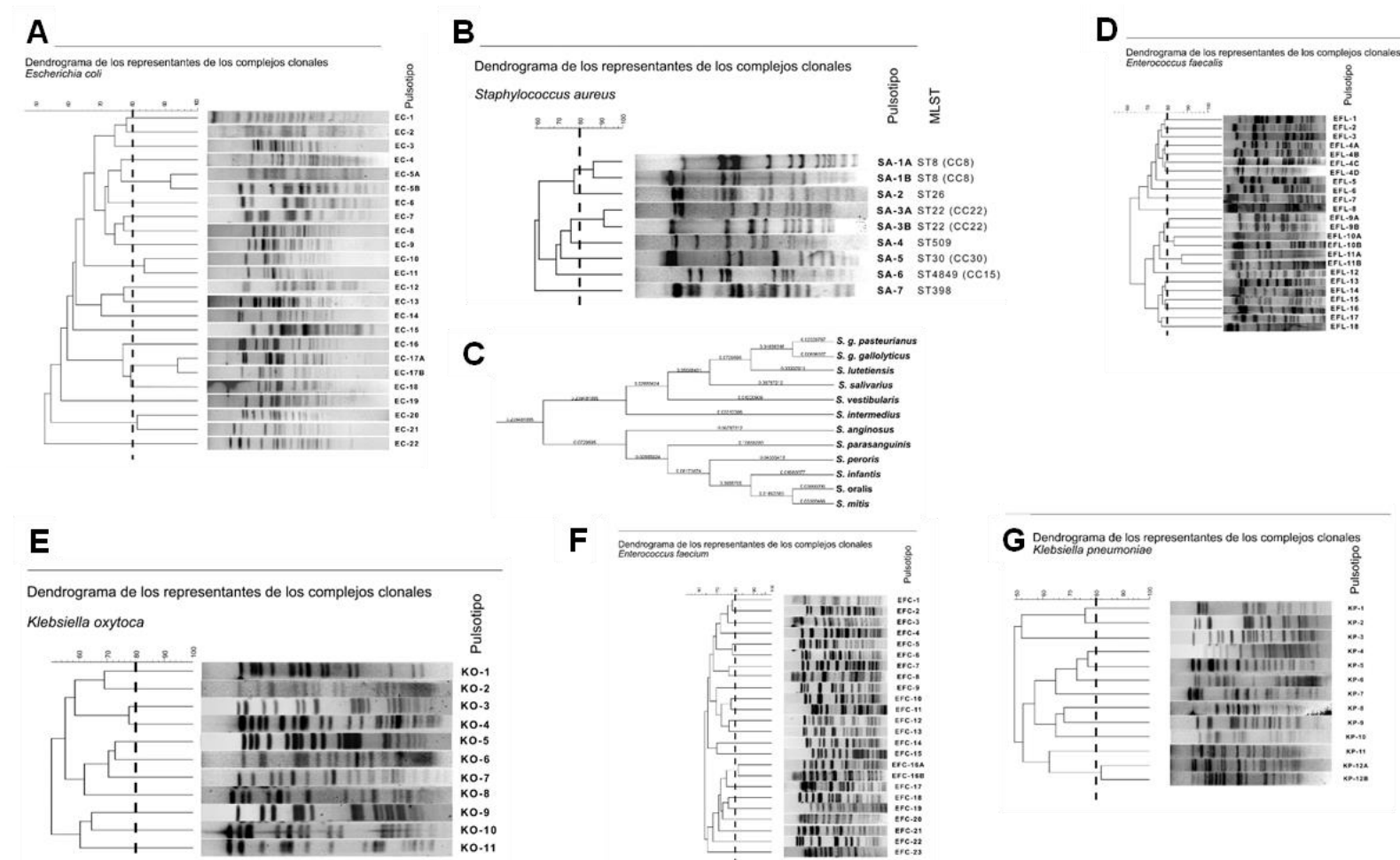


Figura 21. Dendrogramas de los pulsotipos obtenidos por PFGE de las especies Gram negativas *E. coli*, *K. pneumoniae*, *K. oxytoca* y *S. marcescens*. Recuadros: aislados elegidos como representantes de esos pulsotipos.



**Figura 22.** Representantes clonales de los aislados obtenidos de muestras fecales y nasofaríngeas de la cohorte de FQ. Dendrogramas de las especies *E. coli* (A), *S. aureus* (B), *E. faecalis* (D), *K. oxytoca* (E), *E. faecium* (F) y *K. pneumoniae* (G). C: árbol con las relaciones de similitud de las secuencias consenso de las especies de *Streptococcus*. En el dendrograma de *S. aureus* se incluye el MLST de cada pulsotipo.

CAPÍTULO II.  
RESULTADOS

Heces-EFL												LN-EFL											
Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m
#1	7	7	7					7				#1	7		NC					NC			
#2	NC	NC				NC	NC				3	#2						NC	NC	NC		NC	NC
#3	10B	10B			10B						10B	#3	NC	NC			NC						
#4	4D			4D			4A	4A				#4				NC		NC					
#5	10A				10A	10A			NC			#5					NC	NC		NC		NC	
#6	6	6					6		NC		6	#6	NC	NC					NC		NC		NC
#7	16				14	14	15					#7					NC	NC	NC	NC			
#8				16		14						#8	17						NC				
#9	12			12						5		#9	NC			NC		12		NC		NC	1
#10		11B				11B				11A		#10		11B						NC		11A	
#11	NC			18				NC				#11	NC				18						
#12	NC	8					NC					#12	8	NC			NC						
#13	NC											#13	NC										
#14	13			13		13						#14	13					13					
Heces-EFC												LN-EFC											
Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m
#1	23	23										#1											
#2	NC	NC					2				2	#2						NC	NC	NC		NC	NC
#3	NC	NC			NC							#3	NC	NC			NC						NC
#4	NC			NC			20	3			16B	#4				NC			NC				
#5	NC			NC	NC	NC		8			10	#5					NC	NC		NC		NC	
#6	NC	17					16B			16A	16A	#6	NC	NC					NC		NC		NC
#7	NC				NC	NC	1					#7	NC				NC		NC		NC		
#8				NC	NC	NC						#8	NC				NC	NC					
#9	NC			NC	NC	NC				15	14	#9	NC			NC		NC	NC		NC		NC
#10		7				7						#10		NC						NC		NC	
#11				NC				NC				#11	NC				NC						
#12		NC					22					#12	NC	NC			NC						
#13	21											#13											
#14						7						#14											
Heces-SA												LN-SA											
Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m
#1	NC	NC	7									#1	6		6						6		5
#2	NC	NC				NC	NC				NC	#2						NC	NC	3A			5
#3	7	7			7							#3	7	7			7						5
#4	NC			NC			2	1A, 1B				#4				NC			7				
#5	NC			NC	NC	NC						#5					NC	NC				NC	
#6	NC	1A										#6	1B	1A, 1B					4, 1B		1A, 1B		1B
#7	NC			NC	NC	NC						#7					NC	NC	NC		NC		
#8				NC	NC	NC						#8	NC					NC					
#9	3B			NC	NC	NC					NC	#9	3A, 3B			3A	3A	3A	3A		3A		3A
#10	NC	NC				NC					NC	#10		NC						5	5		
#11	NC	NC						NC				#11	NC				NC						
#12	NC	NC										#12	NC	NC			NC						
#13	NC											#13											
#14	NC											#14	NC										
Heces-ST												LN-ST											
Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m
#1	NC	NC	NC									#1	5, 12		12						12		
#2	NC	NC				1	1				NC	#2						4, 7, 9, 12	12	7, 12		7	12
#3	NC	NC			NC						NC	#3	8	4			12						
#4	NC			NC			NC				NC	#4				4, 5, 11			4, 11				
#5	NC			NC	NC	NC		2			2	#5					4, 8	4, 5, 8		5, 12		5, 12	
#6	NC	NC					NC			3		#6	NC	NC					4		4, 8		7, 11
#7	NC			NC	NC	NC						#7					4	4	10		8		
#8				NC	NC	NC						#8	4					8, 9					
#9	3			1, 3		NC					NC	#9	5			5			NC		4		NC
#10		NC					NC				NC	#10		11						NC		11	
#11	NC	NC						NC				#11	NC			11							
#12	NC	NC										#12	NC	5, 12			11						
#13	3											#13											
#14	NC											#14	12			4, 5		NC					

**Figura 23. Detección de clones de las especies Gram positivas identificadas en las muestras fecales y en lavados nasofaríngeos (LN) de los 0 a 12 meses. EFL: *E. faecalis*; EFC: *E. faecium*; SA: *S. aureus*; ST: *Streptococcus*. ST-1: *S. g. pasteurianus*; ST-2: *S. g. gallolyticus*; ST-3: *S. lutetiensis*; ST-4: *S. salivarius*; ST-5: *S. vestibularis*; ST-6: *S. intermedius*; ST-7: *S. anginosus*; ST-8: *S. parasanguinis*; ST-9: *S. peroris*; ST-10: *S. infantis*; ST-11: *S. oralis*; ST-12: *S. mitis*.**

Hececs-EC												LN-EC												
Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	
#1	6	6	6					6				#1	NC	NC	NC					6				
#2	NC	NC				11	3			3, 11		#2	NC					NC	NC	NC				
#3	NC	NC			NC					9	20	#3	NC	NC		NC	NC							
#4	2			NC			1	21				#4			NC				NC					
#5	9			7	7	22		22		8		#5				NC	NC			NC	NC			
#6	NC	9					12, 22			NC		#6	NC	NC						NC	NC			NC
#7	5B				NC	5A	5A					#7					NC	NC			NC			NC
#8					NC	NC	NC					#8	NC					NC	NC					NC
#9	15			15		16				NC	17A	#9	NC			NC			NC			NC		NC
#10	NC	NC		NC		NC				14		#10	NC	NC						NC	NC			
#11	NC							19				#11	NC			NC								
#12	4	4					NC					#12	NC	NC			NC							
#13	18											#13												
#14	NC			NC		19						#14	NC			NC			NC					
Hececs-KO												LN-KO												
Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	
#1	NC	NC	NC					2				#1	NC		NC					NC				
#2	1	1				NC	NC			3		#2						NC	NC	NC			NC	NC
#3	NC	8			NC					NC	NC	#3	NC	NC			NC							
#4	NC			NC				10				#4			NC				NC					
#5	NC			NC	NC	NC				5		#5					NC	NC			NC			NC
#6	NC	NC					NC			NC		#6	NC	NC						1		NC		NC
#7	NC				NC	7	NC					#7					NC	NC		NC				NC
#8				11		NC	NC					#8	NC					NC						
#9	6			NC		NC				3	NC	#9	9			NC		NC		NC				NC
#10		4				NC						#10		NC							NC	NC		
#11	NC			NC				NC				#11	NC			NC								
#12	NC	NC					NC					#12	NC	NC			NC							
#13	NC											#13	NC											
#14	NC			NC		4						#14	NC			NC			NC					
Hececs-KP												LN-KP												
Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	
#1	NC	NC	NC					NC				#1	NC		NC					NC				
#2	NC	NC				8	NC			NC		#2						NC	NC	NC			NC	NC
#3	10	NC			NC					6	NC	#3	NC	NC			NC							
#4	NC			NC			9				NC	#4			NC				NC					
#5	5			NC	7	NC		NC		NC		#5					NC	NC		NC				NC
#6	NC	5					NC			NC	4	#6	NC	NC					NC		NC			NC
#7	1			NC	NC	NC						#7					NC	NC		NC				NC
#8				NC	NC		1					#8	NC					NC						
#9	NC			NC	NC							#9	NC			NC		NC		12A				NC
#10		NC								12A		#10		NC							12B	12B		
#11	12B			12B						NC		#11	12B			NC								
#12	NC	NC					NC					#12	12B	NC			NC							
#13	NC											#13	NC											
#14	NC			11		NC						#14	NC				NC							
Hececs-SM												LN-SM												
Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	Caso	0-1m	1-2m	2m	3m	4m	5m	6-7m	8m	9m	10m	12m	
#1	NC	NC	NC									#1			1						1			
#2	NC	NC				NC	NC			NC		#2						NC	NC	NC				NC
#3	NC	NC				NC				NC	NC	#3	NC	NC			NC							
#4	NC			NC			NC	NC				#4				NC			NC					
#5	NC			NC	NC	NC		NC		NC		#5					NC	NC		NC				NC
#6	NC	NC					NC			NC		#6	NC	NC					NC		NC			NC
#7	NC				NC	NC	NC					#7					NC	NC		NC				NC
#8				NC	NC	NC						#8	NC					NC						
#9	NC			NC	NC					NC		#9	NC			NC		NC		NC				NC
#10		NC				NC				NC		#10		NC							NC	NC		
#11	2			NC						NC		#11	2				NC							
#12	NC	NC					NC					#12	NC	NC				NC						
#13	NC											#13	NC											
#14	NC											#14	NC											

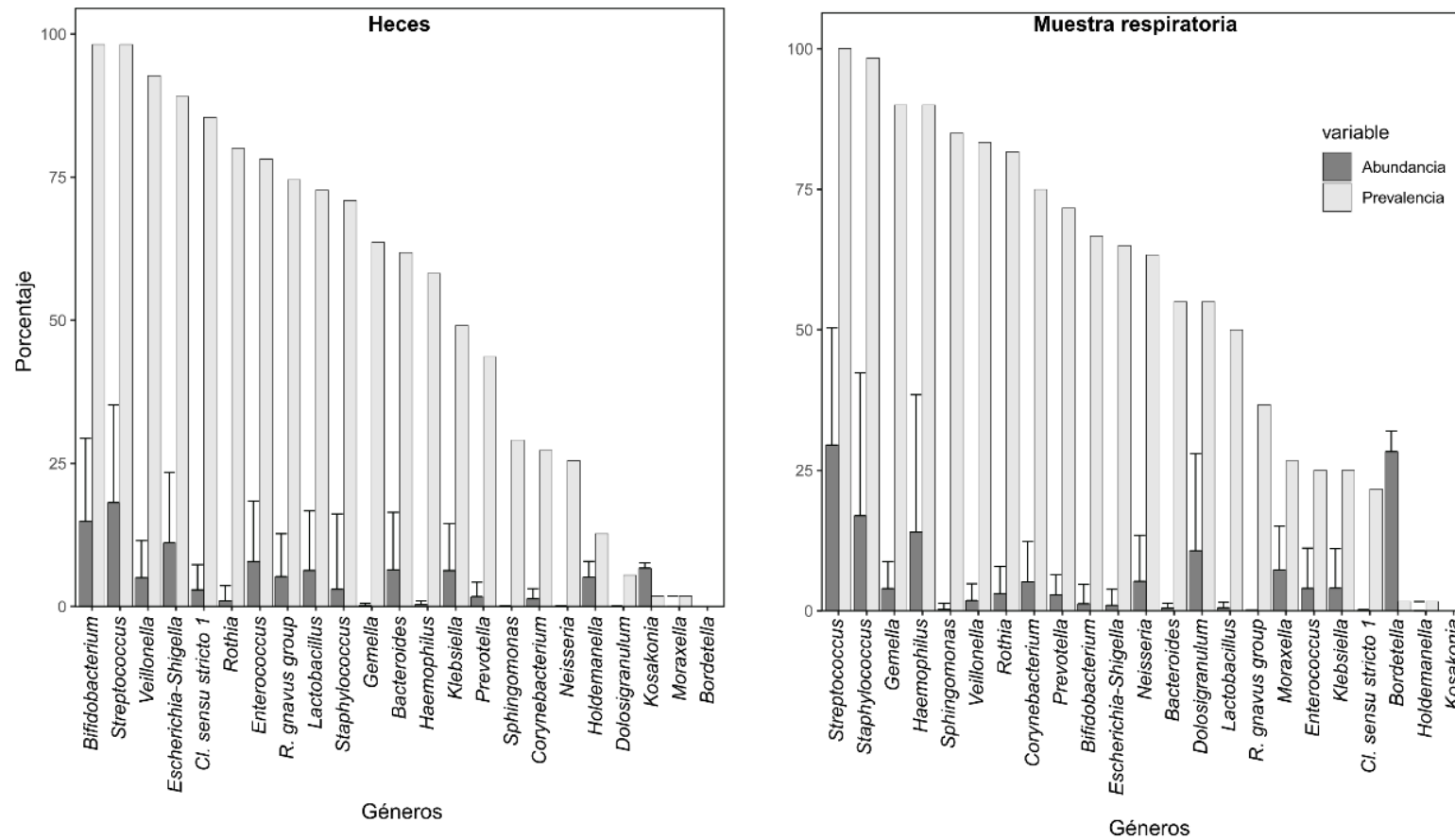
Figura 24. Detección de clones de las especies Gram negativas identificadas en las muestras fecales y en lavados nasofaríngeos (LN) de los 0 a 12 meses. EC: *E. coli*; KP: *K. pneumoniae*; KO: *K. oxytoca*; SM: *S. marcescens*.

Heces-EFL					LN-EFL					Heces-EC					LN-EC								
Caso	14-15m	16-18m	21-22m	23m	24m	Caso	14-15m	16-18m	21-22m	23m	24m	Caso	14-15m	16-18m	21-22m	23m	24m	Caso	14-15m	16-18m	21-22m	23m	24m
#3				4B		#3				10B		#3				21		#3	14				
#4	4B	4B	4B			#4			4C			#4					21	#4					
#5	9A		9B			#5						#5	9	13	21, 22			#5					
#6						#6						#6	10C					#6					
#9	5	2				#9			12			#9	17B					#9					
#10						#10			13			#10						#10					
Heces-EFC					LN-EFC					Heces-KO					LN-KO								
#3				12		#3						#3						#3					
#4	19	18	18		4	#4						#4						#4					
#5	11	9	12			#5						#5						#5					
#6	16A					#6						#6						#6					
#9			13			#9						#9						#9					
#10	6	5				#10						#10						#10					
Heces-SA					LN-SA					Heces-KP					LN-KP								
#3						#3	1B		3A			#3						#3					12B
#4						#4						#4						#4				12B	12B
#5						#5						#5	2					#5					
#6						#6	1B					#6						#6					
#9						#9	3A					#9	3					#9					
#10						#10						#10						#10	12B				
Heces-ST					LN-ST					Heces-SM					LN-SM								
#3						#3	11		5, 8	6, 11		#3						#3					
#4					3	#4				11	11, 12	#4						#4					
#5			2			#5		2, 11				#5						#5					
#6						#6	12					#6						#6					
#9						#9						#9						#9					
#10						#10						#10						#10					

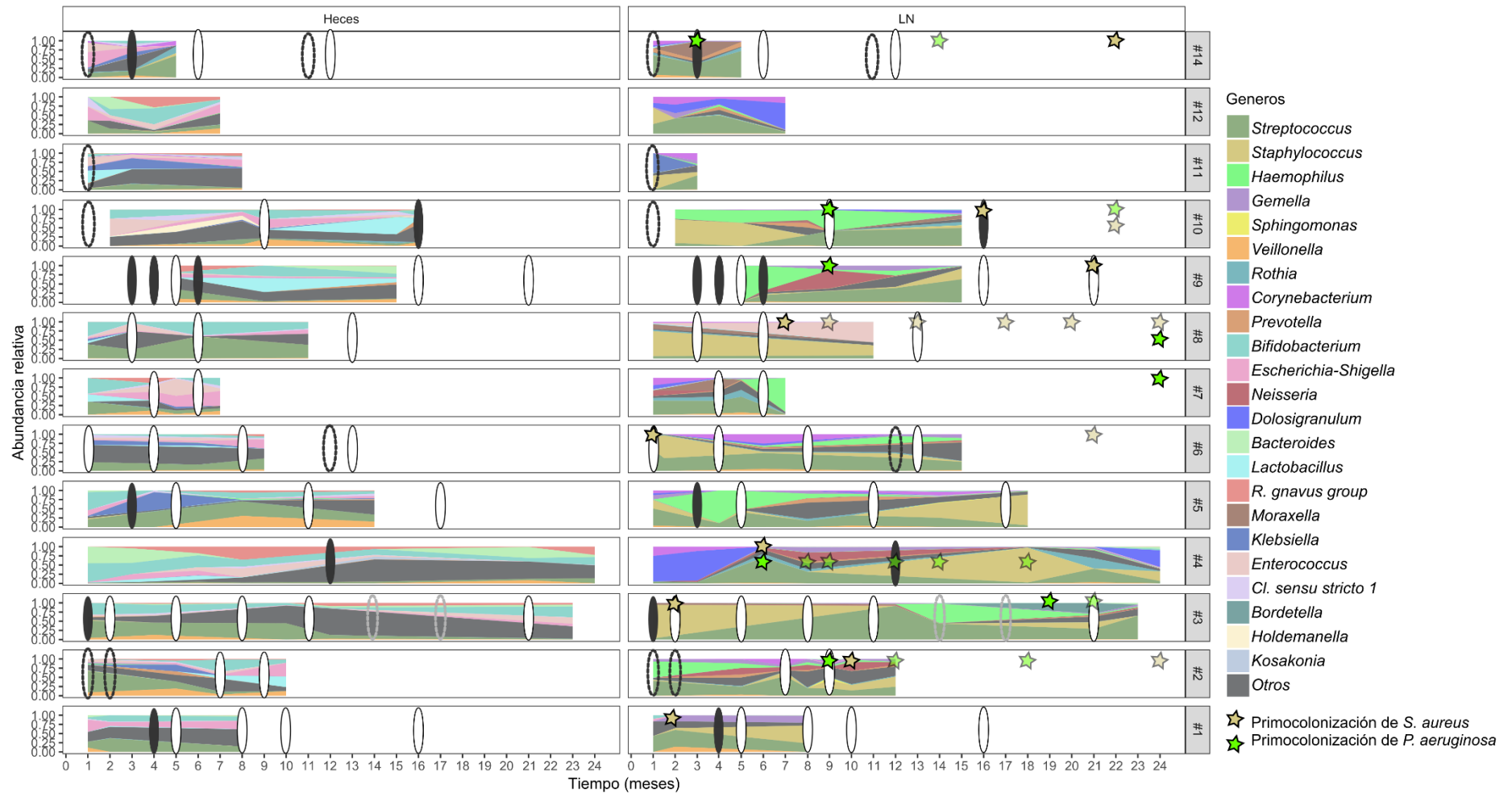
**Figura 25. Detección de clones de las especies recuperadas de heces y lavados nasofaríngeos (LN) entre los 12 y los 24 meses. Sólo los casos #3, #4, #5, #6, #7 y #8 contribuyeron con muestras. ST-1: *S. g. pasteurianus*; ST-2: *S. g. gallolyticus*; ST-3: *S. lutetiensis*; ST-4: *S. salivarius*; ST-5: *S. vestibularis*; ST-6: *S. intermedius*; ST-7: *S. anginosus*; ST-8: *S. parasanguinis*; ST-9: *S. peroris*; ST-10: *S. infantis*; ST-11: *S. oralis*; ST-12: *S. mitis*.**

#### 4. Establecimiento de la microbiota intestinal y respiratoria analizada mediante secuenciación

De las 127 muestras disponibles para la extracción de ADN bacteriano y secuenciación de amplicones, 115 se procesaron exitosamente (60 LN y 55 H). Se detectaron 2.332 ASV correspondientes a 336 géneros bacterianos, de los cuales 130 eran compartidos por todas las muestras, 133 exclusivos de LN y 73 exclusivos de heces. A continuación, se filtraron aquellos con mayor prevalencia y abundancia relativa en toda la cohorte, en función del tipo de muestra (**Figura 26**). Cabe destacar el género *Kosakonia*, presente en sólo una muestra fecal, pero con una abundancia relativa del 7%, y el género *Bordetella*, presente en sólo una muestra nasal, con una abundancia del 28%. Se analizó su cambio composicional en cada paciente a lo largo del seguimiento (al igual que del resto de bacterias), y se observó que, en ambos casos, coincidía temporalmente con la colonización patogénica de *P. aeruginosa* en los pacientes #3 (21 meses) y #4 (14 meses) (**Figura 27**).



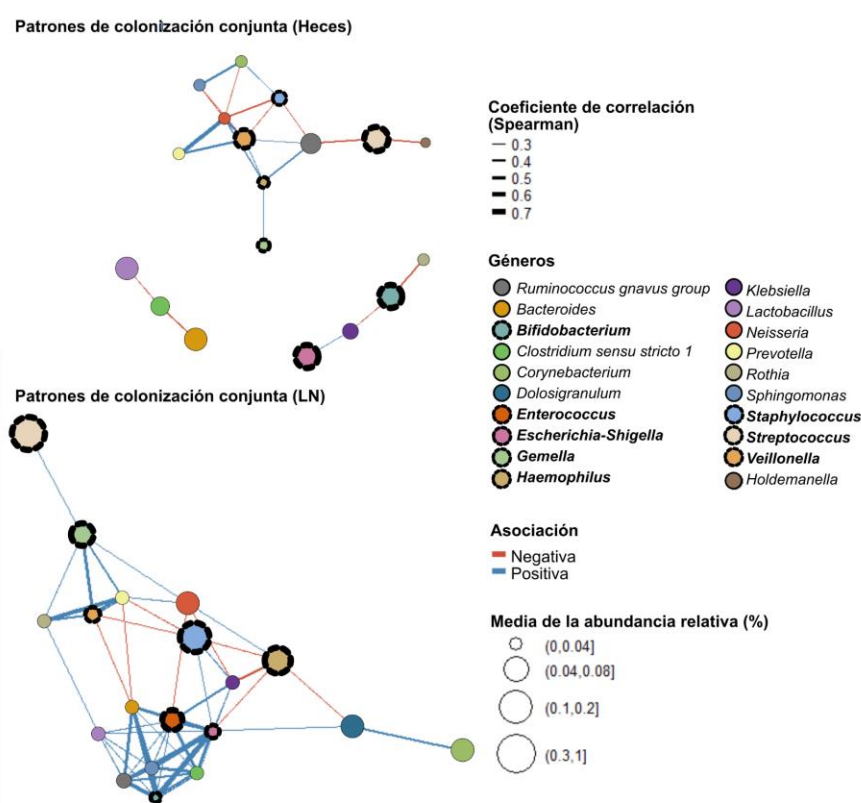
**Figura 26. Prevalencia y abundancia de los principales géneros bacterianos de la cohorte.** Muestras procesadas por secuenciación de amplicones: 55 heces y 60 muestras respiratorias.



**Figura 27. Evolución de los principales géneros bacterianos en muestras de H y LN en todos los pacientes.** Elipse negra: ingreso por exacerbación respiratoria. Elipse blanca: exacerbación respiratoria. Elipse gris discontinua: sintomatología (otro motivo). Elipse negra discontinua: ingreso (otro motivo). El paciente #13 no se representó porque sólo aportó una muestra (un LN).

#### 4.1. Evolución de los principales géneros bacterianos y la alfa diversidad de la microbiota intestinal y respiratoria

El *core* bacteriano (>50% de prevalencia en cada grupo y >1% de abundancia) de heces y LN estaba constituido por los géneros *Escherichia-Shigella*, *Veillonella*, *Haemophilus*, *Gemella*, *Staphylococcus*, *Enterococcus*, *Bifidobacterium* y *Streptococcus*. En cada nicho anatómico aparecieron patrones de colonización conjunta diferentes, siendo más complejos en las muestras respiratorias. En común, se encontraron las asociaciones positivas entre *Escherichia-Shigella* y *Klebsiella*, *Haemophilus* y *Neisseria*, y *Neisseria* con *Veillonella* y *Prevotella*. La asociación negativa común fue la de *Veillonella* y *Staphylococcus* (Figura 28).



**Figura 28. Colonización conjunta en heces y LN de las principales bacterias.** Los géneros que conforman el *core* de heces y muestras respiratorias aparecen resaltados. *Bordetella*, *Kosakonia* y *Moraxella* no aparecen por su baja prevalencia y/o por falta de asociaciones significativas. Valor de  $p=0.05$ .

Al analizar toda la cohorte en su conjunto, y teniendo en cuenta, exclusivamente, la variable tiempo, se observó una correlación positiva, tanto en heces como en muestras

respiratorias, de la edad con la abundancia relativa de *Ruminococcus*, *Haemophilus* y *Neisseria*. En heces, también hubo asociaciones positivas con *Veillonella* y *Gemella*, mientras que *Staphylococcus*, *Corynebacterium* y *Streptococcus* tenían una asociación negativa. En LN, también hubo correlación positiva entre la edad y la abundancia de *Sphingomonas*, *Bacteroides* y *Prevotella*, y no se encontraron asociaciones negativas (Figura 29). Otra prueba estadística, obtenida a partir de la función *trendyspliner* del paquete *splinctomeR* de RStudio, detectó que la tendencia de los géneros minoritarios (“otros”) y *Klebsiella* en heces era distinta de cero (Figura 30), mientras que en LN destacó *Bacteroides*, *Ruminococcus* y los géneros minoritarios (Figura 31). Hay que destacar que estos últimos resultados tienen poca contundencia, debido a que son géneros con abundancias relativas muy pequeñas y la prueba es sensible a estos casos. Los índices de diversidad alfa, en la cohorte en general, fueron significativamente mayores conforme aumentaba la edad en ambos nichos (Figura 32).

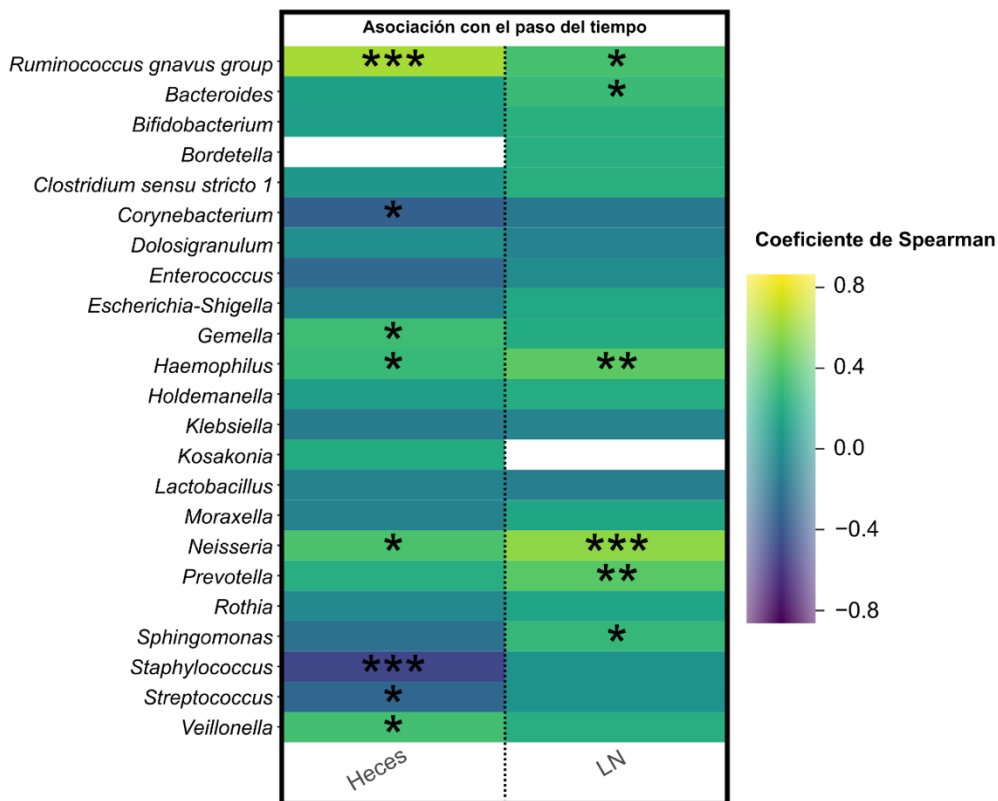
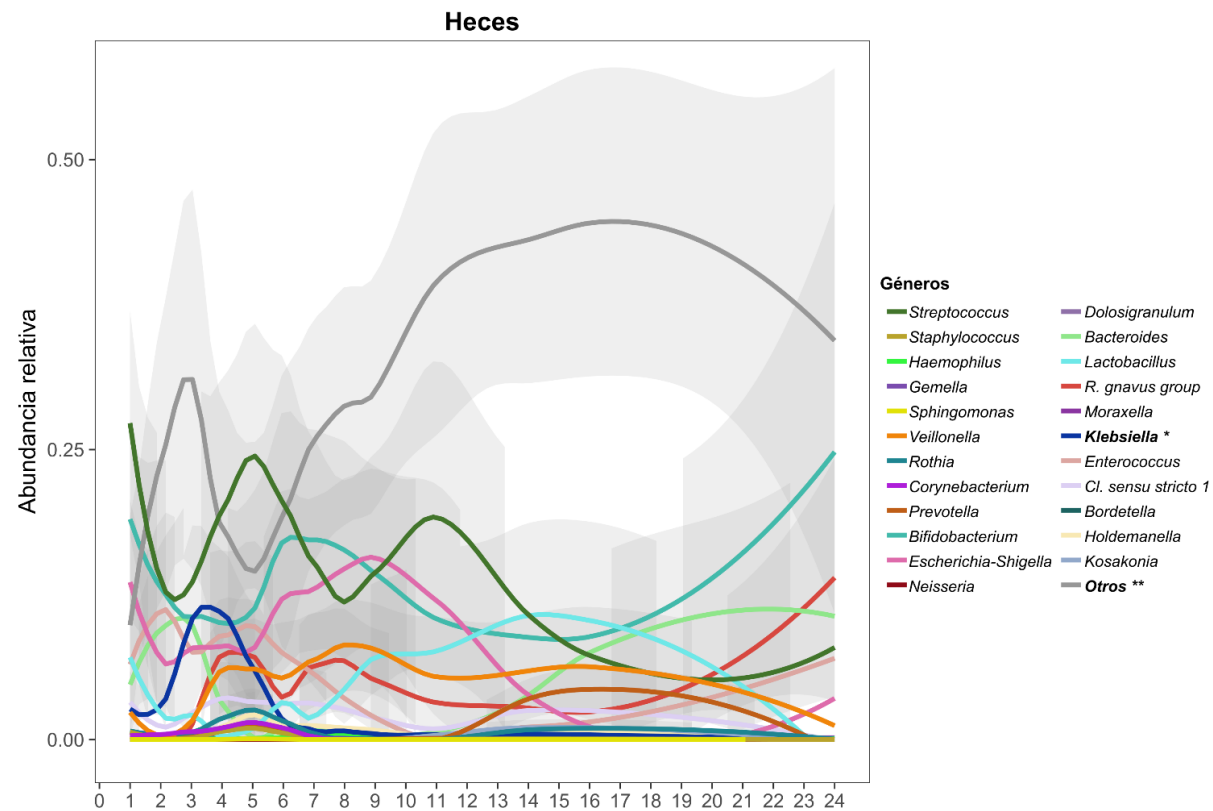
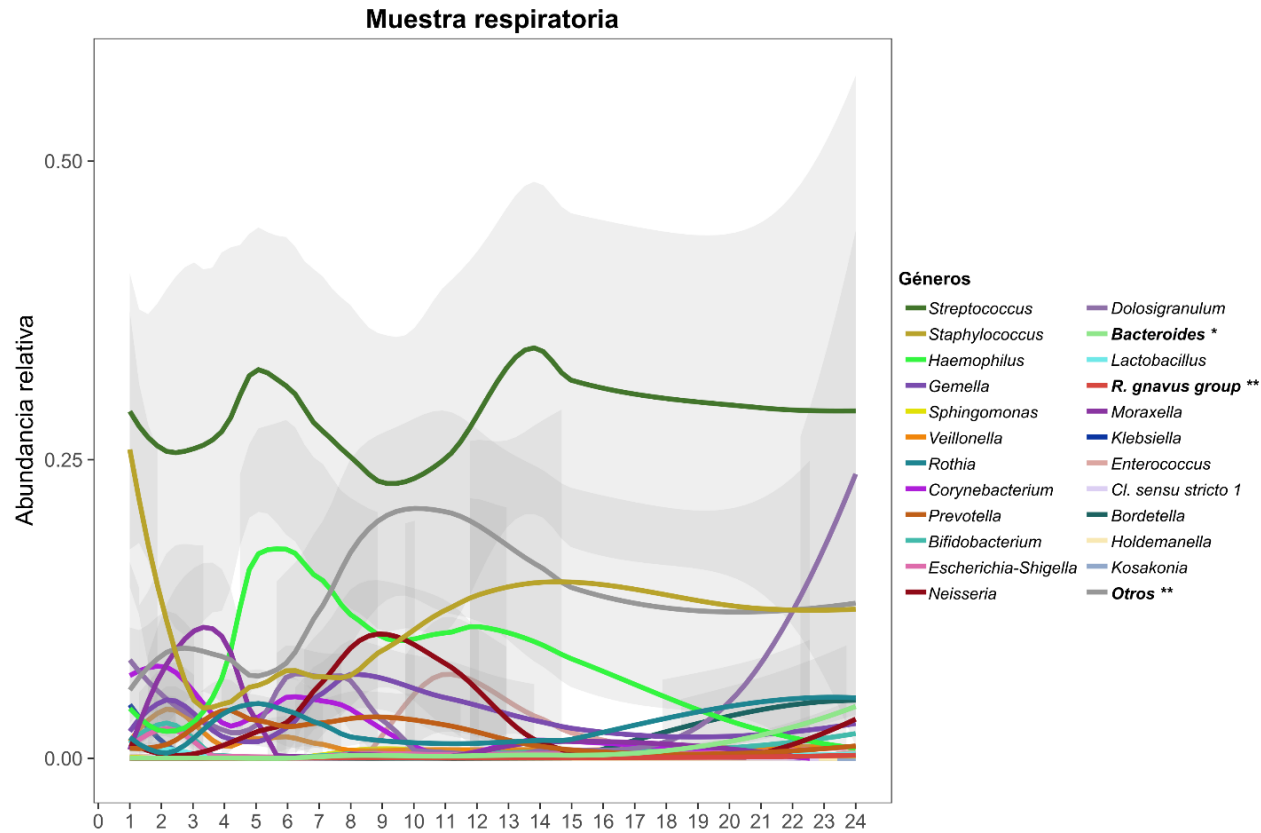


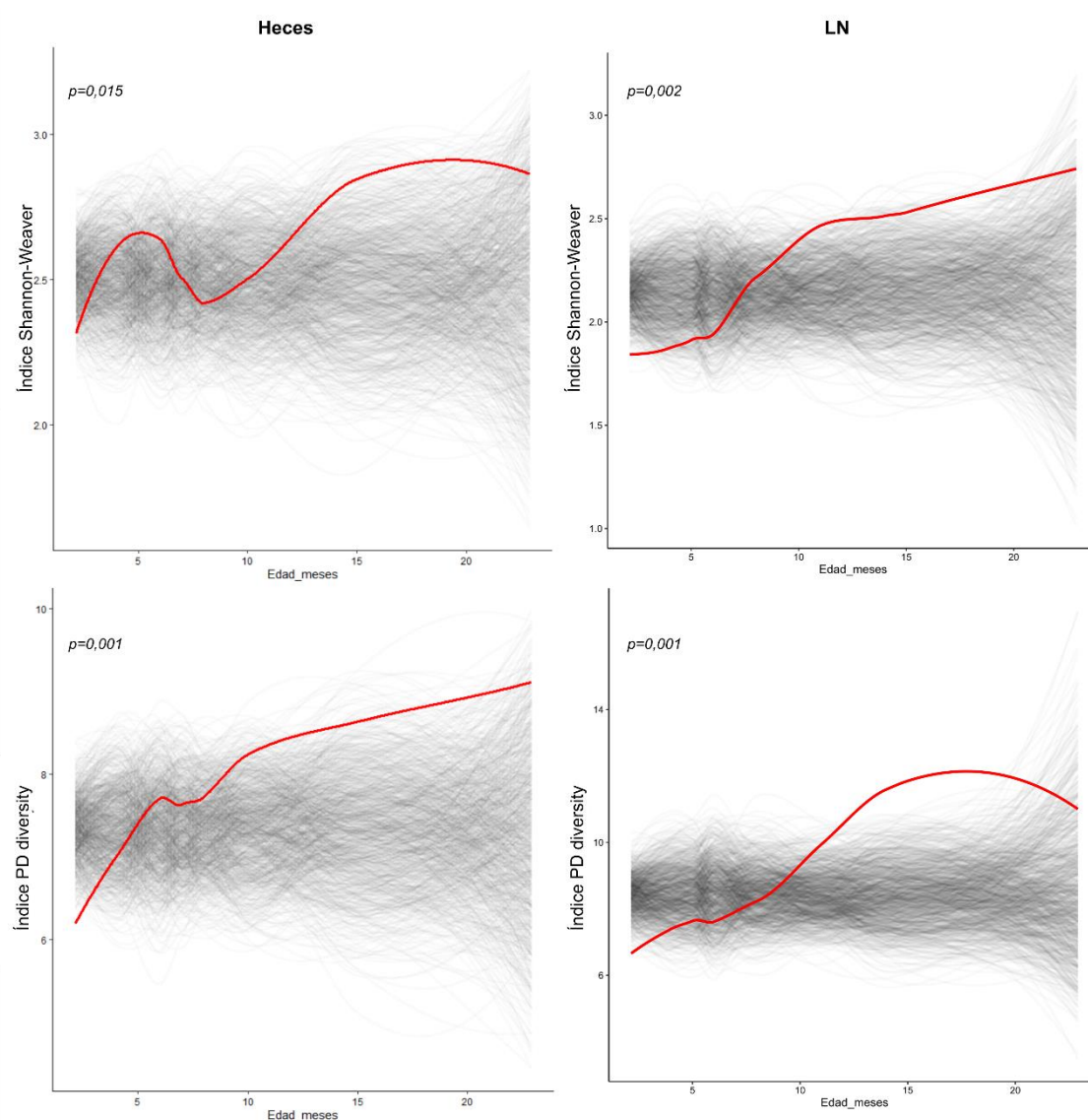
Figura 29. Correlación de los principales géneros con la edad en la cohorte de FQ. Correlación de Spearman. \*\*\*:  $p < 0,001$ ; \*\*:  $p < 0,01$ ; \*:  $p < 0,05$ .



**Figura 30. Evolución a lo largo de 2 años de la microbiota mayoritaria en heces.** Aparecen resaltados en la leyenda los géneros con una tendencia a lo largo del tiempo distinta de cero según la prueba *trendyspliner* (paquete *splinctomeR* de RStudio). \*\*:  $p < 0,01$ ; \*:  $p < 0,05$ .



**Figura 31. Evolución a lo largo de 2 años de la microbiota mayoritaria en muestras LN.** Aparecen resaltados en la leyenda los géneros con una tendencia a lo largo del tiempo distinta de cero según la prueba *trendyspliner* (paquete *splinctomeR* de RStudio). \*\*:  $p < 0,01$ ; \*:  $p < 0,05$ .

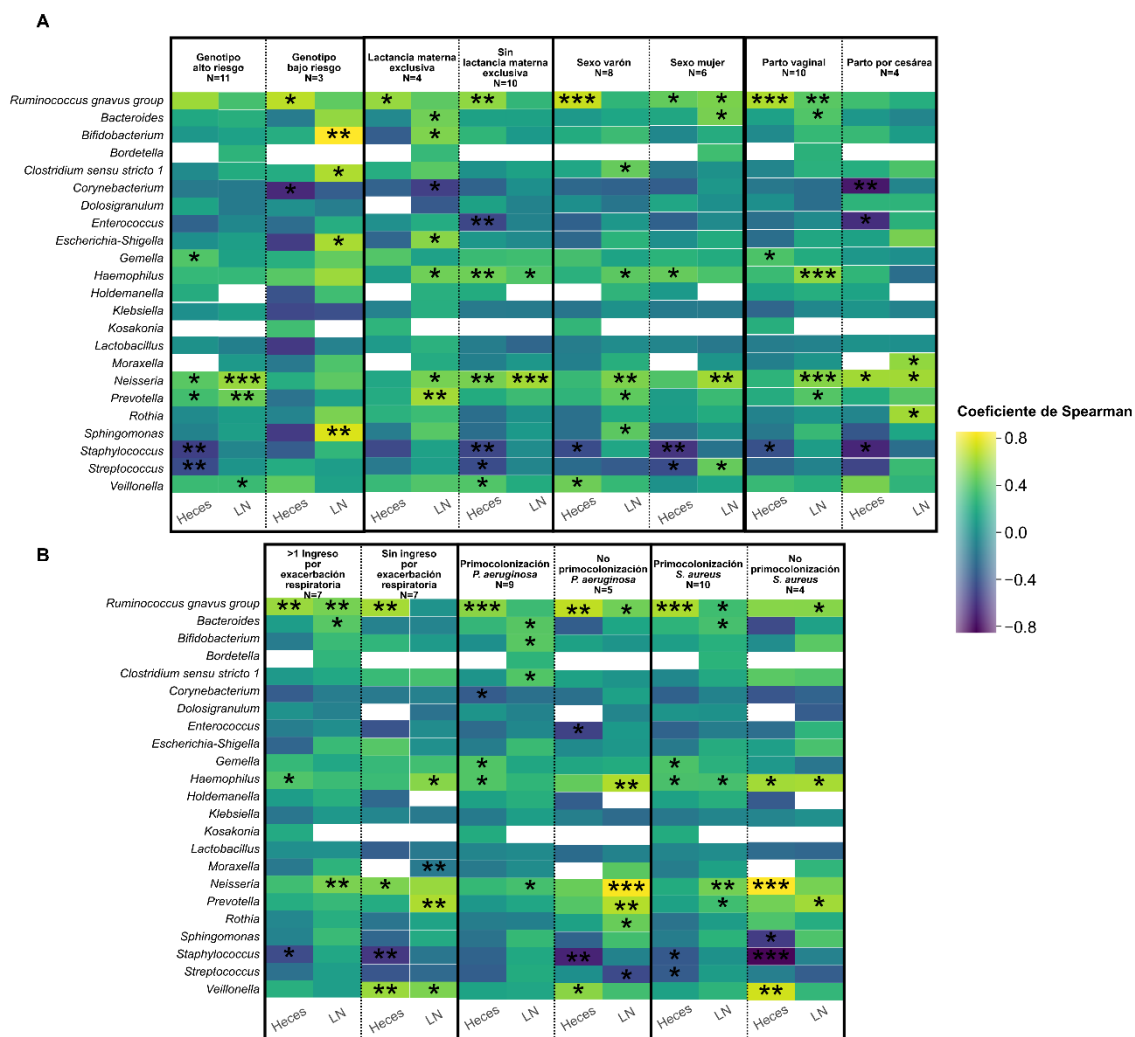


**Figura 32.** Índices de diversidad alfa a lo largo de los dos primeros años de vida en la cohorte de FQ. PD diversity: *phylogenetic diversity index*. Prueba estadística: función *trendyspliner* (paquete *splinctomeR* de RStudio) para valorar si la tendencia de la variable de diversidad alfa es distinta de cero a lo largo del tiempo. Valor  $p=0,05$ .

La composición de los géneros mayoritarios de la microbiota también se analizó en función de diversas variables, tanto basales (genotipo, sexo, tipo de parto o tipo de lactancia) como clínicas (tener o no primocolonización de *S. aureus*, primocolonización de *P. aeruginosa* y exacerbación respiratoria con ingreso en los dos primeros años de vida) (**Figura 33**). En la mayoría de los subgrupos analizados, se repitieron las asociaciones positivas y negativas anteriormente citadas. Además, se detectó que *Bifidobacterium* aumentaba con la edad en los LN de quienes tenían genotipo de bajo riesgo, lactancia materna exclusiva, y se colonizaron por *P. aeruginosa*. *Clostridium*

*sensu stricto* 1 aumentó con la edad en los LN de aquellos con genotipo de bajo riesgo, los varones y los colonizados por *P. aeruginosa*. *Enterococcus* disminuyó su composición relativa con el tiempo en las heces de quienes no tuvieron lactancia materna exclusiva, nacieron por cesárea y no se colonizaron por *P. aeruginosa*. *Rothia*, al contrario, aumentó con el tiempo en los LN de esos subgrupos. *Escherichia* aumentó en los LN de aquellos con genotipo de bajo riesgo y lactancia materna exclusiva.

La prueba de *permuspliner* y *slidingspliner* se llevó a cabo en última instancia para el análisis bivariante entre la composición relativa de cada género y estas variables clínicas (**Figura 34**). El cambio de *R. gnavus* en heces fue diferente en función del genotipo y el tipo de lactancia, predominando en genotipo de bajo riesgo y en lactancia materna exclusiva. *Bifidobacterium* y *Bacteroides* predominaron en las heces de lactancia materna exclusiva. La abundancia relativa de *Neisseria* aumentó en los lavados nasofaríngeos de aquellos que sufrieron primocolonización por *P. aeruginosa* y/o exacerbaciones con ingreso hospitalario, en comparación con quienes no. *Veillonella* aumentó su abundancia relativa en los LN de quienes no sufrieron colonización por *S. aureus* y/o quienes no se colonizaron por *P. aeruginosa*. Por último, destacar que, quienes no sufrieron exacerbaciones ni ingresos por razones respiratorias, experimentaron un mayor aumento de la diversidad alfa en sus LN en comparación con los que sí.



**Figura 33.** Asociaciones positivas y negativas de los principales géneros bacterianos con la edad en la cohorte de FQ, en función de distintas variables. **A:** variables basales. **B:** variables clínicas. Correlación de Spearman. \*\*\*:  $p < 0,001$ ; \*\*:  $p < 0,01$ ; \*:  $p < 0,05$ .

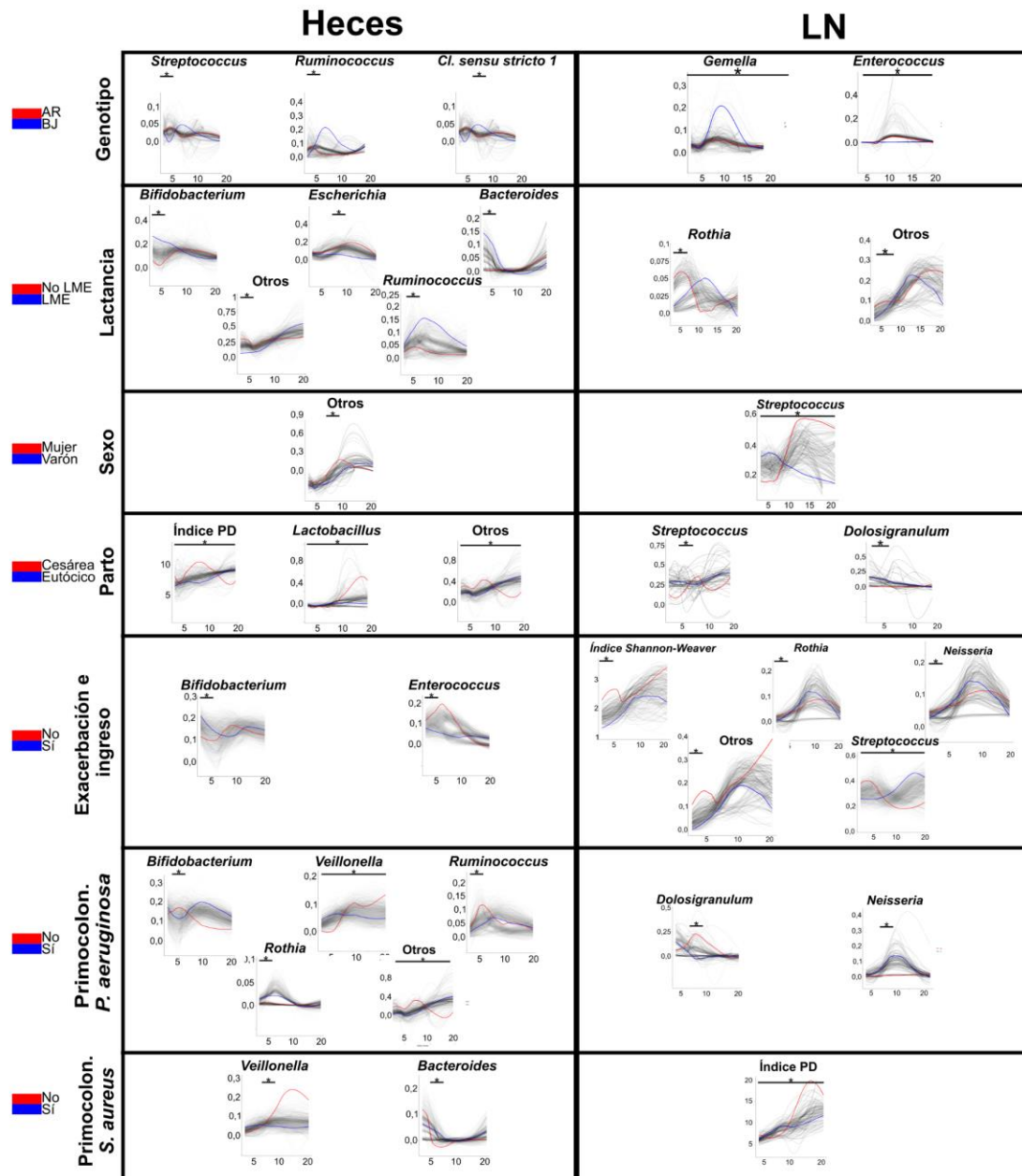
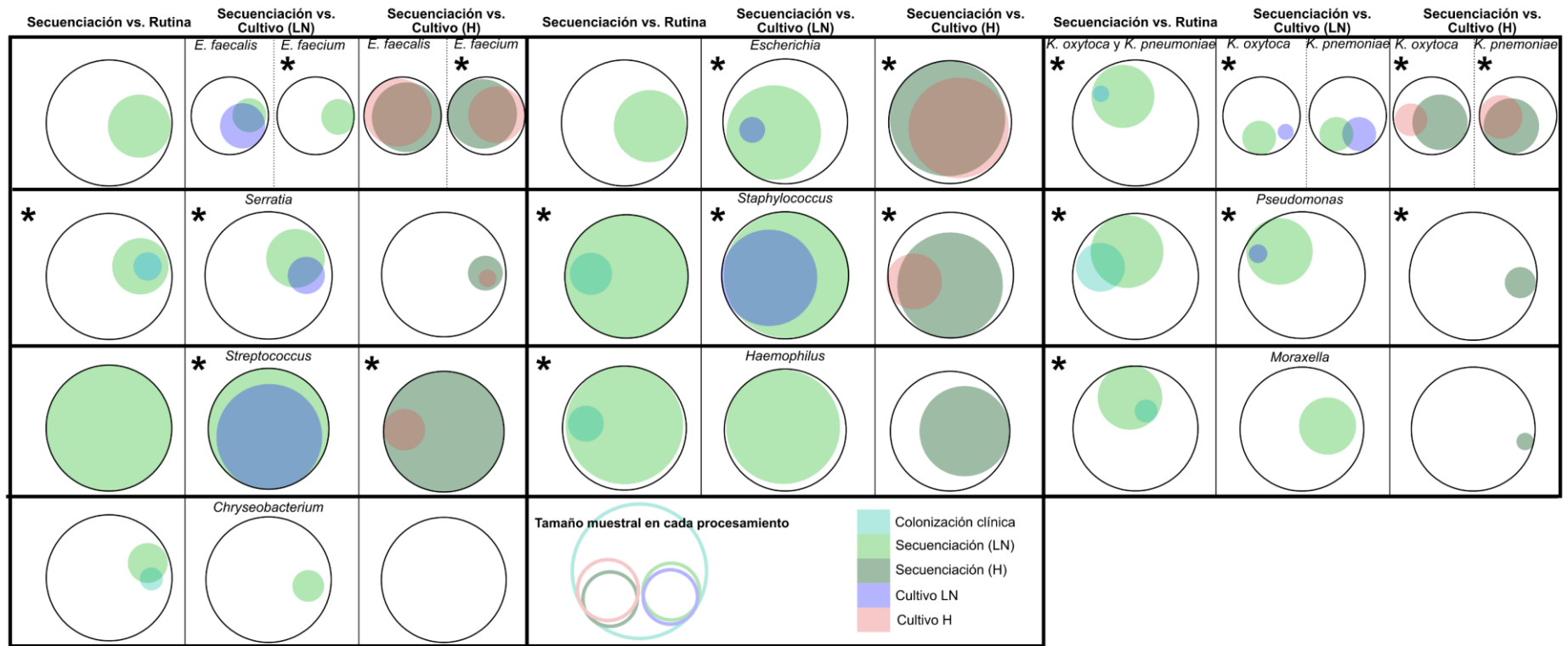


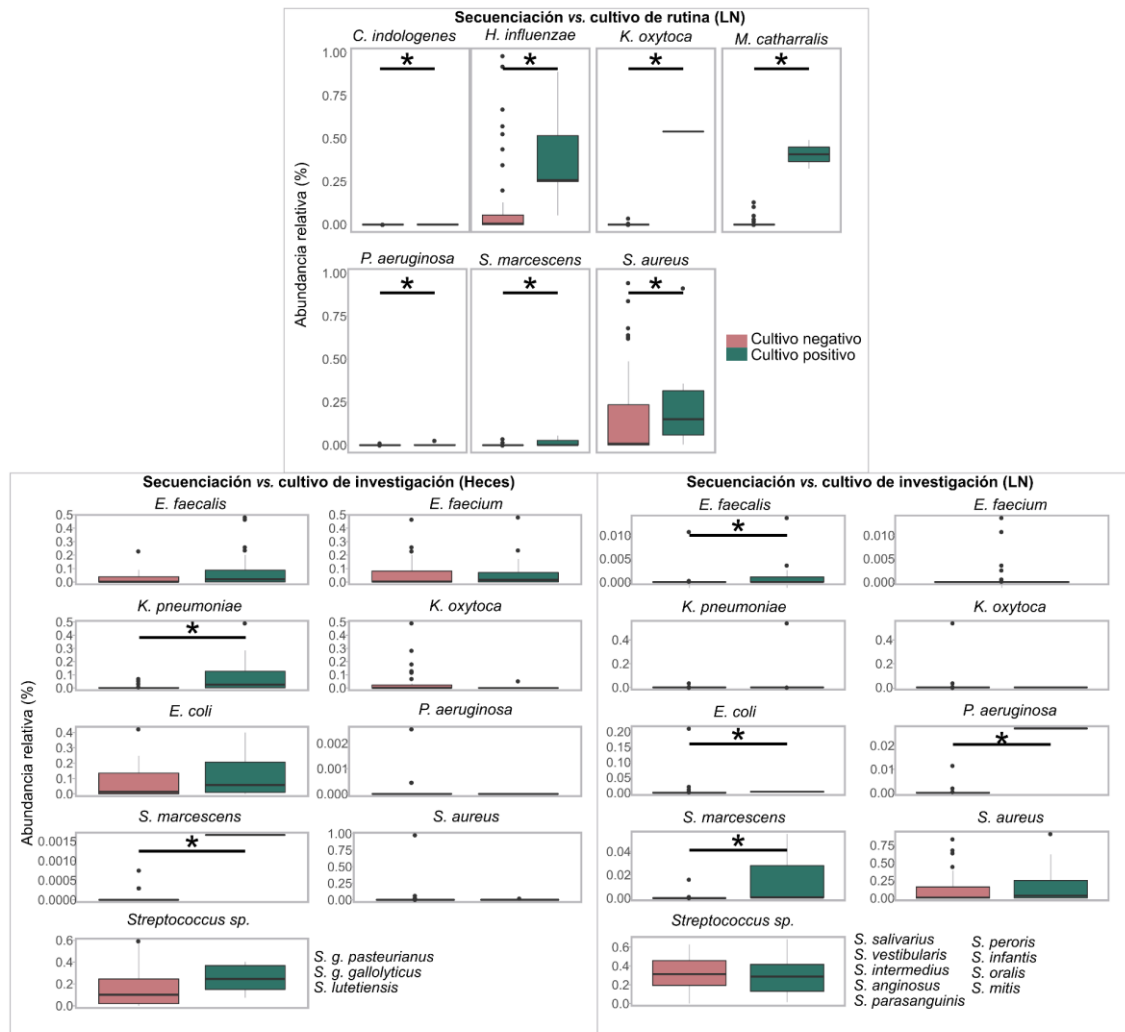
Figura 34. Cambios composicionales significativos a lo largo del tiempo de los principales géneros bacterianos (N=23) en función de distintas variables en las heces y los LN de la cohorte de FQ. AR: genotipo de alto riesgo. BR: genotipo de bajo riesgo. LME: lactancia materna exclusiva. La longitud de la barra de significancia estadística indica que algunos cambios son significativos a nivel general (de 0 a 24 meses), y otros sólo en un momento particular. Eje y: abundancia relativa (%). Eje x: tiempo (meses).

#### 4.2. Concordancia entre los análisis dependientes de cultivo y la secuenciación

En primer lugar, se compararon las prevalencias de los patógenos detectados mediante las técnicas dependientes de cultivo entre sí y con las obtenidas por secuenciación (**Figura 35**). Ambos cultivos de LN (rutina e investigación) detectaban (o no detectaban) en la misma medida a *K. oxytoca* y *S. marcescens* (datos no mostrados). Cuando se comparaba la secuenciación y el cultivo de investigación de LN, sólo *E. faecalis* y *K. pneumoniae* fueron equivalentemente detectadas y, cuando se comparaba secuenciación y el cultivo de heces, coincidían en *E. faecalis* y *S. marcescens*. Por otro lado, la prueba estadística de McNemar mostró que la secuenciación coincidía con el cultivo de rutina en un solo género: *Chryseobacterium*. El análisis bivariante sobre la detección por cultivo de rutina y la abundancia relativa mostró diferencias significativas en la abundancia de todas las especies patógenas, entre las muestras con detección positiva y en las que no se detectó nada de esa bacteria (**Figura 36**). Con el cultivo de investigación, esta asociación fue significativa para *E. faecalis*, *E. coli*, *P. aeruginosa* y *S. marcescens* en LN, y para *K. pneumoniae* y *S. marcescens* en heces.



**Figura 35. Prevalencia en la detección de las bacterias cultivables y patógenas clínicas por secuenciación, cultivo rutinario y cultivo de investigación.** Asterico: prueba de McNemar significativa ( $p < 0,05$ ), lo que quiere decir que las técnicas tienen discordancias a la hora de identificar a la bacteria en el total de muestras (heces o LN) a comparar.



**Figura 36. Diferencias entre la abundancia relativa de las bacterias cultivables en muestras con cultivo positivo y con cultivo negativo.** El cultivo de investigación permitió identificar las distintas especies de *Streptococcus* detectadas en cada nicho. *E. faecium* no fue detectado en cultivo de LN. Asterico: prueba de Mann-Whitney significativa ( $p < 0,05$ ),

## V. DISCUSIÓN

El manejo terapéutico actual de la FQ basado en el diagnóstico perinatal con el cribado ha permitido abordar aspectos como la instauración del microbioma tras el nacimiento y su evolución en los primeros meses de vida. La obtención de estas muestras es siempre compleja por tratarse de recién nacidos. Además, en este caso, hay que destacar algunos aspectos importantes para, en el futuro, poder explorar este tipo de eventos en mayor profundidad y con mayor contundencia estadística. La comunicación con los centros hospitalarios y la recogida de muestras ha sido, en ocasiones, compleja. La pandemia del SARS-Cov2 interrumpió la recogida de lavados nasofaríngeos y retrasó la llegada de muestras al laboratorio. Por otro lado, algunos pacientes fueron desplazados de sus centros de referencia y se perdió su seguimiento. A esto hay que sumarle la reducida cantidad de muestra que podemos obtener de cada niño, lo cual limita la posibilidad de repetir los resultados con varias tecnologías. En este sentido, a pesar de que se recogieron 205 muestras, sólo se pudo trabajar con 148 de ellas, al no disponer de los datos clínicos de todos los niños, porque la secuencia de muestras era incompleta, o la cantidad de la muestra era tan limitada que no se podía usar. También hay que destacar la disparidad en la secuencia de recogida de muestras, ya que coincidía con las visitas rutinarias y esto era variable en función de la evolución de cada paciente. A pesar de las limitaciones metodológicas, el estudio se ha planteado con un robusto análisis bioestadístico incluyendo las variables clínicas y las microbiológicas.

Para clasificar a los recién nacidos con FQ en base al riesgo de su genotipo, nos basamos en diversas publicaciones [248]–[251]. Al igual que en otras cohortes del país, el genotipo de alto riesgo fue el más frecuente en nuestra población. Sin embargo, no se cumplieron las estimaciones en cuanto a la asociación del genotipo de alto riesgo con las variables clínicas, debido, probablemente, a la falta de tamaño muestral. Se observó, no obstante, una primocolonización por *S. aureus* más tardía en el genotipo de bajo riesgo, así como menor número de exacerbaciones respiratorias (sin llegar a la significancia estadística). Dos de los tres niños con genotipo de bajo riesgo presentaron íleo meconial, el cual fue intervenido (#11 y #14), mientras que el otro (#4) sufrió una colonización intermitente de *P. aeuginosa* de 6-18 meses. Esto explicaría el número de ingresos hospitalarios y la elevada toma de antibióticos, pudiendo ser el factor de riesgo que hizo que fueran colonizados precozmente de diversos patógenos (ver Anexos). Ahora que el tiempo de diagnóstico se ha igualado con el cribado neonatal y no hay diferencias entre

los recién nacidos con íleo meconial, que recibían un diagnóstico más temprano, y los que no lo presentan, parece que esta complicación sí está asociada a un retraso del crecimiento y la infección crónica por *P.aeruginosa* [252]. Por otro lado, trabajos anteriores ya advertían de que la asociación entre genotipo y fenotipo en FQ puede ser útil para estudios epidemiológicos, pero no para predecir el pronóstico individual, puesto que el declive en la función pulmonar ocurre, incluso en genotipos con función de CFTR residual [253], [254]. De todas formas, no se duda de su influencia, como han demostrado recientemente con un registro de pacientes de mayor edad [112]. La asociación del genotipo y el perfil del microbioma no se había explorado en profundidad hasta ahora. En este sentido, encontramos en nuestra cohorte que, en los primeros meses, la abundancia de *Ruminococcus* y *Streptococcus* era menor en las heces del genotipo de bajo riesgo, y la abundancia de *Gemella* en LN era muy superior a la del genotipo de alto riesgo, al contrario que el género *Enterococcus*. La composición de *Ruminococcus*, en la cohorte general, aumentó tanto en heces como en LN, confirmando los resultados del Capítulo 1 [255].

Se ha detectado, a través de técnicas de cultivo y secuenciación, paralelismo en la composición y evolución en el tiempo del ecosistema intestinal y el nasofaríngeo, contribuyendo a la hipótesis del eje intestino-pulmón. El niño que padeció una sepsis por *S. marcescens* presentó un enriquecimiento intestinal por esta especie, y se detectó el mismo clon en heces y en LN. El otro recién nacido con infección respiratoria por esta misma especie, también presentó el mismo clon en los dos nichos anatómicos. Todos los indicios apuntan a la conexión microbiológica entre ambos ecosistemas, pues ambos *core* tienen patrones conservados con especies que correlacionan en positivo y en negativo, siendo el núcleo duro *Escherichia-Shigella*, *Veillonella*, *Haemophilus*, *Gemella*, *Staphylococcus*, *Enterococcus*, *Bifidobacterium* y *Streptococcus*, coincidiendo en su mayor parte con otras publicaciones [143]. Los resultados mostraron cómo la diversidad alfa de cada nicho fue aumentando con la edad del niño [255], a pesar de todas las colonizaciones patogénicas, los ingresos hospitalarios y los tratamientos antibióticos que recibieron. Uno de los resultados más llamativos es la alta complejidad del ecosistema respiratorio, con asociaciones positivas entre *Escherichia-Shigella* y *Klebsiella*, *Haemophilus* y *Neisseria*, y *Neisseria* con *Veillonella* y *Prevotella*; mientras que *Veillonella* y *Staphylococcus* parecen tener características excluyentes. Estos resultados están en línea con otros trabajos que ya señalaban la relevancia de las bacterias anaerobias

en la colonización temprana en FQ y en el modelo de ataque-clímax [86], [256].

La composición y abundancia de las bacterias fue diferente según las características de los niños (Figura 34). La lactancia y el tipo de parto son dos de los grandes condicionantes en los primeros momentos de instauración de la microbiota, tanto intestinal como respiratoria [143]. Curiosamente, *Ruminococcus* fue más abundante en las heces de los niños alimentados exclusivamente con lactancia materna. Las exacerbaciones respiratorias por *P. aeruginosa* están inversamente relacionadas con la abundancia de *Dolosigranulum*, bacteria a la que se atribuye un efecto beneficioso en el tracto respiratorio y cuya adición local como microorganismo probiótico ha sido propuesto [257].

Otro de los aspectos relevantes que se analizó fue la transmisión de clones entre pacientes. En el tracto respiratorio, en 6 ocasiones, y en el intestinal, en 15, se han detectado clones idénticos de diversas especies (sobre todo, *E. coli* en heces y *S. aureus* en LN) en pacientes distintos. Este hecho ya se conocía en la FQ [258], y es la base de segregar a los pacientes en consulta en función de su colonización (*P. aeruginosa* sobre todo, pero también *Mycobacterias*), para evitar la contaminación cruzada. Cabe destacar el clon de *S. aureus* ST398, que se detectó en dos niños de centros diferentes, y al que se atribuye un origen porcino. Este clon ya se había descrito previamente en pacientes con FQ [259], pero nunca en niños tan pequeños, y su alarma está justificada por su mayor capacidad de provocar inflamación, y por su resistencia a los antibióticos, que lo hace más difícil de erradicar. Uno de los niños también presentó colonización por *S. aureus* resistente a meticilina (SARM), que también es un hecho destacado en esta patología [260].

Este estudio ha puesto de manifiesto la precocidad de la primocolonización por *P. aeruginosa*, la cual, a pesar del tratamiento erradicador aplicado, es un microorganismo altamente resistente a los antibióticos. Los cultivos realizados desde la investigación de esta tesis permitieron recuperar sólo a una *P. aeruginosa*, donde hubo concordancia total con la secuenciación, puesto que fue la muestra con mayor abundancia relativa de esta bacteria (3%, en LN). No ocurrió lo mismo con otros géneros como se refleja en las Figuras 35 y 36. La secuenciación detecta el ADN de los microorganismos, pero no asegura su viabilidad. Por el momento, no se han establecido los criterios clínicos para estas técnicas, pero sí se han descrito modificaciones del protocolo como la adición de monoazida de propidio (PMA) para evitar amplificar el ADN de bacterias muertas [261].

En *S. aureus*, el otro gran patógeno de la FQ, se observó que no siempre se correlacionaba la abundancia relativa con el cultivo microbiológico, ya que no siempre se recuperaban aislados de muestras con gran proporción relativa de secuenciación, y viceversa. Esto puede deberse, por un lado, a que las condiciones de conservación de la muestra pueden dificultar la viabilidad de las bacterias y su aislamiento en cultivo. Por otro lado, como se comentó en el capítulo anterior de esta tesis, la técnica de secuenciación tiene varios sesgos derivados, principalmente la afinidad de los cebadores y la incapacidad de discriminar ADN de bacterias viables o muertas. Aun así, la secuenciación no llega a la concreción de especie, y lo que se detecta como género *Staphylococcus* puede pertenecer a otras bacterias, como *S. epidermidis*, *S. caprae* o *S. haemolyticus*, que se obtuvieron, pero no se trabajaron. Además, es crucial tener en cuenta que los datos de secuenciación deben tratarse siempre de manera composicional: cada fracción obtenida es referida al total de lecturas que se han obtenido, que serán a su vez sólo una fracción de la composición real del ecosistema de partida.

La diferencia de resultados obtenidos entre los cultivos rutinarios y el de investigación, es que aquí se realizó un enriquecimiento previo de las muestras en caldo BHI, lo cual, por otro lado, puede conllevar una sobrerrepresentación de microorganismos que inicialmente no eran viables o que estaban en bajos recuentos. Sin embargo, y en la mayoría de las veces, la comparación no pudo ser precisa debido a la falta de datos del mismo momento. Asimismo, es más probable obtener muestras de los niños que empeoran o tienen más exacerbaciones que de los que no tienen, y por tanto se tiene un sesgo al disponer solamente de muestras con calidad en estos niños con mayores secreciones.

En resumen, podemos decir que los recién nacidos con FQ experimentan, ya en los primeros dos años de vida, importantes exacerbaciones, ingresos e ingesta de antibióticos, y que son colonizados tempranamente por los patógenos clásicos. Todo ello, impacta en la implementación del microbioma respiratorio e intestinal, que evolucionan de manera muy similar, afectando tanto a la presencia de taxones bacterianos como de su abundancia. Por otro lado, existe un microbioma *core* de los niños con FQ centrado en *Escherichia-Shigella*, *Veillonella*, *Haemophilus*, *Gemella*, *Staphylococcus*, *Enterococcus*, *Bifidobacterium* y *Streptococcus*, con correlaciones positivas y negativas. Se han detectado clones compartidos entre varios pacientes, incluso entre centros, destacando el clon ST398 de *S. aureus* de origen porcino y relevancia epidemiológica.

## **VI. DIVULGACIÓN**

El manuscrito que contiene los resultados de este capítulo se está redactando para mandarlo a publicar.

## **CAPÍTULO 3**

**INFLUENCIA DE UN BROTE DE**  
*Serratia marcescens* **EN EL**  
**ESTABLECIMIENTO DE LA**  
**MICROBIOTA INTESTINAL DE**  
**RECIÉN NACIDOS PREMATUROS Y**  
**ANÁLISIS FILOGENÉTICO DE**  
**AISLADOS NOSOCOMIALES**  
**CONTEMPORÁNEOS E HISTÓRICOS**  
**DE DICHA ESPECIE**

*En la población neonatal prematura ingresada en UCIN...*

*¿Cómo es el establecimiento de su **microbiota intestinal** en el primer mes de vida?*

*¿Cómo les afecta un **contexto epidemiológico** de brote?*

*¿Hay evidencias de **bacterias depredadoras** en este ecosistema?*

*Las bacterias responsables del brote, ¿qué **contenido genético** tienen?*

*¿Cuánto se **diferencian** de las cepas históricas aisladas del mismo lugar con 47 años de diferencia?*

*¿Existen características comunes que expliquen el éxito de la **tasa de transmisión** de cepas de brotes separadas por varias décadas?*

## I. ANTECEDENTES

La relación entre nuestras células eucariotas y las comunidades microbianas de nuestro cuerpo puede clasificarse como mutualista, comensal o patógena. Esta consideración delimita la fina barrera que distingue la colonización de la infección y puede fluctuar a lo largo del tiempo debido a la influencia del huésped o de factores microbianos y ambientales. Se ha confirmado la influencia de la microbiota intestinal en la salud humana global [262]–[264], lo que ha dado lugar a nuevas perspectivas de investigación destinadas a optimizar la composición y la funcionalidad de esta bioed.

El establecimiento adecuado de la microbiota en los recién nacidos es un proceso especialmente relevante [265] y su determinación representa todo un reto científico. La colonización bacteriana humana puede empezar *in utero*, pero el paso crítico comienza con la exposición a las bacterias maternas al nacer y durante el periodo postnatal temprano [266], [267]. Las poblaciones bacterianas fluctúan considerablemente durante los primeros meses de vida, hasta que se establece un ecosistema estable cuando el lactante tiene aproximadamente 2-3 años [70], [268]. Aún no se han establecido los criterios que definen una microbiota intestinal "normal" o "sana", aunque debería incluir una elevada diversidad ecológica, una marcada variabilidad interindividual y una estabilidad intraindividual conservada a lo largo del tiempo. Es importante señalar que la composición de este ecosistema está influenciada por numerosos factores [269], [270], entre ellos la edad gestacional [271], [272].

La prematuridad es la principal causa de morbilidad y mortalidad perinatal, así como un importante factor de riesgo de muerte en los primeros cinco años de vida [94]. En las dos últimas décadas se ha registrado un aumento considerable de las tasas de nacimientos prematuros en todo el mundo, tanto en los países desarrollados como en los que están en vías de desarrollo [94]. En España, la tasa de nacimientos prematuros de todos los nacidos vivos aumentó del 7,1% en 1996 al 8,2% en 2008, lo que supone una de las tasas más altas de Europa [273]. Esta tenencia global puede explicarse por varios factores asociados a la madre, como una edad materna temprana o avanzada, un pequeño intervalo entre embarazos, un índice de masa corporal bajo, un embarazo múltiple, antecedentes de enfermedades infecciosas, estrés, consumo de alcohol y enfermedad periodontal [98]. No obstante, aproximadamente la mitad de los partos prematuros espontáneos no tienen una causa identificada y se ha sugerido que la composición de la microbiota materna podría desempeñar un papel desencadenante [268], [274].

En los recién nacidos prematuros de bajo peso (<2.500 g), la microbiota es aberrante y el establecimiento bacteriano se retrasa por su intensa exposición a los antimicrobianos [272], [275]–[277]. Este hecho explica, al menos en parte, por qué los recién nacidos prematuros tienen un sistema inmunitario inmaduro y suelen sufrir complicaciones infecciosas [278]–[280].

*S. marcescens* es uno de los patógenos emergentes más relevantes que causan brotes graves en esta población [281]. Los primeros informes científicos de *S. marcescens* en infecciones humanas se produjeron en la segunda mitad del siglo XX. En la actualidad, las principales preocupaciones en relación con este patógeno oportunista son su capacidad de propagación por el entorno hospitalario, su capacidad para causar brotes y su potencial para expresar y diseminar la resistencia a los antibióticos, combinando mecanismos intrínsecos y genes adquiridos [282]. Se ha informado con frecuencia de la colonización intestinal patógena durante los brotes nosocomiales; sin embargo, hasta donde sabemos, no se ha estudiado la influencia de un brote en el proceso de establecimiento de la microbiota.

El primer brote nosocomial documentado causado por *S. marcescens* en España se produjo en 1969 el Hospital terciario Infantil La Paz, en la UCIN. En base a las técnicas de fenotipado disponibles en ese momento, que incluían reacciones bioquímicas y pruebas de susceptibilidad a antibióticos, se consideró un brote monoclonal. Este brote se prolongó hasta finales de 1974 (más de tres años), con una alta incidencia de bacteriemia por *S. marcescens* (9,16% de todas las sepsis neonatales tempranas del periodo, frecuentemente asociada a la colonización intestinal durante los tres primeros días de vida (12,5% de 120 niños analizados) [283].

Las nuevas herramientas de secuenciación genética permiten ampliar el conocimiento sobre este tipo de episodios nosocomiales, aportando precisión a los conceptos de clonalidad, portabilidad de resistencias a antimicrobianos y facilidad de propagación. Actualmente, con la secuenciación de genoma completo se puede calcular el resistoma, determinar el genoma completo y el accesorio, realizar árboles filogenéticos con los amplios repositorios disponibles o buscar secuencias de fagos o secuencias de inserción.

## II. JUSTIFICACIÓN Y OBJETIVOS

Los recién nacidos prematuros ingresados en una UCIN están expuestos a un entorno hospitalario donde abundan bacterias resistentes a antibióticos, lo que puede impactar hasta cierto grado a las comunidades bacterianas que adquieren en esta ventana crítica del desarrollo y que luego conformarán su microbiota adulta. Los brotes nosocomiales son frecuentes en estas unidades, siendo *S. marcescens* uno de los patógenos más habituales. Los avances moleculares y los conocimientos bioinformáticos permiten realizar estudios epidemiológicos para caracterizar los brotes, comparar aislados que comparten el mismo ambiente en momentos diferentes y evaluar su impacto en el ecosistema intestinal de pacientes especialmente vulnerables, como los neonatos prematuros.

Los **OBJETIVOS** de este capítulo fueron:

1. Analizar secuencialmente el establecimiento de la microbiota intestinal de neonatos prematuros de bajo peso ingresados en la UCIN del Hospital La Paz durante sus primeras tres semanas de vida, comparando dos escenarios epidemiológicos: un periodo normal y otro con un brote nosocomial de *S. marcescens*.
2. Identificar los clones bacterianos de las principales bacterias cultivables de la microbiota meconial y fecal de neonatos prematuros de bajo peso durante sus tres primeras semanas de vida.
3. Examinar los representantes clonales de los aislados de *S. marcescens* causantes de este brote contemporáneo y comparar su contenido genético con el de los primeros aislados históricos recuperados en la misma UCIN 47 años antes.
4. Actualizar la filogenia de la especie *S. marcescens* empleando el repositorio público de genomas, incluyendo los aislados de ambos brotes.

