

## TRABAJO DE FIN DE MÁSTER

*Máster Universitario Biotecnología Clínica*



# **ANÁLISIS DEL MICROBIOMA INTESTINAL EN PACIENTES PEDIÁTRICOS MEDIANTE SECUENCIACIÓN DE NUEVA GENERACIÓN (NGS) Y SU POSIBLE RELACIÓN CON LA CARGA DE GENES DE RESISTENCIA A CARBAPENÉMICOS.**

**Autor:** Clara Fay Guida

**Tutor:** Elías Dahdouh

## Índice

<b>1. Resumen.</b> .....	2
<b>2. Introducción.</b> .....	3
<b>3. Objetivos.</b> .....	8
<b>4. Metodología.</b> .....	8
<b>4.1. Selección de las muestras y extracción del ADN.</b> .....	8
<b>4.2. Reacción de qPCR.</b> .....	9
<b>4.3. Reacción de PCR punto final.</b> .....	11
<b>4.4. Cuantificación con Qubit.</b> .....	12
<b>4.5. Secuenciación NGS.</b> .....	13
4.5.1. Reparación y fosforilación de extremos. ....	14
4.5.2. Ligación de los bardcodes.....	15
4.5.3. Ligación de los adaptadores y limpieza. ....	15
4.5.4. Preparado e inyección de la muestra en la celda de flujo. ....	16
<b>4.6. Análisis bioinformático.</b> .....	17
<b>4.7. Análisis por Excel.</b> .....	19
<b>4.8. Análisis estadístico.</b> .....	21
<b>5. Resultados.</b> .....	22
<b>5.1. Selección de muestras y extracción del ADN.</b> .....	22
<b>5.2. Reacción de qPCR.</b> .....	22
<b>5.3. Reacción de PCR punto final.</b> .....	23
<b>5.4. Secuenciación NGS.</b> .....	23
<b>5.5. Análisis bioinformático.</b> .....	23
<b>5.6. Análisis en Excel y análisis estadístico.</b> .....	24
<b>6. Discusión.</b> .....	27
<b>7. Conclusiones.</b> .....	29
<b>8. Bibliografía.</b> .....	30
<b>9. Anexo.</b> .....	35

## 1. Resumen.

The human gut microbiome plays a crucial role in pediatric development and immune modulation. Factors such as birth mode, feeding and antibiotic exposure have significant impact on the development and establishment of the microbiome. This study aimed to analyze the diversity of the intestinal microbiome in pediatric patients and evaluate its potential association with the load of carbapenem resistance genes. Rectal swab samples (n=46) from patients under 18 years of age who received healthcare at Hospital Universitario La Paz (HULP, Madrid, Spain) were analyzed. The results of quantitative PCRs were used to determine the relative abundance of resistance genes (*blaKPC*, *blaNDM*, *blaOXA-48* and *blaVIM*) normalized to *16S rRNA* gene levels. Samples were categorized into high, medium and negative intestinal relative load groups based on qPCR results. Following sample selection and classification, DNA was extracted and a full-length *16S rRNA* gene was amplified by end point PCR to enable subsequent library preparation for sequencing using Oxford Nanopore Technology.

The bioinformatic analysis was performed using Galaxy, including quality trimming, taxonomic classification with Kraken2 and visualization of the microbiome composition using Krona. Additionally, the Shannon index was calculated for diversity assessment. Although visual inspection of microbial composition suggested potential differences in diversity, statistical analyses (Shapiro-Wilk, Levene, and Student's t-test) revealed no significant differences in Shannon diversity indices across groups. In most of the samples, *Escherichia coli* was dominant.

The results of this study suggest no direct link between carbapenemase gene loads and microbial diversity. The dominance of *E. coli* raised the possibility of unquantified resistance genes that could be leading to it dominating the intestinal microbiome. However, this will require confirmation in future studies. This study also highlights the value of long read sequencing in microbiome research, the need for a broader set of samples to evaluate the relationship with resistance genes, and patient groups in future investigations.

**Keywords:** pediatric patients, gut microbiome, *16S rRNA*, nanopore sequencing, carbapenemases, antibiotic resistance, microbial diversity and Shannon index.

## **2. Introducción.**

El microbioma humano es una compleja comunidad de microorganismos, entre los cuales hay bacterias, virus, hongos y arqueas. Estas comunidades colonizan diversas partes dentro del cuerpo, como son el intestino, la boca, la piel y el tracto respiratorio. A su vez, desempeñan un papel fundamental en la fisiología humana ya que influyen significativamente en el desarrollo del sistema inmune, nervioso y el crecimiento del individuo (Gaufin, 2018).

Desde el momento del nacimiento, el desarrollo de un microbioma es un proceso dinámico influenciado por múltiples factores. El tipo de parto cumple un papel determinante, ya que los nacidos por vía vaginal adquieren una microbiota similar a la flora vaginal materna, compuesta por *Lactobacillus*, *Prevotella*, *Atopobium* y *Sneathia*. Mientras que los nacidos por cesárea presentan una microbiota más parecida a la cutánea materna, generalmente menos diverso (Dominguez-Bello, 2010). Estas diferencias pueden persistir hasta el primer año de vida y se han asociado con un mayor riesgo de enfermedades alérgicas como enfermedades atópicas (Negele, 2004). La lactancia materna también influye en la composición del microbioma, debido a que aporta microorganismos beneficiosos que contribuyen a la colonización intestinal, diferenciándose del microbioma de los alimentados con fórmulas comerciales. Además, la salud de la madre, como el consumo de antibióticos durante el embarazo, y factores ambientales como la dieta, la presencia de animales en el hogar y el entorno general, también afectan el establecimiento del microbioma (Bokulich, 2016).

En este contexto, investigaciones emergentes se centran en el estudio del microbioma en pacientes pediátricos y esto ha puesto foco en el desarrollo de un microbioma diverso ya que alteraciones en la composición durante las primeras etapas de la vida se pueden asociar al desarrollo de enfermedades crónicas como pueden ser el asma o enfermedades inflamatorias intestinales (Palm, 2014) (Abdel-Aziz, 2025). Además de su papel fundamental en la modulación del sistema inmune, se ha demostrado que el microbioma confiere cierta protección frente a patógenos, puede modular la respuesta a las vacunas y participar en el metabolismo de algunos fármacos. Por lo tanto, entender la composición, la dinámica y el establecimiento del microbioma se ha consagrado como un área de vasta investigación (Abdel-Aziz, 2025) (Bäckhed, 2015).

La relación entre el consumo de antibióticos y el microbioma es también de mucha importancia, aún más en el caso de pacientes pediátricos. Los antibióticos de amplio espectro son de gran

utilidad para tratar infecciones de bacterias patógenas, pero al actuar también van a eliminar otro tipo de bacterias beneficiosas para la diversidad y composición del microbioma (Yassour, 2016). Este tipo de función puede devenir en una disbiosis y por tanto su abuso en una etapa clave de la infancia puede tener consecuencias a largo plazo como problemas metabólicos y un mayor riesgo a desarrollar diabetes tipo II, asma, enfermedades alérgicas y obesidad (Yassour, 2016).

La diversidad microbiana de una muestra se analiza con índices de diversidad alfa como puede ser el índice de Shannon que da información sobre la diversidad y equidad de una muestra, es decir, es un índice que muestra que tan diversas son las especies de una muestra y que representación diferente tiene cada una. Dentro de los diferentes índices de diversidad alfa, este en particular, penaliza especies que están muy representadas y es más preciso para detectar organismos poco abundantes (Shannon, 1997).

Uno de los genes que ha sido protagonista en la caracterización del microbioma es el *16S ARNr*, componente del ribosoma bacteriano. Es un gen de aproximadamente 1400-1500 pares de base y contiene regiones altamente conservadas debido a que es un gen esencial para la función del ribosoma. Estas regiones facilitan el diseño de cebadores y su posterior amplificación por PCR para todas las especies en el dominio Bacteria. Entre las regiones conservadas, intercaladas, se encuentran 9 regiones hipervariables (V1-V9). Dichas regiones son las que permiten la identificación inter-especie ya que presentan diversidad entre ellas (Rozas, 2022). Entonces, si se secuencian estas regiones variables se puede hacer una clasificación de las bacterias presentes en una muestra dada. Las características del gen *16S ARNr* sumado a la existencia de bases de datos con genes de referencia, lo convierten en un candidato ampliamente utilizado en la investigación para estudiar la composición y diversidad de la comunidad bacteriana que conforma el microbioma (Rozas, 2022).

Las bacterias no son lo único que componen el microbioma, pero sí donde se han centrado la mayoría de los estudios en los últimos años. Como fuerza emergente hoy en día se está comenzando a estudiar el micobioma, la comunidad fúngica que se encuentra en ese microbioma (Yan, 2024). La caracterización del micobioma se realiza generalmente mediante el análisis de regiones *ITS* (Internal Transcribed Spacer) del ADN ribosómico o a través del gen *18S ARNr* (Wu, 2021). Las regiones *ITS* son fragmentos de ADN no codificantes que se encuentran entre los genes del ARN ribosómico (*ARNr*) en todos los eucariotas. Existen dos regiones *ITS1* e *ITS2*, las cuales

presentan variabilidad de secuencia comparable con el *16S ARNr* de bacterias. Consecuentemente estas regiones se utilizan como marcadores moleculares para la identificación y clasificación de hongos a nivel de especie (Huseyin, 2017). Sin embargo, lo que retrasa y aumenta la dificultad en el estudio del micobioma es la escasa existencia de bases de datos, y el hecho que las existentes poseen menor información en comparación con genes de referencia bacterianos (Suhr, 2016). No obstante, el intento de estudiar el micobioma y el bacterioma en conjunto proporcionaría una visión más completa del microbioma intestinal y posibilitaría generar relaciones más contundentes.

La reacción en cadena de la polimerasa cuantitativa (qPCR) permite la cuantificación de carga relativa de genes específicos (Klein D, 2002). Es una técnica que permite determinar la cantidad de una región específica de ADN, o amplicón, en tiempo real. La qPCR en el contexto del microbioma, se puede utilizar para determinar la existencia de genes de resistencia a antibióticos comparando con un gen de referencia que posean todas las bacterias como puede ser el gen *16S ARNr* (Livak, 2001) (Dahdouh, 2022) (Dahdouh E., 2023).

Los carbapenémicos están incluidos dentro de una clase de antibióticos betalactámicos ( $\beta$  lactámicos) de amplio espectro. Los cuales se utilizan como último recurso para infecciones bacterianas graves en entornos hospitalarios. No obstante, el aumento de bacterias resistentes a este tipo de fármacos se ha convertido mundialmente en un dilema grave para la salud pública (Jean, 2015). La producción de enzimas carbapenemasas por parte de las bacterias Gram-negativas es uno de los mecanismos principales de resistencia. Los genes *blaVIM* (*Verona Imipenemase*), *blaKPC* (*Klebsiella pneumoniae* Carbapenemase), *blaOXA-48* (*Oxacillinase-48*) y *blaNDM* (*New Delhi Metallo-beta-lactamase*) codifican carbapenemasas, las cuales hidrolizan los carbapenémicos (Hammoudi Halat, 2020). Dichos genes se suelen localizar en elementos genéticos móviles y por lo tanto el microbioma intestinal puede actuar como un reservorio significativo para estos genes de resistencia (Korach-Rechtman, 2020). Los elementos genéticos móviles como puede ser transposones o plásmidos conjugativos son los que cumplen ese papel de diseminación de genes de resistencia. Esto se debe a que transfieren elementos de ADN, entre clones de la misma especie y entre especies distintas, y así facilitar la rápida propagación de genes de resistencia (Touati, 2020).

La detección de genes de resistencia a través de una qPCR tiene como ventaja una mayor rapidez y una reducción de costes por muestra en comparación a una detección por secuenciación masiva

(Li, 2024). A su vez, lo que busca con esta amplificación por qPCR es la posibilidad de cuantificar la carga relativa utilizando los valores de ciclo umbral (Ct). Sin embargo, esta estrategia tiene la desventaja de que no va a poder detectar lo que no se está buscando. Es decir, que solo detectará genes de resistencia para los cuales se cuente con cebadores específicos. Por lo tanto, no será capaz de detectar otros genes de resistencia que no se incluyan en la estrategia planteada.