### III. MATERIALES Y MÉTODOS

#### 1. Diseño metodológico

Para este capítulo, el diseño se dividió en tres apartados:

En primer lugar y para llevar a cabo el objetivo 1, se propuso un análisis transversal, en el que se comparó la composición de la microbiota intestinal de una población de recién nacidos en un contexto de brote epidemiológico con una población de neonatos en un contexto epidemiológico normal. Para ello, se seleccionaron como variables de estudio la abundancia de las bacterias mayoritarias del meconio y de las tres semanas de vida, así como los índices de alfa y beta diversidad, obtenidos a partir de la secuenciación de amplicones.

En segundo lugar y para llevar a cabo el objetivo 2, se realizó un estudio exploratorio en el que se tipificaron los aislados de las principales bacterias cultivables obtenidas del meconio y de las heces de las primeras semanas de vida de los neonatos ingresados durante el brote nosocomial.

En tercer y último lugar, para llevar a cabo los objetivos 3 y 4, se realizó un análisis exhaustivo de los genomas completos de los representantes clonales de *S. marcescens* responsables de este brote nosocomial con los representantes de un brote de características similares ocurrido 47 años antes en la misma UCIN.

#### 2. Participantes, recogida de muestras e inclusión de aislados

El Hospital Universitario La Paz de Madrid cuenta con una UCIN de nivel III de 23 camas, de la que se reclutaron 19 neonatos prematuros de bajo peso al nacer (<32 semanas de gestación) en dos periodos distintos: **A)** durante un período epidemiológicamente normal en 2015, de ahora en adelante grupo control, (n=7) y **B)** durante un brote de *S. marcescens* entre diciembre de 2016 y marzo de 2017, de ahora en adelante grupo brote (n=12). Es importante señalar que en el grupo control no pudimos disponer de datos referentes al consumo de antibióticos, así como otros datos clínicos.

De cada recién nacido prematuro se recogieron cuatro muestras fecales del pañal tras la fecha de nacimiento: el meconio y heces de los días 7, 14 y 21 de vida. Aunque el objetivo era ampliar el periodo de recogida, las limitaciones logísticas lo redujeron a las tres primeras semanas de vida. Las muestras se extrajeron del pañal empleando un hisopo estéril y se almacenaron inmediatamente a -80°C. A pesar de que la intención era recoger

las muestras inmediatamente después de la deposición, no podemos descartar una posible contaminación con la orina, a la que consideramos insignificante respecto del cómputo total de la microbiota intestinal. En los controles, sólo se disponía de ADN extraído de las muestras fecales del grupo control, cedido por la Dra. Laura Moles. Por el contrario, en el grupo brote se disponía de materia biológica y se realizó el cultivo microbiológico además de la extracción del ADN. El Comité Ético de Investigación Clínica del Hospital Universitario La Paz aprobó el estudio (referencia HULP3551) y los datos clínicos de los neonatos se obtuvieron de su historial clínico.

Para la tercera parte de este capítulo, se trabajó con aislados históricos de *S. marcescens* (n=21) causantes de un brote nosocomial ocurrido en el mismo hospital en 1970, que se conservaban a temperatura ambiente en frascos de agar sólido BHI cerrados herméticamente en el hospital Ramón y Cajal. Estos aislados se obtuvieron a partir de muestras fecales de pacientes ingresados en la UCIN el Hospital La Paz entre el 4 y el 21 de noviembre de 1970. El objetivo era compararlos con los obtenidos 47 años después a raíz del brote contemporáneo de 2016-2017.

### 3. Extracción de ADN

La extracción de ADN total de las muestras fecales se realizó con el kit QiaAMP (Qiagen, Alemania). Las muestras se descongelaron lentamente, primero 24 h a -20 °C y luego otras 24 h a 4 °C. Se suspendieron en 500 µL de agua MiliQ y se centrifugaron 5 min a 1.000 rpm para tomar 200 µL de sobrenadante y transferirlo a tubos *ependorf*, donde se procedió a extraer el ADN siguiendo las recomendaciones del fabricante. El ADN bacteriano se conservó en tubos *ependorf* a 4 °C hasta el momento de su utilización. Se determinó la concentración y la calidad de la extracción de ADN en Qubit (Thermo Fisher Scientific, MA, US). En el caso de extraer ADN de cultivo, se procedió a extraer el ADN de la misma forma, pero utilizando el inóculo directamente de un 1 mL de medio de cultivo líquido de 24 h de la bacteria.

### 4. Secuenciación de amplicones del 16S ARNr

Las muestras de ADN fecal se enviaron a FISABIO (Valencia, España) para la secuenciación y procesamiento posterior (análisis primario, secundario y terciario). Se amplificaron y secuenciaron las regiones V3 y V4 del 16S ARNr con la tecnología Mi-Seq 2 × 300 pb de extremos pareados de Illumina (Cod. 15044223 Rev. A). Los cebadores empleados fueron los siguientes [201]:

- Cebador directo:

5'-TCGTCGGCAGCGTCAGATGTGTAAAGAGACCCTACGGNGGCWGCAG.

- Cebador inverso:

5'-GTCTCGTGGCTCGAGATGTGTAAAGACGACTACHVGGGTATCTAATCC.

#### **4.1. Análisis bioinformático**

Los archivos de las secuencias con su calidad (fastq) están depositados en el repositorio SRA del NCBI bajo la referencia *BioProject PRJNA510235*. El análisis estadístico terciario se realizó utilizando el software estadístico Past 3.16 y herramientas en línea como la plataforma Galaxy Huttenhower. La significación estadística se fijó en  $p < 0,05$ .

#### **4.2. Procesamiento de las lecturas**

La calidad de las secuencias se midió según los siguientes parámetros: longitud mínima: 250 pb; tipo de medida de calidad de recorte: media; número de calidad de recorte desde el extremo 3': 30; ventana de calidad de recorte: 10 pb.

#### **4.3. Asignación taxonómica**

La asignación taxonómica se realizó utilizando el clasificador del *Ribosomal Database Project* (RDP) y las lecturas con una puntuación inferior a 0,8 se asignaron al rango taxonómico superior, dejando el último rango como no identificado.

#### **4.4. Cálculos de diversidad y pruebas estadísticas**

Para la estimación de la diversidad bacteriana alfa, se utilizaron los índices Shannon-Weaver y Chao1, que se calcularon eliminando los ASV con menos de tres lecturas y posteriormente a la transformación de la matriz de datos (porcentaje relativo de las lecturas totales de cada muestra). Se empleó la prueba Kruskal-Wallis para evaluar las diferencias en las medianas de los índices de diversidad y otras variables cuantitativas entre grupos de muestras después de asumir la distribución no normal de los datos con la prueba de Shapiro-Wilk. La prueba de Dunn fue empleada cuando se compararon más de dos grupos de muestras, en función de cada variable de estudio.

Para la estimación de la diversidad beta, se calcularon los índices de similitud Bray Curtis y se representaron en un eje de coordenadas mediante Análisis de Componentes Principales (PCA). Asimismo, se realizó la prueba PERMANOVA para evaluar las diferencias estadísticas entre grupos de muestras en función de la diversidad beta.

#### 4.5. Herramientas para detectar biomarcadores característicos

Para el análisis estadístico, se descartaron las muestras con menos de 1.000 lecturas, y se excluyeron todas las muestras de aquellos pacientes de los que faltaba algún dato demográfico.

Los neonatos se clasificaron en función de cuatro variables:

- 1) situación epidemiológica (control vs. brote de *S. marcescens*);
- 2) modalidad de parto (vaginal o cesárea);
- 3) edad gestacional (extremadamente prematuros: <28 semanas; muy prematuro: 28-30 semanas; o moderadamente prematuro: 30-32 semanas; y
- 4) peso al nacer (<1.000 g; 1.000-1.500 g; o >1.500 g).

Las diferencias estadísticas se evaluaron con las pruebas de Chi cuadrado para las variables demográficas y perinatales cualitativas, y la t de *Student* para las variables cuantitativas.

Se utilizó la plataforma Galaxy Huttenhower para realizar un análisis lineal discriminante del tamaño del efecto (LEfSe) y obtener cladogramas en los que se representaron los taxones microbianos que explicaban las diferencias significativas entre grupos de muestras [56].

#### 5. Cultivo de muestras, tipificación mediante electroforesis en gel de campo pulsado (PFGE) y estadística de la culturómica

Se inocularon porciones de entre 0,3-0,5 g de cada muestra de heces en caldo *brain heart infusion* (BHI) (Difco, Detroit, Michigan) y se incubaron a 37 °C durante 24 h a modo de pre-enriquecimiento bacteriano. Las bacterias cultivables se aislaron en medios de agar selectivos y no selectivos del tubo BHI (sembrando distintas diluciones), incluyendo placas de agar de *M-Enterococcus*, De Man, Rogosa y Sharpe (MRS); Manitol-Salt; McConkey y agar Columbia con un 5% de sangre de oveja. Los medios de cultivo se compraron a Difco y las placas se incubaron a 37 °C durante 24-48 h, incluyendo 5% de CO<sub>2</sub> para las placas de agar sangre y condiciones de anaerobiosis para las placas MRS.

Para recuperar los aislados históricos de *S. marcescens* guardados en frascos, el agar se rehidrató con BHI líquido durante 24 h a 37 °C, y después se sembraron en placas de agar sangre Columbia 24 h a 37 °C.

La identificación de las colonias se llevó a cabo mediante espectrometría de masas MALDI-TOF (Bruker, Alemania) y todos los aislados se conservaron a -80 °C en leche semidesnatada esterilizada.

Los aislados se tiparon genéticamente mediante PFGE utilizando los protocolos estandarizados y las enzimas de restricción de acuerdo con la especie bacteriana (*SmaI* para especies de *Staphylococcus* y *Enterococcus*, *XbaI* para *E. coli* y, finalmente, *SpeI* para las especies de *Klebsiella* y *S. marcescens*). El análisis del patrón de bandas se realizó con el programa Phoretix 5.0 (TotalLab, NewCastle, Reino Unido) y la representación de los resultados se hizo a partir de los coeficientes Dice.

Se calculó la frecuencia de cada especie y de cada clon ( $\alpha_h$ ) con la siguiente fórmula, asumiendo que la probabilidad de encontrar cada especie o clon en una muestra fecal es siempre la misma (1), independientemente del periodo de tiempo que represente, ya sea meconio, 7, 14 o 21 días desde el nacimiento.

$$\left( \sum_{k=1}^P \sum_{i=1}^{n_k} \delta_{ik}^h * \frac{1}{n_k} \right) * \frac{1}{P} = \alpha_h$$

Siendo:

- $\alpha_h$ : probabilidad de la presencia del clon o especie  $h$  en  $\rho$ . De 0 a 1
- $P$ : total de neonatos del grupo brote
- $n_k$ : número de muestras (meconio y heces) de neonato  $k$
- $\delta_{ik}^h$ : 1 si el clon o especie  $h$  está presente en muestra  $i$  de neonato  $k$ ; 0 si ausente

## 6. Secuenciación de genoma completo (WGS), ensamblaje y anotación

Se secuenció el genoma completo de cada clon causante de los dos brotes nosocomiales (contemporáneo e histórico). Los genomas ensamblados se depositaron en la base de datos de GenBank y se asociaron al Bioproyecto PRJNA510235 (<https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA510235>).

Las muestras de ADN obtenidas a partir de colonias se enviaron a la Universidad Complutense de Madrid para la preparación de librerías, utilizando el kit Nextera XT de Illumina (Illumina, Inc, U.S.) y la tecnología Illumina MiSeq (2x150 pb). Para el análisis de las secuencias, se utilizó el protocolo bioinformático Nullarbor (versión v.2; disponible en <https://github.com/tseemann/nullarbor>) [284]. Esta herramienta incluye el programa

Trimmomatic v0.38 para el recorte y filtrado de calidad de las lecturas [285], Kraken v1.0 [286] y Centriuge v1.0.4 [287] para la identificación de especies, SPAdes v3.12.0 para el ensamblaje *de novo* [288], Prokka v1.13 y Abricate v0.8. para la anotación [289], la base de datos Resfinder para la identificación del resistoma y los programas Snippy v4.4.1 y Roary v3.12.0 para la identificación del genoma central y el pangenoma [290]. La herramienta en línea Phandango se empleó para la visualización y el trazado del pangenoma [291] y PLACNETw [292] para la detección y reconstrucción gráfica de secuencias de plásmidos putativos que posteriormente fueron confirmados por BLAST. La herramienta en línea Phaster fue utilizada para anotar profagos y las secuencias de inserción (IS) se anotaron utilizando el programa ISEScan [293].

### **7. Análisis filogenético de cepas de estudio y del repositorio global de *S. marcescens* mediante cgMLST y AcCNET**

Se construyó un MLST de genoma central (cgMLST) para la comparación filogenética de las cepas contemporáneas, históricas y los 444 genomas de *S. marcescens* disponibles en ese momento en la base de datos RefSeq del NCBI. Se emplearon las herramientas Kraken y los filtros disponibles de la página web para incluir únicamente los genomas asignados de manera precisa a *S. marcescens* y excluir los genomas parciales y anómalos por debajo de 500 pb. En primer lugar, se estableció una base de datos de alelos del pangenoma utilizando todos los genomas seleccionados (n=452). A continuación, se construyó un árbol MLST a partir de la base de datos, seleccionando las variantes alélicas presentes en el genoma compartido, fijado en >95% de ocurrencia. Todos los análisis se realizaron con la herramienta en línea wgMLST\_BacCompare [294]. Los árboles se editaron en línea con la herramienta iTOL v4.4.2 [295].

Por otro lado, se analizó el genoma accesorio utilizando el programa AcCNET v1.2. Se aplicó el método de agrupación jerárquica de varianza mínima de Ward para agrupar los genomas según su proteoma anotado predicho en diferentes clústeres [296]. Solo las proteínas significativas asignadas de forma robusta a cada clúster (>70% por frecuencia de clúster, <70% por frecuencia total y valor ajustado  $p < 0,001$ ) se tuvieron en cuenta para las comparaciones posteriores. Se utilizó la base de datos COG para asignar categorías funcionales al contenido del proteoma predicho. Se utilizó el software RStudio v.1.2.5001 y Gephi v.0.9.2 para el análisis estadístico y la representación gráfica.

### **8. Análisis de susceptibilidad antibiótica**

Se utilizó el sistema automatizado MicroScan WalkAway (Beckman Coulter, U.S.) para

determinar los valores de concentración mínima inhibitoria (CMI). Se utilizaron los criterios EUCAST para definir los aislados susceptibles, intermedios y resistentes a cada antimicrobiano ([www.eucast.org](http://www.eucast.org)).

#### IV. RESULTADOS

##### 1. Características de los neonatos prematuros

Las características demográficas y clínicas de ambos grupos de recién nacidos se muestran en la **Tabla 11**. Las diferencias más relevantes fueron que los neonatos del grupo control presentaron un mayor peso al nacer, más pacientes de sexo masculino, mayor número de partos eutócicos y una menor incidencia de sepsis (sólo un neonato, el 14,3% de esta cohorte, sufrió sepsis precoz). Por otro lado, cinco neonatos del grupo brote, el 41,7% de esta cohorte, sufrieron sepsis de inicio tardío, la cual fue diagnosticada microbiológica o clínicamente. *S. epidermidis* y *S. marcescens* fueron los microorganismos implicados. Sólo se disponía de datos sobre la terapia antimicrobiana del grupo brote. En esta cohorte, se registró el empleo de cefazolina profiláctica durante el parto en 10 neonatos. Además, ocho lactantes recibieron un tratamiento empírico a base de ampicilina, gentamicina y claritromicina durante la primera semana de vida y durante las semanas subsiguientes se registró el empleo de vancomicina (en siete lactantes), cefotaxima (en dos lactantes) piperacilina/tazobactam (en dos lactantes), meropenem (en un lactante) y amikacina (en tres lactantes) (datos no mostrados).

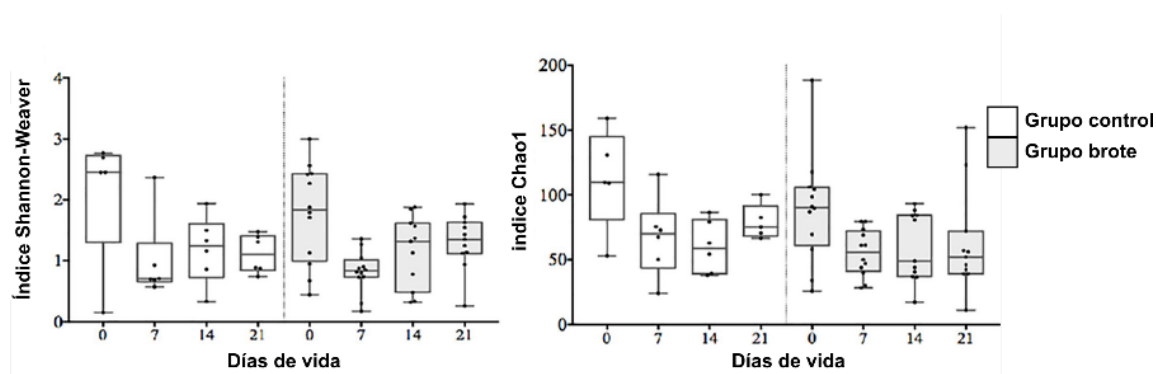
**Tabla 11. Características clínicas y demográficas de las cohortes de neonatos del Hospital La Paz.**

Característica <sup>1</sup>	Grupo control (n=7)	Grupo brote (n=12)	p valor
Peso al nacer (g)	1462 (720-1890) <sup>1</sup>	971 (609-1537)	0,009
Edad gestacional (semanas)	30 (25-31)	28 (25-31)	0,26
Parto vaginal (n, %)	5; 71,4%	3; 25%	0,04
Sexo masculino (n, %)	5; 71,4%	2; 16,6%	0,0001
Sepsis (n, %)	1; 14,3%	6; 50%	0,04
Duración de la estancia (días)	14 (5-140)	55 (7-89)	0,1

<sup>1</sup> Los valores se expresan como mediana (RQ).

## 2. Establecimiento de la microbiota intestinal mediante secuenciación de amplicones

Los índices de diversidad alfa de las muestras de meconio fueron similares a los de las muestras fecales posteriores ( $p > 0,05$ ), indicando un retraso en el establecimiento de la microbiota. Cabe resaltar la gran diversidad bacteriana de las muestras de meconio (**Figura 37**).



**Figura 37. Comparación de la diversidad alfa entre el grupo control y brote.**

Se detectaron diferencias significativas en la composición de la microbiota intestinal de los grupos control y brote en el meconio y en las muestras del día 21 (prueba de PERMANOVA  $p = 0,0024$  a nivel de género y  $p = 0,073$  a nivel de filo). Según el análisis LefSE, las diferencias eran más notables a los 21 días de vida, puesto que existía un mayor número de grupos bacterianos explicando las diferencias entre un grupo y otro. El grupo brote se caracterizó por una mayor proporción de *Proteobacteria*, y, en particular, del género *Serratia*, y por presentar menores índices de *Firmicutes* y *Fusobacteria* (**Figura 38**).

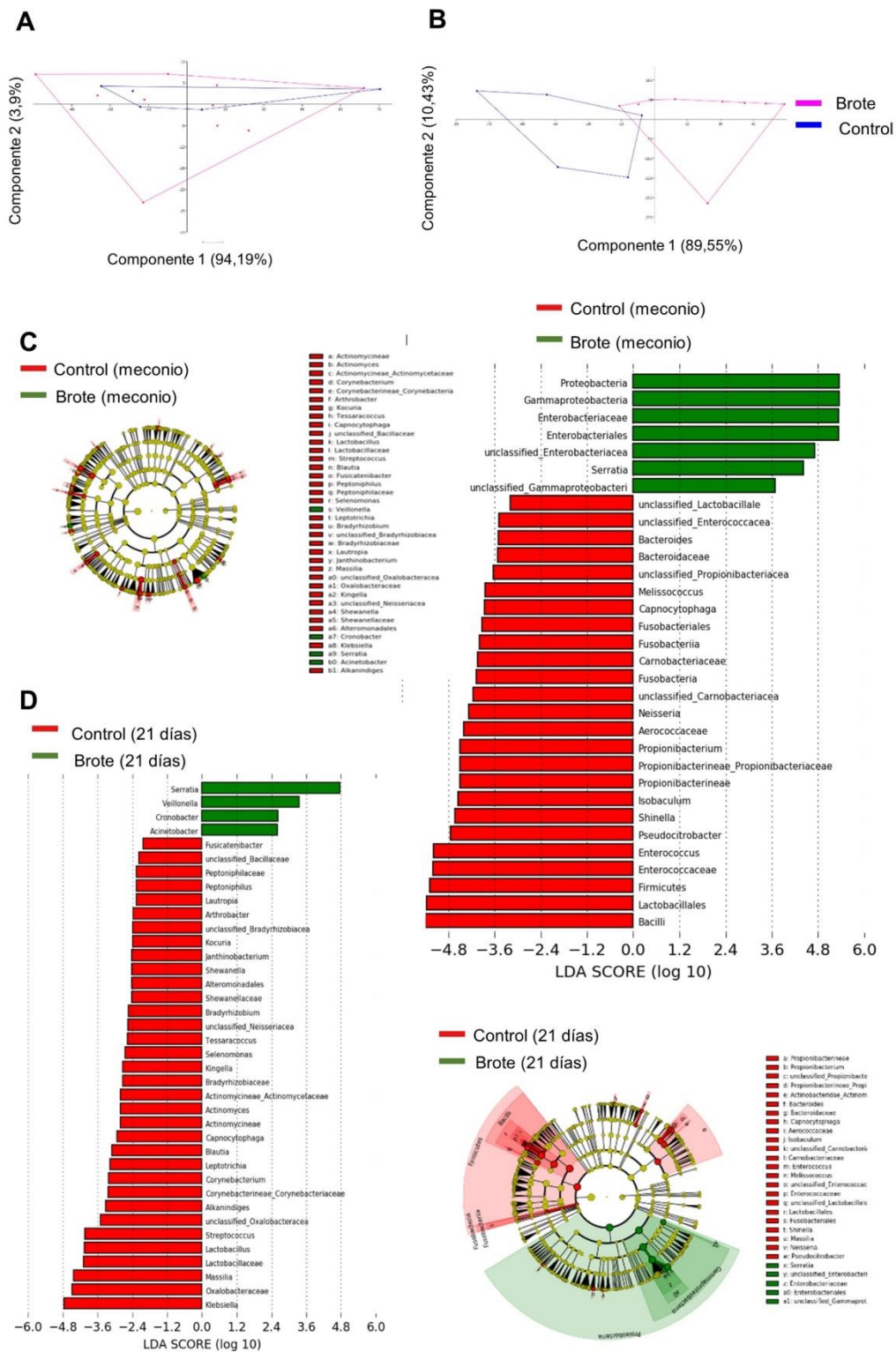
En relación con el modo de parto, las diferencias entre el parto vaginal y la cesárea (juntando ambas cohortes) sólo fueron significativamente diferentes en las muestras de meconio (prueba de PERMANOVA  $p=0,0066$  a nivel de género y  $p=0,0363$  a nivel de filo) (**Figura 39**). No se detectaron diferencias significativas entre las comunidades bacterianas en relación con el sexo, la edad gestacional y el peso al nacer.

Teniendo en cuenta las 55 muestras, el filo predominante fue *Proteobacteria* (mediana (RQ) de 70,1(52,9-84,2)), seguido de *Firmicutes* (mediana (RQ) de 22,1(12,1-41,7)) y, aunque se detectaron hasta otros 23 filios, su contribución fue casi anecdótica

(Figura 40A). A nivel de género, la abundancia de *Escherichia/Shigella* aumentó durante el periodo estudiado, mientras que la de los géneros *Enterococcus* y *Staphylococcus* disminuyó (Figura 40B).

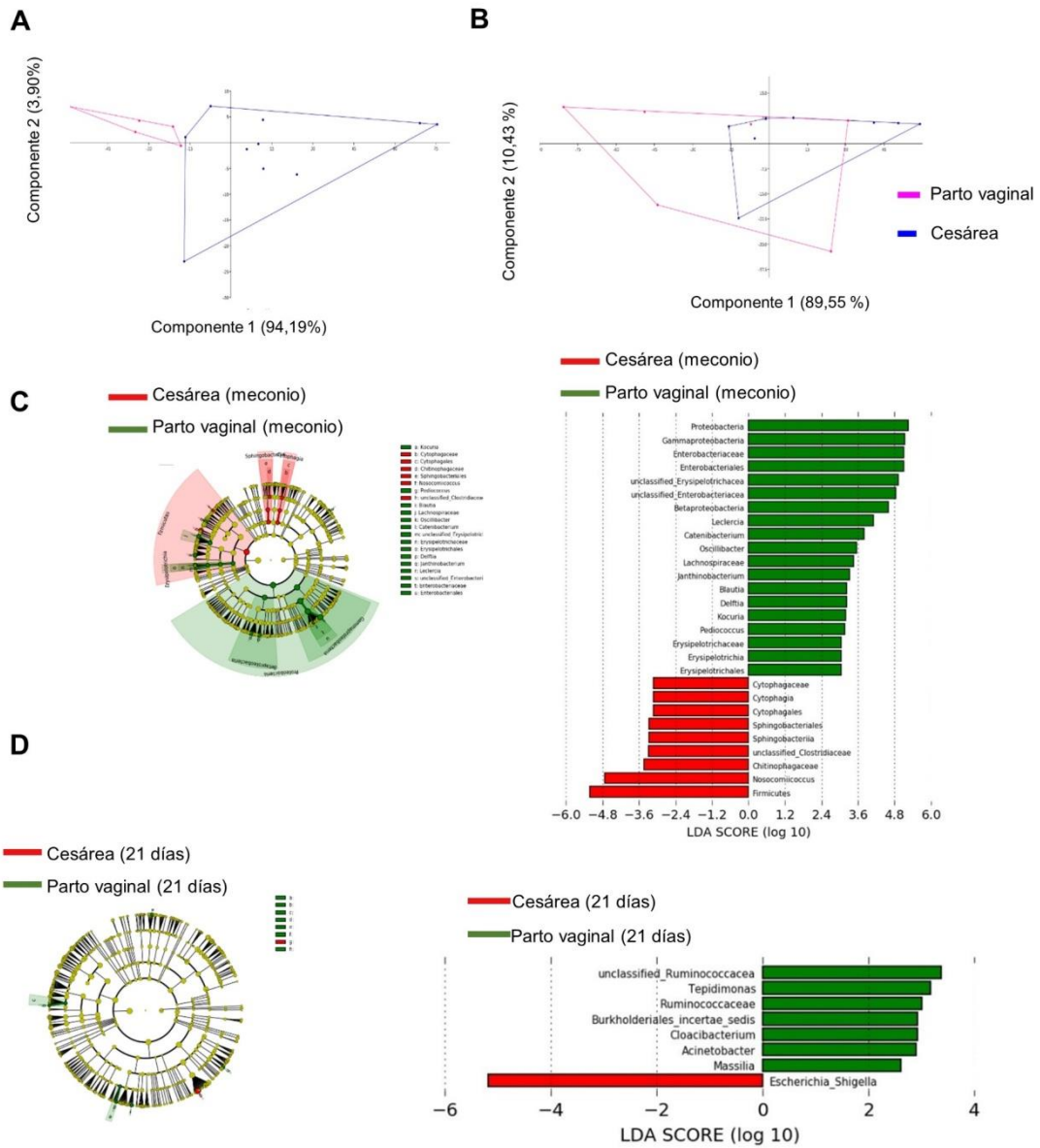
En algunas muestras de meconio del grupo brote (n=3) y del control (n=1) se encontraron lecturas asignadas a los géneros de bacterias depredadoras *Bdellovibrio* (2 neonatos), *Peredibacter* (1) y *Vampirovibrio* (1). La abundancia relativa de especies depredadoras fue extremadamente baja (rango 0,004%-0,11%), y no se pudo detectar ninguna en las muestras fecales posteriores.

Se detectó la presencia de *Serratia* en las muestras de meconio de todos los pacientes del grupo brote (mediana (RQ) de 2,21(0,13-11,9)), mientras que este género era considerablemente menos abundante entre las muestras de meconio del grupo control (mediana (RQ) de 0,005 (0,0003-0,04)). También cabe destacar que las lecturas de *Serratia* aumentaron considerablemente justo antes del diagnóstico de sepsis por este mismo microorganismo, en particular, en el lactante O11 que finalmente murió de sepsis por *S. marcescens* al décimo día de nacer. El enriquecimiento intestinal de las secuencias de *Serratia* en este lactante era destacable, alcanzando niveles del 95% de la microbiota intestinal total en el día 7 (Tabla 12).

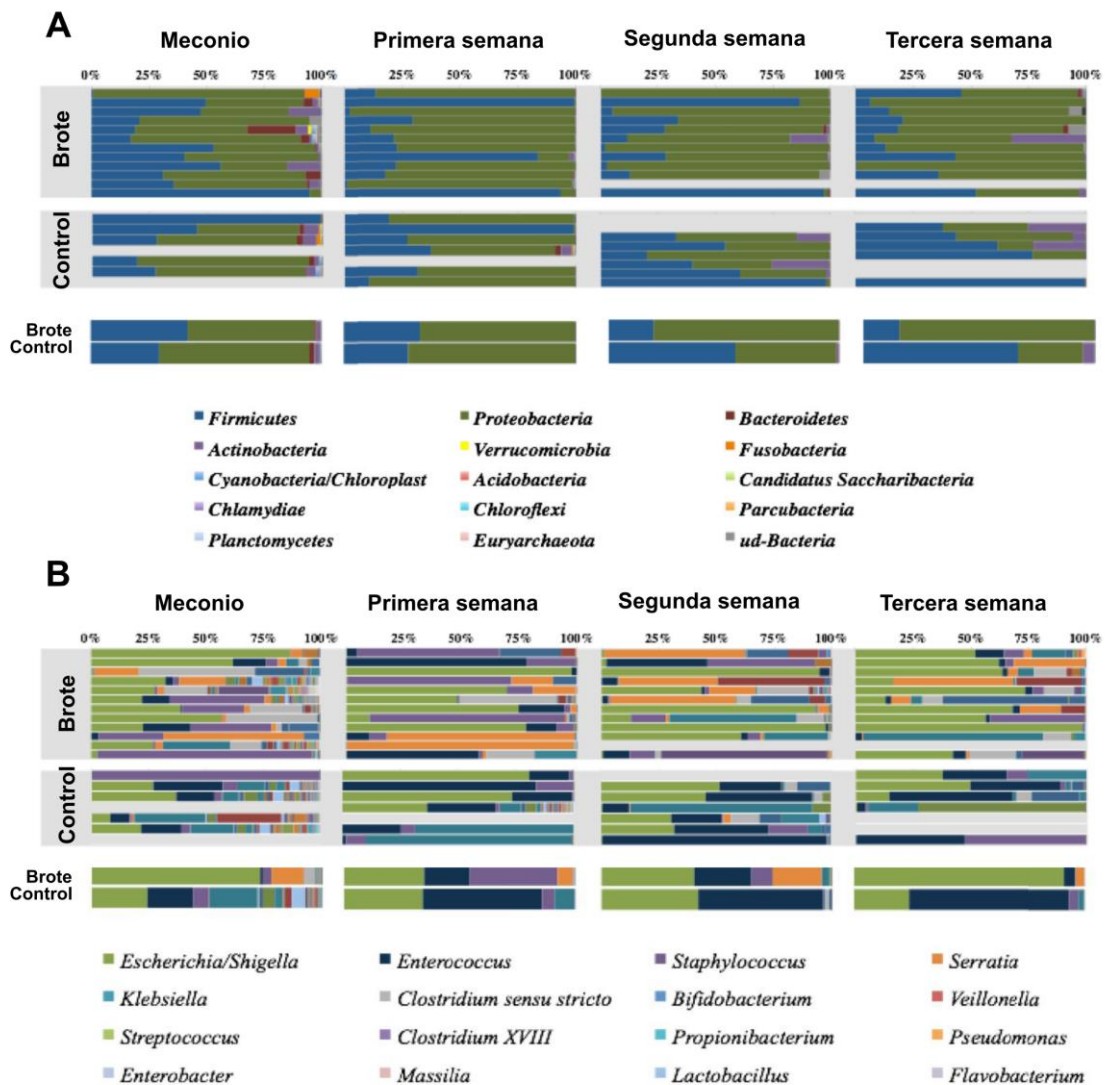


**Figura 38. Diferencias entre el grupo brote y control del hospital La Paz. A:** PCA a nivel de filo de las muestras de meconio. **B:** PCA a nivel de filo de las muestras de 21 días. **C:** cladograma de las muestras de meconio. **D:** cladograma de las muestras del día 21.

CAPÍTULO III.  
RESULTADOS



**Figura 39. Diferencias entre los nacidos por cesárea y por parto vaginal. A:** PCA a nivel de filo de las muestras de meconio. **B:** PCA a nivel de filo de las muestras de 21 días. **C:** cladograma de las muestras de meconio. **D:** cladograma de las muestras del día 21.



**Figura 40. Principales grupos bacterianos y sus abundancias relativas en todos los individuos y en el grupo control y brote del Hospital La Paz (mediana). A:** filos mayoritarios. **B:** géneros mayoritarios.

**Tabla 12. Abundancia relativa del género *Serratia* y aislados de la especie *S. marcescens* en el grupo brote. %: datos de secuenciación de amplicones.**

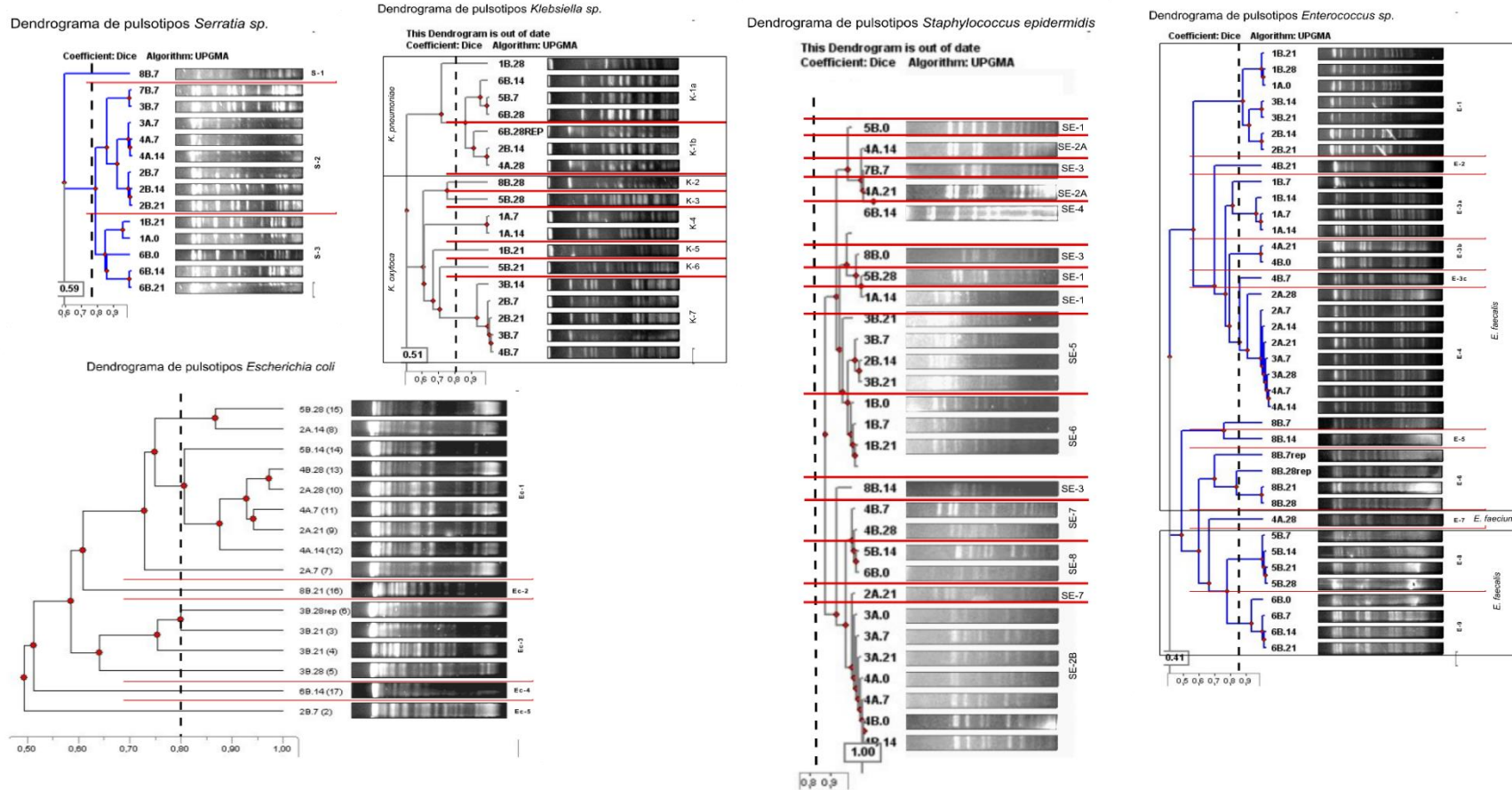
	Neonato	Meconio		Día 7		Día 14		Día 21	
		Cultivable	%	Cultivable	%	Cultivable	%	Cultivable	%
GRUPO BROTE	O1	<b>Clon A</b>	5,4		0,002		12,1		1,0
	O2		12,4		0,003		0,2		1,7
	O3		2,5	Clon B	17,3		19,0		
	O4		2,4	Clon B	2,4	Clon B	4,8		16,9
	O5		1,8		0,003		0,04	Clon A	28,8
	O6		11,8	Clon B	5,1	Clon B	26,5	Clon B	37,4
	O7		2,9	<b>Clon B</b>	0,2		35,5		4,9
	O8		1,9		0,7		1,3		0,02
	O9		2,0		0,004		0,004		0,006
	O10	<b>Clon A</b>	59,5	Clon A	78,5	Clon A	0,8	Clon A	0,4
	O11†		3,1	<b>Clon B</b>	94,5				
	O12		0,3	Clon A	1,3		0,1		1,2
GRUPO CONTROL	C1		0,004		0,005				0,008
	C2		0,006		0,001		0,001		0,006
	C3		0,1		0,0007		0		0
	C4				0,06		0,2		0,01
	C5		0,8				3,1		
	C6		0,06		0		0		
	C7				0,0008		0		0,002

### 3. Caracterización de los aislados cultivables del grupo del brote

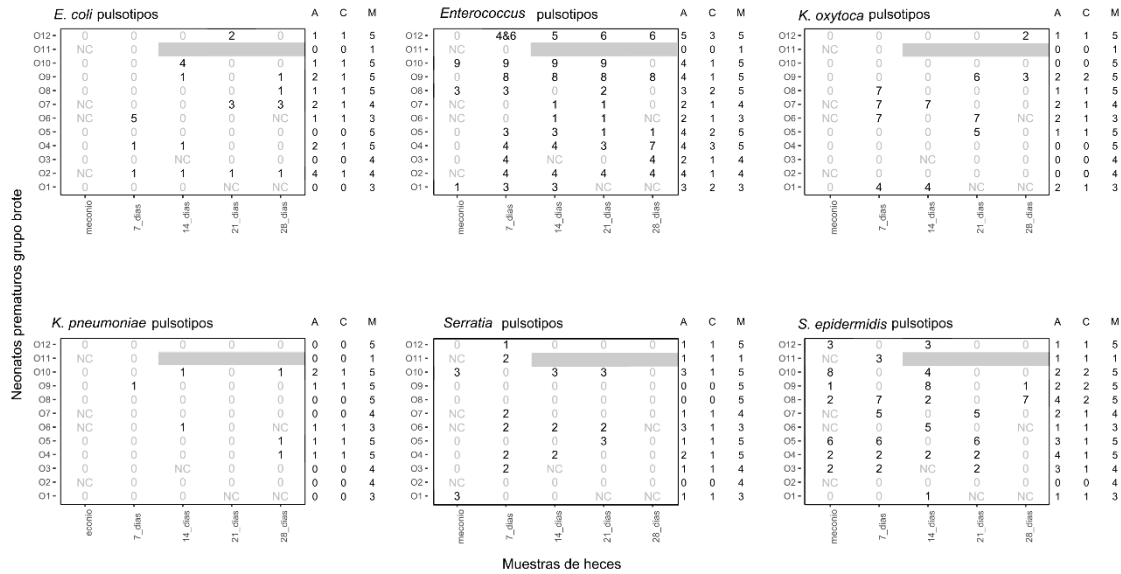
Se recuperaron bacterias de las especies de interés de 49 muestras sembradas en medios selectivos (86%). Sólo en cuatro muestras de meconio y en cuatro de heces (33% y 9% del total de cada grupo, respectivamente) no se aislaron microorganismos viables.

Las especies más frecuentes fueron *E. faecalis* ( $\alpha_h=0,67$ ) y *S. epidermidis* ( $\alpha_h=0,54$ ), seguidas de *S. marcescens* ( $\alpha_h=0,34$ ) y *E. coli* ( $\alpha_h=0,27$ ), mientras que las

minoritarias fueron *E. faecium*, *S. liquefaciens* ( $\alpha_h=0,02$ ), *K. pneumoniae* ( $\alpha_h=0,11$ ) y *K. oxytoca* ( $\alpha_h=0,24$ ). En cuanto a los clones, los más dominantes y compartidos en la cohorte brote fueron *S. marcescens* SM-2 (clon B), *E. faecalis* E-4, *E. coli* Ec-1 y *S. epidermidis* SE-2 (**Figuras 41 y 42**).



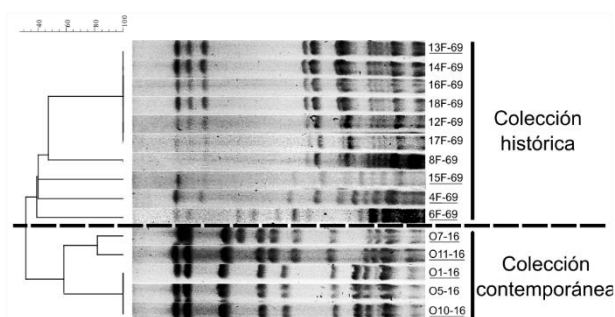
**Figura 41. Dendrogramas de las especies cultivadas.** Especies del mismo género *Klebsiella*, *Enterococcus*, y *Serratia* se analizaron a la vez. El clon S-1 pertenece a *S. liquefaciens*; el clon S-2 corresponde al que llamamos clon B de *S. marcescens*; el clon S-3 corresponde al clon A de *S. marcescens*.



**Figura 42. Esquema de clones detectados en cada paciente del grupo brote en los distintos tiempos de recogida de muestras. NC: no creció ningún microorganismo en los medios selectivos. 0 (gris): muestras no recibidas. A: nº aislados. C: nº clones distintos; M: nº de muestras. \**Enterococcus-7* corresponde a la especie *E. faecium* y el resto a *E. faecalis*.**

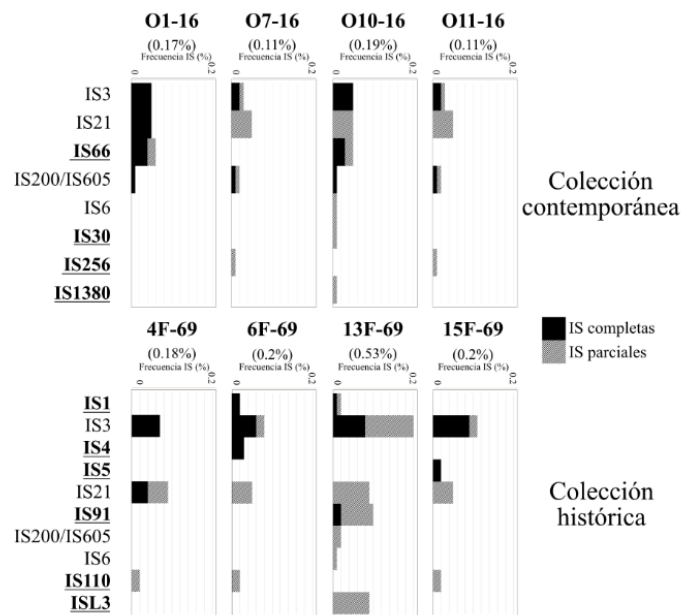
#### 4. Estudio extendido de las cepas causantes del brote contemporáneo y el histórico.

Diez de los 21 aislados de la colección histórica se recuperaron y se analizaron posteriormente. Se seleccionaron ocho aislados como representantes clonales del brote contemporáneo (n=4, los procedentes de los pacientes neonatos con prematuridad O1, O7, O10 y O11) y del brote histórico (n=4), de acuerdo con los resultados de los dendrogramas de pulsotipos, y se secuenció su genoma completo (**Figura 43**).



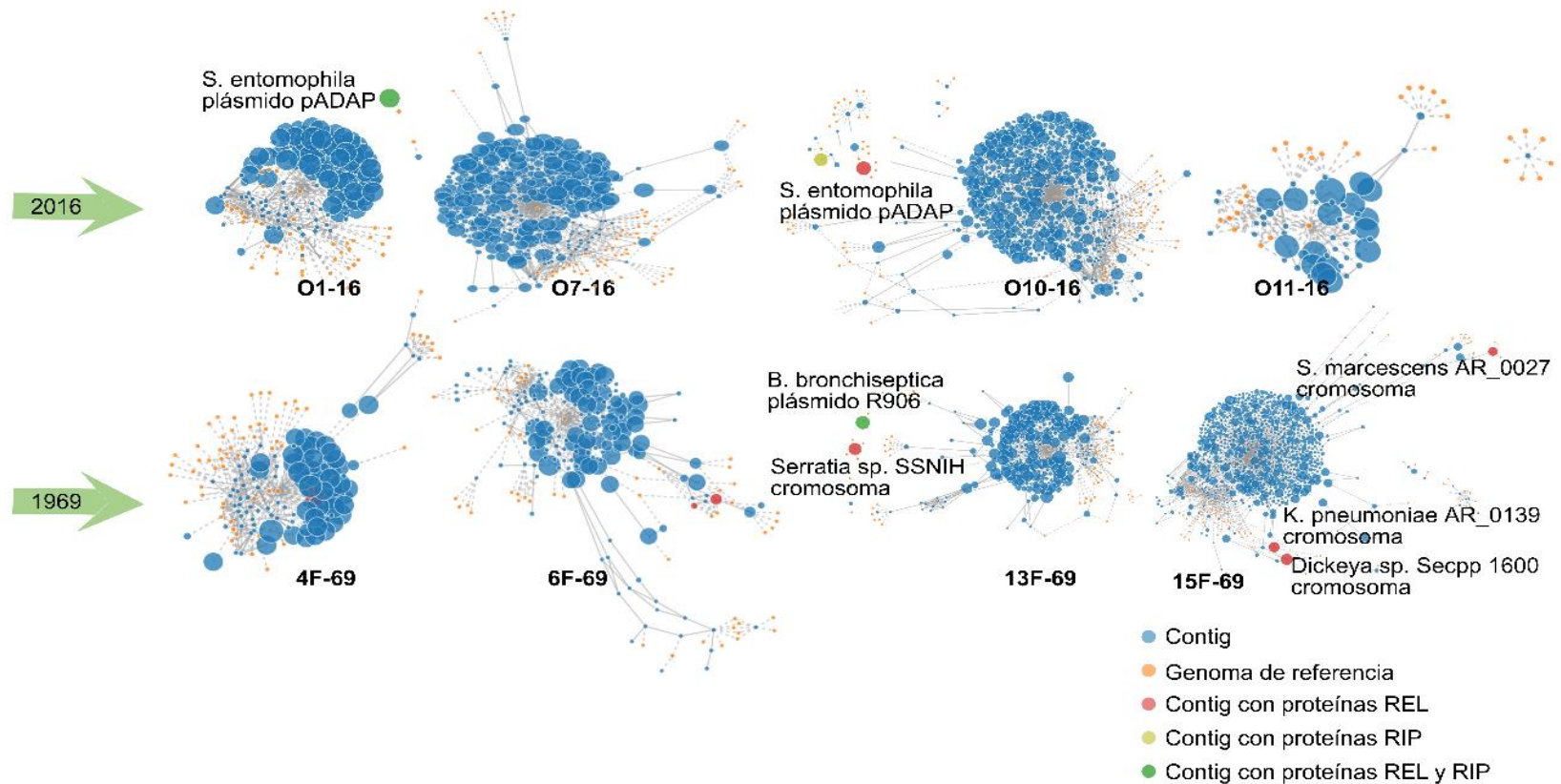
**Figura 43. PFGE conjunto de los aislados de *S. marcescens* encontrados en 2016 (colección contemporánea) y los conservados desde 1969 (colección histórica).** Subrayado: aislados de los que se realizó secuenciación de genoma completo.

Los aislados históricos 13F-69 y 15F-69 eran notablemente diferentes de los demás (mayor contenido de CDS, tRNA y IS) (**Figura 44**). Los aislados contemporáneos parecían presentar una mayor estabilidad genética con un contenido total de IS ligeramente inferior. De las 14 familias de IS diferentes detectadas, ocho se detectaron en un solo aislado, cinco estaban presentes en dos o tres aislados, y sólo IS3, IS21 e IS200/IS605 parecían estar representados de forma regular. Es importante señalar que IS200/605 sólo se detectó en su forma completa en los cuatro aislados contemporáneos, aunque también se observó una secuencia parcial de IS200/605 en la cepa 13F-69. La prueba de Mann-Whitney para muestras independientes mostró diferencias significativas ( $p < 0,05$ ) entre el porcentaje total de IS históricos y contemporáneos. Sin embargo, cuando se extrajo de la comparación la cepa atípica 13F-69, no se obtuvieron diferencias significativas.



**Figura 441.** IS detectadas en los 8 aislados de *S. marcescens* estudiados y su frecuencia en los genomas anotados. Representado como pares de bases (pb) de IS/pb del genoma (%). La frecuencia total de IS en cada genoma se expresa entre paréntesis. El resultado de ISEScan incluyó IS integrales e IS parcialmente presentes (barras negras y grises, respectivamente). Las etiquetas IS subrayadas y en negrita se encontraron sólo en una colección, pero no en la otra.

Según el informe de PlacnetW, se detectaron plásmidos incompletos en las cepas O10-16, 13F-69 y 15F-69 (**Figura 45**). El único plásmido completo, R906, correspondía al que se había notificado anteriormente en *Bordetella bronchiseptica* y se encontró en el aislado histórico 13F-69. Este plásmido portaba genes que codificaban mecanismos de resistencia a antibióticos (*oxa-2*, *sul-1*, *qacEΔ1*, *aph(3')-II*), transportadores de múltiples fármacos *mepC* y *mdtE*, al níquel-cobalto (*cnrA*) y al amonio cuaternario (*sugE*), junto a proteínas del sistema de secreción de tipo IV, toxinas (*fitB*) y transportadores de hierro. Otros fragmentos identificados como plásmidos putativos compartían secuencias con elementos extracromosómicos de *Serratia entomophila*, y *Serratia* spp. o con el cromosoma de *S. marcescens*, *K. pneumoniae*, *Serratia* sp. o *Dickeya* sp. En cuanto a los fagos lisogénicos presentes en los aislados estudiados, se identificaron un total de 5 profagos completos (uno en 13F-69, tres en O7-16, y el último en O11-16), todos ellos pertenecientes al orden *Caudovirales*.



**Figura 45. Resultado de PlacnetW que refleja las secuencias de plásmidos detectadas.** Todos los *contigs* de cada genoma se representan como nodos azules conectados por líneas continuas. Los nodos amarillos representan los genomas de referencia de la base de datos NCBI BLAST. Las líneas discontinuas enlazan cada *contig* con su(s) genoma(s) de referencia más cercano(s). Por último, cada genoma aparece representado por todos sus *contigs* mostrados alrededor de su identidad de genoma de referencia más cercana. Los *contigs* que contienen secuencias de plásmidos están resaltados en nodos de distinto color. REL: relajasa. RIP: proteína de iniciación de la replicación.

Tanto en los brotes históricos como en los contemporáneos, la mayoría de los genomas contenían tres o cuatro genes de resistencia a los antibióticos, excepto los aislados históricos 13F-69 y 15F-69, que portaban 10 y 6 genes, respectivamente (**Tabla 13**). Curiosamente, 13F-69 mantenía su sensibilidad fenotípica a casi todos los antibióticos, excluyendo los afectados por la resistencia constitutiva de *S. marcescens* (**Tabla 14**). Sólo dos aislados, uno de cada periodo, mostraron resistencia a los antibióticos betalactámicos de amplio espectro, presumiblemente debido a mecanismos que implican la desrepresión de *ampC*.

El esquema hecho a partir del pangenoma de todos los aislados de *Serratia* publicados y los de este estudio (cgMLST), reveló que el 7,6% de los genes del pangenoma estaban presentes en >95% de los 452 aislados (genoma central) (**Figura 46A**). La filogenia se dedujo a partir de las variaciones alélicas en este genoma central. Los aislados contemporáneos se agruparon y se relacionaron estrechamente con una muestra clínica humana obtenida de sangre en 2006 en el Reino Unido (cepa: 2880STDY5682913, BioProyecto: PRJEB5065). Los aislamientos históricos se situaron en una rama separada cerca de una cepa del mismo BioProyecto recogida en 2011 (cepa: 2880STDY5683020). El aislado histórico 13F-69 fue el más distante y estaba relacionado con cepas clínicas de Brasil (referencia del BioProyecto: PRJNA420811). Es importante señalar que la mayoría de los genomas públicos anotados tienen un origen clínico u hospitalario (76,5%), mientras que otras fuentes [ambiental (11,7%), animal (3,3%), vegetal (2,4%) y desconocido (6%)] tienen menor representación.

**Tabla 13. Resistoma de los ocho aislados de *S. marcescens* del estudio y principales clases de antibióticos que pueden ser objetivos potenciales de los productos que codifican los genes de resistencia. OQX: olaquinox. CHL: cloranfenicol. QUATS: compuestos de amonio cuaternario. (p): presente en plásmido.**

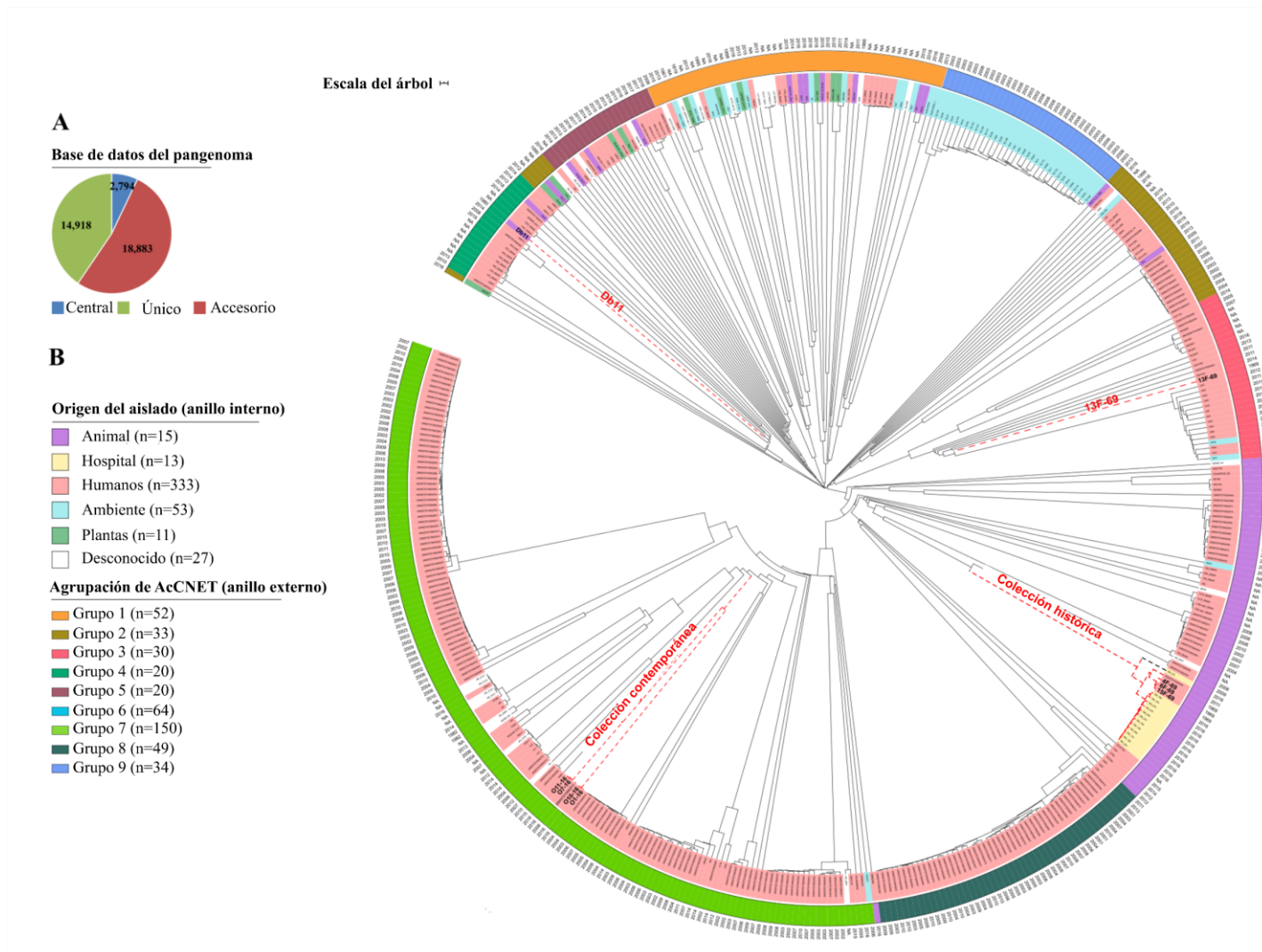
<i>Clase de antibiótico / Resistoma</i>	<i>O1-16</i>	<i>O7-16</i>	<i>O10-16</i>	<i>O11-16</i>	<i>4F-69</i>	<i>6F-69</i>	<i>13F-69</i>	<i>15F-69</i>
<b><i>Beta-lactámicos</i></b>								
<i>blaOXA-2</i>							✓ (p)	
<i>blaSRT-1</i>	✓		✓				✓	
<i>blaSRT-2</i>		✓		✓	✓	✓		✓
<b><i>Aminoglucósidos</i></b>								
<i>aac(6')</i>		✓	✓	✓	✓	✓	✓	✓
<i>aph(3'')-Ib</i>							✓ (p)	✓
<i>aph(6)-Id</i>							✓	✓
<b><i>Quinolonas</i></b>								
<i>qnrB75</i>							✓	
<i>qnrE1</i>	✓		✓		✓	✓		✓
<b><i>Sulfonamida</i></b>								
<i>sul-1</i>							✓ (p)	
<b><i>Multifármacos (oqx, chl)</i></b>								
<i>oqxB24</i>							✓	
<i>oqxB9</i>	✓	✓	✓	✓	✓	✓		✓
<b><i>QUATS</i></b>								
<i>qacEA1</i>							✓ (p)	
<b><i>Tetraciclina</i></b>								
<i>tet(41)</i>							✓	

El análisis del genoma accesorio y el proteoma inferido que se realizó con el programa AcCNET clasificó todos los genomas en nueve clústeres. Las cepas históricas se ubicaron en el clúster 6, con la excepción del aislado 13F-69 que se incluyó en el clúster 3 y las cepas contemporáneas que se ubicaron en el clúster 7. No podemos descartar la

posibilidad de que la sobrerrepresentación del clúster 7 haya sido influenciada por un sesgo en el muestreo (ver Discusión). Esta distribución con respecto al genoma accesorio se comparó con la inferida por el análisis cgMLST en la **Figura 46B** y todas las cepas se agruparon de forma similar. No se encontraron aislamientos de brotes en ninguno de los otros siete clústeres de *S. marcescens*.

**Tabla 14. Concentraciones mínimas inhibitorias (CMI) observadas (mg/L) o diámetro de la zona de inhibición (ZD) (mm) de cada antibiótico para las cepas de *S. marcescens* estudiadas.** Los valores de CMI o los diámetros de zona por encima del corte clínico (basado en los criterios del EUCAST 2019) se destacan con números subrayados y en negrita. NA: no disponible. \*: antibióticos probados solo por prueba de difusión en disco (25 µg para la estreptomina y 30 µg para la tetraciclina).

Antibióticos	Disponible desde	Colección contemporánea				Colección histórica			
		MICs (mg/L) o ZD (mm)				MICs (mg/L) o ZD (mm)			
		O1-16	O7-16	O10-16	O11-16	4F-69	6F-69	13F-69	15F-69
<b>Beta-lactámicos</b>									
Ampicilina	1961	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>
Amoxicilina-Clavulanato	1972	<u>&gt;16/8</u>	<u>&gt;16/8</u>	<u>&gt;16/8</u>	<u>&gt;16/8</u>	<u>&gt;16/8</u>	<u>&gt;16/8</u>	<u>&gt;16/8</u>	<u>&gt;16/8</u>
Piperacilina-Tazobactam	1993	≤8	≤8	≤8	≤8	≤8	<u>&gt;64</u>	≤8	≤8
Cefazolina	1971	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>
Cefuroxima	1978	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>	<u>&gt;16</u>
Cefoxitin	1972	16	<u>&gt;16</u>	<u>&gt;16</u>	16	16	<u>&gt;16</u>	16	16
Ceftazidima	1984	≤1	≤1	<u>&gt;16</u>	≤1	≤1	<u>&gt;16</u>	≤1	≤1
Cefotaxima	1980	≤1	≤1	<u>&gt;32</u>	≤1	≤1	<u>&gt;32</u>	≤1	≤1
Cefepime	1994	≤1	≤1	<u>&gt;16</u>	≤1	≤1	<u>8</u>	≤1	≤1
Aztreonam	1986	≤1	≤1	<u>&gt;16</u>	≤1	≤1	<u>&gt;16</u>	≤1	≤1
Imipenem	1977	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1
Ertapenem	2001	≤0.5	≤0.5	<u>&gt;4</u>	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5
<b>Aminoglucósidos</b>									
Gentamicina	1963	≤2	≤2	<u>&gt;8</u>	≤2	≤2	≤2	≤2	≤2
Tobramicina	1997	≤2	≤2	<u>&gt;8</u>	≤2	≤2	4	4	≤2
Amikacin	1976	≤8	≤8	<u>&gt;32</u>	≤8	≤8	≤8	≤8	≤8
Estreptomina*	1943	15	15	<u>11</u>	16	<u>13</u>	17	<u>0</u>	<u>9</u>
<b>Ciprofloxacina</b>	1987	≤0.5	≤0.5	<u>1</u>	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5
<b>Ácido nalidixico</b>	1965	≤16	≤16	>16	≤16	≤16	≤16	≤16	≤16
<b>TMP/SMX</b>	1968-78	≤2/38	≤2/38	≤2/38	≤2/38	≤2/38	≤2/38	≤2/38	≤2/38
<b>Tetraciclina*</b>	1953	<u>14</u>	<u>12</u>	<u>15</u>	<u>13</u>	21	21	<u>0</u>	<u>12</u>



**Figura 46. Análisis filogenético de 452 genomas de *S. marcescens*.** **A:** Distribución del pangenoma total tras el análisis cgMLST en núcleo (genes presentes en al menos el 95% de los aislados, 7,6%), accesorio (genes presentes en >2 aislados y menos del 95%; 51,6%) y genes únicos (genes exclusivos de un solo aislado; 40,8%). **B:** árbol filogenético inferido a partir del enfoque cgMLST, incluyendo los metadatos de la fuente (colores) y el año (texto). Las cepas de estas colecciones y el genoma de referencia de la especie *S. marcescens* (cepa Db11, referencia GCA\_000513215.1) se destacan en los nodos rojos discontinuos. El círculo externo señala los clústeres inferidos por el análisis AcCNET del genoma accesorio.

## V. DISCUSIÓN

En este capítulo se ha descrito la sucesión de la composición de la microbiota intestinal desde el meconio hasta las tres primeras semanas de vida de recién nacidos prematuros de bajo peso, comparando dos escenarios epidemiológicos de la misma UCIN. El principal resultado de este trabajo está relacionado con la influencia del brote de *S. marcescens* en los patrones de la microbiota intestinal de los lactantes, incluido el meconio, que podría haberse modificado incluso antes del nacimiento en relación con el ingreso hospitalario de las madres para prevenir el parto prematuro. Aunque existe una marcada variabilidad interindividual, la microbiota intestinal de los neonatos prematuros de bajo peso está dominada por el filo *Proteobacteria*, en particular por *E. coli*. Además, la esperada expansión del ecosistema bacteriano tras el nacimiento parece retrasarse, probablemente en relación con la exposición a los antibióticos. En este sentido, es importante señalar la falta de datos sobre la prescripción de antibióticos en el grupo de control, que representa una importante limitación.

En lo que respecta al grupo brote, todos los bebés excepto dos, recibieron terapia antimicrobiana empírica inmediatamente después del nacimiento y la mayoría requirió tratamientos antibióticos guiados por los resultados de los cultivos durante su ingreso, incluyendo hasta tres familias antimicrobianas diferentes. Sin embargo, la elevada incidencia de complicaciones infecciosas entre los neonatos prematuros es el principal argumento utilizado para justificar este amplio uso, especialmente en aquellos con bajo peso al nacer. Para reducir el uso de antibióticos en la UCIN se necesitan biomarcadores tempranos de sepsis en prematuridad, junto con el desarrollo de enfoques basados en el microbioma [76], [297].

La prematuridad es la principal causa de morbilidad y mortalidad neonatal y el establecimiento de una microbiota intestinal adecuada es una de las estrategias más prometedoras para mejorar su salud y reducir el impacto de las secuelas posteriores [298], [299]. Los lactantes prematuros ingresados en una UCIN tienen un alto riesgo de infección y *S. marcescens* es uno de los patógenos nosocomiales más relevantes. En este caso y a diferencia de los brotes en adultos, que suelen estar colonizados en el tracto urinario o respiratorio, el tracto gastrointestinal representa el reservorio más importante para que ocurra la contaminación cruzada [300].

En este estudio se observó un considerable enriquecimiento de *Proteobacteria* en ambos grupos de recién nacidos prematuros, aunque el enriquecimiento fue significativamente mayor en el grupo brote desde la segunda semana de vida, en concordancia con otros autores [301]–[304]. En estudios anteriores, se observó que *Proteobacteria* (principalmente *Enterobacter* y *Photorhabdus*), *Firmicutes* (principalmente *Enterococcus* y *Lactobacillus*) y *Actinobacteria* (*Bifidobacterium*) dominaban la composición microbiana del meconio [305]. *E. coli*, especies de *Staphylococcus* y de *Klebsiella*, así como una alta tasa de anaerobios facultativos, también aparecen habitualmente en el meconio de los neonatos prematuros [106], [306]. Aquí, se observaron diferencias significativas entre el grupo control, dominado por *E. coli* y *Enterococcus*, y el grupo brote, con mayores densidades de *Proteobacteria*. Datos recientes han demostrado que la secuenciación de amplicones tiene importantes limitaciones a la hora de estudiar el ecosistema intestinal de los lactantes prematuros [307]. Ésta es una limitación importante del trabajo, que probablemente refleja una menor proporción de *Bifidobacterium* de la que habría y de la que se podría detectar con técnicas como la metaproteómica, como se argumentó en el Capítulo 1. Un trabajo reciente también demostró que la densidad de *Bifidobacterium* está relacionada con la edad gestacional [272].

Curiosamente, las muestras de meconio de ambos grupos se caracterizaron por una alta variabilidad interindividual y una diversidad alfa similar a la de las muestras fecales posteriores, lo que apunta a un marcado retraso en el establecimiento del ecosistema. Se han publicado resultados similares en recién nacidos prematuros de muy bajo peso [275]. Otros autores han descrito un aumento de la complejidad de la microbiota fecal durante la estancia en la UCIN [308].

Aunque la colonización intestinal fetal intrauterina continúa siendo un tema controvertido [70], [267], [268], la detección de *Serratia* en muestras de meconio sugiere la posibilidad de que su colonización pueda empezar después de la ruptura de membranas y antes del nacimiento. La mayoría de las madres del estudio (7 de 12) fueron ingresadas para prevenir un parto prematuro por diversas causas durante una mediana de 9 días antes del nacimiento, con un rango de 1 a 49 días. En cambio, la madre cuyo bebé murió por sepsis por *S. marcescens* fue ingresada sólo el día anterior al parto. Mientras que este microorganismo apenas estaba representado en el meconio, era muy dominante en el día 7, hecho que probablemente precedió a la translocación sanguínea y al episodio de sepsis

que se produjo en el día 10. Por tanto, explorar sistemáticamente el enriquecimiento potencial de poblaciones bacterianas intestinales específicas podría contribuir a la prevención de la bacteriemia en grupos de pacientes susceptibles, como los bebés prematuros. Recientemente, se ha reportado una novedosa metodología funcional que utiliza compuestos orgánicos volátiles como biomarcadores para la detección temprana del enriquecimiento bacteriano intestinal [309]. Estos resultados también demostraron que algunos microorganismos epidémicos, como *S. marcescens*, son capaces de colonizar y acabar infectando a los neonatos prematuros incluso cuando se han aplicado las medidas preventivas más avanzadas. El análisis PFGE agrupó los aislados que colonizaban a los 12 lactantes del grupo brote en dos complejos principales, mientras que la secuenciación de amplicones reveló una estrecha relación entre ellos, lo que sugiere la existencia de un ancestro común. Estas técnicas moleculares también revelaron que el viruloma de la cepa causante de la bacteriemia y la muerte era idéntico al de otras cepas de pacientes con evolución clínicamente satisfactoria, lo que refuerza la hipótesis de que la barrera poco clara que delimita la colonización de la infección está influida por numerosos factores.

En las tres muestras de meconio analizadas en este estudio se detectaron algunas ASV asignadas a las bacterias depredadoras *Bdellovibrio*, *Vampirovibrio* y *Peredibacter*. Dichas bacterias necesitan a otras bacterias para crecer y reproducirse y se consideran importantes equilibradores ecológicos de las comunidades microbianas [310]. Pocos estudios se han centrado en estas bacterias en los ecosistemas humanos. Sin embargo, su presencia en las muestras de meconio sugiere que podrían no ser tan anecdóticas en la microbiota intestinal. La inoculación de una depredadora se considera una herramienta ecológica para modular las comunidades bacterianas, teniendo en cuenta la especificidad de las parejas depredadora-presa [311], y se abordará en el Capítulo 4 de esta tesis.

La política de nutrición de prematuros promueve que los neonatos sean alimentados con leche materna, aunque ésta suele combinarse con leche humana de donantes y con fórmulas adaptadas a los prematuros. Todos los neonatos prematuros participantes recibieron los tres tipos de leche durante el estudio y, aunque no se detallan estos datos, somos conscientes de que este factor también influye en el establecimiento de la microbiota intestinal. La leche materna puede remodelar la microbiota intestinal del lactante [269], aportando su propia microbiota específica [312]–[314], pero también fomentando el aumento de una población en particular por su acción prebiótica [315]. Por lo tanto, sería conveniente incluir en el esquema de perfiles de la microbiota la

diferenciación entre bacterias vivas y muertas para identificar a los verdaderos colonizadores de los pasajeros bacterianos transitorios asociados a la ingesta de alimentos [316].

Los estudios epidemiológicos combinados con la tipificación bacteriana ayudan a detectar clones de alto riesgo adaptados al entorno hospitalario, que tienen el potencial de propagarse globalmente, así como a mejorar las medidas de contención locales. Las técnicas de tipificación bacteriana han evolucionado a lo largo del tiempo, desde las pruebas fenotípicas (biotipos, antibiotipos) en los años 70, hasta los actuales estudios moleculares basados en el ADN, con una capacidad de discriminación superior, mejor reproducibilidad interlaboratorio y mayor rentabilidad [317], [318]. Las técnicas de tipificación más utilizadas para patógenos relevantes relacionados con el ser humano son PFGE y MLST; sin embargo, aún no se ha implementado un esquema de MLST para *S. marcescens* [281]. Actualmente, la WGS es una opción mucho mejor, dado que es rentable en comparación con el MLST y revela la composición del genoma accesorio. Además, permite un enfoque comparativo gen por gen según una base de datos de alelos del pangenoma previamente construida (PGAdB) y la posterior agrupación de los genomas (wgMLST) [319]. El conjunto de datos también se puede filtrar para seleccionar, por ejemplo, el genoma central (cgMLST), utilizado recientemente para distinguir los linajes de *S. marcescens* adaptados al hospital de los de origen ambiental [320]. Más allá de las investigaciones epidemiológicas, la secuenciación es útil para evaluar la evolución genética de las especies bacterianas, especialmente cuando se analiza el genoma accesorio, como se ha descrito recientemente para *E. faecalis* [321]. Nuevas herramientas de análisis genómico y proteómico como, por ejemplo, AcCNET, han mejorado significativamente dicho análisis [296].

En el presente estudio, se incluyeron en el análisis más de 36.000 grupos de genes. Probablemente, los pangenomas amplios corresponden a organismos ubicuos, sometidos a muchas condiciones ambientales. Por el contrario, los pangenomas de un organismo bacteriano con alto grado de especialización de hábitat, son mucho más limitados. El pangenoma inferido a partir de 50 genomas de *Streptococcus* es considerablemente menor, 5.398 grupos de genes [322], mientras que el tamaño del pangenoma de *Serratia* se acerca más al de otro género ambientalmente promiscuo, *Vibrio*, compuesto por más de 26.000 clústeres de genes [323]. Estos hallazgos revelan la gran diversidad de genes únicos presentes en *S. marcescens*, que muy probablemente corresponden a genes

presentes en elementos genéticos móviles, islas genómicas, transposones y profagos.

Los aislados tanto de las colecciones históricas como de las recientes se agruparon de forma consistente por periodo de muestreo, siendo el histórico 13F-69 el aislado más destacado en todos los análisis realizados. Tanto el análisis del genoma central como el del genoma accesorio agruparon ambas colecciones de forma similar (**Figura 17B**), hecho que se ha descrito previamente en escenarios similares [320], [324]. Esta observación indica que *S. marcescens* parece evolucionar en la misma dirección tanto para los genes centrales como para los accesorios.

*S. marcescens* está particularmente asociada al entorno hospitalario y, particularmente, al de la UCIN, aunque ciertos linajes tienen mayor asociación con sepsis que otros [325]. La coexistencia de varios clones causando un mismo brote de *S. marcescens*, tal y como ocurrió en nuestra colección histórica, había sido reportada previamente [300], [326], [327]. El esquema cgMLST reveló que la mayoría de estos aislados comparten un ancestro común con la mayoría de las cepas clínicas publicadas. Sin embargo, una advertencia de este análisis es la baja representación de aislados no clínicos (menos del 24% del total) en las bases de datos genómicas públicas.

Un hallazgo relevante fue que la cepa 13F-69 alberga un plásmido altamente conservado descrito previamente en *B. bronchiseptica*, un patógeno zoonótico de las vías respiratorias. El intercambio de plásmidos entre especies filogenéticamente distantes (en este caso, de taxones de *Beta-Proteobacteria* a *Gamma-Proteobacteria*) es un acontecimiento poco frecuente. Esta transmisión parece haber ocurrido a pesar de los nichos naturales distantes, aunque el eje intestino-pulmón o tal vez una coinfección pulmonar podrían justificar un nexo común [328]. En cualquier caso, es destacable la plena conservación del plásmido R906 tras un periodo de reposo de 47 años. Por supuesto, el origen del plásmido se dedujo a partir de los datos disponibles y no podemos descartar que *Serratia* sea su huésped primario. Por último, se identificaron cuatro profagos en la colección contemporánea mientras que solo se encontró uno en la colección histórica, un resultado probablemente relacionado con los requisitos actuales de plasticidad del genoma [329].

Las colecciones históricas de bacterias nos proporcionan datos relevantes del periodo pre-antibiótico [330]–[332]. Estos aislados históricos se recogieron entre 1969 y 1970 durante la clásica edad de oro de los antibióticos, pero no habían sido expuestos a

los antimicrobianos recién descubiertos. Sin embargo, los aislados históricos portaban genes que codificaban la resistencia a los antibióticos y desinfectantes que se comercializaron años después. Aunque el olaquinox es un agente antimicrobiano que se utiliza desde 1975 como promotor del crecimiento de los animales de granja, los genes de resistencia al olaquinox *oqx10*, *oqx24* y *oqx9* se encontraron sistemáticamente en los aislados históricos. Este hallazgo podría explicarse por otras funciones que proporcionan las bombas de eflujo de múltiples fármacos codificadas por los genes *oqx*, incluyendo la resistencia a las quinolonas y al cloruro de benzalconio [333]. Los genes *qnr* de resistencia a quinolonas mediada por plásmidos se descubrieron en 2002 [334]. Sin embargo, *qnrB75* y *qnrE1* ya estaban presentes en los aislados históricos, probablemente porque la función *qnr* no está relacionada principalmente con la resistencia a los antibióticos [335]. Alternativamente, podrían haber surgido bajo la selección de quinolonas. Hoy en día, las quinolonas o fluoroquinolonas no se utilizan en pediatría. Sin embargo, en 1969, durante el primer brote, se utilizó ampliamente la primera quinolona, el ácido nalidíxico, para controlar la enteritis epidémica por *E. coli* enteropatógeno en el Hospital Infantil de La Paz, lo que podría haber contribuido a la selección de los genes *qnr* en las cepas intestinales de *Serratia*.

La resistencia intrínseca en *Serratia* incluye penicilinas, macrólidos, lincosaminas, linezolid, glucopéptidos, quinupristina-dalfopristina y rifampicina [282]. La resistencia a la ampicilina y a la tetraciclina se ha considerado normalmente intrínseca, pero esa suposición podría deberse al sesgo impuesto por las cepas clínicas, por lo que varias cepas ambientales podrían ser susceptibles [336]. *S. marcescens* alberga una betalactamasa AmpC cromosómica, constitutiva. Los mutantes en AmpC y AmpD (que restringen la función de AmpR, el represor de AmpC) dan lugar a una expresión constitutiva de AmpC, capaz ahora de inactivar las cefalosporinas de tercera y cuarta generación [337]–[339]. Esto es lo que parece haber ocurrido en la cepa histórica 6F-69 y la cepa contemporánea O10-16, resistentes a la cefotaxima, la ceftazidima y la cefepima. Sin embargo, no pudimos detectar cambios mutacionales en los genes estructurales o reguladores conocidos de la betalactamasa AmpC (*ampC*, *ampD*, *ampE*, *ampG*, *ampR*, y sus secuencias flanqueantes). Esta discrepancia genotipo-fenotipo se ha observado previamente en otras especies [340], por lo que, probablemente, la compleja regulación de la producción de AmpC en *Serratia* debería aclararse más en el futuro. La cepa O10-16 también era resistente a los carbapenems, probablemente debido a la

conurrencia de mutaciones en la porina que impedían la entrada del antibiótico [341]. En la década de 1970, la terapia antimicrobiana de *S. marcescens* se basaba principalmente en la carbenicilina y la gentamicina, a veces combinadas con fosfomicina [342]. Sin embargo, el aislado histórico 6F-69 presentaba una resistencia a las cefalosporinas que no había sido descubierta en 1969, probablemente debido a la exposición hospitalaria y a la selección de mutantes deprimidos por la cefaloridina, introducida en 1962. Asimismo, se detectó resistencia a la gentamicina en el aislado moderno O7-16 a pesar de que la mayoría de los aislados de *S. marcescens* solían ser susceptibles debido al uso poco frecuente de este antimicrobiano [324].