El concepto de carga relativa intestinal cuando se estudia el microbioma va a hacer referencia entonces a esa porción de un especie bacteriana o grupos de especies que poseen un gen específico que confiere cierta resistencia a antibióticos sobre la población total (Livak, 2001). Ese cálculo se puede realizar al comparar los Cts de los genes que confieren resistencia a carbapenémicos, analizándose su abundancia en relación con la población total, la cual se obtiene con los Cts del gen *16S ARNr*. Se ha propuesto en estudios recientes una relación entre la carga relativa intestinal de los genes de resistencia y la diversidad de especies presentes en el microbioma (Migliorini, 2022). La hipótesis sugiere que una elevada carga relativa intestinal de genes de resistencia podría generar una diversidad menor en la población microbiana total debido a que las bacterias que portan ese gen de resistencia se multiplican hasta dominar la población total mientras los antibióticos matan a los que no tienen genes de resistencia (Migliorini, 2022). Es decir, mientras el paciente este con un tratamiento de antibióticos, las bacterias portadoras de estos genes tendrían una ventaja competitiva. A pesar de que una vez finalizado el tratamiento se puede disminuir el predominio de esas bacterias resistentes, clínicamente afecta al paciente y a la posibilidad de tratar futuras infecciones bacterianas. Por lo contrario, una mayor diversidad alfa se asociaría con una menor carga relativa intestinal para genes de resistencia (Dahdouh E., 2023).

Las tecnologías de secuenciación de nueva generación (NGS) han sido un gran paso para comprender el microbioma y han revolucionado el campo de su estudio. Esto se debe a que han permitido el análisis del material genético de comunidades completas (Yen, 2021). Dentro de las distintas técnicas de NGS, la secuenciación de nanoporos creada por Oxford Nanophore Technologies™ en contraposición con otras técnicas como la Illumina™ permite la secuenciación de fragmentos de ADN largos como puede ser el gen *16S ARNr* en su longitud completa (Rozas, 2022). Los amplicones a secuenciar utilizando la tecnología de Illumina™ tienen una longitud máxima de 300 pb. Esta característica intrínseca de la técnica no permite la resolución necesaria para la identificación filogenética hasta el nivel de especie en grupos bacterianos similares (Esberg,

2024). No poder clasificar filogenéticamente con lecturas cortas se debe a que solo cubren una fracción del gen *16S ARNr*. Lo que impide captar la variabilidad de todas las regiones hipervariables necesarias para poder hacer una identificación precisa a nivel de especie. Sin embargo, esta tecnología ha sido el *gold standard* durante mucho tiempo gracias a alta resolución en la asignación de bases debido al tipo de química que utiliza la metodología. A pesar de que la tecnología de nanoporos aún está intentando aumentar la resolución en el *basecalling* posee una ventaja clave para ser la elegida en estudios de identificación bacteriana. Presenta una mayor resolución taxonómica que puede llegar hasta nivel de especie, cosa que se complica con tecnologías que generan lecturas de mayor precisión, pero más cortas (Cha, 2023). A su vez, la tecnología de Oxford presenta un costo relativamente menor lo que la vuelve una opción interesante para investigaciones del microbioma (Oehler, 2023). A pesar de que muchos estudios caracterizan el microbioma a través de muestras como heces o biopsias de la mucosa, diversos estudios han demostrado que las torundas rectales también son una alternativa válida. Varios trabajos han reportado alta similitud de la composición del microbioma entre heces y torundas rectales, considerándolas intercambiables. A su vez, las torundas rectales presentan la ventaja de poder ser recolectadas de manera inmediata en el momento que se necesiten (Budding, 2014) (Bassis, 2017).

El marco de este trabajo de fin de máster (TFM) se encuadra dentro del Hospital Universitario La Paz (HULP) situada en la Comunidad de Madrid. Se utilizó la secuenciación por nanoporos para analizar el gen *16S ARNr* de bacterias en muestras de torundas rectales de pacientes pediátricos. La cuantificación de carga relativa intestinal de genes de resistencia a carbapenémicos (*blaVIM*, *blaOXA48*, *blaKPC* y *blaNDM*) así como el gen *16S ARNr*, se llevó a cabo mediante datos de qPCRs. A su vez, se añadió el dato de si estas muestras tenían presencia de hongos, gracias a una qPCR utilizando cebadores en la región del *ITS2*. El estudio incluyó 46 muestras de pacientes pediátricos menores a 18 años, incluyendo neonatos. Mediante este enfoque se busca explorar en profundidad los datos del microbioma intestinal pediátrico y si hay correlación entre éste y las cargas relativas intestinales de genes de resistencia.

### 3. Objetivos.

- Objetivo general:
  - Analizar la diversidad del microbioma intestinal en pacientes pediátricos con el fin de relacionarlo con genes de resistencia, para a futuro implementar estrategias que predicen cambios precisos en el microbioma.
- Objetivos específicos:
  - Analizar datos de microbioma de 46 torundas rectales provenientes de pacientes pediátricos.
  - Relacionar la diversidad del microbioma con las cargas relativas intestinales para genes de carbapenemasas.

### 4. Metodología.

#### 4.1. Selección de las muestras y extracción del ADN.

Las muestras de torundas rectales incluidas en este estudio fueron de pacientes pediátricos que se atendieron en el HULP entre enero y marzo del 2025. Las muestras utilizadas fueron originalmente tomadas de estos pacientes para realizar una vigilancia epidemiológica de rutina y no específicamente para este trabajo. Es decir, las muestras fueron reutilizadas para este estudio. A su vez, los resultados de este trabajo no fueron informados a los pacientes y tampoco influyeron en el tratamiento de los mismos. Uno de los criterios excluyentes para seleccionarlas eran la edad de los pacientes, debían ser pediátricos. Por lo tanto, solo se utilizaron muestras de pacientes que tenían entre 0 y 18 años de edad.

Las muestras se seleccionaron con diferentes criterios: teniendo en cuenta los resultados de una qPCR, es parte de otro estudio y se realiza para determinar la presencia de genes de resistencia como *blaVIM*, *blaKPC*, *blaNDM* y *blaOXA-48*. A cada una se le calculó la carga intestinal relativa, utilizando la fórmula:

$$\text{Carga intestinal relativa} = \log (2^{(-(Ct_{carb} - Ct_{16S}))})$$

Donde el  $Ct_{carb}$  es el ciclo umbral al cual amplifica uno de los genes de resistencia y el  $Ct_{16S}$  es el ciclo umbral al cual amplifica el gen *16S ARNr*. Para el caso de muestras donde amplificaron más de uno de los genes se calculó entonces más de una carga intestinal relativa.

Para extraer el ADN de las muestras, se tomaron 100  $\mu$ L de las suspensiones obtenidas a partir de hisopados rectales y fueron diluidas en 900  $\mu$ L buffer TE (dilución 1:10) con el fin de evitar inhibiciones en las reacciones de PCR. Se incluyó extracción de agua (Fresenius Kabi<sup>TM</sup>) distribuida una posición en cada *run* para analizar posibles contaminaciones del proceso.

Luego esta dilución fue sometida a un hervido a 95 °C durante 10 minutos y una lisis mecánica a 7000 rpm durante 70 segundos utilizando tubos de MagNA Lyser Green Beads<sup>TM</sup> y el equipo Magna Lyser<sup>TM</sup> (Roche<sup>®</sup>). El último paso fue la extracción propiamente dicha en el equipo MagMAX<sup>TM</sup> express (Thermo Fisher Applied Biosystem<sup>®</sup>) utilizando los reactivos de la misma casa comercial y siguiendo los pasos especificados por el fabricante. Para ello se utilizaron 4 placas de 96 pocillos a las cuales se les agregó: a una de ellas 250  $\mu$ L de Wash 1. A otra placa 250  $\mu$ L de Wash 2. En otra placa, en la cual van a las muestras se agregó 175  $\mu$ L de Lysis Buffer, 175  $\mu$ L de Biding Buffer y 20  $\mu$ L de Magnetic Beads. Por último, en una placa de elución se agregó 90  $\mu$ L de Elution Buffer. Estas placas se colocaron en las posiciones que el equipo indica para cada una de ellas.

Una vez con el ADN extraído este se almacenó a -20°C hasta su uso.

#### **4.2. Reacción de qPCR.**

A pesar de tener los resultados de la qPCR que detectó genes de carbapenemasas y el gen *16S ARNr*, se realizó también una qPCR posterior a seleccionar las muestras.

El fin de esta qPCR era asegurar la presencia del gen *16S* y a su vez, en pocillos distintos, pero utilizando las mismas muestras amplificar el fragmento *ITS2* presente en hongos.

Los cebadores utilizados fueron *16S\_qPCR\_F* y *16S\_qPCR\_R* para detectar el gen *16S* y *ITS2\_qPCR\_F* y *ITS2\_qPCR\_R* para detectar el gen *ITS2* (tabla 1). Los volúmenes y concentraciones de los reactivos se encuentran en la tabla 2 y las condiciones de ciclado en la tabla 3.

**Tabla nº1.** Nombres y secuencias de los cebadores utilizados en la qPCR.

Cebadores	Secuencia
16S_qPCR_F	5'- TGGAGCATGTGGTTAACCGA -3'
16S_qPCR_R	5'- TGCAGGACTAACCCAAACA -3'
ITS2_qPCR_F	5'- TCCTCCGCTTATTGATATGC -3'

ITS2_qPCR_R	5' - GGAAGTAAAAGTCGTAACAAGG- 3'
-------------	---------------------------------

**Tabla nº2.** Componentes, concentración y volumen de la qPCR para amplificar el gen *16S* e *ITS2*, por separado. Los volúmenes son para una sola reacción.

Componentes y concentración	Casa comercial	Volumen (μL)
Agua	Fresenius Kabi™	7
Syber Master mix (2X)	Thermo Scientific™	10
Cebadores mezcla FR (c/u 0,5 μL)	Sigma-Aldrich	1
ADN extraído	-	2
Total por pocillo	-	20

La mezcla de cebadores FR se preparó llevando inicialmente a una concentración 1 μM a cada uno por separado. Esto se debe a que la solución madre se encuentra almacenada a 10 μM. Para ello se toman 10 μL de primer F o R y se diluye en 90 μL de agua. Luego, se tomaron 10 μL de primer F con 10 μL de primer R y se diluyen en 80 μL.

Se prepararon dos mezclas de reacción, una con los cebadores para *ITS2* y otra con los del *16S*. Para preparar cada mezcla de reacción se multiplican los valores de la tabla 2 por la cantidad de muestras a procesar más 2 muestras para corregir los errores de pipeteo. En este caso se calculaban para sets de 26 muestras a pesar de que se largaron 24 muestras.

**Tabla nº 3.** Protocolo diseñado para la qPCR. El termociclador utilizado es el CFX Connect.

Fase de la PCR	Temperatura (°C)	Tiempo	Nº de ciclo
Desnaturalización inicial	95	15 seg	1
Desnaturalización	95	15 seg	40 ciclos
Anillado	58	15 seg	
Elongación	72	30 seg	

Una vez finalizado este proceso, los amplicones se almacenaban a -20°C. A su vez se registraron los datos de los Ct a los cual amplificaban o no las muestras para ambos genes.

#### **4.3. Reacción de PCR punto final.**

Para cada muestra se realizó una PCR punto final con el fin de amplificar el gen *16S ARNr* completo para luego secuenciarlo. Los cebadores utilizados fueron 27F\_16S y 1492R\_16S (tabla 4) mientras que los reactivos utilizados, sus volúmenes y concentraciones se presentan en la tabla 5, y los condiciones de ciclado en la tabla 6.

**Tabla nº4.** Nombre y secuencia de los cebadores utilizados en la PCR punto final.

Cebadores	Secuencia
27F_16S	5'- AGAGTTGATCMTGGCTCAG -3'
1492R_16S	5'- TACGGYTACCTTGTACGACTT -3'

**Tabla nº5.** Componentes, concentración y volumen de la PCR punto final para amplificar el gen *16S*. Los volúmenes son para una sola reacción.

Componentes y concentración	Casa comercial	Volumen (μL)
Agua	Fresenius Kabi™	6,4
Rx mix (Phusion high-fidelity PCR master mix)	Thermo Scientific™ REF: F-532S	10
DMSO	Thermo Scientific™ REF: F-532S	0,6
Cebadores mezcla FR (c/u 0,5 μL)	Sigma-Aldrich	1
ADN extraído	-	2
Total por pocillo	-	20

La enzima polimerasa utilizada en este caso, es una de alta fidelidad debido a que es necesario obtener un fragmento amplificado de buena calidad ya que es el producto que va a ser secuenciado. La fidelidad de la misma es 52 X comparado con una *Taq* comercial, que se traduce en una tasa de error 50 veces menor. A su vez, se agrega el DMSO con la intención de potenciar el trabajo de la polimerasa y aumentar su precisión, especialmente en las zonas ricas en guanina y citosina (GC).