Por último, es pertinente discutir la posible influencia del prolongado periodo de reposo que experimentaron las cepas históricas. En una publicación clásica [343], el grupo de Werner Arber comparó una cepa de *E. coli* K-12 (W3110) descubierta en 1960 con subclones recuperados a partir de la original (que era guardada inmediatamente) en 1965 y 1972. A lo largo de esos 30 años, el cultivo en reposo parecía haber desarrollado un alto grado de diversidad genética, lo cual era un hallazgo inesperado teniendo en cuenta la reducida tasa de propagación durante el almacenamiento. La fuente de la mayor parte de dicha variabilidad se atribuyó a la transposición de IS [343], [344], lo que sugiere que el genoma de las bacterias en reposo es más dinámico de lo que se creía durante mucho tiempo. En el presente estudio, se encontraron varias IS en las cepas, con más diversidad y una proporción ligeramente mayor en la colección histórica. Como ya se ha mencionado, la WGS demostró el transporte en el resistoma de la cepa histórica 13F-69 de varios genes resistentes a los antibióticos, pero de algunos de ellos, como *qnrB*, *qnrE* que actúan sobre las quinolonas, o *tet(41)* que codifica una bomba de eflujo de tetraciclina, no se produjo el fenotipo de resistencia esperado. No se puede descartar una disminución de la expresión como estrategia para reducir el coste de *fitness* de los organismos sometidos a un periodo de descanso muy prolongado [345], [346].

En resumen, el análisis filogenético realizado con grandes bases de datos de genomas nos permitió demostrar que tanto las estrategias de tipificación del genoma central como las del genoma accesorio agruparon las cepas de *S. marcescens* de forma similar. Tanto las cepas históricas como las de brotes recientes se agruparon en dos grupos en los que se encontraban la mayoría de las cepas clínicas. La única excepción fue el aislado histórico 13F-69, más cercano a las cepas con un estilo de vida más "ambiental", probablemente abierto a un alto intercambio genético incluyendo la adquisición de un

plásmido completo de otro género ambiental, *Bordetella* [347]. Las cepas históricas portaban genes de resistencia hacia antimicrobianos aún no comercializados en 1969, lo que sugiere la presencia de factores que impulsan la evolución y transmisión de esos genes que no están necesariamente relacionados con la terapia antibiótica. Por último, los aislados históricos mostraban un cierto dinamismo evolutivo durante su periodo de reposo, aunque mucho menor que el descrito anteriormente para *E. coli*.

Globalmente, estos resultados indican que, independientemente de su entorno perinatal, los neonatos prematuros ingresados en la misma UCIN están inicialmente colonizados por comunidades microbianas similares que posteriormente evolucionan según las condiciones individuales. La influencia de un brote de *Serratia* en el establecimiento de la microbiota intestinal parece ser relevante ya desde los primeros días de ingreso. Esto pone de manifiesto la importancia del entorno en relación con el patrón de colonización intestinal de los lactantes prematuros hospitalizados. Por último, se ha demostrado que la WGS es una herramienta fundamental para la tipificación bacteriana, permitiendo el análisis de los genomas centrales y accesorios con una potente discriminación. Además, amplía nuestra comprensión de la historia filogenética de los aislados de brotes epidemiológicos, ofreciendo una perspectiva evolutiva que será útil para desarrollar intervenciones de Salud Pública específicas en el futuro.

## VI. DIVULGACIÓN

Los resultados de este capítulo se han publicado en la revista *PLoS One*. bajo el título: *Influence of a Serratia marcescens outbreak on the gut microbiota establishment process in low-weight preterm neonates* (DOI: [10.1371/journal.pone.0216581](https://doi.org/10.1371/journal.pone.0216581)), y en la revista *Frontiers in microbiology* bajo el título *Genomics of Serratia marcescens Isolates Causing Outbreaks in the Same Pediatric Unit 47 Years Apart: Position in an Updated Phylogeny of the Species* (DOI: [10.3389/fmicb.2020.00451](https://doi.org/10.3389/fmicb.2020.00451)), y se incluyen en **Anexos**.

# **CAPÍTULO 4**

## **BACTERIAS DEPREDADORAS COMO CONTROL DE *Pseudomonas aeruginosa* y *Serratia marcescens***

*Cuando **B. bacteriovorus** depreda sobre diferentes colecciones de presas clínicas...*

*¿Cuál es el **rango de depredación** para cada especie?*

*¿Cuál es la **eficiencia de depredación** dentro de cada colección?*

*¿Existen **diferencias** en la depredación que puedan asociarse a la depredadora o la presa?*

*¿Pueden asociarse estas diferencias a algún **factor ambiental o genético**?*

## I. ANTECEDENTES

A la edad de oro de los antibióticos a mediados del siglo XX le siguió la crisis de las bacterias multirresistentes (MDR) [348]. Para identificar terapias alternativas al uso de antimicrobianos, es necesario comprender las reglas ecológicas que rigen la microbiota comensal la cual incluye, tal y como acostumbramos a ver en la fauna y flora macroscópica, agentes competidores, colaboradores y depredadores [349]–[352]. En la naturaleza, se sabe que las bacterias depredadoras desempeñan un papel importante en el mantenimiento de las poblaciones bacterianas presa. Por un lado, su acción depredadora favorece la producción de biomasa procedente de las comunidades microbianas a las que atacan y la ponen a disposición de otros agentes y, por otro, favorecen su eliminación a través de su propio consumo, fomentando la diversidad de microorganismos y contribuyendo a la estabilización global del ecosistema [167], [168]. Por ello, se ha planteado la posibilidad de aprovechar este papel ecológico de las bacterias depredadoras en la lucha contra los patógenos clínicos, ya que son microorganismos dinámicos que experimentan (al igual que sus oponentes) una continua adaptación física, morfológica y metabólica al ecosistema. Esta reciprocidad evolutiva es la base de la coevolución, en la que la adaptación de un actor no sólo promueve el cambio en su oponente, sino que la adaptación de éste también genera selección como respuesta evolutiva al primer actor [178]. Estaríamos, por tanto, ante la ocasión de emplear un organismo vivo, capaz de adaptarse y superar las resistencias que plantee su presa, como agente terapéutico modulador de poblaciones patógenas.

Para desarrollar el potencial de las bacterias depredadoras como “agentes antimicrobianos vivos” es necesario comprender su mecanismo de depredación, lo que implica estudiar a fondo sus preferencias y afinidades, es decir, su rango de depredación (a quién puede depredar) y su eficiencia (con cuánto éxito) entre las diferentes especies. No obstante, dentro de una misma especie bacteriana, la depredación no es homogénea ni universal, sino que varía de un linaje genético a otro, algo que se ha denominado rango de depredación intraespecífico (I-PR, por sus siglas en inglés). Asimismo, para cada cepa o linaje genético, la bacteria depredadora puede tener una eficiencia de depredación (PE, por sus siglas en inglés) diferente, dependiendo de la proporción de células que haya depredado. Este comportamiento de depredación polimórfico se basa, fundamentalmente, en la composición de la envoltura celular de la especie o de la presa individual, que a su vez está influenciada por las condiciones ambientales [166], [180], [353]. Por el

momento, el estudio acerca de los factores que influyen en la preferencia de las bacterias depredadoras por las presas y las diferencias en su actividad depredadora entre y dentro de las distintas especies ha sido limitado y, sobre todo, se ha centrado en colecciones bacterianas ambientales y no tanto en las clínicas [179], [354].

Las bacterias depredadoras más estudiadas son *Bdellovibrio* y los organismos similares (conocidos por el acrónimo BALO por sus siglas en inglés), las cuales son pequeñas bacterias aerobias Gram negativas vibrioides, recientemente reclasificadas en la clase *Oligoflexia* que pertenece al filo *Pseudomonadota* (anteriormente conocido como *Proteobacteria*) [355]. Aunque las BALO se aislaron por primera vez del suelo, son ubicuas y pueden encontrarse, viables o por presencia de su material genético, en entornos acuáticos y terrestres, incluidos los sistemas hipersalinos [169], biopelículas [356], intestinos de mamíferos [170], [310], [357] y, como ha demostrado un grupo de investigación cercano, los pulmones de los pacientes con FQ [142]. Asimismo, existen evidencias de la presencia de material genético de bacterias depredadoras en recién nacidos prematuros (Capítulo 3 de esta tesis). Además de la detección de material genético de las BALO, varios autores han documentado el fenómeno *in vivo* de la depredación en los ecosistemas microbianos humanos [174], [175], [358], [359].

*P. aeruginosa* es la presa más estudiada de las BALO y tanto la depredadora como la presa suelen compartir un origen ambiental común. Este género es también uno de los patógenos humanos más importantes incluido en la categoría ESKAPE, que incluye *E. faecium*, *S. aureus*, *K. pneumoniae*, *A. baumannii*, *P. aeruginosa* y especies del género *Enterobacter*, por su alta prevalencia en infecciones invasivas graves como la bacteriemia y su capacidad para adquirir multirresistencia a los antibióticos [360]. *P. aeruginosa* es también el microorganismo más relevante y abundante implicado en la colonización pulmonar patogénica de la FQ, en particular los denominados clones de alto riesgo con multirresistencia a los antibióticos y los de morfotipo mucoide por hiperproducción de alginato. La caracterización genética y molecular de los aislados de *P. aeruginosa* procedentes de FQ [361] y de bacteriemias [362] han sido abordadas previamente por nuestro grupo, demostrando que los aislados de FQ están relacionados con un origen ambiental [363].

Como se ha comentado anteriormente, *S. marcescens* es una especie que ha sido infravalorada en la esfera clínica hasta mediados del siglo pasado, cuando comenzaron los primeros brotes nosocomiales [282]. A pesar de tener una virulencia relativamente

baja, causa infecciones graves en pacientes inmunocomprometidos, especialmente en las unidades de cuidados intensivos y, en particular, en las de neonatos (UCIN) [364]. El torrente sanguíneo, seguido del aparato respiratorio y del tracto gastrointestinal, son los sitios más frecuentes de infección por *S. marcescens* en la población de recién nacidos, causando una amplia variedad de manifestaciones clínicas [104].

## II. JUSTIFICACIÓN Y OBJETIVOS

Para evaluar la capacidad de las bacterias depredadoras para servir como terapia antibacteriana alternativa, nos planteamos una serie de ensayos *in vitro* para determinar la depredación de *B. bacteriovorus* en diversas colecciones de patógenos clínicos pertenecientes a las especies *P. aeruginosa* y *S. marcescens*. Este campo de investigación se encuentra en sus primeros pasos dentro del área clínica y ya existen algunos resultados prometedores probados tanto *in vitro* como en animales de experimentación, si bien aún no se han realizado estudios tan extendidos en cuanto a las preferencias de las depredadoras ni a la eficiencia de su depredación en colecciones amplias de patógenos clínicos [365], [366].

Los **OBJETIVOS** para este capítulo fueron:

1. Determinar el rango de depredación intraespecífico (I-PR) y la eficiencia de depredación (PE) de la cepa de referencia *B. bacteriovorus* 109J frente a una amplia colección de aislados clínicos de *P. aeruginosa* (procedentes de adultos con FQ o con bacteriemia) y de *S. marcescens* (procedentes de adultos y prematuros de bajo peso con bacteriemia) de diversos orígenes, antecedentes genéticos, morfotipos y perfiles de resistencia antibiótica.
2. Evaluar la posibilidad de aumentar el I-PR y/o la PE en un subgrupo de aislados clínicos de *P. aeruginosa* modificando el genoma de la bacteria depredadora, empleando como prueba de concepto la cepa mutante bd2637 de *B. bacteriovorus*.

### III. MATERIALES Y MÉTODOS

#### 1. Cepas y condiciones de crecimiento

La cepa de tipo salvaje *B. bacteriovorus* 109J (bd109J) se utilizó como depredadora contra 78 cepas de *P. aeruginosa* seleccionadas de colecciones previamente bien caracterizadas obtenidas de muestras respiratorias de pacientes con FQ (n=30) [361] y aislados de bacteriemia (BACT) (n=48) [362]. Las 78 cepas de *P. aeruginosa* se recogieron y caracterizaron en estudios anteriores, para los que se obtuvieron la aprobación ética y el consentimiento correspondientes [361], [362]. La selección de las presas se basó en los tipos de secuencia (ST) representativos determinados por MLST, tanto los ST frecuentes como los raros, su fenotipo de susceptibilidad a los antibióticos, incluidos los aislados multirresistentes, así como su morfología en los aislados de FQ, diferenciando los aislados mucoides (n=7) de los no mucoides (n=23). Las metodologías utilizadas para la caracterización molecular (ST, perfil de resistencia a los antibióticos, morfotipo) de las colecciones de *P. aeruginosa* se detallaron en las publicaciones anteriores [361], [362]. Por otro lado, se escogieron 13 aislados de *S. marcescens* (SM) de adultos (n=8) y neonatos (n=5) ingresados en UCI por bacteriemia o infecciones invasivas, previamente caracterizadas por nuestro grupo [367], [368].

La actividad de depredación del mutante de la PHA-depolimerasa *B. bacteriovorus* bd2637 [171] sólo se probó contra los 48 aislados de bacteriemia. Las mutaciones se inspeccionaron manualmente con Artemis [369] utilizando los números de acceso de Genbank SRX10641169 y SRX10641170 para *B. bacteriovorus* 109J y el mutante bd2637, respectivamente (**Tabla 15**).

**Tabla 15. Mutaciones cromosómicas acumuladas durante la delección génica.** La posición hace referencia a las coordenadas en el cromosoma. El genoma de referencia es el de *B. bacteriovorus* 109J (registrado en SRA con el identificador de Bioproyecto PRJNA723206). El efecto hace referencia a la consecuencia anotada de esta variante.

Gen	Producto génico	Tipo	Posición	Referencia	Alternativa	Efecto
<i>bd2637</i>	Depolimerasa PHA de cadena corta	Inserción	1.948.957	G	GAT	Cambio del marco de lectura
Intergénico	-	Delección	1.043.270	ACTTCTT	A	Sin sentido

Las depredadoras se cultivaron de forma rutinaria como se ha descrito anteriormente [370] mediante el cocultivo con *P. putida* KT2440 en lugar de *P. aeruginosa*, para evitar el sesgo que podría resultar de la posible influencia adaptativa en el crecimiento de un tipo particular de ST de *P. aeruginosa*. Se utilizó un caldo nutritivo diluido (DNB) (0,8 g/l de caldo nutritivo a pH 7,4) para recuperar las células depredadoras a partir de las existencias congeladas en glicerol y un tampón HEPES (25 mM a pH 7,8) para realizar los experimentos de depredación. Tanto el tampón DNB como el HEPES se complementaron con 2 mM de CaCl<sub>2</sub> · 2H<sub>2</sub>O y 3 mM de MgCl<sub>2</sub> · 3H<sub>2</sub>O. Para purificar la depredadora, los cocultivos se filtraron dos veces a través de un filtro de 0,45 µm. Finalmente, las cepas de *P. aeruginosa* se cultivaron en caldo LB durante 16 h a 37 °C y se diluyeron de nuevo en tampón HEPES hasta alcanzar una densidad óptica a 600 nm de 1 (DO<sub>600</sub>=1) para los posteriores experimentos de depredación.

## 2. Experimentos de depredación

Los cocultivos de presa (100 µL hasta la DO final<sub>600</sub> 0,3) y depredadora (100 µL) se realizaron en placas de 96 micropocillos durante 24 h a 30 °C con agitación orbital en un Synergy HTX (BioTek) con monitorización de la densidad celular (DO<sub>600</sub>) cada 15 min. Cada presa se ensayó en al menos 2-3 réplicas biológicas independientes, correspondiendo los resultados expresados a los valores medios de todos los experimentos. También se incluyó un pozo de control de crecimiento sin la depredadora para cada aislado de *P. aeruginosa*.

Para asegurar la homogeneidad del inóculo de la depredadora, cada experimento tenía un cocultivo de control paralelo de *B. bacteriovorus* (100 µL) y de *P. putida* (100 µL de 10<sup>9</sup> UFC/mL). Después de la incubación, se determinó la supervivencia de las presas sembrando diluciones seriadas en placas de agar, descartando aquellos

conjuntos en los que la disminución de la viabilidad de las presas tras la depredación del control de *P. putida* era diferente de 3 y 4 logaritmos para las depredadoras de tipo salvaje y bd2637, respectivamente, como se ha descrito previamente [171], [371].

Así, los experimentos de depredación incluyeron un control al principio del experimento tanto del inóculo de la depredadora como de la presa (recuentos viables correspondientes a la  $DO_{600}=0,3$ ). El recuento de células viables de la presa se realizó a las 0 y a las 24 h de la depredación. Diez microlitros de la dilución correspondiente del cocultivo (incluyendo el control sin depredadora y el control de depredación con *P. putida*) se sembraron en medio sólido LB y se contaron las unidades formadoras de colonias (UFC/mL) tras 24 h de incubación a 37 °C.

La actividad de depredación de una depredadora sobre una colección de presas se definió mediante dos parámetros: i) el rango I-PR, que corresponde a la proporción de cepas de presas susceptibles dentro de una colección de la misma especie, y ii) la PE, específica para cada combinación presa/depredadora, que se determina como la relación entre las células viables de la presa tras la exposición de la depredadora en comparación con el inóculo inicial. La PE se expresó en valores  $\log_{10}$ . Sólo se consideró que la depredación era positiva cuando el PE era  $>0,5$ , lo que correspondía a la depredación del 50% de la población ( $p<0,046$ ).

### 3. Análisis estadísticos

Se empleó la prueba de Dunn para evaluar las diferencias en las medianas de los valores de PE entre las tres colecciones de aislados que sirvieron de presas para la depredadora bd109J, después de asumir la distribución no normal de los datos con la prueba de Shapiro-Wilk. Por otro lado, se empleó la prueba no paramétrica de suma de rangos de Wilcoxon pareada para comparar los valores de PE de la colección BACT en función del tipo de depredadora que se empleó (bd109J vs. bd2637), ya que se trataba de dos series de datos relacionados. Además, se realizaron asociaciones entre las características cualitativas de las colecciones de *Pseudomonas* y la variable cuantitativa PE mediante regresión lineal simple.

Las diferencias de los valores I-PR con bd109J entre las colecciones se exploraron mediante la prueba de Chi-cuadrado. Las diferencias de los valores I-PR en la colección BACT (bd109J vs. bd2637) se exploraron mediante la prueba de McNemar emparejada. Todos los conjuntos de datos se analizaron utilizando la versión 3.5.0 de R, y los gráficos se realizaron utilizando la versión 2.2.1 de *ggplot* y el software RStudio v.1.2.5001.

## IV. RESULTADOS

### 1. La fuente clínica de la presa influye en la actividad de depredación de *B. bacteriovorus* 109J

Se observaron diferencias significativas entre el I-PR de las tres colecciones de presas ( $p=5,2^{-05}$ ), siendo *S. marcescens* la que mayores valores obtuvo (13/13, 100%), seguida de la colección FQ (20/30, 67%) y la colección BACT (17/48, 35,4%) (**Figuras 47-49**).

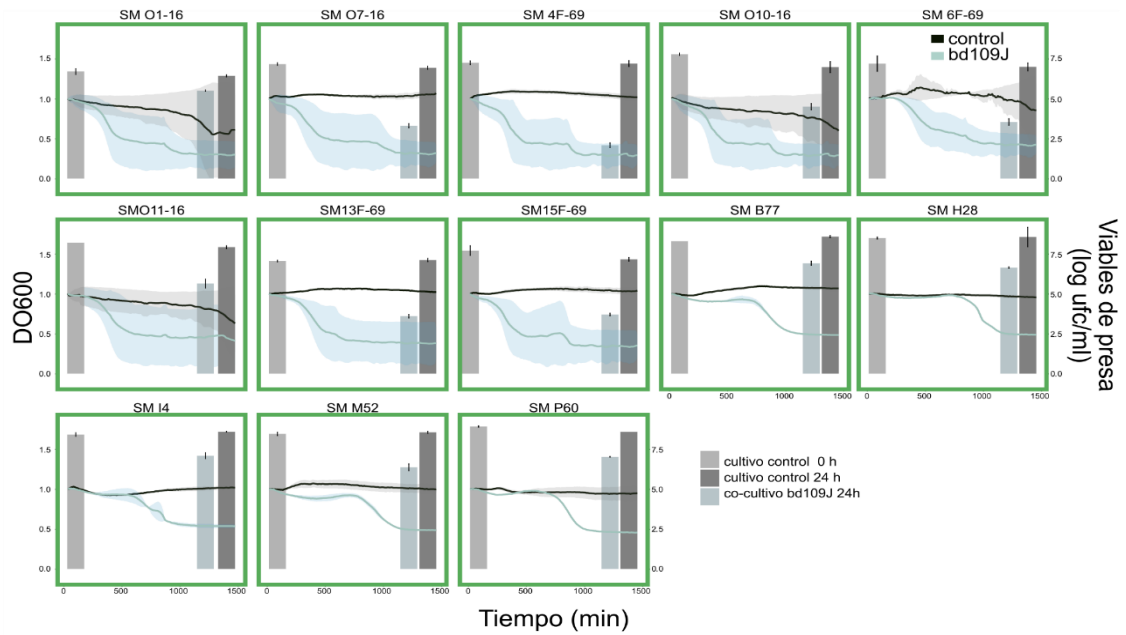
La asociación entre la actividad depredadora y el origen de las colecciones de *P. aeruginosa* también se evaluó mediante regresión lineal (lm,  $p<0,0001$ ). No se observó correlación entre la actividad de depredación y el linaje genético (según el tipo de secuencia o ST obtenido por MLST) de *P. aeruginosa* ( $p>0,05$ ), con discrepancias en 6 de los 12 ST que agrupaban a más de un aislado de ambas colecciones [ST175 (1 presa susceptible de un total de 3), ST253 (2/3), ST274 (1/2), ST532 (1/2), ST646 (2/3) y ST1017 (1/2)]. Tampoco se detectó correlación con el fenotipo mucóide (4/7 aislados mucóides susceptibles) o no mucóide de los aislados de FQ (16/23 aislados no mucóides susceptibles) ( $p>0,05$ ) ni con el fenotipo de susceptibilidad a los antibióticos ( $p>0,05$ ) (**Figura 50A**).

Además, la mediana de PE fue significativamente mayor en los aislados de FQ (mediana=3,91) y de *S. marcescens* (mediana=2,2) respecto a la colección BACT (mediana=0,23) (prueba de Dunn  $p<0,0001$ ). Entre *S. marcescens* y FQ no se encontraron diferencias significativas (**Figura 51A**).

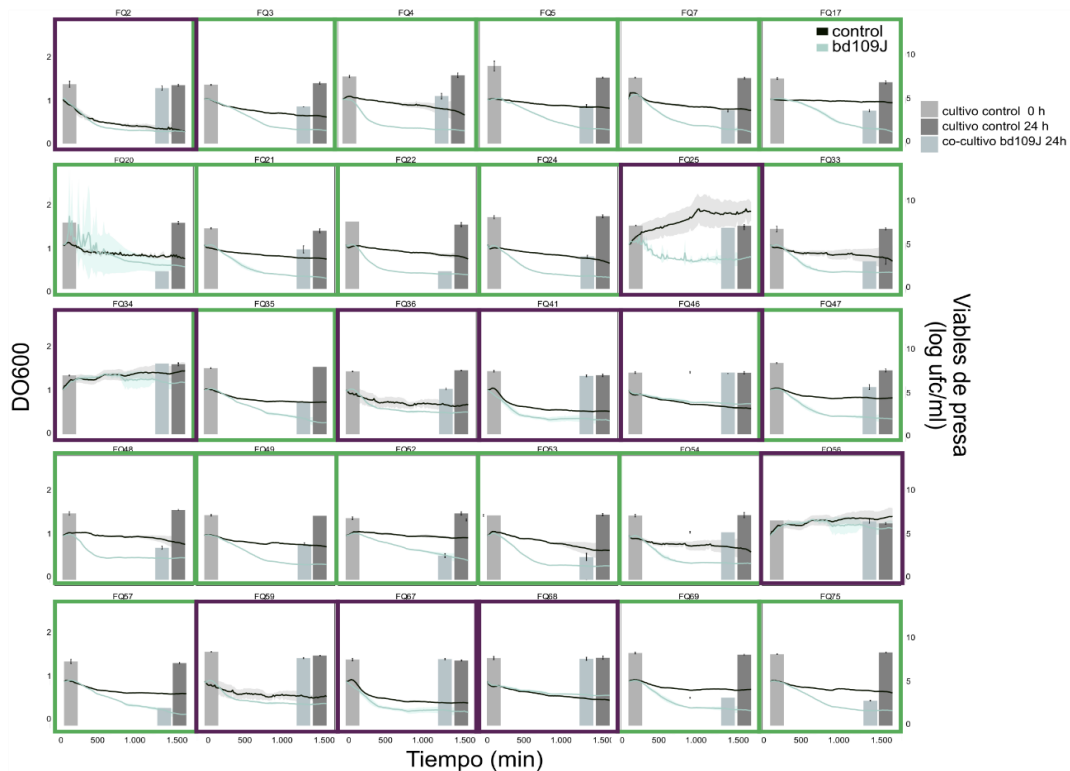
### 2. Influencia de la variación genética de la depredadora

La capacidad de depredación de la cepa de tipo salvaje de *B. bacteriovorus* 109J se comparó con la del mutante de un solo gen bd2637 en los 48 aislados de la colección BACT, la colección menos susceptible. No hubo diferencias significativas entre los valores de I-PR de la cepa mutante y la de tipo salvaje (21/48, 43,8% frente a 17/48, 35,4%, respectivamente  $p>0,05$ ) (**Figuras 52**). Curiosamente, el patrón de depredación fue diferente entre ambas depredadoras: 16 aislados (33,3%) fueron resistentes y 6 (12,5%) susceptibles a ambas depredadoras, mientras que 15 (31,2%) fueron susceptibles sólo al mutante bd2637 y, finalmente, 11 aislados (22,9%) fueron susceptibles sólo a la depredadora de tipo salvaje (prueba McNemar  $p>0,05$ ) (**Figura 50B**). Finalmente, tampoco se detectaron diferencias significativas en los valores globales de la mediana de PE (0,23 para *B. bacteriovorus* 109J frente a 0,29 para el mutante bd2637,  $p>0,05$ ) (**Figura 51A**).

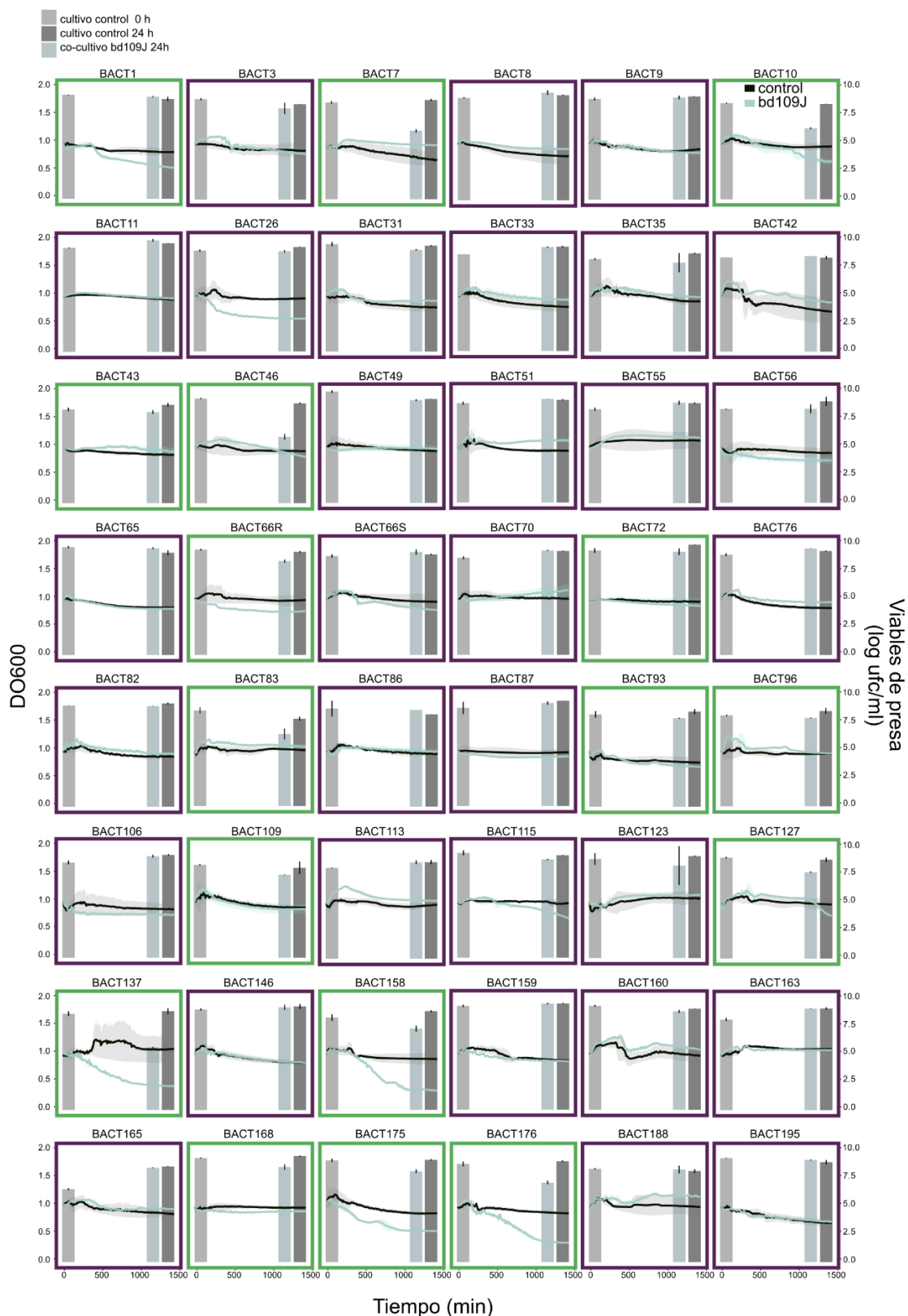
No se observó correlación entre la depredación del mutante *bd2637* y la susceptibilidad a los antibióticos o cualquier otra característica de la presa ( $p > 0,05$ ) (**Figura 50B**). Estos resultados se observaron sistemáticamente en las réplicas de cada aislado y no se asociaron a ninguna característica de presa en particular. Así pues, sólo la supresión de la actividad catalítica del gen *bd2637* y los efectos asociados fueron los responsables de los cambios en la susceptibilidad de las presas.



**Figura 47. Curvas de depredación de bd109J en la colección *S. marcescens* y células viables antes y después del experimento. Recuadro verde: susceptibilidad a la depredación ( $PE \geq 0,5$ ).**

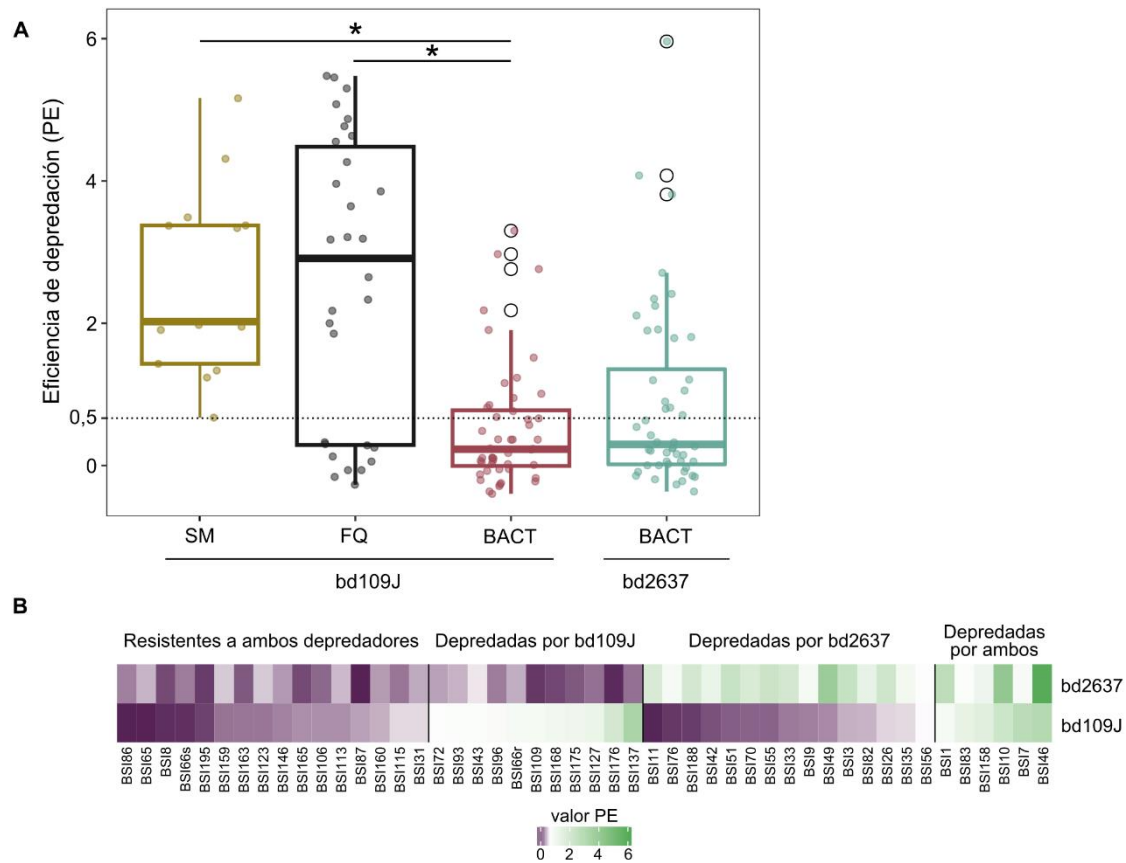


**Figura 48. Curvas de depredación de bd109J en la colección *P. aeruginosa* FQ y células viables antes y después del experimento. Recuadro verde: susceptibilidad a la depredación ( $PE \geq 0,5$ ). Recuadro morado: resistencia a la depredación ( $PE < 0,5$ ).**

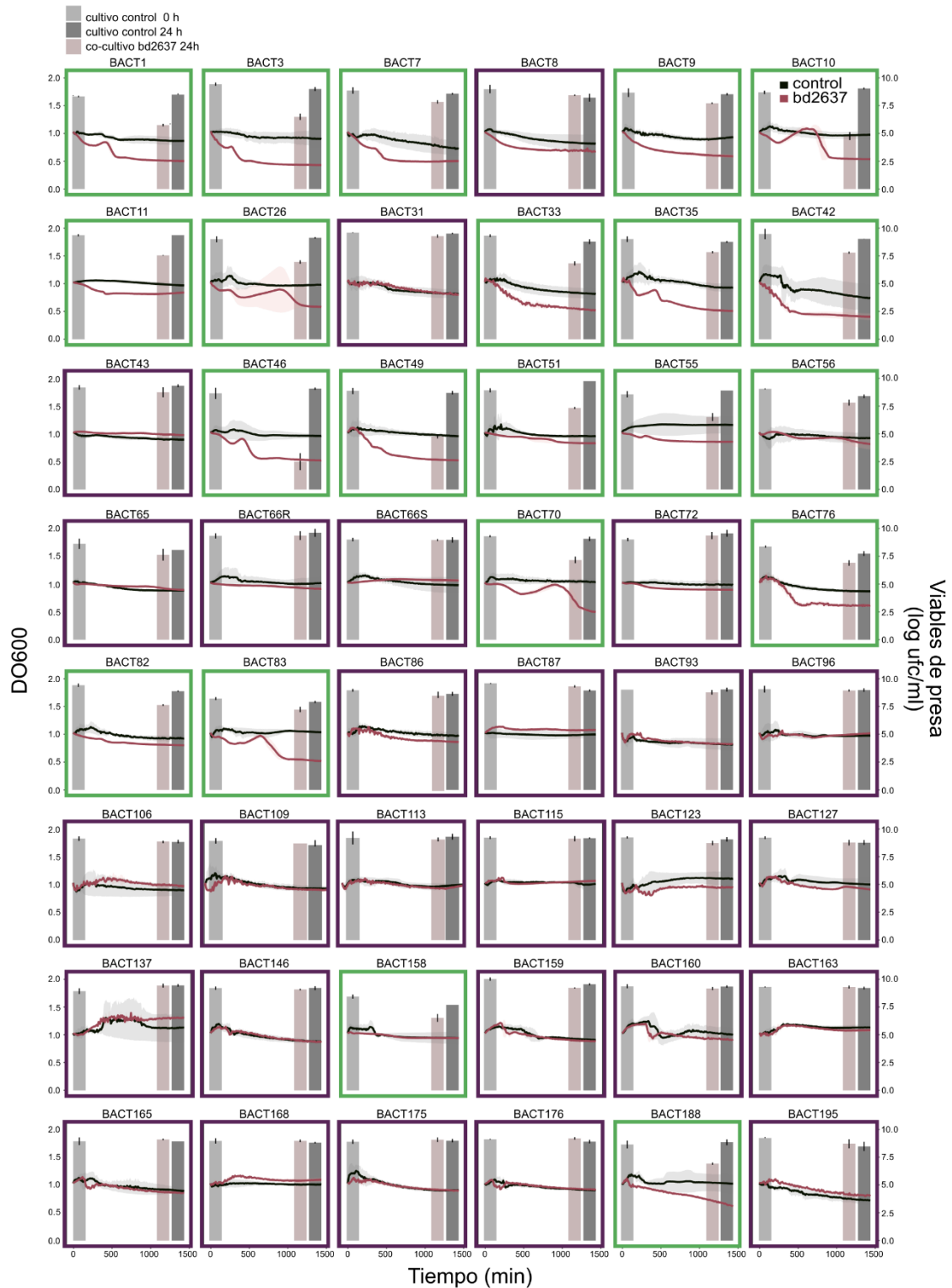


**Figura 49.** Curvas de depredación de bd109J en la colección *P. aeruginosa* BACT y células viables antes y después del experimento. Recuadro verde: susceptibilidad a la depredación ( $PE \geq 0,5$ ). Recuadro morado: resistencia a la depredación ( $PE < 0,5$ ).





**Figura 51. Susceptibilidad a la depredación de las colecciones de presas clínicas.** **A:** PE de bd109J en las colecciones *S. marcescens*, FQ y BACT ( $p < 0,0001$ ) y PE del mutante bd2637 en la colección BACT ( $p > 0,05$ ). **B:** perfil de los valores de PE de bd109J y de bd2637 en la colección BACT. La susceptibilidad a la depredación se consideró cuando  $PE > 0,5$  (línea discontinua). Cada cepa de depredadora tiene un patrón específico de depredación. Círculos blancos: *outliers*.



**Figura 52. Curvas de depredación de bd2637 en la colección *P. aeruginosa* BACT y células viables antes y después del experimento. Recuadro verde: susceptibilidad a la depredación ( $PE \geq 0,5$ ). Recuadro morado: resistencia a la depredación ( $PE < 0,5$ ).**

## V. DISCUSIÓN

El uso de las BALO como agentes de control biológico [357], [372] se ha sugerido en base a su falta de interacción con las células humanas [366], [373], [374]. Al igual que ocurre con los antibióticos, la comprobación de la susceptibilidad individual *in vitro* para las parejas de cepas de presa y depredadora es un requisito, principalmente cuando la depredación podría verse afectada sustancialmente por las condiciones ambientales o biológicas, como se postula en este trabajo. Las depredadoras se han estudiado principalmente en condiciones controladas (medio de cultivo, pH y temperatura) y los conocimientos sobre la susceptibilidad a la depredación de los aislados clínicos siguen siendo muy limitados.

Por un lado, se quiso incluir una colección de *S. marcescens* procedente de pacientes que sufrían de septicemia, tanto adultos como neonatos prematuros ingresados en UCIN, por tratarse de casos de gran vulnerabilidad y con cierta prevalencia en nuestro hospital. Una fracción de estas cepas fue estudiada en capítulos anteriores de esta tesis doctoral como parte de dos brotes de carácter policlonal, uno histórico y otro contemporáneo. Sorprendentemente, el 100% de los aislados fue depredado por la cepa salvaje bd109J, la cual mostró, además, una eficiencia relativamente alta (mediana PE=2,2). En estudios anteriores se había notificado la habilidad de esta bacteria depredadora frente a *S. marcescens*, así como a otras bacterias patógenas similares [164], aunque la gran mayoría de los estudios empleaban cepas de colección de cultivos de laboratorio en lugar de aislados clínicos [375]. Sin embargo, en experimentos *in vitro* como el presente, a menudo se muestra una depredación exitosa que luego puede no corresponderse en experimentos *in vivo*. Se ha apuntado que sería necesaria la intervención tanto de la bacteria depredadora como del sistema inmunitario del hospedador para que la erradicación de la infección sea efectiva [376]. Aun así, estos resultados son prometedores para el futuro de la intervención terapéutica contra bacterias multirresistentes, muy adaptadas al entorno hospitalario y con facilidad de propagación, como es el caso de *S. marcescens*.

Por otro lado, esta investigación se ha centrado en la especie *P. aeruginosa*, ya que se ha informado de que esta especie tiene una capacidad limitada para servir de presa a *B. bacteriovorus* [365] y, por tanto, se esperaba una mayor diversidad en la respuesta depredadora-presa en comparación con la representante Gram negativa *S. marcescens*. Además, *P. aeruginosa* es ubicua en la naturaleza (donde se espera que se produzcan la

mayoría de las interacciones entre *Bdellovibrio* y *Pseudomonas*, además de un importante patógeno humano que frecuentemente alberga mecanismos de resistencia a los antibióticos. Por último, los estudios anteriores de nuestro grupo facilitaron la creación de una colección completa y bien caracterizada que incluye linajes frecuentes e infrecuentes de FQ [361] y causantes de bacteriemia [362] de acuerdo con su ST. Hay que tener en cuenta que aislados de *P. aeruginosa* pertenecientes a un mismo ST pueden encontrarse en diferentes hábitats, como nichos ambientales y nosocomiales, pero algunos linajes parecen estar muy adaptados a entornos clínicos concretos. Este es el caso de la colonización crónica de las vías respiratorias de los pacientes con FQ, donde los linajes caracterizados se encuentran próximos a cepas de origen ambiental, como se ha demostrado anteriormente [363]. En este estudio, no se encontró una correlación entre ST y la I-PR.

Un resultado notable de esta investigación es la susceptibilidad cualitativa (I-PR, proporción de cepas depredadas en las colecciones) y cuantitativa (PE, proporción de células depredadas por cepa individual) significativamente mayor de los aislados de FQ en comparación con los de bacteriemia. Además, no existió correlación entre la depredación y la morfología de las colonias (mucoide frente a no mucoide), según informes anteriores [377], ni entre los patrones de susceptibilidad a los antibióticos o, incluso, entre los linajes genéticos de ST. No podemos excluir que las diferencias de depredación en el mismo linaje de ST puedan derivarse de otras características genéticas, como la presencia de islas de patogenicidad o de profagos.

No podemos descartar la existencia de otras depredadoras altamente especializadas en la microbiota humana, ya que todos los datos disponibles se han obtenido de depredadoras ambientales. Sin embargo, es probable que los ecosistemas microbianos humanos tengan las mismas reglas de control poblacional. Las depredadoras de origen humano podrían ser más adecuadas para limitar los patógenos humanos bien adaptados [142] a pesar de que, en un ecosistema complejo y diverso, la preferencia por una presa concreta sería una característica dinámica y cambiante [375]. Aunque los experimentos suelen realizarse con linajes individuales, el uso de poblaciones mixtas debería ser un objetivo futuro para validar la especificidad de las presas en una comunidad y las consecuencias en la estructura de la población microbiana. Este trabajo indica que la depredación cuantitativa, determinada como PE relativa, podría diferir entre los miembros de una población aparentemente homogénea y podría ser fundamental para entender la dinámica de conjuntos bacterianos compuestos por diferentes presas y

depredadoras.

Sólo se han introducido en las bases de datos públicas 8 genomas de *B. bacteriovorus*, todos ellos de origen ambiental (suelo, ríos, agua de mar, aguas residuales), por lo que la contribución del componente genético de las depredadoras en los valores de PE es una cuestión pendiente. Para dilucidar esto, se utilizó el mutante bd2637 para probar la posibilidad de diseñar depredadoras con un I-PR y una PE más eficientes y amplios [171] frente a los aislados de bacteriemia. Se demostró que este mutante tiene un espectro diferente al de la cepa de tipo salvaje. Sin embargo, ¿cómo puede ocurrir eso? El genotipo de la cepa bd2637 corresponde a una única delección en el gen *bd2637*, que codifica una depolimerasa de PHA responsable de la degradación de los biopolímeros. Hipotéticamente, este mutante debería tener menos acceso a los ricos recursos nutricionales del polihidroxicanoato (PHA) y, en consecuencia, un menor *fitness*. Sin embargo, la realidad es probablemente más compleja. Dentro del "metabolismo común", cuando la depredadora se coloca en la presa, la degradación masiva de PHA produce un exceso de ATP que aumenta las especies reactivas de oxígeno que pueden oxidar y dañar el ADN, los carbohidratos, las proteínas o los lípidos [371], [378]. El exceso de ATP también desencadena la respiración pero, en las células pequeñas (como *Bdellovibrio*), la cuestión es si los altos niveles de ATP son incompatibles con el mantenimiento de la respiración si la superficie de la célula no puede asignar suficientes proteínas respiratorias [379], [380]. En resumen, los mutantes defectuosos de PHA podrían tener consecuencias metabólicas inesperadas [381]–[385] que influyen en la actividad de depredación y la investigación en este campo podría ayudar a identificar y, eventualmente, superar los factores de resistencia a la depredación en *P. aeruginosa*. Cabe destacar que el alcance de la presa (rango intra o interespecífico) no depende sólo de la susceptibilidad de la presa, sino también de la especificidad de la depredadora, lo que pone de manifiesto la importancia de las pruebas de susceptibilidad específicas entre depredadora y presa, ya que su interacción y coevolución pueden superar la depredación y la resistencia, respectivamente [386], [387]. Una perspectiva prometedora para futuros trabajos es tratar de ver si la especificidad presa-depredadora y la tasa de depredación pueden diferir cuando las depredadoras se propagan a costa de especies o clones bacterianos aislados de nichos estrechamente relacionados o diferentes. La evolución de la depredadora en comunidades bacterianas en coevolución podría dar lugar a una evolución genómica acelerada de la depredadora y de la presa [388].

Los determinantes de la presa o de la depredadora responsables de la especificidad de la depredación no han sido completamente dilucidados. Sin embargo, el arsenal hidrolítico de *B. bacteriovorus* desempeña un papel crucial en la actividad de depredación, dado que determina el éxito de su ciclo vital [171], [389]. La caracterización del espectro de presas específico de cada depredadora es un requisito para el uso clínico de estas bacterias como agentes antimicrobianos vivos. Los estudios han informado de que diferentes linajes de BALO y depredadoras aisladas de diferentes nichos tienen diferentes espectros de presas [390], [391]. Por lo tanto, la selección de la depredadora apropiada para fines clínicos requeriría estudios más amplios y profundos sobre la depredación para superar la resistencia a la depredación en especies de presas específicas y cepas individuales dentro de la especie.

## VI. DIVULGACIÓN

La parte de este capítulo referida a la depredación de *B. bacteriovorus* sobre *P. aeruginosa*, específica para cada cepa, con un rango de depredación mayor para los aislados de fibrosis quística que para los de bacteriemia, se ha publicado recientemente en la revista indexada *Scientific Reports* bajo el título *Strain-specific predation of Bdellovibrio bacteriovorus on Pseudomonas aeruginosa with a higher range for cystic fibrosis than for bacteremia isolates* (DOI: [10.1038/s41598-022-14378-5](https://doi.org/10.1038/s41598-022-14378-5)).

# **CONCLUSIONES**



## CONCLUSIONES

**Primera.** La secuenciación de amplicones y la metaproteómica proporcionan resultados concordantes de la composición de la microbiota, más de forma cualitativa que cuantitativa, detectándose únicamente una discordancia en grupos parafiléticos por las reclasificaciones taxonómicas. La proteómica adolece de desactualización de las bases de datos y presenta limitaciones para el cálculo de la diversidad alfa. Su principal ventaja es la detección de proteínas del huésped en respuesta a la colonización bacteriana.

**Segunda.** La metodología condiciona el resultado de la detección de microorganismos; existen disparidades entre cultivos microbiológicos rutinarios, cultivos extensivos con pre-enriquecimiento y la secuenciación de amplicones.

**Tercera.** La implantación de la microbiota intestinal de los recién nacidos con fibrosis quística está marcada por un enriquecimiento de *Ruminococcus gnavus* dentro de un contexto de respuesta inflamatoria humana, mientras que la microbiota respiratoria presenta patrones complejos de cocolonización, donde predominan anaerobias facultativas a pesar de la primocolonización patogénica temprana.

**Cuarta.** La variabilidad microbiológica inter-individuo es elevada; cada persona posee unos linajes genéticos bacterianos propios que, en ocasiones, son compartidos por el ecosistema intestinal y el respiratorio.

**Quinta.** El contexto epidemiológico de una institución sanitaria condiciona el establecimiento de la microbiota intestinal al nacer. Durante el brote de *Serratia marcescens* se constató la predominancia de *Proteobacteria* en la microbiota intestinal, mientras que en un contexto sin brote suele estar en una proporción minoritaria.

**Sexta.** Todos los ecosistemas microbiológicos están conectados: el enriquecimiento intestinal de *S. marcescens* es el primer paso de la sepsis. Un ecosistema intestinal saludable se caracteriza por la adquisición progresiva de una diversidad de especies que están equilibradamente distribuidas.

**Séptima.** Se detecta ADN de *Serratia* y de bacterias depredadoras en meconio, continuando la controversia sobre el inicio de la colonización bacteriana durante la gestación y el nacimiento.

**Octava.** Se ha descrito por primera vez el pangenoma de *S. marcescens*, lo que ha puesto de manifiesto el carácter ambiental de este género con abundancia de elementos genéticos móviles. Los aislados actuales de *S. marcescens* son diferentes de los que producían las infecciones hace 50 años, si bien existe una especialización de las cepas clínicas adaptadas al ambiente hospitalario respecto de las cepas ambientales.

**Novena.** El rango de depredación de *Bdellovibrio bacteriovorus* es mayor en cepas de *S. marcescens*, así como en cepas de *Pseudomonas aeruginosa* aisladas del tracto respiratorio de sujetos con fibrosis quística respecto a aislados de *P. aeruginosa* causantes de bacteriemia. Este hecho parece estar condicionado por el origen ambiental de las cepas, sin que la depredación esté condicionada por el morfotipo, linaje genético o perfil de sensibilidad a los antibióticos de la presa.

**Décima.** Existe una especificidad presa-depredadora, por lo que el uso de esta arma biológica debe estar condicionada a la realización de pruebas *in vitro* que aseguren el éxito de cada combinación.

# **BIBLIOGRAFÍA**



## BIBLIOGRAFÍA

- [1] R. Vemuri *et al.*, “Gut Microbial Changes, Interactions, and Their Implications on Human Lifecycle: An Ageing Perspective,” *Biomed Res Int*, vol. 2018, pp. 1–13, 2018, doi: 10.1155/2018/4178607.
- [2] T. Requena and M. Velasco, “The human microbiome in sickness and in health,” *Revista Clínica Española (English Edition)*, vol. 221, no. 4, pp. 233–240, Apr. 2021, doi: 10.1016/j.rceng.2019.07.018.
- [3] M. G. Dominguez-Bello, F. Godoy-Vitorino, R. Knight, and M. J. Blaser, “Role of the microbiome in human development,” *Gut*, vol. 68, no. 6, pp. 1108–1114, Jun. 2019, doi: 10.1136/gutjnl-2018-317503.
- [4] R. Vemuri, E. M. Shankar, M. Chieppa, R. Eri, and K. Kavanagh, “Beyond Just Bacteria: Functional Biomes in the Gut Ecosystem Including Virome, Mycobiome, Archaeome and Helminths,” *Microorganisms*, vol. 8, no. 4, p. 483, Mar. 2020, doi: 10.3390/microorganisms8040483.
- [5] C. E. Huseyin, P. W. O’Toole, P. D. Cotter, and P. D. Scanlan, “Forgotten fungi—the gut mycobiome in human health and disease,” *FEMS Microbiol Rev*, vol. 41, no. 4, pp. 479–511, Jul. 2017, doi: 10.1093/femsre/fuw047.
- [6] G. Borrel, J.-F. Brugère, S. Gribaldo, R. A. Schmitz, and C. Moissl-Eichinger, “The host-associated archaeome,” *Nat Rev Microbiol*, vol. 18, no. 11, pp. 622–636, Nov. 2020, doi: 10.1038/s41579-020-0407-y.
- [7] A. Chikina and D. Matic Vignjevic, “At the right time in the right place: How do luminal gradients position the microbiota along the gut?,” *Cells & Development*, vol. 168, p. 203712, Dec. 2021, doi: 10.1016/j.cdev.2021.203712.
- [8] I. Sekirov, S. L. Russell, L. C. M. Antunes, and B. B. Finlay, “Gut Microbiota in Health and Disease,” *Physiol Rev*, vol. 90, no. 3, pp. 859–904, Jul. 2010, doi: 10.1152/physrev.00045.2009.
- [9] L. Mancabelli *et al.*, “Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations,” *Environ Microbiol*, vol. 19, no. 4, pp. 1379–1390, Apr. 2017, doi: 10.1111/1462-2920.13692.
- [10] A. Moya and M. Ferrer, “Functional Redundancy-Induced Stability of Gut Microbiota Subjected to Disturbance,” *Trends Microbiol*, vol. 24, no. 5, pp. 402–413, May 2016, doi: 10.1016/j.tim.2016.02.002.
- [11] I. Semin, J. Ninnemann, M. Bondareva, I. Gimaev, and A. A. Kruglov, “Interplay Between Microbiota, Toll-Like Receptors and Cytokines for the Maintenance of Epithelial Barrier Integrity,” *Front Med (Lausanne)*, vol. 8, May 2021, doi: 10.3389/fmed.2021.644333.
- [12] I. Rowland *et al.*, “Gut microbiota functions: metabolism of nutrients and other food components,” *Eur J Nutr*, vol. 57, no. 1, pp. 1–24, Feb. 2018, doi: 10.1007/s00394-017-1445-8.
- [13] D. J. Morrison and T. Preston, “Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism,” *Gut Microbes*, vol. 7, no. 3, pp. 189–200, May 2016, doi:

- 10.1080/19490976.2015.1134082.
- [14] S. Magnúsdóttir, D. Ravcheev, V. de Crécy-Lagard, and I. Thiele, “Systematic genome assessment of B-vitamin biosynthesis suggests cooperation among gut microbes,” *Front Genet*, vol. 6, Apr. 2015, doi: 10.3389/fgene.2015.00148.
- [15] M. Parvaneh, G. Karimi, R. Jamaluddin, A. M. H. NG, Z. Ibrahim, and S. I. Muhammad, “*Lactobacillus helveticus* (ATCC 27558) upregulates Runx2 and Bmp2 and modulates bone mineral density in ovariectomy-induced bone loss rats,” *Clin Interv Aging*, vol. Volume 13, pp. 1555–1564, Aug. 2018, doi: 10.2147/CIA.S169223.
- [16] Y. Sugiyama *et al.*, “Gut bacterial aromatic amine production: aromatic amino acid decarboxylase and its effects on peripheral serotonin production,” *Gut Microbes*, vol. 14, no. 1, Dec. 2022, doi: 10.1080/19490976.2022.2128605.
- [17] Y. Wan and T. Zuo, “Interplays between drugs and the gut microbiome,” *Gastroenterol Rep (Oxf)*, vol. 10, Jan. 2022, doi: 10.1093/gastro/goac009.
- [18] J. Doroszkiewicz, M. Groblewska, and B. Mroczko, “The Role of Gut Microbiota and Gut–Brain Interplay in Selected Diseases of the Central Nervous System,” *Int J Mol Sci*, vol. 22, no. 18, p. 10028, Sep. 2021, doi: 10.3390/ijms221810028.
- [19] P. Strandwitz, “Neurotransmitter modulation by the gut microbiota,” *Brain Res*, vol. 1693, pp. 128–133, Aug. 2018, doi: 10.1016/j.brainres.2018.03.015.
- [20] M. Carabotti, A. Scirocco, M. A. Maselli, and C. Sever, “The gut-brain axis: Interactions between enteric microbiota, central and enteric nervous systems,” *Ann. Gastroenterol.*, vol. 28, pp. 203–209, 2015.
- [21] W. Barcik, R. C. T. Boutin, M. Sokolowska, and B. B. Finlay, “The Role of Lung and Gut Microbiota in the Pathology of Asthma,” *Immunity*, vol. 52, no. 2, pp. 241–255, Feb. 2020, doi: 10.1016/j.immuni.2020.01.007.
- [22] M. Torcia, “Interplay among Vaginal Microbiome, Immune Response and Sexually Transmitted Viral Infections,” *Int J Mol Sci*, vol. 20, no. 2, p. 266, Jan. 2019, doi: 10.3390/ijms20020266.
- [23] D. Pistone *et al.*, “A Journey on the Skin Microbiome: Pitfalls and Opportunities,” *Int J Mol Sci*, vol. 22, no. 18, p. 9846, Sep. 2021, doi: 10.3390/ijms22189846.
- [24] F. Baquero *et al.*, “Epidermis as a Platform for Bacterial Transmission,” *Front Immunol*, vol. 12, Dec. 2021, doi: 10.3389/fimmu.2021.774018.
- [25] E. Rosenberg and I. Zilber-Rosenberg, “Microbiotas are Transmitted Between Holobiont Generations,” in *The Hologenome Concept: Human, Animal and Plant Microbiota*, Cham: Springer International Publishing, 2013, pp. 41–54. doi: 10.1007/978-3-319-04241-1\_4.
- [26] I. Chattopadhyay *et al.*, “Exploring the Role of Gut Microbiome in Colon Cancer,” *Appl Biochem Biotechnol*, vol. 193, no. 6, pp. 1780–1799, Jun. 2021, doi: 10.1007/s12010-021-03498-9.
- [27] F. Kallinowski *et al.*, “Prevalence of enteropathogenic bacteria in surgically treated chronic inflammatory

- bowel disease.,” *Hepatogastroenterology*, vol. 45, no. 23, pp. 1552–8.
- [28] D. Rojo *et al.*, “Exploring the human microbiome from multiple perspectives: factors altering its composition and function,” *FEMS Microbiol Rev*, vol. 41, no. 4, pp. 453–478, Jul. 2017, doi: 10.1093/femsre/fuw046.
- [29] J. Libertucci and V. B. Young, “The role of the microbiota in infectious diseases,” *Nat Microbiol*, vol. 4, no. 1, pp. 35–45, Jan. 2019, doi: 10.1038/s41564-018-0278-4.
- [30] P. Luczynski, K.-A. McVey Neufeld, C. S. Oriach, G. Clarke, T. G. Dinan, and J. F. Cryan, “Growing up in a Bubble: Using Germ-Free Animals to Assess the Influence of the Gut Microbiota on Brain and Behavior,” *International Journal of Neuropsychopharmacology*, vol. 19, no. 8, p. pyw020, Aug. 2016, doi: 10.1093/ijnp/pyw020.
- [31] M. A. Sze *et al.*, “Changes in the Bacterial Microbiota in Gut, Blood, and Lungs following Acute LPS Instillation into Mice Lungs,” *PLoS One*, vol. 9, no. 10, p. e111228, Oct. 2014, doi: 10.1371/journal.pone.0111228.
- [32] W. H. W. Tang, D. Y. Li, and S. L. Hazen, “Dietary metabolism, the gut microbiome, and heart failure,” *Nat Rev Cardiol*, vol. 16, no. 3, pp. 137–154, Mar. 2019, doi: 10.1038/s41569-018-0108-7.
- [33] Y. Jiao, L. Wu, N. D. Huntington, and X. Zhang, “Crosstalk Between Gut Microbiota and Innate Immunity and Its Implication in Autoimmune Diseases,” *Front Immunol*, vol. 11, Feb. 2020, doi: 10.3389/fimmu.2020.00282.
- [34] T. Vatanen *et al.*, “Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans,” *Cell*, vol. 165, no. 4, pp. 842–853, May 2016, doi: 10.1016/j.cell.2016.04.007.
- [35] X. Zhang, L. Li, J. Butcher, A. Stintzi, and D. Figeys, “Advancing functional and translational microbiome research using meta-omics approaches,” *Microbiome*, vol. 7, no. 1, p. 154, Dec. 2019, doi: 10.1186/s40168-019-0767-6.
- [36] S. Quainoo *et al.*, “Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis,” *Clin Microbiol Rev*, vol. 30, no. 4, pp. 1015–1063, Oct. 2017, doi: 10.1128/CMR.00016-17.
- [37] A. van Belkum *et al.*, “Guidelines for the validation and application of typing methods for use in bacterial epidemiology,” *Clinical Microbiology and Infection*, vol. 13, pp. 1–46, 2007, doi: 10.1111/j.1469-0691.2007.01786.x.
- [38] P. Seng *et al.*, “Ongoing Revolution in Bacteriology: Routine Identification of Bacteria by Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry,” *Clinical Infectious Diseases*, vol. 49, no. 4, pp. 543–551, Aug. 2009, doi: 10.1086/600885.
- [39] E. G. Parizad, “The Application of Pulsed Field Gel Electrophoresis in Clinical Studies,” *JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH*, 2016, doi: 10.7860/JCDR/2016/15718.7043.
- [40] M. C. J. Maiden *et al.*, “MLST revisited: the gene-by-gene approach to

- bacterial genomics,” *Nat Rev Microbiol*, vol. 11, no. 10, pp. 728–736, Oct. 2013, doi: 10.1038/nrmicro3093.
- [41] D. Jiang *et al.*, “Microbiome Multi-Omics Network Analysis: Statistical Considerations, Limitations, and Opportunities,” *Front Genet*, vol. 10, Nov. 2019, doi: 10.3389/fgene.2019.00995.
- [42] “The Integrative Human Microbiome Project,” *Nature*, vol. 569, no. 7758, pp. 641–648, May 2019, doi: 10.1038/s41586-019-1238-8.
- [43] M. A. Malla, A. Dubey, A. Kumar, S. Yadav, A. Hashem, and E. F. Abd Allah, “Exploring the Human Microbiome: The Potential Future Role of Next-Generation Sequencing in Disease Diagnosis and Treatment,” *Front Immunol*, vol. 9, Jan. 2019, doi: 10.3389/fimmu.2018.02868.
- [44] D. Sims, I. Sudbery, N. E. Ilott, A. Heger, and C. P. Ponting, “Sequencing depth and coverage: key considerations in genomic analyses,” *Nat Rev Genet*, vol. 15, no. 2, pp. 121–132, Feb. 2014, doi: 10.1038/nrg3642.
- [45] K. Korpela and W. M. de Vos, “Infant gut microbiota restoration: state of the art,” *Gut Microbes*, vol. 14, no. 1, Dec. 2022, doi: 10.1080/19490976.2022.2118811.
- [46] A. M. Fricker, D. Podlesny, and W. F. Fricke, “What is new and relevant for sequencing-based microbiome research? A mini-review,” *J Adv Res*, vol. 19, pp. 105–112, Sep. 2019, doi: 10.1016/j.jare.2019.03.006.
- [47] K. E. Nelson *et al.*, “A Catalog of Reference Genomes from the Human Microbiome,” *Science (1979)*, vol. 328, no. 5981, pp. 994–999, May 2010, doi: 10.1126/science.1183605.
- [48] J. Peterson *et al.*, “The NIH Human Microbiome Project,” *Genome Res*, vol. 19, no. 12, pp. 2317–2323, Dec. 2009, doi: 10.1101/gr.096651.109.
- [49] R. Ranjan, A. Rani, A. Metwally, H. S. McGee, and D. L. Perkins, “Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing,” *Biochem Biophys Res Commun*, vol. 469, no. 4, pp. 967–977, Jan. 2016, doi: 10.1016/j.bbrc.2015.12.083.
- [50] Y. W. Lim *et al.*, “Clinical Insights from Metagenomic Analysis of Sputum Samples from Patients with Cystic Fibrosis,” *J Clin Microbiol*, vol. 52, no. 2, pp. 425–437, Feb. 2014, doi: 10.1128/JCM.02204-13.
- [51] C. Quast *et al.*, “The SILVA ribosomal RNA gene database project: improved data processing and web-based tools,” *Nucleic Acids Res*, vol. 41, no. D1, pp. D590–D596, Nov. 2012, doi: 10.1093/nar/gks1219.
- [52] J. R. Cole *et al.*, “Ribosomal Database Project: data and tools for high throughput rRNA analysis,” *Nucleic Acids Res*, vol. 42, no. D1, pp. D633–D642, Jan. 2014, doi: 10.1093/nar/gkt1244.
- [53] T. Z. DeSantis *et al.*, “Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB,” *Appl Environ Microbiol*, vol. 72, no. 7, pp. 5069–5072, Jul. 2006, doi: 10.1128/AEM.03006-05.
- [54] C.-C. Liao *et al.*, “MetaSquare: an integrated metadatabase of 16S rRNA gene amplicon for microbiome

- taxonomic classification,” *Bioinformatics*, vol. 38, no. 10, pp. 2930–2931, May 2022, doi: 10.1093/bioinformatics/btac184.
- [55] S. Agnihotry, A. Sarangi, and R. Aggarwal, “Construction & assessment of a unified curated reference database for improving the taxonomic classification of bacteria using 16S rRNA sequence data,” *Indian Journal of Medical Research*, vol. 151, no. 1, p. 93, 2020, doi: 10.4103/ijmr.IJMR\_220\_18.
- [56] N. Segata *et al.*, “Metagenomic biomarker discovery and explanation,” *Genome Biol*, vol. 12, no. 6, p. R60, 2011, doi: 10.1186/gb-2011-12-6-r60.
- [57] J. G. Caporaso *et al.*, “QIIME allows analysis of high-throughput community sequencing data,” *Nat Methods*, vol. 7, no. 5, pp. 335–336, May 2010, doi: 10.1038/nmeth.f.303.
- [58] S. Graspentner, N. Loeper, S. Künzel, J. F. Baines, and J. Rupp, “Selection of validated hypervariable regions is crucial in 16S-based microbiota studies of the female genital tract,” *Sci Rep*, vol. 8, no. 1, p. 9678, Dec. 2018, doi: 10.1038/s41598-018-27757-8.
- [59] J.-C. Lagier *et al.*, “Culturing the human microbiota and culturomics,” *Nat Rev Microbiol*, vol. 16, no. 9, pp. 540–550, Sep. 2018, doi: 10.1038/s41579-018-0041-0.
- [60] J.-C. Lagier *et al.*, “Culture of previously uncultured members of the human gut microbiota by culturomics,” *Nat Microbiol*, vol. 1, no. 12, p. 16203, Dec. 2016, doi: 10.1038/nmicrobiol.2016.203.
- [61] P. Hardouin, R. Chiron, H. Marchandin, J. Armengaud, and L. Grenga, “Metaproteomics to Decipher CF Host-Microbiota Interactions: Overview, Challenges and Future Perspectives,” *Genes (Basel)*, vol. 12, no. 6, p. 892, Jun. 2021, doi: 10.3390/genes12060892.
- [62] S. Bashiardes, G. Zilberman-Schapira, and E. Elinav, “Use of Metatranscriptomics in Microbiome Research,” *Bioinform Biol Insights*, vol. 10, p. BBLS34610, Jan. 2016, doi: 10.4137/BBLS34610.
- [63] K. Cheng *et al.*, “MetaLab 2.0 Enables Accurate Post-Translational Modifications Profiling in Metaproteomics,” *J Am Soc Mass Spectrom*, vol. 31, no. 7, pp. 1473–1482, Jul. 2020, doi: 10.1021/jasms.0c00083.
- [64] Y. Wang, Y. Zhou, X. Xiao, J. Zheng, and H. Zhou, “Metaproteomics: A strategy to study the taxonomy and functionality of the gut microbiota,” *J Proteomics*, vol. 219, p. 103737, May 2020, doi: 10.1016/j.jprot.2020.103737.
- [65] J. Lee and D. Banerjee, “Metabolomics and the Microbiome as Biomarkers in Sepsis,” *Crit Care Clin*, vol. 36, no. 1, pp. 105–113, Jan. 2020, doi: 10.1016/j.ccc.2019.08.008.
- [66] X. Liu and J. W. Locasale, “Metabolomics: A Primer,” *Trends Biochem Sci*, vol. 42, no. 4, pp. 274–284, Apr. 2017, doi: 10.1016/j.tibs.2017.01.004.
- [67] D. Podlesny and W. F. Fricke, “Strain inheritance and neonatal gut microbiota development: A meta-analysis,” *International Journal of Medical Microbiology*, vol. 311, no. 3, p. 151483, Apr. 2021, doi: 10.1016/j.ijmm.2021.151483.

- [68] E. B. Hollister *et al.*, “Structure and function of the healthy pre-adolescent pediatric gut microbiome,” *Microbiome*, vol. 3, no. 1, p. 36, Dec. 2015, doi: 10.1186/s40168-015-0101-x.
- [69] A. Sarkar, J. Y. Yoo, S. Valeria Ozorio Dutra, K. H. Morgan, and M. Groer, “The Association between Early-Life Gut Microbiota and Long-Term Health and Diseases,” *J Clin Med*, vol. 10, no. 3, p. 459, Jan. 2021, doi: 10.3390/jcm10030459.
- [70] C. Milani *et al.*, “The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota,” *Microbiology and Molecular Biology Reviews*, vol. 81, no. 4, Dec. 2017, doi: 10.1128/MMBR.00036-17.
- [71] K. M. Kennedy *et al.*, “Fetal meconium does not have a detectable microbiota before birth,” *Nat Microbiol*, vol. 6, no. 7, pp. 865–873, Jul. 2021, doi: 10.1038/s41564-021-00904-0.
- [72] K. Korpela *et al.*, “Selective maternal seeding and environment shape the human gut microbiome,” *Genome Res*, vol. 28, no. 4, pp. 561–568, Apr. 2018, doi: 10.1101/gr.233940.117.
- [73] M. C. Collado, E. Isolauri, K. Laitinen, and S. Salminen, “Effect of mother’s weight on infant’s microbiota acquisition, composition, and activity during early infancy: a prospective follow-up study initiated in early pregnancy,” *Am J Clin Nutr*, vol. 92, no. 5, pp. 1023–1030, Nov. 2010, doi: 10.3945/ajcn.2010.29877.
- [74] D. M. Chu *et al.*, “The early infant gut microbiome varies in association with a maternal high-fat diet,” *Genome Med*, vol. 8, no. 1, p. 77, Dec. 2016, doi: 10.1186/s13073-016-0330-z.
- [75] T. K. Soderborg *et al.*, “Gestational Diabetes Is Uniquely Associated With Altered Early Seeding of the Infant Gut Microbiota,” *Front Endocrinol (Lausanne)*, vol. 11, Nov. 2020, doi: 10.3389/fendo.2020.603021.
- [76] M. W. Groer, A. A. Luciano, L. J. Dishaw, T. L. Ashmeade, E. Miller, and J. A. Gilbert, “Development of the preterm infant gut microbiome: a research priority,” *Microbiome*, vol. 2, no. 1, p. 38, Dec. 2014, doi: 10.1186/2049-2618-2-38.
- [77] D. M. Chu, J. Ma, A. L. Prince, K. M. Antony, M. D. Seferovic, and K. M. Aagaard, “Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery,” *Nat Med*, vol. 23, no. 3, pp. 314–326, Mar. 2017, doi: 10.1038/nm.4272.
- [78] E. Rutayisire, K. Huang, Y. Liu, and F. Tao, “The mode of delivery affects the diversity and colonization pattern of the gut microbiota during the first year of infants’ life: a systematic review,” *BMC Gastroenterol*, vol. 16, no. 1, p. 86, Dec. 2016, doi: 10.1186/s12876-016-0498-0.
- [79] B. Brooks *et al.*, “Microbes in the neonatal intensive care unit resemble those found in the gut of premature infants,” *Microbiome*, vol. 2, no. 1, p. 1, Dec. 2014, doi: 10.1186/2049-2618-2-1.
- [80] D. H. Taft *et al.*, “Bacterial colonization and antimicrobial resistance genes in neonatal enteral feeding tubes,” *FEMS Microbiol Ecol*, vol. 95, no. 4, Apr. 2019, doi: 10.1093/femsec/fiz039.

- [81] N. A. Bokulich, D. A. Mills, and M. A. Underwood, "Surface Microbes in the Neonatal Intensive Care Unit: Changes with Routine Cleaning and over Time," *J Clin Microbiol*, vol. 51, no. 8, pp. 2617–2624, Aug. 2013, doi: 10.1128/JCM.00898-13.
- [82] J. M. Rodríguez, "The Origin of Human Milk Bacteria: Is There a Bacterial Entero-Mammary Pathway during Late Pregnancy and Lactation?," *Advances in Nutrition*, vol. 5, no. 6, pp. 779–784, Nov. 2014, doi: 10.3945/an.114.007229.
- [83] R. E. Moore and S. D. Townsend, "Temporal development of the infant gut microbiome," *Open Biol*, vol. 9, no. 9, p. 190128, Sep. 2019, doi: 10.1098/rsob.190128.
- [84] E. C. Davis, A. M. Dinsmoor, M. Wang, and S. M. Donovan, "Microbiome Composition in Pediatric Populations from Birth to Adolescence: Impact of Diet and Prebiotic and Probiotic Interventions," *Dig Dis Sci*, vol. 65, no. 3, pp. 706–722, Mar. 2020, doi: 10.1007/s10620-020-06092-x.
- [85] E. Nissilä *et al.*, "C4B gene influences intestinal microbiota through complement activation in patients with paediatric-onset inflammatory bowel disease," *Clin Exp Immunol*, vol. 190, no. 3, pp. 394–405, Nov. 2017, doi: 10.1111/cei.13040.
- [86] T. A. Laguna *et al.*, "Airway Microbiota in Bronchoalveolar Lavage Fluid from Clinically Well Infants with Cystic Fibrosis," *PLoS One*, vol. 11, no. 12, p. e0167649, Dec. 2016, doi: 10.1371/journal.pone.0167649.
- [87] A. Pulliero, D. Traversi, E. Franchitti, M. Barchitta, A. Izzotti, and A. Agodi, "The Interaction among Microbiota, Epigenetic Regulation, and Air Pollutants in Disease Prevention," *J Pers Med*, vol. 12, no. 1, p. 14, Dec. 2021, doi: 10.3390/jpm12010014.
- [88] K. Andrews and A. Gonzalez, "Contextual risk factors impacting the colonization and development of the intestinal microbiota: Implications for children in low- and middle-income countries," *Dev Psychobiol*, vol. 61, no. 5, pp. 714–728, Jul. 2019, doi: 10.1002/dev.21823.
- [89] J. L. Sonnenburg and E. D. Sonnenburg, "Vulnerability of the industrialized microbiota," *Science (1979)*, vol. 366, no. 6464, Oct. 2019, doi: 10.1126/science.aaw9255.
- [90] L. F. Stinson, "Establishment of the early-life microbiome: a DOHaD perspective," *J Dev Orig Health Dis*, vol. 11, no. 3, pp. 201–210, Jun. 2020, doi: 10.1017/S2040174419000588.
- [91] C. Michel and H. M. Blottière, "Neonatal Programming of Microbiota Composition: A Plausible Idea That Is Not Supported by the Evidence," *Front Microbiol*, vol. 13, Jun. 2022, doi: 10.3389/fmicb.2022.825942.
- [92] P. Vangay *et al.*, "US Immigration Westernizes the Human Gut Microbiome," *Cell*, vol. 175, no. 4, pp. 962–972.e10, Nov. 2018, doi: 10.1016/j.cell.2018.10.029.
- [93] A. R. Jha *et al.*, "Gut microbiome transition across a lifestyle gradient in Himalaya," *PLoS Biol*, vol. 16, no. 11, p. e2005396, Nov. 2018, doi: 10.1371/journal.pbio.2005396.
- [94] H. Blencowe *et al.*, "Born Too Soon: The global epidemiology of 15

- million preterm births,” *Reprod Health*, vol. 10, no. S1, p. S2, Nov. 2013, doi: 10.1186/1742-4755-10-S1-S2.
- [95] C. J. Kim, R. Romero, P. Chaemsaitong, N. Chaiyasit, B. H. Yoon, and Y. M. Kim, “Acute chorioamnionitis and funisitis: definition, pathologic features, and clinical significance,” *Am J Obstet Gynecol*, vol. 213, no. 4, pp. S29–S52, Oct. 2015, doi: 10.1016/j.ajog.2015.08.040.
- [96] J. Challis, J. Newnham, F. Petraglia, M. Yeganegi, and A. Bocking, “Fetal sex and preterm birth,” *Placenta*, vol. 34, no. 2, pp. 95–99, 2013, doi: 10.1016/j.placenta.2012.11.007.
- [97] G. Yamey, H. Horváth, L. Schmidt, J. Myers, and C. D. Brindis, “Reducing the global burden of Preterm Birth through knowledge transfer and exchange: a research agenda for engaging effectively with policymakers,” *Reprod Health*, vol. 13, no. 1, p. 26, Dec. 2016, doi: 10.1186/s12978-016-0146-8.
- [98] H. Blencowe *et al.*, “National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications,” *The Lancet*, vol. 379, no. 9832, pp. 2162–2172, Jun. 2012, doi: 10.1016/S0140-6736(12)60820-4.
- [99] M. García-Reymundo, X. Demestre, M. J. Calvo, G. Ginovart, A. Jiménez, and J. A. Hurtado, “Prematuro tardío en España: experiencia del Grupo SEN34-36,” *An Pediatr (Engl Ed)*, vol. 88, no. 5, pp. 246–252, May 2018, doi: 10.1016/j.anpedi.2017.05.006.
- [100] P. S. Sutton and G. L. Darmstadt, “Preterm Birth and Neurodevelopment: A Review of Outcomes and Recommendations for Early Identification and Cost-effective Interventions,” *J Trop Pediatr*, vol. 59, no. 4, pp. 258–265, Aug. 2013, doi: 10.1093/tropej/fmt012.
- [101] J. K.-F. Lee, L. T. Hern Tan, A. Ramadas, N.-S. Ab Mutalib, and L.-H. Lee, “Exploring the Role of Gut Bacteria in Health and Disease in Preterm Neonates,” *Int J Environ Res Public Health*, vol. 17, no. 19, p. 6963, Sep. 2020, doi: 10.3390/ijerph17196963.
- [102] M. A. Carl *et al.*, “Sepsis From the Gut: The Enteric Habitat of Bacteria That Cause Late-Onset Neonatal Bloodstream Infections,” *Clinical Infectious Diseases*, vol. 58, no. 9, pp. 1211–1218, May 2014, doi: 10.1093/cid/ciu084.
- [103] P. S. X. Yap *et al.*, “Neonatal intensive care unit (NICU) exposures exert a sustained influence on the progression of gut microbiota and metabolome in the first year of life,” *Sci Rep*, vol. 11, no. 1, p. 1353, Dec. 2021, doi: 10.1038/s41598-020-80278-1.
- [104] P. Gastmeier *et al.*, “Outbreaks in neonatal intensive care units—They are not like others,” *Am J Infect Control*, vol. 35, no. 3, pp. 172–176, Apr. 2007, doi: 10.1016/j.ajic.2006.07.007.
- [105] J. Johnson and C. Quach, “Outbreaks in the neonatal ICU,” *Curr Opin Infect Dis*, vol. 30, no. 4, pp. 395–403, Aug. 2017, doi: 10.1097/QCO.0000000000000383.
- [106] L. Moles *et al.*, “Bacterial Diversity in Meconium of Preterm Neonates and Evolution of Their Fecal Microbiota during the First Month of

- Life,” *PLoS One*, vol. 8, no. 6, p. e66986, Jun. 2013, doi: 10.1371/journal.pone.0066986.
- [107] A. C. Masi and C. J. Stewart, “The role of the preterm intestinal microbiome in sepsis and necrotising enterocolitis,” *Early Hum Dev*, vol. 138, p. 104854, Nov. 2019, doi: 10.1016/j.earlhumdev.2019.104854.
- [108] S. Arboleya *et al.*, “Establishment and development of intestinal microbiota in preterm neonates,” *FEMS Microbiol Ecol*, vol. 79, no. 3, pp. 763–772, Mar. 2012, doi: 10.1111/j.1574-6941.2011.01261.x.
- [109] K. Korpela *et al.*, “Intestinal microbiota development and gestational age in preterm neonates,” *Sci Rep*, vol. 8, no. 1, p. 2453, Dec. 2018, doi: 10.1038/s41598-018-20827-x.
- [110] S. M. Rowe, S. Miller, and E. J. Sorscher, “Cystic Fibrosis,” *New England Journal of Medicine*, vol. 352, no. 19, pp. 1992–2001, May 2005, doi: 10.1056/NEJMra043184.
- [111] K. de Boeck and M. D. Amaral, “Progress in therapies for cystic fibrosis,” *Lancet Respir Med*, vol. 4, no. 8, pp. 662–674, Aug. 2016, doi: 10.1016/S2213-2600(16)00023-0.
- [112] S. Rueda-Nieto *et al.*, “Analysis of the genotypic profile and its relationship with the clinical manifestations in people with cystic fibrosis: study from a rare disease registry,” *Orphanet J Rare Dis*, vol. 17, no. 1, p. 222, Dec. 2022, doi: 10.1186/s13023-022-02373-y.
- [113] G. Mehta, M. Macek, and A. Mehta, “Cystic fibrosis across Europe: EuroCareCF analysis of demographic data from 35 countries,” *Journal of Cystic Fibrosis*, vol. 9, pp. S5–S21, Dec. 2010, doi: 10.1016/j.jcf.2010.08.002.
- [114] A. Orenti, A. Zolin, A. Jung, and J. van Rens, “ECFSR Annual Report 2020,” 2022.
- [115] J. S. Elborn *et al.*, “Report of the European Respiratory Society/European Cystic Fibrosis Society task force on the care of adults with cystic fibrosis,” *European Respiratory Journal*, vol. 47, no. 2, pp. 420–428, Feb. 2016, doi: 10.1183/13993003.00592-2015.
- [116] “Registro Español de Fibrosis Quística: Informe anual 2016,” 2018.
- [117] L. S. Hanssens, J. Duchateau, and G. J. Casimir, “CFTR Protein: Not Just a Chloride Channel?,” *Cells*, vol. 10, no. 11, p. 2844, Oct. 2021, doi: 10.3390/cells10112844.
- [118] A. A. Pezzulo *et al.*, “Reduced airway surface pH impairs bacterial killing in the porcine cystic fibrosis lung,” *Nature*, vol. 487, no. 7405, pp. 109–113, Jul. 2012, doi: 10.1038/nature11130.
- [119] J. K. Gustafsson *et al.*, “Bicarbonate and functional CFTR channel are required for proper mucin secretion and link cystic fibrosis with its mucus phenotype,” *Journal of Experimental Medicine*, vol. 209, no. 7, pp. 1263–1272, Jul. 2012, doi: 10.1084/jem.20120562.
- [120] M. Lévêque, S. le Trionnaire, P. del Porto, and C. Martin-Chouly, “The impact of impaired macrophage functions in cystic fibrosis disease progression,” *Journal of Cystic Fibrosis*, vol. 16, no. 4, pp. 443–453, Jul. 2017, doi: 10.1016/j.jcf.2016.10.011.