**Tabla nº6.** Protocolo diseñado para la PCR punto final. El termociclador utilizado es el 2720 Thermal Cycler de Applied Biosystem™.

Fase de la PCR	Temperatura (°C)	Tiempo	Nº de ciclo
Desnaturalización inicial	95	5 min	1
Desnaturalización	95	30 seg	35 ciclos
Anillado	50	30 seg	
Elongación	72	1,5 min	
Extensión final	72	7 min	1
Resguardo de los amplicones	4	∞	1

Previo a entrar a la etapa de secuenciación, se debió realizar un gel de electroforesis para visualizar que efectivamente la PCR se realizó correctamente. Paso que sirve para confirmar si había producto de todas las muestras sin amplificaciones inespecíficas (e.j. bandas de diferentes tamaños visualizados en el gel). Mientras que el control negativo, agua, no se debería visualizar nada.

La presencia del producto esperado (única banda de 1500 pb) fue corroborado por electroforesis en geles de agarosa al 1%, teñidos con 3µL de Gel Red (10,000X) (LabNet Biotécnica®). En cada pocillo del gel se depositó 5 µL del producto de PCR junto a 2 µL de buffer de carga. Para definir el tamaño del amplicón se utilizó un marcador de peso molecular 1 kb. Los geles se corrieron 30 minutos a 90 V y 500 mA en cubetas de electroforesis horizontales (LabNet Biotécnica®) y luego se revelaron en un transiluminador de luz UV (Spectroline®). El volumen remanente de la reacción de PCR se almacenó a -20 °C.

#### **4.4. Cuantificación con Qubit.**

Una vez confirmado la presencia del producto único con el gel de agarosa se realizó una cuantificación en el fluorómetro Qubit 4 (Thermo Fisher Scientifc®) con el protocolo “Broad Range” y utilizando el kit Qubit™ dsDNA HS assay. Para ello se debió preparar los dos estándares de ADN que vienen con el kit, uno de concentración nula de ADN (0 ng/µL) y otro de una concentración alta (100 ng/µL) para construir una curva de calibración. Así cuando se ingresó la

muestra en el fluorómetro, en función de la fluorescencia obtenida el propio equipo te dice la concentración de cada muestra.

Se tomó 1  $\mu$ L de cada muestra y en tubos de 0,5 mL (Qubit<sup>TM</sup> assay tubes) se agregaron 99  $\mu$ L de working solution 1X preparada previamente con el kit.

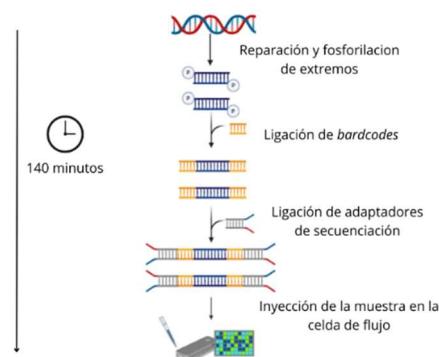
Este paso previo a la secuenciación es para asegurar que se esté ingresando al proceso con una cantidad de ADN similar de cada muestra. El rango de concentración para este tipo de muestras y tipo de proceso era aceptable valores entre 15-60 ng/ $\mu$ L.

Una vez que finalizó la PCR y los controles por Qubit y electroforesis en geles, se almacenó los amplicones a -20 °C hasta la preparación de las librerías para secuenciación.

#### **4.5.Secuenciación NGS.**

Para la preparación de las librerías se utilizó el kit Native Barcoding kit 24 V14 (SQK-NBD114.24) (Oxford Nanopore Technologies, 2025) según indicaciones del fabricante con algunas modificaciones. El equipo de secuenciación utilizado fue el MinION Mk1C (Oxford Nanopore Technologies<sup>TM</sup>).

Un esquema general de los pasos se presenta en la imagen 1, cada paso se va a detallar a continuación. Mientras que los volúmenes que se mencionan en los próximos apartados son para 24 muestras debido a que es la cantidad máxima para el tipo de kit utilizado.



**Imagen 1.** Preparación de las librerías e inyección de la muestra en la celda de flujo. Imagen de elaboración propia, Biorender. (Adaptado de Oxford Nanopore Technologies, 2025)

#### 4.5.1. Reparación y fosforilación de extremos.

El primer paso es la reparación de los extremos para generar extremos romos y para que los extremos se fosforilen.

Se preparó la mezcla de reacción con los reactivos NEBNEx® Companion Module E7180 según las instrucciones del fabricante. Se mezclaron 39  $\mu$ L de Ultra II End-prep Reaction Buffer y 18.2  $\mu$ L de Ultra II End-prep Enzyme Mix (cálculo para 26 muestras para tener volumen en exceso). Se agregó 2.2  $\mu$ L de la mezcla a cada muestra.

Se tomaron 12.5  $\mu$ L (de concentración entre 15-60 ng/ $\mu$ L) de amplicón de cada una de las muestras y se le añadieron 2  $\mu$ L de mezcla mencionada anteriormente. A esos 14  $\mu$ L de reacción se la incubó en un termociclador (2720 Thermal Cycler de Applied Biosystem™) 10 minutos a 20°C y luego otros 10 minutos a 65°C. Los 20°C son la temperatura óptima para que la enzima funcione. Mientras que los 65°C finales son para desactivar la enzima.

Inmediatamente, una vez finalizada la fosforilación de los extremos en el termociclador, se realizó un paso de limpieza. Para ello se preparó 5 mL de etanol al 80%, el cual se debe preparar el mismo día de su uso para asegurar que no hubo evaporación de etanol.

Se añadió 14.5  $\mu$ L AMPure XP Beads (que son bolitas magnéticas) a cada una de las muestras y se resuspendió suavemente con pipeta, incubando 5 minutos a temperatura ambiente. Luego, se colocaron los tubos en un imán esperando hasta que la solución se aclare. Ya con la solución aclarada se tomó con una pipeta el sobrenadante y se descartó. De esta manera, en el pellet quedó el ADN unido a las bolitas magnéticas gracias a la interacción catiónica que se genera por la carga negativa del ADN y la positiva de las bolitas magnéticas. Se agregó 180  $\mu$ L del etanol 80% homogenizando de arriba a abajo 10 veces, suavemente utilizando la pipeta de 200  $\mu$ L. Nuevamente, se descartó el sobrenadante. A esos tubos se les dió un spin de 5 segundos para que caiga el etanol remanente, se volvió a colocar en el imán y aspiró con pipeta los restos de etanol.

Por último, se quitaron los tubos del imán, resuspendiendo el pellet con 6  $\mu$ L de H<sub>2</sub>O free (Fresenius Kabi™) sin tocar el pellet y dispensando el agua por encima de éste. Se dejó 2 minutos a temperatura ambiente y para finalizar se volvieron a colocar los tubos en el imán y se recuperó 5  $\mu$ L del eluido que tiene el ADN purificado.

#### 4.5.2. Ligación de los bardcodes.

El producto del paso 4.5.1 se cataloga como End-prepped ADN (5  $\mu$ L). A cada uno de los tubos se les añadió 1.5  $\mu$ L de cada *barcode*, en este caso se contaba con los *barcode*s del 1 al 24 por *run*. Por lo tanto, a cada muestra se le añadió un *barcode* correspondiente en orden creciente. Es importante en este paso, saber a qué muestra se le añadió que número de *barcode* ya que éste sirve como una “etiqueta” para la muestra. Por lo tanto, permite que luego las muestras se junten en un mismo tubo porque ya fueron etiquetadas. Luego, a cada tubo se le añadió 4.5  $\mu$ L de la mezcla de reacción de la blunt/TA ligasa, la enzima encargada de ligar los *barcode*s a los extremos romos fosforilados del amplicón. Ya con la ligasa y los *barcode*s en cada tubo, se incubó en termociclador (2720 Thermal Cycler de Applied Biosystem<sup>TM</sup>) a 25 °C durante 25 minutos. Una vez pasados los 25 minutos, se agregó 2  $\mu$ L de EDTA a cada tubo para parar la reacción de ligación.

En este punto cada muestra se encontraba asociada a su *barcode*, por lo tanto, ya fue posible juntar 11  $\mu$ L de cada tubo en un tubo eppendorf de 1,5 mL. Con todas las muestras en un mismo tubo, se realizó la segunda limpieza. Para ello, se agregan 5 veces de bolitas magnéticas por muestras en la mezcla (24x5) y se agregaron al tubo. Mientras se incubaba 10 minutos a temperatura ambiente, se preparó 1mL de etanol al 80%. Ya pasados los minutos, se colocó el tubo en el imán, se esperó que aclare y se descartó el sobrenadante. Sobre el pellet, se agregó 800  $\mu$ L de etanol 80%, se aspiró de arriba a abajo 10 veces y se descartó el sobrenadante.

Luego se les dio un spin a las muestras para que los restos de etanol queden abajo y aspirarlos. Con el tubo fuera del imán se agregaron 21  $\mu$ L de agua free y se incubó 10 minutos a 37 °C en termociclador (2720 Thermal Cycler de Applied Biosystem<sup>TM</sup>). Una vez terminado, se colocó en el imán y se recuperó 20  $\mu$ L del eluido.

#### 4.5.3. Ligación de los adaptadores y limpieza.

Sobre el tubo con 20  $\mu$ L del pool del paso 4.5.2 se agregó 4  $\mu$ L del *native adapter*, 8  $\mu$ L de la mezcla de ligación con el buffer para que actúe la ligasa y 4  $\mu$ L de la ligasa de ADN T4 quick. El *native adapter* se une a los *barcode*s que se encuentran en los extremos de cada amplicón. La función de este adaptador es unir cada amplicón al poro que se encuentra en la celda de flujo del MinION. Estos adaptadores son proteínas que se unen al amplicón y al poro. La unión del adaptador al *barcode* lo realiza la ligasa T4.

Para catalizar esta reacción de unión se debió incubar esta mezcla en un termociclador 25 minutos a temperatura ambiente (25°C). Una vez pasados estos minutos, se realizó la tercera y última limpieza. Sobre los 36 µL anteriores se agregó 13 µL de AMPure beads y se dejó 10 minutos a temperatura ambiente. Luego se colocó el tubo en el imán, se esperó a que aclare y luego se descartó el sobrenadante. Sobre el pellet se agregó 125 µL de Buffer de Fragmentos Largos (LFB), buffer que se encarga de beneficiar la presencia de amplicones largos en el eluido y que se descarten los cortos. En este caso no se utiliza etanol y se utiliza el LFB porque es la última limpieza previa a la secuenciación. No es posible utilizar etanol debido a que este dañara tanto los adaptadores como el poro de la proteína que se encuentra en la celda de flujo. Se dispensó el buffer y se aspiró de arriba abajo 10 veces y luego se descartó el sobrenadante. Se le dió un spin al tubo y se descartaron los restos del LFB.

Finalmente se quitó el tubo del imán, se agregaron 7 µL del buffer de elución (incluido en el kit de secuenciación) y se incubó a 37 °C durante 10 minutos en termociclador (2720 Thermal Cycler de Applied Biosystem). Ya pasados esos minutos se colocó el tubo en el imán y se recuperaron 6 µL del eluido en un nuevo tubo.

De esos 6 µL se tomó 1 µL y se cuantificó en el fluorómetro Qubit 4 (Thermo Fisher Sciemtific®) con el protocolo “Broad Range”. Para ello se debió preparar los dos estándares de ADN que vienen con el kit, como se explicó en el apartado 4.4.

#### 4.5.4. Preparado e inyección de la muestra en la celda de flujo.

Como paso previo a cargar la mezcla de preparado y la muestra en la celda de flujo, se debe atemperar la celda de flujo y hacer un chequeo de poros de la misma, la cual debe tener como mínimo 50 poros para obtener un número aceptable de lecturas.

La celda de flujo se colocó en el MinION y se despegó la pegatina negra, de esta manera se deja descubierto el puerto de carga. Siempre tener en cuenta que no se puede generar presión alguna sobre la celda ya que se pueden dañar los poros.