- [121] B. P. O'Sullivan and S. D. Freedman, "Cystic fibrosis," *The Lancet*, vol. 373, no. 9678, pp. 1891–1904, May 2009, doi: 10.1016/S0140-6736(09)60327-5.
- [122] R. Cantón *et al.*, "Consenso español para la prevención y el tratamiento de la infección bronquial por *Pseudomonas aeruginosa* en el paciente con fibrosis quística," *Arch Bronconeumol*, vol. 51, no. 3, pp. 140–150, Mar. 2015, doi: 10.1016/j.arbres.2014.09.021.
- [123] E. Hatziaorou *et al.*, "Changing epidemiology of the respiratory bacteriology of patients with cystic fibrosis—data from the European cystic fibrosis society patient registry," *Journal of Cystic Fibrosis*, vol. 19, no. 3, pp. 376–383, May 2020, doi: 10.1016/j.jcf.2019.08.006.
- [124] C. H. Goss, "Acute Pulmonary Exacerbations in Cystic Fibrosis," *Semin Respir Crit Care Med*, vol. 40, no. 06, pp. 792–803, Dec. 2019, doi: 10.1055/s-0039-1697975.
- [125] R. M. Girón Moreno, M. García-Clemente, L. Diab-Cáceres, A. Martínez-Vergara, M. Á. Martínez-García, and R. M. Gómez-Punter, "Treatment of Pulmonary Disease of Cystic Fibrosis: A Comprehensive Review," *Antibiotics*, vol. 10, no. 5, p. 486, Apr. 2021, doi: 10.3390/antibiotics10050486.
- [126] M. Wilschanski and I. Novak, "The Cystic Fibrosis of Exocrine Pancreas," *Cold Spring Harb Perspect Med*, vol. 3, no. 5, pp. a009746–a009746, May 2013, doi: 10.1101/cshperspect.a009746.
- [127] M. Sathe and R. Houwen, "Meconium ileus in Cystic Fibrosis," *Journal of Cystic Fibrosis*, vol. 16, pp. S32–S39, Nov. 2017, doi: 10.1016/j.jcf.2017.06.007.
- [128] M. B. Toledano *et al.*, "The emerging burden of liver disease in cystic fibrosis patients: A UK nationwide study," *PLoS One*, vol. 14, no. 4, p. e0212779, Apr. 2019, doi: 10.1371/journal.pone.0212779.
- [129] L. S. Mott *et al.*, "Progression of early structural lung disease in young children with cystic fibrosis assessed using CT," *Thorax*, vol. 67, no. 6, pp. 509–516, Jun. 2012, doi: 10.1136/thoraxjnl-2011-200912.
- [130] S. C. Ranganathan, G. L. Hall, P. D. Sly, S. M. Stick, and T. A. Douglas, "Early Lung Disease in Infants and Preschool Children with Cystic Fibrosis. What Have We Learned and What Should We Do about It?," *Am J Respir Crit Care Med*, vol. 195, no. 12, pp. 1567–1575, Jun. 2017, doi: 10.1164/rccm.201606-1107CI.
- [131] M. Shteinberg, I. J. Haq, D. Polineni, and J. C. Davies, "Cystic fibrosis," *The Lancet*, vol. 397, no. 10290, pp. 2195–2211, Jun. 2021, doi: 10.1016/S0140-6736(20)32542-3.
- [132] G. G. Hernández, "Cribado neonatal de fibrosis quística," *Anales de Pediatría Continuada*, vol. 12, no. 1, pp. 34–38, Jan. 2014, doi: 10.1016/S1696-2818(14)70165-9.
- [133] N. L. Turcios, "Cystic Fibrosis Lung Disease: An Overview," *Respir Care*, vol. 65, no. 2, pp. 233–251, Feb. 2020, doi: 10.4187/respcare.06697.
- [134] D. S. Armstrong *et al.*, "Lower Airway Inflammation in Infants with Cystic Fibrosis Detected by Newborn

- Screening,” *Pediatr Pulmonol*, vol. 40, no. 6, pp. 500–510, Dec. 2005, doi: 10.1002/ppul.20294.
- [135] B. Ahmed *et al.*, “Longitudinal development of the airway microbiota in infants with cystic fibrosis,” *Sci Rep*, vol. 9, no. 1, p. 5143, Dec. 2019, doi: 10.1038/s41598-019-41597-0.
- [136] N. Acosta *et al.*, “Sputum microbiota is predictive of long-term clinical outcomes in young adults with cystic fibrosis,” *Thorax*, vol. 73, no. 11, pp. 1016–1025, Nov. 2018, doi: 10.1136/thoraxjnl-2018-211510.
- [137] D. Conrad, M. Haynes, P. Salamon, P. B. Rainey, M. Youle, and F. Rohwer, “Cystic Fibrosis Therapy: A Community Ecology Perspective,” *Am J Respir Cell Mol Biol*, vol. 48, no. 2, pp. 150–156, Feb. 2013, doi: 10.1165/rcmb.2012-0059PS.
- [138] R. A. Quinn *et al.*, “Ecological networking of cystic fibrosis lung infections,” *NPJ Biofilms Microbiomes*, vol. 2, no. 1, p. 4, Dec. 2016, doi: 10.1038/s41522-016-0002-1.
- [139] J. S. Brown, K. L. Zeman, and W. D. Bennett, “Regional Deposition of Coarse Particles and Ventilation Distribution in Healthy Subjects and Patients with Cystic Fibrosis,” *Journal of Aerosol Medicine*, vol. 14, no. 4, pp. 443–454, Dec. 2001, doi: 10.1089/08942680152744659.
- [140] M. Dmitrijeva *et al.*, “Strain-Resolved Dynamics of the Lung Microbiome in Patients with Cystic Fibrosis,” *mBio*, vol. 12, no. 2, Apr. 2021, doi: 10.1128/mBio.02863-20.
- [141] D. Willner *et al.*, “Case Studies of the Spatial Heterogeneity of DNA Viruses in the Cystic Fibrosis Lung,” *Am J Respir Cell Mol Biol*, vol. 46, no. 2, pp. 127–131, Feb. 2012, doi: 10.1165/rcmb.2011-0253OC.
- [142] J. de Dios Caballero *et al.*, “Individual Patterns of Complexity in Cystic Fibrosis Lung Microbiota, Including Predator Bacteria, over a 1-Year Period,” *mBio*, vol. 8, no. 5, Nov. 2017, doi: 10.1128/mBio.00959-17.
- [143] R. Enaud *et al.*, “The Gut-Lung Axis in Health and Respiratory Diseases: A Place for Inter-Organ and Inter-Kingdom Crosstalks,” *Front Cell Infect Microbiol*, vol. 10, Feb. 2020, doi: 10.3389/fcimb.2020.00009.
- [144] G. Debyser *et al.*, “Faecal proteomics: A tool to investigate dysbiosis and inflammation in patients with cystic fibrosis,” *Journal of Cystic Fibrosis*, vol. 15, no. 2, pp. 242–250, Mar. 2016, doi: 10.1016/j.jcf.2015.08.003.
- [145] P. Vonaesch, M. Anderson, and P. J. Sansonetti, “Pathogens, microbiome and the host: emergence of the ecological Koch’s postulates,” *FEMS Microbiol Rev*, vol. 42, no. 3, pp. 273–292, May 2018, doi: 10.1093/femsre/fuy003.
- [146] A. L. Byrd and J. A. Segre, “Adapting Koch’s postulates,” *Science (1979)*, vol. 351, no. 6270, pp. 224–226, Jan. 2016, doi: 10.1126/science.aad6753.
- [147] E. J. Stevens, K. A. Bates, and K. C. King, “Host microbiota can facilitate pathogen infection,” *PLoS Pathogens*, vol. 17, no. 5. Public Library of Science, May 01, 2021. doi: 10.1371/journal.ppat.1009514.
- [148] J. Stripling and M. Rodriguez,

- “Current Evidence in Delivery and Therapeutic Uses of Fecal Microbiota Transplantation in Human Diseases—Clostridium difficile Disease and Beyond,” *Am J Med Sci*, vol. 356, no. 5, pp. 424–432, Nov. 2018, doi: 10.1016/j.amjms.2018.08.010.
- [149] M. Ponce-Alonso, S. García-Fernández, and R. del Campo, “Challenges of landscaping and modulation strategies in human intestinal microbiota,” *Enferm Infecc Microbiol Clin*, vol. 39, no. 4, pp. 165–167, Apr. 2021, doi: 10.1016/j.eimc.2021.02.001.
- [150] Y. Xiao, M. T. Angulo, S. Lao, S. T. Weiss, and Y.-Y. Liu, “An ecological framework to understand the efficacy of fecal microbiota transplantation,” *Nat Commun*, vol. 11, no. 1, p. 3329, Dec. 2020, doi: 10.1038/s41467-020-17180-x.
- [151] P. Belvoncikova, M. Maronek, and R. Gardlik, “Gut Dysbiosis and Fecal Microbiota Transplantation in Autoimmune Diseases,” *Int J Mol Sci*, vol. 23, no. 18, p. 10729, Sep. 2022, doi: 10.3390/ijms231810729.
- [152] T. S. B. Schmidt *et al.*, “Drivers and determinants of strain dynamics following fecal microbiota transplantation,” *Nat Med*, vol. 28, no. 9, pp. 1902–1912, Sep. 2022, doi: 10.1038/s41591-022-01913-0.
- [153] S. J. Ott *et al.*, “Efficacy of Sterile Fecal Filtrate Transfer for Treating Patients With Clostridium difficile Infection,” *Gastroenterology*, vol. 152, no. 4, pp. 799–811.e7, Mar. 2017, doi: 10.1053/j.gastro.2016.11.010.
- [154] A. York, “FMT in the clinic,” *Nat Rev Microbiol*, vol. 17, no. 3, pp. 127–127, Mar. 2019, doi: 10.1038/s41579-019-0157-x.
- [155] T. S. Dhillon, E. K. Dhillon, H. C. Chau, W. K. Li, and A. H. Tsang, “Studies on bacteriophage distribution: virulent and temperate bacteriophage content of mammalian feces,” *Appl Environ Microbiol*, vol. 32, no. 1, pp. 68–74, Jul. 1976, doi: 10.1128/aem.32.1.68-74.1976.
- [156] K. M. Wylie, G. M. Weinstock, and G. A. Storch, “Emerging view of the human virome,” *Translational Research*, vol. 160, no. 4, pp. 283–290, Oct. 2012, doi: 10.1016/j.trsl.2012.03.006.
- [157] P. Manrique, B. Bolduc, S. T. Walk, J. van der Oost, W. M. de Vos, and M. J. Young, “Healthy human gut phageome,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 37, pp. 10400–10405, Sep. 2016, doi: 10.1073/pnas.1601060113.
- [158] S. R. Modi, H. H. Lee, C. S. Spina, and J. J. Collins, “Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome,” *Nature*, vol. 499, no. 7457, pp. 219–222, Jul. 2013, doi: 10.1038/nature12212.
- [159] K. E. Kortright, B. K. Chan, J. L. Koff, and P. E. Turner, “Phage Therapy: A Renewed Approach to Combat Antibiotic-Resistant Bacteria,” *Cell Host Microbe*, vol. 25, no. 2, pp. 219–232, Feb. 2019, doi: 10.1016/j.chom.2019.01.014.
- [160] J.-Y. Liu *et al.*, “Decolonization of carbapenem-resistant Klebsiella pneumoniae from the intestinal microbiota of model mice by phages targeting two surface structures,” *Front Microbiol*, vol. 13, Aug. 2022, doi: 10.3389/fmicb.2022.877074.

- [161] E. M. Waters *et al.*, “Phage therapy is highly effective against chronic lung infections with *Pseudomonas aeruginosa*,” *Thorax*, vol. 72, no. 7, pp. 666–667, Jul. 2017, doi: 10.1136/thoraxjnl-2016-209265.
- [162] C. Buttmer *et al.*, “Impact of a phage cocktail targeting *Escherichia coli* and *Enterococcus faecalis* as members of a gut bacterial consortium in vitro and in vivo,” *Front Microbiol*, vol. 13, Jul. 2022, doi: 10.3389/fmicb.2022.936083.
- [163] K. M. Caflisch, G. A. Suh, and R. Patel, “Biological challenges of phage therapy and proposed solutions: a literature review,” *Expert Rev Anti Infect Ther*, vol. 17, no. 12, pp. 1011–1041, Dec. 2019, doi: 10.1080/14787210.2019.1694905.
- [164] A. Dashiff, R. A. Junka, M. Libera, and D. E. Kadouri, “Predation of human pathogens by the predatory bacteria *Micavibrio aeruginosavorus* and *Bdellovibrio bacteriovorus*,” *J Appl Microbiol*, vol. 110, no. 2, pp. 431–444, Feb. 2011, doi: 10.1111/j.1365-2672.2010.04900.x.
- [165] H. Stolp and M. P. Starr, “*Bdellovibrio bacteriovorus* gen. et sp. n., a predatory, ectoparasitic, and bacteriolytic microorganism,” *Antonie Van Leeuwenhoek*, vol. 29, no. 1, pp. 217–248, Dec. 1963, doi: 10.1007/BF02046064.
- [166] E. Jurkevitch and Y. Davidov, “Phylogenetic Diversity and Evolution of Predatory Prokaryotes,” *ACS Division of Fuel Chemistry*. 2006.
- [167] J. A. Fuhrman and D. A. Caron, “Heterotrophic Planktonic Microbes: Virus, Bacteria, Archaea, and Protozoa,” in *Manual of Environmental Microbiology*, Washington, DC, USA: ASM Press, 2015, pp. 4.2.2-1-4.2.2-34. doi: 10.1128/9781555818821.ch4.2.2.
- [168] J. N. Thompson, “The Evolution of Species Interactions,” *Science (1979)*, vol. 284, no. 5423, pp. 2116–2118, Jun. 1999, doi: 10.1126/science.284.5423.2116.
- [169] S. A. Pineiro, H. N. Williams, and O. C. Stine, “Phylogenetic relationships amongst the saltwater members of the genus *Bacteriovorax* using *rpoB* sequences and reclassification of *Bacteriovorax stolpii* as *Bacteriolyticum stolpii* gen. nov., comb. nov.,” *Int J Syst Evol Microbiol*, vol. 58, no. 5, pp. 1203–1209, May 2008, doi: 10.1099/ijs.0.65710-0.
- [170] D. Schwudke, E. Strauch, M. Krueger, and B. Appel, “Taxonomic Studies of Predatory *Bdellovibrios* Based on 16S rRNA Analysis, Ribotyping and the hit Locus and Characterization of Isolates from the Gut of Animals,” *Syst Appl Microbiol*, vol. 24, no. 3, pp. 385–394, Jan. 2001, doi: 10.1078/0723-2020-00042.
- [171] V. Martínez, C. Herencias, E. Jurkevitch, and M. A. Prieto, “Engineering a predatory bacterium as a proficient killer agent for intracellular bio-products recovery: The case of the polyhydroxyalkanoates,” *Sci Rep*, vol. 6, no. 1, p. 24381, Apr. 2016, doi: 10.1038/srep24381.
- [172] E. Bratanis, T. Andersson, R. Lood, and E. Bukowska-Faniband, “Biotechnological Potential of *Bdellovibrio* and Like Organisms and Their Secreted Enzymes,” *Front Microbiol*, vol. 11, Apr. 2020, doi: 10.3389/fmicb.2020.00662.

- [173] F. M. Cavallo, L. Jordana, A. W. Friedrich, C. Glasner, and J. M. van Dijl, “*Bdellovibrio bacteriovorus*: a potential ‘living antibiotic’ to control bacterial pathogens,” *Crit Rev Microbiol*, vol. 47, no. 5, pp. 630–646, Sep. 2021, doi: 10.1080/1040841X.2021.1908956.
- [174] K. Shatzkes *et al.*, “Predatory Bacteria Attenuate *Klebsiella pneumoniae* Burden in Rat Lungs,” *mBio*, vol. 7, no. 6, Dec. 2016, doi: 10.1128/mBio.01847-16.
- [175] R. M. Q. Shanks *et al.*, “An Eye to a Kill: Using Predatory Bacteria to Control Gram-Negative Pathogens Associated with Ocular Infections,” *PLoS One*, vol. 8, no. 6, p. e66723, Jun. 2013, doi: 10.1371/journal.pone.0066723.
- [176] R. J. Atterbury and J. Tyson, “Predatory bacteria as living antibiotics – where are we now?,” *Microbiology (N Y)*, vol. 167, no. 1, Jan. 2021, doi: 10.1099/mic.0.001025.
- [177] C. Ghosh, P. Sarkar, R. Issa, and J. Haldar, “Alternatives to Conventional Antibiotics in the Era of Antimicrobial Resistance,” *Trends Microbiol*, vol. 27, no. 4, pp. 323–338, Apr. 2019, doi: 10.1016/j.tim.2018.12.010.
- [178] R. Gallet, T. Tully, and M. E. K. Evans, “ECOLOGICAL CONDITIONS AFFECT EVOLUTIONARY TRAJECTORY IN A PREDATOR-PREY SYSTEM,” *Evolution (N Y)*, vol. 63, no. 3, pp. 641–651, Mar. 2009, doi: 10.1111/j.1558-5646.2008.00559.x.
- [179] M. C. Duncan *et al.*, “High-Throughput Analysis of Gene Function in the Bacterial Predator *Bdellovibrio bacteriovorus*,” *mBio*, vol. 10, no. 3, Jun. 2019, doi: 10.1128/mBio.01040-19.
- [180] C. Lambert *et al.*, “Interrupting peptidoglycan deacetylation during *Bdellovibrio* predator-prey interaction prevents ultimate destruction of prey wall, liberating bacterial-ghosts,” *Sci Rep*, vol. 6, no. 1, p. 26010, May 2016, doi: 10.1038/srep26010.
- [181] M. Selma-Royo *et al.*, “Perinatal environment shapes microbiota colonization and infant growth: impact on host response and intestinal function,” *Microbiome*, vol. 8, no. 1, p. 167, Dec. 2020, doi: 10.1186/s40168-020-00940-8.
- [182] L. Cuthbertson *et al.*, “Lung function and microbiota diversity in cystic fibrosis,” *Microbiome*, vol. 8, no. 1, p. 45, Dec. 2020, doi: 10.1186/s40168-020-00810-3.
- [183] G. Dayama, S. Priya, D. E. Niccum, A. Khoruts, and R. Blekhan, “Interactions between the gut microbiome and host gene regulation in cystic fibrosis,” *Genome Med*, vol. 12, no. 1, p. 12, Dec. 2020, doi: 10.1186/s13073-020-0710-2.
- [184] E. R. Andrinopoulou, J. P. Clancy, and R. D. Szczesniak, “Multivariate joint modeling to identify markers of growth and lung function decline that predict cystic fibrosis pulmonary exacerbation onset,” *BMC Pulm Med*, vol. 20, no. 1, p. 142, Dec. 2020, doi: 10.1186/s12890-020-1177-z.
- [185] D. G. Burke *et al.*, “The altered gut microbiota in adults with cystic fibrosis,” *BMC Microbiol*, vol. 17, no. 1, p. 58, Dec. 2017, doi: 10.1186/s12866-017-0968-8.
- [186] C. E. Price and G. A. O’Toole, “The Gut-Lung Axis in Cystic Fibrosis,” *J Bacteriol*, vol. 203, no. 20, Sep. 2021,

doi: 10.1128/JB.00311-21.

[187] G. Duytschaever, G. Huys, M. Bekaert, L. Boulanger, K. de Boeck, and P. Vandamme, “Cross-Sectional and Longitudinal Comparisons of the Predominant Fecal Microbiota Compositions of a Group of Pediatric Patients with Cystic Fibrosis and Their Healthy Siblings,” *Appl Environ Microbiol*, vol. 77, no. 22, pp. 8015–8024, Nov. 2011, doi: 10.1128/AEM.05933-11.

[188] S. Schippa *et al.*, “Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Allelic Variants Relate to Shifts in Faecal Microbiota of Cystic Fibrosis Patients,” *PLoS One*, vol. 8, no. 4, p. e61176, Apr. 2013, doi: 10.1371/journal.pone.0061176.

[189] K. M. Antosca *et al.*, “Altered Stool Microbiota of Infants with Cystic Fibrosis Shows a Reduction in Genera Associated with Immune Programming from Birth,” *J Bacteriol*, vol. 201, no. 16, Aug. 2019, doi: 10.1128/JB.00274-19.

[190] R. Knight *et al.*, “Best practices for analysing microbiomes,” *Nat Rev Microbiol*, vol. 16, no. 7, pp. 410–422, Jul. 2018, doi: 10.1038/s41579-018-0029-9.

[191] A. Rintala *et al.*, “Gut Microbiota Analysis Results Are Highly Dependent on the 16S rRNA Gene Target Region, Whereas the Impact of DNA Extraction Is Minor,” *J Biomol Tech*, vol. 28, no. 1, pp. 19–30, Apr. 2017, doi: 10.7171/jbt.17-2801-003.

[192] M. D. Mayers, C. Moon, G. S. Stupp, A. I. Su, and D. W. Wolan, “Quantitative Metaproteomics and Activity-Based Probe Enrichment Reveals Significant Alterations in

Protein Expression from a Mouse Model of Inflammatory Bowel Disease,” *J Proteome Res*, vol. 16, no. 2, pp. 1014–1026, Feb. 2017, doi: 10.1021/acs.jproteome.6b00938.

[193] X. Zhang, L. Li, J. Mayne, Z. Ning, A. Stintzi, and D. Figeys, “Assessing the impact of protein extraction methods for human gut metaproteomics,” *J Proteomics*, vol. 180, pp. 120–127, May 2018, doi: 10.1016/j.jprot.2017.07.001.

[194] N. Issa Isaac, D. Philippe, A. Nicholas, D. Raoult, and C. Eric, “Metaproteomics of the human gut microbiota: Challenges and contributions to other OMICS,” *Clinical Mass Spectrometry*, vol. 14, pp. 18–30, Sep. 2019, doi: 10.1016/j.clinms.2019.06.001.

[195] B. A. Petriz and O. L. Franco, “Metaproteomics as a Complementary Approach to Gut Microbiota in Health and Disease,” *Front Chem*, vol. 5, Jan. 2017, doi: 10.3389/fchem.2017.00004.

[196] G. Debyser *et al.*, “A Method for Comprehensive Proteomic Analysis of Human Faecal Samples to Investigate Gut Dysbiosis in Patients with Cystic Fibrosis,” 2019, pp. 137–160. doi: 10.1007/978-3-030-12298-0\_6.

[197] A. J. Lee, G. G. Einarsson, D. F. Gilpin, and M. M. Tunney, “Multi-Omics Approaches: The Key to Improving Respiratory Health in People With Cystic Fibrosis?,” *Front Pharmacol*, vol. 11, Sep. 2020, doi: 10.3389/fphar.2020.569821.

[198] C. T. Brown *et al.*, “Hospitalized Premature Infants Are Colonized by Related Bacterial Strains with Distinct Proteomic Profiles,” *mBio*, vol. 9, no. 2,

- May 2018, doi: 10.1128/mBio.00441-18.
- [199] A. Bevivino, G. Bacci, P. Drevinek, M. T. Nelson, L. Hoffman, and A. Mengoni, “Deciphering the Ecology of Cystic Fibrosis Bacterial Communities: Towards Systems-Level Integration,” *Trends Mol Med*, vol. 25, no. 12, pp. 1110–1122, Dec. 2019, doi: 10.1016/j.molmed.2019.07.008.
- [200] C. Mirzayi *et al.*, “Strengthening The Organization and Reporting of Microbiome Studies (STORMS): A Reporting Checklist for Human Microbiome Research,” *bioRxiv*, p. 2020.06.24.167353, Jan. 2020, doi: 10.1101/2020.06.24.167353.
- [201] A. Klindworth *et al.*, “Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies,” *Nucleic Acids Res*, vol. 41, no. 1, pp. e1–e1, Jan. 2013, doi: 10.1093/nar/gks808.
- [202] E. Bolyen *et al.*, “Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2,” *Nat Biotechnol*, vol. 37, no. 8, pp. 852–857, Aug. 2019, doi: 10.1038/s41587-019-0209-9.
- [203] M. S. Robeson *et al.*, “RESCRIPt: Reproducible sequence taxonomy reference database management,” *PLoS Comput Biol*, vol. 17, no. 11, p. e1009581, Nov. 2021, doi: 10.1371/journal.pcbi.1009581.
- [204] N. A. Bokulich *et al.*, “Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2’s q2-feature-classifier plugin,” *Microbiome*, vol. 6, no. 1, p. 90, Dec. 2018, doi: 10.1186/s40168-018-0470-z.
- [205] J. J. Werner *et al.*, “Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys,” *ISME J*, vol. 6, no. 1, pp. 94–103, Jan. 2012, doi: 10.1038/ismej.2011.82.
- [206] Y. Perez-Riverol *et al.*, “The PRIDE database and related tools and resources in 2019: improving support for quantification data,” *Nucleic Acids Res*, vol. 47, no. D1, pp. D442–D450, Jan. 2019, doi: 10.1093/nar/gky1106.
- [207] J. Li *et al.*, “An integrated catalog of reference genes in the human gut microbiome,” *Nat Biotechnol*, vol. 32, no. 8, pp. 834–841, Aug. 2014, doi: 10.1038/nbt.2942.
- [208] “UniProt: a worldwide hub of protein knowledge,” *Nucleic Acids Res*, vol. 47, no. D1, pp. D506–D515, Jan. 2019, doi: 10.1093/nar/gky1049.
- [209] B. Liao *et al.*, “iMetaLab 1.0: a web platform for metaproteomics data analysis,” *Bioinformatics*, vol. 34, no. 22, pp. 3954–3956, Nov. 2018, doi: 10.1093/bioinformatics/bty466.
- [210] P. J. McMurdie and S. Holmes, “phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data,” *PLoS One*, vol. 8, no. 4, p. e61217, Apr. 2013, doi: 10.1371/journal.pone.0061217.
- [211] S. M. Meeker *et al.*, “CFTR dysregulation drives active selection of the gut microbiome,” *PLoS Pathog*, vol. 16, no. 1, p. e1008251, Jan. 2020, doi: 10.1371/journal.ppat.1008251.
- [212] A. Martinović, R. Cocuzzi, S. Arioli, and D. Mora, “Streptococcus thermophilus: To Survive, or Not to

- Survive the Gastrointestinal Tract, That Is the Question!,” *Nutrients*, vol. 12, no. 8, p. 2175, Jul. 2020, doi: 10.3390/nu12082175.
- [213] M. Egan, E. Dempsey, C. A. Ryan, R. P. Ross, and C. Stanton, “The Sporobiota of the Human Gut,” *Gut Microbes*, vol. 13, no. 1, Jan. 2021, doi: 10.1080/19490976.2020.1863134.
- [214] X. Zhang *et al.*, “Deep Metaproteomics Approach for the Study of Human Microbiomes,” *Anal Chem*, vol. 89, no. 17, pp. 9407–9415, Sep. 2017, doi: 10.1021/acs.analchem.7b02224.
- [215] X. Zhang and D. Figeys, “Perspective and Guidelines for Metaproteomics in Microbiome Studies,” *J Proteome Res*, vol. 18, no. 6, pp. 2370–2380, Jun. 2019, doi: 10.1021/acs.jproteome.9b00054.
- [216] C. García-Durán *et al.*, “Distinct Human Gut Microbial Taxonomic Signatures Uncovered With Different Sample Processing and Microbial Cell Disruption Methods for Metaproteomic Analysis,” *Front Microbiol*, vol. 12, Jul. 2021, doi: 10.3389/fmicb.2021.618566.
- [217] S. Martínez Arbas *et al.*, “Challenges, Strategies, and Perspectives for Reference-Independent Longitudinal Multi-Omic Microbiome Studies,” *Front Genet*, vol. 12, Jun. 2021, doi: 10.3389/fgene.2021.666244.
- [218] L. Mancabelli *et al.*, “The Impact of Primer Design on Amplicon-Based Metagenomic Profiling Accuracy: Detailed Insights into Bifidobacterial Community Structure,” *Microorganisms*, vol. 8, no. 1, p. 131, Jan. 2020, doi: 10.3390/microorganisms8010131.
- [219] M. J. Coffey *et al.*, “Gut Microbiota in Children With Cystic Fibrosis: A Taxonomic and Functional Dysbiosis,” *Sci Rep*, vol. 9, no. 1, p. 18593, Dec. 2019, doi: 10.1038/s41598-019-55028-7.
- [220] H.-H. Chua *et al.*, “Intestinal Dysbiosis Featuring Abundance of *Ruminococcus gnavus* Associates With Allergic Diseases in Infants,” *Gastroenterology*, vol. 154, no. 1, pp. 154–167, Jan. 2018, doi: 10.1053/j.gastro.2017.09.006.
- [221] H. Wu *et al.*, “Fucosidases from the human gut symbiont *Ruminococcus gnavus*,” *Cellular and Molecular Life Sciences*, vol. 78, no. 2, pp. 675–693, Jan. 2021, doi: 10.1007/s00018-020-03514-x.
- [222] C. Gren, M. R. Spiegelhauer, E. C. Rotbain, B. K. Ehmsen, P. Kampmann, and L. P. Andersen, “*Ruminococcus gnavus* bacteraemia in a patient with multiple haematological malignancies,” *Access Microbiol*, vol. 1, no. 8, Oct. 2019, doi: 10.1099/acmi.0.000048.
- [223] E. H. Crost *et al.*, “The mucin-degradation strategy of *Ruminococcus gnavus*: The importance of intramolecular *trans*-sialidases,” *Gut Microbes*, vol. 7, no. 4, pp. 302–312, Jul. 2016, doi: 10.1080/19490976.2016.1186334.
- [224] D. Azzouz *et al.*, “Lupus nephritis is linked to disease-activity associated expansions and immunity to a gut commensal,” *Ann Rheum Dis*, vol. 78, no. 7, pp. 947–956, Jul. 2019, doi: 10.1136/annrheumdis-2018-214856.
- [225] A. Resch, R. Rosenstein, C. Nerz, and F. Götz, “Differential Gene

- Expression Profiling of *Staphylococcus aureus* Cultivated under Biofilm and Planktonic Conditions,” *Appl Environ Microbiol*, vol. 71, no. 5, pp. 2663–2676, May 2005, doi: 10.1128/AEM.71.5.2663-2676.2005.
- [226] K. R. Fourie and H. L. Wilson, “Understanding GroEL and DnaK Stress Response Proteins as Antigens for Bacterial Diseases,” *Vaccines (Basel)*, vol. 8, no. 4, p. 773, Dec. 2020, doi: 10.3390/vaccines8040773.
- [227] D. Legrand, A. Pierce, E. Ellass, M. Carpentier, C. Mariller, and J. Mazurier, “Lactoferrin Structure and Functions,” in *Bioactive Components of Milk*, New York, NY: Springer New York, pp. 163–194. doi: 10.1007/978-0-387-74087-4\_6.
- [228] Y. Aratani, “Myeloperoxidase: Its role for host defense, inflammation, and neutrophil function,” *Arch Biochem Biophys*, vol. 640, pp. 47–52, Feb. 2018, doi: 10.1016/j.abb.2018.01.004.
- [229] R. C. de Lisle and D. Borowitz, “The Cystic Fibrosis Intestine,” *Cold Spring Harb Perspect Med*, vol. 3, no. 9, pp. a009753–a009753, Sep. 2013, doi: 10.1101/cshperspect.a009753.
- [230] D. Gelfond *et al.*, “Impact of CFTR Modulation on Intestinal pH, Motility, and Clinical Outcomes in Patients With Cystic Fibrosis and the G551D Mutation,” *Clin Transl Gastroenterol*, vol. 8, no. 3, p. e81, Mar. 2017, doi: 10.1038/ctg.2017.10.
- [231] C. A. Youngberg *et al.*, “Comparison of gastrointestinal pH in cystic fibrosis and healthy subjects,” *Dig Dis Sci*, vol. 32, no. 5, pp. 472–480, May 1987, doi: 10.1007/BF01296029.
- [232] C. Y. Ooi *et al.*, “Impact of CFTR modulation with Ivacaftor on Gut Microbiota and Intestinal Inflammation,” *Sci Rep*, vol. 8, no. 1, p. 17834, Dec. 2018, doi: 10.1038/s41598-018-36364-6.
- [233] V. Sergeev, F. Y. Chou, G. Y. Lam, C. M. Hamilton, P. G. Wilcox, and B. S. Quon, “The Extrapulmonary Effects of Cystic Fibrosis Transmembrane Conductance Regulator Modulators in Cystic Fibrosis,” *Ann Am Thorac Soc*, vol. 17, no. 2, pp. 147–154, Feb. 2020, doi: 10.1513/AnnalsATS.201909-671CME.
- [234] S. M. Rowe *et al.*, “Clinical Mechanism of the Cystic Fibrosis Transmembrane Conductance Regulator Potentiator Ivacaftor in G551D-mediated Cystic Fibrosis,” *Am J Respir Crit Care Med*, vol. 190, no. 2, pp. 175–184, Jul. 2014, doi: 10.1164/rccm.201404-0703OC.
- [235] C. Y. Ooi and P. R. Durie, “Cystic fibrosis from the gastroenterologist’s perspective,” *Nat Rev Gastroenterol Hepatol*, vol. 13, no. 3, pp. 175–185, Mar. 2016, doi: 10.1038/nrgastro.2015.226.
- [236] J. D. Lloyd-Still, “Crohn’s disease and cystic fibrosis,” *Dig Dis Sci*, vol. 39, no. 4, pp. 880–885, Apr. 1994, doi: 10.1007/BF02087437.
- [237] J. Konrad, E. Eber, and V. Stadlbauer, “Changing paradigms in the treatment of gastrointestinal complications of cystic fibrosis in the era of cystic fibrosis transmembrane conductance regulator modulators,” *Paediatr Respir Rev*, vol. 42, pp. 9–16, Jun. 2022, doi: 10.1016/j.prrv.2020.12.001.

- [238] P. Scott, K. Anderson, M. Singhanian, and R. Cormier, “Cystic Fibrosis, CFTR, and Colorectal Cancer,” *Int J Mol Sci*, vol. 21, no. 8, p. 2891, Apr. 2020, doi: 10.3390/ijms21082891.
- [239] K. F. Budden *et al.*, “Emerging pathogenic links between microbiota and the gut–lung axis,” *Nat Rev Microbiol*, vol. 15, no. 1, pp. 55–63, Jan. 2017, doi: 10.1038/nrmicro.2016.142.
- [240] D. Zhang, S. Li, N. Wang, H.-Y. Tan, Z. Zhang, and Y. Feng, “The Cross-Talk Between Gut Microbiota and Lungs in Common Lung Diseases,” *Front Microbiol*, vol. 11, Feb. 2020, doi: 10.3389/fmicb.2020.00301.
- [241] B. R. Loman *et al.*, “Age and environmental exposures influence the fecal bacteriome of young children with cystic fibrosis,” *Pediatr Pulmonol*, vol. 55, no. 7, pp. 1661–1670, Jul. 2020, doi: 10.1002/ppul.24766.
- [242] A. G. Hoen *et al.*, “Associations between Gut Microbial Colonization in Early Life and Respiratory Outcomes in Cystic Fibrosis,” *J Pediatr*, vol. 167, no. 1, pp. 138–147.e3, Jul. 2015, doi: 10.1016/j.jpeds.2015.02.049.
- [243] L. A. Carmody *et al.*, “Fluctuations in airway bacterial communities associated with clinical states and disease stages in cystic fibrosis,” *PLoS One*, vol. 13, no. 3, p. e0194060, Mar. 2018, doi: 10.1371/journal.pone.0194060.
- [244] A. García-Caballero *et al.*, “Evaluation of different phenotypic methods to detect methicillin resistance in *Staphylococcus aureus* isolates recovered from cystic fibrosis patients,” *Diagn Microbiol Infect Dis*, vol. 102, no. 1, p. 115559, Jan. 2022, doi: 10.1016/j.diagmicrobio.2021.115559.
- [245] M. Kosecka-Strojek, M. Wolska, D. Żabicka, E. Sadowy, and J. Międzobrodzki, “Identification of Clinically Relevant *Streptococcus* and *Enterococcus* Species Based on Biochemical Methods and 16S rRNA, *sodA*, *tuf*, *rpoB*, and *recA* Gene Sequencing,” *Pathogens*, vol. 9, no. 11, p. 939, Nov. 2020, doi: 10.3390/pathogens9110939.
- [246] K. A. Jolley, J. E. Bray, and M. C. J. Maiden, “Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications,” *Wellcome Open Res*, vol. 3, p. 124, Sep. 2018, doi: 10.12688/wellcomeopenres.14826.1.
- [247] L. Beule and P. Karlovsky, “Improved normalization of species count data in ecology by scaling with ranked subsampling (SRS): application to microbial communities,” *PeerJ*, vol. 8, p. e9593, Aug. 2020, doi: 10.7717/peerj.9593.
- [248] S. Beck *et al.*, “Cystic fibrosis patients with the 3272-26A>G mutation have mild disease, leaky alternative mRNA splicing, and CFTR protein at the cell membrane,” *Hum Mutat*, vol. 14, no. 2, pp. 133–144, 1999, doi: 10.1002/(SICI)1098-1004(1999)14:2<133::AID-HUMU5>3.0.CO;2-T.
- [249] D. Bozon, J. Zielenski, F. Rininsland, and L.-C. Tsui, “Identification of four new mutations in the cystic fibrosis transmembrane conductance regulator gene: I148T, L1077P, Y1092X, 2183AA→G,” *Hum Mutat*, vol. 3, no. 3, pp. 330–332, 1994, doi: 10.1002/humu.1380030329.

- [250] L. Masvidal *et al.*, “The p.Arg258Gly Mutation in Intracellular Loop 2 of CFTR is Associated with CFTR -Related Disorders,” *Genet Test Mol Biomarkers*, vol. 13, no. 6, pp. 765–768, Dec. 2009, doi: 10.1089/gtmb.2009.0070.
- [251] E. F. McKone, C. H. Goss, and M. L. Aitken, “CFTR Genotype as a Predictor of Prognosis in Cystic Fibrosis,” *Chest*, vol. 130, no. 5, pp. 1441–1447, Nov. 2006, doi: 10.1378/chest.130.5.1441.
- [252] M. Sathe and R. Houwen, “Is meconium ileus associated with worse outcomes in cystic fibrosis?,” *Journal of Cystic Fibrosis*, vol. 18, no. 6, p. 746, Nov. 2019, doi: 10.1016/j.jcf.2019.10.003.
- [253] C. Castellani *et al.*, “Consensus on the use and interpretation of cystic fibrosis mutation analysis in clinical practice,” *Journal of Cystic Fibrosis*, vol. 7, no. 3, pp. 179–196, May 2008, doi: 10.1016/j.jcf.2008.03.009.
- [254] G. S. Sawicki *et al.*, “Rate of Lung Function Decline in People with Cystic Fibrosis Having a Residual Function Gene Mutation,” *Pulm Ther*, Nov. 2022, doi: 10.1007/s41030-022-00202-y.
- [255] C. Saralegui *et al.*, “Statistical Evaluation of Metaproteomics and 16S rRNA Amplicon Sequencing Techniques for Study of Gut Microbiota Establishment in Infants with Cystic Fibrosis,” *Microbiol Spectr*, Oct. 2022, doi: 10.1128/spectrum.01466-22.
- [256] P. Soret *et al.*, “Respiratory mycobiome and suggestion of inter-kingdom network during acute pulmonary exacerbation in cystic fibrosis,” *Sci Rep*, vol. 10, no. 1, p. 3589, Dec. 2020, doi: 10.1038/s41598-020-60015-4.
- [257] D. Henares *et al.*, “Differential nasopharyngeal microbiota composition in children according to respiratory health status,” *Microb Genom*, vol. 7, no. 10, Oct. 2021, doi: 10.1099/mgen.0.000661.
- [258] F. Festini *et al.*, “Isolation measures for prevention of infection with respiratory pathogens in cystic fibrosis: a systematic review,” *Journal of Hospital Infection*, vol. 64, no. 1, pp. 1–6, Sep. 2006, doi: 10.1016/j.jhin.2006.02.021.
- [259] D. F. Lima, R. W. Cohen, G. A. Rocha, R. M. Albano, E. A. Marques, and R. S. Leão, “Genomic information on multidrug-resistant livestock-associated methicillin-resistant *Staphylococcus aureus* ST398 isolated from a Brazilian patient with cystic fibrosis,” *Mem Inst Oswaldo Cruz*, vol. 112, no. 1, pp. 79–80, Jan. 2017, doi: 10.1590/0074-02760160342.
- [260] A. Molina *et al.*, “High prevalence in cystic fibrosis patients of multiresistant hospital-acquired methicillin-resistant *Staphylococcus aureus* ST228-SCCmecI capable of biofilm formation,” *Journal of Antimicrobial Chemotherapy*, vol. 62, no. 5, pp. 961–967, Jul. 2008, doi: 10.1093/jac/dkn302.
- [261] R. Escudero-Sánchez *et al.*, “Long-Term Impact of Suppressive Antibiotic Therapy on Intestinal Microbiota,” *Genes (Basel)*, vol. 12, no. 1, p. 41, Dec. 2020, doi: 10.3390/genes12010041.
- [262] F. Sommer and F. Bäckhed, “The gut microbiota — masters of host

- development and physiology,” *Nat Rev Microbiol*, vol. 11, no. 4, pp. 227–238, Apr. 2013, doi: 10.1038/nrmicro2974.
- [263] J. R. Marchesi *et al.*, “The gut microbiota and host health: a new clinical frontier,” *Gut*, vol. 65, no. 2, pp. 330–339, Feb. 2016, doi: 10.1136/gutjnl-2015-309990.
- [264] S. v. Lynch and O. Pedersen, “The Human Intestinal Microbiome in Health and Disease,” *New England Journal of Medicine*, vol. 375, no. 24, pp. 2369–2379, Dec. 2016, doi: 10.1056/NEJMra1600266.
- [265] W. A. Walker, “The importance of appropriate initial bacterial colonization of the intestine in newborn, child, and adult health,” *Pediatr Res*, vol. 82, no. 3, pp. 387–395, Sep. 2017, doi: 10.1038/pr.2017.111.
- [266] S. Rautava, R. Luoto, S. Salminen, and E. Isolauri, “Microbial contact during pregnancy, intestinal colonization and human disease,” *Nat Rev Gastroenterol Hepatol*, vol. 9, no. 10, pp. 565–576, Oct. 2012, doi: 10.1038/nrgastro.2012.144.
- [267] M. E. Perez-Muñoz, M.-C. Arrieta, A. E. Ramer-Tait, and J. Walter, “A critical assessment of the ‘sterile womb’ and ‘in utero colonization’ hypotheses: implications for research on the pioneer infant microbiome,” *Microbiome*, vol. 5, no. 1, p. 48, Dec. 2017, doi: 10.1186/s40168-017-0268-4.
- [268] K. Aagaard, J. Ma, K. M. Antony, R. Ganu, J. Petrosino, and J. Versalovic, “The Placenta Harbors a Unique Microbiome,” *Sci Transl Med*, vol. 6, no. 237, May 2014, doi: 10.1126/scitranslmed.3008599.
- [269] F. Bäckhed *et al.*, “Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life,” *Cell Host Microbe*, vol. 17, no. 5, pp. 690–703, May 2015, doi: 10.1016/j.chom.2015.04.004.
- [270] J. C. Madan, S. F. Farzan, P. L. Hibberd, and M. R. Karagas, “Normal neonatal microbiome variation in relation to environmental factors, infection and allergy,” *Curr Opin Pediatr*, vol. 24, no. 6, pp. 753–759, Dec. 2012, doi: 10.1097/MOP.0b013e32835a1ac8.
- [271] P. S. la Rosa *et al.*, “Patterned progression of bacterial populations in the premature infant gut,” *Proceedings of the National Academy of Sciences*, vol. 111, no. 34, pp. 12522–12527, Aug. 2014, doi: 10.1073/pnas.1409497111.
- [272] D. A. Chernikova *et al.*, “The premature infant gut microbiome during the first 6 weeks of life differs based on gestational maturity at birth,” *Pediatr Res*, vol. 84, no. 1, pp. 71–79, Jul. 2018, doi: 10.1038/s41390-018-0022-z.
- [273] J. Zeitlin *et al.*, “Preterm birth time trends in Europe: a study of 19 countries,” *BJOG*, vol. 120, no. 11, pp. 1356–1365, Oct. 2013, doi: 10.1111/1471-0528.12281.
- [274] D. B. DiGiulio *et al.*, “Temporal and spatial variation of the human microbiota during pregnancy,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 35, pp. 11060–11065, Sep. 2015, doi: 10.1073/pnas.1502875112.
- [275] A. L. Patel *et al.*, “Longitudinal Survey of Microbiota in Hospitalized Preterm Very-Low-Birth-Weight Infants,” *J Pediatr Gastroenterol Nutr*,

- vol. 62, no. 2, pp. 292–303, Feb. 2016, doi: 10.1097/MPG.0000000000000913.
- [276] S. Unger, A. Stintzi, P. Shah, D. Mack, and D. L. O’Connor, “Gut microbiota of the very-low-birth-weight infant,” *Pediatr Res*, vol. 77, no. 1–2, pp. 205–213, Jan. 2015, doi: 10.1038/pr.2014.162.
- [277] L. Moles *et al.*, “Preterm infant gut colonization in the neonatal ICU and complete restoration 2 years later,” *Clinical Microbiology and Infection*, vol. 21, no. 10, pp. 936.e1–936.e10, Oct. 2015, doi: 10.1016/j.cmi.2015.06.003.
- [278] T. G. Elgin, S. L. Kern, and S. J. McElroy, “Development of the Neonatal Intestinal Microbiome and Its Association With Necrotizing Enterocolitis,” *Clin Ther*, vol. 38, no. 4, pp. 706–715, Apr. 2016, doi: 10.1016/j.clinthera.2016.01.005.
- [279] M. W. Groer, K. E. Gregory, A. Louis-Jacques, S. Thibeau, and W. A. Walker, “The very low birth weight infant microbiome and childhood health,” *Birth Defects Res C Embryo Today*, vol. 105, no. 4, pp. 252–264, Dec. 2015, doi: 10.1002/bdrc.21115.
- [280] D. Yusef, T. Shalakhti, S. Awad, H. Algharaibeh, and W. Khasawneh, “Clinical characteristics and epidemiology of sepsis in the neonatal intensive care unit in the era of multi-drug resistant organisms: A retrospective review,” *Pediatr Neonatol*, vol. 59, no. 1, pp. 35–41, Feb. 2018, doi: 10.1016/j.pedneo.2017.06.001.
- [281] C. Martineau *et al.*, “*Serratia marcescens* Outbreak in a Neonatal Intensive Care Unit: New Insights from Next-Generation Sequencing Applications,” *J Clin Microbiol*, vol. 56, no. 9, Sep. 2018, doi: 10.1128/JCM.00235-18.
- [282] S. D. Mahlen, “*Serratia* Infections: from Military Experiments to Current Practice,” *Clin Microbiol Rev*, vol. 24, no. 4, pp. 755–791, Oct. 2011, doi: 10.1128/CMR.00017-11.
- [283] F. Baquero, “*Serratia*: un nuevo problema en microbiología hospitalaria,” *Proc. 2nd Congr. Nacional de Microbiología. Soc. Española de Microbiología*. p. 126, 1969.
- [284] T. Seemann, A. Goncalves da Silva, D. Bulach, M. Schultz, J. Kwong, and B. Howden, “Nullarbor.” Github, 2020.
- [285] A. M. Bolger, M. Lohse, and B. Usadel, “Trimmomatic: a flexible trimmer for Illumina sequence data,” *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, Aug. 2014, doi: 10.1093/bioinformatics/btu170.
- [286] D. E. Wood and S. L. Salzberg, “Kraken: ultrafast metagenomic sequence classification using exact alignments,” *Genome Biol*, vol. 15, no. 3, p. R46, 2014, doi: 10.1186/gb-2014-15-3-r46.
- [287] D. Kim, L. Song, F. P. Breitwieser, and S. L. Salzberg, “Centrifuge: rapid and sensitive classification of metagenomic sequences,” *Genome Res*, vol. 26, no. 12, pp. 1721–1729, Dec. 2016, doi: 10.1101/gr.210641.116.
- [288] A. Bankevich *et al.*, “SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing,” *Journal of Computational Biology*, vol. 19, no. 5, pp. 455–477, May 2012, doi: 10.1089/cmb.2012.0021.

- [289] T. Seemann, “Prokka: rapid prokaryotic genome annotation,” *Bioinformatics*, vol. 30, no. 14, pp. 2068–2069, Jul. 2014, doi: 10.1093/bioinformatics/btu153.
- [290] A. J. Page *et al.*, “Roary: rapid large-scale prokaryote pan genome analysis,” *Bioinformatics*, vol. 31, no. 22, pp. 3691–3693, Nov. 2015, doi: 10.1093/bioinformatics/btv421.
- [291] J. Hadfield, N. J. Croucher, R. J. Goater, K. Abudahab, D. M. Aanensen, and S. R. Harris, “Phandango: an interactive viewer for bacterial population genomics,” *Bioinformatics*, vol. 34, no. 2, pp. 292–293, Jan. 2018, doi: 10.1093/bioinformatics/btx610.
- [292] L. Vielva, M. de Toro, V. F. Lanza, and F. de la Cruz, “PLACNETw: a web-based tool for plasmid reconstruction from bacterial genomes,” *Bioinformatics*, vol. 33, no. 23, pp. 3796–3798, Dec. 2017, doi: 10.1093/bioinformatics/btx462.
- [293] Z. Xie and H. Tang, “ISEScan: automated identification of insertion sequence elements in prokaryotic genomes,” *Bioinformatics*, vol. 33, no. 21, pp. 3340–3347, Nov. 2017, doi: 10.1093/bioinformatics/btx433.
- [294] Y.-Y. Liu, J.-W. Lin, and C.-C. Chen, “cano-wgMLST\_BacCompare: A Bacterial Genome Analysis Platform for Epidemiological Investigation and Comparative Genomic Analysis,” *Front Microbiol*, vol. 10, Jul. 2019, doi: 10.3389/fmicb.2019.01687.
- [295] I. Letunic and P. Bork, “Interactive Tree Of Life (iTOL) v4: recent updates and new developments,” *Nucleic Acids Res*, vol. 47, no. W1, pp. W256–W259, Jul. 2019, doi: 10.1093/nar/gkz239.
- [296] V. F. Lanza, F. Baquero, F. de la Cruz, and T. M. Coque, “AcCNET (Accessory Genome Constellation Network): comparative genomics software for accessory genome analysis using bipartite networks,” *Bioinformatics*, vol. 33, no. 2, pp. 283–285, Jan. 2017, doi: 10.1093/bioinformatics/btw601.
- [297] D. Sharma, N. Farahbakhsh, S. Shastri, and P. Sharma, “Biomarkers for diagnosis of neonatal sepsis: a literature review,” *The Journal of Maternal-Fetal & Neonatal Medicine*, vol. 31, no. 12, pp. 1646–1659, Jun. 2018, doi: 10.1080/14767058.2017.1322060.
- [298] L. Ruiz, L. Moles, M. Gueimonde, and J. M. Rodriguez, “Perinatal Microbiomes’ Influence on Preterm Birth and Preterms’ Health: Influencing Factors and Modulation Strategies,” *J Pediatr Gastroenterol Nutr*, vol. 63, no. 6, pp. e193–e203, Dec. 2016, doi: 10.1097/MPG.0000000000001196.
- [299] L. F. Stinson, M. S. Payne, and J. A. Keelan, “Planting the seed: Origins, composition, and postnatal health significance of the fetal gastrointestinal microbiota,” *Crit Rev Microbiol*, vol. 43, no. 3, pp. 352–369, May 2017, doi: 10.1080/1040841X.2016.1211088.
- [300] C. Montagnani *et al.*, “*Serratia marcescens* outbreak in a neonatal intensive care unit: crucial role of implementing hand hygiene among external consultants,” *BMC Infect Dis*, vol. 15, no. 1, p. 11, Dec. 2015, doi: 10.1186/s12879-014-0734-6.
- [301] X. Cong *et al.*, “Gut Microbiome Developmental Patterns in Early Life of

- Preterm Infants: Impacts of Feeding and Gender,” *PLoS One*, vol. 11, no. 4, p. e0152751, Apr. 2016, doi: 10.1371/journal.pone.0152751.
- [302] C. J. Hill *et al.*, “Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort,” *Microbiome*, vol. 5, no. 1, p. 4, Dec. 2017, doi: 10.1186/s40168-016-0213-y.
- [303] M. A. Underwood and K. Sohn, “The Microbiota of the Extremely Preterm Infant,” *Clin Perinatol*, vol. 44, no. 2, pp. 407–427, Jun. 2017, doi: 10.1016/j.clp.2017.01.005.
- [304] A. Nogacka *et al.*, “Impact of intrapartum antimicrobial prophylaxis upon the intestinal microbiota and the prevalence of antibiotic resistance genes in vaginally delivered full-term neonates,” *Microbiome*, vol. 5, no. 1, p. 93, Dec. 2017, doi: 10.1186/s40168-017-0313-3.
- [305] R. Hansen *et al.*, “First-Pass Meconium Samples from Healthy Term Vaginally-Delivered Neonates: An Analysis of the Microbiota,” *PLoS One*, vol. 10, no. 7, p. e0133320, Jul. 2015, doi: 10.1371/journal.pone.0133320.
- [306] A. N. Ardissonne *et al.*, “Meconium Microbiome Analysis Identifies Bacteria Correlated with Premature Birth,” *PLoS One*, vol. 9, no. 3, p. e90784, Mar. 2014, doi: 10.1371/journal.pone.0090784.
- [307] C. Alcon-Giner *et al.*, “Optimisation of 16S rRNA gut microbiota profiling of extremely low birth weight infants,” *BMC Genomics*, vol. 18, no. 1, p. 841, Dec. 2017, doi: 10.1186/s12864-017-4229-x.
- [308] A. Jacquot *et al.*, “Dynamics and Clinical Evolution of Bacterial Gut Microflora in Extremely Premature Patients,” *J Pediatr*, vol. 158, no. 3, pp. 390–396, Mar. 2011, doi: 10.1016/j.jpeds.2010.09.007.
- [309] D. J. C. Berkhout *et al.*, “Late-onset Sepsis in Preterm Infants Can Be Detected Preclinically by Fecal Volatile Organic Compound Analysis: A Prospective, Multicenter Cohort Study,” *Clinical Infectious Diseases*, Jun. 2018, doi: 10.1093/cid/ciy383.
- [310] V. Iebba *et al.*, “Higher Prevalence and Abundance of *Bdellovibrio bacteriovorus* in the Human Gut of Healthy Subjects,” *PLoS One*, vol. 8, no. 4, p. e61608, Apr. 2013, doi: 10.1371/journal.pone.0061608.
- [311] D. E. Kadouri, K. To, R. M. Q. Shanks, and Y. Doi, “Predatory Bacteria: A Potential Ally against Multidrug-Resistant Gram-Negative Pathogens,” *PLoS One*, vol. 8, no. 5, p. e63397, May 2013, doi: 10.1371/journal.pone.0063397.
- [312] L. Fernández *et al.*, “The human milk microbiota: Origin and potential roles in health and disease,” *Pharmacol Res*, vol. 69, no. 1, pp. 1–10, Mar. 2013, doi: 10.1016/j.phrs.2012.09.001.
- [313] P. S. Pannaraj *et al.*, “Association Between Breast Milk Bacterial Communities and Establishment and Development of the Infant Gut Microbiome,” *JAMA Pediatr*, vol. 171, no. 7, p. 647, Jul. 2017, doi: 10.1001/jamapediatrics.2017.0378.
- [314] A. J. Macpherson, M. G. de Agüero, and S. C. Ganal-Vonarburg, “How nutrition and the maternal microbiota shape the neonatal immune system,” *Nat Rev Immunol*, vol. 17, no.

- 8, pp. 508–517, Aug. 2017, doi: 10.1038/nri.2017.58.
- [315] S. Moukarzel and L. Bode, “Human Milk Oligosaccharides and the Preterm Infant,” *Clin Perinatol*, vol. 44, no. 1, pp. 193–207, Mar. 2017, doi: 10.1016/j.clp.2016.11.014.
- [316] G. R. Young *et al.*, “Reducing Viability Bias in Analysis of Gut Microbiota in Preterm Infants at Risk of NEC and Sepsis,” *Front Cell Infect Microbiol*, vol. 7, Jun. 2017, doi: 10.3389/fcimb.2017.00237.
- [317] A. Mellmann *et al.*, “High Interlaboratory Reproducibility and Accuracy of Next-Generation-Sequencing-Based Bacterial Genotyping in a Ring Trial,” *J Clin Microbiol*, vol. 55, no. 3, pp. 908–913, Mar. 2017, doi: 10.1128/JCM.02242-16.
- [318] M. Pérez-Losada, M. Arenas, and E. Castro-Nallar, “Microbial sequence typing in the genomic era,” *Infection, Genetics and Evolution*, vol. 63, pp. 346–359, Sep. 2018, doi: 10.1016/j.meegid.2017.09.022.
- [319] A. C. Schürch, S. Arredondo-Alonso, R. J. L. Willems, and R. V. Goering, “Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches,” *Clinical Microbiology and Infection*, vol. 24, no. 4, pp. 350–354, Apr. 2018, doi: 10.1016/j.cmi.2017.12.016.
- [320] E. Abreo and N. Altier, “Pangenome of *Serratia marcescens* strains from nosocomial and environmental origins reveals different populations and the links between them,” *Sci Rep*, vol. 9, no. 1, p. 46, Dec. 2019, doi: 10.1038/s41598-018-37118-0.
- [321] R. León-Sampedro *et al.*, “Phylogenomics of *Enterococcus faecalis* from wild birds: new insights into host-associated differences in core and accessory genomes of the species,” *Environ Microbiol*, vol. 21, no. 8, pp. 3046–3062, Aug. 2019, doi: 10.1111/1462-2920.14702.
- [322] B. Contreras-Moreira and P. Vinuesa, “GET\_HOMOLOGUES, a Versatile Software Package for Scalable and Robust Microbial Pangenome Analysis,” *Appl Environ Microbiol*, vol. 79, no. 24, pp. 7696–7701, Dec. 2013, doi: 10.1128/AEM.02411-13.
- [323] C. C. Thompson *et al.*, “Genomic taxonomy of vibrios,” *BMC Evol Biol*, vol. 9, no. 1, p. 258, Dec. 2009, doi: 10.1186/1471-2148-9-258.
- [324] D. Moradigaravand, C. J. Boinett, V. Martin, S. J. Peacock, and J. Parkhill, “Recent independent emergence of multiple multidrug-resistant *Serratia marcescens* clones within the United Kingdom and Ireland,” *Genome Res*, vol. 26, no. 8, pp. 1101–1109, Aug. 2016, doi: 10.1101/gr.205245.116.
- [325] L. Moles *et al.*, “*Serratia marcescens* colonization in preterm neonates during their neonatal intensive care unit stay,” *Antimicrob Resist Infect Control*, vol. 8, no. 1, p. 135, Dec. 2019, doi: 10.1186/s13756-019-0584-5.
- [326] K. Dawczynski *et al.*, “Intensified colonisation screening according to the recommendations of the German Commission for Hospital Hygiene and Infectious Diseases Prevention (KRINKO): identification and containment of a *Serratia marcescens* outbreak in the neonatal

- intensive care unit, Jena, Germany, 2013–2014,” *Infection*, vol. 44, no. 6, pp. 739–746, Dec. 2016, doi: 10.1007/s15010-016-0922-y.
- [327] F. Fleisch *et al.*, “Three Consecutive Outbreaks of *Serratia marcescens* in a Neonatal Intensive Care Unit,” *Clinical Infectious Diseases*, vol. 34, no. 6, pp. 767–773, Mar. 2002, doi: 10.1086/339046.
- [328] E. Mathieu *et al.*, “Paradigms of Lung Microbiota Functions in Health and Disease, Particularly, in Asthma,” *Front Physiol*, vol. 9, Aug. 2018, doi: 10.3389/fphys.2018.01168.
- [329] J. Shen, J. Zhou, Y. Xu, and Z. Xiu, “Prophages contribute to genome plasticity of *Klebsiella pneumoniae* and may involve the chromosomal integration of ARGs in CG258,” *Genomics*, vol. 112, no. 1, pp. 998–1010, Jan. 2020, doi: 10.1016/j.ygeno.2019.06.016.
- [330] K. S. Baker *et al.*, “The extant World War 1 dysentery bacillus NCTC1: a genomic analysis,” *The Lancet*, vol. 384, no. 9955, pp. 1691–1697, Nov. 2014, doi: 10.1016/S0140-6736(14)61789-X.
- [331] K. S. Baker *et al.*, “The Murray collection of pre-antibiotic era Enterobacteriaceae: a unique research resource,” *Genome Med*, vol. 7, no. 1, p. 97, Dec. 2015, doi: 10.1186/s13073-015-0222-7.
- [332] A. M. Devault *et al.*, “Second-Pandemic Strain of *Vibrio cholerae* from the Philadelphia Cholera Outbreak of 1849,” *New England Journal of Medicine*, vol. 370, no. 4, pp. 334–340, Jan. 2014, doi: 10.1056/NEJMoal308663.
- [333] L. H. Hansen, L. B. Jensen, H. I. Sørensen, and S. J. Sørensen, “Substrate specificity of the OqxAB multidrug resistance pump in *Escherichia coli* and selected enteric bacteria,” *Journal of Antimicrobial Chemotherapy*, vol. 60, no. 1, pp. 145–147, Jul. 2007, doi: 10.1093/jac/dkm167.
- [334] J. H. Tran and G. A. Jacoby, “Mechanism of plasmid-mediated quinolone resistance,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 8, pp. 5638–5642, Apr. 2002, doi: 10.1073/pnas.082092899.
- [335] A. Hernández, M. B. Sánchez, and J. L. Martínez, “Quinolone Resistance: Much More than Predicted,” *Front Microbiol*, vol. 2, 2011, doi: 10.3389/fmicb.2011.00022.
- [336] N. Joel Ehrenkranz, Elizabeth A. Bolyard, M. Wiener, and Timothy J. Cleary, “antibiotic-sensitive *Serratia marcescens* infections complicating cardiopulmonary operations: contaminated disinfectant as a reservoir,” *The Lancet*, vol. 316, no. 8207, pp. 1289–1292, Dec. 1980, doi: 10.1016/S0140-6736(80)92349-1.
- [337] K. Kaneko, R. Okamoto, R. Nakano, S. Kawakami, and M. Inoue, “Gene Mutations Responsible for Overexpression of AmpC  $\beta$ -Lactamase in Some Clinical Isolates of *Enterobacter cloacae*,” *J Clin Microbiol*, vol. 43, no. 6, pp. 2955–2958, Jun. 2005, doi: 10.1128/JCM.43.6.2955-2958.2005.
- [338] S. Meini, C. Tascini, M. Cei, E. Sozio, and G. M. Rossolini, “AmpC  $\beta$ -lactamase-producing Enterobacterales: what a clinician should know,” *Infection*, vol. 47, no. 3, pp. 363–375, Jun. 2019, doi: 10.1007/s15010-019-01291-9.

- [339] A. Raimondi, F. Sisto, and H. Nikaido, "Mutation in *Serratia marcescens* AmpC  $\beta$ -Lactamase Producing High-Level Resistance to Ceftazidime and Cefpirome," *Antimicrob Agents Chemother*, vol. 45, no. 8, pp. 2331–2339, Aug. 2001, doi: 10.1128/AAC.45.8.2331-2339.2001.
- [340] J. I. Campbell, O. Ciofu, and N. Høiby, "Pseudomonas aeruginosa isolates from patients with cystic fibrosis have different beta-lactamase expression phenotypes but are homogeneous in the ampC-ampR genetic region," *Antimicrob Agents Chemother*, vol. 41, no. 6, pp. 1380–1384, Jun. 1997, doi: 10.1128/AAC.41.6.1380.
- [341] H. Mammeri, P. Nordmann, A. Berkani, and F. Eb, "Contribution of extended-spectrum AmpC (ESAC)  $\beta$ -lactamases to carbapenem resistance in *Escherichia coli*," *FEMS Microbiol Lett*, vol. 282, no. 2, pp. 238–240, Apr. 2008, doi: 10.1111/j.1574-6968.2008.01126.x.
- [342] F. Baquero *et al.*, "Antibiotherapy of *Serratia marcescens* Septicemia in Children," *Chemotherapy*, vol. 23, no. 1, pp. 416–422, 1977, doi: 10.1159/000222084.
- [343] T. Naas, M. Blot, W. M. Fitch, and W. Arber, "Insertion sequence-related genetic variation in resting *Escherichia coli* K-12.," *Genetics*, vol. 136, no. 3, pp. 721–730, Mar. 1994, doi: 10.1093/genetics/136.3.721.
- [344] T. Naas, M. Blot, W. M. Fitch, and W. Arber, "Dynamics of IS-related genetic rearrangements in resting *Escherichia coli* K-12.," *Mol Biol Evol*, Mar. 1995, doi: 10.1093/oxfordjournals.molbev.a040198.
- [345] R. C. Allen, J. Engelstädter, S. Bonhoeffer, B. A. McDonald, and A. R. Hall, "Reversing resistance: different routes and common themes across pathogens," *Proceedings of the Royal Society B: Biological Sciences*, vol. 284, no. 1863, p. 20171619, Sep. 2017, doi: 10.1098/rspb.2017.1619.
- [346] V. I. Enne, A. A. Delsol, J. M. Roe, and P. M. Bennett, "Evidence of Antibiotic Resistance Gene Silencing in *Escherichia coli*," *Antimicrob Agents Chemother*, vol. 50, no. 9, pp. 3003–3010, Sep. 2006, doi: 10.1128/AAC.00137-06.
- [347] I. Hamidou Soumana, B. Linz, and E. T. Harvill, "Environmental Origin of the Genus *Bordetella*," *Front Microbiol*, vol. 8, Jan. 2017, doi: 10.3389/fmicb.2017.00028.
- [348] F. Baquero, "Threats of antibiotic resistance: an obliged reappraisal," *International Microbiology*, vol. 24, no. 4, pp. 499–506, Nov. 2021, doi: 10.1007/s10123-021-00184-y.
- [349] A. I. Celis and D. A. Relman, "Competitors versus Collaborators: Micronutrient Processing by Pathogenic and Commensal Human-Associated Gut Bacteria," *Mol Cell*, vol. 78, no. 4, pp. 570–576, May 2020, doi: 10.1016/j.molcel.2020.03.032.
- [350] M. Legein *et al.*, "Modes of Action of Microbial Biocontrol in the Phyllosphere," *Front Microbiol*, vol. 11, Jul. 2020, doi: 10.3389/fmicb.2020.01619.
- [351] M. Madhav, D. Baker, J. A. T. Morgan, S. Asgari, and P. James, "Wolbachia: A tool for livestock ectoparasite control," *Vet Parasitol*, vol. 288, p. 109297, Dec. 2020, doi: 10.1016/j.vetpar.2020.109297.

- 10.1016/j.vetpar.2020.109297.
- [352] Y. Shen and M. J. Loessner, “Beyond antibacterials – exploring bacteriophages as antivirulence agents,” *Curr Opin Biotechnol*, vol. 68, pp. 166–173, Apr. 2021, doi: 10.1016/j.copbio.2020.11.004.
- [353] J. Pernthaler, “Predation on prokaryotes in the water column and its ecological implications,” *Nat Rev Microbiol*, vol. 3, no. 7, pp. 537–546, Jul. 2005, doi: 10.1038/nrmicro1180.
- [354] N. Li *et al.*, “Analysis of gene gain and loss in the evolution of predatory bacteria,” *Gene*, vol. 598, pp. 63–70, Jan. 2017, doi: 10.1016/j.gene.2016.10.039.
- [355] D. W. Waite *et al.*, “Proposal to reclassify the proteobacterial classes Deltaproteobacteria and Oligoflexia, and the phylum Thermodesulfobacteria into four phyla reflecting major functional capabilities,” *Int J Syst Evol Microbiol*, vol. 70, no. 11, pp. 5972–6016, Nov. 2020, doi: 10.1099/ijsem.0.004213.
- [356] D. Kadouri and G. A. O’Toole, “Susceptibility of Biofilms to *Bdellovibrio bacteriovorus* Attack,” *Appl Environ Microbiol*, vol. 71, no. 7, pp. 4044–4051, Jul. 2005, doi: 10.1128/AEM.71.7.4044-4051.2005.
- [357] L. Hobley *et al.*, “Genome analysis of a simultaneously predatory and prey-independent, novel *Bdellovibrio bacteriovorus* from the River Tiber, supports in silico predictions of both ancient and recent lateral gene transfer from diverse bacteria,” *BMC Genomics*, vol. 13, no. 1, p. 670, Dec. 2012, doi: 10.1186/1471-2164-13-670.
- [358] E. G. Romanowski *et al.*, “Predatory bacteria are nontoxic to the rabbit ocular surface,” *Sci Rep*, vol. 6, no. 1, p. 30987, Aug. 2016, doi: 10.1038/srep30987.
- [359] P. H. F. Silva *et al.*, “The impact of predatory bacteria on experimental periodontitis,” *J Periodontol*, vol. 90, no. 9, pp. 1053–1063, Sep. 2019, doi: 10.1002/JPER.18-0485.
- [360] D. M. P. de Oliveira *et al.*, “Antimicrobial Resistance in ESKAPE Pathogens,” *Clin Microbiol Rev*, vol. 33, no. 3, Jun. 2020, doi: 10.1128/CMR.00181-19.
- [361] C. López-Causapé *et al.*, “Antibiotic resistance and population structure of cystic fibrosis *Pseudomonas aeruginosa* isolates from a Spanish multi-centre study,” *Int J Antimicrob Agents*, vol. 50, no. 3, pp. 334–341, Sep. 2017, doi: 10.1016/j.ijantimicag.2017.03.034.
- [362] M. García-Castillo *et al.*, “Wide Dispersion of ST175 Clone despite High Genetic Diversity of Carbapenem-Nonsusceptible *Pseudomonas aeruginosa* Clinical Strains in 16 Spanish Hospitals,” *J Clin Microbiol*, vol. 49, no. 8, pp. 2905–2910, Aug. 2011, doi: 10.1128/JCM.00753-11.
- [363] A. Fernández-Olmos *et al.*, “Population Structure and Antimicrobial Susceptibility of Both Nonpersistent and Persistent *Pseudomonas aeruginosa* Isolates Recovered from Cystic Fibrosis Patients,” *J Clin Microbiol*, vol. 51, no. 8, pp. 2761–2765, Aug. 2013, doi: 10.1128/JCM.00802-13.
- [364] M. Cristina, M. Sartini, and A. Spagnolo, “*Serratia marcescens* Infections in Neonatal Intensive Care Units (NICUs),” *Int J Environ Res*

- Public Health*, vol. 16, no. 4, p. 610, Feb. 2019, doi: 10.3390/ijerph16040610.
- [365] N. Y. Markelova, “Predacious bacteria, *Bdellovibrio* with potential for biocontrol,” *Int J Hyg Environ Health*, vol. 213, no. 6, pp. 428–431, Nov. 2010, doi: 10.1016/j.ijheh.2010.08.004.
- [366] G. Bonfiglio, B. Neroni, G. Radocchia, M. Marazzato, F. Pantanella, and S. Schippa, “Insight into the Possible Use of the Predator *Bdellovibrio bacteriovorus* as a Probiotic,” *Nutrients*, vol. 12, no. 8, p. 2252, Jul. 2020, doi: 10.3390/nu12082252.
- [367] C. Saralegui *et al.*, “Genomics of *Serratia marcescens* Isolates Causing Outbreaks in the Same Pediatric Unit 47 Years Apart: Position in an Updated Phylogeny of the Species,” *Front Microbiol*, vol. 11, Mar. 2020, doi: 10.3389/fmicb.2020.00451.
- [368] B. Pérez-Viso *et al.*, “Characterization of carbapenemase-producing *Serratia marcescens* and whole-genome sequencing for plasmid typing in a hospital in Madrid, Spain (2016–18),” *Journal of Antimicrobial Chemotherapy*, vol. 76, no. 1, pp. 110–116, Jan. 2021, doi: 10.1093/jac/dkaa398.
- [369] K. Rutherford *et al.*, “Artemis: sequence visualization and annotation,” *Bioinformatics*, vol. 16, no. 10, pp. 944–945, Oct. 2000, doi: 10.1093/bioinformatics/16.10.944.
- [370] C. Herencias, M. A. Prieto, and V. Martínez, “Determination of the predatory capability of *Bdellovibrio bacteriovorus* HD100,” *Bio Protoc*, vol. 7, no. 06, pp. 2–10, 2017, doi: 10.21769/BioProtoc.2177.
- [371] V. Martínez, E. Jurkevitch, J. L. García, and M. A. Prieto, “Reward for *Bdellovibrio bacteriovorus* for preying on a polyhydroxyalkanoate producer,” *Environ Microbiol*, vol. 15, no. 4, pp. 1204–1215, Apr. 2013, doi: 10.1111/1462-2920.12047.
- [372] J. M. Westergaard and T. T. Kramer, “*Bdellovibrio* and the intestinal flora of vertebrates,” *Appl Environ Microbiol*, vol. 34, no. 5, pp. 506–511, Nov. 1977, doi: 10.1128/aem.34.5.506-511.1977.
- [373] S. Gupta, C. Tang, M. Tran, and D. E. Kadouri, “Effect of Predatory Bacteria on Human Cell Lines,” *PLoS One*, vol. 11, no. 8, p. e0161242, Aug. 2016, doi: 10.1371/journal.pone.0161242.
- [374] A. K. Monnappa, W. Bari, S. Y. Choi, and R. J. Mitchell, “Investigating the Responses of Human Epithelial Cells to Predatory Bacteria,” *Sci Rep*, vol. 6, no. 1, p. 33485, Dec. 2016, doi: 10.1038/srep33485.
- [375] A. M. Rogosky, P. L. Moak, and E. A. B. Emmert, “Differential Predation by *Bdellovibrio bacteriovorus* 109J,” *Curr Microbiol*, vol. 52, no. 2, pp. 81–85, Feb. 2006, doi: 10.1007/s00284-005-0038-6.
- [376] E. G. Romanowski, S. Gupta, A. Pericleous, D. E. Kadouri, and R. M. Q. Shanks, “Clearance of Gram-Negative Bacterial Pathogens from the Ocular Surface by Predatory Bacteria,” *Antibiotics*, vol. 10, no. 7, p. 810, Jul. 2021, doi: 10.3390/antibiotics10070810.
- [377] S. F. Koval and M. E. Bayer, “Bacterial capsules: no barrier against *Bdellovibrio*,” *Microbiology (N Y)*, vol. 143, no. 3, pp. 749–753, Mar. 1997, doi:

10.1099/00221287-143-3-749.

[378] A. v. Akhova and A. G. Tkachenko, “ATP/ADP alteration as a sign of the oxidative stress development in *Escherichia coli* cells under antibiotic treatment,” *FEMS Microbiol Lett*, vol. 353, no. 1, pp. 69–76, Apr. 2014, doi: 10.1111/1574-6968.12405.

[379] M. Szenk, K. A. Dill, and A. M. R. de Graff, “Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis,” *Cell Syst*, vol. 5, no. 2, pp. 95–104, Aug. 2017, doi: 10.1016/j.cels.2017.06.005.

[380] M. A. Lobritz *et al.*, “Increased energy demand from anabolic-catabolic processes drives  $\beta$ -lactam antibiotic lethality,” *Cell Chem Biol*, vol. 29, no. 2, pp. 276–286.e4, Feb. 2022, doi: 10.1016/j.chembiol.2021.12.010.

[381] D. Sutton, P. G. Livingstone, E. Furness, M. T. Swain, and D. E. Whitworth, “Genome-Wide Identification of Myxobacterial Predation Genes and Demonstration of Formaldehyde Secretion as a Potentially Predation-Resistant Trait of *Pseudomonas aeruginosa*,” *Front Microbiol*, vol. 10, Nov. 2019, doi: 10.3389/fmicb.2019.02650.

[382] M. Knoll, T. M. Hamm, F. Wagner, V. Martinez, and J. Pleiss, “The PHA Depolymerase Engineering Database: A systematic analysis tool for the diverse family of polyhydroxyalkanoate (PHA) depolymerases,” *BMC Bioinformatics*, vol. 10, no. 1, p. 89, Dec. 2009, doi: 10.1186/1471-2105-10-89.

[383] M. Dori-Bachash, B. Dassa, S. Pietrokovski, and E. Jurkevitch,

“Proteome-Based Comparative Analyses of Growth Stages Reveal New Cell Cycle-Dependent Functions in the Predatory Bacterium *Bdellovibrio bacteriovorus*,” *Appl Environ Microbiol*, vol. 74, no. 23, pp. 7152–7162, Dec. 2008, doi: 10.1128/AEM.01736-08.

[384] S. C. Potter, A. Luciani, S. R. Eddy, Y. Park, R. Lopez, and R. D. Finn, “HMMER web server: 2018 update,” *Nucleic Acids Res*, vol. 46, no. W1, pp. W200–W204, Jul. 2018, doi: 10.1093/nar/gky448.

[385] M.-T. Manoli, J. Nogales, and A. Prieto, “Synthetic Control of Metabolic States in *Pseudomonas putida* by Tuning Polyhydroxyalkanoate Cycle,” *mBio*, vol. 13, no. 1, Feb. 2022, doi: 10.1128/mbio.01794-21.