La mezcla de preparado (*priming mix*) que se encuentra en el kit de expansión (Oxford Nanopore) sirve para preparar la celda antes de inyectar la muestra. Se mezclaron 117 µL de Flow Cell Flush y 3 µL de Flow Cell Tether. Esos 120 µL se debe dispensar sobre el puerto de carga de la “Flonge

Flow cell”, sin generar ningún tipo de burbuja ya que estas van a dañar los poros de la membrana. Para pipetear sin generar burbujas se aspiró el volumen con una pipeta de 200  $\mu\text{L}$ , la cual se seteó en 120  $\mu\text{L}$ . Antes de apoyar la punta sobre la celda de flujo se aseguró que no había aire en la punta, para ello se giró la rueda de la pipeta hasta el líquido quede bien en el final de la punta. Ya con esto asegurado, se colocó la pipeta perpendicularmente sobre el puerto de carga de la celda. Una vez allí, se comenzó a emitir el volumen sin presionar émbolo si no que girando la rueda de manera tal que el volumen disminuya e ingrese el líquido a la celda. Al cargar esta solución se consiguió hacer vacío sobre el puerto de carga.

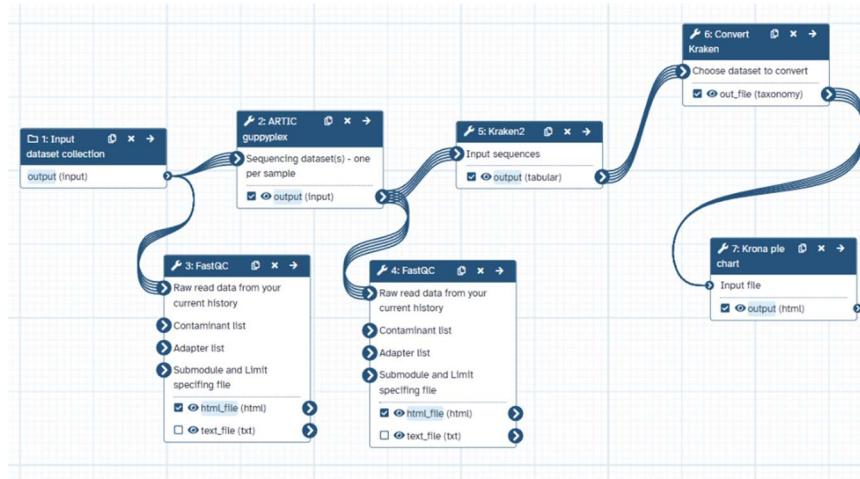
Por separado, en otro tubo, sobre los 5  $\mu\text{L}$  de librería de ADN obtenidos en el paso 4.5.3 se agregó 15  $\mu\text{L}$  de buffer de secuenciación y 10  $\mu\text{L}$  de beads de librería. Gracias al vacío que se generó en el puerto de carga, estos 30  $\mu\text{L}$  se inyectan a la celda pipeteando gota a gota. Ya con todo cargado se debe tapar el puerto de carga con la pegatina que previamente se retiró y se programó el MinION.

Se debe seleccionar correctamente el tipo de celda de flujo utilizada, el tipo de *basecalling* que se quiere obtener en la secuenciación y en qué valor de calidad realizar el corte. En este caso, el tipo de celda es FLO-FLG114 y el tipo de *basecalling* se utilizó “high sensitivity”, es decir, alta sensibilidad y el valor de calidad el 9. Este valor es equivalente a 1 error en el *basecalling* por cada 1000 bases. Todas las lecturas que no entran en esta calidad alta van a entrar dentro de las lecturas que el secuenciador identifica como *failed*.

#### **4.6. Análisis bioinformático.**

Una vez finalizado el proceso de secuenciación en el secuenciador MinION, se descargaron los archivos fastq. Este tipo de archivos los genera el propio secuenciador al realizar el *basecalling*, convirtiendo las señales de cambio de electricidad en nucleótidos con valores de calidad asignando a su vez mientras los fragmentos de ADN pasen por los poros en la celda. Es decir, que estos archivos tienen datos de secuencia y calidad de cada lectura. Dentro de MinKnow, el programa del secuenciador de Oxford Nanopore, se puede ver diferentes métricas del *run* analítico para evaluar el desempeño de la misma. Entre estos parámetros se pueden destacar el número de lecturas por *barcode* y el valor de calidad de las lecturas (Q).

Las muestras se analizaron en la nube Galaxy (Galaxy Community, 2024), que es una interfase que conecta un sistema operativo de Linux con Windows y es de acceso gratuito. Para poder llegar a un resultado de clasificación taxonómica y visualización de la composición de cada muestra se debieron seguir una serie de pasos, mencionados en la imagen 2. El flujo permite realizar los pasos de manera secuencial. Por lo tanto, allí figuran todas las herramientas en orden donde los archivos de salida del paso anterior son el archivo de entrada al siguiente.



**Imagen 2.** Flujo de trabajo realizado en Galaxy (Galaxy Community, 2024).

- *Trimming*: es el proceso por el cual se limpian las lecturas crudas obtenidas en el secuenciador. Es decir, que se eliminan lecturas de baja calidad porque no son confiables, y por el tamaño de las lecturas que no coinciden con lo del gen completo *16S ARNr*. Los archivos de entrada a este proceso son los que tienen formato fastq. Hay diferentes herramientas para realizar esto en Galaxy (Galaxy Community, 2024) pero en este caso se utilizó Artic Guppyplex (versión 1.6.0) que esta optimizado para lecturas de nanopore. En este caso, la calidad se determinó en 8 y el largo de las lecturas máximo en 1800 pb y mínimo en 1200 pb ya que el gen completo de *16S* es de apropiadamente 1500 pb. El máximo se seleccionó por encima de este valor ya que hay que considerar que en los extremos del amplicón se encuentran tanto los *barcode*s como los adaptadores.
- FASTQC (versión 0.12.1) (Andrews, n.d.): es otra herramienta de Galaxy que permite tener un reporte de la calidad de las lecturas. En este caso se utilizó tanto antes como después del proceso de limpieza de lecturas, lo que permite comparar y sacar conclusiones de si esa limpieza elimina muchas lecturas y de la calidad de los resultados.

- Kraken2 (versión 2.1.3) (Wood, 2014): es el programa que se encarga de clasificar las lecturas en etiquetas taxonómicas. El archivo de entrada de esta herramienta debe ser el archivo de salida del Artic Guppyplex. La función entonces es tomar las secuencias de ADN y le asigna etiquetas taxonómicas. Para realizar la clasificación es necesario que el programa cuente con una base de datos de referencia para poder mapear las lecturas. Kraken2 entonces incluye una base de datos predeterminada basada en el *RefSeq* del “National Center Of Biotechnology Information” (NCBI) de Estados Unidos (Wood, 2014). La herramienta “Basic Local Aligment Search Tool” (BLAST) creada por el Instituto Nacional de Salud de Estados Unidos (Zhang, 2000), funciona de manera de similar al alinear datos de secuencias con bases de datos. Una de las ventajas novedosas de Kraken2 es que gracias a su algoritmo realiza esta clasificación de manera más rápida (Wood, 2014).

En este caso se seleccionó la base de datos “Prebuilt Refseq indexes: PlusPF (standard plus protozoa and fungi) (Version: 2022-06-07 – Downloaded: 2022-09-04T165121Z)”. Es una de las bases de datos en BLAST más completas donde hay genomas de referencia de bacterias, arqueas, humanos, virales, plásmidos, hongos y protozoos. La herramienta no corrige errores, todas las lecturas que encuentra las clasifica. Por eso, es importante que el archivo de entrada sea el archivo de salida del Artic Guppyplex. Lo que hace es clasificarlas en documentos de hoja de cálculo, pero no permite visualizarlos, para ello se usa Krona.

- Convert Kraken (versión 5.1.0) (Wood, 2014) (Kosakovsky Pond, 2009): el archivo de salida del Kraken2 no es compatible con la entrada de la versión de Krona del Galaxy. Por lo tanto, se utiliza esta herramienta intermedia entre un paso y otro. Todos los archivos de Kraken2 pasan por esta herramienta y quedan en formato “taxonomy”.
- Krona (versión 2.7.1) (Ondov, 2011) (Cuccuru, 2014): con esos archivos en formato taxonomy, la herramienta Krona genera gráficos y permite visualizar los datos. En este caso se seleccionó que la herramienta clasifique desde raíz hasta especie.

#### **4.7. Análisis por Excel.**

A pesar de utilizar Galaxy y finalizar con los datos visualizados en Krona, se realizó posteriormente una limpieza de lecturas y se agruparon con el resto de los datos que se había recolectado durante su selección. Para ello, con los archivos de hojas de cálculo que generó Krona, se descargaron y

se analizaron manualmente. Para el análisis se siguieron varios pasos, primero sobre la hoja de cálculo de cada muestra se contaron (fórmula contar de Excel) las lecturas totales de la muestra y luego se contaron (fórmula contar.si de Excel) condicionalmente cuantas lecturas tenían la misma etiqueta de identificación sobre las lecturas totales. Una vez con estos dos datos, se borraron las fórmulas y se dejaron los valores, de tal manera de eliminar los duplicados de esas etiquetas taxonómicas. Debido a que Krona clasifica las lecturas haciendo un BLAST de cada una, las especies en la clasificación se repiten ya que el programa utiliza bases de datos que no están al 100% curadas. Para solucionar esto, se realizó una tabla dinámica de cada muestra utilizando los datos de especies en la muestra y el número de lecturas. A partir de cada una de las tablas que se realizaron para todas las muestras, se extrajeron las especies mayoritarias de cada una y se juntaron todos los datos en una sola tabla. De manera de hacer un resumen del *run*, las lecturas que eran menor al número de lecturas de la especie mayoritaria del agua se catalogaron como “otros”. Mientras que las mayores a ese número de lecturas, si se dejaron descritas, tanto con su número de lecturas y el número total de lecturas de la muestra.

De tal manera, que el resultado final fue una hoja de cálculo con cada una de las muestras con todos los datos relevantes agrupados. Dentro de ellos, las especies que conformaban la muestra acompañados de datos como sexo del paciente, edad, número de lecturas. Para el caso de las muestras que eran positivas para alguno de los genes de resistencia se agregaron los datos de la carga intestinal relativa. A su vez, se calculó el índice de Shannon utilizando las ecuaciones mencionadas a continuación.

$$\text{Índice de Shannon} = \sum(\ln P_i \times P_i)$$

Siendo:

$$P_i = \frac{n^{\circ} \text{ de lecturas por especie}}{n^{\circ} \text{ total de lecturas}}$$

Es relevante resaltar que, para este cálculo no se consideraron las lecturas que Kraken2 no pudo clasificar hasta especies. Es decir, que había lecturas que quedaban clasificadas como “n” para especies porque la herramienta solo llegaba a una clasificación a nivel de género o incluso, solo hasta familia. El criterio que se tomó fue que para esas lecturas que eran una suma de “n” a nivel especie, solo se les anotó que tipo de Familia era la mayoritaria acompañada del porcentaje.

#### **4.8. Análisis estadístico.**

Como se mencionó en el punto 4.1 se calcularon las cargas intestinales relativas. En función del valor de la misma, las muestras se clasificaron en 3 categorías: carga alta (0 a -2), carga media (menor a -2) y negativos para muestras que no presentaron amplificación para ningún gen de carbapenemasas. A su vez, se recolectaron datos como el sexo y la edad de los pacientes.

La clasificación se basa en que las muestras que se encuentran en una categoría de carga alta van a tener entre un 100 y un 10 % de bacterias con genes de resistencia, carga media entre un 10 y un 1% mientras que carga baja sería entre un 1 y un 0,1%, en este caso no había ninguna muestra que cumpla con esta categoría. Lo que se espera en una persona adulta sana, con un microbioma variado y sin enfermedad, es una carga entre media baja.

Se realizó aparte un archivo de Excel donde de cada muestra se contaba con el índice de Shannon correspondiente y la categoría a la cual pertenecía esa muestra según su carga intestinal relativa. También con el sexo y edad. Con este nuevo archivo se realizaron diferentes análisis utilizando la herramienta R Studio® (R Core Team, 2024). Análisis más descriptivos como distribuciones de los índices Shannon en las categorías de sexo y edad. Para ver distribuciones de ambos sexos se realizó un *violin-plot*. Mientras que, con respecto a la edad, la división se realizó en 4 categorías y se realizó un gráfico de barras donde primero se debió calcular: media, desviación y número de datos por categoría.

Mientras que para realizar un análisis inferencial se comenzó con una prueba de Shapiro-Wilk (Shapiro, 1965). Para ello también se utilizó R (R Core Team, 2024). Una vez que se confirmó normalidad, se realizó una prueba de Levene para evaluar homogeneidad de las varianzas (Levene, 1960). De tal manera con la prueba de normalidad y de análisis de las varianzas, se pudo realizar una prueba de Student para comparar las medias entre los distintos grupos (Student, 1908). Esta prueba estadística permitió analizar si había cambios significativos entre las diferentes categorías de cargas intestinales y los índices de Shannon calculados. Es decir, analizar si había diferencias entre la diversidad del microbioma de las muestras dependiendo de las diferentes categorías a la cual pertenecían.

Por último, utilizando R se realizó una imagen de resumen con la composición de especies de cada muestra. Se dividió el gráfico en las categorías teniendo en cuenta carga intestinal relativa. El

gráfico de barras apiladas tuvo en cuenta las lecturas de cada especie dentro de cada muestra. Es importante destacar que las lecturas clasificadas como “n” donde no se logra clasificar hasta especie, no se consideraron para este gráfico.

## 5. Resultados.

### 5.1. Selección de muestras y extracción del ADN.

Dentro de las 46 muestras de torundas rectales seleccionadas, 22 eran positivas para genes de resistencia a carbapenémicos y 24 eran negativas para dichos genes. Cada *run* de secuenciación constaba de 24 muestras y en cada uno hubo una posición de muestra que se ocupó con agua. Agua la cual se extrajo de la misma manera que las muestras. En a tabla 1 del anexo, se detallan los códigos de cada muestra y la categoría a la cual pertenecen.

### 5.2. Reacción de qPCR.

A las 48 muestras se les realizó la reacción de qPCR para detectar el gen *16S ARNr* y la región *ITS2*. En la tabla 7 se ven los resultados y cuantas muestras amplificaron en cada reacción. Teniendo en cuenta que un Ct mayor a 35 se consideró negativo, ya que valores por encima de este umbral suelen asociarse a amplificaciones inespecíficas o cantidades de ADN por debajo del límite de detección confiables.

**Tabla nº7.** Resultados qPCR.

Categoría muestras	Cantidad de positivas qPCR <i>16S ARNr</i>	Cantidad de positivas qPCR <i>ITS2</i>
Muestras positivas para genes de resistencia a carbapenémicos	22	5
Muestras negativas para genes de resistencia a carbapenémicos	23	3
Agua	0	0

Por lo tanto, como se indica en la tabla 7 todas las muestras que provenían de torundas rectales amplificaron para el gen característico de bacterias menos una de ellas. Dicha muestra (Mb358) con un Ct por encima de 35, negativa, provenía de un paciente neonato y se consideraron dos hipótesis: un error en la toma de la muestra porque al diluirla no mejoró la amplificación y otra opción, sería que por consumo de antibióticos y al ser un paciente neonato, el antibiótico haya erradicado a las bacterias que se podían detectar del recién nacido.

Como se esperaba, ambas muestras de agua no amplificaron en ningún caso.

### **5.3. Reacción de PCR punto final.**

Para los resultados de la PCR punto final se realizaron dos geles de agarosa al 1%. En la imagen 1 del anexo se puede observar ambas fotos de los geles. Como detalle se menciona que la muestra proveniente de neonato que no amplificó en la qPCR tampoco lo hizo para este caso. Razón por lo cual en esa posición (marcada con flecha roja en la imagen 1 del anexo) en el gel se ve sin amplificación.

### **5.4. Secuenciación NGS.**

A partir de los reportes realizados por el MiKNOW se consideró que ambos *runs* pasaron los parámetros de calidad del programa. En la imagen 2 del anexo, se puede observar el gráfico que realiza el programa para presentar si la mayoría de las lecturas están cerca de lo estimado o si fallaron.

A su vez, con los reportes de cada *run* se decidió que 3 muestras negativas para genes de resistencia a carbapenémicos (del *run* 2) se iban a dejar fuera de los análisis. Esto se debe a que el número de lecturas por muestra era considerablemente menor al resto. En el *run* 2, la media de las lecturas totales de cada muestra fue de aproximadamente 1900 lecturas. Sin embargo, las muestras que se dejaron por fuera de los análisis tenían 389 lecturas (Mb358), 1218 lecturas (Mb365), y 1062 lecturas (Mb375).

### **5.5. Análisis bioinformático.**

Con el análisis realizado con Galaxy se obtienen como resultado final, de cada muestra de cada *run*, un gráfico que realiza la herramienta Krona. Esos gráficos interactivos que se generan tienen la composición de especies de cada muestra. A modo de ejemplo se adjuntan 4 imágenes en el anexo (imágenes 3, 4, 5 y 6 del anexo).

Dos de las muestras adjuntadas en el anexo son Mb4 y Mb14 (imágenes 3 y 4 del anexo, respectivamente) pertenecen a la categoría de muestras que tenía una carga alta para genes de resistencia a carbapenémicos. Mientras que Mb355 y Mb373 (imágenes 5 y 6 del anexo, respectivamente) pertenecen a la categoría de muestras negativas para dichos genes. A pesar de las categorías a la cual pertenece cada una, observando los gráficos generados por Krona se puede

destacar que la diversidad de especies es mayor en los casos de Mb14 y Mb373 comparado con las otras dos.

### **5.6. Análisis en Excel y análisis estadístico.**

En la tabla 2 del anexo se puede ver todas las muestras analizadas con la categoría a la cual pertenecen, índice de Shannon calculado, sexo y edad.

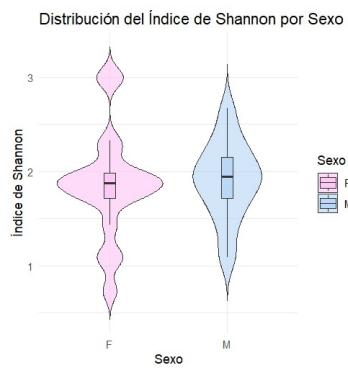
Como parte de un análisis estadístico inferencial, se comenzó realizando la prueba de Shapiro-Wilk donde se contaba con 3 categorías: alta, media y negativos. Con los resultados de la prueba no se pudo rechazar la hipótesis nula, que sugiere que las muestras siguen una distribución normal. Esto se debe a que los valores de p valor eran mayores a 0,05 y los valores de las estadísticas w son cercanas a 1 (resultados adjuntados en la tabla 3 del anexo). Por lo tanto, la distribución de los índices de Shannon se considera normal en las 3 categorías. En la imagen 7 del anexo se puede observar la distribución de los índices de Shannon en cada una de las categorías.

Una vez confirmada la distribución normal, con la prueba de Levenne se confirmó la homogeneidad de las varianzas. Ya que el p valor obtenido era de 0,1756, valor mayor a 0,05. Por lo tanto, se pudieron considerar que las varianzas de las 3 categorías son homogéneas entre sí.

Gracias a la normalidad de los datos y la homogeneidad de las varianzas, fue posible realizar las pruebas de Student para comparar las 3 categorías entre sí (resultados en la tabla 4 del anexo). Por los resultados obtenidos no se pudo considerar que haya diferencias significativas entre los índices de Shannon de cada categoría. Esto se debe a que las medias son muy similares entre sí y por lo tanto, la diferencia entre ellas es casi nula. Siendo la media de la categoría alta 1.86, la media de la categoría media 1.84 y de los negativos 1.92.

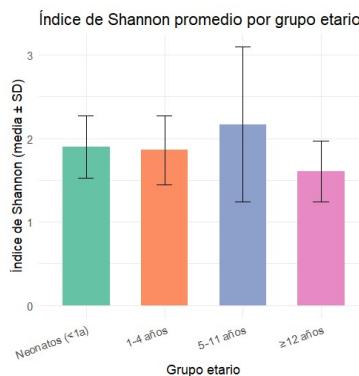
Como parte de un análisis descriptivo con los datos de la tabla 2 del anexo se realizaron gráficos con R para mostrar las distribuciones de los índices de Shannon en ambos sexos. Del total de muestras, había 25 masculinos y 21 femeninos. Se realizó un *violin-plot* (imagen 4) donde se ve la distribución para ambos sexos. Ésta sugiere que en ambos sexos las medianas de los índices de Shannon son muy similares entre sí, para femenino esta aproximadamente entre 1.9-2 y para masculino en 2.0. No obstante, la forma de los gráficos sugiere que para el grupo femenino hay

mayor heterogeneidad en la distribución que en el caso de los masculino donde los datos están más centrados en su mediana.



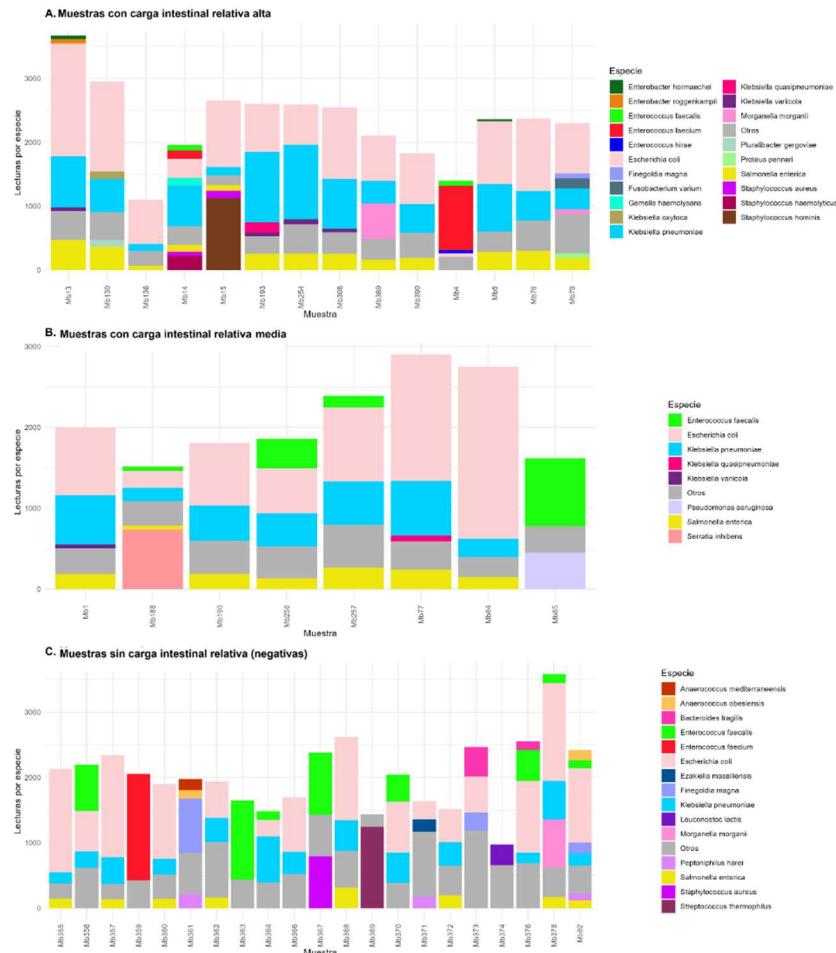
**Imagen 4.** Violin-plot para la distribución de los índices de Shannon según el sexo de los pacientes.

También de manera descriptiva se realizó un gráfico de barras para las distribuciones de los índices de Shannon en los diferentes grupos etarios (imagen 5). Neonatos (menores de 1 año) con un  $n=15$  muestras, entre 1 y 4 años ( $n=19$ ). Mientras que, entre 5 y 11 años,  $n=5$  y mayores a 12 años solamente  $n=4$ . Teniendo en cuenta el grafico de la imagen 5, se observa que la variación del índice de Shannon no es lineal con el aumento de la edad. Para el grupo entre 5 y 11 años se observa una mayor diversidad que para los otros grupos, pero también la barra de error es mayor, sugiriendo heterogeneidad en los datos. Mientras que, para el grupo de mayores de 12 años, la diversidad se observa disminuida. A pesar de estas observaciones que sugieren los gráficos, las diferencias en la cantidad de muestras de cada grupo no permite establecer que estas distribuciones se deban a una razón concreta entre las edades de los pacientes y la diversidad del microbioma, y no a una aleatoriedad de los datos.