[386] N. Sydney, M. T. Swain, J. M. T. So, E. Hoiczkyk, N. P. Tucker, and D. E. Whitworth, “The Genetics of Prey Susceptibility to Myxobacterial Predation: A Review, Including an Investigation into *Pseudomonas aeruginosa* Mutations Affecting Predation by *Myxococcus xanthus*,” *Microb Physiol*, vol. 31, no. 2, pp. 57–66, 2021, doi: 10.1159/000515546.

[387] R. J. Mitchell, W. Mun, S. S. Mabekou, H. Jang, and S. Y. Choi, “Compounds affecting predation by and viability of predatory bacteria,” *Appl Microbiol Biotechnol*, vol. 104, no. 9, pp. 3705–3713, May 2020, doi: 10.1007/s00253-020-10530-1.

[388] R. R. Nair, M. Vasse, S. Wielgoss, L. Sun, Y.-T. N. Yu, and G. J. Velicer, “Bacterial predator-prey coevolution accelerates genome

evolution and selects on virulence-associated prey defences,” *Nat Commun*, vol. 10, no. 1, p. 4301, Dec. 2019, doi: 10.1038/s41467-019-12140-6.

[389] S. Rendulic *et al.*, “A Predator Unmasked: Life Cycle of *Bdellovibrio bacteriovorus* from a Genomic Perspective,” *Science (1979)*, vol. 303, no. 5658, pp. 689–692, Jan. 2004, doi: 10.1126/science.1093027.

[390] E. Jurkevitch, D. Minz, B. Ramati, and G. Barel, “Prey Range

Characterization, Ribotyping, and Diversity of Soil and Rhizosphere *Bdellovibrio* spp. Isolated on Phytopathogenic Bacteria,” *Appl Environ Microbiol*, vol. 66, no. 6, pp. 2365–2371, Jun. 2000, doi: 10.1128/AEM.66.6.2365-2371.2000.

[391] M. P. Starr and R. J. Seidler, “The *Bdellovibrios*,” *Annu Rev Microbiol*, vol. 25, no. 1, pp. 649–678, Oct. 1971, doi: 10.1146/annurev.mi.25.100171.003245.




# **ANEXOS**



## ANEXOS

## Comité Ético de Investigación para los Capítulos 1 y 2 (3 páginas)

 <b>Hospital Universitario Ramón y Cajal</b> <small>SaludMadrid</small>	COMITÉ DE ÉTICA DE LA INVESTIGACIÓN
---	-------------------------------------

## DICTAMEN DEL COMITE DE ÉTICA DE LA INVESTIGACIÓN

Dña. M<sup>a</sup>. Ángeles Gálvez Mugica, Vicepresidenta del COMITÉ DE ÉTICA DE LA INVESTIGACIÓN DEL HOSPITAL UNIVERSITARIO RAMÓN Y CAJAL

## CERTIFICA

Que se ha evaluado la propuesta referida al Proyecto de Investigación:

TÍTULO:

**ESTABLECIMIENTO DE LA MICROBIOTA INTESTINAL Y PULMONAR EN NEONATOS CON FIBROSIS QUÍSTICA COMO FACTOR PRONÓSTICO PARA EL TRATAMIENTO ANTIMICROBIANO Y ANTIINFLAMATORIO.**

Versión Protocolo Julio 2017  
HIP/Ci versión Agosto 2017 tras aclaraciones CEI

- El estudio se plantea siguiendo los requisitos legalmente establecidos, y su realización es pertinente.
- Se cumplen los requisitos necesarios de idoneidad del protocolo en relación con los objetivos del estudio y están justificados los riesgos y molestias previsibles para el sujeto.
- El alcance de las compensaciones económicas previstas no interfiere con el respeto a los postulados éticos.
- La capacidad del investigador y sus colaboradores, y las instalaciones y medios disponibles, tal y como ha sido informado, son apropiados para llevar a cabo el estudio.

Este CEI actuando como comité evaluador, emite dictamen **FAVORABLE** y acepta que dicho estudio sea realizado en los centros siguientes por los investigadores principales que se relacionan a continuación:

**Dra. Rosa del Campo Moreno / Hospital Universitario Ramón y Cajal**

Y HACE CONSTAR QUE:


- 1° En la reunión celebrada el día **27/07/2017 ACTA 328** se decidió emitir el informe correspondiente al estudio de referencia.
- 2° En dicha reunión se cumplieron los requisitos establecidos en la legislación vigente -Real Decreto 1090/2015 y Decreto CAM 39/1994 de la Comunidad de Madrid- para que la decisión del citado CEI sea válida.
- 3° El CEI, tanto en su composición, como en los PNT cumple con las normas de BPC (CPMP/ ICH/ 135/95)
- 4° La composición actual del CEI es la siguiente:

<b>PRESIDENTE</b>	Sonsoles Sancho García.	Oncología Radioterápica
<b>VICEPRESIDENTE</b>	M <sup>a</sup> Ángeles Gálvez	Farmacología Clínica
<b>SECRETARIO TECNICO</b>	Itziar de Pablo López de Abechuco	Farmacología Clínica
<b>VOCALES</b>	M <sup>a</sup> Jesús Blanchard Rodríguez	Hematología
	Irene Cabrera Troye	Miembro Lego, ajeno a la investigación biomédica/asistencia clínica

Ctra. Colmenar, km. 9,100 - 28034 MADRID (Planta -2 Dcha)  
Tlf.: 91 336 8322

E-mail: [ceic.hrc@salud.madrid.org](mailto:ceic.hrc@salud.madrid.org)

1

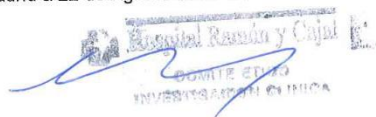
 <b>Hospital Universitario Ramón y Cajal</b> <small>SaludMadrid</small>	<b>COMITÉ DE ÉTICA DE LA INVESTIGACIÓN</b>
---	--


	Yolanda de Blas Fernández	DUE Servicio de Urgencias
	Guillermo García Ribas	Neurología
	Carmen Guillén Ponce	Oncología Médica. Comité Ética Asistencial. Miembro de la Comisión de Investigación
	Rosa Ana Hernández Antolín	Cardiología
	Manuel Luque Ramírez	Endocrinología y Nutrición. Miembro de la Comisión de Investigación
	María Muñoz García	Farmacia Hospitalaria.
	Alfonso Muriel García	Bioestadística Clínica
	Cristina Pueyo López	Farmacia Hospitalaria. Máster en Bioética
	Carmen Ruiz Ballesteros	Miembro Lego, ajeno a la investigación biomédica/asistencia clínica. Licenciada en Derecho
	Javier Sánchez Casajús	Miembro Lego, ajeno a la investigación biomédica/asistencia clínica
	José Miguel Sánchez Torres	Oncología Médica
	José Luis San Millán López	Genética Molecular
	Sonia Soto Díaz	Medicina Familiar y Comunitaria, Comisión Central de Investigación de la GAP
	Marta del Álamo Camuñas	Biología-IRYCIS

En el caso de que algún miembro participe en el estudio o declare algún conflicto de interés no habrá participado en la evaluación, ni en el dictamen de la solicitud de autorización del mismo.

Para que conste donde proceda, y a petición del promotor

Madrid a 22 de Agosto de 2017.



  
 COMITÉ DE ÉTICA  
 INVESTIGACIÓN CLÍNICA

Fdo.: D<sup>a</sup>. M<sup>a</sup>. Ángeles Gálvez Mugica  
Vicepresidenta CEI

**Conformidad de la Dirección del Centro**

**Dº. Juan José Equiza Escudero**, Director Gerente del Hospital "Ramón y Cajal" y vista la autorización del Comité de la Ética de la Investigación que ha evaluado el estudio

**CERTIFICA**

Que conoce la propuesta para que sea realizado en este Hospital el Proyecto de Investigación:

**ESTABLECIMIENTO DE LA MICROBIOTA INTESTINAL Y PULMONAR EN NEONATOS CON FIBROSIS QUÍSTICA COMO FACTOR PRONÓSTICO PARA EL TRATAMIENTO ANTIMICROBIANO Y ANTIINFLAMATORIO.**

**Versión Protocolo Julio 2017  
HIP/CI versión Agosto 2017 tras aclaraciones CEI**

Y que será realizado por la **Dra. Rosa del Campo Moreno**, como Investigadora Principal.

Servicio: **Microbiología.**

Que acepta la realización de dicho Proyecto de Investigación en este Hospital siempre y cuándo no implique ningún gasto adicional para el centro.

Lo que firma en Madrid a 22 de Agosto de 2017.

Hospital Universitario  
Ramón y Cajal  
CIENCIA

Fdo. Dº. Juan José Equiza Escudero

## Consentimiento informado a las familias del Capítulo 1 y 2 (6 páginas)

### HOJA DE INFORMACIÓN AL PACIENTE Y CONSENTIMIENTO INFORMADO

**Título del estudio:** Establecimiento de la Microbiota Intestinal y Pulmonar en Neonatos con Fibrosis Quística como Factor Pronóstico para el Tratamiento Antimicrobiano y Antiinflamatorio.

**Investigador principal:**

Rosa del Campo Moreno

Servicio de Microbiología, Hospital Universitario Ramón y Cajal, Madrid.

**Investigadores colaboradores:**

Rafael María Cantón Moreno, Jefe de Servicio, Hospital Universitario Ramón y Cajal

Marta Cobo Álvarez, Técnico de Laboratorio, Hospital Universitario Ramón y Cajal

María Garriga García. Nutricionista, Hospital Universitario Ramón y Cajal

Alejandro López Neyra, FEA Pediatría, Hospital Universitario Ramón y Cajal

Saioa Vicente Santamaría, FEA Pediatría, Hospital Universitario Ramón y Cajal

María Isabel Barrio, Jefe de Sección, Hospital Universitario La Paz

Marta Ruiz de Valbuena, FEA Pediatría, Hospital Universitario La Paz

Manuel Molina, Jefe de Sección, Hospital Universitario La Paz

Verónica Sanz Santiago, FEA Pediatría Hospital Niño Jesús

Carmen Luna Paredes, FEA Pediatría, Hospital Universitario 12 de Octubre

Enrique Salcedo Lobato, FEA Pediatría, Hospital Universitario 12 de Octubre

Antonio Salcedo Posadas, Jefe de Unidad de Neumología Pediátrica, Hospital Universitario Gregorio Marañón.

Isabel Delgado Pellecín, FEA Pediatría, Hospital Universitario Virgen del Rocío, Sevilla

Estela Pérez Ruiz, FEA Pediatría, Hospital Universitario Carlos Haya

**INTRODUCCIÓN**

Nos dirigimos a usted para informarle sobre el estudio en el que se le propone participar a su hijo. Nuestra intención es tan solo que usted reciba la información correcta y suficiente para que pueda evaluar y juzgar si quiere o no participar en este estudio. Para ello lea esta hoja informativa con atención y nosotros le aclararemos las dudas que le puedan surgir después de la explicación.

Su participación es voluntaria y puede revocar su decisión y retirar el consentimiento en cualquier momento sin que por ello se altere la relación con su médico ni se produzca perjuicio en sus cuidados médicos. En caso de retirar el consentimiento para participar en el estudio, ningún dato nuevo será añadido a la base de datos y puede exigir la destrucción de todas las muestras identificables previamente retenidas.

**FUNDAMENTO**

Las primeras bacterias que colonizan el tracto respiratorio y el intestinal de los recién nacidos establecen un consorcio microbiano que suele permanecer durante toda nuestra vida. En Fibrosis Quística, el pulmón se coloniza por patógenos como *Pseudomonas aeruginosa* y ello determina tanto la calidad como la esperanza de vida. El fundamento de este estudio es estudiar como se establece el microbioma en el pulmón y el intestino y determinar la colonización por patógenos para intentar aplicar medidas de modulación (antibióticos, probióticos y otros) antes de que se establezca el consorcio final, y optimar los tratamientos antibióticos y antiinflamatorios crónicos.

**¿QUÉ VAMOS A HACER?**

Lo que se le va a pedir por su participación en el estudio son muestras heces (para estudiar las bacterias del intestino) y de aspirados nasofaríngeos (para estudiar las bacterias del pulmón). Estas muestras se recogerán cuando usted venga a la consulta, aproximadamente cada dos meses, y hasta que su hijo cumpla 2 años de edad. La muestra de heces la podrá traer usted de su casa, mientras que el aspirado nasofaríngeo se recogerá durante la consulta y puede ser un poco molesto para su hijo. Todas las muestras serán identificadas mediante un código, congeladas inmediatamente y después se procesarán en tandas para determinar las bacterias

presentes en cada una de ellas. También le preguntaremos datos sobre la alimentación del niño que introduciremos en una base de datos junto con los datos clínicos de otras pruebas. Las muestras de este estudio no formarán parte de ninguna colección, y no se utilizarán con otros fines.

#### ***BENEFICIOS ESPERADOS E INCONVENIENTES***

En principio su hijo seguirá el mismo proceso de visitas a la consulta que realizaría si no participara en el estudio. El único inconveniente que tendrá será que se le recogerán las muestras de heces y de aspirado faríngeo. En el caso de que observemos que la composición de su microbiota está muy alterada, o detectemos patógenos, el doctor que le lleva puede adelantar el tratamiento antibiótico para intentar erradicarlos.

#### ***CONFIDENCIALIDAD***

Todos los datos recogidos para el estudio, procedentes de su Historia Clínica o facilitados por usted mismo, serán tratados con las medidas de seguridad establecidas en cumplimiento de la Ley Orgánica 15/1999 de Protección de Datos de carácter personal. Debe saber que tiene derecho de acceso, rectificación y cancelación de los mismos en cualquier momento.

Sólo aquellos datos de la historia clínica que estén relacionados con el estudio serán objeto de comprobación. Esta comprobación la realizará el Investigador Principal o los Investigadores Colaboradores, responsables de garantizar la confidencialidad de todos los datos de las historias clínicas pertenecientes a los sujetos participantes en el estudio. Los datos recogidos para el estudio estarán identificados mediante un código y sólo el Investigador Principal o los Investigadores Colaboradores podrán relacionar dichos datos con su nombre y con su historia clínica. Todos los procedimientos del estudio se llevarán a cabo según lo estipulado en la Ley de investigación biomédica 14/2007.

Debe usted saber que tiene derecho de acceso, rectificación y cancelación de todos los datos que se recojan para este estudio. Y para ejercer este derecho debe ponerse en contacto con el investigador principal en cualquier momento.

Cabe la posibilidad de que su caso sea interesante para nuestras investigaciones y que podamos solicitarle alguna muestra adicional de heces una vez que haya

finalizado el estudio. También puede que el algún momento solicitemos más información sobre su estado de salud o sobre el de alguno de sus familiares. En este caso se le contactará de nuevo y se le propondrá la ampliación del estudio.

Es importante que usted tenga claro que en todo momento se garantizará la máxima confidencialidad de sus datos, que no podrán pasar a personas ajenas a este proyecto. Además usted tiene el derecho de revocar su consentimiento a participar en este estudio en cualquier momento, de forma parcial o total.

Las muestras de su hijo, y fundamentalmente las bacterias presentes en ellas, serán almacenadas en el Servicio de Microbiología del Hospital Universitario Ramón y Cajal y utilizadas únicamente con fines de investigación. Mediante este consentimiento usted cede sus muestras sin ningún tipo de beneficio directo o indirecto sobre ellas. Sus muestras y las bacterias que de ellas se aíslen, una vez finalizado el presente proyecto pasarán a integrar colecciones de investigaciones que únicamente se utilizarán en proyectos relacionados con la propuesta inicial y nunca serán cedidas a personas ajenas a este proyecto.

#### **OTROS ASPECTOS DE INTERÉS**

Este proyecto está por la Fundación Mutua Madrileña (2017-2019), y por el Fondo de Investigaciones Sanitarias (PI17/00115) y en el participan Hospitales de la Comunidad de Madrid: H. U. 12 de Octubre, H. U. Gregorio Marañón, H. U. La Paz y H. U. Ramón y Cajal, y de otras comunidades: H. U. Virgen del Rocío de Sevilla, y H. U. Carlos Haya de Málaga. Todas las muestras serán procesadas en el H. U. Ramón y Cajal por el mismo equipo de investigación la finalidad de estas investigaciones es ampliar el conocimiento científico de la interacción entre los microorganismos y la salud humana. Todos los resultados que se obtengan serán publicados en congresos especializados en el tema y en revistas científicas de la mayor repercusión posible, para que estos conocimientos contribuyan al mejor tratamientos de otros niños que padezcan la misma enfermedad. En caso de necesitar cualquier información o por cualquier motivo no dude en contactar con la investigadora principal del estudio (Dra. Rosa del Campo, Teléfono: 913368542 de lunes a viernes de 8.00-15.00). En caso de dudas respecto a sus derechos debe dirigirse al Servicio de Atención al Paciente del Hospital Universitario Ramón y Cajal.

**Firma del paciente**

**Firma del investigador**

**Nombre:**

**Nombre:**

**Fecha:**

**Fecha:**

**Este documento se firmará por duplicado quedándose una copia el investigador y otra el paciente**

**MODELO DE CONSENTIMIENTO INFORMADO POR ESCRITO**

Título del Estudio: Establecimiento de la Microbiota Intestinal y Pulmonar en Neonatos con Fibrosis Quística como Factor Pronóstico para el Tratamiento Antimicrobiano y Antiinflamatorio.

Promotor: Fundación Mutua Madrileña

Yo (nombre y apellidos participante).....

He leído la hoja de información que se me ha entregado.

He podido hacer preguntas sobre el estudio.

He recibido suficiente información sobre el estudio.

He hablado con:

.....(nombre del investigador)

Comprendo que mi participación es voluntaria.

Comprendo que puedo retirarme del estudio:

1º Cuando quiera

2º Sin tener que dar explicaciones.

3º Sin que esto repercuta en mis cuidados médicos.

Presto libremente mi conformidad para participar en el estudio.

Estoy de acuerdo en que mis muestras sean utilizadas junto con los datos clínicos y analíticos de mi enfermedad SOLO PARA ESTE PROYECTO: SI  NO .

Indique si tiene alguna restricción concreta para el uso de las muestras donadas:

NO  SI  ¿Cuál?.....

¿Querría usted recibir información derivada de los datos del estudio que pudiera ser relevante para su salud o la de sus familiares? SI  NO

Teléfono:.....Correo electrónico:.....

FECHA:

FIRMA DEL PARTICIPANTE

FECHA:

FIRMA DEL INVESTIGADOR

## Datos sobre los tratamientos recibidos por la cohorte del capítulo 1

Hospital	Paciente	Tratamiento	Vía	Posología	Comienzo	Fin	Días	Probióticos	Ciclos totales	Total días en tratamiento
1	Paciente_1	TMP-SMX	oral	4,5ml/12h	10/12/2018	24/12/2018	14	sí	2	28
1	Paciente_1	TMP-SMX	oral	4,5ml/12h	12/04/2019	26/04/2019	14	sí	2	28
3	Paciente_2	HTSS	inhulado		27/02/2019			no	2	28
3	Paciente_2	BA	inhulado		27/02/2019			no	2	28
3	Paciente_2	MEM	intravenoso	160mg/8h	13/03/2019	27/03/2019	14	no	2	28
3	Paciente_2	TB	intravenoso	45mg/24h	13/03/2019	27/03/2019	14	no	2	28
3	Paciente_3	HTSS	inhulado		08/04/2018			sí	6	86
3	Paciente_3	BA	inhulado		08/04/2018			sí	6	86
3	Paciente_3	AMX-CLV	oral	100mg/8h	28/04/2018	01/05/2018	3	sí	6	80
3	Paciente_3	CEC	oral	50mg/8h	02/05/2018	30/05/2018	28	sí	6	80
3	Paciente_3	GCE	inhulado		11/05/2018	17/05/2018	6	sí	6	86
3	Paciente_3	AMX-CLV	oral	100mg/8h	02/07/2018	16/07/2018	14	sí	6	80
3	Paciente_3	AMX-CLV	oral	160mg/8h	10/10/2018	24/10/2018	14	sí	6	80
3	Paciente_3	AMX-CLV	oral	160mg/8h	10/01/2019	24/01/2019	14	sí	6	80
3	Paciente_3	AMX-CLV	oral	180mg/8h	01/04/2019	08/04/2019	7	sí	6	80
3	Paciente_4	HTSS	inhulado		28/08/2018			no	2	43
3	Paciente_4	BA	inhulado		28/08/2018			no	2	43
3	Paciente_4	AMX-CLV	oral	160mg/8h	28/08/2018	12/09/2018	15	no	2	43
3	Paciente_4	AZT	inhulado	75mg/8h	20/09/2018	18/10/2018	28	no	2	43
3	Paciente_5	HTSS	inhulado		06/11/2018			no	2	30
3	Paciente_5	BA	inhulado		06/11/2018			no	2	30
3	Paciente_5	AMX-CLV	oral	150mg/8h	14/11/2018	29/11/2018	15	no	2	30
3	Paciente_5	AMX-CLV	oral	150mg/8h	02/01/2019	17/01/2019	15	no	2	30
3	Paciente_6	AMX-CLV	oral	140mg/8h	09/10/2018	23/10/2018	14	no	4	58
3	Paciente_6	AMX-CLV	oral	140mg/8h	07/11/2018	22/11/2018	15	no	4	58
3	Paciente_6	TMP-SMX	oral	32mg/12h	09/01/2019	23/01/2019	14	no	4	58
3	Paciente_6	TMP-SMX	oral	35mg/12h	15/05/2019	30/05/2019	15	no	4	58
3	Paciente_6	HTSS	inhulado		21/08/2019			no	4	58
3	Paciente_6	BA	inhulado		21/08/2019			no	4	58
2	Paciente_7	BA	inhulado		01/03/2018			no	2	20
2	Paciente_7	HTSS	inhulado		01/03/2018			no	2	20
2	Paciente_7	CXM	oral	250mg/12h	01/06/2018	15/06/2018	14	no	2	20
2	Paciente_7	CXM	oral	250mg/12h	31/07/2018	06/08/2018	6	no	2	20
2	Paciente_8	BA	inhulado		11/04/2018			no	3	39
2	Paciente_8	HTSS	inhulado		11/04/2018			no	3	39
2	Paciente_8	AMX-CLV	oral	100mg/8h	30/04/2018	14/05/2018	14	no	3	39
2	Paciente_8	AMX-CLV	oral	160mg/8h	15/07/2018	25/07/2018	10	no	3	39
2	Paciente_8	TMP-SMX	oral	32mg/12h	01/03/2019	16/03/2019	15	no	3	39

TMP-SMX: trimetoprim-sulfametoxazol; HTSS: suero salino hipertónico; BA: beta-adrenérgicos MEM: meropenem; TB: trobramicina; AMX-CLV: amoxicilina-ácido clavulánico; AZT: aztreonam; CXM: cefuroxima

## Datos clínicos sobre exacerbaciones, ingresos y tratamientos antibióticos de los pacientes del capítulo 1 y 2

Color de relleno	Significado
Ingreso	Ingreso por exacerbación respiratoria
Exac. Resp.	Exacerbación respiratoria
Ingreso	Ingreso por otra razón
Malestar	Otro tipo de síntomas no respiratorios
	Ciclos de tobramicina-colisimetato

Caso	Edad (meses)	Tipo de evento	Duración (días)	Comentario	Tto. 1	Tto. 2	Tto. 3
#1	2	patógeno		<i>S. marcescens</i>			
#1	2	patógeno		<i>S. aureus</i>			
#1	4	patógeno, ingreso exac. resp.	8	<i>S. marcescens</i>	TMP-SMX (oral)		
#1	5	exac. resp.					
#1	8	patógeno, exac. resp.		<i>S. marcescens</i>	GCE (INH)	TMP-SMX (oral)	
#1	9	tratamiento		infección en la boca	GCE (INH)	Miconazol (topica)	
#1	10	exac. resp.			TMP-SMX (oral)		
#1	13	tratamiento			AMX-CLV (oral)		
#1	14	tratamiento			AMX-CLV (oral)		
#1	15	tratamiento			AMX-CLV (oral)		
#1	16	exac. resp.			TMP-SMX (oral)	Tobramicina (ocular)	
#1	19	tratamiento			AMX-CLV (oral)		
#2	1	ingreso, exac. no resp.	14	fallo de medro, tos	Teicoplanina (IV)	Tobramicina (IV)	
#2	2	ingreso, exac. no resp.	14	fallo de medro, tos			
#2	7	exac. resp.	1	fiebre, tos	AMX-CLV (oral)		
#2	9	patógeno, exac. resp.		<i>P. aeruginosa</i>	Cefaclor (oral)	Ciprofloxacina (oral)	
#2	10	patógeno		<i>S. aureus</i> , <i>H. influenzae</i> , <i>H. parainfluenzae</i>	Ciclo tobra-colisimetato inhalados (6 meses)		
#2	12	patógeno		<i>P. aeruginosa</i>	Cefaclor (oral)		
#2	18	patógeno		<i>P. aeruginosa</i>	Ciprofloxacina (oral)		
#2	23	patógeno		SARS-CoV-2			
#2	25	patógeno		<i>S. aureus</i>	AMX-CLV (oral)		

(Continuación)

Caso	Edad (meses)	Tipo de evento	Duración (días)	Comentario	Tto. 1	Tto. 2	Tto. 3
#3	1	patógeno, ingreso exac. resp.	5	VRS			
#3	2	patógeno, exac. resp.		fiebre, sibilancia, VRS, <i>S. aureus</i>			
#3	3	tratamiento			AMX-CLV (oral)	Cefaclor (oral)	GCE (INH)
#3	5	patógeno, exac. resp.		fiebre, sibilancia, VRS	AMX-CLV (oral)		
#3	8	patógeno, exac. resp.		fiebre, sibilancia, VRS	AMX-CLV (oral)		
#3	11	patógeno, exac. resp.		fiebre, sibilancia, VRS	AMX-CLV (oral)		
#3	14	virus, otro malestar		fiebre, VRS	AMX-CLV (oral)		
#3	17	virus, otro malestar		fiebre, otalgia, VRS	AMX-CLV (oral)		
#3	19	patógeno		<i>P. aeruginosa</i>	AMX-CLV (oral)	Ciprofloxacina (oral)	Tobramicina (INH)
#3	21	patógeno, exac. resp.		<i>P. aeruginosa</i>	Ciprofloxacina (oral)	Colismetato (INH)	
#4	6	patógeno		<i>S. aureus</i> , <i>P. aeruginosa</i>	AMX-CLV (oral)	Aztreonam (INH)	
#4	8	patógeno		<i>P. aeruginosa</i>	Ciprofloxacina (oral)	Colismetato (INH)	
#4	9	patógeno		<i>P. aeruginosa</i>	Ciprofloxacina (oral)	Tobramicina (INH)	
#4	10	tratamiento			Ciclo tobra-colismetato inhalados (13 meses)		
#4	12	patógeno, ingreso exac. resp.	1	<i>P. aeruginosa</i> , cefta, tobra, IV	Ceftazidima (IV)	Colismetato (INH)	Tobramicina (IV)
#4	15	tratamiento			Ciprofloxacina (oral)	Colismetato (INH)	
#5	3	patógeno, ingreso exac. resp.	1	<i>H. influenzae</i> , moco, deshidratacion	AMX-CLV (oral)		
#5	5	patógeno, exac. resp.		<i>H. influenzae</i>	AMX-CLV (oral)		
#5	11	exac. resp.			AMX-CLV (oral)		
#5	17	exac. resp.		tos, conjuntivitis	AMX-CLV (oral)		
#6	1	patógeno, exac. resp.		<i>S. aureus</i>	AMX-CLV (oral)	AMX-CLV (oral)	
#6	4	exac. resp.			TMP-SMX (oral)		
#6	8	exac. resp.			TMP-SMX (oral)		
#6	12	ingreso, exac. no resp.	2	deshidratacion	Cefaclor (oral)		
#6	13	exac. resp.					
#6	13	exac. resp.					
#6	21	patógeno		<i>S. aureus</i> , <i>E. coli</i>			

(Continuación)

Caso	Edad (meses)	Tipo de evento	Duración (días)	Comentario	Tto. 1	Tto. 2	Tto. 3
#7	4	exac. resp.			Cefuroxima (oral)		
#7	6	exac. resp.			Cefuroxima (oral)		
#7	24	patógeno		<i>P. aeruginosa</i>			
#8	3	exac. resp.			AMX-CLV (oral)		
#8	6	exac. resp.			AMX-CLV (oral)		
#8	7	patógeno		<i>S. aureus</i>			
#8	9	patógeno		<i>S. aureus</i>			
#8	13	patógeno, exac. resp.		<i>S. aureus</i>	TMP-SMX (oral)		
#8	17	patógeno		<i>S. aureus</i>			
#8	20	patógeno		<i>S. aureus</i>			
#8	24	patógeno		<i>S. aureus, P. aeruginosa</i>			
#9	3	ingreso exac. resp.	5	febril, broncoespasmo	Cefotaxima (IV)		
#9	4	ingreso exac. resp.	3	febril, broncoespasmo	AMX-CLV (oral)	Azitromicina (oral)	
#9	5	patógeno, exac. resp.		<i>C. indologenes, afebril</i>	AMX-CLV (oral)		
#9	6	ingreso exac. resp.	3	riesgo vital, hipotonía, tos, emetizante	Cefaclor (oral)		
#9	7	tratamiento			Cefaclor (oral)		
#9	8	tratamiento			Cefaclor (oral)		
#9	9	patógeno		<i>P. aeruginosa</i>	Cefaclor (oral)	Ciprofloxacina (oral)	Colismetato (INH)
#9	10	tratamiento			Cefaclor (oral)		
#9	11	tratamiento			Cefaclor (oral)		
#9	12	tratamiento			Cefaclor (oral)		
#9	13	tratamiento			Cefaclor (oral)		
#9	14	tratamiento			Cefaclor (oral)		
#9	15	tratamiento			Cefaclor (oral)		
#9	16	exac. resp.		afebril	Cefaclor (oral)	Ciprofloxacina (oral)	
#9	21	patógeno, exac. resp.		<i>S. aureus</i>	Azitromicina (oral)	AMX-CLV (oral)	
#9	22	patógeno		<i>P. aeruginosa, S. aureus</i>	Colismetato (INH)	Ciprofloxacina (oral)	

(Continuación)

Caso	Edad (meses)	Tipo de evento	Duración (días)	Comentario	Tto. 1	Tto. 2	Tto. 3
#10	1	patógeno, ingreso, exac. no resp.	1	<i>H. influenzae</i> , <i>M. catarrhalis</i> , fallo de medro, vomitos	Cefaclor (oral)		
#10	2	tratamiento			Cefaclor (oral)		
#10	6	tratamiento			Cefaclor (oral)		
#10	8	patógeno		<i>H. influenzae</i>			
#10	9	patógeno, exac. resp.		<i>C. indologenes</i> , <i>P. aeruginosa</i>	Cefaclor (oral)		
#10	10	tratamiento			Cefaclor (oral)	Ciprofloxacina (oral)	Colismetato (INH)
#10	11	tratamiento			Cefaclor (oral)	Colismetato (INH)	
#10	12	tratamiento			Cefaclor (oral)	Colismetato (INH)	
#10	13	tratamiento			Cefaclor (oral)	Colismetato (INH)	
#10	14	tratamiento			Cefaclor (oral)	Colismetato (INH)	
#10	15	tratamiento			Colismetato (INH)	Colismetato (INH)	
#10	16	patógeno	1	<i>S. aureus</i>	RF-SMX (oral)	Colismetato (INH)	
#10	17	tratamiento			Colismetato (INH)		
#10	18	tratamiento			Colismetato (INH)		
#10	19	tratamiento			Colismetato (INH)		
#10	24	patógeno		<i>S. aureus</i>	AMX-CLV (oral)		
#11	1	ingreso, exac. no resp.	30	ileo, meconial	AMP-GENTA (IV)	AMX-CLV (oral)	
#11	1	ingreso, exac. no resp.	6	fallo de medro			
#11	1	patógeno, ingreso, exac. no resp.	2	<i>K. oxytoca</i> , <i>K. pneumoniae</i> , <i>S. marcescens</i> , diarrea	Meropenem (IV)	VAN-AMK (IV)	
#11	3	tratamiento			Reuteri (oral)		
#11	4	tratamiento			Reuteri (oral)		
#11	5	tratamiento			Reuteri (oral)		
#11	6	tratamiento			Reuteri (oral)		
#12	24	patógeno		<i>S. marcescens</i>			

(Continuación)

Caso	Edad (meses)	Tipo de evento	Duración (días)	Comentario	Tto. 1	Tto. 2	Tto. 3
#13	1	patógeno		<i>S. aureus</i>	AMX-CLV (oral)	Cefaclor (oral)	
#13	13	patógeno		<i>S. aureus</i>	AMX-CLV (oral)		
#13	17	patógeno		<i>S. aureus</i>	AMX-CLV (oral)	AMX-CLV (oral)	
#13	20	patógeno, exac. resp.		<i>P. aeruginosa</i>	AMX-CLV (oral)	Colismetato (INH)	
#13	20	exac. resp.			Ciprofloxacina (oral)	AMX-CLV (oral)	
#13	21	patógeno, exac. resp.		<i>S. maltophilia</i>	AMX-CLV (oral)		
#13	22	patógeno, exac. resp.		<i>P. aeruginosa</i>	AMX-CLV (oral)		
#13	23	patógeno, exac. resp.		<i>E. coli</i>	AMX-CLV (oral)		
#14	1	ingreso, exac. no resp.	18	ileo, meconial			
#14	3	patógeno, ingreso exac. resp.	7	<i>M. catarrhalis, P. aeruginosa, dificultad respiratoria</i>	Cefaclor (oral)	Ceftazidima (IV)	Ciprofloxacina (oral)
#14	3	tratamiento			Colismetato (INH)	Tobramicina (IV)	
#14	4	tratamiento			Colismetato (INH)		
#14	5	tratamiento			Colismetato (INH)		
#14	6	exac. resp.			Colismetato (INH)		
#14	7	tratamiento			Colismetato (INH)		
#14	8	tratamiento			Colismetato (INH)		
#14	11	patógeno, ingreso, exac. no resp.		<i>C. jejuni, deshidratación</i>			
#14	12	exac. resp.			AMX-CLV (oral)		
#14	14	patógeno		<i>P. aeruginosa, S. maltophilia</i>	Ciclo tobra-colismetato inhalados (9 meses)		
#14	15	tratamiento			Ciprofloxacina (oral)		
#14	22	patógeno		<i>S. aureus, S. maltophilia</i>			

## Informe STORMS del CAPÍTULO 1

Version: 1,03						
#	Item	Recommendation	Item Source	Additional Guidance	Yes/No/NA	Comments or location in manuscript
<b>Abstract</b>						
1,0	Structured or Unstructured Abstract	Abstract should include information on background, methods, results, and conclusions in structured or unstructured format.	STORMS		Yes	p.2
1,1	Study Design	State study design in abstract.	STORMS	See 3.0 for additional information on study design.	Yes	p.2
1,2	Sequencing methods	State the strategy used for metagenomic classification.	STORMS	For example, targeted 16S by qPCR or sequencing, shotgun metagenomics, metatranscriptomics, etc.	Yes	p.2
1,3	Specimens	Describe body site(s) studied.	STORMS		Yes	p.2
<b>Introduction</b>						
2,0	Background and Rationale	Summarize the underlying background, scientific evidence, or theory driving the current hypothesis as well as the study objectives.	STORMS		Yes	p.3-5
2,1	Hypotheses	State the pre-specified hypothesis. If the study is exploratory, state any pre-specified study objectives.	STORMS		Yes	p.5
<b>Methods</b>						
3,0	Study Design	Describe the study design.	STORMS	Observational (Case-Control, Cohort, Cross-sectional survey, etc.) or Experimental (Randomized controlled trial, Non-randomized controlled trial, etc.). For a brief description of common study designs see: DOI: 10.11613/BM.2014.022  If applicable, describe any blinding (e.g. single or double-blinding) used in the course of the study.	Yes	p.5
3,1	Participants	State what the population of interest is, and the method by which participants are sampled from that population. Include relevant information on physiological state of the subjects or stage in the life history of disease under study when participants were sampled.	STORMS	Examples of the population of interest could be: adults with no chronic health conditions, adults with type II diabetes, newborns, etc. This is the total population to whom the study is hoped to be generalizable to. The sampling method describes how potential participants were selected from that population.  If the participants are from a substudy of a larger study, provide a brief description of that study and cite that study.  Clearly state how cases and controls are defined.  An example of relevant physiological state might be pre/post menopausal for a vaginal microbiome study; examples of stage in the life history of disease could be whether specimens were collected during active or dormant disease, or before or after treatment.	Yes	p.5

3,2	Geographic location	State the geographic region(s) where participants were sampled from.	MixS: geographic location (country and/or sea, region)	Geographic coordinates can be reported to prevent potential ambiguities if necessary.	Yes	p.5 (country level)
3,3	Relevant Dates	State the start and end dates for recruitment, follow-up, and data collection.	STORMS	Recruitment is the period in which participants are recruited for the study. In longitudinal studies, follow-up is the date range in which participants are asked to complete a specific assessment. Finally, data collection is the total period in which data is being collected from participants including during initial recruitment through all follow-ups.	NA	
3,4	Eligibility criteria	List any criteria for inclusion and exclusion of recruited participants.	Modified STROBE	Among potential recruited participants, how were some chosen and others not? This could include criteria such as sex, diet, age, health status, or BMI.  If there is a primary and validation sample, describe inclusion/exclusion criteria for each.	Yes	p.5 (inclusion only)
3,5	Antibiotics Usage	List what is known about antibiotics usage before or during sample collection.	STORMS	If participants were excluded due to current or recent antibiotics usage, state this here.  Other factors (e.g. proton pump inhibitors, probiotics, etc.) that may influence the microbiome should also be described as well.	Yes	Additional file 1
3,6	Analytic sample size	Explain how the final analytic sample size was calculated, including the number of cases and controls if relevant, and reasons for dropout at each stage of the study. This should include the number of individuals in whom microbiome sequencing was attempted and the number in whom microbiome sequencing was successful.	STORMS	Consider use of a flow diagram (see template at <a href="https://stormsmicrobiome.org/figures">https://stormsmicrobiome.org/figures</a> ). Also state sample size in abstract.  If power analysis was used to calculate sample size, describe those calculations.	Yes	p.5-6
3,7	Longitudinal Studies	For longitudinal studies, state how many follow-ups were conducted, describe sample size at follow-up by group or condition, and discuss any loss to follow-up.	STORMS	If there is loss to follow-up, discuss the likelihood that drop-out is associated with exposures, treatments, or outcomes of interest.	Yes	p.5
3,8	Matching	For matched studies, give matching criteria.	Modified STROBE	"Matched" refers to matching between comparable study participants as cases and controls or exposed / unexposed.  Indicate whether participants were individual or frequency matched and in what ratio were they matched (e.g. 1 case to 1 control).	Yes	p.5-9
3,9	Ethics	State the name of the institutional review board that approved the study and protocols, protocol number and date of approval, and procedures for obtaining informed consent from participants.	STORMS		No	

4,0	Laboratory methods	State the laboratory/center where laboratory work was done.	STORMS	Provide a reference to complete lab protocols if previously published elsewhere such as on protocols.io. Note any modifications of lab protocols and the reason for protocol modifications.	Yes	p.6
4,1	Specimen collection	State the body site(s) sampled from and how specimens were collected.	MixS: sample collection device or method; host body site	Use terms from the Uber-anatomy Ontology ( <a href="https://www.ebi.ac.uk/ols/ontologies/uberon">https://www.ebi.ac.uk/ols/ontologies/uberon</a> ) to describe body sites in a standardized format.	Yes	p.5
4,2	Shipping	Describe how samples were stored and shipped to the laboratory.	STORMS	Include length of time from collection to receipt by the lab and if temperature control was used during shipping.	Yes	p.5
4,3	Storage	Describe how the laboratory stored samples, including time between collection and storage and any preservation buffers or refrigeration used.	STORMS	State where each procedure or lot of samples was done if not all in the same place.  Include reagent/lot/catalogue #s for storage buffers.	Yes	p.5
4,4	DNA extraction	Provide DNA extraction method, including kit and version if relevant.	MixS: nucleic acid extraction	If any DNA quantification methods were used prior to DNA amplification or at the pooling step of library preparation, state so here.	Yes	p.6
4,5	Human DNA sequence depletion or microbial DNA enrichment	Describe whether human DNA sequence depletion or enrichment of microbial or viral DNA was performed.	STORMS		Yes	p.6-7
4,6	Primer selection	Provide primer selection and DNA amplification methods as well as variable region sequenced (if applicable).	MixS: pcr primers		Yes	p.6
4,7	Positive Controls	Describe any positive controls (mock communities) if used.	STORMS	If used, should be deposited under guidance provided in the 8.X items.	NA	
4,8	Negative Controls	Describe any negative controls if used.	STORMS	If used, should be deposited under guidance provided in the 8.X items.	NA	
4,9	Contaminant mitigation and identification	Provide any laboratory or computational methods used to control for or identify microbiome contamination from the environment, reagents, or laboratory.	STORMS	Includes filtering of reagents and other steps to minimize contamination. It is relevant to state whether the specimens of interest have low microbial load, which makes contamination especially relevant.	NA	
4,10	Replication	Describe any biological or technical replicates included in the sequencing, including which steps were replicated between them.	STORMS	Replication may be biological (redundant biological specimens) or technical (aliquots taken at different stages of analysis) and used in extraction, sequencing, preprocessing, and/or data analysis.	No	
4,11	Sequencing strategy	Major divisions of strategy, such as shotgun or amplicon sequencing.	MixS: sequencing method	For amplicon sequencing (for example, 16S variable region), state the region selected. State the model of sequencer used.	Yes	p.6

4,12	Sequencing methods	State whether experimental quantification was used (QMP/cell count based, spike-in based) or whether relative abundance methods were applied.	STORMS	These include read length, sequencing depth per sample (average and minimum), whether reads are paired, and other parameters.	Yes	p.6-9
4,13	Batch effects	Detail any blocking or randomization used in study design to avoid confounding of batches with exposures or outcomes. Discuss any likely sources of batch effects, if known.	STORMS	Sources of batch effects include sample collection, storage, library preparation, and sequencing and are commonly unavoidable in all but the smallest of studies.	No	
4,14	Metatranscriptomics	Detail whether any mRNA enrichment was performed and whether/how retrotranscription was performed prior to sequencing. Provide size range of isolated transcripts. Describe whether the sequencing library was stranded or not. Provide details on sequencing methods and platforms.	STORMS	Provide details on any internal standards which may have been used as well as parameters and versions of any software or databases used.	NA	
4,15	Metaproteomics	Detail which protease was used for digestion. Provide details on proteomic methods and platforms (e.g. LC-MS/MS, instrument type, column type, mass range, resolution, scan speed, maximum injection time, isolation window, normalised collision energy, and resolution).	STORMS	Provide details on any internal standards which may have been used as well as parameters and versions of any software or databases used.	Yes	p.6-7
4,16	Metabolomics	Specify the analytic method used (such as nuclear magnetic resonance spectroscopy or mass spectrometry). For mass spectrometry, detail which fractions were obtained (polar and/or non polar) and how these were analyzed. Provide details on metabolomics methods and platforms (e.g. derivatization, instrument type, injection type, column type and instrument settings).	STORMS	Provide details on any internal standards which may have been used as well as parameters and versions of any software or databases used.	NA	

5,0	Data sources/ measurement	For each non-microbiome variable, including the health condition, intervention, or other variable of interest, state how it was defined, how it was measured or collected, and any transformations applied to the variable prior to analysis.	MixS: host disease status	State any sources of potential bias in measurements, for example multiple interviewers or measurement instruments, and whether these potential biases were assessed or accounted for in study design.  Use terms from a standardized ontology such as the Experimental Factor Ontology ( <a href="https://www.ebi.ac.uk/efo/">https://www.ebi.ac.uk/efo/</a> ) to describe variables of interest in a standardized format.	Yes	p.5
6,0	Research design for causal inference	Discuss any potential for confounding by variables that may influence both the outcome and exposure of interest. State any variables controlled for and the rationale for controlling for them.	STORMS	For causal inference, this item refers to describing the assumptions that would be required to draw causal inferences from observational data. See Vujkovic-Cvijin, I., Sklar, J., Jiang, L. et al. Host variables confound gut microbiota studies of human disease. <i>Nature</i> 587, 448–454 (2020). <a href="https://doi.org/10.1038/s41586-020-2881-9">https://doi.org/10.1038/s41586-020-2881-9</a> for more details on confounding in observational microbiome studies.  For example, hypothesized confounders may be controlled for by multivariable adjustment. Consider using a directed acyclic graph (DAG) to describe your causal model and justify any variables controlled for. DAGs can be made using <a href="http://www.dagitty.net">www.dagitty.net</a> .	NA	
6,1	Selection bias	Discuss potential for selection or survival bias.	STORMS	Selection bias can occur when some members of the target study population are more likely to be included in the study/final analytic sample than others. Some examples include survival bias (where part of the target study population is more likely to die before they can be studied), convenience sampling (where members of the target study population are not selected at random), and loss to follow-up (when probability of dropping out is related to one of the things being studied).	NA	
7,0	Bioinformatic and Statistical Methods	Describe any transformations to quantitative variables used in analyses (e.g. use of percentages instead of counts, normalization, rarefaction, categorization).	STORMS	If a variable is analyzed using different transformations, state rationale for the transformation and for each analyses which version of the variable is used.  In case of any complex or multistep transformations, give enumerated instructions for reproducing those transformations.	Yes	p.8-9
7,1	Quality Control	Describe any methods to identify or filter low quality reads or samples.	MixS: sequence quality check	If samples were excluded based on quality or read depth, list the criteria used, the number of samples excluded, and the final sample size after quality control.	Yes	p.6-9
7,2	Sequence analysis	Describe any taxonomic, functional profiling, or other sequence analysis performed.	MixS: feature prediction; similarity search method		Yes	p.8-9

7,3	Statistical methods	Describe all statistical methods.	Modified STROBE	Describe any statistical tests used, exploratory data analysis performed, dimension reduction methods/unsupervised analysis, alpha/beta metrics, and/or methods for adjusting for measurement bias.  If multiple statistical methods are possible, discuss why the methods used were selected.  If a multiple hypothesis testing correction method was used, describe the type of correction used.  State which taxonomic levels are analyzed.	Yes	p.8-9
7,4	Longitudinal analysis	If the study is longitudinal, include a section that explicitly states what analysis methods were used (if any) to account for grouping of measurements by individual or patterns over time.	STORMS		Yes	p.8-9
7,5	Subgroup analysis	Describe any methods used to examine subgroups and interactions.	STROBE		NA	
7,6	Missing data	Explain how missing data were addressed.	STROBE	"Missing data" refers to participant measurements such as covariates, exposures, outcomes, or time points that should have been collected but were not, not to zeros in taxonomic abundance tables or data points not applicable to that observation.	Yes	p.5
7,7	Sensitivity analyses	Describe any sensitivity analyses.	STROBE		NA	
7,8	Findings	State criteria used to select findings for reporting.	STORMS	For example, false discovery rate with total number of tests, effect size threshold, significance threshold, microbes of interest.	Yes	p.8-9
7,9	Software	Cite all software (including read mapping software) and databases (including any used for taxonomic reference or annotating amplicons, if applicable) used. Include version numbers.	Modified STREGA	Installed packages, add-ons or libraries should be stated and cited in addition to the software used.  All parameters employed that differ from the default of that software/version should be provided.  This is in addition to, not a replacement for, publishing of code as outlined in the section Reproducible Research.	Yes	p.6-9
8,0	Reproducible research	Make a statement about whether and how others can reproduce the reported analysis.	STORMS	Any protected information that has been excluded or provided under controlled access should be listed along with any relevant data access procedures. "On request from authors" is not sufficiently detailed; formal data access procedures and conditions should be defined.  If data are unavailable, state so clearly.  Consider using a specialized rubric for reproducible research (such as: <a href="https://mbio.asm.org/content/9/3/e00525-18.short">https://mbio.asm.org/content/9/3/e00525-18.short</a> ).  Consider preregistering the study protocol (such as on <a href="https://osf.io">osf.io</a> or <a href="https://plos.org/open-science/preregistration/">https://plos.org/open-science/preregistration/</a> ).	No	
8,1	Raw data access	State where raw data may be accessed including demultiplexing information.	STORMS	Robust, long-term databases such as those hosted by NCBI and EBI are preferred. If using a private repository, provide rationale.	Yes	p.17 section "Availability of data and materials"



8,2	Processed data access	State where processed data may be accessed.	STORMS	Unfiltered data should be provided. Robust, long-term databases such as those hosted by NCBI and EBI-EMBL are preferred. Repositories like zenodo ( <a href="https://zenodo.org/">https://zenodo.org/</a> ) or publisso ( <a href="https://www.publisso.de/en/working-for-you/doi-service/">https://www.publisso.de/en/working-for-you/doi-service/</a> ) can be used to provide a DOI and long-term storage for processed datasets, even those which cannot be published openly.	Yes	Additional files 1 and 2
8,3	Participant data access	State where individual participant data such as demographics and other covariates may be accessed, and how they can be matched to the microbiome data.	STORMS	If re-categorized, transformed, or otherwise derived variables were used in the analysis, these variables or code for deriving them should be provided.  Examples of how participant data can be matched to microbiome data are: using the same set of anonymized identifiers, or using different anonymized identifiers but providing a map.  Provided data should be sufficient to independently replicate the current analysis.	Yes	Table 1 and additional file XX
8,4	Source code access	State where code may be accessed.	STORMS	If a standard or formalized workflow was employed, reference it here.	No	
8,5	Full results	Provide full results of all analyses, in computer-readable format, in supplementary materials.	STORMS	For example, any fold-changes, p-values, or FDR values calculated, provided as a spreadsheet.  Use a machine-readable, plain-text format such as csv or tsv.	Yes	Additional files
<b>Results</b>						
9,0	Descriptive data	Give characteristics of study participants (e.g. dietary, demographic, clinical, social) and information on exposures and potential confounders.	STROBE	Typically reported in a table included in the paper or as a supplementary table. Indicate number of participants with missing data for each variable of interest.  This includes environmental and lifestyle factors that may affect the relationship between the microbiome and the condition of interest. Participant diet and medication use should be summarized, if known.  At minimum, age and sex of all participants should be summarized.	No	
10,0	Microbiome data	Report descriptive findings for microbiome analyses with all applicable outcomes and covariates.	STORMS	This includes measures of diversity as well as relative abundances. These descriptive findings should be reported both for the sample overall and for individual groups.	Yes	p.9-12
10,1	Taxonomy	Identify taxonomy using standardized taxon classifications that are sufficient to uniquely identify taxa.	STORMS	If not using full taxonomic hierarchy, make sure it is clear whether names stated are species, genera, family, etc.  Italicize genus/species pairs. Consult journal guidelines or standardized references on taxonomic nomenclature. For instance, <a href="https://wwwnc.cdc.gov/eid/page/scientific-nomenclature">https://wwwnc.cdc.gov/eid/page/scientific-nomenclature</a>	Yes	p.9-12

10,2	Differential abundance	Report results of differential abundance analysis by the variable of interest and (if applicable) by time, clearly indicating the direction of change and total number of taxa tested.	STORMS	If there are more than two groups, include omnibus (multigroup) test results if applicable to the research question.  If applicable, reported effect sizes should include a measure of uncertainty such as the confidence interval.	Yes	p.9-12
10,3	Other data types	Report other data analyzed--e.g. metabolic function, functional potential, MAG assembly, and RNAseq.	STORMS		Yes	p.12
10,4	Other statistical analysis	Report any statistical data analysis not covered above.	STORMS	This could include subgroup analysis, sensitivity analyses, and cluster analysis.  Visualizations should be easily interpretable and colorblind-friendly. The caption and/or main text should provide a detailed description of visualizations for visually-impaired readers.	NA	
<b>Discussion</b>						
11,0	Key results	Summarise key results with reference to study objectives	STROBE		Yes	p.13-16
12,0	Interpretation	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	STROBE	Define or clarify any subjective terms such as "dominant," "dysbiosis," and similar words used in interpretation of results.  When interpreting the findings, consider how the interpretation of the findings may be summarized or quoted for the general public such as in press releases or news articles.  If causal language is used in the interpretation (such as "alters," "affects," "results in," "causes," or "impacts"), assumptions made for causal inference should be explicitly stated as part of 6.0 and 13.0.  Distinguish between function potential (ie inferred from metagenomics) and observed activity (ie metatranscriptomic, metabolomic, proteomic) if discussing microbial function.	NA	
13,0	Limitations	Discuss limitations of the study, taking into account sources of potential bias or imprecision.	STROBE	Also consider limitations resulting from the methods (especially novel methods), the study design, and the sample size.	Yes	p.13
13,1	Bias	Discuss any potential for bias to influence study findings.	STORMS	May include sampling method, representativeness of study participants, or potential confounding.	No	
13,2	Generalizability	Discuss the generalisability (external validity) of the study results	STROBE	To what populations or other settings do you expect the conclusions to generalize?		
14,0	Ongoing/future work	Describe potential future research or ongoing research based on the study's findings.	STORMS			
<b>Other information</b>						

15,0	Funding	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	STROBE		Yes	Funding section
15,1	Acknowledgements	Include acknowledgements of those who contributed to the research but did not meet criteria for authorship.	STORMS	For general guidelines on authorship, see <a href="http://www.icmje.org">http://www.icmje.org</a> and <a href="https://www.elsevier.com/authors/journal-authors/policies-and-ethics/credit-author-statement">https://www.elsevier.com/authors/journal-authors/policies-and-ethics/credit-author-statement</a>	Yes	Acknowledgements section
15,2	Conflicts of Interest	Include a conflicts of interest statement.	STORMS		Yes	
16,0	Supplements	Indicate where supplements may be accessed and what materials they contain.	STORMS		Yes	
17,0	Supplementary data	Provide supplementary data files of results with for all taxa and all outcome variables analyzed. Indicate the taxonomic level of all taxa.	STORMS	Depending on the analysis performed, examples of the supplemental results included could be mean relative abundance, differential abundance, raw p-value, multiple hypothesis testing-adjusted p-values, and standard error.  All discussed taxa should include the taxonomic level (e.g. class, order, genus).	Yes	



# Statistical Evaluation of Metaproteomics and 16S rRNA Amplicon Sequencing Techniques for Study of Gut Microbiota Establishment in Infants with Cystic Fibrosis

Claudia Saralegui,<sup>a,b,c</sup> Carmen García-Durán,<sup>d</sup> Eduardo Romeu,<sup>e</sup> María Luisa Hernández-Sánchez,<sup>e</sup> Ainhize Maruri,<sup>a,b,c</sup> Natalia Bastón-Paz,<sup>a,b,c</sup> Adelaida Lamas,<sup>c,f</sup> Saioa Vicente,<sup>c,f</sup> Estela Pérez-Ruiz,<sup>g</sup> Isabel Delgado,<sup>h</sup> Carmen Luna-Paredes,<sup>i</sup> Juan de Dios Caballero,<sup>a,b,c</sup> Javier Zamora,<sup>j</sup>  Lucía Monteoliva,<sup>d,e</sup> Concepción Gil,<sup>d,e</sup>  Rosa del Campo<sup>a,b,c,k</sup>

<sup>a</sup>Servicio de Microbiología, Hospital Universitario Ramón y Cajal and Instituto Ramón y Cajal de Investigación Sanitaria, Madrid, Spain

<sup>b</sup>CIBERINFEC, Madrid, Spain

<sup>c</sup>Unidad de Fibrosis Quística, Hospital Universitario Ramón y Cajal, Madrid, Spain

<sup>d</sup>Departamento de Microbiología y Parasitología, Universidad Complutense de Madrid and Instituto Ramón y Cajal de Investigación Sanitaria, Madrid, Spain

<sup>e</sup>Unidad de Proteómica, Universidad Complutense de Madrid, Madrid, Spain

<sup>f</sup>Servicio de Pediatría, Hospital Universitario Ramón y Cajal and Instituto Ramón y Cajal de Investigación Sanitaria, Madrid, Spain

<sup>g</sup>Unidad de Fibrosis Quística, Hospital Regional Universitario de Málaga, Málaga, Spain

<sup>h</sup>Unidad de Fibrosis Quística, Hospital Virgen del Rocío, Seville, Spain

<sup>i</sup>Sección de Neumología y Alergia Infantil, Unidad Multidisciplinar Fibrosis Quística, Hospital Doce de Octubre, Madrid, Spain

<sup>j</sup>Unidad de Bioestadística, Hospital Universitario Ramón y Cajal and Instituto Ramón y Cajal de Investigación Sanitaria and CIBERESP, Madrid, Spain

<sup>k</sup>Universidad Alfonso X El Sabio, Madrid, Spain

Claudia Saralegui and Carmen García-Durán contributed equally to this work. The order of the authors was decided by their contributions.

**ABSTRACT** Newborn screening for cystic fibrosis (CF) can identify affected but asymptomatic infants. The selection of omic technique for gut microbiota study is crucial due to both the small amount of feces available and the low microorganism load. Our aims were to compare the agreement between 16S rRNA amplicon sequencing and metaproteomics by a robust statistical analysis, including both presence and abundance of taxa, to describe the sequential establishment of the gut microbiota during the first year of life in a small size sample (8 infants and 28 fecal samples). The taxonomic assignments by the two techniques were similar, whereas certain discrepancies were observed in the abundance detection, mostly the lower predicted relative abundance of *Bifidobacterium* and the higher predicted relative abundance of certain *Firmicutes* and *Proteobacteria* by amplicon sequencing. During the first months of life, the CF gut microbiota is characterized by a significant enrichment of *Ruminococcus gnavus*, the expression of certain virulent bacterial traits, and the detection of human inflammation-related proteins. Metaproteomics provides information on composition and functionality, as well as data on host-microbiome interactions. Its strength is the identification and quantification of *Actinobacteria* and certain classes of *Firmicutes*, but alpha diversity indices are not comparable to those of amplicon sequencing. Both techniques detected an aberrant microbiota in our small cohort of infants with CF during their first year of life, dominated by the enrichment of *R. gnavus* within a human inflammatory environment.

**IMPORTANCE** In recent years, some techniques have been incorporated for the study of microbial ecosystems, being 16S rRNA gene sequencing being the most widely used. Metaproteomics provides the advantage of identifying the interaction between microorganisms and human cells, but the available databases are less extensive as well as imprecise. Few studies compare the statistical differences between the two techniques to define the composition of an ecosystem. Our work shows that the two methods are comparable in terms of microorganism identification but provide different results in alpha diversity analysis. On the other hand, we have studied newborns with cystic fibrosis, for whom we have

**Editor** Joanna B. Goldberg, Emory University School of Medicine

**Copyright** © 2022 Saralegui et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Concepción Gil, [conchagil@ucm.es](mailto:conchagil@ucm.es), or Rosa del Campo, [rosa.campo@salud.madrid.org](mailto:rosa.campo@salud.madrid.org).

The authors declare a conflict of interest. Rosa del Campo is recipient of a Vertex grant. The remaining authors have no conflict of interest.

**Received** 21 April 2022

**Accepted** 23 September 2022

described the establishment of an intestinal ecosystem marked by the inflammatory response of the host and the enrichment of *Ruminococcus gnavus*.

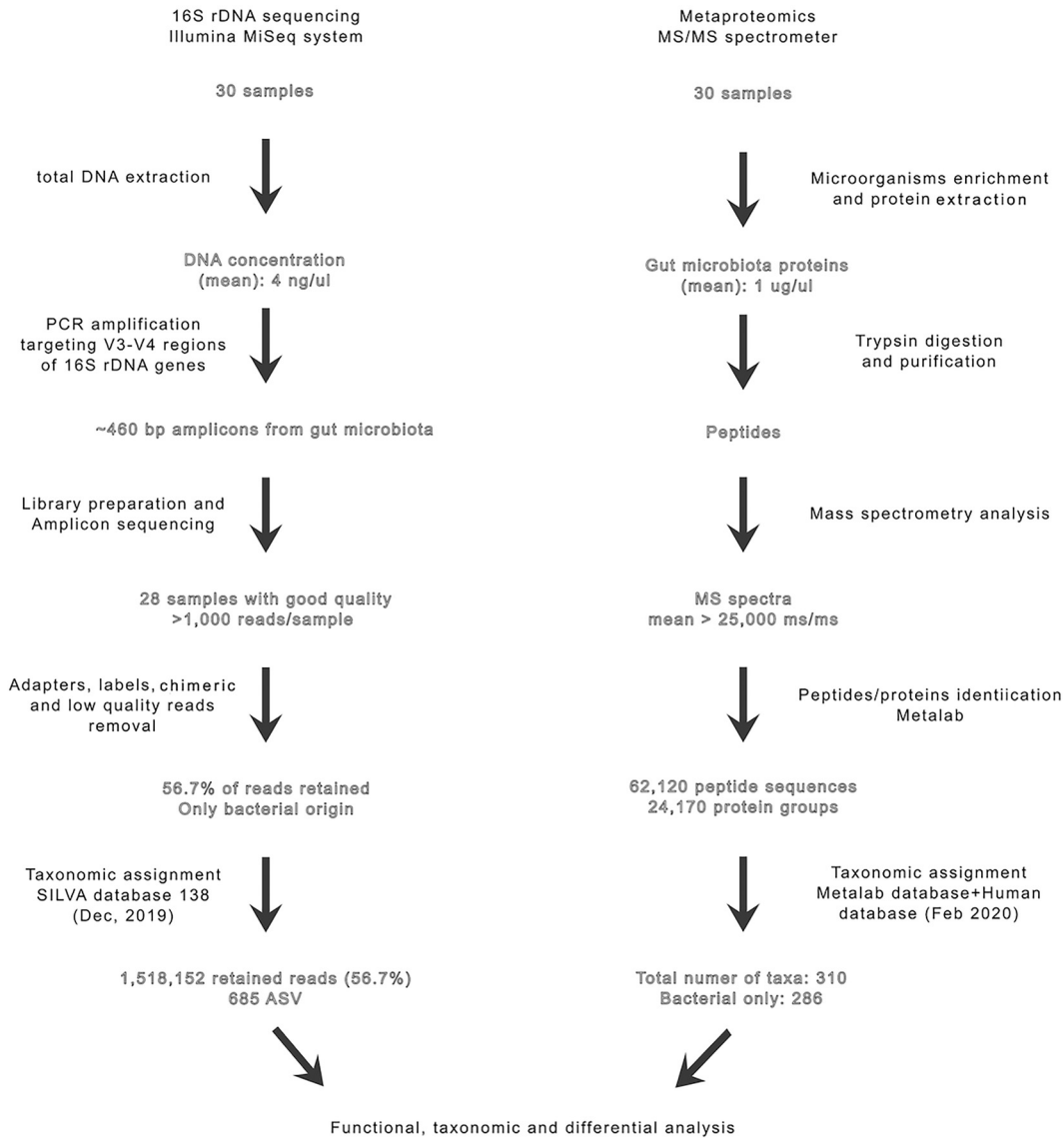
**KEYWORDS** amplicon sequencing, metaproteomics, cystic fibrosis, gut microbiota establishment, *Ruminococcus gnavus*, Bland-Altman test

The establishment of the microbiota at birth affects human health during the rest of life (1). In cystic fibrosis (CF), a genetic disease diagnosed by neonatal screening, the colonization of the lung airways by pathogenic bacteria impacts survival (2). The life expectancy of individuals with CF has increased considerably in recent years, and comorbidities beyond those of the respiratory tract have been revealed, mostly related to the digestive tract (3). The metabolic activity of the gut microbiota has a direct impact on nutritional status, a relevant aspect in CF since most of the patients are malnourished, and proper nutrition has a major effect on their infectious lung disease (4).

In CF, the gut microbiota is conditioned by the following factors: transmembrane regulator dysfunction, which creates a thick mucus layer; the acidic environment caused by impaired bicarbonate secretion; maldigestion; malabsorption; steatorrhea due to pancreatic insufficiency; and the high intake of antibiotics for bacterial infections during pulmonary exacerbations (5, 6). Significant differences in the microbiota's bacterial composition and alpha diversity have been reported compared with that in healthy infants (7–9).

Currently, 16S rRNA amplicon sequencing (AS) is the most extensive, rapid, and cost-effective technique for characterizing complex bacterial communities, such as the gut microbiota (10). The technique's main weakness results from the variable affinity of the primers for flanking regions; consequently, a higher predicted relative abundance has been found for certain genera, whereas others, mainly *Bifidobacterium*, are consistently reported to have a lower predicted relative abundance than other methods (11). Even with optimized primers, amplicon sequencing often provides deficient taxonomic assignment at the genus and species levels due to the short length of the amplicons. On the other hand, metaproteomics applies high-resolution mass spectrometry (MS) for the detection and quantification of proteins, allowing both taxonomic identification and functional characterization of all present microorganisms, not just bacteria. The recent introduction of updated bioinformatics tools, such as MetaLab, has considerably improved data analysis (12); for lesser-represented species, however, the analysis remains inaccurate (13). Available data suggest that only 10% to 20% of the total protein content of a clinical sample is identified, corresponding to only one-third of the entire microbiota (14). Additional strategies have recently been suggested as prefractionation steps to circumvent instrument sensitivity problems and access low-abundance proteins of biological importance (15). The strengths of metaproteomics include a high resolution in bacterial species identification, the overview of functional and metabolic activities of microorganisms, and, most notably, the detection of host proteins, expanding the knowledge of host-microbiome interactions (16, 17), which is particularly interesting in CF (18). However, metaproteomics also has relevant limitations (19): standardization of the technique is needed to make it reproducible, and integration of meta-omics platforms is necessary for in-depth characterization. Moreover, although its cost has been reduced, it is still far from those for amplicon sequencing.

The combination of amplicon sequencing and metaproteomics techniques provides a wider characterization of microbial communities and their interaction with the host (20–22), and the integration of the results could support accurate predictions (23). However, few studies have included a statistical analysis comparing the results by the two methodologies, especially for challenging samples such as those from neonates, characterized by limited sample quantity and potentially low bacterial load. The aims of this study were to perform a robust statistical analysis on the concordance of amplicon sequencing and metaproteomics to study and describe the sequential establishment of the gut microbiota of a small representative cohort of infants with CF.



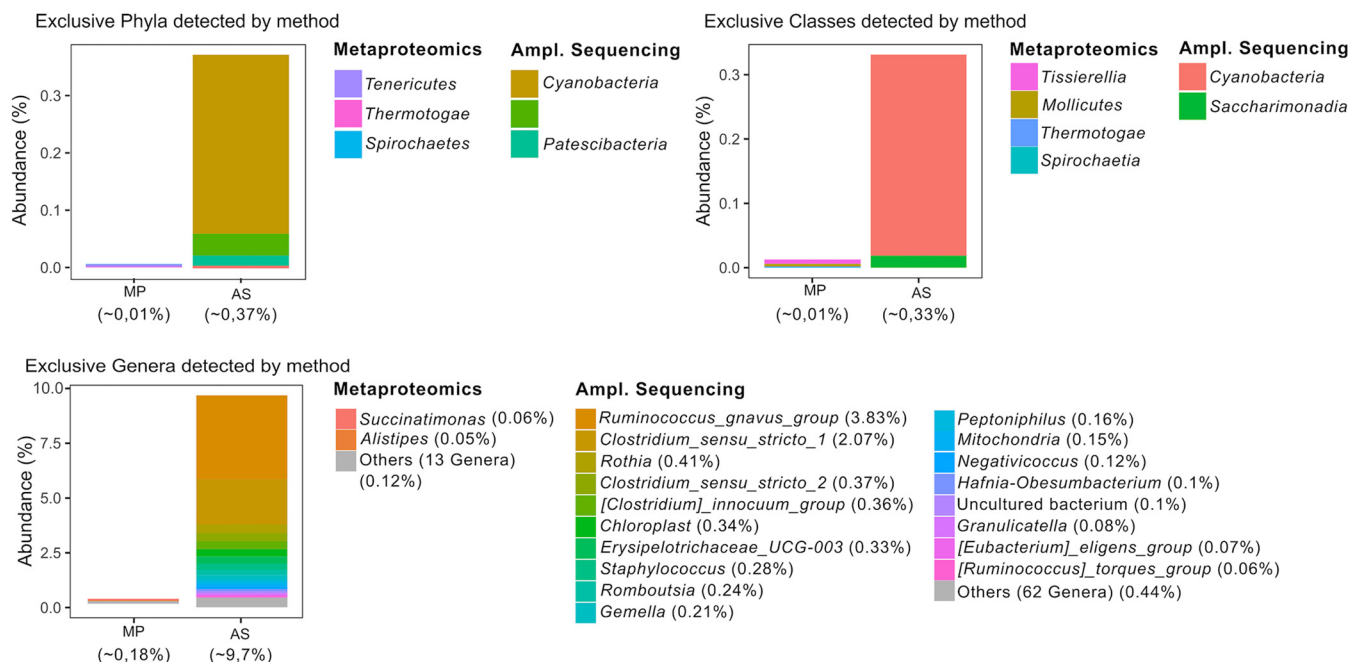
**FIG 1** Workflow and main results obtained from the two methodologies used in this study.

## RESULTS

### Statistical concordance of the results of amplicon sequencing and metaproteomics.

We analyzed fecal samples ( $n = 28$ ) using both techniques for bacterial identification (presence/absence) and quantification (abundance) and then respectively applied the McNemar and Wilcoxon tests to detect significant discrepancies (Fig. 1; see also “Amplicon sequencing data” and “Metaproteomics” in the supplemental material). Globally, 685 amplicon sequence variants (ASVs) (9 phyla, 13 classes, and 130 genera) were identified by amplicon sequencing (AS), and 310 taxa (8 phyla, 17 classes, and 65 genera) were identified by metaproteomics.

Taxa detected exclusively by a single method accounted for less than 10% of the total bacterial composition, being the genus level the one with most of the unique bacteria (Fig. 2). Minor statistical discrepancies were observed in taxon presence/absence by McNemar test (Table 1), most of which were related to the genus level (Fig. 2, Table 2, and “Common taxonomy” in the supplemental material). Furthermore, we detected certain nominal inconsistencies that should be accounted for related to the taxonomic classification by each database, such as the inclusion of the *Tenericutes* phylum from metaproteomics in the *Firmicutes* phylum



**FIG 2** Bacterial phyla, classes, and genera uniquely identified by each of the methodologies. AS, amplicon sequencing; MP, metaproteomics. White spaces correspond to taxa correctly identified by both methods.

according to amplicon sequencing and the recent taxonomic reclassification of *Proteobacteria* classes (Table 2). Most of the reclassified taxa were at higher taxonomic levels.

Figure 3 summarizes the results of the compositional analysis from the top common taxa. Wilcoxon signed-rank paired tests showed significant differences in the relative abundances of the following taxa: the phylum *Actinobacteriota*, the classes *Bacilli* and *Actinobacteria*, and the genera *Blautia*, *Haemophilus*, *Escherichia*, *Enterococcus*, *Streptococcus*, and *Bifidobacterium* (Fig. 3). These top genera were separately explored to consider within-subject effect across the time series (Fig. 4A). The only genera that had a differential composition across time according to AS were *Ruminococcus* ( $P = 0.022$ ) and *Eubacterium* ( $P = 0.007$ ) (Fig. 4B, left). According to metaproteomics, the genera that have changed over time are *Fusobacterium*

**TABLE 1** Contingency table for McNemar test accounting for the number of samples in which different bacteria (class and genus levels) were or were not identified by one, both, or neither of the technologies

Organism(s)	No. of samples detected by:				P value <sup>a</sup>
	Amplicon sequencing only	Metaproteomics only	Both	Neither	
<b>Class</b>					
<i>Alphaproteobacteria</i>	8	1	9	10	0.02
<b>Genus</b>					
<i>Eggerthella</i>	12	0	1	15	0.001
<i>Anaerococcus</i>	11	0	1	16	0.002
<i>Haemophilus</i>	15	2	5	6	0.003
<i>Coprococcus</i>	1	10	1	16	0.007
<i>Lachnospirillum</i>	0	9	2	17	0.007
<i>Roseburia</i>	0	9	2	17	0.007
<i>Eubacterium</i>	1	14	6	7	0.01
<i>Enterobacter</i>	10	1	3	14	0.01
<i>Enterococcus</i>	10	1	14	3	0.01
<i>Escherichia</i>	11	1	14	2	0.01
<i>Actinomyces</i>	14	4	6	4	0.03
<i>Collinsella</i>	8	1	5	14	0.04

<sup>a</sup>Significance level < 0.05.

**TABLE 2** Discrepancies in nomenclature or taxonomic classification at different ranks (phylum/class/order/family/genus) by amplicon sequencing and metaproteomics<sup>a</sup>

Taxa identified by amplicon sequencing	Taxa identified by metaproteomics
<i>Campilobacterota/Campylobacteria/Campylobacterales</i>	<i>Proteobacteria<sup>a</sup>/Epsilonproteobacteria/Campylobacterales</i>
<i>Firmicutes/Bacilli/Mycoplasmatales</i>	<i>Tenericutes/Mollicutes/Mycoplasmatales</i>
<i>Actinobacteriota</i>	<i>Actinobacteria</i>
<i>Bacteroidota</i>	<i>Bacteroidetes</i>
.../Bacilli/Erysipelotrichales	.../Erysipelotrichia/Erysipelotrichales
.../Gammaproteobacteria/Burkholderiales	.../Betaproteobacteria/Burkholderiales
<i>Desulfobacterota/Desulfovibrionia/Desulfovibrionales</i>	<i>Proteobacteria/Deltaproteobacteria/Desulfovibrionales</i>
.../Clostridia/Peptostreptococcales-Tissierellales	.../Tissierellia/Tissierellales
.../.../.../Escherichia-Shigella	.../.../.../Escherichia
.../.../.../Lachnospirales/Lachnospiraceae/Ruminococcus_gnavus_group	.../.../.../Clostridiales/Ruminococcaceae/Ruminococcus
.../.../.../Lachnospirales/Lachnospiraceae/Ruminococcus_torques_group	

<sup>a</sup>Names of taxonomic ranks that are identical are expressed with ".../..." The taxonomic assignment was made in 2019 before the latest nominal changes in the bacterial phyla. *Actinobacteriota* is now *Actinomycetota*, *Firmicutes* is *Bacillota*, and *Proteobacteria* is now *Pseudomonadota*.

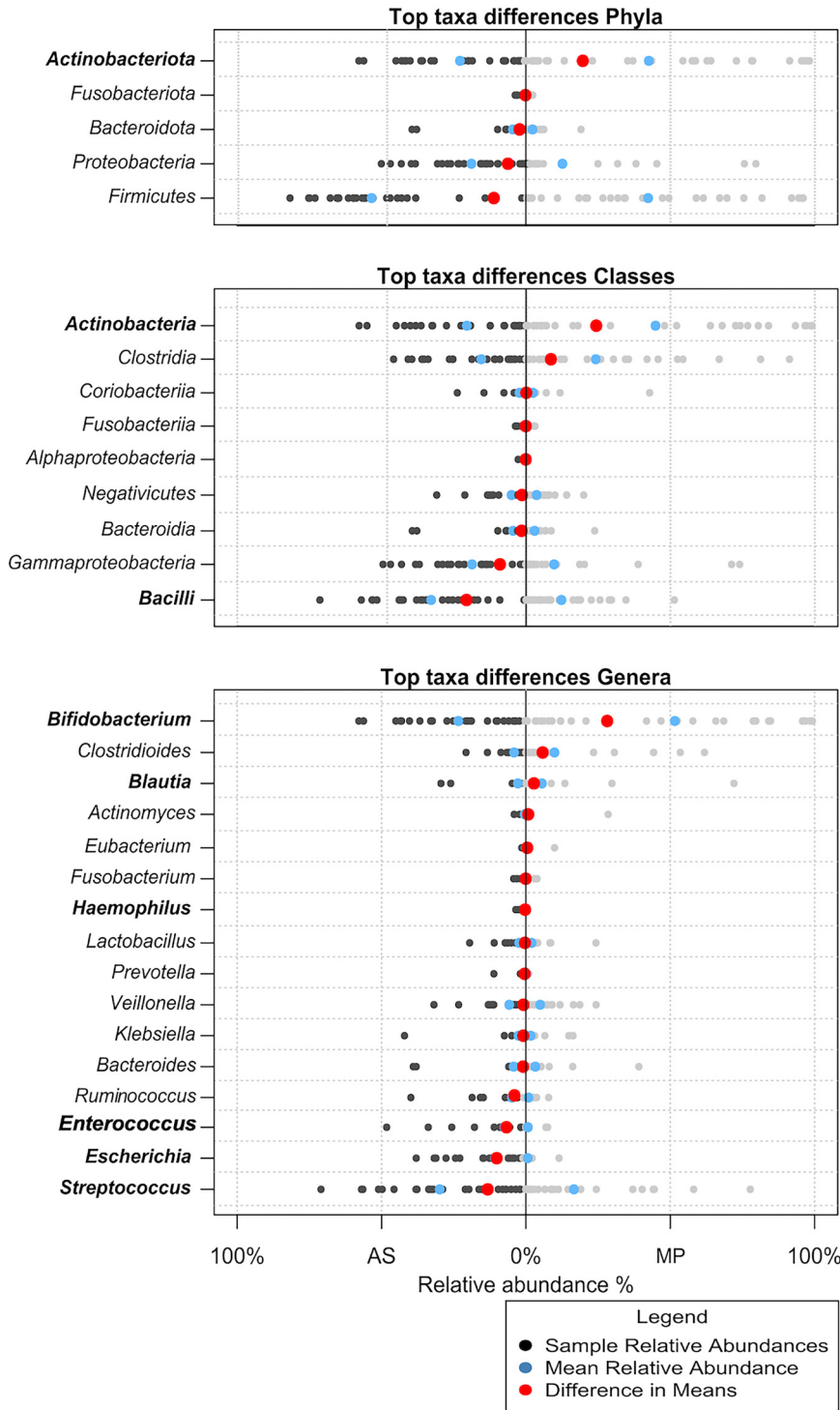
( $P = 0.02$ ), *Haemophilus* ( $P = 0.029$ ) and *Veillonella* ( $P = 0.03$ ) (Fig. 4B, right). The two methods showed similar trends for most genera, except for *Bifidobacterium*, *Escherichia*, *Enterococcus*, and *Ruminococcus*, which almost reach statistical significance ( $P < 0.05$  but with a false-discovery rate [FDR] of 0.13). *Bifidobacterium* was the genus with the highest distance between methods (Fig. 4C to E). It is worth mentioning that significance could be affected by the small sample size and that sometimes changes are mostly detected for the lowest-abundance genera.

Alpha diversity metrics were significantly different between methodologies (Wilcoxon signed-rank paired test,  $P < 0.05$ ). The methods also showed differences in temporal trends (Fig. 5), with only amplicon sequencing showing an increasing trend for the Shannon index in the latest samples. It must be mentioned that alpha diversity metrics calculations did not work well with metaproteomics data, as phyloseq software warned that there were no singletons in metaproteomics data frames. Calculating Shannon indices from rank-collapsed data and label free quantifications (LFQ) intensities as the ones that were available in metaproteomics workflow output is not common and should be considered for further works. Amplicon sequencing remained the best approach to calculate diversity metrics such as the aforementioned.

**Functional assignment by metaproteomics.** Functional information on the detected proteins was obtained by MetaLab software, using the sum of the peptide intensities for each protein (total intensity of the protein). Both the number and intensity of the proteins in the initial (5,718 and 4.86e12, respectively) and the early-CF (7,029 and 4.70e12, respectively) fecal samples were comparable. MetaLab assigned the functional profile by matching proteins with the Cluster of Orthologous Groups (COG) database with a 97% successful assignment rate for name/category. A previous process of normalization of the protein intensities (intensity of the function/sum of the intensities in that sample) was performed, and 2-way analysis of variance detected 14 COGs (6 from the initial and 8 from the early-CF samples) with statistical relevance (Table 3). However, and despite the small sample size of our work, there were no significant differences in the general functional profiles in the enrichment analysis, of which "energy production and conversion" (C), "carbohydrate transport and metabolism" (G), and "translation, ribosomal structure, and biogenesis" (J) were the COG categories with greater intensity (Fig. 6A).