**Imagen 5.** Gráfico de barras con la distribución del índice de Shannon por grupos etarios.

Para analizar la estructura de la comunidad microbioma a nivel de especie, se generaron gráficos de barras apiladas (imagen 6) utilizando RStudio, una forma de resumir la composición taxonómica de cada muestra. En todos los casos (A, B y C), las muestras tienen lecturas indicadas como “Otros” (especies que por sí solas tenían muy pocas lecturas).



**Imagen 6.** Gráfico de barras apiladas para representar la composición de cada muestra con sus lecturas por especie. Eje X, código de identificación de cada muestra. Eje Y, lecturas por especie. A. Muestras con carga intestinal alta para genes de resistencia a carbapenémicos. B. Muestras con carga intestinal media para genes de resistencia a carbapenémicos. C. Muestras negativas para genes de resistencia a carbapenémicos.

A partir del grafico de barras apiladas (imagen 6) se observan diferencias en los patrones de composición según la categoría a la cual pertenecen las muestras. De manera general, para las categorías, hay muchas lecturas ocupadas por “Otros” (color gris). En las tablas 5, 6 y 7 del

anexo se encuentra cuantas especies se encuentran agrupadas en “Otros” de cada una de las muestras en las imágenes 6 A, B y C, respectivamente.

Para el caso de las muestras con carga intestinal alta (A), se destacan que en la mayoría hay presencia de especies como *Escherichia coli*, *Klebsiella pneumoniae* y *Salmonella enterica*. Las muestras con carga intestinal media (B) también hay en la mayoría presencia de *E. coli* y *K. pneumoniae* pero sugiere una menor diversidad de especies que para el caso A. Con respecto al caso de muestras negativas para genes de resistencia (C), también la mayoría presentan *E. coli* y *K. pneumoniae* con la diferencia que tienen muchas lecturas ocupadas por “Otros”.

## **6. Discusión.**

En este trabajo, se propuso explorar el microbioma intestinal en 46 muestras provenientes de pacientes pediátricos a través de la secuenciación del gen *16S ARNr* completo. A su vez, como criterios de clasificación de las muestras se utilizaron los resultados de una qPCR para genes de resistencia a carbapenémicos (*blaVIM*, *blaOXA48*, *blaKPC* y *blaNDM*). Resultados a partir de los cuales se clasificaron en 3 categorías basadas en las cargas intestinales relativas. La hipótesis planteada en este trabajo consistía en que cargas intestinales relativas altas se verían asociados a una menor diversidad alfa. Diversidad asociada en este caso con el cálculo del índice de Shannon, es decir, a mayor índice, mayor diversidad del microbioma.

A partir de los resultados obtenidos en la secuenciación existirían diferencias con la hipótesis inicial. En el grafico de barras apiladas de la imagen 6 se puede ver un buen resumen de los resultados. A simple vista, la cantidad de lecturas varían bastante de muestra a muestra y esto podría estar relacionado con las diferentes concentraciones de ADN con la cual se inicia el proceso de secuenciación. Otra observación importante es que, indiferente de la categoría de las muestras representadas, hay muchas lecturas ocupadas por la categoría “Otros”, lecturas de especies que por sí solas eran bajas, pero al tomarlas en conjunto tienen impacto. Lo que en la imagen 6 no se representa es cuantas especies hay agrupadas en “Otros”, información que se encuentra en el anexo (tablas 5, 6 y 7). Este dato no es menor, ya que al guiarse al visualmente, sugiere que la mayor diversidad se encuentra en las muestras con carga intestinal relativa alta (A). Sin embargo, en las muestras negativas para genes de resistencia (C) y en las muestras con carga intestinal relativa media (B), la media de especies incluidas en “Otros” es mayor que para el caso A.

A su vez, otra observación importante es que la mayoría de las muestras tiene muchas lecturas identificadas como *Escherichia coli*. Al momento de hacer un estudio de microbioma de este tipo y querer sacar conclusiones, es importante contextualizar la epidemiología local. En este caso, las muestras fueron tomadas del HULP, Hospital ubicado en la ciudad de Madrid, España. El estudio se centró en enzimas carbapenemasas, pero existen muchas otras como las  $\beta$ -lactamasas de espectro extendido (BLEE) que son prevalentes en esta región y en la población de estudio (Villa, 2014). Especialmente *E. coli* productora de BLEE con el gen *blaCTX-M-1-Family* en tanto entornos hospitalarios como comunitarios en España (López-Siles, 2023). Por lo tanto, la estrategia planteada no encuentra lo que no busca, es una limitación de detectar genes de resistencia a antibióticos con una qPCR.

Adicionalmente, dejando de lado lo que visualmente se observa en la imagen 6, resultados relevantes fueron los obtenidos con la prueba de Student. Ya que un visualmente puede relacionar las alturas de la imagen 6 y la cantidad de colores con la diversidad, pero con la prueba se buscaron diferencias significativas entre las categorías y lo índices de Shannon. Si la hipótesis inicial se cumpliría, debería haber diferencias significativas entre los 3 grupos. De hecho, la media de los índices de Shannon debería ser menor en el caso de muestras con carga intestinal alta y la media de las muestras negativas para genes de resistencia a carbapenemasas debería ser la mayor. Esto no solo no ocurre si no que no hay diferencias significativas, ya que las medias son muy similares entre sí. Una posible explicación es la toma de antibióticos por parte de los pacientes que ha resultado en la dominancia intestinal por pocas especies resistentes a estos antibióticos, resultando en poca diversidad, aunque estas no tienen genes de carbapenemasas (pero esto no tiene que decir que no tendrán otros genes de resistencia a antibióticos).

A su vez, la disparidad de la cantidad muestras en los diferentes grupos etarios es una limitación de la estrategia de trabajo. Siendo los grupos por encima de 5 años (de 5 a 11 y mayores de 12) con una baja representatividad.

En conclusión, este estudio exploratorio y descriptivo no se ha detectado diferencias significativas entre las cargas intestinales de carbapenemasas y la diversidad del microbioma intestinal. Pero, ha sido un estudio que da lugar a futuras investigaciones, donde sería interesante adicionar una detección por qPCR para genes de resistencia a otros tipos de antibióticos, teniendo en cuenta la epidemiología local. Como sumar una detección del gen *blaCTX-M-1-Family* considerando que la

mayoría de las muestras analizadas presentan muchas lecturas dominadas por *E. coli*. También sería interesante aumentar el número de muestras analizadas con distribuciones equitativas entre los diferentes grupos etarios. Esto permitiría aumentar la complejidad de los análisis estadísticos a realizar y conclusiones del estudio más contundentes.

Este trabajo se alinea con la Agenda 2030 para el Desarrollo Sostenible de las Naciones Unidas. Específicamente con el Objetivo de Desarrollo Sostenible (ODS) 3 de Salud y Bienestar. El trabajo contribuye conocimiento y experiencia para a futuro poder predecir cambios y efectos en el microbioma producidos, por ejemplo, por el consumo de antibióticos. El consumo de antibióticos y su efecto en el microbioma intestinal de pediátricos tiene impacto crucial en la salud y bienestar de esta población.

## 7. Conclusiones.

En base a los objetivos del presente estudio y los resultados obtenidos se concluye que:

- La relación entre la diversidad microbiana y las cargas intestinales relativas no coincidieron con la hipótesis planteada por la bibliografía en 46 muestras de torundas rectales obtenidos de pacientes pediatricos.
- La presencia de *Escherichia coli* en conjunto con el contexto de la epidemiología local de BLEE podría sugerir que la diversidad del microbioma esta influenciada por otros genes de resistencia a antibióticos no cuantificados en este estudio.
- Las limitaciones en el tamaño muestral por grupo etario impiden establecer correlaciones conclusivas entre los grupos etarios y la diversidad del microbioma.
- Para poder implementar a futuro estrategias que predigan cambios precisos en el microbioma sería interesante aumentar el tamaño de muestras y analizar otros genes de resistencia a antibióticos.

## 8. Bibliografía.

Abdel-Aziz, M. I., Hashimoto, S., Neerincx, A. H., Haarman, E. G., Cecil, A., Lintelmann, J., Witting, M., Hauck, S. M., Kerssemakers, N., Verster, J. C., Bang, C., Franke, A., Dierdorp, B. S., Dekker, T., Metwally, N. K. A., Duitman, J. W., Lutter, R., Gorenjak, M., Toncheva, A. A., Kheiroddin, P., ... SysPharmPediA consortium (2025). Metabotypes are linked to uncontrolled childhood asthma, gut microbiota, and systemic inflammation. *The Journal of allergy and clinical immunology*, S0091-6749(25)00457-9. Advance online publication.

<https://doi.org/10.1016/j.jaci.2025.04.017>

Andrews, S. (n.d.). *FastQC A Quality Control tool for High Throughput Sequence Data*. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y., Xie, H., Zhong, H., Khan, M. T., Zhang, J., Li, J., Xiao, L., Al-Aama, J., Zhang, D., Lee, Y. S., Kotowska, D., Colding, C., Tremaroli, V., ... Wang, J. (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell host & microbe*, 17(6), 852.

<https://doi.org/10.1016/j.chom.2015.05.012>

Bassis, C. M., Moore, N. M., Lolans, K., Seekatz, A. M., Weinstein, R. A., Young, V. B., Hayden, M. K., & CDC Prevention Epicenters Program (2017). Comparison of stool versus rectal swab samples and storage conditions on bacterial community profiles. *BMC microbiology*, 17(1), 78. <https://doi.org/10.1186/s12866-017-0983-9>

Bokulich, N. A., Chung, J., Battaglia, T., Henderson, N., Jay, M., Li, H., D Lieber, A., Wu, F., Perez-Perez, G. I., Chen, Y., Schweizer, W., Zheng, X., Contreras, M., Dominguez-Bello, M. G., & Blaser, M. J. (2016). Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Science translational medicine*, 8(343), 343ra82.

<https://doi.org/10.1126/scitranslmed.aad7121>

Budding, A. E., Grasman, M. E., Eck, A., Bogaards, J. A., Vandenbroucke-Grauls, C. M., van Bodegraven, A. A., & Savelkoul, P. H. (2014). Rectal swabs for analysis of the intestinal microbiota. *PLoS one*, 9(7), e101344. <https://doi.org/10.1371/journal.pone.0101344>

Cha, T., Kim, H. H., Keum, J., Kwak, M. J., Park, J. Y., Hoh, J. K., Kim, C. R., Jeon, B. H., & Park, H. K. (2023). Gut microbiome profiling of neonates using Nanopore MinION and Illumina MiSeq sequencing. *Frontiers in microbiology*, 14, 1148466.

<https://doi.org/10.3389/fmicb.2023.1148466>

Cuccuru, G., Orsini, M., Pinna, A., Sbardellati, A., Soranzo, N., Travaglione, A., Uva, P., Zanetti, G., & Fotia, G. (2014). Orione, a web-based framework for NGS analysis in microbiology. *Bioinformatics (Oxford, England)*, 30(13), 1928–1929.  
<https://doi.org/10.1093/bioinformatics/btu135>

Dahdouh, E., Fernández-Tomé, L., Cendejas-Bueno, E., Ruiz-Carrascoso, G., Schüffelmann, C., Alós-Díez, M., Lázaro-Perona, F., Castro-Martínez, M., Escosa-García, L., Jiménez-Rodríguez, S., Hierro-Llanillo, L., & Mingorance, J. (2022). Intestinal Dominance by Multidrug-Resistant Bacteria in Pediatric Liver Transplant Patients. *Microbiology spectrum*, 10(6), e0284222.  
<https://doi.org/10.1128/spectrum.02842-22>

Dahdouh, E., Cendejas-Bueno, E., Ruiz-Carrascoso, G., Schüffelmann, C., Lázaro-Perona, F., Castro-Martínez, M., Moreno-Ramos, F., Escosa-García, L., Alguacil-Guillén, M., & Mingorance, J. (2023). Intestinal loads of extended-spectrum beta-lactamase and Carbapenemase genes in critically ill pediatric patients. *Frontiers in cellular and infection microbiology*, 13, 1180714. <https://doi.org/10.3389/fcimb.2023.1180714>

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., & Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences of the United States of America*, 107(26), 11971–11975. <https://doi.org/10.1073/pnas.1002601107>

Esberg, A., Fries, N., Haworth, S., & Johansson, I. (2024). Saliva microbiome profiling by full-gene 16S rRNA Oxford Nanopore Technology versus Illumina MiSeq sequencing. *NPJ biofilms and microbiomes*, 10(1), 149. <https://doi.org/10.1038/s41522-024-00634-1>

Galaxy Community (2024). The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic acids research*, 52(W1), W83–W94.  
<https://doi.org/10.1093/nar/gkae410>

Gauffin, T., Tobin, N. H., & Aldrovandi, G. M. (2018). The importance of the microbiome in pediatrics and pediatric infectious diseases. *Current opinion in pediatrics*, 30(1), 117–124.  
<https://doi.org/10.1097/MOP.0000000000000576>

Hammoudi Halat, D., & Ayoub Moubareck, C. (2020). The Current Burden of Carbapenemases: Review of Significant Properties and Dissemination among Gram-Negative Bacteria. *Antibiotics (Basel, Switzerland)*, 9(4), 186. <https://doi.org/10.3390/antibiotics9040186>

Huseyin, C. E., Rubio, R. C., O'Sullivan, O., Cotter, P. D., & Scanlan, P. D. (2017). The Fungal Frontier: A Comparative Analysis of Methods Used in the Study of the Human Gut Mycobiome. *Frontiers in microbiology*, 8, 1432. <https://doi.org/10.3389/fmicb.2017.01432>

Jean, S. S., Lee, W. S., Lam, C., Hsu, C. W., Chen, R. J., & Hsueh, P. R. (2015). Carbapenemase-producing Gram-negative bacteria: current epidemics, antimicrobial susceptibility and treatment options. *Future microbiology*, 10(3), 407–425. <https://doi.org/10.2217/fmb.14.135>

Klein D. (2002). Quantification using real-time PCR technology: applications and limitations. *Trends in molecular medicine*, 8(6), 257–260. [https://doi.org/10.1016/s1471-4914\(02\)02355-9](https://doi.org/10.1016/s1471-4914(02)02355-9)

Korach-Rechtman, H., Hreish, M., Fried, C., Gerassy-Vainberg, S., Azzam, Z. S., Kashi, Y., & Berger, G. (2020). Intestinal Dysbiosis in Carriers of Carbapenem Resistant *Enterobacteriaceae*. *mSphere*, 5(2), e00173-20. <https://doi.org/10.1128/mSphere.00173-20>

Kosakovsky Pond, S., Wadhawan, S., Chiaromonte, F., Ananda, G., Chung, W. Y., Taylor, J., Nekrutenko, A., & Galaxy Team (2009). Windshield splatter analysis with the Galaxy metagenomic pipeline. *Genome research*, 19(11), 2144–2153. <https://doi.org/10.1101/gr.094508.109>

Levene, H. (1960). Robust tests for equality of variances. En I. Olkin (Ed.), *Contributions to probability and statistics: Essays in honor of Harold Hotelling* (pp. 278–292). Stanford University Press.

Li, F., Liu, J., Maldonado-Gómez, M. X., Frese, S. A., Gänzle, M. G., & Walter, J. (2024). Highly accurate and sensitive absolute quantification of bacterial strains in human fecal samples. *Microbiome*, 12(1), 168. <https://doi.org/10.1186/s40168-024-01881-2>

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif.)*, 25(4), 402–408. <https://doi.org/10.1006/meth.2001.1262>

López-Siles, M., Moure, Z., Muadica, A. S., Sánchez, S., Cruces, R., Ávila, A., Lara, N., Köster, P. C., Dashti, A., Oteo-Iglesias, J., Carmena, D., & McConnell, M. J. (2023). Fecal carriage of extended-spectrum beta-lactamase-producing Enterobacteriales in healthy Spanish schoolchildren. *Frontiers in microbiology*, 14, 1035291. <https://doi.org/10.3389/fmicb.2023.1035291>

Migliorini, L. B., Leaden, L., de Sales, R. O., Correa, N. P., Marins, M. M., Koga, P. C. M., Toniolo, A. D. R., de Menezes, F. G., Martino, M. D. V., Mingorance, J., & Severino, P. (2022). The Gastrointestinal Load of Carbapenem-Resistant Enterobacteriaceae Is Associated With the Transition From Colonization to Infection by *Klebsiella pneumoniae* Isolates Harboring the *blaKPC* Gene. *Frontiers in cellular and infection microbiology*, 12, 928578. <https://doi.org/10.3389/fcimb.2022.928578>

Negele, K., Heinrich, J., Borte, M., von Berg, A., Schaaf, B., Lehmann, I., Wichmann, H. E., Bolte, G., & LISA Study Group (2004). Mode of delivery and development of atopic disease during the first 2 years of life. *Pediatric allergy and immunology: official publication of the European Society of Pediatric Allergy and Immunology*, 15(1), 48–54.

<https://doi.org/10.1046/j.0905-6157.2003.00101.x>

Oehler, J. B., Wright, H., Stark, Z., Mallett, A. J., & Schmitz, U. (2023). The application of long-read sequencing in clinical settings. *Human genomics*, 17(1), 73. <https://doi.org/10.1186/s40246-023-00522-3>

Ondov, B. D., Bergman, N. H., & Phillippy, A. M. (2011). Interactive metagenomic visualization in a Web browser. *BMC bioinformatics*, 12, 385. <https://doi.org/10.1186/1471-2105-12-385>

Oxford Nanopore Technologies. (2025). *Ligation sequencing amplicons - Native Barcoding Kit 24 V14 (SQK-NBD114.24)* <https://nanoporetech.com/es/document/ligation-sequencing-amplicons-native-barcoding-v14-sqk-nbd114-24>

Palm, N. W., de Zoete, M. R., Cullen, T. W., Barry, N. A., Stefanowski, J., Hao, L., Degnan, P. H., Hu, J., Peter, I., Zhang, W., Ruggiero, E., Cho, J. H., Goodman, A. L., & Flavell, R. A. (2014). Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell*, 158(5), 1000–1010. <https://doi.org/10.1016/j.cell.2014.08.006>

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>

Rozas, M., Brillet, F., Callewaert, C., & Paetzold, B. (2022). MinION™ Nanopore Sequencing of Skin Microbiome 16S and 16S-23S rRNA Gene Amplicons. *Frontiers in cellular and infection microbiology*, 11, 806476. <https://doi.org/10.3389/fcimb.2021.806476>

Shannon C. E. (1997). The mathematical theory of communication. 1963. *M.D. computing: computers in medical practice*, 14(4), 306–317.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3–4), 591–611. <https://doi.org/10.2307/2333709>

Student. (1908). The probable error of a mean. *Biometrika*, 6(1), 1–25. <https://doi.org/10.2307/2331554>

Suhr, M. J., Banjara, N., & Hallen-Adams, H. E. (2016). Sequence-based methods for detecting and evaluating the human gut mycobiome. *Letters in applied microbiology*, 62(3), 209–215. <https://doi.org/10.1111/lam.12539>

Thänert, R., Schwartz, D. J., Keen, E. C., Hall-Moore, C., Wang, B., Shaikh, N., Ning, J., Rouggly-Nickless, L. C., Thänert, A., Ferreiro, A., Fishbein, S. R. S., Sullivan, J. E., Radmacher, P., Escobedo, M., Warner, B. B., Tarr, P. I., & Dantas, G. (2024). Clinical sequelae of gut

microbiome development and disruption in hospitalized preterm infants. *Cell host & microbe*, 32(10), 1868–1869. <https://doi.org/10.1016/j.chom.2024.09.012>

Touati, A., & Mairi, A. (2020). Carbapenemase-Producing Enterobacterales in Algeria: A Systematic Review. *Microbial drug resistance (Larchmont, N.Y.)*, 26(5), 475–482. <https://doi.org/10.1089/mdr.2019.0320>

Villa, J., Viedma, E., Brañas, P., Mingorance, J., & Chaves, F. (2014). Draft Whole-Genome Sequence of OXA-48-Producing Multidrug-Resistant Klebsiella pneumoniae KP\_ST11\_OXA-48. *Genome announcements*, 2(4), e00737-14. <https://doi.org/10.1128/genomeA.00737-14>

Wood, D. E., & Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome biology*, 15(3), R46. <https://doi.org/10.1186/gb-2014-15-3-r46>

Wu, X., Xia, Y., He, F., Zhu, C., & Ren, W. (2021). Intestinal mycobiota in health and diseases: from a disrupted equilibrium to clinical opportunities. *Microbiome*, 9(1), 60. <https://doi.org/10.1186/s40168-021-01024-x>

Yan, Q., Li, S., Yan, Q., Huo, X., Wang, C., Wang, X., Sun, Y., Zhao, W., Yu, Z., Zhang, Y., Guo, R., Lv, Q., He, X., Yao, C., Li, Z., Chen, F., Ji, Q., Zhang, A., Jin, H., Wang, G., ... Ma, X. (2024). A genomic compendium of cultivated human gut fungi characterizes the gut mycobiome and its relevance to common diseases. *Cell*, 187(12), 2969–2989.e24. <https://doi.org/10.1016/j.cell.2024.04.043>

Yassour, M., Vatanen, T., Siljander, H., Hämäläinen, A. M., Härkönen, T., Ryhänen, S. J., Franzosa, E. A., Vlamakis, H., Huttenhower, C., Gevers, D., Lander, E. S., Knip, M., DIABIMMUNE Study Group, & Xavier, R. J. (2016). Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. *Science translational medicine*, 8(343), 343ra81. <https://doi.org/10.1126/scitranslmed.aad0917>

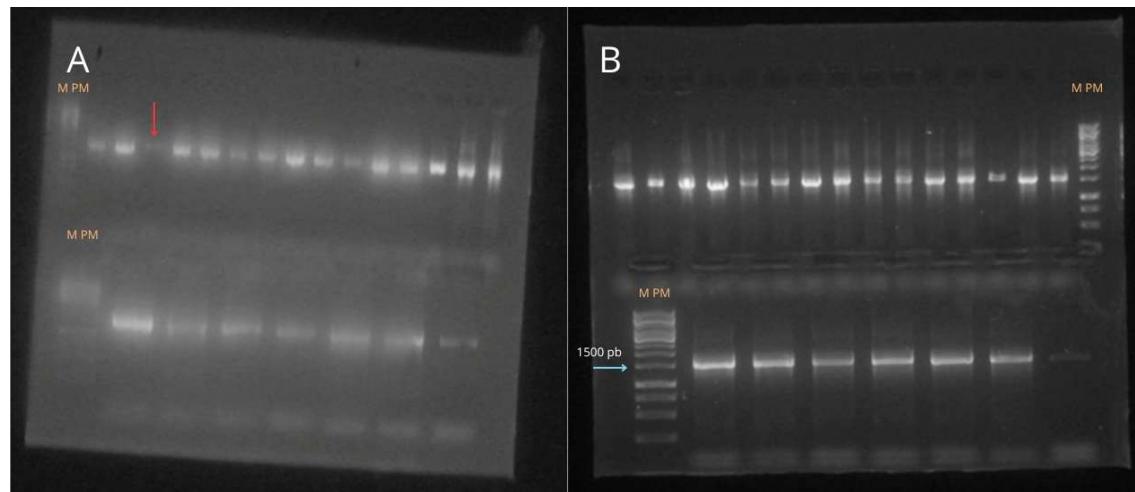
Yen, S., & Johnson, J. S. (2021). Metagenomics: a path to understanding the gut microbiome. *Mammalian genome: official journal of the International Mammalian Genome Society*, 32(4), 282–296. <https://doi.org/10.1007/s00335-021-09889-x>

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of computational biology : a journal of computational molecular cell biology*, 7(1-2), 203–214. <https://doi.org/10.1089/10665270050081478>