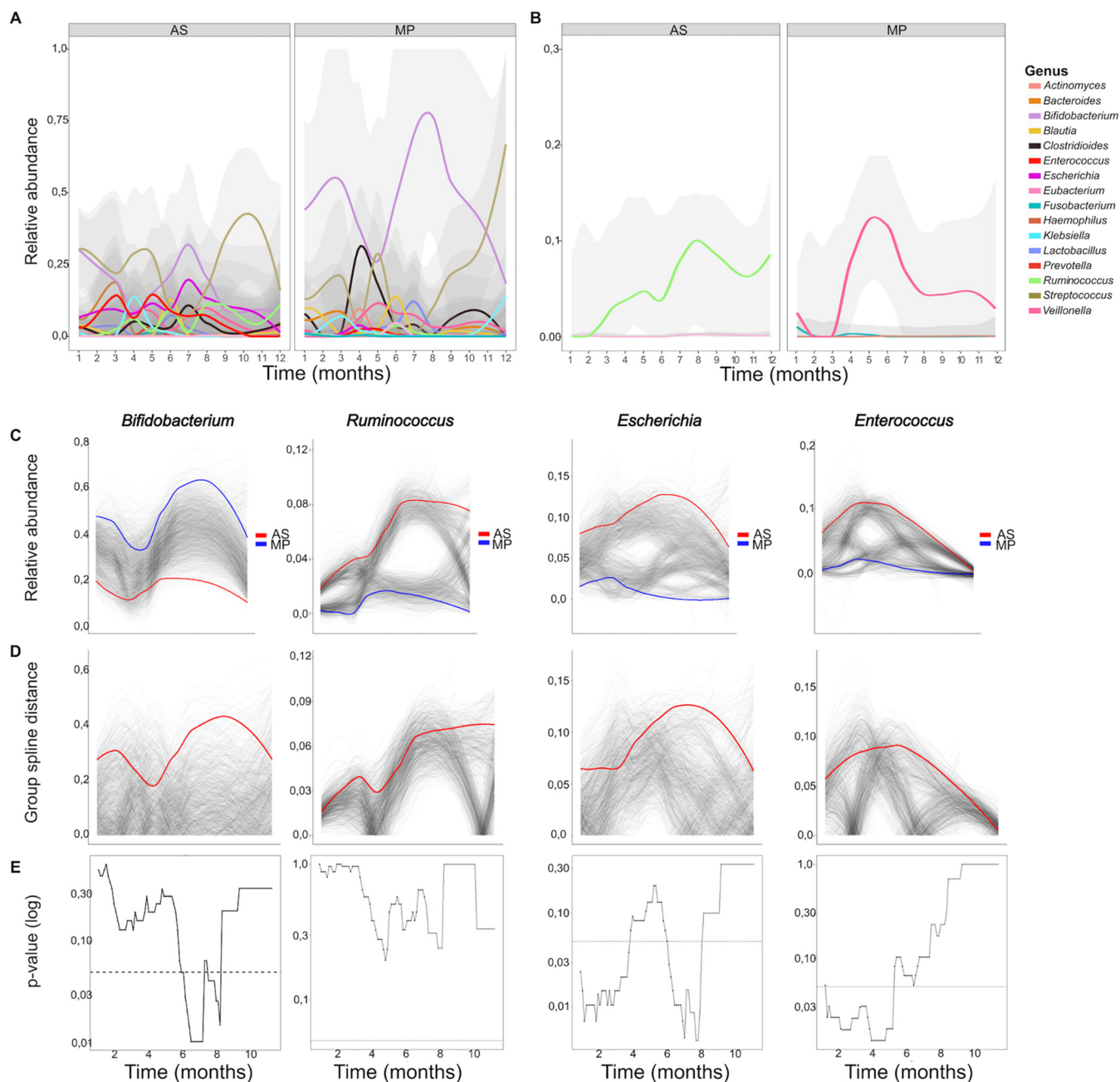
The correlation of each bacterial taxon and its function was also performed with iMetaLab, using the iMetaShiny apps. *Bifidobacterium* spp. were the most abundant taxa and their proteins remained abundant in the early-CF samples, in which most of the functional pathways were promoted by *Ruminococcus gnavus* (Fig. 6B). A significant enrichment of *R. gnavus* proteins was observed from initial to early-CF times (2 versus 69 proteins), most of them related to translational processes (Fig. 6C). In contrast, the functions associated with *Clostridium perfringens* and *Blautia hansenii* showed a statistically significant decrease in the early-CF samples.

Metaproteomics also allowed the identification of 293 human proteins: 85 exclusively



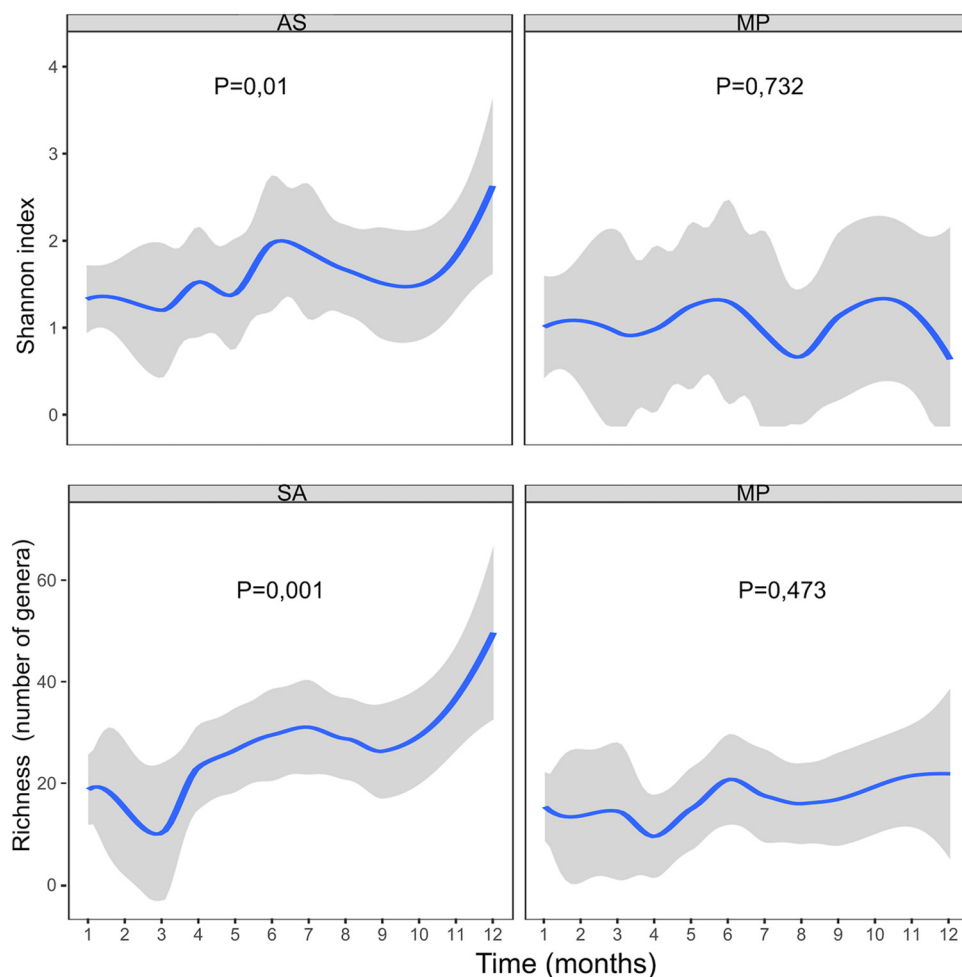
**FIG 3** Relative abundances of the main taxa detected by both methodologies. Taxa in bold are those for which the paired Wilcoxon test result was significant ( $P < 0.05$ ).

present in initial samples, 43 exclusively in early-CF samples, and 165 in both samples (“Human proteins” in the supplemental material). Six of the 165 shared proteins had significantly differential expression levels (Table 4). A STRING analysis (STRING version 11.5) was performed with the 3 groups of human proteins (initial, early CF, and shared proteins), including for the first 2 groups those proteins exclusively identified and with a significantly higher abundance in each of the samples (Fig. 7). The analysis of the proteins shared by the two groups



**FIG 4** Compositional change of the major genera in each of the methodologies. (A) Evolution of the relative abundance of the main bacterial genera ( $n = 16$ ). (B) Evolution of bacterial genera with a significant change over time detected by AS (*Ruminococcus* and *Eubacterium*) and by MP (*Fusobacterium*, *Haemophilus*, and *Veillonella*). (C) Graphs obtained by the permutspliner function (999 permutations) showing the temporal evolution of relative abundance in those genera with significant differences between the two methodologies. (D) Plots of distances between both methodologies. The difference (solid red line) is not significant if it is not above 95% of the permuted values (translucent gray). (E) Plots obtained by sliding spliner showing the  $P$  value at each specified interval (shown with 100 intervals by default). The dotted line indicates a  $P$  value of 0.05. At certain intervals, the differences became significant.

showed a significant protein-protein interaction ( $1.87 \times 10^{-16}$ ). Shared proteins were mainly related to neutrophil-mediated immunity (42 proteins;  $P = 1.87 \times 10^{-25}$ ), such as calprotectin (S100-A9 and S100-A8), neutrophil elastase (ELANE), cathepsin G, and myeloperoxidase, or related to the intestinal barrier, such as mucins (MUC2, MUC4, and MUC13) or proteins present in cell junctions (40 proteins;  $P = 1.24 \times 10^{-5}$ ) (Fig. 7A). The initial samples contained significantly more proteins associated with chylomicron assembly and the immune system than the early-CF samples (Fig. 7B), in which a cluster of muscle proteins stood out (Fig. 7C).



**FIG 5** Evolution of alpha diversity (measured by Shannon index and richness) detected by one method and the other. The  $P$  value was obtained with the trendspliner spliner function (shown with 100 intervals by default), which evaluates the differences over time for each group of samples, separately.

The names of the proteins belonging to each function are listed in “Microbial data” in the supplemental material.

## DISCUSSION

A major clinical feature of CF are digestive comorbidities, such as steatorrhea, mal-digestion/malabsorption, and pancreatic insufficiency, which also affect the quality of life and nutritional status of these individuals. A new classification of CF genetic mutations has been published (24), differentiating between high and low risk according to their clinical manifestations. Seven out of the 8 infants of this study carried high-risk mutations. The CF environment strongly conditions the gut microbiota composition (25), also contributing to its malfunction, even at the age of 6 weeks (9). Even though amplicon sequencing is the most widely used technique for determining the composition of gut microbiota, there is a growing demand for techniques that determine the global functionality of the ecosystem, such as metaproteomics.

The taxonomical incongruities detected were attributed to paraphyletic groups, which are continuously reclassified according to the latest advances in phylogeny. In this sense, SILVA 138 is a more accurate database with which to establish phylogenetic relationships, because protein databases are less frequently updated. By the time this paper is published, certain phylum names will have changed: *Firmicutes* is now *Bacillota*, *Proteobacteria* is now *Pseudomonadota*, and *Actinobacteriota* is currently named *Actinomycetota*.

**TABLE 3** Functional categories of bacterial proteins with significantly different values between the initial group of samples and the early-CF group

COG name <sup>a</sup>	Mean difference	P value	COG category
Formyltetrahydrofolate synthetase (2,402)	2.64	<0.0001	Nucleotide transport and metabolism
ABC-type phosphate transport system, periplasmic component (2,605)	1.20	<0.0001	Inorganic ion transport and metabolism
Superfamily II DNA and RNA helicase (3,843)	1.25	<0.0001	Replication, recombination, and repair
Translation elongation factor EF-Tu, a GTPase (2,272)	0.74	0.0009	Translation, ribosomal structure, and biogenesis
F <sub>0</sub> F <sub>1</sub> -type ATP synthase, delta subunit (1,785)	0.65	0.01	Energy production and conversion
Pyruvate-formate lyase (4,465)	0.63	0.03	Energy production and conversion
Chaperonin GroEL (HSP60 family) (2,675)	-2.06	<0.0001	Posttranslational modification, protein turnover, and chaperones
Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase (2,333)	-1.03	<0.0001	Carbohydrate transport and metabolism
Carboxysome shell and ethanolamine utilization microcompartment protein CcmL/EutN (994)	-0.98	<0.0001	Secondary-metabolite biosynthesis, transport, and catabolism
Phosphoenolpyruvate-protein kinase (PTS system component in bacteria) (4,879)	-0.96	<0.0001	Carbohydrate transport and metabolism
Alcohol dehydrogenase, class IV (6,245)	-0.94	<0.0001	Energy production and conversion
Rubryerythrin (4,121)	-0.86	<0.0001	Energy production and conversion
Threonine dehydrogenase or related Zn-dependent dehydrogenase (5,594)	-0.79	<0.0001	General function prediction only
ABC-type sugar transport system, periplasmic component, contains N-terminal Xre family HTH domain (4,787)	-0.62	0.03	Carbohydrate transport and metabolism

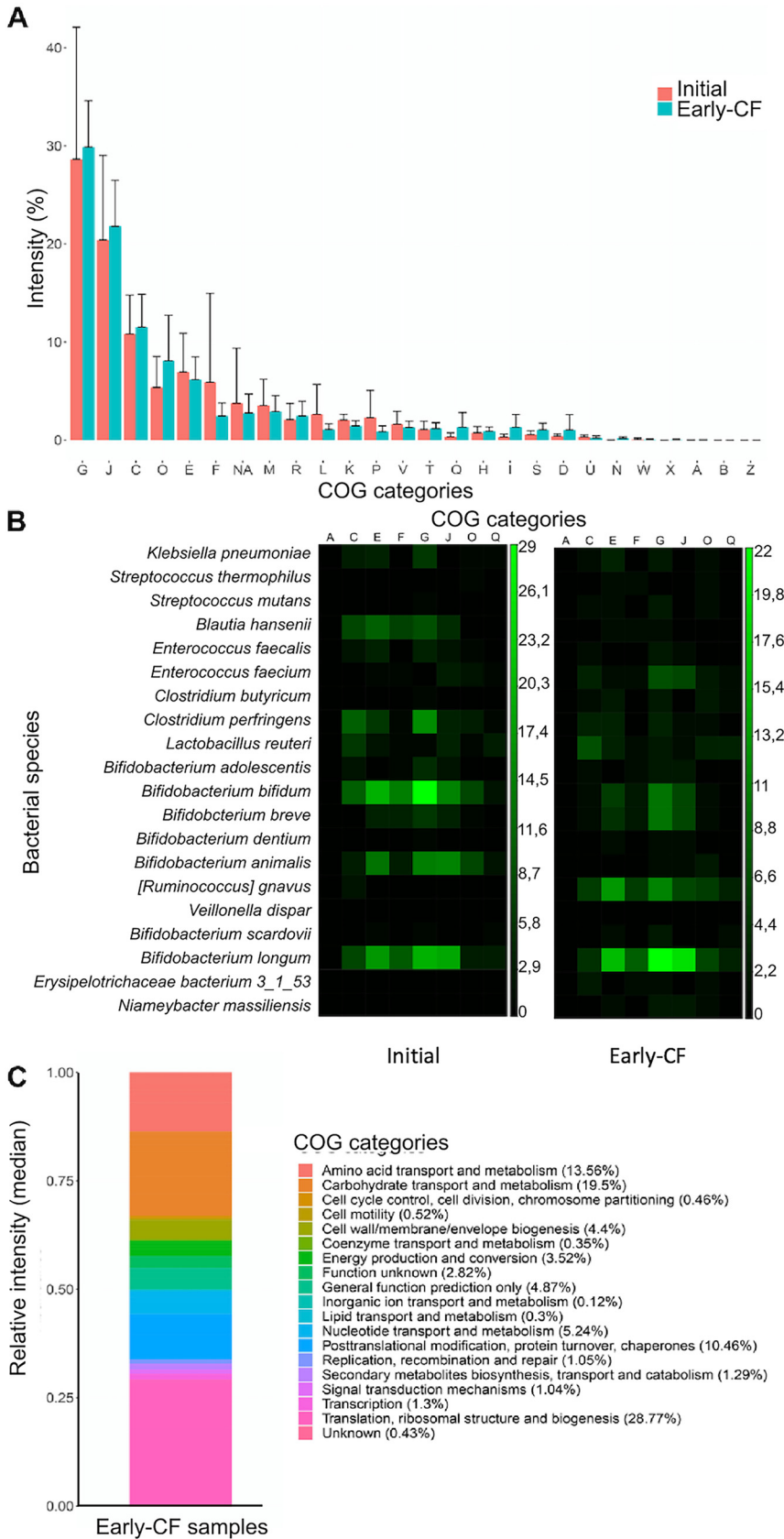
<sup>a</sup>PTS, phosphotransferase system; HTH, helix-turn-helix.

Curiously, metaproteomics detected proteins attributed to *Streptococcus thermophilus*, whereas when using amplicon sequencing, the proteins appeared to correspond to *Streptococcus salivarius*. Poor resolution of streptococcal species using the V3-V4 regions of the 16S rRNA gene has been previously reported (26), whereas the accuracy of metaproteomics in bacterial species identification has been demonstrated (12). Another amplicon sequencing limitation is the detection of what is known as sporobiota (27), most of the endospore-forming species belonging to *Firmicutes*. Spores are underrepresented because of their ability to resist DNA extraction techniques. Other bacterial phyla, such as *Actinobacteria* (particularly *Bifidobacterium*), have genomes with a high GC content, and this likewise contributes to their underestimation by sequencing techniques.

Metaproteomics has also failed to detect or properly estimate certain bacterial groups typically found in the human gut microbiota due to their resistance to cell lysis, especially Gram-positive bacteria (28). Moreover, certain microbial components and detergents used for cell lysis can affect postenzymatic digestion (14). Our workflow includes optimized steps, such as serial centrifugations for microbial pre-enrichment, sonication with bead beating, and lysis with sodium dodecyl sulfate (SDS)-based buffers. While our method improves the number of peptide and protein identifications and detection of the relative abundance of *Actinobacteria*, other protocols have shown a higher identification of proteins from other taxonomic groups, such as *Proteobacteria*, but a lower number of peptide and protein identifications (29). Metaproteomic analysis presents several bioinformatics challenges, the most common being adequate peptide identification, taxonomic assignment, and the quality of functional annotation (14).

The statistical analysis of our results demonstrated equivalent bacterial taxonomic assignments by both techniques, while for abundance detection, some discrepancies were found. The lower predicted relative abundance of *Bifidobacterium* by amplicon sequencing has been repeatedly reported, as well as the higher predicted relative abundance we detected for some *Firmicutes* and *Proteobacteria* taxa in comparison to other methodologies (30).

The small size of our cohort limits comparison with other studies. Alpha diversity indices of our samples did not reach those observed in other CF cohorts or in the control cohort of Antosca et al. (9), although other work showed lower values in CF patients than in healthy controls (31). We conclude that it is not possible to compare alpha diversity indices between metaproteomics and amplicon sequencing with the approach we followed in this study,



**FIG 6** Functional analysis of bacterial proteins from this cohort, comparing the group of initial and final samples (early CF). (A) Relative intensity of each of the COG functional categories in each group of (Continued on next page)

**TABLE 4** Human proteins enriched in each group of the cohort

Higher expression in initial sample	Higher expression in early-CF sample	P value
Lactotransferrin		<0.0001
Myeloperoxidase		<0.0001
	Actin cytoplasmic 2	<0.0001
	Chymotrypsin-C	<0.0001
	Carboxypeptidase A1	0.0409
	Myosin-1	<0.0001

although comparison between samples processed with identical proteomic techniques could be possible. In amplicon sequencing, several ASVs (each with a certain number of reads, typically on the order of 1 to 10<sup>4</sup>) may correspond to the same bacterial taxon. In metaproteomics, on the other hand, the sum of intensities (on the order of 10<sup>6</sup> to 10<sup>11</sup>) of all peptides corresponding to the same taxon is evaluated to calculate the taxon's relative abundance in every sample. This makes the calculation of metrics such as the Shannon index, which considers the proportion of each bacterium, and their comparison between techniques not entirely reliable; nonetheless, species richness (number of different species) could still be used.

Fecal proteomics for adults with CF has pointed to the dominance of *R. gnavus*, *Enterobacteriaceae*, and *Clostridia* species, combined with decreased butyrate producers, such as *Faecalibacterium prausnitzii* (18). We detected similar compositional changes in a very reduced cohort of 8 infants followed for a year right after being diagnosed with CF. *R. gnavus* has been associated with several human gastrointestinal diseases and is a mucolytic bacterium capable of degrading the human colonic mucin MUC2, which is a major form of secretory glycosylated mucin coating the epithelia of the intestine and therefore has a close relationship with the mucus layer (32, 33). It has been suggested that the bacterium can penetrate the mucus layer and lead to an activation of the immune system and inflammatory responses (34, 35). Nevertheless, the involvement of this species in allergic or autoimmune diseases has been also questioned (32, 36).

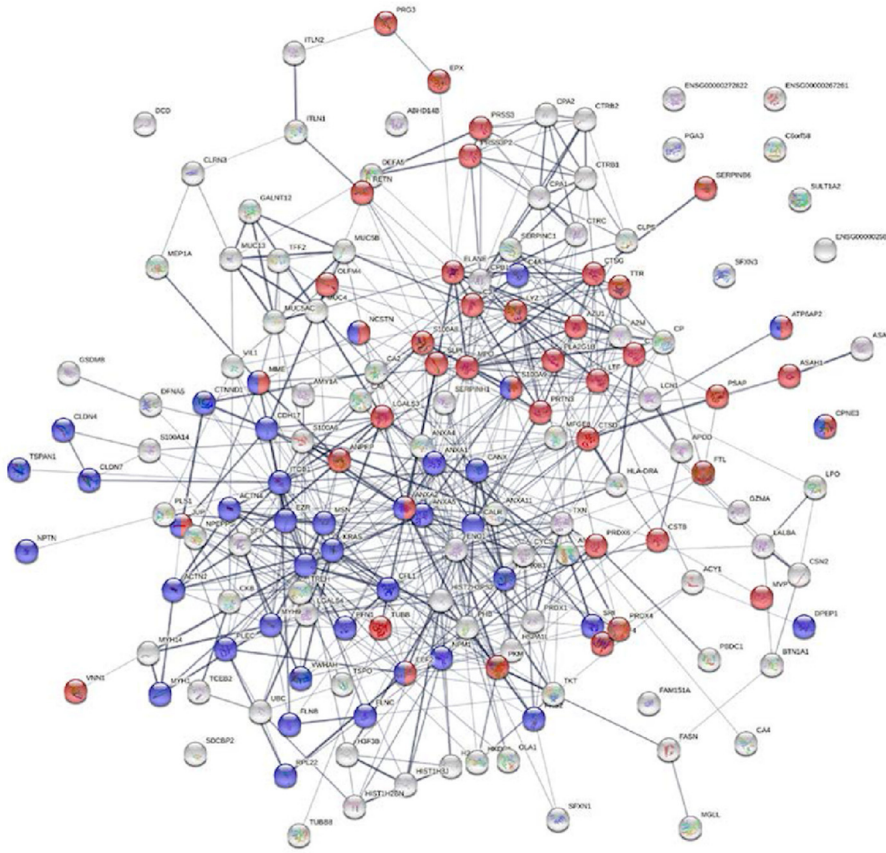
Drastic changes in the bacterial functional profiles of the initial and early-CF samples were not observed; the most abundant proteins in the initial samples were a formyltetrahydrofolate synthetase and a pyruvate-formate lyase previously related to biofilm upregulation (37). In the early-CF samples, the most relevant enrichment was that of the chaperonin GroEL, a stress response protein whose main role is protein folding, which has recently been observed to have roles in virulence and pathogenesis, such as bacterial adherence and immune invasion (38).

One of the most interesting advantages of metaproteomics is the detection of human proteins. In our samples, several proteins associated with the maintenance of the intestinal epithelium and the immune response were detected. Lactotransferrin, an iron sequestration protein abundant in human breast milk; and myeloperoxidase, regulator of inflammation in granulocytes of neutrophils, were significantly enriched in the initial samples (39, 40). All our results point to an altered establishment of the

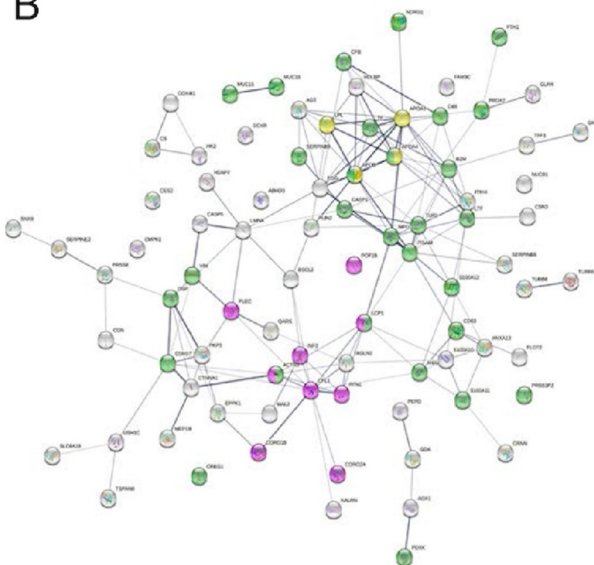
**FIG 6** Legend (Continued)

samples. (B) Correlation between the most enriched functional categories and the bacterial species to which they are assigned. The color gradient indicates higher relative intensity (light green). (C) Proportions of COG functional categories of proteins assigned to *R. gnavus*. COG category nomenclature: A, RNA processing and modification; B, chromatin structure and dynamics; C, energy production and conversion; D, cell cycle control, cell division, and chromosome partitioning; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; J, translation, ribosomal structure, and biogenesis; K, transcription; L, replication, recombination, and repair; M, cell wall/membrane/envelope biogenesis; N, cell motility; O, posttranslational modification, protein turnover, and chaperones; P, inorganic ion transport and metabolism; Q, secondary-metabolite biosynthesis, transport, and catabolism; R, general function prediction only; S, function unknown; T, signal transduction mechanisms; U, intracellular trafficking, secretion, and vesicular transport; V, defense mechanisms; W, extracellular structures; X, mobilome: prophages and transposons; Z, cytoskeleton; NA, not assigned to any category.

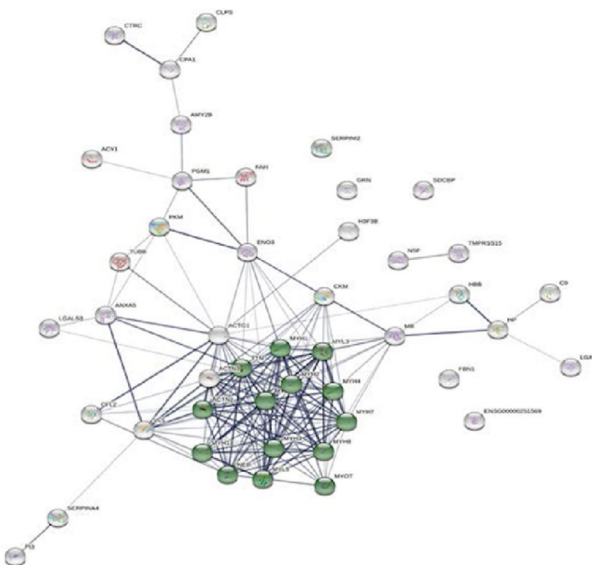
A



B



C



- Neutrophil mediated immunity
- Cell junction
- Immune system process
- Chylomicron remodeling
- Actin binding
- Muscle protein and sarcomere organization

**FIG 7** Human proteins found interacting with the gut microbiome of the cohort. (A) Group of proteins unique to the initial samples; (B) group of proteins unique to the early-CF samples; (C) proteins shared by both groups.

**TABLE 5** Relevant clinical characteristics of infants and sampling<sup>a</sup>

Infant	Sex	Pancreatic insufficiency	Delivery mode	Birth feeding	CF mutation <sup>b</sup>	Sampling at mo (no. of sample)											
						1	2	3	4	5	6	7	8	9	10	11	12
1	Female	Yes	Vaginal	Formula	G542X/R1162X	1	2						3				
2	Male	Yes	Caesarean section	Formula	F508del/Y1092X	4				5		6					
3	Female	Yes	Vaginal	Exc. <sup>c</sup> breast milk	F508del/F508del	7			8						9		10
4	Male	No	Vaginal	Exc. breast milk	F508del/3272-26A->G	11		12			13		14				
5	Male	Yes	Vaginal	Exc. breast milk	F508del/F508del	15			16				17				
6	Male	Yes	Vaginal	Formula	F508del/N1303K	18					19			20			
7	Female	Yes	Vaginal	Formula	F508del/F508del	21			22	23		24					
8	Female	Yes	Vaginal	Formula	F508del/R1162X	25		26		27							28

<sup>a</sup>All samples ( $n = 28$ ) were included in the methodological comparative analysis, whereas samples in gray ( $n = 16$ ; 8 initial samples and 8 early-CF samples) were considered to analyze gut microbiota and human proteomic content in this cohort.

<sup>b</sup>All CFTR alleles except 3272-26A>G are classified as high risk, and therefore, all infants had a high-risk CF genotype except infant 4.

<sup>c</sup>Exc., exclusively breast milk. At endpoint, all infants had included solid food in their feeding status.

gut microbiota in infants with CF, with a dominance of *R. gnavus* within a human inflammatory state.

The main limitation of our study is the small sample size. Due to that, we have included unadjusted, slightly higher *P* values to inform about bacterial phyla, classes, and genera and about microbial and human proteins which appeared to change but did not reach statistical significance. These features could be considered in future studies with larger sample sizes or more exhaustive CF metabolic characterization.

Our study provides an extended comparative analysis with robust statistical support that could optimize the use of both approaches for gut microbiota research. Metaproteomics provides information on composition and functionality, as well as data on host-microbiome interactions (41). Its strength is the identification and quantification of *Actinobacteria* and certain classes of *Firmicutes*, whereas its main limitation is the lack of comparability of alpha diversity indices. Taking all the results into account, both techniques detected an aberrant microbiota in infants with CF during their first year of life, dominated by the enrichment of *R. gnavus* within a human inflammatory environment.

## MATERIALS AND METHODS

**Participants, sample collection, and medical records.** Subjects included in this study had CF diagnosed by neonatal screening, sweat chloride test, and mutation sequencing and were recruited during their first month of life, when the first fecal sample was collected, between January 2018 and January 2019. The infants were attended at 3 Spanish CF reference hospitals, geographically distant. Ethical approval for the study was granted by the Ramón y Cajal Ethics Committee in 2017 with (file number 163/17), and all infants' parents gave informed consent. Subjects were excluded if they were born after that date or had fewer than 3 collected samples or if important clinical data were missing. Eight infants with CF were finally followed up, and the relevant data, including feeding habits, delivery mode, and CF mutations, are shown in Table 5. Data related to antibiotic and other treatments are summarized in "Antibiotic treatments" in the supplemental material. Fecal samples were sequentially collected from the diaper of each infant during the first year of life over routine medical checkups. Samples were stored at  $-80^{\circ}\text{C}$  in 2 aliquots for subsequent amplicon sequencing and meta-proteomics analysis.

**DNA extraction and 16S rRNA amplicon sequencing.** Our study followed guidelines for the Strengthening The Organization and Reporting of Microbiome Studies (STORMS) reporting ("STORM checklist" in the supplemental material) (42). Total DNA from the fecal samples was obtained with the QiaAMP kit (Qiagen, Germany), and further MiSeq 2  $\times$  300-bp paired-end (Illumina) 16S rRNA sequencing of the V3 and V4 regions (43) was performed at the Central Unit for Translational Genomics Support (Ramón y Cajal Health Research Institute). The sequencing data analysis was performed using the Qiime pipeline (44), which includes DADA2 for sequence quality filtering, and the SILVA 138 database (released December 2019) was used for taxonomical assignment, discharging those samples with fewer than 1,000 reads ( $n = 2$ ). Nucleotide sequences were deposited in the National Center for Biotechnology Information's Sequence Read Archive (SRA) repository, BioProject identifier (ID) [PRJNA719717](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA719717).

**Metaproteomics.** Samples were prepared using differential centrifugation, and the microbial pellet was processed by sonication according to a previously published protocol. In brief, 0.1 to 0.3 g of feces was suspended in 10 mL of phosphate-buffered saline (PBS) and mixed in a tube rotator for 45 min at  $4^{\circ}\text{C}$ , as previously published (45, 46). The samples were centrifuged at  $500 \times g$  for 5 min, and the supernatant was collected in a 50-mL tube. Ten milliliters of PBS was added again, repeating the process twice, and the 3 supernatants ( $\sim 30$  mL) were centrifuged at  $11,000 \times g$ . Microbial pellets were suspended in 500  $\mu\text{L}$  of lysis buffer (4% sodium dodecyl sulfate [SDS], 50 mM Tris-HCl [pH 8.0]) and heated for 10 min at  $95^{\circ}\text{C}$ . Four sonication cycles (30 s with a 1-min interval on ice) were performed, with an amplitude of 40%. Silica beads were then added

(0.3 g) to each sample, and 5 rounds of bead beating were performed (30 s with a 5-min interval on ice) at a speed of  $6.5 \text{ ms}^{-1}$ . Centrifugation at  $14,000 \times g$  was performed to remove the beads and cell debris. To remove the SDS, proteins were precipitated using methanol-chloroform and suspended in 8 M urea for in-solution trypsin digestion. Lastly, the peptides were quantified in a Qubit fluorimeter (Thermo Scientific), and 1  $\mu\text{g}$  of peptides was loaded for reverse-phase nano-liquid chromatography electrospray ionization tandem mass spectrometric analysis on an EASY-nLC 1000 system (Proxeon) coupled to a Q-Exactive HF mass spectrometer (Thermo Scientific). Peptides were loaded on-line onto an Acclaim PepMap 100 trapping column (75  $\mu\text{m}$  [inside diameter] by 20 mm, 3  $\mu\text{m}$ ,  $C_{18}$  resin with 100-Å pore size; Thermo Scientific) using buffer A (0.1% formic acid) and then separated on a  $C_{18}$  resin analytical column (75  $\mu\text{m}$  [inside diameter] by 500 mm, 2  $\mu\text{m}$ , 100-Å pore size; EASY-Spray column; Thermo Scientific). A 240-min gradient from 2% to 40% buffer B (0.1% formic acid in 100% acetonitrile) in buffer A was performed to separate the peptides.

Data were obtained by data-dependent acquisition in positive mode. From each MS scan (between 350 and 2,000 Da), the 15 most intense precursors (charges between 2+ and 5+) were selected for their high collision energy dissociation fragmentation, with a dynamic exclusion of 10 s and a normalized collision energy of 20, and the corresponding MS/MS spectra were acquired. Peptides were eluted using a 240-min gradient. The MS proteomics raw data have been submitted to the ProteomeXchange Consortium (<http://www.proteomexchange.org>) via the Proteomics Identifications Database partner repository with the database identifier PXD029284 (47). For data processing, we employed MetaLab software, which provides a human gut microbial database (12, 48). A human database downloaded from UniProt DB (<http://www.uniprot.org>) (49), restricted to human taxonomy (downloaded on 18 February 2020 with 74,451 sequences), was also used to identify human proteins. For peptide and protein identification, the FDR was set to 0.01. Only taxa identified with at least 2 peptides were considered and were manually filtered to eliminate human peptides. The sum of the intensities of all the distinctive peptides assigned to a taxon was used as the relative abundance of that taxon. A taxon-function analysis was also performed, using taxonomic information from the enrichment analysis from the iMetaLab platform (<http://shiny.imetalab.ca/>) (50).

**Graphs and statistical analysis.** The data were trimmed by selecting only the taxa identified by both methods (shared taxa), solving annotation incongruences. First, we evaluated the dichotomous dependent variable, "taxa presence," by a McNemar test with continuity correction at the phylum, class, and genus levels, considering in each sample any amplicon sequence variant (of at least 2 peptides) per taxon. Significance was set to a *P* value of  $<0.05$ . Second, a differential abundance analysis was performed with taxa detected in more than 50% of the samples by at least one of the approaches. Compositional abundance of each phylum, class, and genus was normalized per sample, dividing by total bacterial read counts assigned to the specific taxonomical level and by total sum of LFQ intensity detected at that bacterial taxonomical level for AS and metaproteomics, respectively. Alpha diversity metrics (Shannon index and species richness) were calculated with `estimate_richness` function from the `phyloseq` package (version 1.40.0) of RStudio using genus abundances (LFQ or read counts). Package `ggpurb` (version 0.4.0) was used for plotting paired data. GraphPad Prism (version 9.1.1.225) was used for the functional statistical analysis of the microbial and human proteins.

Nonparametric Wilcoxon signed-rank paired tests were performed to compare the populations' mean ranks (individual taxa and alpha diversity metrics), pairing by sample ID. Significance values in Wilcoxon tests were adjusted via a Benjamini-Hochberg (BH) FDR of 0.1 for AS and metaproteomics abundances.

The CF infant microbiota through the first months of life was profiled using the `splinctomeR` package (version 0.1.0) in RStudio (version 4.2.1). Differences in the top taxon relative abundances over the time series detected by each method were evaluated with the `permuspliner` function (999 permutations). Likewise, each taxon was investigated individually to detect any divergence between methods at a particular time point along the time course with the `slidingspliner` function. Finally, we investigated the time effect for the top taxon relative abundance trends along the time series detected by each method separately using the `trendyspliner` function. The same approach was used to evaluate differences between methods for alpha diversity metrics. Significance values were adjusted by an FDR of 0.1.

Significant differences were verified with linear models adjusting for age, birth feeding (maternal only/formula), cystic fibrosis transmembrane conductance regulator (CFTR) mutation type (high risk/low risk), and respiratory exacerbations (yes/no), which were considered if one or more episodes were suffered during the sample collection.

**Data availability.** The sequence data from this study are deposited in the NCBI's SRA repository, BioProject ID [PRJNA719717](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA719717). The proteomics data set from this paper has been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier [PXD029284](https://www.ebi.ac.uk/pride/archive/study/PXD029284).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, XLSX file, 0.01 MB.

**SUPPLEMENTAL FILE 2**, XLSX file, 0.1 MB.

**SUPPLEMENTAL FILE 3**, XLSX file, 0.1 MB.

**SUPPLEMENTAL FILE 4**, XLSX file, 0.2 MB.

**SUPPLEMENTAL FILE 5**, XLSX file, 0.02 MB.

**SUPPLEMENTAL FILE 6**, XLSX file, 0.03 MB.

**SUPPLEMENTAL FILE 7**, XLSX file, 0.01 MB.

## ACKNOWLEDGMENTS

We thank the families of the children for their participation and all the health personnel of the CF units, and we thank Marta Cobo for technical assistance.

R.D.C. is the recipient of a Vertex Pharmaceuticals grant. The other authors declare no conflicts of interest related to this work.

This work was supported by the Instituto de Salud Carlos III (PI17/00115 and PI20/00164 to R.D.C.), REIPI 307 (RD16/0016/0011) actions, cofinanced by the European Development Regional Fund “A way to achieve Europe” (ERDF), Vertex Pharmaceuticals. C.S. has grant support from “Fundación Mutua Madrileña” in Atenas Internacional (ATH) 017 call to R.D.C. (AP165902017). Proteomic experiments were supported by projects RTI2018-094004-B-I00 and InGEMICS-CM B2017/BMD3691 via funds to C.G.

## REFERENCES

- Selma-Royo M, Calatayud Arroyo M, García-Mantrana I, Parra-Llorca A, Escuriel R, Martínez-Costa C, Collado MC. 2020. Perinatal environment shapes microbiota colonization and infant growth: impact on host response and intestinal function. *Microbiome* 8:167. <https://doi.org/10.1186/s40168-020-00940-8>.
- Cuthbertson L, Walker AW, Oliver AE, Rogers GB, Rivett DW, Hampton TH, Ashare A, Elborn JS, De Soyza A, Carroll MP, Hoffman LR, Lanyon C, Moskowitz SM, O'Toole GA, Parkhill J, Planet PJ, Teneback CC, Tunney MM, Zuckerman JB, Bruce KD, van der Gast CJ. 2020. Lung function and microbiota diversity in cystic fibrosis. *Microbiome* 8:45. <https://doi.org/10.1186/s40168-020-00810-3>.
- Dayama G, Priya S, Niccum DE, Khoruts A, Blekhnman R. 2020. Interactions between the gut microbiome and host gene regulation in cystic fibrosis. *Genome Med* 12:12. <https://doi.org/10.1186/s13073-020-0710-2>.
- Andrinopoulou ER, Clancy JP, Szczesniak RD. 2020. Multivariate joint modeling to identify markers of growth and lung function decline that predict cystic fibrosis pulmonary exacerbation onset. *BMC Pulm Med* 20:142. <https://doi.org/10.1186/s12890-020-1177-z>.
- Burke DG, Fouhy F, Harrison MJ, Rea MC, Cotter PD, O'Sullivan O, Stanton C, Hill C, Shanahan F, Plant BJ, Ross RP. 2017. The altered gut microbiota in adults with cystic fibrosis. *BMC Microbiol* 17:58. <https://doi.org/10.1186/s12866-017-0968-8>.
- Price CE, O'Toole GA. 2021. The gut-lung axis in cystic fibrosis. *J Bacteriol* 203:e00311-21. <https://doi.org/10.1128/JB.00311-21>.
- Duytschaever G, Huys G, Bekaert M, Boulanger L, De Boeck K, Vandamme P. 2011. Cross-sectional and longitudinal comparisons of the predominant fecal microbiota compositions of a group of pediatric patients with cystic fibrosis and their healthy siblings. *Appl Environ Microbiol* 77:8015–8024. <https://doi.org/10.1128/AEM.05933-11>.
- Schipa S, Iebba V, Santangelo F, Gagliardi A, De Biase RV, Stamato A, Bertasi S, Lucarelli M, Conte MP, Quattrucci S. 2013. Cystic fibrosis transmembrane conductance regulator (CFTR) allelic variants relate to shifts in faecal microbiota of cystic fibrosis patients. *PLoS One* 8:e61176. <https://doi.org/10.1371/journal.pone.0061176>.
- Antoska KM, Chernikova DA, Price CE, Ruoff KL, Li K, Guill MF, Sontag NR, Morrison HG, Hao S, Drumm ML, MacKenzie TA, Dorman DB, Feenan LM, Williams MA, Dessaint J, Yuan IH, Aldrich BJ, Moulton LA, Ting L, Martinez-del Campo A, Stewart EJ, Karagas MR, O'Toole GA, Madan JC. 2019. Altered stool microbiota of infants with cystic fibrosis shows a reduction in genera associated with immune programming from birth. *J Bacteriol* 201:e00274-19. <https://doi.org/10.1128/JB.00274-19>.
- Knight R, Vrbancic A, Taylor BC, Aksenov A, Callewaert C, Debelius J, Gonzalez A, Kosciolk T, McCall L-I, McDonald D, Melnik AV, Morton JT, Navas J, Quinn RA, Sanders JG, Swafford AD, Thompson LR, Tripathi A, Xu ZZ, Zaneveld JR, Zhu Q, Caporaso JG, Dorrestein PC. 2018. Best practices for analysing microbiomes. *Nat Rev Microbiol* 16:410–422. <https://doi.org/10.1038/s41579-018-0029-9>.
- Rintala A, Pietilä S, Munukka E, Eerola E, Pursiheimo J-P, Laiho A, Pekkala S, Huovinen P. 2017. Gut microbiota analysis results are highly dependent on the 16S rRNA gene target region, whereas the impact of DNA extraction is minor. *J Biomol Tech* 28:19–30. <https://doi.org/10.1171/jbt.17-2801-003>.
- Cheng K, Ning Z, Zhang X, Li L, Liao B, Mayne J, Figeys D. 2020. MetaLab 2.0 enables accurate post-translational modifications profiling in metaproteomics. *J Am Soc Mass Spectrom* 31:1473–1482. <https://doi.org/10.1021/jasms.0c00083>.
- Zhang X, Li L, Butcher J, Stintzi A, Figeys D. 2019. Advancing functional and translational microbiome research using meta-omics approaches. *Microbiome* 7:154. <https://doi.org/10.1186/s40168-019-0767-6>.
- Zhang X, Figeys D. 2019. Perspective and guidelines for metaproteomics in microbiome studies. *J Proteome Res* 18:2370–2380. <https://doi.org/10.1021/acs.jproteome.9b00054>.
- Mayers MD, Moon C, Stupp GS, Su AI, Wolan DW. 2017. Significant alterations in protein expression from a mouse model of inflammatory bowel disease. *J Proteome Res* 16:1014–1026. <https://doi.org/10.1021/acs.jproteome.6b00938>.
- Zhang X, Li L, Mayne J, Ning Z, Stintzi A, Figeys D. 2018. Assessing the impact of protein extraction methods for human gut metaproteomics. *J Proteomics* 180:120–127. <https://doi.org/10.1016/j.jprot.2017.07.001>.
- Isaac NI, Philippe D, Nicholas A, Raouf D, Eric C. 2019. Metaproteomics of the human gut microbiota: challenges and contributions to other OMICS. *Clin Mass Spect* 14:18–30. <https://doi.org/10.1016/j.clinms.2019.06.001>.
- Debyser G, Mesuere B, Clement L, Van de Weygaert J, Van Hecke P, Duytschaever G, Aerts M, Dawyndt P, De Boeck K, Vandamme P, Devreese B. 2016. Faecal proteomics: a tool to investigate dysbiosis and inflammation in patients with cystic fibrosis. *J Cyst Fibros* 15:242–250. <https://doi.org/10.1016/j.jcf.2015.08.003>.
- Petriz BA, Franco OL. 2017. Metaproteomics as a complementary approach to gut microbiota in health and disease. *Front Chem* 5:4. <https://doi.org/10.3389/fchem.2017.00004>.
- Brown CT, Xiong W, Olm MR, Thomas BC, Baker R, Firek B, Morowitz MJ, Hettich RL, Banfield JF. 2018. Hospitalized premature infants are colonized by related bacterial strains with distinct proteomic profiles. *mBio* 9:e00441-18. <https://doi.org/10.1128/mBio.00441-18>.
- Debyser G, Aerts M, Van Hecke P, Mesuere B, Duytschaever G, Dawyndt P, De Boeck K, Vandamme P, Devreese B. 2019. A method for comprehensive proteomic analysis of human faecal samples to investigate gut dysbiosis in patients with cystic fibrosis. *Adv Exp Med Biol* 1073:137–160. [https://doi.org/10.1007/978-3-030-12298-0\\_6](https://doi.org/10.1007/978-3-030-12298-0_6).
- Lee AJ, Einarsson GG, Gilpin DF, Tunney MM. 2020. Multi-omics approaches: the key to improving respiratory health in people with cystic fibrosis? *Front Pharmacol* 11:569821. <https://doi.org/10.3389/fphar.2020.569821>.
- Bevivino A, Bacci G, Drevinek P, Nelson MT, Hoffman L, Mengoni A. 2019. Deciphering the ecology of cystic fibrosis bacterial communities: towards systems-level integration. *Trends Mol Med* 25:1110–1122. <https://doi.org/10.1016/j.molmed.2019.07.008>.
- Rueda-Nieto S, Mondejar-Lopez P, Mira-Escolano MP, Cutillas-Tolín A, Maceda-Roldán LA, Arense-Gonzalo JJ, Palomar-Rodríguez JA. 2022. Analysis of the genotypic profile and its relationship with the clinical manifestations in people with cystic fibrosis: study from a rare disease registry. *Orphanet J Rare Dis* 17:222. <https://doi.org/10.1186/s13023-022-02373-y>.
- Meeker SM, Mears KS, Sangwan N, Brittnacher MJ, Weiss EJ, Treuting PM, Tolley N, Pope CE, Hager KR, Vo AT, Paik J, Frevert CW, Hayden HS, Hoffman LR, Miller SI, Hajjar AM. 2020. CFTR dysregulation drives active selection of the gut microbiome. *PLoS Pathog* 16:e1008251. <https://doi.org/10.1371/journal.ppat.1008251>.
- Martinović A, Cocuzzi R, Arioli S, Mora D. 2020. *Streptococcus thermophilus*: to survive, or not to survive the gastrointestinal tract, that is the question! *Nutrients* 12:2175. <https://doi.org/10.3390/nu12082175>.
- Egan M, Dempsey E, Ryan CA, Ross RP, Stanton C. 2021. The sporobiota of the human gut. *Gut Microbes* 13:1. <https://doi.org/10.1080/19490976.2020.1863134>.
- Zhang X, Chen W, Ning Z, Mayne J, Mack D, Stintzi A, Tian R, Figeys D. 2017. Deep metaproteomics approach for the study of human microbiomes. *Anal Chem* 89:9407–9415. <https://doi.org/10.1021/acs.analchem.7b02224>.

29. García-Durán C, Martínez-López R, Zapico I, Pérez E, Romeu E, Arroyo J, Hernáez ML, Pitarch A, Monteoliva L, Gil C. 2021. Distinct human gut microbial taxonomic signatures uncovered with different sample processing and microbial cell disruption methods for metaproteomic analysis. *Front Microbiol* 12:618566. <https://doi.org/10.3389/fmicb.2021.618566>.
30. Mancabelli L, Milani C, Lugli GA, Fontana F, Turrone F, van Sinderen D, Ventura M. 2020. The impact of primer design on amplicon-based metagenomic profiling accuracy: detailed insights into bifidobacterial community structure. *Microorganisms* 8:131. <https://doi.org/10.3390/microorganisms8010131>.
31. Coffey MJ, Nielsen S, Wemheuer B, Kaakoush NO, Garg M, Needham B, Pickford R, Jaffe A, Thomas T, Ooi CY. 2019. Gut microbiota in children with cystic fibrosis: a taxonomic and functional dysbiosis. *Sci Rep* 9: 18593. <https://doi.org/10.1038/s41598-019-55028-7>.
32. Chua HH, Chou HC, Tung YL, Chiang BL, Liao CC, Liu HH, Ni YH. 2018. Intestinal dysbiosis featuring abundance of *Ruminococcus gnavus* associates with allergic diseases in infants. *Gastroenterology* 154:154–167. <https://doi.org/10.1053/j.gastro.2017.09.006>.
33. Wu H, Rebello O, Crost EH, Owen CD, Walpole S, Bennati-Granier C, Ndeh D, Monaco S, Hicks T, Colville A, Urbanowicz PA, Walsh MA, Angulo J, Spencer DIR, Juge N. 2021. Fucosidases from the human gut symbiont *Ruminococcus gnavus*. *Cell Mol Life Sci* 78:675–693. <https://doi.org/10.1007/s00018-020-03514-x>.
34. Gren C, Spiegelhauer MR, Rotbain EC, Ehmsen BK, Kampmann P, Andersen LP. 2019. *Ruminococcus gnavus* bacteraemia in a patient with multiple haematological malignancies. *Access Microbiol* 1:e000048. <https://doi.org/10.1099/acmi.0.000048>.
35. Crost EH, Tailford LE, Monestier M, Swarbreck D, Henrissat B, Crossman LC, Juge N. 2016. The mucin-degradation strategy of *Ruminococcus gnavus*: the importance of intramolecular trans-sialidases. *Gut Microbes* 7: 302–312. <https://doi.org/10.1080/19490976.2016.1186334>.
36. Azzouz D, Omarbekova A, Heguy A, Schwudke D, Gisch N, Rovin BH, Caricchio R, Buyon JP, Alekseyenko AV, Silverman GJ. 2019. Lupus nephritis is linked to disease-activity associated expansions and immunity to a gut commensal. *Ann Rheum Dis* 78:947–956. <https://doi.org/10.1136/annrheumdis-2018-214856>.
37. Resch A, Rosenstein R, Nerz C, Götz F. 2005. Differential gene expression profiling of *Staphylococcus aureus* cultivated under biofilm and planktonic conditions. *Appl Environ Microbiol* 71:2663–2676. <https://doi.org/10.1128/AEM.71.5.2663-2676.2005>.
38. Fourie KR, Wilson HL. 2020. Understanding GroEL and DnaK stress response proteins as antigens for bacterial diseases. *Vaccines (Basel)* 8:773. <https://doi.org/10.3390/vaccines8040773>.
39. Legrand D, Pierce A, Elass E, Carpentier M, Mariller C, Mazurier J. 2008. Lactoferrin structure and functions. *Adv Exp Med Biol* 606:163–194. [https://doi.org/10.1007/978-0-387-74087-4\\_6](https://doi.org/10.1007/978-0-387-74087-4_6).
40. Aratani Y. 2018. Myeloperoxidase: its role for host defense, inflammation, and neutrophil function. *Arch Biochem Biophys* 640:47–52. <https://doi.org/10.1016/j.abb.2018.01.004>.
41. Subramanian I, Verma S, Kumar S, Jere A, Anamika K. 2020. Multi-omics data integration, interpretation, and its application. *Bioinform Biol Insights* 14: 1177932219899051. <https://doi.org/10.1177/1177932219899051>.
42. Mirzayi C, Renson A, Zohra F, Elsaforay S, Geistlinger L, Kasselman L, Eckenrode K, van de Wijgert J, Loughman A, Marques FZ, STORMS Consortium, Genomic Standards Consortium, Massive Analysis and Quality Control Society, Segata N, Huttenhower C, Dowd JB, Jones HE, Levi Waldron L. 2021. Strengthening The Organization and Reporting of Microbiome Studies (STORMS): a reporting checklist for human microbiome research. *bioRxiv*. <https://doi.org/10.1101/2020.06.24.167353>.
43. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41:e1. <https://doi.org/10.1093/nar/gks808>.
44. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Gregory Caporaso J. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6:90. <https://doi.org/10.1186/s40168-018-0470-z>.
45. Long S, Yang Y, Shen C, Wang Y, Deng A, Qin Q, Qiao L. 2020. Metaproteomics characterizes human gut microbiome function in colorectal cancer. *NPJ Biofilms Microbiomes* 6:14. <https://doi.org/10.1038/s41522-020-0123-4>.
46. Abbondio M, Palomba A, Tanca A, Fraumene C, Pagnozzi D, Serra M, Marongiu F, Laconi E, Uzzau S. 2019. Fecal metaproteomic analysis reveals unique changes of the gut microbiome functions after consumption of sourdough *carasau* bread. *Front Microbiol* 10:1733. <https://doi.org/10.3389/fmicb.2019.01733>.
47. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz S, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, Ternent T, Brazma A, Vizcaino JA. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 47:D442–D450. <https://doi.org/10.1093/nar/gky1106>.
48. Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, Arumugam M, Kultima JR, Prifti E, Nielsen T, Juncker AS, Manichanh C, Chen B, Zhang W, Levenez F, Wang J, Xu X, Xiao L, Liang S, Zhang D, Zhang Z, Chen W, Zhao H, Al-Aama JY, Edris S, Yang H, Wang J, Hansen T, Nielsen HB, Brunak S, Kristiansen K, Guarner F, Pedersen O, Doré J, Ehrlich SD, Bork P, Wang J, MetaHIT Consortium. 2014. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* 32:834–841. <https://doi.org/10.1038/nbt.2942>.
49. UniProt Consortium. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47:D506–D515. <https://doi.org/10.1093/nar/gky1049>.
50. Liao B, Ning Z, Cheng K, Zhang X, Li L, Mayne J, Figeys D. 2018. iMetaLab 1.0: a web platform for metaproteomics data analysis. *Bioinformatics* 34: 3954–3956. <https://doi.org/10.1093/bioinformatics/bty466>.

## RESEARCH ARTICLE

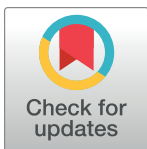
# Influence of a *Serratia marcescens* outbreak on the gut microbiota establishment process in low-weight preterm neonates

Esperanza Escribano<sup>1</sup>✉, Claudia Saralegui<sup>2</sup>✉, Laura Moles<sup>2,3</sup>, María Teresa Montes<sup>1</sup>, Claudio Alba<sup>4</sup>, Teresa Alarcón<sup>4</sup>, Fernando Lázaro-Perona<sup>5</sup>, Juan Miguel Rodríguez<sup>3</sup>, Miguel Sáenz de Pipaón<sup>1,6</sup>, Rosa del Campo<sup>1,6</sup>✉\*

**1** Servicio de Neonatología, Hospital Universitario La Paz, and Universidad Autónoma de Madrid, Madrid, Spain, **2** Servicio de Microbiología y Parasitología, Hospital Universitario Ramón y Cajal, and Instituto Ramón y Cajal de Investigaciones Sanitarias (IRYCIS), Madrid, Spain, **3** Departamento de Bromatología, Facultad de Veterinaria Nutrición y Ciencia de los Alimentos, Universidad Complutense de Madrid, Madrid, Spain, **4** Servicio de Microbiología, Hospital Universitario La Princesa, and Universidad Autónoma de Madrid, Madrid, Spain, **5** Servicio de Microbiología, Hospital Universitario La Paz, Madrid, Spain, **6** Red de Salud Materno Infantil y del Desarrollo, Instituto de Salud Carlos III, Madrid, Spain

✉ These authors contributed equally to this work.

\* [rosacampo@yahoo.com](mailto:rosacampo@yahoo.com)



## OPEN ACCESS

**Citation:** Escribano E, Saralegui C, Moles L, Montes MT, Alba C, Alarcón T, et al. (2019) Influence of a *Serratia marcescens* outbreak on the gut microbiota establishment process in low-weight preterm neonates. PLoS ONE 14(5): e0216581. <https://doi.org/10.1371/journal.pone.0216581>

**Editor:** Anatoly V. Grishin, Children's Hospital Los Angeles, UNITED STATES

**Received:** November 10, 2018

**Accepted:** April 25, 2019

**Published:** May 21, 2019

**Copyright:** © 2019 Escribano et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The clinical data have been uploaded as a Supporting Information file. Additionally, the genome sequences are deposited in the European Nucleotide Archive database with the accession numbers QYRU00000000 and QYSA00000000 for Clone A (submission numbers SUB4493714 and SUB4514092); and QYRV00000000 and QYSB00000000 for Clone B (submission numbers SUB4510090 and SUB4525283). Fasta files are deposited in the NCBI

## Abstract

Adequate gut microbiota establishment is important for lifelong health. The aim was to sequentially analyze the gut microbiota establishment in low-birth-weight preterm neonates admitted to a single neonatal intensive care unit during their first 3 weeks of life, comparing two epidemiological scenarios. Seven control infants were recruited, and another 12 during a severe *S. marcescens* outbreak. Meconium and feces from days 7, 14, and 21 of life were collected. Gut microbiota composition was determined by 16S rDNA massive sequencing. Cultivable isolates were genotyped by pulsed-field gel electrophoresis, with four *S. marcescens* submitted for whole-genome sequencing. The expected bacterial ecosystem expansion after birth is delayed, possibly related to antibiotic exposure. The *Proteobacteria* phylum dominates, although with marked interindividual variability. The outbreak group considerably differed from the control group, with higher densities of *Escherichia coli* and *Serratia* to the detriment of *Enterococcus* and other *Firmicutes*. Curiously, obligate predators were only detected in meconium and at very low concentrations. Genotyping of cultivable bacteria demonstrated the high bacterial horizontal transmission rate that was confirmed with whole-genome sequencing for *S. marcescens*. Preterm infants admitted at NICU are initially colonized by homogeneous microbial communities, most of them from the nosocomial environment, which subsequently evolve according to the individual conditions. Our results demonstrate the hospital epidemiology pressure, particularly during outbreak situations, on the gut microbiota establishing process.

web site under the Bioproject PRJNA510235 reference.

**Funding:** Claudio Alba was supported by “Consejería de Educación, Juventud y Deporte de la Comunidad de Madrid” and European Social Funds. Claudia Saralegui was supported by “Fundación Mutua Madrileña” grant to Rosa del Campo achieved in 2017 call with reference number AP165902017.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

The relationship between our human cells and the microbial communities living inside us can be classified as mutualistic, commensal, or pathogenic. This consideration delimits the fine barrier that distinguishes colonization from infection, and can fluctuate over time due to influence from the host or from microbial and environmental factors. The influence of the intestinal microbiota on global human health has been confirmed [1–3], forming new research perspectives aimed to optimize the composition and the functionality of this bionetwork.

Adequate microbiota establishment in newborns is a process particularly relevant for their lifelong health [4], and its management represents a scientific challenge. Human bacterial colonization might start *in utero*, but the critical step begins with the exposure to maternal bacteria at birth and during the early postnatal period [5–6]. Bacterial populations fluctuate considerably during the first months of life, until a stable ecosystem is established when the infant is approximately 2–3 years of age [7–8]. Universal criteria defining a “normal” or “healthy” gut microbiota have not yet been established, which should be characterized by a high diversity, marked inter-individual variability, and conserved intraindividual stability. However, the composition of this ecosystem is influenced by numerous factors [9–10], such as the gestational age at which the neonate is born [11–12].

Preterm birth is the main cause of perinatal morbidity and mortality, as well as an important risk factor for death in the first 5 years of life [13]. A considerable increase in preterm birth rates over the past two decades has been reported worldwide, in both developed and developing countries [14]. In Spain, the preterm birth rate of all live births increased from 7.1% in 1996 to 8.2% in 2008, which is one of the highest rates in Europe [15]. This global tendency can be explained by several factors, including an early or advanced mother’s age, a small gap between pregnancies, low body mass index, multiple pregnancy, history of infectious diseases, stress, alcohol consumption, and periodontal disease [13]. Nevertheless, approximately half of the spontaneous preterm births have an unidentified cause, and it has been suggested that the composition of the maternal microbiota could play a relevant triggering role [7,16].

In low-birth-weight preterm infants (<2500 g), the gut microbiota composition and their biodiversity are aberrant, given the bacterial establishment is delayed by their prolonged hospital stay and their intense exposure to antimicrobials [12,17–19]. This fact explains, at least partially, why preterm infants have a very immature immune system and typically experience infectious complications [20–22]. In this context, *Serratia marcescens* is one of the most relevant emerging pathogens causing severe outbreaks in this population [23–24]. Pathogenic gut colonization during nosocomial outbreaks has frequently been reported; to our knowledge, however, the influence of an outbreak on the microbiota establishment process has not thus far been studied.

The aim of the present study was to sequentially analyze the gut microbiota establishment of low-birth-weight preterm neonates admitted to a single neonatal intensive care unit (NICU) during their first 3 weeks of life, comparing two epidemiological scenarios: a normal period and a period with a nosocomial *S. marcescens* outbreak.

## Materials and methods

### Preterm neonate inclusion criteria and sampling procedure

La Paz University Hospital (Madrid, Spain) has a 23-bed level III NICU, from which 19 low-birth-weight preterm neonates (<32 weeks gestational age) were recruited in two separate periods: (A) during an epidemiologically normal period in 2015 (control group, n = 7); and (B) during a severe *S. marcescens* outbreak from December 2016 to March 2017 (outbreak

group,  $n = 12$ ). Despite the different sampling periods, there were no significant changes in the NICU. It is important to note the data lack in the control group about antibiotic consumption, and clinical data, as a limitation of our work. From each preterm infant, four fecal samples were collected after birth: meconium, and feces from 7, 14, and 21 days of life. Although our initial aim was to extend the recruitment period, logistical limitations limited the study to the first three weeks. The samples were directly recovered from the diaper using a sterile plastic stick and immediately stored at  $-80^{\circ}\text{C}$ . Although our intention was to collect fecal samples immediately after deposition, we cannot rule out the possible contact and contamination with urine. However, the contribution of the urinary microbiota should be insignificant.

For the control group, only DNA from fecal samples was available, whereas for the outbreak group bacterial growth was obtained by culture-dependent techniques in addition to DNA. The ethics committee “Comité Ético de Investigación Clínica del Hospital Universiatrio La Paz” approved the study (reference HULP3551), and the data of all the neonates were obtained from their clinical chart. The infants were categorized according to four variables: 1) epidemiological situation (normal or *S. marcescens* outbreak); 2) delivery mode (vaginal or C-section); 3) gestational age (extremely preterm:  $<28$  weeks; very preterm: 28–30 weeks; or moderately preterm: 30–32 weeks); and 4) birth weight ( $<1000$  g; 1000–1500 g; or  $>1500$  g).

### Sample processing

Fecal samples from the *S. marcescens* outbreak group were slowly defrosted at  $-20^{\circ}\text{C}$  for 24 h and  $4^{\circ}\text{C}$  for another 24 h, in order to avoid bacterial death. Portions between 0.3–0.5 g of each sample were inoculated into Brain Heart Infusion (BHI) broth (Difco, Detroit, Michigan) and incubated at  $37^{\circ}\text{C}$  for 24 h as a bacterial pre-enrichment. Cultivable bacteria were isolated in selective and nonselective agar media from the BHI tube, including agar plates of *M-Enterococcus*; De Man, Rogosa and Sharpe (MRS); mannitol salt; McConkey; and Columbia, with 5% sheep blood. The culture media were purchased from Difco, and the plates were incubated at  $37^{\circ}\text{C}$  for 24–48 h, including 5%  $\text{CO}_2$  for the blood agar plates, and anaerobic conditions for the MRS plates. Colony identification was performed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (Bruker, Germany), and all the isolates were conserved at  $-80^{\circ}\text{C}$  in semi-skimmed milk. In parallel to bacterial cultures, total DNA was obtained from fecal aliquots of 0.3–0.5 g with the QiaAMP kit (Qiagen, Germany), determining their concentration and quality by Qubit fluorometer (Invitrogen, USA).

### 16S rDNA next-generation sequencing

The fecal DNA samples were sent to FISABIO (Valencia, Spain) for massive Mi-Seq  $2 \times 300$  bp paired-end Illumina 16S rDNA sequencing (Cod. 15044223 Rev. A) from the V3 and V4 regions, which were amplified with the following primers: (Forward Primer: 5' -TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG), (Reverse Primer: 5' -GTC TCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVG GGTATCTAATCC) [25]. Shannon and Chao1 indexes were used for alpha bacterial diversity estimation and were calculated, eliminating taxa with fewer than three lectures. Taxonomic affiliations were assigned using the Ribosomal Database Project (RDP) classifier, and reads with an RDP score below 0.8 were assigned to the upper taxonomic rank, leaving the last rank as unidentified. Sequence quality was measured according to the following parameters: minimum length, 250 bp; trimming quality measure type, mean; trimming quality number from 3' extreme, 30; trimming quality window, 10 bp. Relative abundance and contingency tables included singletons and very low-represented taxa. The statistical analysis was performed using R statistical software and several open source libraries. The quantitative data of the reads were homogenized using their

relative percentage from the total reads of each sample to facilitate the comparison between samples. Finally, the Galaxy Huttenhower Platform (<http://huttenhower.sph.harvard.edu/galaxy>) was used in order to calculate the Linear Discriminant Effect Size Analysis (LEfSe) algorithm and to obtain cladograms in which microbial taxa that explain significant differences among groups of samples were represented. A free software platform was used according to paper instructions [26]. For the statistical analysis, samples with fewer than 1,000 reads were dismissed, and all the samples from patient 7B were excluded due to a lack of some demographic and perinatal information. Fasta files are deposited in the NCBI web site under the Bioproject PRJNA510235 reference.

### Pulse-field gel electrophoresis typing

Cultivable bacterial isolates were genotyped by pulse-field gel electrophoresis (PFGE), using the habitual particular settings for the PFGE protocol and also for the restriction enzymes (*Sma*I for staphylococci and enterococci, *Xba*I for *Escherichia coli*, and finally *Spe*I for *Serratia* and *Klebsiella*). The PFGE pattern analysis was made with Phoretix 5.0 software (TotalLab, Newcastle upon Tyne, UK), and the representation of the results was made based on Dice coefficients and the unweighted pair group method with arithmetic mean algorithm.

### Whole genome sequencing

Four *S. marcescens* isolates from different infants were submitted to whole genome sequencing (WGS) by MiSeq technology (Illumina), and the genetic relationships were analyzed in the Galaxy Huttenhower Platform. The genome sequences are deposited in the European Nucleotide Archive database with the accession numbers QYRU00000000 and QYSA00000000 for Clone A, and QYRV00000000 and QYSB00000000 for Clone B.

### Fungi identification

The internal transcribed spacer (ITS)-1 region was amplified from the total fecal DNA using polymerase chain reaction (PCR) in order to analyze fungi diversity, using the primers ITS1-F (5' -CGCCCGCCGCGCGCGCGGGCGG GCGGGGGCACGGGGGGCTTGGTCATTTAGAGGAAGTAA-3') and ITS1-R (5' -TCCTCCGCTTATTGATATGC-3') [27]. Afterward, the amplicons were separated with denaturing gradient gel electrophoresis (DGGE), using the D-CODE system (BioRad Laboratories, USA). The gel bands were cut, reamplified, and sequenced in an AQBI Prism 7000 apparatus.

### Statistical analysis

The Kruskal-Wallis index was used for differences between two groups of samples (comparing medians) for each variable of study, and Dunn's test was used when more than two groups were defined regarding a variable of study. The principal component analysis was applied for the multivariate analysis regarding taxonomical data in order to see differences between groups according to the variables of study. Statistical significance was adjusted to  $p < .005$ .

## Results

### Characteristics of the preterm infants

The demographic and clinical characteristics of both newborn groups are shown in [Table 1](#) and in [S1 Dataset](#). The most relevant differences were that the control infants presented a higher weight at birth and a lower incidence of sepsis. One neonate from the control group (14.3%) suffered from early-onset sepsis, whereas late-onset sepsis was microbiologically or

Table 1. Clinical and demographic characteristics of the preterm infants of both groups.

Characteristic	Control Group (7 infants)	Outbreak Group (12 infants)	p value
Weight at birth (g)	1462 (720–1890) <sup>a</sup>	971 (600–1537) <sup>a</sup>	0.009
Gestational Age (weeks)	30 (25–31) <sup>a</sup>	28 (25–31) <sup>a</sup>	0.26
Vaginal delivery (n, %)	5, 71.4%	3, 25%	0.04
Male sex (n, %)	5, 71.4%	2, 16.6%	0.0001
Sepsis (n, %)	1, 14.3%	6, 50%	0.04
Length of stay (days)	14 (5–140) <sup>a</sup>	55 (7–89) <sup>a</sup>	0.1

<sup>a</sup>Values expressed as the median value and the range (between parentheses).

<https://doi.org/10.1371/journal.pone.0216581.t001>

clinically diagnosed in five neonates from the outbreak group (41.7%), being *Staphylococcus epidermidis* and *S. marcescens* the microorganisms implicated. Data on antimicrobial therapy were only available for the outbreak group and were included the administration of prophylactic cefazolin during labor (10 infants), empiric treatment based on ampicillin, gentamicin, and clarithromycin during the first week of life (8 infants), and treatment that included vancomycin (7 infants), cefotaxime (2 infants) piperacillin/tazobactam (2 infants), meropenem (1 infant), and amikacin (3 infants) for the second and third weeks of life.

### Gut microbiota establishment by next-generation sequencing

The number of operational taxonomic units (OTUs) and the alpha diversity indexes of the meconium samples were similar to the further fecal samples (Fig 1), indicating a delayed bacterial establishment process. LEfSe analysis allowed us to explore the differences in microbiota composition between the control and the outbreak groups for the four variables stated above (epidemiological situation, delivery mode, gestational age, and birth weight).

Significant differences in the gut microbiota composition of the control and outbreak groups were detected in meconium and in the day 21 samples and were more relevant in those found at day 21 ( $p = .0024$  at the genus level and  $p = .073$  at the phylum level) (Fig 2). The outbreak group was characterized by a higher proportion of  $\gamma$ -Proteobacteria, related to a higher density of *Serratia*, and with lower proportions of the Firmicutes and Fusobacteria phyla.

In relation to the delivery mode, differences between vaginal delivery and C-section delivery were only significantly different at day 0 ( $p = .0066$  at the genus level and  $p = .0363$  at the phylum level) (Fig 3). Significant differences among bacterial communities regarding gestational age and birth weight were not detected.

Taking into account all 55 samples, the predominant phylum during the first weeks of life of the low-weight preterm infants was Proteobacteria (median $\pm$ SD 70.1%  $\pm$  26.9%, range 0.3%–99.4%), followed by Firmicutes (median $\pm$ SD 22.1%  $\pm$  26.8%; range 0.05–99.4); and although up to another 23 phyla were detected, their contribution was nearly anecdotic (Fig 4A). At the genus level, the abundance of *Escherichia/Shigella* sp. increased over the studied period while that of *Enterococcus* sp. and *Staphylococcus* sp. decreased (Fig 4B).

Reads accounting for the predator bacteria *Bdellovibrio* (2 infants), *Peredibacter* (1 infant), and *Vampirovibrio* (1 infant) were found in meconium samples from the outbreak (3 infants) and the control (1 infant) groups. The relative abundance of predatory species was extremely low (0.004%–0.11%), and none could be detected in the subsequent fecal samples.

*Serratia* was detected in the meconium samples from all the patients in the outbreak group (median value, 5558 OTUs), whereas this genus was considerably less abundant among the meconium samples from the control group (median value, 180 OTUs) (Table 2). Also noticeable was

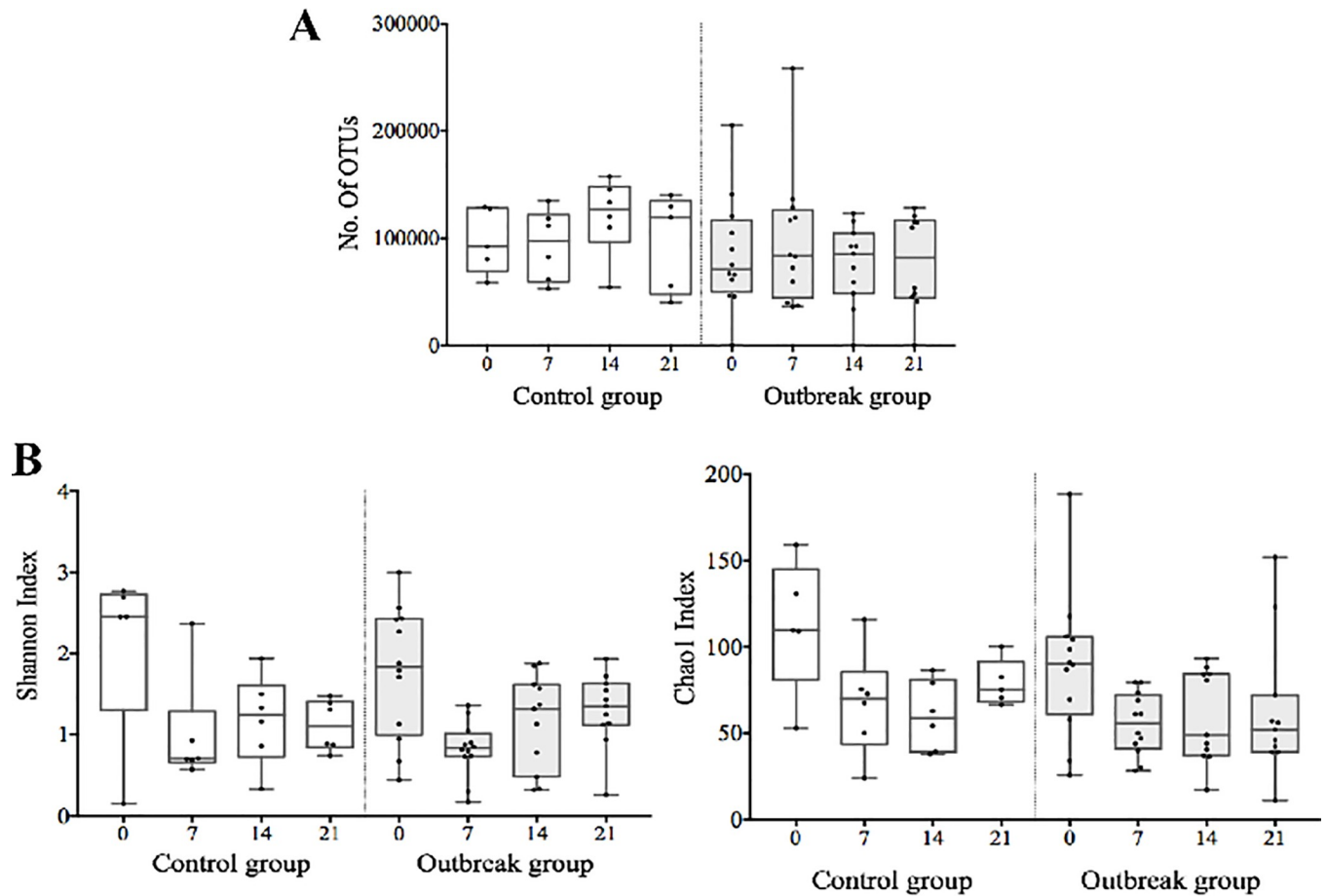


Fig 1. Number of operational taxonomic units (OTUs) (A), and alpha diversity measured by the Chao1 index (B) in all samples studied.

<https://doi.org/10.1371/journal.pone.0216581.g001>

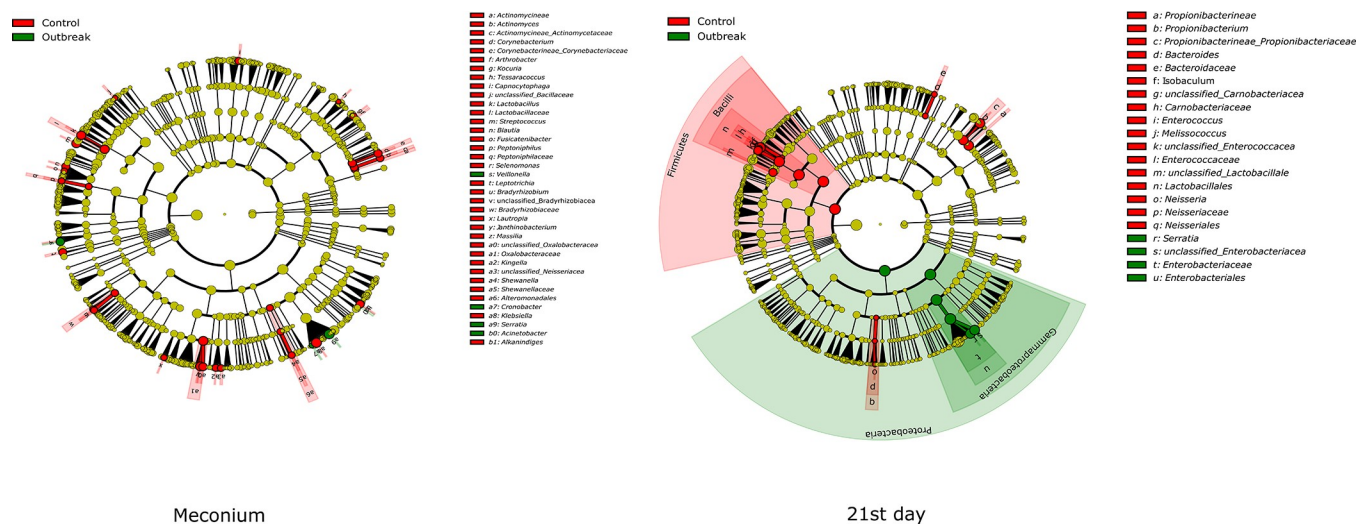


Fig 2. Cladograms showing the significant differences of gut microbiota composition in meconium and 21 days feces between control and outbreak groups.

<https://doi.org/10.1371/journal.pone.0216581.g002>

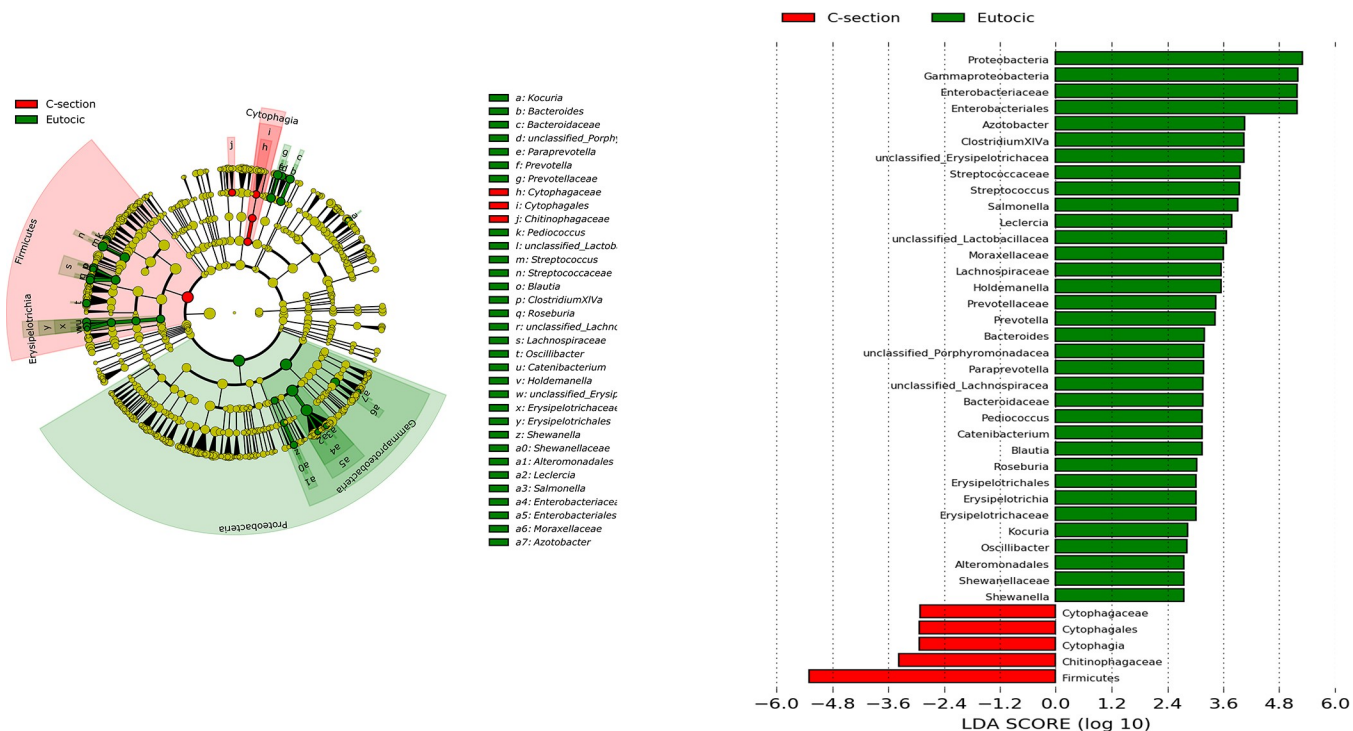


Fig 3. Significant differences in the gut microbiota of meconium by the delivery mode.

<https://doi.org/10.1371/journal.pone.0216581.g003>

that the *Serratia* reads considerably increased just before the diagnosis of *Serratia* sepsis in some patients from the outbreak group, and particularly in infant O11, who finally died from *S. marcescens* sepsis at day 10 after birth. The gut enrichment of *Serratia* sequences in this infant was manifest, reaching levels of 95% of the total intestinal microbiota at day 7 (Table 2).

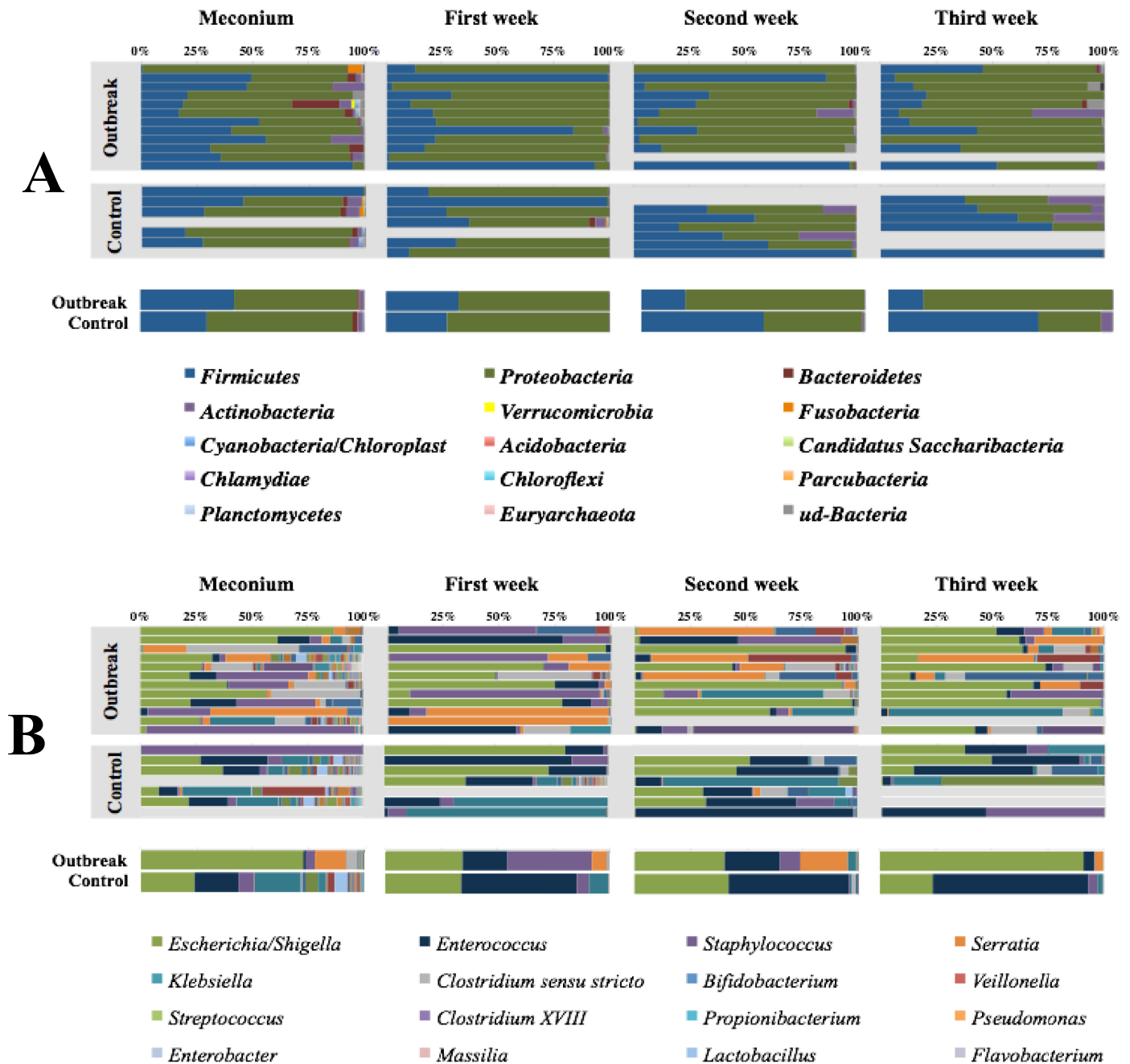
### Characterization of cultivable isolates from the outbreak group

A total of 16 *E. coli* (8 infants), 35 *Enterococcus faecalis* (11 infants), 2 *Enterococcus faecium* (1 infant), 12 *Klebsiella oxytoca* (7 infants), 6 *Klebsiella pneumoniae* (5 infants), 32 *S. epidermidis* (all 12 infants), 1 *Serratia liquefaciens* and 14 *S. marcescens* (8 infants) isolates were recovered. Regarding the meconium samples, the cultivable microorganisms were *S. epidermidis* (7 infants), *E. faecalis* (3 infants), and *S. marcescens* (2 infants). Only four meconium samples (33.3%) did not yield viable microorganisms. PFGE analysis showed that isolates recovered from different samples from the same infant were identical or closely related (Fig 5).

Similarly, a single *E. faecalis* pulsotype was detected, colonizing three infants and two major clones of *K. pneumoniae* in two infants. Among the *Serratia* isolates, two genetically unrelated clones were detected, affecting four and five infants each. The relative abundance of *Serratia* in the gut microbiota of each infant is presented in Table 2, which compares both next-generation sequencing (NGS) and microbiological culture techniques. It shows the high counts of *Serratia* in all samples from the outbreak group, confirming the results previously observed in the NGS analysis.

### WGS of *Serratia* clones

Four *Serratia* isolates representing the two dominant PFGE clones associated with sepsis in the outbreak group were submitted to WGS (Fig 6). The genome analysis confirmed two



**Fig 4.** (A) *Phyla* percentage in each sample and infant, and summary of both groups expressed as the medians values. (B). *Genera* percentage in each sample and infant, and summary of both groups expressed as the median values. The 16 most abundant *genera* are highlighted in the figure, although up to 215 *genera* were detected in the samples analyzed in this study.

<https://doi.org/10.1371/journal.pone.0216581.g004>

phylogenetic lineages unrelated to others previously published. It is important to note that the isolate causing the sepsis and death of infant O11 was identical to the isolate obtained from infant O7, who had a satisfactory clinical evolution. WGS allowed the characterization of the resistome (*aac(6')-Ic-1*, *bla<sub>SRT-2-1</sub>*, *bla<sub>SRT-1-1</sub>*, and *qnrE*) and the virulome (*cheB*, *cheR*, *cheW*, *cheY*, *flgB*, *flgC*, *flgG*, *flgI*, *flhA*, *flhC*, *flhD*, *fliA*, *fliG*, *fliI*, *fliM*, *fliN*, *fliP*, *fliQ*, and *fliZ*), which

**Table 2. *Serratia* abundance detected by molecular tools and distribution of the two major clones detected in the outbreak group.** The underlined isolates were submitted to whole genome sequencing. High abundance of *Serratia* by NGS is marked in light grey color, whereas the dark grey means a clear dominance of the *Serratia* genera.

	INFANT	Meconium		Day 7		Day 14		Day 21	
		Cultivable	NGS (%)	Cultivable	NGS (%)	Cultivable	NGS (%)	Cultivable	NGS (%)
OUTBREAK GROUP	O1	<u>Clone A</u>	5.4		0.002		12.1		1.0
	O2		12.4		0.003		0.2		1.7
	O3		2.5	Clone B	17.3		19.0		
	O4		2.4	Clone B	2.4	Clone B	4.8		16.9
	O5		1.8		0.003		0.04	Clone A	28.8
	O6		11.8	Clone B	5.1	Clone B	26.5	Clone B	37.4
	O7		2.9	<u>Clone B</u>	0.2		35.5		4.9
	O8		1.9		0.7		1.3		0.02
	O9		2.0		0.004		0.004		0.006
	O10	<u>Clone A</u>	59.5	Clone A	78.5	Clone A	0.8	Clone A	0.4
	O11		3.1	<u>Clone B</u>	94.5				
	O12		0.3	Clone A	1.3		0.1		1.2
CONTROL GROUP	C1		0.004		0.005				0.008
	C2		0.006		0.001		0.001		0.006
	C3		0.1		0.0007		0		0
	C4				0.06		0.2		0.01
	C5		0.8				3.1		
	C6		0.06		0		0		
	C7				0.0008		0		0.002

<https://doi.org/10.1371/journal.pone.0216581.t002>

was uniform in all four isolates. Nevertheless, our isolates remained susceptible to most antibiotics, given the resistant genes detected were located on the chromosome and not transferable.

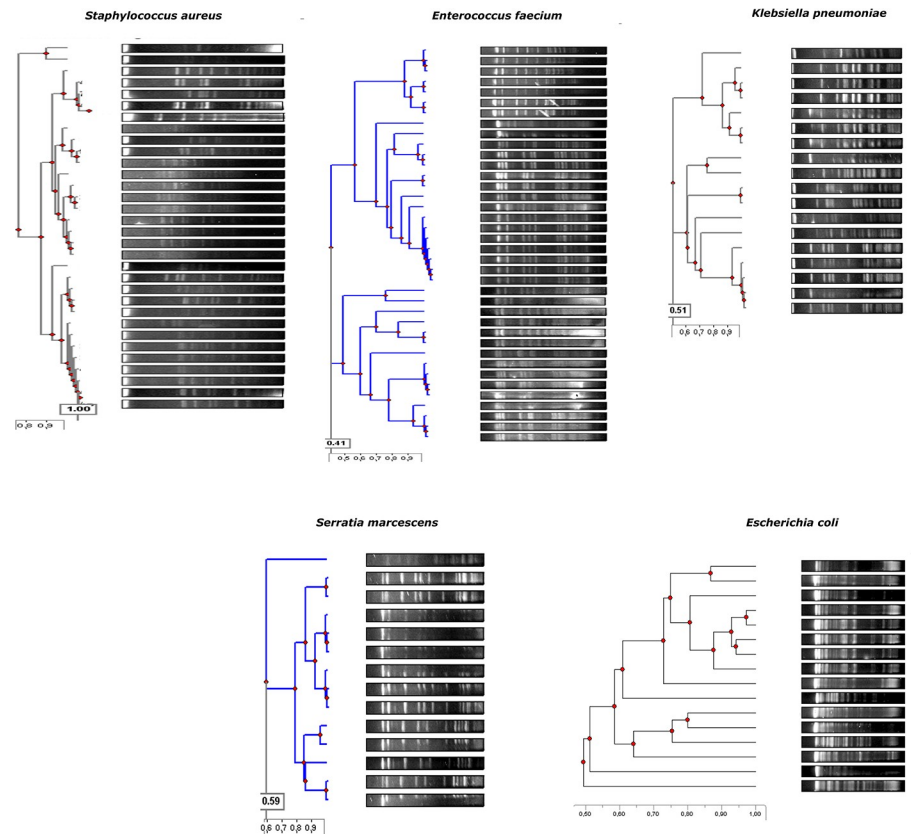
### Fungi detection

PCR-DGGE yielded positive amplifications in infant 6 (days 21 and 28) and infant 8 (days 14 and 21), and the nucleotide sequences of these amplicons corresponded to *Candida albicans*.

### Discussion

In the present study, we have described the succession of the gut microbiota composition from meconium to the first 3 weeks of life of low-weight preterm infants, comparing two epidemiological scenarios from the same NICU. The primary result of our work is related to the *S. marcescens* outbreak influence on infant gut microbiota patterns, including meconium, which might have been modified even before birth in relation to the hospital admission of mothers to prevent premature birth. Although there is marked inter-individual variability, the gut microbiota of low-weight premature infants is dominated by *Proteobacteria phylum*, particularly *E. coli*. Moreover, the expected bacterial ecosystem expansion after birth appears to be delayed, probably in relation to antibiotic exposition. In that sense it is important to remark the data lack about antibiotic prescription on the control group, that represent an important limitation of our work.

Regarding to the outbreak group, all but two infants received empiric antibiotics immediately after birth, and most required culture-guided antibiotic treatments during their admittance, including up to three different antimicrobial families. However, the high incidence of infectious complications among preterm neonates is the main argument used to justify the



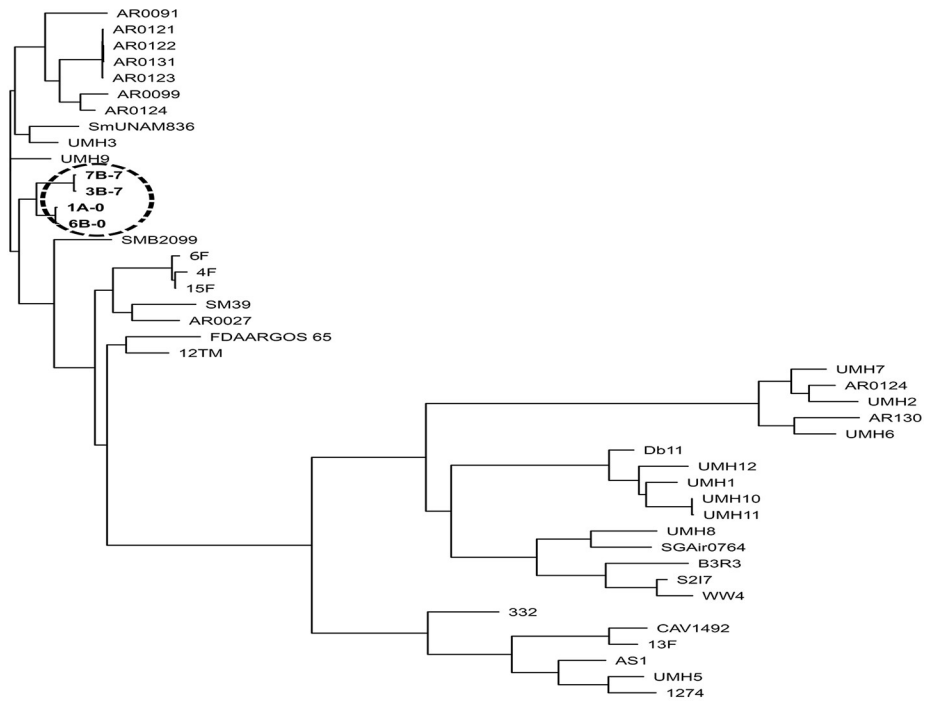
**Fig 5. Dendrograms showing the genetic relationship among the cultivable isolates based on the Dice's coefficient.**

<https://doi.org/10.1371/journal.pone.0216581.g005>

wide empiric and culture-guided use of antimicrobials in preterm infants, particularly in those with low birth weight, as in our case. Early biomarkers of preterm sepsis, together with the development of microbiome-based approaches, are urgently required to reduce antibiotic use in NICUs [28–29].

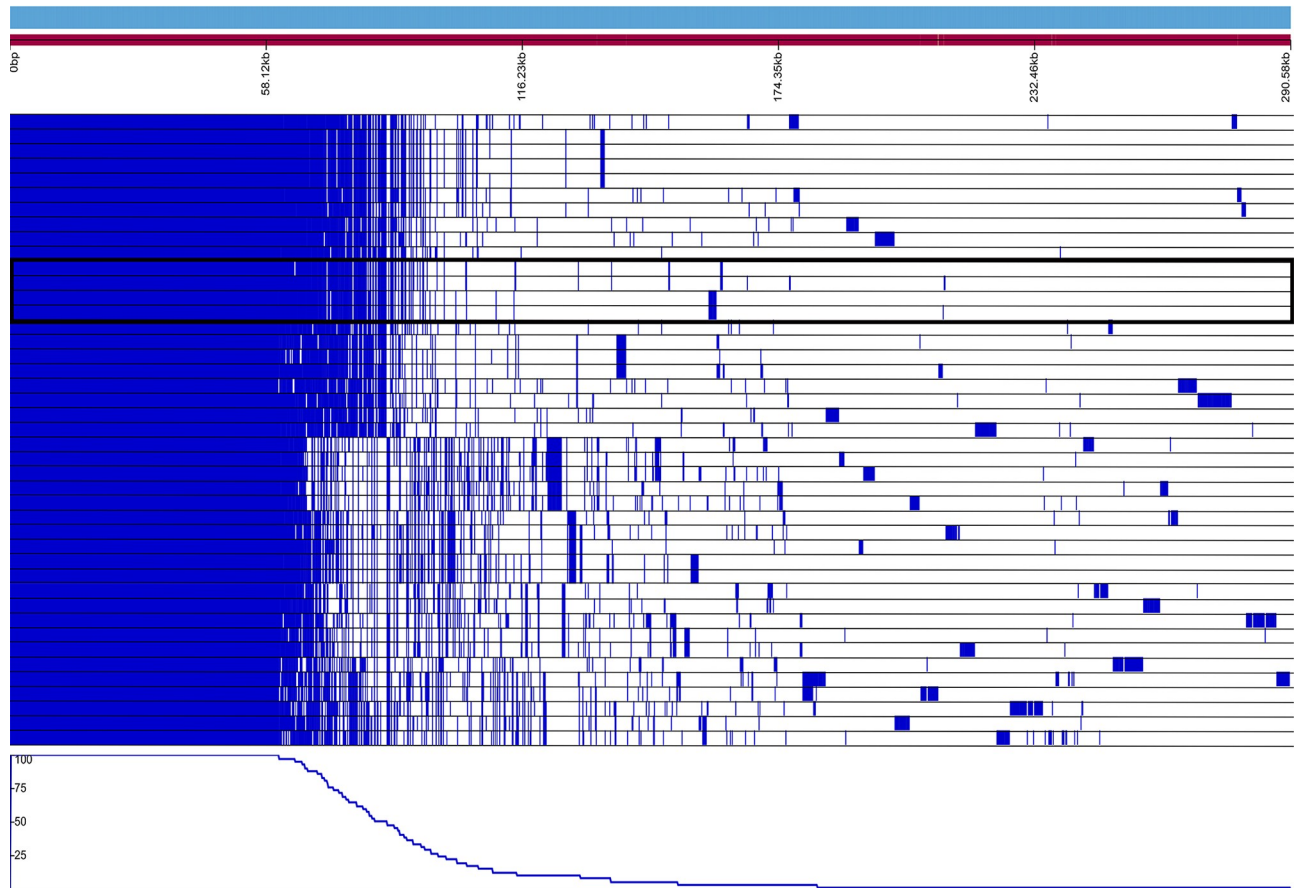
Prematurity is the main cause of neonatal morbidity and mortality, and the establishment of an adequate gut microbiota appears to be one of the most promising strategies to improve preterm infants' health and to reduce the impact of sequelae later in life [30–31]. Preterm infants admitted to an NICU have a high risk of infection, and *S. marcescens* is one of the most relevant nosocomial pathogens;<sup>23</sup> their intestinal carriage has been identified as a potential reservoir [32].

Our results demonstrated considerable *Proteobacteria* enrichment in both preterm infant groups, although the enrichment was significantly higher in the outbreak group from the second week of life, in concordance with other authors [33–36]. Previous studies found that *Proteobacteria* (mostly *Enterobacter* and *Photobacterium*), *Firmicutes* (mostly *Enterococcus* and *Lactobacillus*), and *Actinobacteria* (*Bifidobacterium*) dominated the microbial composition of meconium [37]. *E. coli*, *Staphylococcus sp.*, *Klebsiella sp.*, and a high rate of facultative anaerobes also commonly appear in the meconium of preterm neonates [38–39]. Our study demonstrated significant differences between the control group dominated by *E. coli* and *Enterococcus* and the outbreak group with higher densities of *E. coli* and *Serratia*. Recent data have shown that current 16S rDNA technology is not applicable for the gut ecosystem of premature infants [40]; this is an important limitation of our work, probably showing lower



**A**

**B**



**Fig 6.** Phylogenic representation of the 4 *S.marcescens* genomes sequenced in this study and comparison with other 32 public *S. marcescens* genomes (A). The second part of the figure (B) represents the common core of all 36 genomes versus the isolate-specific genes. Our 4 isolates grouped together in both analysis and are marked.

<https://doi.org/10.1371/journal.pone.0216581.g006>

detection of the *Bifidobacterium* population. A recent work also demonstrated that *Bifidobacterium* density is related to the gestational age [12].

Curiously, the meconium samples from both groups were characterized by high interindividual variability and similar alpha diversity as the subsequent fecal samples, pointing to a marked delay in the establishment of the ecosystem. Similar results have been reported in very-low-birth-weight preterm infants [17]. Other authors have described an increase of fecal microbiota complexity during the NICU stay [41].

Although intrauterine fetal gut colonization is still a controversial issue [6–8], the detection of the *Serratia* microorganism in meconium samples suggests the possibility that its preterm colonization could start before birth. Most of the mothers in our study (7 of 12) were admitted to prevent a premature birth by various causes for a median period of 9 days before birth, with a range from 1 to 49 days. In contrast, the mother whose infant died from *S. marcescens* sepsis was admitted only on the day before delivery. Whereas this microorganism was scarcely represented in the meconium, it was very dominant at day 7, a fact that probably preceded the blood translocation and the sepsis episode that occurred at day 10. Therefore, a systematic routine exploration for potential enrichment of specific gut bacterial populations could possibly contribute to prevention of bacteremia in susceptible groups of patients, such as preterm infants. Recently, a novel functional methodology using volatile organic compounds as biomarkers for early detection of gut bacterial enrichment was reported [42]. Our results also demonstrated that some epidemic microorganisms, such as *S. marcescens*, are able to colonize and eventually infect preterm neonates even when state-of-the-art preventive measures have been applied. The PFGE analysis grouped isolates colonizing the 12 infants in the outbreak group into two major clones, whereas the WGS revealed a close relationship between them, suggesting the existence of a common ancestor. These molecular techniques also revealed that the virulome of the strain causing bacteremia and death was identical to other strains with clinical successful evolution, reinforcing the hypothesis that the unclear barrier delimiting colonization from infection is influenced by numerous factors.

Some OTUs assigned to the mandatory predator bacteria *Bdellovibrio*, *Vampirovibrio*, and *Peredibacter* were detected in the three meconium samples analyzed in this study. Such bacteria need to predate other bacteria to grow and reproduce and are considered to be important ecological balancers of the microbial communities [43]. Few studies have focused on these bacteria in human ecosystems; however, their presence in meconium samples suggests that they might not be infrequent in the gut microbiota. A predator's inoculation could represent an ecological tool to modulate bacterial communities, taking into account predator-prey specificity [44].

The preterm nutrition policy of the hospital specifies neonates be fed maternal milk, although this is typically combined with human milk from donors and with preterm-adapted formulas. All the participating preterm infants received all three types of milk during the study, and although such data are not detailed, we are aware that this factor also influences the gut microbiota establishment. Maternal milk can reshape the infant gut microbiota [9], contributing its own site-specific microbiota [45–47], but also promoting the increase of a precise population by its prebiotic action [48]. Thus, it would be suitable to include in the microbiota profiling scheme the differentiation between living and dead bacteria in order to identify real colonizers from casual bacterial passengers associated with food intake [49].

Globally, our results indicate that, regardless of their perinatal settings, preterm neonates admitted to the same NICU are initially colonized by similar microbial communities that later evolve according to individual conditions. A *Serratia* outbreak influence on the establishment of the gut microbiota appears to be universal from the first days of admission; however, our results might also be applied to outbreaks caused by other microorganisms. This highlights the importance of the environment regarding the pattern of gut colonization of hospitalized pre-term infants.

## Supporting information

**S1 Dataset. Supplementary clinical data.**  
(XLSX)

## Acknowledgments

The author would like to thank Dr. Manuel Ponce-Alonso for his helpful with genomes analysis.

## Author Contributions

**Conceptualization:** Esperanza Escribano, Claudia Saralegui, Juan Miguel Rodríguez, Miguel Sáenz de Pipaón, Rosa del Campo.

**Data curation:** Laura Moles.

**Formal analysis:** Claudia Saralegui, María Teresa Montes, Claudio Alba, Teresa Alarcón, Fernando Lázaro-Perona.

**Funding acquisition:** Miguel Sáenz de Pipaón.

**Investigation:** Esperanza Escribano, Claudia Saralegui, Laura Moles, María Teresa Montes.

**Methodology:** Claudio Alba, Teresa Alarcón, Rosa del Campo.

**Project administration:** Juan Miguel Rodríguez.

**Supervision:** Miguel Sáenz de Pipaón, Rosa del Campo.

**Validation:** Teresa Alarcón, Fernando Lázaro-Perona.

**Writing – original draft:** Rosa del Campo.

**Writing – review & editing:** Juan Miguel Rodríguez, Miguel Sáenz de Pipaón.

## References

1. Sommer F, Bäckhed F. The gut microbiota: masters of host development and physiology. *Nat Rev Microbiol.* 2013, 11:227–38. <https://doi.org/10.1038/nrmicro2974> PMID: 23435359
2. Marchesi JR, Adams DH, Fava F, Hermes GD, Hirschfield GM, Hold G, et al. The gut microbiota and host health: a new clinical frontier. *Gut.* 2016, 65:330–9. <https://doi.org/10.1136/gutjnl-2015-309990> PMID: 26338727
3. Lynch SV, Pedersen O. The human intestinal microbiome in health and disease. *N Engl J Med.* 2016, 375:2369–79. <https://doi.org/10.1056/NEJMra1600266> PMID: 27974040
4. Walker WA. The importance of appropriate initial bacterial colonization of the intestine in newborn, child, and adult health. *Pediatr Res.* 2017, 82:387–95. <https://doi.org/10.1038/pr.2017.111> PMID: 28426649
5. Rautava S, Luoto R, Salminen S, Isolauri E. Microbial contact during pregnancy, intestinal colonization and human disease. *Nat Rev Gastroenterol Hepatol.* 2012, 9:565–76. <https://doi.org/10.1038/nrgastro.2012.144> PMID: 22890113

6. Perez-Muñoz ME, Arrieta MC, Ramer-Tait AE, Walter J. A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: implications for research on the pioneer infant microbiome. *Microbiome*. 2017, 5:48. <https://doi.org/10.1186/s40168-017-0268-4> PMID: 28454555
7. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. The placenta harbors a unique microbiome. *Sci Transl Med*. 2014, 6(237). <https://doi.org/10.1126/scitranslmed.3008599> PMID: 24848255
8. Milani C, Duranti S, Bottacini F, Casey E, Turrone F, Mahony J, et al. The first microbial colonizers of the human gut: composition, activities, and health implications of the infant gut microbiota. *Microbiol Mol Biol Rev*. 2017, 81(4). <https://doi.org/10.1128/MMBR.00036-17> PMID: 29118049
9. Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, et al. Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe*. 2015, 17:690–703. <https://doi.org/10.1016/j.chom.2015.04.004> PMID: 25974306
10. Madan JC, Farzan SF, Hibberd PL, Karagas MR. Normal neonatal microbiome variation in relation to environmental factors, infection and allergy. *Curr Opin Pediatr*. 2012, 24:753–9. <https://doi.org/10.1097/MOP.0b013e32835a1ac8> PMID: 23111681
11. La Rosa PS, Warner BB, Zhou Y, Weinstock GM, Sodergren E, Hall-Moore CM, et al. Patterned progression of bacterial populations in the premature infant gut. *Proc Natl Acad Sci U S A*. 2014, 111:12522–7. <https://doi.org/10.1073/pnas.1409497111> PMID: 25114261
12. Chernikova DA, Madan JC, Housman ML, Zain-Ul-Abideen M, Lundgren SN, Morrison HG, et al. The premature infant gut microbiome during the first 6 weeks of life differs based on gestational maturity at birth. *Pediatr Res*. 2018, 84:71–9. <https://doi.org/10.1038/s41390-018-0022-z> PMID: 29795209
13. Blencowe H, Cousens S, Chou D, et al. Born too soon: the global epidemiology of 15 million preterm births. *Reprod Health*. 2013, 10:Suppl1:S2.
14. Blencowe H, Cousens S, Oestergaard MZ, Oestergaard M, Say L, Moller AB, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet*. 2012, 379:2162–72. [https://doi.org/10.1016/S0140-6736\(12\)60820-4](https://doi.org/10.1016/S0140-6736(12)60820-4) PMID: 22682464
15. Zeitlin J, Szamatulska K, Drewniak N, Mohangoo AD, Chalmers J, Sakkeus L, et al. Preterm birth time trends in Europe: a study of 19 countries. *BJOG*. 2013, 120:1356–65. <https://doi.org/10.1111/1471-0528.12281> PMID: 23700966
16. DiGiulio DB, Callahan BJ, McMurdie PJ, Costello EK, Lyell DJ, Robaczewska A, et al. Temporal and spatial variation of the human microbiota during pregnancy. *Proc Natl Acad Sci U S A*. 2015, 112:11060–5. <https://doi.org/10.1073/pnas.1502875112> PMID: 26283357
17. Patel AL, Mutlu EA, Sun Y, Koenig L, Green S, Jakubowicz A, et al. Longitudinal survey of microbiota in hospitalized preterm very low birth weight infants. *J Pediatr Gastroenterol Nutr*. 2016, 62:292–303. <https://doi.org/10.1097/MPG.0000000000000913> PMID: 26230901
18. Unger S, Stintzi A, Shah P, Mack D, O'Connor DL. Gut microbiota of the very-low-birth-weight infant. *Pediatr Res*. 2015, 77:205–13. <https://doi.org/10.1038/pr.2014.162> PMID: 25310760
19. Moles L, Gómez M, Jiménez E, Fernández L, Bustos G, Chaves F, et al. Preterm infant gut colonization in the neonatal ICU and complete restoration 2 years later. *Clin Microbiol Infect*. 2015, 21:936.e1–10. <https://doi.org/10.1016/j.cmi.2015.06.003> PMID: 26086569
20. Elgin TG, Kern SL, McElroy SJ. Development of the neonatal intestinal microbiome and its association with necrotizing enterocolitis. *Clin Ther*. 2016, 38:706–15. <https://doi.org/10.1016/j.clinthera.2016.01.005> PMID: 26852144
21. Groer MW, Gregory KE, Louis-Jacques A, Thibeau S, Walker WA. The very low birth weight infant microbiome and childhood health. *Birth Defects Res C Embryo Today*. 2015, 105:252–64. <https://doi.org/10.1002/bdrc.21115> PMID: 26663857
22. Yusef D, Shalakhti T, Awad S, Algharaibeh H, Khasawneh W. Clinical characteristics and epidemiology of sepsis in the neonatal intensive care unit in the era of multi-drug resistant organisms: A retrospective review. *Pediatr Neonatol*. 2017, 59:35–41. <https://doi.org/10.1016/j.pedneo.2017.06.001> PMID: 28642139
23. Martineau C, Li X, Lalancette C, Perreault T, Fournier E, Tremblay J, et al. *Serratia marcescens* outbreak in a neonatal intensive care unit (NICU): new insights from next-generation sequencing applications. *J Clin Microbiol*. 2018, 27; 56(9). pii: e00235–18. <https://doi.org/10.1128/JCM.00235-18> PMID: 29899005
24. Johnson J, Quach C. Outbreaks in the neonatal ICU: a review of the literature. *Curr Opin Infect Dis*. 2017, 30:395–403. <https://doi.org/10.1097/QCO.0000000000000383> PMID: 28582313
25. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res*. 2013, 41:1–11. <https://doi.org/10.1093/nar/gks1039>

26. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011, 12:R60. <https://doi.org/10.1186/gb-2011-12-6-r60> PMID: 21702898
27. Liu J, Yu Y, Cai Z, Bartlam M, Wang Y. Comparison of ITS and 18S rDNA for estimating fungal diversity using PCR-DGGE. *World J Microbiol Biotechnol.* 2015, 31:1387–95. <https://doi.org/10.1007/s11274-015-1890-6> PMID: 26081603
28. Groer MW, Luciano AA, Dishaw LJ, Ashmeade TL, Miller E, Gilbert JA. Development of the preterm infant gut microbiome: a research priority. *Microbiome.* 2014, 2:38. <https://doi.org/10.1186/2049-2618-2-38> PMID: 25332768
29. Sharma D, Farahbakhsh N, Shastri S, Sharma P. Biomarkers for diagnosis of neonatal sepsis: a literature review. *J Matern Fetal Neonatal Med.* 2018, 31:1646–59. <https://doi.org/10.1080/14767058.2017.1322060> PMID: 28427289
30. Ruiz L, Moles L, Gueimonde M, Rodriguez JM. Perinatal microbiomes' influence on preterm birth and preterms' health: influencing factors and modulation strategies. *J Pediatr Gastroenterol Nutr.* 2016, 63:e193–e20. <https://doi.org/10.1097/MPG.0000000000001196> PMID: 27019409
31. Stinson LF, Payne MS, Keelan JA. Planting the seed: origins, composition, and postnatal health significance of the fetal gastrointestinal microbiota. *Crit Rev Microbiol.* 2017, 43:352–69. <https://doi.org/10.1080/1040841X.2016.1211088> PMID: 27931152
32. Montagnani C, Cocchi P, Lega L, Campana S, Biermann KP, Braggion C, et al. *Serratia marcescens* outbreak in a neonatal intensive care unit: crucial role of implementing hand hygiene among external consultants. *BMC Infect Dis.* 2015, 15:11. <https://doi.org/10.1186/s12879-014-0734-6> PMID: 25582674
33. Cong X, Xu W, Janton S, Henderson WA, Matson A, McGrath JM, et al. Gut microbiome developmental patterns in early life of preterm infants: impacts of feeding and gender. *PLoS One.* 2016, 11:e0152751. <https://doi.org/10.1371/journal.pone.0152751> PMID: 27111847
34. Hill CJ, Lynch DB, Murphy K, Ulaszewska M, Jeffery IB, O'Shea CA, et al. Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort. *Microbiome.* 2017, 5:4. <https://doi.org/10.1186/s40168-016-0213-y> PMID: 28095889
35. Underwood MA, Sohn K. The microbiota of the extremely preterm infant. *Clin Perinatol.* 2017, 44:407–27. <https://doi.org/10.1016/j.clp.2017.01.005> PMID: 28477669
36. Nogacka A, Salazar N, Suárez M, Milani C, Arboleya S, Solís G, et al. Impact of intrapartum antimicrobial prophylaxis upon the intestinal microbiota and the prevalence of antibiotic resistance genes in vaginally delivered full-term neonates. *Microbiome.* 2017, 5:93. <https://doi.org/10.1186/s40168-017-0313-3> PMID: 28789705
37. Hansen R, Scott KP, Khan S, Martin JC, Berry SH, Stevenson M, et al. First-pass meconium samples from healthy term vaginally-delivered neonates: an analysis of the microbiota. *PLoS One.* 2015, 10:e0133320. <https://doi.org/10.1371/journal.pone.0133320> PMID: 26218283
38. Ardisson AN, de la Cruz DM, Davis-Richardson AG, Rechcigl KT, Li N, Drew JC, et al. Meconium microbiome analysis identifies bacteria correlated with premature birth. *PLoS One.* 2014, 9:e90784. <https://doi.org/10.1371/journal.pone.0090784> PMID: 24614698
39. Moles L, Gómez M, Heilig H, Bustos G, Fuentes S, de Vos W, et al. Bacterial diversity in meconium of preterm neonates and evolution of their fecal microbiota during the first month of life. *PLoS One.* 2013, 8:e66986. <https://doi.org/10.1371/journal.pone.0066986> PMID: 23840569
40. Alcon-Giner C, Caim S, Mitra S, Ketskemety J, Wegmann U, Wain J, et al. Optimisation of 16S rRNA gut microbiota profiling of extremely low birth weight infants. *BMC Genomics.* 2017, 18:841. <https://doi.org/10.1186/s12864-017-4229-x> PMID: 29096601
41. Jacquot A, Neveu D, Aujoulat F, Mercier G, Marchandin H, Jumas-Bilak E, et al. Dynamics and clinical evolution of bacterial gut microflora in extremely premature patients. *J Pediatr.* 2011, 158:390–6. <https://doi.org/10.1016/j.jpeds.2010.09.007> PMID: 20961563
42. Berkhout DJC, van Keulen BJ, Niemarkt HJ, Bessem JR, de Boode WP, Cossey V, et al. Late-onset sepsis in preterm infants can be detected preclinically by fecal volatile organic compound analysis: a prospective, multicenter cohort study. *Clin Infect Dis.* 2018, Jun 21, *In press.* <https://doi.org/10.1093/cid/ciy383> PMID: 29931245
43. Iebba V, Santangelo F, Totino V, Nicoletti M, Gagliardi A, De Biase RV, et al. Higher prevalence and abundance of *Bdellovibrio bacteriovorus* in the human gut of healthy subjects. *PLoS One.* 2013, 8:e61608. <https://doi.org/10.1371/journal.pone.0061608> PMID: 23613881
44. Kadouri DE, To K, Shanks RMQ, Doi Y. Predatory bacteria: a potential ally against multidrug-resistant gram-negative pathogens. *PLoS One.* 2013, 8:e63397. <https://doi.org/10.1371/journal.pone.0063397> PMID: 23650563

45. Fernández L, Langa S, Martín V, Maldonado A, Jiménez E, Martín R, et al. The human milk microbiota: origin and potential roles in health and disease. *Pharmacol Res.* 2013, 69:1–10. <https://doi.org/10.1016/j.phrs.2012.09.001> PMID: 22974824
46. Pannaraj PS, Li F, Cerini C, Bender JM, Yang S, Rollie A, et al. Association between breast milk bacterial communities and establishment and development of the infant gut microbiome. *JAMA Pediatr.* 2017, 171:647–54. <https://doi.org/10.1001/jamapediatrics.2017.0378> PMID: 28492938
47. Macpherson AJ, de Agüero MG, Ganai-Vonarburg SC. How nutrition and the maternal microbiota shape the neonatal immune system. *Nat Rev Immunol.* 2017, 17:508–17. <https://doi.org/10.1038/nri.2017.58> PMID: 28604736
48. Moukarzel S, Bode L. Human milk oligosaccharides and the preterm infant: a journey in sickness and in health. *Clin Perinatol.* 2017, 44:193–207. <https://doi.org/10.1016/j.clp.2016.11.014> PMID: 28159206
49. Young GR, Smith DL, Embleton ND, Berrington JE, Schwalbe EC, Cummings SP, et al. Reducing viability bias in analysis of gut microbiota in preterm infants at risk of NEC and sepsis. *Front Cell Infect Microbiol.* 2017, 7:237. <https://doi.org/10.3389/fcimb.2017.00237> PMID: 28634574



# Genomics of *Serratia marcescens* Isolates Causing Outbreaks in the Same Pediatric Unit 47 Years Apart: Position in an Updated Phylogeny of the Species

Claudia Saralegui<sup>1,2†</sup>, Manuel Ponce-Alonso<sup>1,2†</sup>, Blanca Pérez-Viso<sup>1</sup>, Laura Moles Alegre<sup>3</sup>, Esperanza Escribano<sup>4</sup>, Fernando Lázaro-Perona<sup>5</sup>, Val F. Lanza<sup>6,7</sup>, Miguel Sáenz de Pipaón<sup>4</sup>, Juan Miguel Rodríguez<sup>8</sup>, Fernando Baquero<sup>1,7</sup> and Rosa del Campo<sup>1,2\*</sup>

## OPEN ACCESS

### Edited by:

Iain Sutcliffe,  
Northumbria University,  
United Kingdom

### Reviewed by:

Nora Adriana Altier,  
Instituto Nacional de Investigación  
Agropecuaria (INIA), Uruguay  
Daria Van Tyne,  
University of Pittsburgh, United States

### \*Correspondence:

Rosa del Campo  
rosacampo@yahoo.com

†These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Evolutionary and Genomic  
Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 08 October 2019

**Accepted:** 02 March 2020

**Published:** 31 March 2020

### Citation:

Saralegui C, Ponce-Alonso M, Pérez-Viso B, Moles Alegre L, Escribano E, Lázaro-Perona F, Lanza VF, de Pipaón MS, Rodríguez JM, Baquero F and del Campo R (2020) Genomics of *Serratia marcescens* Isolates Causing Outbreaks in the Same Pediatric Unit 47 Years Apart: Position in an Updated Phylogeny of the Species. *Front. Microbiol.* 11:451. doi: 10.3389/fmicb.2020.00451

<sup>1</sup> Servicio de Microbiología, Hospital Universitario Ramón y Cajal and Instituto Ramón y Cajal de Investigación Sanitaria, Madrid, Spain, <sup>2</sup> Red Española de Investigación en Patología Infecciosa, Madrid, Spain, <sup>3</sup> Unidad de Esclerosis Múltiple, Instituto de Investigación Sanitaria Bionostia, San Sebastián, Spain, <sup>4</sup> Servicio de Neonatología, Hospital Universitario La Paz, and Universidad Autónoma de Madrid, Madrid, Spain, <sup>5</sup> Servicio de Microbiología, Hospital Universitario La Paz, Madrid, Spain, <sup>6</sup> Unidad de Bioinformática del IRYCIS, Madrid, Spain, <sup>7</sup> Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública, Madrid, Spain, <sup>8</sup> Departamento de Nutrición y Ciencia de los Alimentos, Universidad Complutense de Madrid, Madrid, Spain

The first documented nosocomial outbreak caused by *Serratia marcescens* in Spain occurred in 1969 at the neonatal intensive care unit (NICU) of the tertiary La Paz Children's Hospital in Madrid, Spain, and based on the available phenotyping techniques at this time, it was considered as a monoclonal outbreak. Only 47 years later, another *S. marcescens* outbreak of an equivalent dimension occurred at the same NICU. The aim of the present study was to study isolates from these historical and contemporary outbreaks by phenotypic analysis and whole-genome sequencing techniques and to position these strains along with 444 publicly available *S. marcescens* genomes, separately comparing core genome and accessory genome contents. Clades inferred by both approaches showed high correlation, indicating that core and accessory genomes seem to evolve in the same manner for *S. marcescens*. Nine *S. marcescens* clusters were identified, and isolates were grouped in two of them according to sampling year. One exception was isolate 13F-69, the most genetically distant strain, located in a different cluster. Categorical functions in the annotated accessory genes of both collections were preserved among all isolates. No significant differences in frequency of insertion sequences in historical (0.18–0.20)—excluding the outlier strain—versus contemporary isolates (0.11–0.19) were found despite the expected resting effect. The most dissimilar isolate, 13F-69, contains a highly preserved plasmid previously described in *Bordetella bronchiseptica*. This strain exhibited a few antibiotic resistance genes not resulting in a resistant phenotype, suggesting the value of gene down expression in adaptation to long-term starvation.

**Keywords:** *Serratia marcescens*, phylogeny, resistome, antibiotic susceptibility, nosocomial outbreak

## INTRODUCTION

*Serratia marcescens* is a ubiquitous environmental microorganism, but also a relevant nosocomial pathogen able to cause a broad spectrum of infections, particularly in neonates (Zingg et al., 2008). The first scientific reports involving *S. marcescens* in human infections occurred in the second half of the 20th century. Today, the main concerns regarding this opportunistic pathogen are its capability to spread in the hospital environment, to cause outbreaks, and its potential for expressing and disseminating antibiotic resistance, combining intrinsic mechanisms and acquired antimicrobial genes (Mahlen, 2011).

The first documented *S. marcescens* outbreak affecting pediatric patients in Spain occurred in 1969 at the neonatal intensive care unit (NICU) of La Paz Children's Hospital, the main pediatric university hospital in Madrid. The outbreak lasted, with oscillations, until late 1974 (>5 years), with a high incidence of *S. marcescens* bacteremia (9.16% of all early neonatal sepsis during the period), frequently associated with gut colonization during the first 3 days of life (12.5% of 120 children tested) (Baquero et al., 1969). However, the microorganism had been found not to have colonized the mothers' vaginal cavity (37 tested) (Baquero et al., 1977). The recovered isolates were considered as a single clone causing a monoclonal outbreak according to the available bacterial typing techniques, based on phenotypic tests that included biochemical reactions and antimicrobial susceptibility testing.

Forty-seven years later (2016), in the same NICU and after several intermittent mild episodes, a new severe *S. marcescens* outbreak was declared with an extremely high incidence of bacteremia, also preceded by gut *S. marcescens* enrichment (Escribano et al., 2019). The environmental surveillance of the affected NICU did not obtain conclusive results, and bacterial typing with molecular tools demonstrated the coexistence of several genetic lineages involved in this outbreak (Redondo-Bravo et al., 2019). During the interval between the first and second outbreaks, *S. marcescens* was isolated from patients but without causing outbreaks of comparable magnitude. The aim of the present study was to examine the genomes of *S. marcescens* isolates causing historical and contemporary prolonged outbreaks that occurred in the same NICU but separated by almost half a century (from 1969 to 2016). A new, updated phylogeny of the species allowed to position old and contemporary outbreak isolates in a broader evolutionary perspective.

## MATERIALS AND METHODS

### Sample Collection and Processing

Outbreak-causing *S. marcescens* isolates were obtained from blood and/or fecal samples from preterm neonates admitted to the NICU of Hospital La Paz in Madrid, Spain, and were grouped into historical (21 isolates from 1969) and contemporary (five isolates from 2016) collections. The contemporary outbreak from which contemporary collection belonged to started in December 2016 and ended in March 2017. The historical isolates were

conserved at room temperature as stabs (short slant) in 4 ml brain–heart infusion (BHI) agar in screw-threaded small neutral glass vials (1 cm wide, 5 cm high), tightly sealed with parafilm at room temperature, and had not been opened before 2016. To refresh the cultures, the agar was rehydrated for 2 h with BHI broth (Difco, United States), gently shaken, and seeded in Columbia blood agar plates. Ten out of 21 historical strains were able to grow overnight. Species identification was confirmed by matrix-assisted laser desorption ionization-time of flight mass spectrometry (Bruker, Germany), and all recovered isolates were re-stored at  $-80^{\circ}\text{C}$  in sterilized half-skimmed milk.

### Pulse-Field Gel Electrophoresis Typing

Pulse-field gel electrophoresis (PFGE) was developed according to a previously described protocol (Villa et al., 2012) using the *SpeI* restriction enzyme. The band patterns were analyzed using BioNumerics software<sup>1</sup>, and the results representation was built based on Dice coefficients and the unweighted pair group method with arithmetic mean algorithm.

### Whole-Genome Sequencing, Assembly, and Annotation

Eight isolates (four from each period) were selected according to PFGE and submitted to whole-genome sequencing (WGS). DNA was obtained using the QIAamp kit (QIAAMP, Germany), determining their concentration and quality by Qubit fluorometer. DNA samples were sent to the Complutense University of Madrid for library preparation using the Nextera XT kit (Illumina, Inc., United States). Paired-end sequencing was performed using a MiSeq system (2 × 150 pb) (Illumina). Nullarbor bioinformatics pipeline was used for sequence analyses (v2; Seemann, T.)<sup>2</sup>. This tool includes Trimmomatic v0.38 software for trimming and quality filtering of reads (Bolger et al., 2014), Kraken software v1.0 (Wood and Salzberg, 2014) and Centrifuge software v1.0.4 (Kim et al., 2016) for species identification, SPAdes v3.12.0 software for *de novo* assembly (Bankevich et al., 2012), Prokka software v1.13 for annotation (Seemann, 2014), Abricate software v0.8.8 (Seemann, T.)<sup>3</sup> using the Resfinder database for resistome identification, and Snippy software v4.4.1 (Seemann, T.)<sup>4</sup> and Roary software v3.12.0 (Page et al., 2015) for core genome and pangenome calculation. The Phandango web-based tool (Hadfield et al., 2017) was used for pangenome visualization and plotting. PLACNETw free-access software (Vielva et al., 2017) was used for detection and graphical reconstruction of contigs harbored by putative plasmids (containing relaxases and/or replication initiator proteins), which were subsequently confirmed by the basic local alignment search tool. The Phaster web-based tool (Arndt et al., 2016) was used for prophage annotation. Insertion sequences (ISs) were annotated using ISEScan software (Xie and Tang, 2017).

<sup>1</sup><http://www.applied-maths.com/bionumerics/bionumerics.htm>

<sup>2</sup><https://github.com/tseemann/nullarbor>

<sup>3</sup><https://github.com/tseemann/abicate>

<sup>4</sup><https://github.com/tseemann/snippy>

## Phylogenetic Analysis of Historical and Contemporary Strains by Single-Nucleotide Polymorphism

A single-nucleotide polymorphism (SNP) calling “all against all” approach was applied to calculate how many SNPs were present between each pair of isolates. Subsequently, the SNP calling approach was used to evaluate the phylogenetic relationship among eight strains, both contemporary and historical ones. First, each genome was mapped against the reference genome (O1-16, accession number QYRU00000000) by Snippy-core. Next, SNP variant calling was performed (Snippy v4.0-dev2) on the mapped sequences, and the SNP’s phylogeny was inferred (FastTree v2.1.10). All steps were performed using Nullarbor bioinformatics pipeline.

## Phylogenetic Analysis of the Global *S. marcescens* Population by Core Genome Multilocus Sequence Typing

We built a core genome multilocus sequence typing (cgMLST) scheme for phylogenetic comparison among eight contemporary or historical strains and 444 *S. marcescens* genomes assembled as contigs, scaffolds, or complete genomes that were available in the RefSeq National Center for Biotechnology Information (NCBI) database<sup>5</sup> (Supplementary Data). We used the KRAKEN toolkit and all available filters of the NCBI web tool to only include genomes robustly assigned to *Serratia* and exclude partial and anomalous genomes. Likewise, we only considered contigs of >500 bp. First, a pangenome allele database was set using all selected genomes ( $n = 452$ ). Then, we built one MLST tree from the pangenome allele database, selecting allelic variations only present in the core genome (cgMLST), set at >95% of occurrence. All analyses were performed using the cano-wgMLST\_BacCompare web-based tool (Liu et al., 2019). Trees were edited using the iTOL v4.4.2 web-based tool (Letunic and Bork, 2019).

## Accessory Genome Functional Analysis

We compared the accessory genome of the contemporary or historical outbreak isolates and all *S. marcescens* annotated genomes from the previously mentioned RefSeq NCBI database using AcCNET v1.2 software. Ward’s minimum variance hierarchical clustering method was applied to cluster genomes according to their predicted annotated proteome (Lanza et al., 2017). Only significant proteins robustly assigned to each cluster (>70% per cluster frequency, <70% per total frequency, and <0.001 adjusted  $p$ -value) were taken into account for further comparisons. The Clusters of Orthologous Groups of proteins (COGs) database was used to assign functional categories to the predicted proteome content. RStudio software v.1.2.5001 and Gephi v.0.9.2 were used for statistical analysis and plotting.

## Antimicrobial Susceptibility Testing

The MicroScan WalkAway automated system (Beckman Coulter, United States) was used to determine the

minimal inhibitory concentration (MIC) values. EUCAST criteria for each antimicrobial were used to define the susceptible, intermediate, and resistant isolates<sup>6</sup>.

## RESULTS

### *Serratia marcescens* Pulse-Field Gel Electrophoresis Typing and Whole-Genome Sequencing

Ten out of 21 isolates from the historical collection were recovered after 24 h of overnight incubation and further analyzed. These historical strains were grouped in five PFGE patterns, and the five contemporary isolates were grouped into three pulsotypes (Figure 1). Four representative isolates from each period were finally submitted to WGS analysis ( $n = 8$  strains). Data regarding isolation date and source of selected strains are included in Table 1. The relevant data of WGS throughput are summarized in Table 2, including accession numbers. The assembled genomes were deposited in the GenBank database and linked to BioProject PRJNA510235. The 13F-69 strain presented the lowest GC% content and the highest number of contigs.

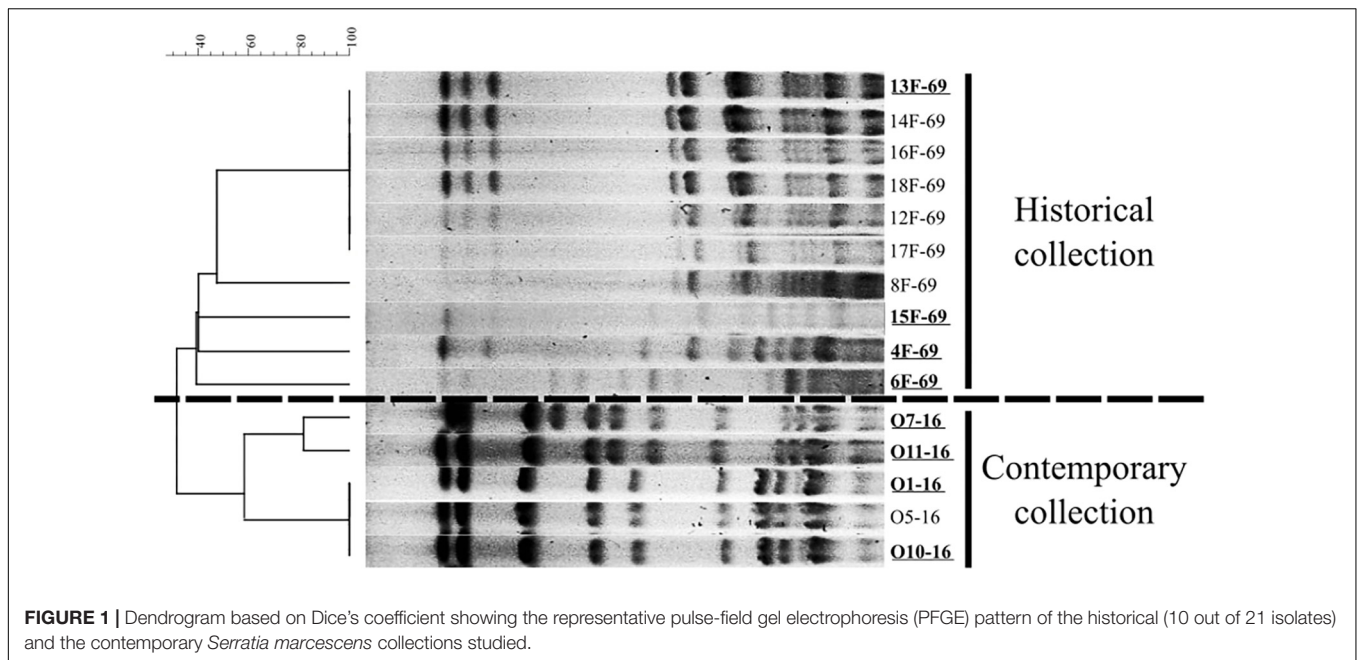
### Functional Annotation, Plasmids, and Prophage Identification

A total of 6,390 different clusters of orthologous genes constituted the pangenome of the eight studied outbreak isolates (four isolates from the historical and four isolates from the contemporary collection). The core genome (genes present in >99% of strains as established by the Nullarbor pipeline) constituted 3,642 (57%) gene clusters, whereas 1,237 (19.4%) gene clusters corresponded to the accessory genome (genes present in <15% of strains). All isolates presented a similar number of protein-coding sequences (CDSs) and tRNA copies. A summary of the gene distribution and additional data are presented in Figure 2, whereas IS profiles are displayed in Figure 3.

The historical 13F-69 and 15F-69 isolates were remarkably different from the others (highest CDS, tRNA, and IS content), and all data indicated that numerous genetic exchange events had occurred in these strains. The contemporary isolates appeared to present higher genetic stability with slightly lower total IS content. From the 14 different IS families detected, eight were detected in a single isolate, five were present in two or three isolates, and only IS3, IS21, and IS200/IS605 appeared to be regularly represented. It is important to note that IS200/605 was only detected in its complete form in the four contemporary isolates, although a partial IS200/605 sequence was also observed in the 13F-69 strain. Mann–Whitney U test for independent samples showed significant differences ( $p < 0.05$ ) between historical and contemporary total IS percentage. Nevertheless, when the

<sup>5</sup><https://www.ncbi.nlm.nih.gov/refseq/>

<sup>6</sup>[www.eucast.org](http://www.eucast.org)



outlier 13F-69 was extracted from the comparison, no significant differences resulted.

According to the PlacnetW report, incomplete plasmids were detected in the O10-16, 13F-69, and 15F-69 strains (Figure 4 and Table 3). The only complete plasmid, R906, corresponded to the one previously reported in *Bordetella bronchiseptica* and was found in the historical 13F-69 isolate. This plasmid carried genes encoding antibiotic resistance [*oxa-2*, *sul-1*, *qacEΔ1*, *aph(3')*-II, and multidrug transporters *mepC* and *mdtE*], nickel-cobalt resistance (*cnrA*), quaternary ammonium resistance (*sugE*), type IV secretion system proteins, toxins (*fitB*), and iron transporters. Other fragments identified as putative plasmids shared sequences with extra-chromosomal elements of *Serratia entomophila* and *Serratia* spp. or with the chromosome of *S. marcescens*, *Klebsiella pneumoniae*, *Serratia* sp., or *Dickeya* sp. (Figure 4). Regarding lysogenic phages present in the studied isolates, a total of five complete prophages were identified, all belonging to the *Caudovirales* order (Table 3).

**TABLE 1 |** Isolation date and source of the eight *Serratia marcescens* studied isolates.

Collection	Isolates	Isolation Date	Isolation Source
Historical	4F-69	4/11/1969	Gut
	6F-69	6/11/1969	Gut
	13F-69	13/11/1969	Gut
	15F-69	15/11/1969	Gut
	Contemporary	O1-16	21/11/2016
	O7-16	28/11/2016	Gut
	O10-16	8/12/2016	Gut
	O11-16	16/11/2016	Gut

## Antibiotic-Resistant Genes and Susceptibility

Both in the historical and contemporary outbreaks, the majority of genomes contained three or four antibiotic resistance genes, except the historical 13F-69 and 15F-69 isolates, which carried 10 and six genes, respectively (Table 4). Interestingly, 13F-69 remained phenotypically susceptible to almost all tested antibiotics, excluding those affected by *S. marcescens* constitutive resistance (Table 5). Only two isolates, one from each period, exhibited resistance to broad-spectrum beta-lactam antibiotics, presumably due to mechanisms involving *ampC* de-repression.

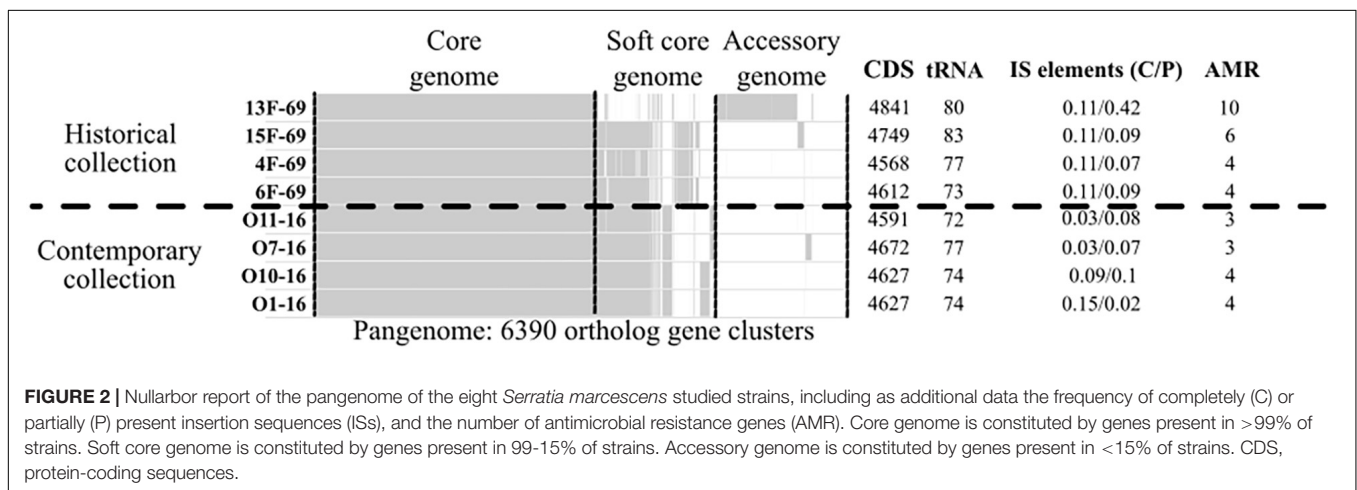
## Phylogenetic Analysis Within Contemporary and Historical Outbreak Strain Collections by the Single-Nucleotide Polymorphism Approach

The “all against all” SNP matrix showed the isolates had a certain degree of genetic polymorphism within each outbreak. In the contemporary outbreak, there were two clones, one represented by O1-16 and O10-16 and another represented by O7-16 and O11-16, whereas more SNPs were identified within isolates from the historical collection (Table 6). Then, one isolate from the contemporary collection (O1-16) was taken as the reference genome for constructing a phylogenetic tree based on SNP core distance. A total of 173,766 SNPs were found in the mapped regions of the eight studied isolates with O1-16 as a reference. Phylogenetic distances from reference strain O1-16 were 0–4,962 in the four strains of contemporary collection and 8,738–8,862 in three out of four in the historical collection. The outlier strain 13F-69 had a value of 32,536 and constituted a separate distant branch (Figure 5).

**TABLE 2** | Sequencing information of the eight *Serratia marcescens* bacterial genomes compared in the study from Nullarbor pipeline report.

Collection	Isolates	Accession Number	Contigs	Reads	Size (bp)	GC (%)	Seq. Depth	N50
Historical	4F-69	QZVX00000000	20	2,374,810	5,035,544	59.6	63	2,820,768
	6F-69	RAIC00000000	23	1,394,190	5,095,011	59.6	37	518,765
	13F-69	SNQH00000000	65	2,457,194	5,279,779	59.1	66	163,753
	15F-69	RAID00000000	22	1,840,822	5,238,206	59.5	49	2,937,699
Contemporary	O1-16	QYRU00000000	26	3,232,838	5,102,926	59.7	87	600,674
	O7-16	QYRV00000000	21	2,215,906	5,127,038	59.7	59	3,175,227
	O10-16	QYSA00000000	22	1,950,698	5,105,131	59.7	52	600,674
	O11-16	QYSB00000000	23	1,925,266	5,057,421	59.8	51	580,213

N50, minimum contig length needed to cover 50% of the genome; GC, percentage of guanine-cytosine content.



**FIGURE 2** | Nullarbor report of the pangenome of the eight *Serratia marcescens* studied strains, including as additional data the frequency of completely (C) or partially (P) present insertion sequences (ISs), and the number of antimicrobial resistance genes (AMR). Core genome is constituted by genes present in >99% of strains. Soft core genome is constituted by genes present in 99-15% of strains. Accessory genome is constituted by genes present in <15% of strains. CDS, protein-coding sequences.

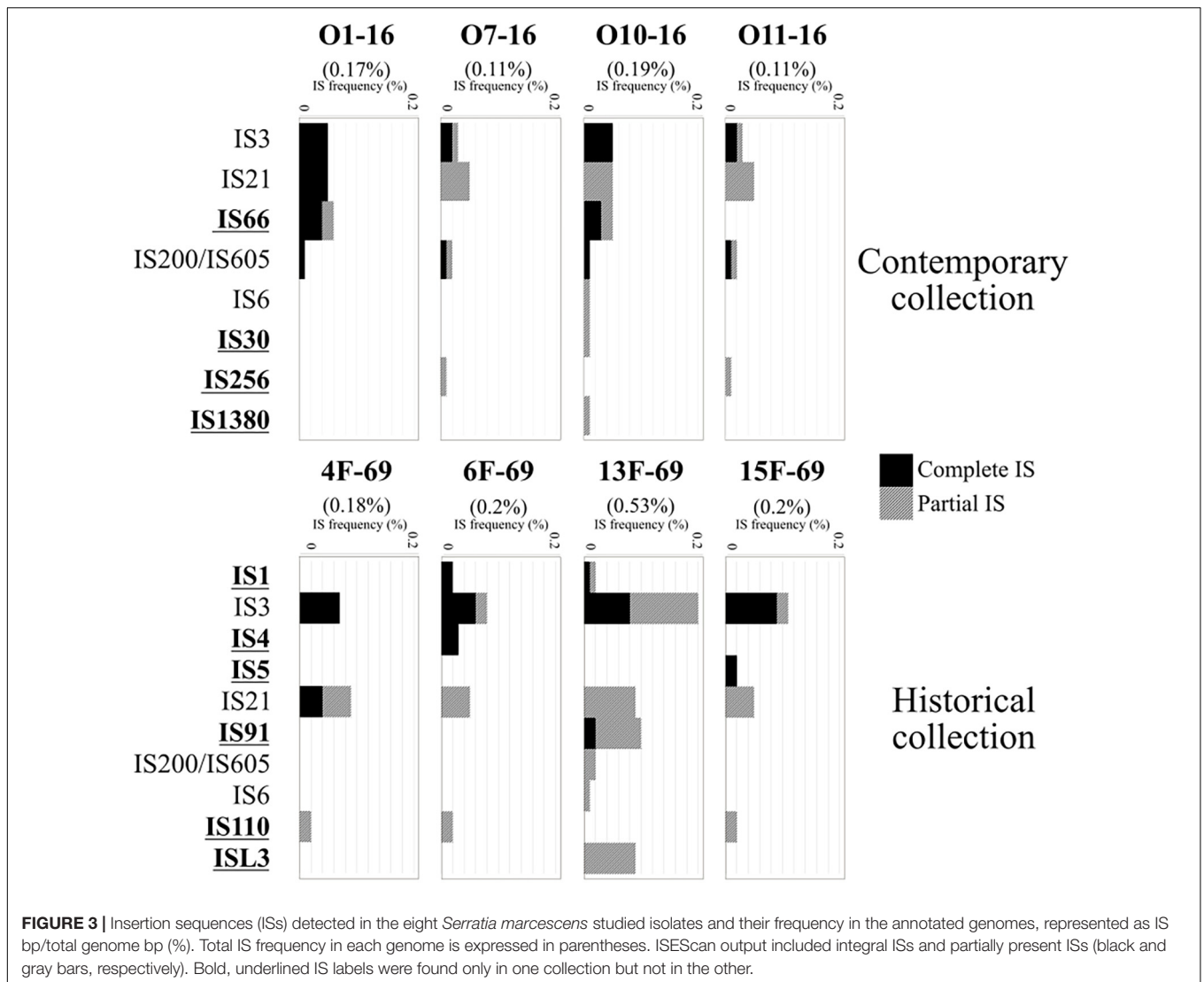
## *Serratia marcescens* Core Genome Multilocus Sequence Typing Scheme

The aim of this analysis was to compare all eight historical or contemporary isolates along with 444 complete genomes in the *S. marcescens* database by constructing a cgMLST scheme. This analysis showed that the *S. marcescens* pangenome comprised a total of 36,595 genes, 7.6% of which were shared by >95% of the 452 isolates (core genome) (Figure 6A). Phylogeny was inferred based on allelic variations across this core genome (Figure 6B). Contemporary isolates clustered together and were closely related to a human clinical specimen obtained from blood in 2006 in the United Kingdom (strain: 2880STDY5682913, BioProject: PRJEB5065). This 2006 strain showed 26,242 nucleotide variants when compared to the O1-16 strain and 703 variants when compared to the O7-16 strain. Historical isolates were located in a separate branch close to a strain from the same BioProject collected in 2011 (strain: 2880STDY5683020). Again, the historical 13F-69 isolate was the most distant and was related to clinical strains from Brazil (BioProject reference: PRJNA420811). It is important to note that most of the annotated public genomes have a clinical or hospital origin (76.5%), whereas other sources [environmental (11.7%), animal (3.3%), plant (2.4%), and miscellaneous (6%)] are considerably less represented.

## Accessory Genome Comparison Using Publicly Available Databases

The content of the accessory genome of the eight studied outbreak isolates was compared with the 444 previously annotated *S. marcescens* genomes deposited in public databases using AcCNET software. This analysis differentiated all genomes into nine clusters according to their predicted accessory proteome content. Historical strains were located in cluster 6, with the exception of isolate 13F-69, which was included in cluster 3, and contemporary strains were located in cluster 7 (Figure 7A). This distribution regarding the accessory genome was compared to that inferred by cgMLST analysis in Figure 6B, and all strains were clustered in a similar manner. No outbreak isolates were found in any of the other seven clusters of *S. marcescens*.

The hierarchical clustering method used in this work minimizes the total within-cluster variance, and the nine resulting clusters were the most homogeneous possible. Cluster 7 was the most frequently represented among the *S. marcescens* genome collection, grouping 33.2% of the total isolates, followed by cluster 6 (14.2%) and cluster 1 (11.5%). The other clusters are less represented probably because they include infrequent accessory genes that can only be found sporadically. We cannot rule out the possibility that overrepresentation of cluster 7 could have been influenced by sampling bias (see section “Discussion”).



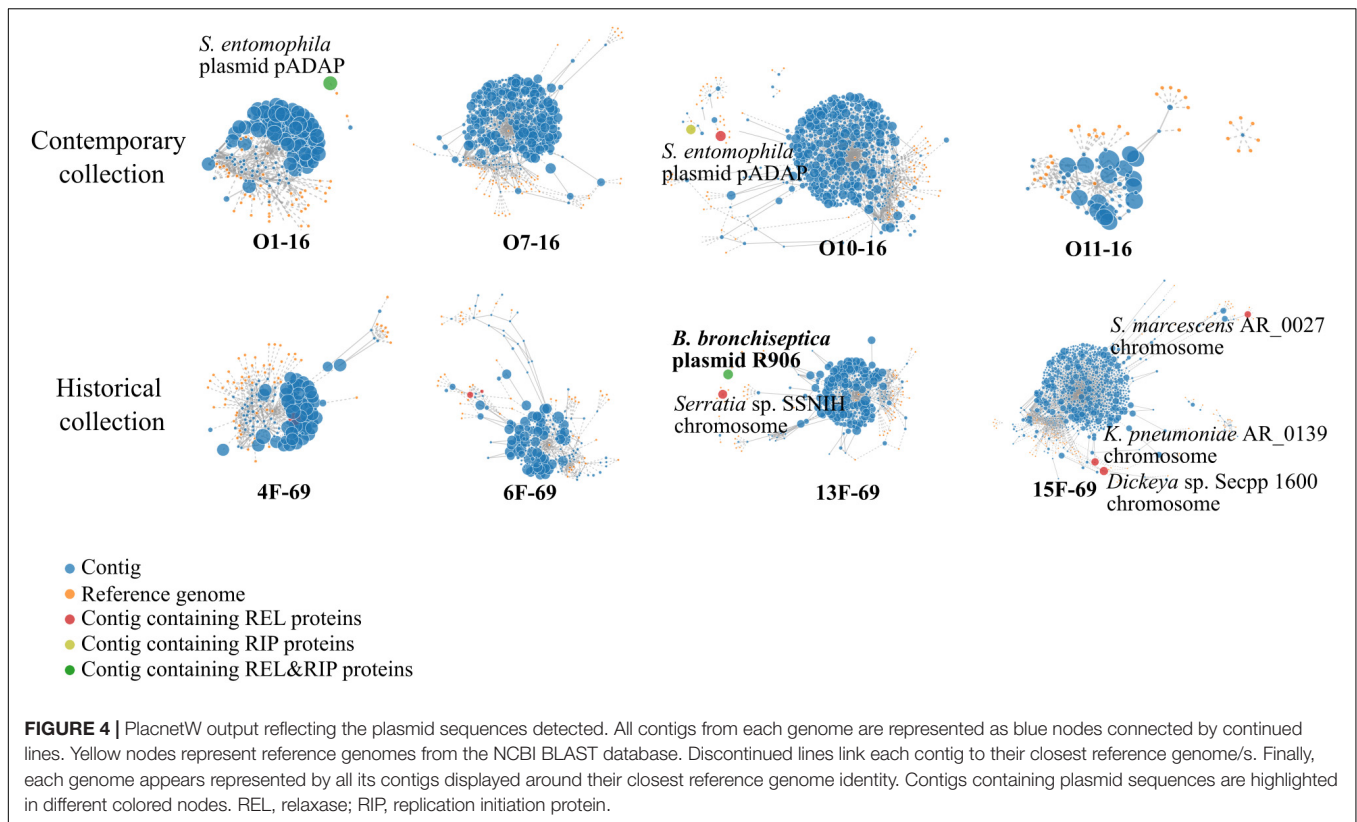
The functional category of each accessory gene was assigned using the COGs database. Interestingly, the genes of the nine clusters had analogous functions, even though the genetic content was different (Figure 7B). The top significant proteins of the accessory genome separating clusters 3, 6, and 7 from the rest are summarized in Table 7.

## DISCUSSION

Epidemiological studies combined with bacterial typing are essential for controlling nosocomial outbreaks, given they help to detect high-risk hospital-adapted clones, which have the potential to spread globally, and to improve local containment measures. Bacterial typing techniques have evolved over time—from phenotypic tests (biotypes, antibiograms) in the 70s of the previous century to current molecular studies based on DNA, with a superior discrimination capacity, better interlaboratory reproducibility, and greater cost-effectiveness

(Mellmann et al., 2017; Pérez-Losada et al., 2017). In this study, we had the opportunity to re-type by WGS a historical collection of *S. marcescens* causing a NICU outbreak. Those isolates had initially been classified as forming part of a monoclonal outbreak; however, we were able to distinguish the coexistence of two lineages by PFGE, and WGS demonstrated the apparent individuality of each one of the sequenced isolates. In addition, within the same NICU settings, we compared historical isolates with their contemporary counterparts trying to decipher whether there were common features that explained the successful transmission rate of these outbreak strains separated by 47 years.

The most frequently used typing techniques for relevant human-related pathogens are PFGE and MLST; however, for *S. marcescens*, the MLST scheme has not yet been implemented (Martineau et al., 2018). Currently, WGS is a much better option given it is cost-effective in comparison with MLST and it reveals the composition of the accessory genome. In addition, WGS allows (i) the detection of SNP mapping against a reference genome and (ii) a gene by gene-based comparative approach



**TABLE 3 |** Plasmid and prophage sequences annotated by PlacnetW and Phaster software, respectively.

Isolate	Plasmids and prophage sequences	Similarity <sup>a</sup>	Contigs <sup>b</sup>	Size (bp)	CDS
13F-69	<i>Serratia</i> sp. SSNIH <sup>c</sup>	50/99.7	1	119,213	115
	<b><i>B. bronchiseptica</i> R906 plasmid</b>	<b>99/99.6</b>	<b>1</b>	<b>57,271</b>	<b>64</b>
	<i>Siphoviridae</i> phage	100	–	17,500	18
15F-69	<i>S. marcescens</i> AR_0027 <sup>c</sup>	30/98	1	53,451	52
	<i>Dickeya</i> sp. Secpp 1600 <sup>c</sup>	41/98	1	26,298	12
	<i>K. pneumoniae</i> AR_0139 <sup>c</sup>	80/96.5	1	83,980	92
	<i>Siphoviridae</i> phage	110	–	43,600	61
O7-16	<i>Myoviridae</i> phage	150	–	34,400	40
	<i>Myoviridae</i> phage	110	–	20,400	25
	<i>Myoviridae</i> phage	110	–	20,400	25
O1-16	pADAP plasmid	55/76.9	1	83,980	92
O10-16	pADAP plasmid	64.3/75.2	2	66,877	85
O11-16	<i>Myoviridae</i> phage	150	–	34,400	40

The unique complete plasmid was highlighted in bold. CDS, protein-coding sequence. <sup>a</sup>Identity score of phages (minimum threshold = 60) and plasmids (expressed as query cover/% identity from BLAST report). <sup>b</sup>Number of contigs separately plotted in PlacnetW that matched with the same BLAST identification. <sup>c</sup>Contigs plotted as putative plasmids by PlacnetW software which sequences matched with fragments of bacterial chromosomes.

according to a previously constructed pangenome allele database (PGAdb) and further clustering of the genomes [whole genome multilocus sequence typing (wgMLST)] (Schürch et al., 2018). The dataset can also be curated to include only a set of shared genes, e.g., the core genome (cgMLST) given it has recently been used to distinguish *S. marcescens* hospital-adapted lineages from those of environmental origin (19) (Abreo and Altier, 2019). The SNP approach is more sensitive than cg/wgMLST analysis but lacks interlaboratory reproducibility due to the need to establish a concrete reference. Beyond epidemiological investigations,

WGS is useful for assessing the genetic evolution of bacterial species, particularly when the accessory genome is analyzed, as it has recently been described for *Enterococcus faecalis* (León-Sampedro et al., 2019). New genomic and proteomic analysis tools as, for example, AccNET, have significantly improved such analysis (Lanza et al., 2017).

In the present study, both SNP and cgMLST (taking a set of genes shared by >95% of the whole database) were performed, together with an accessory genome analysis. The pangenome composition obtained in our study included over

**TABLE 4** | Resistome derived from WGS analysis of the eight *Serratia marcescens* isolates of the study and major antibiotic classes which can be potential targets of the products the resistance genes encode.

Antibiotic Class/Resistome	O1-16	O7-16	O10-16	O11-16	4F-69	6F-69	13F-69	15F-69
<b>Beta-lactams</b>								
<i>bla</i> OXA-2							✓ (p)	
<i>bla</i> SRT-1	✓		✓				✓	
<i>bla</i> SRT-2		✓		✓	✓	✓		✓
<b>Aminoglycosides</b>								
<i>aac</i> (6')		✓	✓	✓	✓	✓	✓	✓
<i>aph</i> (3'')-Ib							✓ (p)	✓
<i>aph</i> (6)-Id							✓	✓
<b>Quinolones</b>								
<i>qnr</i> B75							✓	
<i>qnr</i> E1	✓		✓		✓	✓		✓
<b>Sulfonamide</b>								
<i>sul</i> -1							✓ (p)	
<b>Multidrug (oqx, chl)</b>								
<i>oqx</i> B24							✓	
<i>oqx</i> B9	✓	✓	✓	✓	✓	✓		✓
<b>QUATS</b>								
<i>qacE</i> Δ1							✓ (p)	
<b>Tetracycline</b>								
<i>tet</i> (41)							✓	

CHL, chloramphenicol; OQX, olaquinox; WGS, whole-genome sequencing; (p), gene located in plasmid.

**TABLE 5** | Observed minimal inhibitory concentrations (MICs) (mg/L) or zone of inhibition diameter (ZD) (mm) of each antibiotic for the eight studied *Serratia marcescens* strains.