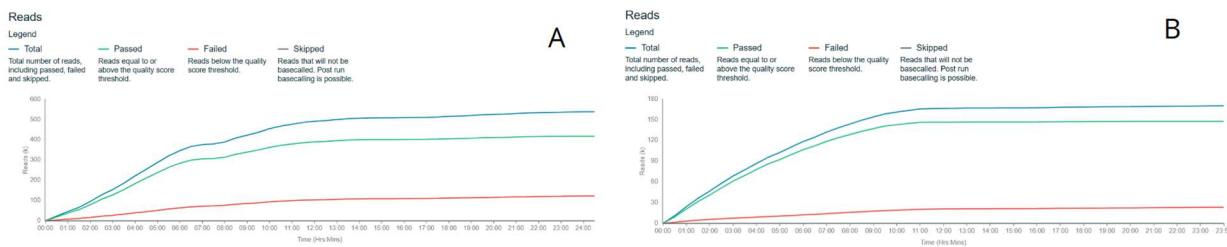
**9. Anexo.**

**Tabla nº1 Anexo.** Listado de las muestras procesadas y a que categoría pertenecen según la clasificación realizada.

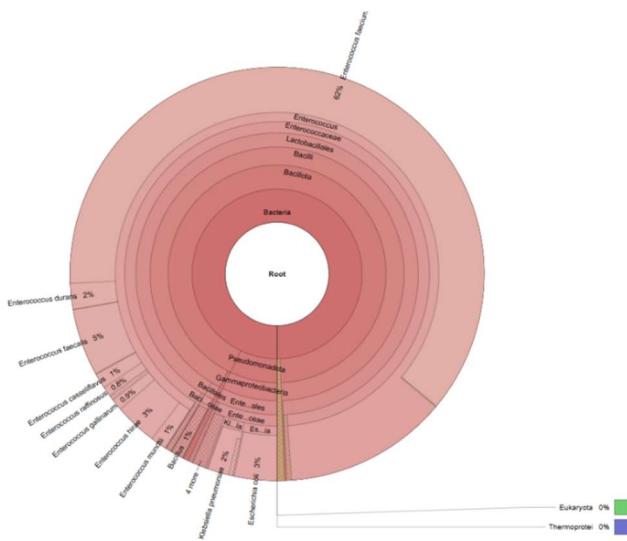
Muestras con carga alta para para genes de resistencia a carbapenémicos	Muestras con carga media para para genes de resistencia a carbapenémicos	Muestras negativas para para genes de resistencia a carbapenémicos
Mb4, Mb5, Mb13, Mb14, Mb15, Mb76, Mb79, Mb130, Mb136, Mb193, Mb254, Mb308, Mb389 y Mb390.	Mb1, Mb77, Mb84, Mb85, Mb188, Mb190, Mb256 y Mb257.	Mb82, Mb355, Mb356, Mb357, Mb358, Mb359, Mb360, Mb361, Mb362, Mb363, Mb364, Mb365, Mb366, Mb367, Mb368, Mb369, Mb370, Mb371, Mb372, Mb373, Mb374, Mb375, Mb376 y Mb378.
Total de muestras con carga: 22		Total negativas: 24



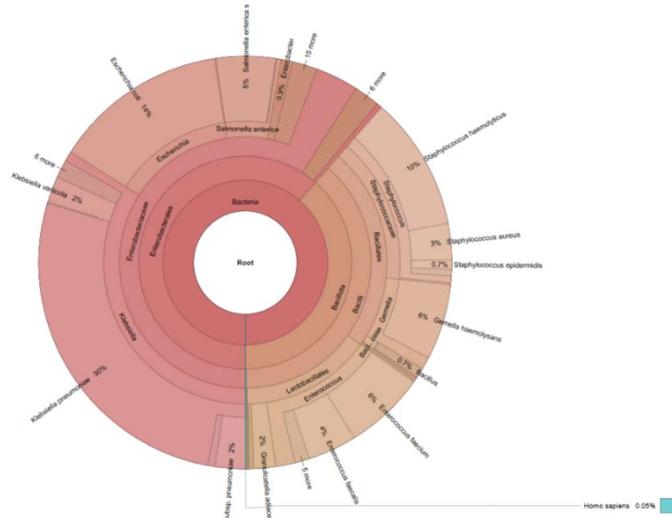
**Imagen 1 Anexo.** Geles de agarosa. En ambos geles, tanto en la parte superior como inferior se encuentra el marcador de peso molecular (marcado como “M PM”). En el resto de los pocillos se observa un solo producto del tamaño deseado (flecha azul, 1500pb), excepto en el control negativo (última posición de ambos geles). A. Muestras negativas para genes de resistencia a carbapenémicos, la flecha roja marca una muestra que no amplificó. B. Muestras positivas para genes de resistencia a carbapenémicos. Vale aclarar que no se sembraron en estos geles todas las muestras procesadas ya que hay 24 posiciones en ambos, pero 4 de ellas se ocuparon con los marcadores de peso molecular.



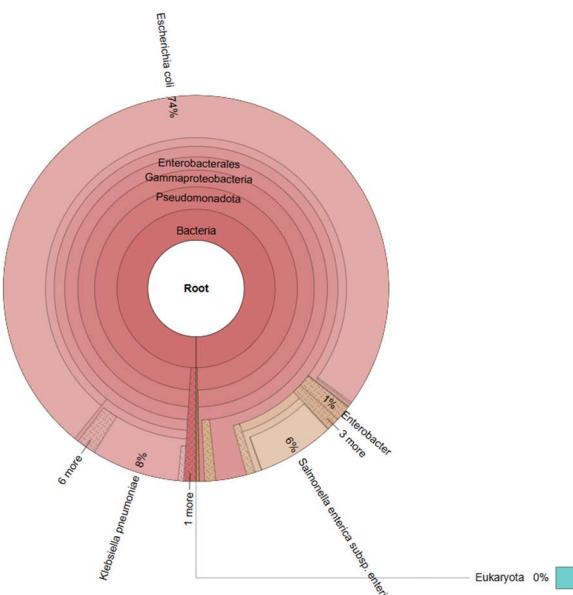
**Imagen 2 Anexo.** Grafica de las lecturas que han pasado y no han pasado el filtro de calidad. A. Gráfico para el *run* 1. B. Gráfico para el *run* 2.



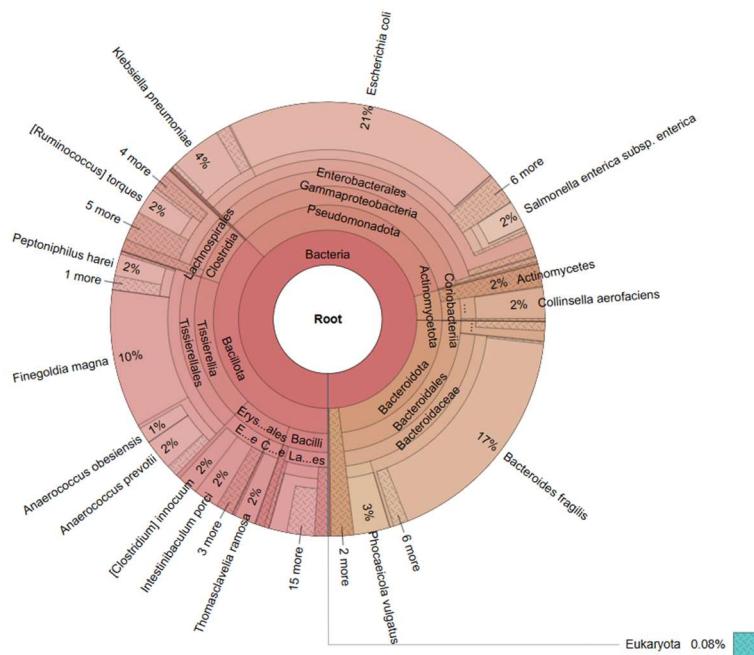
**Imagen 3 Anexo.** Resultado de la representación de la composición taxonómica con Krona de la muestra Mb4 (categoría: carga alta para genes de resistencia a carbapenémicos).



**Imagen 4 Anexo.** Resultado de la representación de la composición taxonómica con Krona de la muestra Mb14 (categoría: carga alta para genes de resistencia a carbapenémicos).



**Imagen 5 Anexo.** Resultado de la representación de la composición taxonómica con Krona de la muestra Mb355 (categoría: negativa para genes de resistencia a carbapenémicos).



**Imagen 6 Anexo.** Resultado de la representación de la composición taxonómica con Krona de la muestra Mb373 (categoría: negativa para genes de resistencia a carbapenémicos).

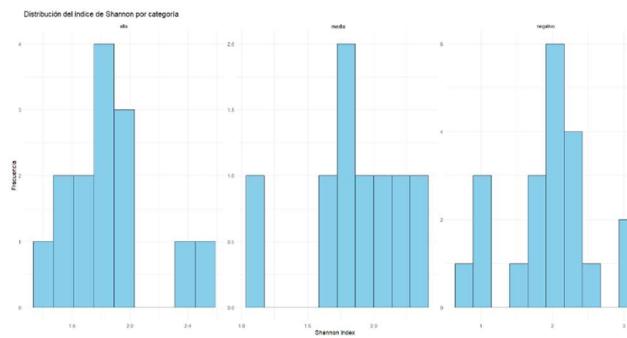
**Tabla nº2 Anexo.** Listado de cada muestra con la categoría a la cual pertenecen, índice de Shannon, sexo y edad.

Código muestra	Índice de Shannon	Categoría a la cual pertenece	Sexo	Edad
Mb4	1.33	alta	M	3
Mb5	1.73	alta	F	1
Mb13	1.76	alta	F	1
Mb14	2.37	alta	M	3
Mb15	1.49	alta	M	1
Mb76	1.88	alta	F	0
Mb79	2.59	alta	M	3
Mb 130	1.88	alta	F	1
Mb 136	1.55	alta	M	14
Mb 193	1.78	alta	M	2
Mb 254	1.96	alta	F	0
Mb 308	1.74	alta	M	3
Mb389	2	alta	M	17
Mb390	1.94	alta	M	0
Mb1	1.84	media	M	1
Mb77	1.63	media	M	2
Mb84	1.09	media	M	0

Mb85	1.74	media	M	17
Mb 188	1.95	media	M	0
Mb 190	2	media	M	1
Mb 256	2.33	media	M	1
Mb 257	2.14	media	M	0
Mb82	2.27	negativo	M	0
Mb355	1.05	negativo	F	1
Mb356	2.23	negativo	F	0
Mb357	1.1	negativo	M	0
Mb359	1.13	negativo	F	15
Mb360	1.7	negativo	F	0
Mb361	2.11	negativo	M	9
Mb362	2.67	negativo	M	4
Mb363	1.43	negativo	F	1
Mb364	1.85	negativo	F	0
Mb366	1.81	negativo	F	2
Mb367	1.96	negativo	M	4
Mb368	1.93	negativo	F	0
Mb369	0.73	negativo	F	11
Mb370	1.93	negativo	M	0
Mb371	2.97	negativo	F	6
Mb372	2.19	negativo	M	0
Mb373	3.02	negativo	F	9
Mb374	2.01	negativo	F	9
Mb376	2.33	negativo	F	0
Mb378	1.94	negativo	F	1

**Tabla nº3 Anexo.** Resultados de la prueba Shapiro-Wilk.

Categoría	n (total de muestras por categoría)	P_valor	W_statistic
alta	14	0.379	0.937
media	8	0.638	0.943
negativo	21	0.519	0.960



**Imagen 7 Anexo.** Distribución de los índices de Shannon en las 3 categorías, alta, media y negativo respectivamente.

**Tabla nº4 Anexo.** Resultados de la prueba de Student, realizado con RStudio.

Comparaciones	t	df	P_valor	Media del grupo 1	Media del grupo 2	IC_95_bajo	IC_95_alto
Alta vs Media	0.11	20	0.91	1.86	1.84	-0.30	0.34
Alta vs Negativo	-0.37	33	0.71	1.86	1.92	-0.42	0.29
Media vs Negativo	-0.36	27	0.72	1.84	1.92	-0.55	0.39

**Tabla nº5 Anexo.** Cantidad de especies que fueron agrupadas en “Otros” en cada muestra de la imagen 5 A.

Muestras	Cantidad de especies resumidas en “Otros”
Mb4	49
Mb5	65
Mb13	69
Mb14	91
Mb15	61
Mb76	60
Mb79	106
Mb130	76
Mb136	52
Mb193	68
Mb254	95

Mb308	77
Mb389	73
Mb390	77

**Tabla nº6 Anexo.** Cantidad de especies que fueron agrupadas en “Otros” en cada muestra de la imagen 5 B.

Muestras	Cantidad de especies resumidas en “Otros”
Mb1	75
Mb77	77
Mb84	58
Mb85	92
Mb188	81
Mb190	75
Mb256	111
Mb257	97

**Tabla nº7 Anexo.** Cantidad de especies que fueron agrupadas en “Otros” en cada muestra de la imagen 5 C.

Muestras	Cantidad de especies resumidas en “Otros”
Mb82	99
Mb355	56
Mb356	110
Mb357	56
Mb359	69
Mb360	79
Mb361	74
Mb362	97
Mb363	79
Mb364	68
Mb366	68

Mb367	110
Mb368	86
Mb369	51
Mb370	92
Mb371	113
Mb372	71
Mb373	142
Mb374	82
Mb376	151
Mb378	98