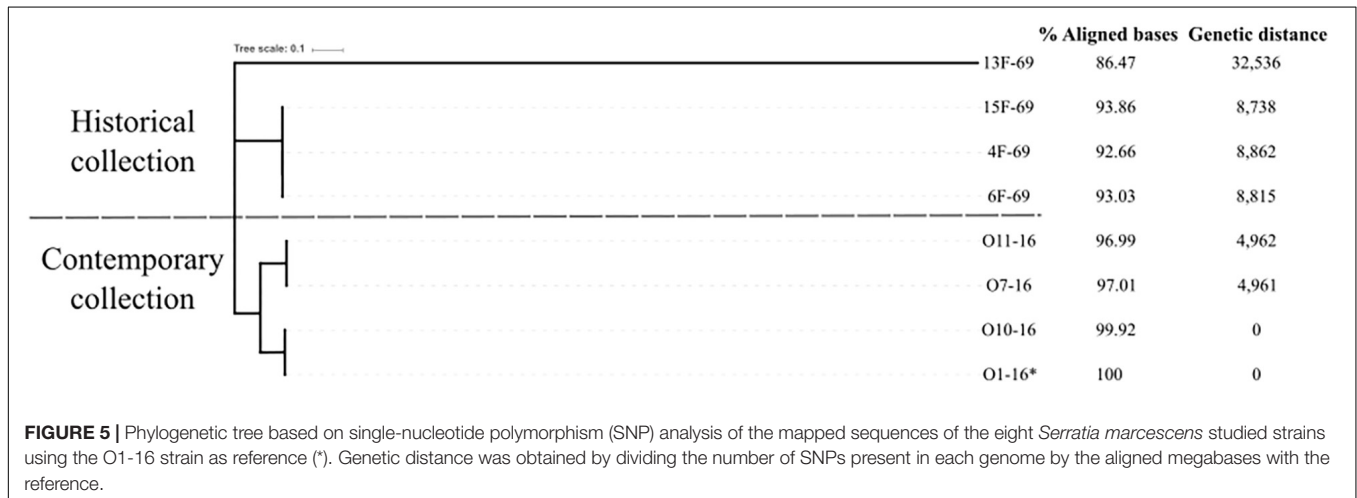
Antibiotics and Resistance genes	Available since	Contemporary collection				Historical collection			
		MICs (mg/L) or ZD (mm)				MICs (mg/L) or ZD (mm)			
		O1-16	O7-16	O10-16	O11-16	4F-69	6F-69	13F-69	15F-69
<b>Beta-lactams</b>									
Ampicillin	1961	<b>&gt;16</b>	<b>&gt;16</b>	<b>&gt;16</b>	<b>&gt;16</b>	<b>&gt;16</b>	<b>&gt;16</b>	<b>&gt;16</b>	<b>&gt;16</b>
Amoxicillin–Clavulanate	1972	<b>&gt;16/8</b>	<b>&gt;16/8</b>	<b>&gt;16/8</b>	<b>&gt;16/8</b>	<b>&gt;16/8</b>	<b>&gt;16/8</b>	<b>&gt;16/8</b>	<b>&gt;16/8</b>
Piperacillin–Tazobactam	1993	≤8	≤8	≤8	≤8	≤8	<b>≥64</b>	≤8	≤8
Cefazolin	1971	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>
Cefuroxime	1978	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>	<b>≥16</b>
Cefoxitin	1972	16	<b>≥16</b>	<b>≥16</b>	16	16	<b>≥16</b>	16	16
Ceftazidime	1984	≤1	≤1	<b>≥16</b>	≤1	≤1	<b>≥16</b>	≤1	≤1
Cefotaxime	1980	≤1	≤1	<b>≥32</b>	≤1	≤1	<b>≥32</b>	≤1	≤1
Cefepime	1994	≤1	≤1	<b>≥16</b>	≤1	≤1	<b>8</b>	≤1	≤1
Aztreonam	1986	≤1	≤1	<b>≥16</b>	≤1	≤1	<b>≥16</b>	≤1	≤1
Imipenem	1977	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1
Ertapenem	2001	≤0.5	≤0.5	<b>≥4</b>	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5
<b>Aminoglycosides</b>									
Gentamicin	1963	≤2	≤2	<b>≥8</b>	≤2	≤2	≤2	≤2	≤2
Tobramycin	1997	≤2	≤2	<b>≥8</b>	≤2	≤2	4	4	≤2
Amikacin	1976	≤8	≤8	<b>≥32</b>	≤8	≤8	≤8	≤8	≤8
Streptomycin*	1943	15	15	<b>11</b>	16	<b>13</b>	17	<b>0</b>	<b>9</b>
<b>Ciprofloxacin</b>									
Ciprofloxacin	1987	≤0.5	≤0.5	<b>1</b>	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5
<b>Nalidixic acid</b>									
Nalidixic acid	1965	≤16	≤16	>16	≤16	≤16	≤16	≤16	≤16
<b>TMP/SMX</b>									
TMP/SMX	1968-78	≤2/38	≤2/38	≤2/38	≤2/38	≤2/38	≤2/38	≤2/38	≤2/38
<b>Tetracycline*</b>									
Tetracycline*	1953	<b>14</b>	<b>12</b>	<b>15</b>	<b>13</b>	21	21	<b>0</b>	<b>12</b>

MIC values or zone diameters over the clinical cutoff (based on EUCAST 2019 criteria) are highlighted in bold, underlined numbers. NA, not available; TMP/SMX, trimethoprim/sulfamethoxazole. Asterisk, antibiotics tested only by disk diffusion test (25 μg for streptomycin and 30 μg for tetracycline).

**TABLE 6** | SNP matrix including all SNPs present between pairs of isolates.

	O1-16	O7-16	O10-16	O11-16	4F-69	6F-69	13F-69	15F-69
O1-16	0	24,557	0	24,524	41,900	41,895	141,431	41,916
O7-16	24,554	0	24,555	1	42,382	42,366	141,588	42,403
O10-16	0	24,562	0	24,516	41,898	41,878	141,320	41,914
O11-16	24,562	0	24,556	0	42,392	42,371	141,574	42,405
4F-69	41,864	42,376	41,865	42,340	0	37	141,842	53
6F-69	41,779	42,295	41,788	42,239	36	0	141,382	23
13F-69	143,127	143,429	143,332	143,324	143,904	143,990	0	143,868
15F-69	41,821	42,330	41,817	42,287	49	19	141,762	0

Isolates in columns are the ones to map the SNP content to. SNP, single-nucleotide polymorphism.



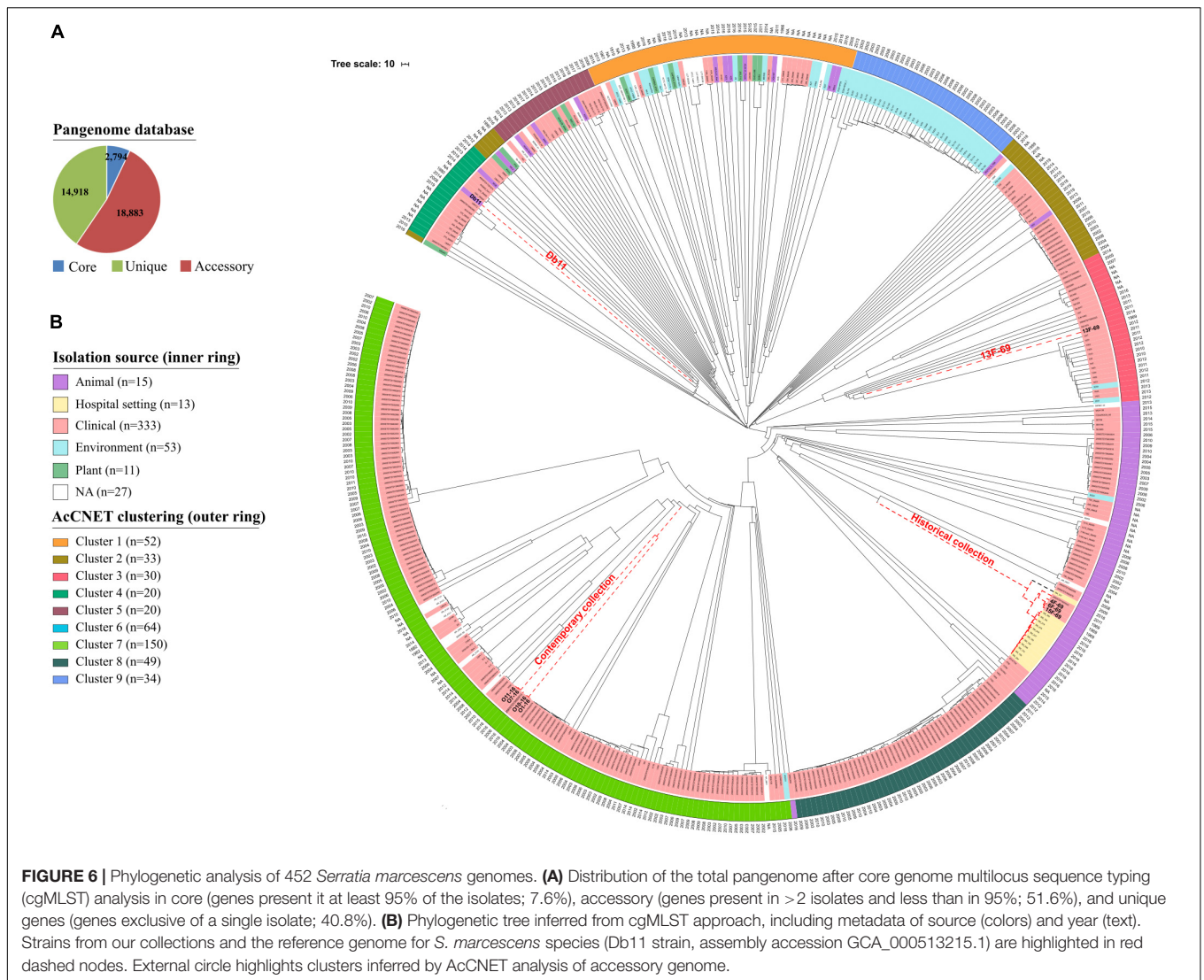
36,000 gene clusters. Probably wide pangenomes correspond to ubiquitous organisms submitted to many environmental conditions. In contrast, pangenomes from a bacterial organism with a high degree of habitat specialization are much more limited. The pangenome inferred from 50 *Streptococcus* genomes is considerably lower, 5,398 gene clusters (Contreras-Moreira and Vinuesa, 2013). The size of *Serratia* pangenome is closer to that of another environmental-promiscuous genus, *Vibrio*, composed of over 26,000 gene clusters in a study involving 32 genomes (Thompson et al., 2009). These findings reveal the high diversity of unique genes present in *S. marcescens*, most probably corresponding to genes present in mobile genetic elements, genomic islands, transposons, and prophages.

Isolates from both historical and recent collections were consistently grouped by sampling period, being the historical 13F-69 the most outstanding isolate in all analyses performed. Both the core genome and accessory genome analyses grouped both collections in a similar way (Figure 6B), a fact that has been previously described in similar scenarios (Moradigaravand et al., 2016; Abreo and Altier, 2019). This observation indicates that *S. marcescens* appears to evolve in the same direction for both core and accessory genes.

*Serratia marcescens* is particularly associated to NICU environment, being certain lineages more associated to sepsis than others (Moles et al., 2019). The coexistence of several clones in the same *S. marcescens* outbreak, as occurred in our historical

collection, had previously been reported (Fleisch et al., 2002; Montagnani et al., 2015; Dawczynski et al., 2016). The cgMLST scheme revealed that most of our isolates share a common ancestor with most of the published clinical strains. However, a caveat of this analysis is the low representation of non-clinical isolates (<24% of the total) in the public genome databases. Comparing our isolates with the previously published genomes, the contemporary collection was located next to a multidrug-resistant nosocomial strain collected in the United Kingdom in 2006; the historical collection next to several strains collected from ICU surfaces in Pakistan in 2016; and finally, the 13F-69 isolate was grouped in a separate branch with nosocomial isolates from Brazil isolated in 2012. This analysis, to our knowledge, is the most extensive study of the phylogeny of *S. marcescens* to date; however, as previously mentioned, an important lack of isolates from non-hospital sources was observed.

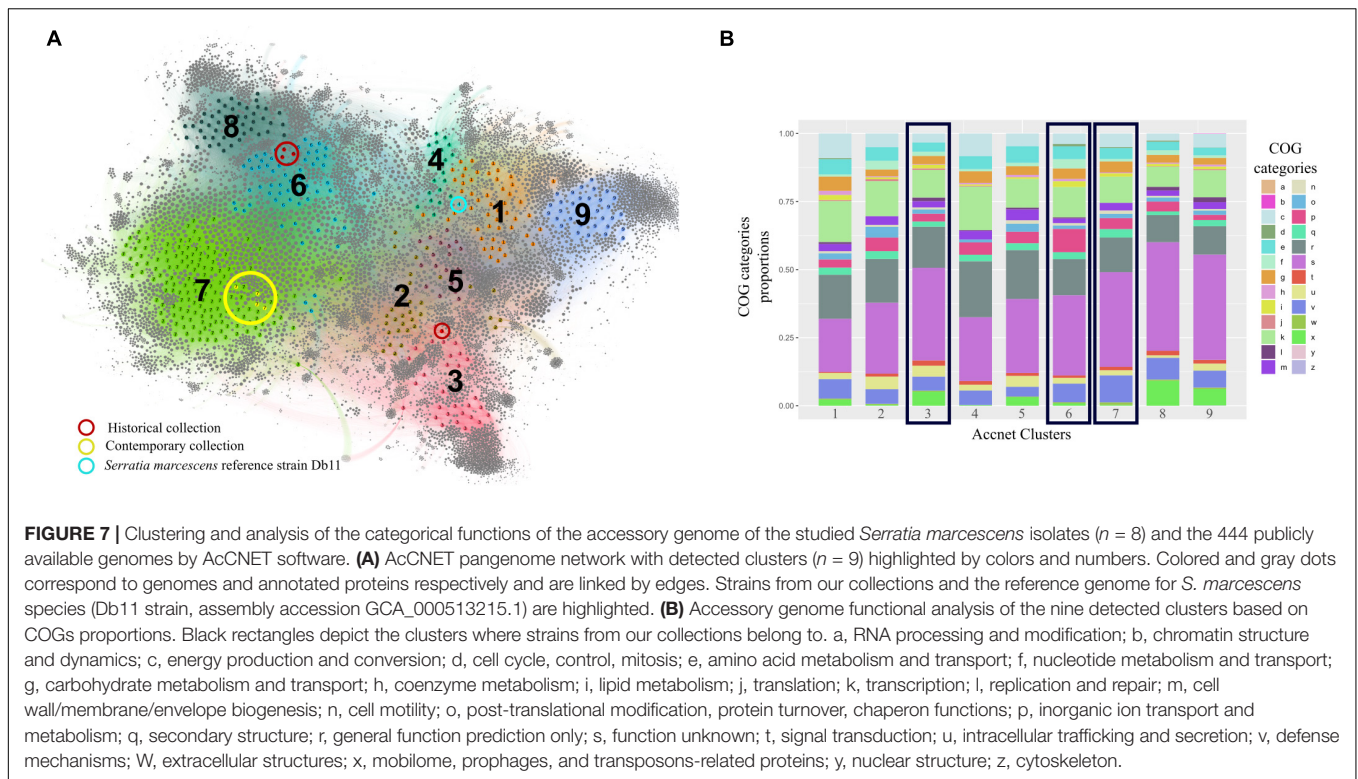
The accessory genome determines the idiosyncrasy of each strain, but it appears to correlate with the core genome content throughout evolution, as previously observed (Moradigaravand et al., 2016). Each one of the nine clades that we were able to discriminate had a unique combination of accessory genes compared with the whole collection. On one hand, clusters 6 and 7 appeared to be more related to each other than to cluster 3, where 13F-69 isolate was located. Both clusters share proteins involved in defense mechanisms such as type II toxin/antitoxin systems and multidrug efflux pumps, and



also in the cell cycle, intracellular trafficking and secretion, and inorganic ion transport and metabolism. Cluster 6 seems to be enriched in proteins involved in nucleotide metabolism and transport, such as aminohydrolases and dehydrogenases for anaerobic nitrogen utilization, whereas cluster 7 is enriched in proteins related to energy production and conversion, and they are more faithfully assigned to the taxonomic range of *S. marcescens* species. On the other hand, proteins in cluster 3 are more frequently assigned to upper taxonomic ranges such as phylum or order (*Proteobacteria*, *Enterobacterales*). This cluster is enriched in proteins related to translation processes, intracellular trafficking and secretion, and proteins that constituted the mobilome, prophages, and transposons-related proteins such as conjugative systems, transposases, and plasmid replication initiators. Overall, the accessory genome of our strains carried typical features that allowed them to survive in a clinical setting. In a previous publication, several efflux pumps had been detected in a collection of 32 *Serratia* genomes from various sources, with the clinical strains the ones carrying the highest number

of efflux pumps and acquired resistance genes *via* horizontal gene transfer elements (Sandner-Miranda et al., 2018). Resistance genes coding for efflux pumps, as well as their regulatory elements, should be taken into consideration as influencing the success of antibiotic therapy.

A relevant finding was that the 13F-69 strain harbors a highly conserved plasmid previously described in *B. bronchiseptica*, a zoonotic pathogen of the respiratory airways. Plasmid exchange between phylogenetically distant species (in our case from Beta-*Proteobacteria* to Gamma-*Proteobacteria* taxa) is a rare event. Such transmission appears to have occurred despite distant natural niches, although the gut-lung axis or maybe a lung coinfection could justify a common nexus (Matthieu et al., 2018). In any case, the full preservation of the R906 plasmid after a rest period of 47 years is noteworthy. Of course, the origin of the plasmid was inferred from the available data, and we cannot rule out *Serratia* as its primary host. Finally, this plasmid contains the YafQ-DinJ toxin-antitoxin system, which could be involved in long-term plasmid maintenance. Four prophages were identified



in the contemporary collection, whereas only one was found in the historical collection, a result probably related to current genome plasticity requirements (Shen et al., 2019).

Historical collections of bacteria provide us with relevant data from the pre-antibiotic period (Baker et al., 2014, 2015; Devault et al., 2014). Our historical isolates were collected from 1969 to 1970 during the antibiotics classic golden age, but they had not been exposed to newly discovered antimicrobials. However, these historical isolates carried genes coding for resistance to antibiotics and disinfectants that were commercialized years later. Although olaquinox is an antimicrobial agent only used since 1975 as a growth promoter for farm animals, the olaquinox resistance genes *oqx10*, *oqx24*, and *oqx9* were consistently found in our historical isolates. This finding could possibly be explained by other functions that are provided by the multidrug efflux pumps encoded by *oqx* genes, including quinolone and benzalkonium chloride resistance (Hansen et al., 2007). Plasmid-mediated quinolone-resistance *qnr* genes were discovered in 2002 (Tran and Jacoby, 2002); however, *qnrB75* and *qnrE1* were already present in our historical isolates probably as the *qnr* function is not related primarily with antibiotic resistance (Hernández et al., 2011). Alternatively, they could have emerged under quinolone selection. In our days, quinolones or fluoroquinolones are not used in pediatrics. However, in 1969, at the time of the first *S. marcescens* outbreak, the first quinolone, nalidixic acid, was extensively used to control epidemic enteritis by enteropathogenic *Escherichia coli* at La Paz Children's Hospital, which could have contributed to the selection of *qnr* genes in intestinal *Serratia* strains.

Intrinsic resistance in *Serratia* includes penicillins, macrolides, lincosamines, linezolid, glycopeptides, quinupristin-dalfopristin, and rifampin (Mahlen, 2011). Ampicillin and tetracycline resistance has typically been considered intrinsic, but that assumption could be due to the bias imposed by clinical strains, and a number of environmental strains could be susceptible (Ehrenkranz et al., 1980). Probably all *S. marcescens* strains harbor a chromosomal AmpC beta-lactamase, normally repressed. Mutants in AmpC and AmpD (restricting the function of AmpR, the AmpC repressor) result in a constitutive expression of AmpC, now able to deactivate third- and fourth-generation cephalosporins (Raimondi et al., 2001; Kaneko et al., 2005; Meini et al., 2018). This appears to have occurred in the historical strain 6F-69 and the contemporary strain O10-16, resistant to cefotaxime, ceftazidime, and cefepime. However, we were unable to detect mutational changes in the structural or known regulatory genes of the AmpC beta-lactamase (*ampC*, *ampD*, *ampE*, *ampG*, *ampR*, and their flanking sequences); this genotype-phenotype discrepancy has been previously observed in other species (Campbell et al., 1997) so that probably the complex regulation of AmpC production in *Serratia* should be further clarified in the future. Strain O10-16 was also resistant to carbapenems probably due to concurrence with porin mutations impairing antibiotic entry (Mammeri et al., 2008). In the 1970s, *S. marcescens* antimicrobial therapy was based mostly on carbenicillin and gentamicin, sometimes combined with fosfomycin (Baquero et al., 1977). However, the historical 6F-69 isolate exhibited resistance to cephalosporins that had not been discovered in 1969 probably because of the hospital

**TABLE 7** | Top relevant proteins of the accessory genome that significantly explained the discrimination of each of the three clusters to which the eight *Serratia marcescens* studied strains belong to, with respect to the overall population, analyzed by AcCNET software.

COGs Categories	Cluster 3 (13F-69 isolate)	Cluster 6 (Historical collection)	Cluster 7 (Contemporary collection)
<b>C</b>	Gfo/Idh/MocA family [ <i>Enterobacterales</i> ] Taurine dioxygenase TauD [ <i>Serratia</i> ] Tartrate dehydrogenase [ <i>Enterobacterales</i> ]	CbbBc protein Nitrite reductase	NADP-oxidoreductase [ <i>Enterobacterales</i> ] Short-chain reductase [ <i>Enterobacterales</i> ] (2Fe-2S)-binding protein Electron transport complex
<b>D</b>	NA	Cell division protein Fic/DedD	Cell division protein Fic
<b>F</b>	(S)-ureidoglycine aminohydrolase [ <i>Enterobacterales</i> ] Ureidoglycolate dehydrogenase [ <i>Enterobacterales</i> ]	Pyridine nucleotide-disulfide oxidoreductase [ <i>Serratia</i> ] (S)-ureidoglycine aminohydrolase [ <i>Enterobacterales</i> ] Ureidoglycolate dehydrogenase [ <i>Enterobacterales</i> ] 5'-nucleotidase, lipoprotein e(P4) family	Pyridine nucleotide-disulfide oxidoreductase [ <i>Serratia</i> ] Tyrosine-protein kinase Wzc [ <i>Serratia</i> ]
<b>J</b>	Methylase [ <i>Enterobacterales</i> ] tRNA-binding protein [ <i>Enterobacterales</i> ] tRNA(fMet) endonuclease [ <i>Serratia</i> ]	NA	NA
<b>L</b>	Recombinase family protein Antirestriction protein Conjugative transfer relaxase DNA-damage inducible protein J [ <i>Serratia</i> ]	DNA polymerase III [ <i>Enterobacterales</i> ]	NA
<b>P</b>	TonB-dependent receptor Ions translocating ATPase [ <i>Serratia</i> ] Siderophore biosynthesis protein [ <i>Serratia</i> ]	TonB-dependent receptor Ions translocating ATPase [ <i>Serratia</i> ] Multicopper oxidase CueO	TonB-dependent receptor Ions translocating ATPase [ <i>Serratia</i> ] Multicopper oxidase CueO Ferrichrome porin FhuA
<b>U</b>	Hemolysin secretion protein Type VI secretion system [ <i>Serratia</i> ] Type IV secretion system [ <i>Enterobacterales</i> ] Microcin H47 secretion protein [ <i>Serratia</i> ] LysE family translocator [ <i>Serratia</i> ]	Hemolysin secretion protein Type II secretion system HlyD family secretion protein	Hemolysin secretion protein Type II secretion system HlyD family secretion protein
<b>V</b>	Type 1 fimbrial protein Type II toxin-antitoxin system Colicin V biosynthesis protein Tetracycline efflux transporter Aminoglycoside transporter MBL fold metallo-hydrolase Stress protection protein MarC [ <i>Serratia</i> ] Tellurite-resistance protein TehA Exotoxin	Type 1 fimbrial protein Type II toxin-antitoxin system [ <i>Serratia</i> ] Multidrug efflux pump RND [ <i>Serratia</i> ] Toxin-activating lys-acyltransferase [ <i>Serratia</i> ] Antibiotic biosynthesis monooxygenase [ <i>Enterobacterales</i> ] Class A beta-lactamase-related serine hydrolase Addiction module antidote protein [ <i>Enterobacterales</i> ]	Type 1 fimbrial protein Type II toxin-antitoxin system [ <i>Serratia</i> ] Multidrug efflux pump RND [ <i>Serratia</i> ] Toxin-activating lys-acyltransferase [ <i>Serratia</i> ] Antibiotic biosynthesis monooxygenase [ <i>Enterobacterales</i> ] Exotoxin Filamentous hemagglutinin Antitoxin [ <i>Serratia</i> ]
<b>X</b>	Type-F conjugative transfer system [ <i>Klebsiella</i> ] IS110 transposase [ <i>Proteobacteria</i> ] IS200/IS605 transposase Holin [ <i>Enterobacterales</i> ] Plasmid replication initiator [ <i>Enterobacterales</i> ] IS5 transposase Competence protein TfoX	Killer protein [ <i>Serratia</i> ] Prevent-host-death protein [ <i>Serratia</i> ]	Transposase

Note that some classes of proteins are shared among clusters. Some taxonomic ranges are included when the annotated proteins did not reach the category of *S. marcescens* species. COGs, Clusters of Orthologous Groups of proteins; NA, not applicable.

exposure and selection of de-repressed mutants by cephaloridine, introduced in 1962. Likewise, gentamicin resistance was detected in the modern O7-16 isolate even though most *S. marcescens* isolates used to be susceptible due to the infrequent use of this antimicrobial (Moradigaravand et al., 2016).

Finally, it is relevant to discuss the potential influence of the prolonged rest period experienced by our “old” strains. In a classic publication (Naas et al., 1994), Werner Arber’s group compared a strain of *E. coli* K-12 (W3110) stabbed in 1960 with subclones recovered in 1965, 1972, and 1990 from small portions of the original stab, which was resealed immediately. Over these 30 years, the resting culture appeared to have developed a high degree of genetic diversity, which was an unexpected finding, considering the reduced rate of propagation during storage. The source of most of such variability was attributed to IS transposition (Naas et al., 1994, 1995), suggesting that

the genome of resting bacteria is more dynamic than was long believed. In the present study, we found several ISs in our strains, with more diversity and a slightly higher proportion in the historical collection; however, major differences could not be established. This result could indicate less dynamism in resting strains of *S. marcescens* compared with those observed in *E. coli* (Naas et al., 1994). As mentioned before, WGS demonstrated the carriage in the resistome of the historical strain 13F-69 of several antibiotic-resistant genes, but from some of them, as *qnrB*, *qnrE*, acting on quinolones, or *tet(41)*, encoding a tetracycline efflux pump, failed to produce the expected resistance phenotype. We cannot discard a decrease in expression as a strategy to reduce the fitness cost of the organisms under a very prolonged rest period to decrease fitness costs (Enne et al., 2006; Allen et al., 2017).

In summary, phylogenetic analysis using large genome databases allowed us to demonstrate that both core genome and

accessory genome typing strategies grouped *S. marcescens* strains in a similar manner. Both historical and recent outbreak isolates clustered in two groups where most of the clinical strains were found. The only exception was the historical 13F-69 isolate, closer to strains with a more “environmental” style of life, probably open to high genetic exchange, including the acquisition of a complete plasmid from another environmental genus, *Bordetella* (Hamidou Soumana et al., 2017). Historical strains carried resistance genes toward antimicrobials not yet commercialized in 1969, suggesting the presence of factors driving the evolution and transmission of those genes that are not necessarily related to the antibiotic therapy. Finally, our historical isolates appeared to show a certain evolutionary dynamism during their resting period, although much lower than previously described for *E. coli*.

As shown in this study, WGS is becoming a critical tool for bacterial typing, permitting the analysis of core and accessory genomes with powerful discrimination. Even more importantly, WGS enlarges our understanding of the phylogenetic history of our outbreak isolates, offering an evolutionary perspective that will be useful for developing focused public health interventions (Baquero et al., 2015).

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the BioProject PRJNA510235.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Spanish Legislation for Good Practices and the Ley Orgánica 3/2018 on Personal Data Protection. The protocol was approved by the Ethical Committee of the University Hospital La Paz. All subjects or their legal guardians gave written informed consent in accordance with the Declaration of Helsinki.

## AUTHOR CONTRIBUTIONS

FB and RC conceived the experiments and supervised the final manuscript. EE, FL-P, MP, and JR provided the data

## REFERENCES

- Abreo, E., and Altier, N. (2019). Pangenome of *Serratia marcescens* strains from nosocomial and environmental origins reveals different populations and the links between them. *Sci. Rep.* 9:46. doi: 10.1038/s41598-018-37118-0
- Allen, R. C., Engelstädter, J., Bonhoeffer, S., McDonald, B. A., and Hall, A. R. (2017). Reversing resistance: different routes and common themes across pathogens. *Proc. Biol. Sci.* 284, 20171619. doi: 10.1098/rspb.2017.1619
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387

or strains. CS, MP-A, BP-V, VL, and LM performed the experiments, bioinformatic analysis, and drafted the initial manuscript. All authors read and approved the final version of the manuscript.

## FUNDING

CS was supported by “Fundación Mutua Madrileña” grant to RC achieved in 2017 call with reference number AP165902017. MP-A was supported by the Programa Operativo de Empleo Juvenil, cofinanced by the European Social Fund Investing in your future (ESF) and ERDF (PEJD-2018-PRE/BMD-8237). BP-V was funded by H2020 FTIPilot 2016 project no. 730713 “FAST-bact “A novel fast and automated test for antibiotic susceptibility testing for Gram positive and negative bacteria” and co-funded by Instituto de Salud Carlos III and the European Regional Development Fund (ERDF, “A way to achieve Europe”). FB was supported by grants from the Madrid Regional Government (InGEMICS-C; S2017/BMD-3691) and CIBER (CIBER in Epidemiology and Public Health, CIBERESP; CB06/02/0053), co-funded by Instituto de Salud Carlos III and the European Regional Development Fund (ERDF, “A way to achieve Europe”). This work was supported by the Instituto de Salud Carlos III, PI17/00115 (RC), and REIPI (RD16/0016/0011) actions, cofinanced by the European Development Regional Fund “A way to achieve Europe” (ERDF).

## ACKNOWLEDGMENTS

We are grateful to María Isabel Morosini for antimicrobial susceptibility determination. We would like to thank Mr. Liu and colleagues for their support and help with cgMLST analysis.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.00451/full#supplementary-material>

**DATA SHEET S1** | Metadata of the 452 genomes of *Serratia marcescens* available in the database of the National Center for Biotechnology Information (NCBI) RefSeq that have been used in this work.

- Baker, K. S., Burnett, E., McGregor, H., Deheer-Graham, A., Boinett, C., Langridge, G. C., et al. (2015). The Murray collection of pre-antibiotic era *Enterobacteriaceae*: a unique research resource. *Genome Med.* 7:97. doi: 10.1186/s13073-015-0222-7
- Baker, K. S., Mather, A. E., McGregor, H., Coupland, P., Langridge, G. C., Day, M., et al. (2014). The extant World War 1 dysentery *Bacillus* NCTC1: a genomic analysis. *Lancet (Lond. Engl.)* 384, 1691–1697. doi: 10.1016/S0140-6736(14)61789-X
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

- Baquero, F., Carvajal, A., de Rafael, L., Reig, M., and Sales, R. (1969). "Serratia: un nuevo problema en microbiología hospitalaria," in *Proceedings of the 2nd Congreso Nacional de Microbiología* (Madrid: Sociedad Española de Microbiología), 126.
- Baquero, F., Hortelano, J. G., Navarro, M., Scarpellini, A., Jara, P., Cañedo, T., et al. (1977). Antibiotherapy of *Serratia marcescens* septicemia in children. *Chemother* 23, 416–422. doi: 10.1159/000222084
- Baquero, F., Lanza, V. F., Cantón, R., and Coque, T. M. (2015). Public health evolutionary biology of antimicrobial resistance: priorities for intervention. *Evol. Appl.* 8, 223–239. doi: 10.1111/eva.12235
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Campbell, J. I., Ciofu, O., and Høiby, N. (1997). *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis have different beta-lactamase expression phenotypes but are homogeneous in the ampC-ampR genetic region. *Antimicrob. Agents Chemother* 41, 1380–1384. doi: 10.1128/aac.41.6.1380
- Contreras-Moreira, B., and Vinuesa, P. (2013). GET\_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl. Environ. Microbiol.* 79, 7696–7701. doi: 10.1128/AEM.02411-13
- Dawczynski, K., Proquitté, H., Roedel, J., Edel, B., Pfeifer, Y., Hoyer, H., et al. (2016). Intensified colonisation screening according to the recommendations of the german commission for hospital hygiene and infectious diseases prevention (KRINKO): identification and containment of a *Serratia marcescens* outbreak in the neonatal intensive care unit, Jena, Germany, 2013–2014. *Infection* 44, 739–746. doi: 10.1007/s15010-016-0922-y
- Devault, A. M., Golding, G. B., Waglechner, N., Enk, J. M., Kuch, M., Tien, J. H., et al. (2014). Second-pandemic strain of *Vibrio cholerae* from the Philadelphia cholera outbreak of 1849. *N. Engl. J. Med.* 370, 334–340. doi: 10.1056/NEJMoa1308663
- Ehrenkranz, N. J., Bolyard, E. A., Wiener, M., and Cleary, T. J. (1980). Antibiotic-sensitive *Serratia marcescens* infections complicating cardiopulmonary operations: contaminated disinfectant as a reservoir. *Lancet* 2, 1289–1292. doi: 10.1016/s0140-6736(80)92349-1
- Enne, V. I., DelSol, A. A., Roe, J. M., and Bennet, P. M. (2006). Evidence of antibiotic resistance gene silencing in *Escherichia coli*. *Antimicrob. Agents Chemother* 50, 3003–3010. doi: 10.1128/AAC.00137-06
- Escribano, E., Saralegui, C., Moles, L., Montes, M. T., Alba, C., Alarcón, T., et al. (2019). Influence of a *Serratia marcescens* outbreak on the gut microbiota establishment process in low-weight preterm neonates. *PLoS One* 14:e0216581. doi: 10.1371/journal.pone.0216581
- Fleisch, F., Zimmermann-Baer, U., Zbinden, R., Bischoff, G., Arlettaz, R., Waldvogel, K., et al. (2002). Three consecutive outbreaks of *Serratia marcescens* in a neonatal intensive care unit. *Clin. Infect. Dis.* 34, 767–773. doi: 10.1086/339046
- Hadfield, J., Croucher, N. J., Goater, R. J., Abudahab, K., Aanensen, D. M., and Harris, S. R. (2017). Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics* 34, 292–293. doi: 10.1093/bioinformatics/btx610
- Hamidou Soumana, I., Linz, B., and Harvill, E. T. (2017). Environmental origin of the genus *Bordetella*. *Front. Microbiol.* 8:28. doi: 10.3389/fmicb.2017.00028
- Hansen, L. H., Jensen, L. B., Sørensen, H. I., and Sørensen, S. J. (2007). Substrate specificity of the OqxAB multidrug resistance pump in *Escherichia coli* and selected enteric bacteria. *J. Antimicrob. Chemother.* 60, 145–147. doi: 10.1093/jac/dkm167
- Hernández, A., Sanchez, M. B., and Martínez, J. L. (2011). Quinolone resistance: much more than predicted. *Front. Microbiol.* 2:22. doi: 10.3389/fmicb.2011.00022
- Kaneko, K., Okamoto, R., Nakano, R., Kawakami, S., and Inoue, M. (2005). Gene mutations responsible for overexpression of AmpC  $\beta$ -lactamase in some clinical isolates of *Enterobacter cloacae*. *J. Clin. Microbiol.* 43, 2955–2958. doi: 10.1128/JCM.43.6.2955-2958.2005
- Kim, D., Song, L., Breitwieser, F. P., and Salzberg, S. L. (2016). Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res.* 26, 1721–1729. doi: 10.1101/gr.210641.116
- Lanza, V. F., Baquero, F., de la Cruz, F., and Coque, T. M. (2017). AcCNET (Accessory Genome Constellation Network): comparative genomics software for accessory genome analysis using bipartite networks. *Bioinformatics* 33, 283–285. doi: 10.1093/bioinformatics/btw601
- León-Sampedro, R., Del Campo, R., Rodríguez-Baños, M., Lanza, V. F., Pozuelo, M. J., Francés-Cuesta, C., et al. (2019). Phylogenomics of *Enterococcus faecalis* from wild birds: new insights into host-associated differences in core and accessory genomes of the species. *Env. Microb.* 21, 3046–3062. doi: 10.1111/1462-2920.14702
- Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239
- Liu, Y. Y., Lin, J. W., and Chen, C. C. (2019). cano-wgMLST\_BacCompare: a bacterial genome analysis platform for epidemiological investigation and comparative genomic analysis. *Front. Microbiol.* 10:1687. doi: 10.3389/fmicb.2019.01687
- Mahlen, S. D. (2011). *Serratia* infections: from military experiments to current practice. *Clin. Microbiol. Rev.* 24, 755–791. doi: 10.1128/CMR.00017-11
- Mammeri, H., Nordmann, P., Berkani, A., and Eb, F. (2008). Contribution of extended-spectrum AmpC (ESAC) beta-lactamases to carbapenem resistance in *Escherichia coli*. *FEMS Microbiol. Lett.* 282, 238–240. doi: 10.1111/j.1574-6968.2008.01126.x
- Martineau, C., Li, X., Lalancette, C., Perreault, T., Fournier, E., Tremblay, J., et al. (2018). *Serratia marcescens* outbreak in a neonatal intensive care unit: new insights from next-generation sequencing applications. *J. Clin. Microbiol.* 56:e00235-18. doi: 10.1128/JCM.00235-18
- Matthieu, E., Escribano-Vázquez, U., Descamps, D., Cherbuy, C., Langella, P., Riffault, S., et al. (2018). Paradigms of lung microbiota functions in health and disease, particularly, in asthma (2018). *Front. Physiol.* 9:1168. doi: 10.3389/fphys.2018.01168
- Meini, S., Tascini, C., Cei, M., Sozio, E., and Rossolini, G. M. (2018). AmpC  $\beta$ -lactamase-producing *Enterobacteriales*: what a clinician should know. *Infection* 47, 363–375. doi: 10.1007/s15010-019-01291-9
- Mellmann, A., Andersen, P. S., Bletz, S., Friedrich, W., Kohl, T. A., Lilje, B., et al. (2017). High interlaboratory reproducibility and accuracy of next-generation-sequencing-based bacterial genotyping in a ring trial. *J. Clin. Microbiol.* 55, 908–913. doi: 10.1128/JCM.02242-16
- Moles, L., Gómez, M., Moroder, E., Jiménez, E., Escuder, D., Bustos, G., et al. (2019). *Serratia marcescens* colonization in preterm neonates during their neonatal intensive care unit stay. *Antimicrob. Resist. Infect. Control.* 8:135. doi: 10.1186/s13756-019-0584-5
- Montagnani, C., Cocchi, P., Lega, L., Campana, S., Biermann, K. P., Braggion, C., et al. (2015). *Serratia marcescens* outbreak in a neonatal intensive care unit: crucial role of implementing hand hygiene among external consultants. *BMC Infect. Dis.* 15:11. doi: 10.1186/s12879-014-0734-6
- Moradigaravand, D., Boinett, C. J., Martin, V., Peacock, S. J., and Parkhill, J. (2016). Recent independent emergence of multiple multidrug-resistant *Serratia marcescens* clones within the United Kingdom and Ireland. *Genome Res.* 26, 1101–1109. doi: 10.1101/gr.205245.116
- Naas, T., Blot, M., Fitch, W. M., and Arber, W. (1994). Insertion sequence-related genetic variation in resting *Escherichia coli* K-12. *Genetics* 136, 721–730.
- Naas, T., Blot, M., Fitch, W. M., and Arber, W. (1995). Dynamics of IS-related genetic rearrangements in resting *Escherichia coli* K-12. *Mol. Biol. Evol.* 12, 198–207. doi: 10.1093/oxfordjournals.molbev.a040198
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., et al. (2015). Roary: rapid large-scale prokaryote pangenome analysis. *Bioinformatics (Oxf. Engl.)* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421
- Pérez-Losada, M., Arenas, M., and Castro-Nallar, E. (2017). Microbial sequence typing in the genomic era. *Infect. Genet. Evol.* 63, 346–359. doi: 10.1016/j.meegid.2017.09.022
- Raimondi, A., Sisto, F., and Nikaido, H. (2001). Mutation in *Serratia marcescens* AmpC  $\beta$ -lactamase producing high-level resistance to ceftazidime and ceftiprome. *Antimicrob. Agents Chemother* 45, 2331–2339. doi: 10.1128/AAC.45.8.2331-2339.2001
- Redondo-Bravo, L., Gutiérrez-González, E., San Juan-Sanz, I., Fernández-Jiménez, I., Ruiz-Carrasco, G., Gallego-Lombardo, S., et al. (2019). *Serratia marcescens* outbreak in a neonatology unit of a Spanish tertiary hospital: risk factors and

- control measures. *Am. J. Infect. Control.* 47, 271–279. doi: 10.1016/j.ajic.2018.08.026
- Sandner-Miranda, L., Vinuesa, P., Cravioto, A., and Morales-Espinosa, R. (2018). The genomic basis of intrinsic and acquired antibiotic resistance in the genus *Serratia*. *Front. Microbiol.* 9:828. doi: 10.3389/fmicb.2018.00828
- Schürch, A. C., Arreondo-Alonso, S., Willems, R. J. L., and Goering, R. V. (2018). Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches. *Clin. Microbiol. Infect.* 24, 350–354. doi: 10.1016/j.cmi.2017.12.016.350e354
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Shen, J., Zhou, J., Xu, Y., and Xiu, Z. (2019). Prophages contribute to genome plasticity of *Klebsiella pneumoniae* and may involve the chromosomal integration of ARGs in CG258. *Genomics* 112, 998–1010. doi: 10.1016/j.ygeno.2019.06.016
- Thompson, C. C., Vicente, A. C., Souza, R. C., Vasconcelos, A. T., Vesth, T., Alves, N. Jr., et al. (2009). Genomic taxonomy of *Vibriosis*. *BMC Evol. Biol.* 9:258. doi: 10.1186/1471-2148-9-258
- Tran, J. H., and Jacoby, G. A. (2002). Mechanism of plasmid-mediated quinolone resistance. *Proc. Natl. Acad. Sci. U.S.A.* 99, 5638–5642. doi: 10.1073/pnas.082092899
- Vielva, L., de Toro, M., Lanza, V. F., and de la Cruz, F. (2017). PLACNETw: a web-based tool for plasmid reconstruction from bacterial genomes. *Bioinformatics* 33, 3796–3798. doi: 10.1093/bioinformatics/btx462
- Villa, J., Alba, C., Barrado, L., Sanz, F., del Castillo, E. G., Viedma, E., et al. (2012). Long-term evolution of multiple outbreaks of *Serratia marcescens* bacteraemia in a neonatal intensive care unit. *Pediatr. Infect. Dis. J.* 12, 298–300. doi: 10.1097/INF.0b013e318267f441
- Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46
- Xie, Z., and Tang, H. (2017). ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. *Bioinformatics* 33, 3340–3347. doi: 10.1093/bioinformatics/btx433
- Zingg, W., Posfay-Barbe, K. M., and Pittet, D. (2008). Healthcare-associated infections in neonates. *Curr. Opin. Infect. Dis.* 3, 228–234. doi: 10.1097/QCO.0b013e3282fcec5f

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Saralegui, Ponce-Alonso, Pérez-Viso, Moles Alegre, Escribano, Lázaro-Perona, Lanza, de Pipaón, Rodríguez, Baquero and del Campo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN

## Strain-specific predation of *Bdellovibrio bacteriovorus* on *Pseudomonas aeruginosa* with a higher range for cystic fibrosis than for bacteremia isolates

Claudia Saralegui<sup>1,4</sup>, Cristina Herencias<sup>1,4</sup>✉, Ana Verónica Halperin<sup>1</sup>, Juan de Dios-Caballero<sup>1</sup>, Blanca Pérez-Viso<sup>1</sup>, Sergio Salgado<sup>2</sup>, Val F. Lanza<sup>1</sup>, Rafael Cantón<sup>1</sup>, Fernando Baquero<sup>1,3</sup>, M. Auxiliadora Prieto<sup>2</sup> & Rosa del Campo<sup>1</sup>✉

This work aimed to evaluate the predatory activity of *Bdellovibrio bacteriovorus* 109J on clinical isolates of *Pseudomonas aeruginosa* selected from well-characterized collections of cystic fibrosis (CF) lung colonization (n = 30) and bloodstream infections (BSI) (n = 48) including strains selected by genetic lineage (frequent and rare sequence types), antibiotic resistance phenotype (susceptible and multidrug-resistant isolates), and colony phenotype (mucoïd and non-mucoïd isolates). The intraspecies predation range (I-PR) was defined as the proportion of susceptible strains within the entire collection. In contrast, the predation efficiency (PE) is the ratio of viable prey cells remaining after predation compared to the initial inoculum. I-PR was significantly higher for CF (67%) than for BSI *P. aeruginosa* isolates (35%) probably related to an environmental origin of CF strains whereas invasive strains are more adapted to humans. I-PR correlation with bacterial features such as mucoïd morphotype, genetic background, or antibiotic susceptibility profile was not detected. To test the possibility of increasing I-PR of BSI isolates, a polyhydroxyalkanoate depolymerase deficient *B. bacteriovorus* bd2637 mutant was used. Global median I-PR and PE values remained constant for both predators, but 31.2% of 109J-resistant isolates were susceptible to the mutant, and 22.9% of 109J-susceptible isolates showed resistance to predation by the mutant, pointing to a predator-prey specificity process. The potential use of predators in the clinical setting should be based on the determination of the I-PR for each species, and the PE of each particular target strain.

The “golden age of antibiotics” in the mid-twentieth century was followed by the emergence of pathogens resistant to almost all available antibiotics, leading to the current global crisis of multidrug-resistant (MDR) bacteria<sup>1,2</sup>. Identifying alternative interventions to antimicrobials requires a complete understanding of the ecological rules governing the commensal microbiota, which includes natural competitors and predators<sup>3–6</sup>. In nature, predatory bacteria play an important role in maintaining population sizes by linking the production and removal of biomass in microbial communities, which in turn promotes the diversity of microorganisms and contributes to the global stabilization of the ecosystem<sup>7,8</sup>. The ecological role of predators could also be exploited in the fight against clinical pathogens since they are dynamic microorganisms that undergo (as do their opponents) continuous physical, morphological, and metabolic adaptation to the ecosystem. This evolutionary reciprocity

<sup>1</sup>Department of Microbiology, Hospital Universitario Ramón Y Cajal, Instituto Ramón Y Cajal de Investigación Sanitaria (IRYCIS), CIBERINFEC, Madrid, Spain. <sup>2</sup>Microbial and Plant Biotechnology Department, Biological Research Center-Margarita Salas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain. <sup>3</sup>CIBERESP, Madrid, Spain. <sup>4</sup>These authors contributed equally: Claudia Saralegui and Cristina Herencias. ✉email: cheroedr@gmail.com; rosacampo@yahoo.com

is the basis of coevolution, in which the adaptation of a player not only promotes change in its opponent, but the opponent's adaptation also generates selection as an evolutionary response to the first player<sup>9</sup>.

The successful development of predatory bacteria as “living antimicrobial agents” and a complete understanding of the predation mechanism depend on the characterization of their predation preferences, mainly their predation range and efficiency among different species. However, within a bacterial species, predation appears to be species-specific but not universal for all its lineages, which can be referred to as intraspecific predation range (I-PR). Again, within each strain or clone, the predator may have different predation efficient (PE), depending on the proportion of predated cells. This polymorphic predation behavior is fundamentally based on the composition of the species or individual prey cell envelope, eventually influenced by environmental conditions<sup>10–12</sup>. The identification of the factors driving prey preference and differences in predatory activity between and within species has so far remained limited, particularly in non-environmental bacterial collections<sup>13, 14</sup>.

The most studied bacterial predators are *Bdellovibrio* and -like organisms (BALOs), which are small vibrioid gram-negative aerobic bacteria, recently reclassified to the class of *Oligoflexia*, which belongs to the *Pseudomonadota* phylum<sup>15</sup>. Although BALOs were first isolated from soil, they are ubiquitous in nature and can be found in aquatic and terrestrial environments, including hypersaline systems<sup>16</sup>, biofilms<sup>17</sup>, mammalian intestines<sup>18–20</sup>, and, as we have shown, the lungs of cystic fibrosis (CF) patients<sup>21</sup>. In addition to the genetic detection of BALOs, several authors have documented the in vivo phenomenon of predation in human microbial ecosystems<sup>22–25</sup>.

*Pseudomonas aeruginosa* is the most studied prey of BALOs, and both predator and prey frequently share a common environmental origin. This genus is also one of the most significant human pathogens, included in the ESKAPE category (including *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.) by its high prevalence in severe invasive infections, such as bacteremia, and their ability to acquire antibiotic multiresistance<sup>26</sup>. *P. aeruginosa* is also the most relevant and abundant microorganism involved in CF pathogenic lung colonization; particularly the so-called high-risk clones with multiresistance to antibiotics and those with mucoid morphotype due to alginate hyperproduction. The genetic and molecular characterization of *P. aeruginosa* isolates from CF<sup>27</sup> and bacteremia<sup>28</sup> origins have been previously addressed by our group, showing that CF isolates are more related to an environmental origin<sup>29</sup>.

Herein, we determined the I-PR of the reference strain *B. bacteriovorus* 109J against a comprehensive collection of 78 *P. aeruginosa* clinical isolates, of diverse patient origins, genetic backgrounds, colonial, and antibiotic-resistant phenotypes. We also explore the possibility of increasing I-PR on the BSI collection by modifying the predator genome, presenting as proof of concept a single-gene *B. bacteriovorus* mutant<sup>30</sup>, the strain bd2637 which harbors a deficient polyhydroxyalkanoate depolymerase (PHA) which was able to prey on some 109J -resistant strains, but in contrast, other 109J -sensitive strains showed resistance to predation by the mutant. We discuss possible mechanisms that might explain this effect, opening the door to genetic engineering of predators to overcome reduced predator susceptibility in *P. aeruginosa*.

## Results

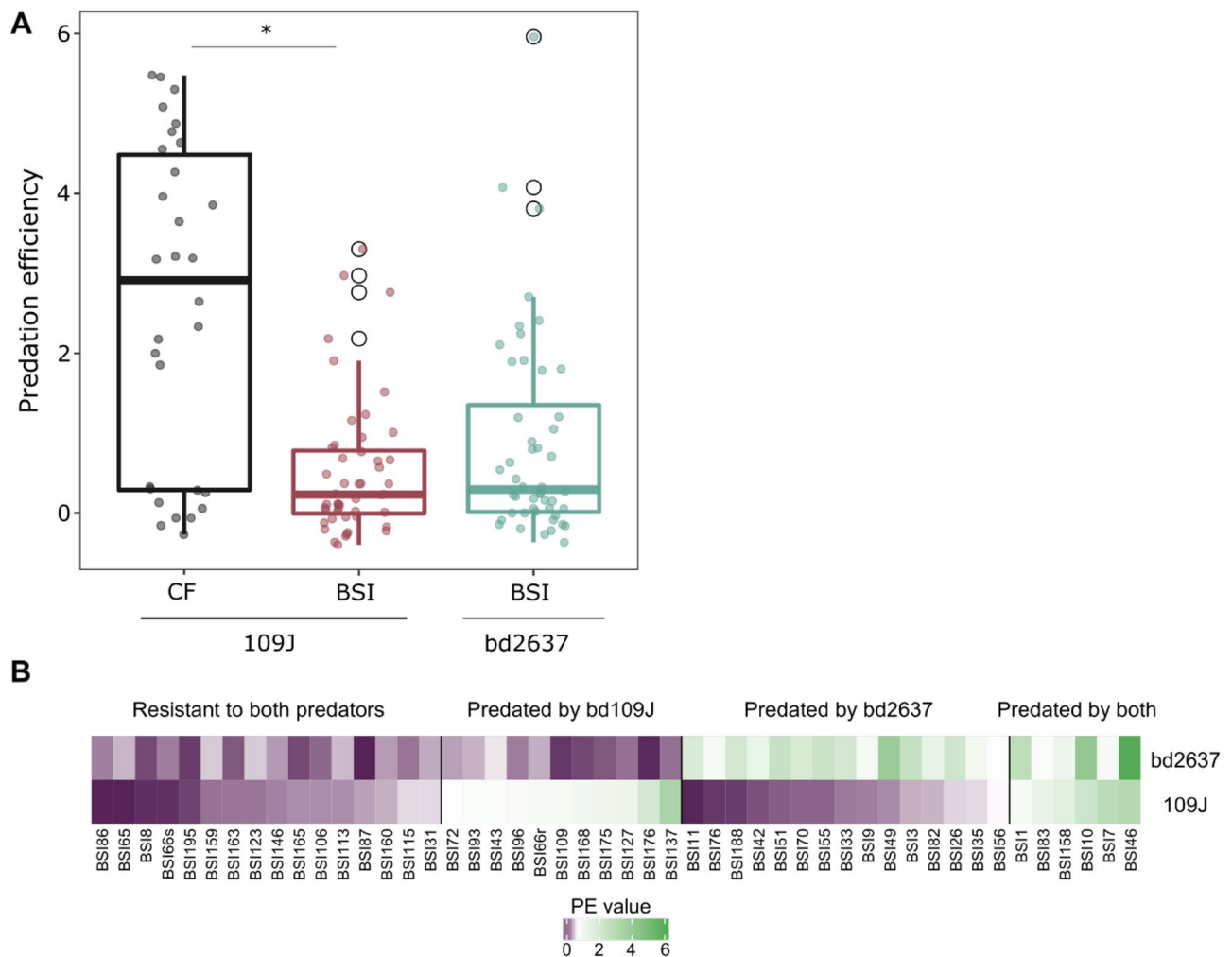
**Prey's clinical source influences predation activity by *B. bacteriovorus* 109J.** A significantly higher I-PR value was observed in CF (20/30, 67.0%, Fig. S1) than in BSI isolates (17/48, 35.4%,  $p = 0.02$ , Fig. S2). In addition, median PE was higher for CF (median PE 3.91) than for BSI strains (median PE 1.10) (Mann–Whitney U test  $p < 0.0001$ ) (Fig. 1A, Tables S1, and S2). The association of I-PR and the origin of the collection was also performed using linear model (lm,  $p < 0.0001$ ). Correlation between predation activity and *P. aeruginosa* genetic lineage (Sequence Type; ST, obtained by MLST) was not observed ( $p > 0.05$ ), with discrepancies in 6 of the 12 STs grouping more than one isolate from both collections [ST175 (1 susceptible prey out of a total of 3), ST253 (2/3), ST274 (1/2), ST532 (1/2), ST646 (2/3), and ST1017 (1/2)] (Tables S1 and S2). There was also no correlation with the mucoid (4/7 mucoid isolates) or non-mucoid phenotype of CF isolates (16/23 mucoid isolates) ( $p > 0.05$ ) or with the antibiotic susceptibility phenotype ( $p > 0.05$ ) (Fig. S3A and Table S2).

**Influence of the predator genetic variation.** The predation ability of the *B. bacteriovorus* 109J wildtype strain was compared to those of the single-gene mutant bd2637 in the 48 BSI isolates, the less susceptible collection. There were no significant differences between I-PR values of mutant and wildtype strains (21/48, 43.8% vs. 17/48, 35.4%,  $p > 0.05$ ); nor in the global median PE values (1.10 for *B. bacteriovorus* 109J vs. 1.8 for bd2637 mutant,  $p > 0.05$ ) (Fig. 1A, Fig. S4, and Table S4). Interestingly, the range between both predators within the BSI collection differed: 16 isolates (33.3%) were resistant and 6 (12.5%) susceptible to both predators, whereas 15 (31.2%) were susceptible only to the bd2637 mutant, and finally 11 isolates (22.9%) were susceptible only to the wildtype predator (Fig. 1B).

There was no correlation between predation of bd2637 mutant and antibiotic susceptibility or any other prey feature ( $p > 0.05$ ) (Fig. S3B and Table S3). These results were consistently observed in the replicates of each isolate and were not associated with any detected particular prey characteristic. Thus, only the deletion of the catalytic activity from the *bd2637* gene and the associated effects were responsible for the changes in prey susceptibility.

## Discussion

The use of BALOs as biological control agents in environmental and medical microbiological settings<sup>18, 31</sup> has been suggested based on their lack of interaction with human cells<sup>32–34</sup>. As occurs with antibiotics, testing the individual in vitro susceptibility for prey and predator pairs of strains is a requirement, mainly when predation could be substantially affected by environmental or biological conditions, as we postulate herein. Predators have been studied primarily under controlled conditions (medium, pH, and temperature), and the knowledge on predation susceptibility of clinical isolates remains very limited. We focused our research on *P. aeruginosa* as this species was reported to have a limited ability to serve as a prey for *B. bacteriovorus*<sup>35</sup>, and therefore we



**Figure 1.** Predation susceptibility of *P. aeruginosa* collections. **(A)** Predation efficiency (PE) of *B. bacteriovorus* 109J on CF and BSI collection (\*represents significance  $p < 0.0001$ ) and PE of mutant strain bd2637 upon BSI collection ( $p > 0.05$ ). **(B)** PE profile of the wildtype *B. bacteriovorus* 109J and mutant strain bd2637 upon BSI collection. Positive predation was considered when  $PE > 0.5$ . Each strain has a specific predation profile. White circles in BSI collection represent outliers.

expected a higher diversity in predator–prey response. In addition, *P. aeruginosa*, ubiquitous in nature (where most *Bdellovibrio-Pseudomonas* interactions are expected to occur), is a significant human pathogen frequently harboring mechanisms of antibiotic resistance. Also, previous studies of our group facilitated the building-up of a comprehensive and well-characterized collections including both frequent and infrequent lineages from CF<sup>27</sup> and BSI<sup>28</sup>, accordingly to their ST. Note that the same *P. aeruginosa* ST can be found in different habitats, such as environmental and nosocomial niches (human microbiota of patients, built environments), but some lineages seem to be highly adapted to particular clinical environments. This is the case of chronic colonization of CF-patients airways, where the associated lineages are close to strains of environmental origin, as we have previously shown<sup>29</sup>. We did not find a correlation between ST (sequence types) and I-PR.

A notable result of our study is the significantly higher qualitative (I-PR, proportion of predated strains in the collections) and quantitative (PE, proportion of cells predated per individual strain) susceptibility of CF isolates in comparison with BSI ones. In addition, there was no correlation between predation and the colonial morphology, mucoid vs. non-mucoid, according to previous reports<sup>36</sup>, or antibiotic susceptibility patterns, or even genetic ST lineages. We cannot exclude that predation differences in the same ST lineage might derive from other genetic features such as the presence of pathogenicity islands or prophages.

The existence of highly specialized predators in human microbiota cannot be ruled out since all available data have been obtained from environmental predators. Human microbial ecosystems, however, probably have the same rules of population control based on predation. Predators with a human origin could be more suitable for limiting the well-adapted human pathogens<sup>21</sup>, despite that in a complex and diverse ecosystem, the preference for particular prey would be a dynamic feature<sup>37</sup>. Although experiments are often conducted using individual lineages, the use of mixed populations should be a future goal to validate prey specificity in a community and the consequences on microbial population structure. This work indicates that quantitative predation, determined as

relative PEs might differ among members of an (apparently) homogeneous population and could be critical to understand the dynamics of bacterial ensembles composed by different prey and predators.

Only 8 *B. bacteriovorus* genomes have been entered into public databases, all of them from an environmental origin (soil, rivers, seawater, wastewater), but the contribution of the predators' genetic background in the PE is a pending issue. To elucidate that, we used the mutant bd2637, to test the possibility of engineering more efficient, wider I-PR and PE predators<sup>30</sup> on BSI isolates. We show that this mutant has a different spectrum than the wildtype strain, but how might that occur? The genotype of the bd2637 strain corresponds to a single deletion of the *bd2637* gene, which encodes a putative PHA depolymerase responsible for the degradation of biopolymers. Hypothetically, this mutant should have less access to the rich nutritional resources of polyhydroxyalkanoate (PHA) and consequently less fitness. However, the reality is probably more complex. Inside the "common metabolism" when the predator is placed into the prey, massive PHA degradation yields an ATP excess, which is known to increase reactive oxygen species, which can oxidize and damage DNA, carbohydrates, proteins, or lipids<sup>38,39</sup>. Excess ATP also triggers respiration, but in small cells (such as *Bdellovibrio*) the question is whether high ATP levels are incompatible with the maintenance of respiration if the surface of the cell cannot allocate enough respiratory proteins<sup>40,41</sup>. In short, PHA defective mutants might have unexpected metabolic consequences<sup>42–46</sup> influencing predation activity, and research in this field could help to identify and eventually overcome predation resistance factors in *P. aeruginosa*. It is worth noting that the prey range (intra- or interspecies range) does not depend only on the prey susceptibility but also on the predator specificity, which highlights the importance of specific predator–prey susceptibility testing, as their interaction and co-evolution may overcome predation and resistance, respectively<sup>47,48</sup>. A promising perspective for future work is trying to see if predation host-specificity and predation rate might differ when predators are propagated in bacterial species/clones isolated from closely related or different niches. The evolution of the predator in coevolving bacterial communities might result in an accelerated genomic predator and host evolution<sup>49</sup>.

The prey or predator determinants responsible for predation specificity have not been completely elucidated. However, the hydrolytic arsenal of *B. bacteriovorus* plays a crucial role in predation activity, given that it determines the success of their lifecycle<sup>30,50</sup>. Characterization of each predator's specific prey spectrum is a requirement for the clinical use of predators as "living antimicrobial agents". Studies have reported that different BALOs lineages and predators isolated from different niches have different prey spectra<sup>51,52</sup>. Thus, the selection of the appropriate predator for clinical purposes would require larger and more in-depth studies on predation to overcome predation resistance in specific prey species and individual strains within the species.

## Materials and methods

**Strains and growth conditions.** The *B. bacteriovorus* 109J wildtype strain was used as predator against 78 *P. aeruginosa* strains selected from previously well-characterized collections obtained from respiratory samples of CF patients (n = 30)<sup>27</sup>, and invasive BSI isolates (n = 48)<sup>28</sup>. All 78 *P. aeruginosa* isolates were collected and characterized in earlier studies, for which the appropriate ethical approval and consent were obtained<sup>27,28</sup>. The selection of prey was based on representative sequence types (STs) determined by MLST, both frequent and rare STs, their antibiotic susceptibility phenotype, including multidrug-resistant isolates, as well as their morphology in CF isolates, differentiating mucoid (n = 7) and non-mucoid isolates (n = 23) (Tables S1 and S2). The methodologies used for the molecular characterization (ST, antibiotic resistance profile, morphotype) of *P. aeruginosa* collections were detailed in the previous publication<sup>27,28</sup>.

The predation activity of the PHA-depolymerase mutant *B. bacteriovorus* bd2637<sup>30</sup> was only tested against the 48 BSI isolates. Mutations were manually inspected with Artemis<sup>53</sup> using the Genbank accession numbers: SRX10641169 and SRX10641170 for *B. bacteriovorus* 109J and the bd2637 mutant, respectively (Table S5).

Predators were routinely grown as previously described<sup>54,55</sup> by co-cultivation with *Pseudomonas putida* KT2440 instead of *P. aeruginosa* to prevent the bias that might result from possible adaptive influence on the growth of a particular ST type of *P. aeruginosa*. Diluted Nutrient Broth (DNB) (0.8 g/l nutrient broth at pH 7.4) was used to recover the predator cells from glycerol stocks and HEPES buffer (25 mM at pH 7.8) was used to perform the predation experiments. Both DNB and HEPES buffer were supplemented with 2 mM CaCl<sub>2</sub>·2H<sub>2</sub>O and 3 mM MgCl<sub>2</sub>·3H<sub>2</sub>O. To purify the predator, co-cultures were filtered twice through a 0.45- $\mu$ m filter. Finally, *P. aeruginosa* strains were cultivated on Luria Broth (LB) for 16 h at 37 °C and were further diluted in HEPES buffer to an optical density at 600 nm of 1 (OD<sub>600</sub> 1) for the subsequent predation experiments.

**Predation experiments.** Prey (100  $\mu$ l to final OD<sub>600</sub> 0.3) and predator (100  $\mu$ l) co-cultures were performed on 96-microwell plates for 24 h at 30 °C with orbital shaking in a Synergy HTX (BioTek) with cell density monitoring (OD<sub>600</sub>) each 15 min. Each prey was tested in at least 2–3 independent biological replicates, with the results expressed as mean values of these replicates for all experiments. Growth control wells without the predator were also included for each *P. aeruginosa* isolate.

To assure the homogeneity of the predator inoculum, each experiment had a parallel control co-culture of *B. bacteriovorus* (100  $\mu$ l) and *P. putida* (100  $\mu$ l of 10<sup>9</sup> CFU/ml); and after incubation, prey survival was determined by seeding serial dilutions in agar plates, discarding those sets in which the viable prey decrease upon the control *P. putida* predation was different than 3 log<sub>10</sub> and 4 log<sub>10</sub> for wildtype and bd2637 predators, respectively, as previously described<sup>30,39</sup>.

Thus, predation experiments included control at the beginning of the experiment of both predator and prey inoculum (viable counts corresponding to OD<sub>600</sub> = 0.3). Viable cell count of the prey was performed at 0 h and 24 h of predation. 10  $\mu$ l of the corresponding dilution of co-culture (including the control without predator and the predation control with *P. putida*) was plated on LB solid medium and colony-forming units (CFU/ml) were counted after 24 h of incubation at 37 °C.

The predation activity of a predator upon a prey collection is defined by two parameters: (1) the intraspecies predation range (I-PR), which corresponds to the proportion of susceptible prey strains within a species collection, and (2) the predation efficiency (PE), specific for each prey/predator combination, which is determined as the ratio between viable prey cells after predator exposure compared to the initial inoculum. PE was expressed in log<sub>10</sub> values. Positive predation was only considered when PE was > 0.5, corresponding with the predation of 50% of the population ( $p < 0.046$ , Table S1).

**Statistical analysis.** The Mann–Whitney U test was used to compare PE values between CF and BSI collections, after assuming non-normality with the Shapiro–Wilk normality test. Additionally, associations between qualitative features of *Pseudomonas* collections and the quantitative variable I-PR were performed using simple linear model. The differences in I-PR between the CF and BSI collections were explored using the chi-squared test. All data sets were analyzed using R version 3.5.0, and plotting was performed using ggplot version 2.2.1. and RStudio software v.1.2.5001 (R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>).

Received: 22 February 2022; Accepted: 6 June 2022

Published online: 22 June 2022

## References

1. Baquero, F. Threats of antibiotic resistance: An obliged reappraisal. *Int. Microbiol.* **24**(4), 499–506 (2021).
2. Cavallo, F. M., Jordana, L., Friedrich, A. W., Glasner, C. & van Dijk, J. M. *Bdellovibrio bacteriovorus*: A potential 'living antibiotic' to control bacterial pathogens. *Crit. Rev. Microbiol.* **47**(5), 630–646 (2021).
3. Celis, A. I. & Relman, D. A. Competitors versus collaborators: Micronutrient processing by pathogenic and commensal human-associated gut bacteria. *Mol. Cell* **78**, 570–576 (2020).
4. Legein, M. *et al.* Modes of action of microbial biocontrol in the phyllosphere. *Front. Microbiol.* **11**, 1619 (2020).
5. Madhav, M., Baker, D., Morgan, J. A., Asgari, S. & James, P. *Wolbachia*: A tool for livestock ectoparasite control. *Vet. Parasitol.* **288**, 109297 (2020).
6. Shen, Y. & Loessner, M. J. Beyond antibacterials—exploring bacteriophages as antivirulence agents. *Curr. Opin. Biotechnol.* **68**, 166–173 (2021).
7. Fuhrman, J. A. & Caron, D. A. Heterotrophic planktonic microbes: Virus, bacteria, archaea, and protozoa. <https://doi.org/10.1128/9781555818821.ch4.2.2> (2016).
8. Thompson, J. N. The evolution of species interactions. *Science (Washingt. D C)* **284**, 2116–2118 (1999).
9. Gallet, R., Tully, T. & Evans, M. E. K. Ecological conditions affect evolutionary trajectory in a predator–prey system. *Evolution (N. Y.)* **63**, 641–651 (2009).
10. Lambert, C. *et al.* Interrupting peptidoglycan deacetylation during *Bdellovibrio* predator–prey interaction prevents ultimate destruction of prey wall, liberating bacterial-ghosts. *Sci. Rep.* **6**, 26010 (2016).
11. Jurkevitch, E. & Davidov, Y. Phylogenetic diversity and evolution of predatory prokaryotes. In *ACS Division of Fuel Chemistry, Preprints* (2006). <https://doi.org/10.1007/7171>.
12. Pernthaler, J. Predation on prokaryotes in the water column and its ecological implications. *Nat. Rev. Microbiol.* **3**, 537–546 (2005).
13. Li, N. *et al.* Analysis of gene gain and loss in the evolution of predatory bacteria. *Gene* **598**, 63–70 (2017).
14. Duncan, M. C. *et al.* High-throughput analysis of gene function in the bacterial predator *Bdellovibrio bacteriovorus*. *MBio* **10**, e01040-19 (2019).
15. Waite, D. W. *et al.* Proposal to reclassify the proteobacterial classes *Deltaproteobacteria* and *Oligoflexia*, and the phylum *Thermodesulfobacteria* into four phyla reflecting major functional capabilities. *Int. J. Syst. Evol. Microbiol.* **70**, 5972–6016 (2020).
16. Piñero, S. A., Williams, H. N. & Stine, O. C. Phylogenetic relationships amongst the saltwater members of the genus *Bacteriovorax* using *rpoB* sequences and reclassification of *Bacteriovorax stolpii* as *Bacteriolyticum stolpii* gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* **58**, 1203–1209 (2008).
17. Kadouri, D. & O'Toole, G. A. Susceptibility of biofilms to *Bdellovibrio bacteriovorus* attack. *Appl. Environ. Microbiol.* **71**, 4044–4051 (2005).
18. Hobbey, L. *et al.* Genome analysis of a simultaneously predatory and prey-independent, novel *Bdellovibrio bacteriovorus* from the River Tiber, supports in silico predictions of both ancient and recent lateral gene transfer from diverse bacteria. *BMC Genom.* **13**, 670 (2012).
19. Schwudke, D., Strauch, E., Krueger, M. & Appel, B. Taxonomic studies of predatory *Bdellovibrios* based on 16S rRNA analysis, ribotyping and the *hit* locus and characterization of isolates from the gut of animals. *Syst. Appl. Microbiol.* **24**, 385–394 (2001).
20. Iebba, V. *et al.* Higher prevalence and abundance of *Bdellovibrio bacteriovorus* in the human gut of healthy subjects. *PLoS ONE* **8**, e61608 (2013).
21. Caballero, J. D. D. *et al.* Individual patterns of complexity in including predator bacteria, over a 1-year period. *mBio* **8**, e00959-17 (2017).
22. Shatzkes, K. *et al.* Predatory bacteria attenuate *Klebsiella pneumoniae* burden in rat lungs. *MBio* **7**, 1–9 (2016).
23. Shanks, R. M. Q. *et al.* An eye to a kill: Using predatory bacteria to control gram-negative pathogens associated with ocular infections. *PLoS ONE* **8**, e66723 (2013).
24. Romanowski, E. G. *et al.* Predatory bacteria are nontoxic to the rabbit ocular surface. *Sci. Rep.* **6**, 30987 (2016).
25. Silva, P. H. F. *et al.* The impact of predatory bacteria on experimental periodontitis. *J. Periodontol.* <https://doi.org/10.1002/JPER.18-0485> (2019).
26. De Oliveira, D. M. P. *et al.* Antimicrobial resistance in ESKAPE pathogens. *Clin. Microbiol. Rev.* **33**, 1–49 (2020).
27. López-Causapé, C. *et al.* Antibiotic resistance and population structure of cystic fibrosis *Pseudomonas aeruginosa* isolates from a Spanish multi-centre study. *Int. J. Antimicrob. Agents* **50**, 334–341 (2017).
28. García-Castillo, M. *et al.* Wide dispersion of ST175 clone despite high genetic diversity of carbapenem-nonsusceptible *Pseudomonas aeruginosa* clinical strains in 16 Spanish hospitals. *J. Clin. Microbiol.* **49**, 2905–2910 (2011).
29. Fernández-Olmos, A. *et al.* Population structure and antimicrobial susceptibility of both nonpersistent and persistent *Pseudomonas aeruginosa* isolates recovered from cystic fibrosis patients. *J. Clin. Microbiol.* **51**, 2761–2765 (2013).
30. Martínez, V., Herencias, C., Jurkevitch, E. & Prieto, M. A. Engineering a predatory bacterium as a proficient killer agent for intracellular bio-products recovery: The case of the polyhydroxyalkanoates. *Sci. Rep.* **6**, 24381 (2016).
31. Westergaard, J. M. & Kramer, T. T. *Bdellovibrio* and the intestinal flora of vertebrates. *Appl. Environ. Microbiol.* **34**, 506–511 (1977).

32. Bonfiglio, G. *et al.* Insight into the possible use of the predator *Bdellovibrio bacteriovorus* as a probiotic. *Nutrients* **12**, 2252 (2020).
33. Gupta, S., Tang, C., Tran, M. & Kadouri, D. E. Effect of predatory bacteria on human cell lines. *PLoS ONE* **11**, 1–15 (2016).
34. Monnappa, A. K., Bari, W., Choi, S. Y. & Mitchell, R. J. Investigating the responses of human epithelial cells to predatory bacteria. *Sci. Rep.* **6**, 1–14 (2016).
35. Markelova, N. Y. Predacious bacteria, *Bdellovibrio* with potential for biocontrol. *Int. J. Hyg. Environ. Health* **213**, 428–431 (2010).
36. Koval, S. F. & Bayer, M. E. Bacterial capsules: No barrier against *Bdellovibrio*. *Microbiology* **143**, 749–753 (1997).
37. Rogosky, A. M., Moak, P. L. & Emmert, E. A. B. Differential predation by *Bdellovibrio bacteriovorus* 109J. *Curr. Microbiol.* **52**, 81–85 (2006).
38. Akhova, A. V. & Tkachenko, A. G. ATP/ADP alteration as a sign of the oxidative stress development in *Escherichia coli* cells under antibiotic treatment. *FEMS Microbiol. Lett.* **353**, 69–76 (2014).
39. Martínez, V., Jurkevitch, E., García, J. L. & Prieto, M. A. Reward for *Bdellovibrio bacteriovorus* for preying on a polyhydroxyalkanoate producer. *Environ. Microbiol.* <https://doi.org/10.1111/1462-2920.12047> (2013).
40. Szenk, M., Dill, K. A. & de Graff, A. M. R. Why do fast-growing bacteria enter overflow metabolism? testing the membrane real estate hypothesis. *Cell Syst.* **5**, 95–104 (2017).
41. Lobritz, M. A. *et al.* Increased energy demand from anabolic-catabolic processes drives  $\beta$ -lactam antibiotic lethality. *Cell Chem. Biol.* <https://doi.org/10.1016/j.chembiol.2021.12.010> (2022).
42. Sutton, D., Livingstone, P. G., Furness, E., Swain, M. T. & Whitworth, D. E. Genome-wide identification of myxobacterial predation genes and demonstration of formaldehyde secretion as a potentially predation-resistant trait of *Pseudomonas aeruginosa*. *Front. Microbiol.* **10**, 1–9 (2019).
43. Knoll, M., Hamm, T. M., Wagner, F., Martínez, V. & Pleiss, J. The PHA depolymerase engineering database: A systematic analysis tool for the diverse family of polyhydroxyalkanoate (PHA) depolymerases. *BMC Bioinform.* **10**, 89 (2009).
44. Dori-Bachash, M., Dassa, B., Pietrokovski, S. & Jurkevitch, E. Proteome-based comparative analyses of growth stages reveal new cell cycle-dependent functions in the predatory bacterium *Bdellovibrio bacteriovorus*. *Appl. Environ. Microbiol.* **74**, 7152–7162 (2008).
45. Potter, S. C. *et al.* HMMER web server: 2018 update. *Web Serv. issue Publ. online* **46** (2018).
46. Manoli, M.-T., Nogales, J. & Prieto, A. Synthetic control of metabolic states in *Pseudomonas putida* by tuning polyhydroxyalkanoate Cycle. *MBio* <https://doi.org/10.1128/mbio.01794-21> (2022).
47. Sydney, N. *et al.* The genetics of prey susceptibility to myxobacterial predation: A review, including an investigation into *Pseudomonas aeruginosa* mutations affecting predation by *Myxococcus xanthus*. *Microb. Physiol.* <https://doi.org/10.1159/000515546> (2021).
48. Mitchell, R. J., Mun, W., Mabekou, S. S., Jang, H. & Choi, S. Y. Compounds affecting predation by and viability of predatory bacteria. *Appl. Microbiol. Biotechnol.* **104**, 3705–3713 (2020).
49. Nair, R. R. *et al.* Bacterial predator-prey coevolution accelerates genome evolution and selects on virulence-associated prey defences. *Nat. Commun.* **10**, 4301 (2019).
50. Rendulic, S. *et al.* A predator unmasked: Life cycle of *Bdellovibrio bacteriovorus* from a genomic perspective. *Science* **303**, 689–692 (2004).
51. Jurkevitch, E., Minz, D., Ramati, B. & Barel, G. Prey range characterization, ribotyping, and diversity of soil and rhizosphere *Bdellovibrio* spp. isolated on phytopathogenic bacteria. *Appl. Environ. Microbiol.* **66**, 2365–2371 (2000).
52. Starr, M. P. & Seidler, R. J. The Bdellovibrios. *Annu. Rev. Microbiol.* **25**, 649–678 (1971).
53. Rutherford, K. *et al.* Artemis: Sequence visualization and annotation. *Bioinformatics* **16**, 944–945 (2000).
54. Herencias, C., Prieto, M. A. & Martínez, V. Determination of the predatory capability of *Bdellovibrio bacteriovorus* HD100. *Bio Protoc.* **7**, 2–10 (2017).
55. Lambert, C. & Sockett, R. E. Laboratory maintenance of *Bdellovibrio*. *Curr. Protoc. Microbiol.* <https://doi.org/10.1002/9780471729259.mc07b02s9> (2008).

## Acknowledgements

We appreciate the technical support of Natalia Huertas. CH is supported by Comunidad Autónoma de Madrid (PEJD-2018-POST/BMD-8016). CS is granted by “Fundación Mutua Madrileña” achieved in 2017 call by RDC (AP165902017). SSB is a recipient of a predoctoral FPU grant (FPU17/03978) from the Spanish Ministry of Universities. This work was supported by the Instituto de Salud Carlos III, PI17/00115 and PI20/00164 to RdC, REIPI (RD16/0016/0011) actions, co-financed by the European Development Regional Fund “A way to achieve Europe” (ERDF), and Vertex Pharmaceuticals. We acknowledge the opinion of James J. Collins, from the Department of Biological Engineering in the Massachusetts Institute of Technology, Wyss Institute, Harvard University, concerning the ATP excess toxicity.

## Author contributions

C.H., A.P. and R.dC. conceived the presented idea. C.S. and C.H. carried out experiments, data curation, formal analysis, and validation. A.H., J.dD., B.P. and S.S. provided experimental resources (strains). V.F.L. performed sequences analysis. C.H. and R.dC. supervised the project. C.S., C.H., and R.dC. wrote, reviewed, and edited the manuscript. C.H., A.P., R.dC., R.C. and F.B. authors discussed and contributed to the final version of the manuscript. A.P. and R.dC. obtained funds for the project.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-14378-5>.

**Correspondence** and requests for materials should be addressed to C.H. or R.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022