

**UNIVERSIDAD INTERNACIONAL SEK**

**FACULTAD DE CIENCIAS DE LA SALUD**

Trabajo de Fin de Máster Titulado:

**Estandarización de pipelines para la identificación de genes de resistencia a los antimicrobianos a partir de metagenomas con lecturas obtenidas por NGS**

Realizado por:

**BRIGITTE ARACELLY GARZÓN PROAÑO**

Director del proyecto:

**Ing. Andrés Herrera, MSc.**

Como requisito para la obtención del título de:

**MAGISTER EN BIOMEDICINA**

Quito, 26 de marzo del 2025

## **DECLARACIÓN JURAMENTADA**

Yo, BRIGHITTE ARACELLY GARZÓN PROANO, con cédula de identidad # 172326939-3, declaro bajo juramento que el trabajo aquí desarrollado es de mi autoría, que no ha sido previamente presentado para ningún grado a calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración, cedo mis derechos de propiedad intelectual que correspondan relacionados a este trabajo, a la UNIVERSIDAD INTERNACIONAL SEK, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normativa institucional vigente.

---

**Brigitte Aracelly Garzón Proaño**  
C.I. 1723269393

## **DECLARACIÓN DEL DIRECTOR DE TESIS**

Declaro haber dirigido este trabajo a través de reuniones periódicas con el estudiante, orientando sus conocimientos y competencias para un eficiente desarrollo del tema escogido y dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación.

---

**Ing. Andrés Herrera, MSc.**

Los Profesores Informantes:

**José Rubén Ramirez Iglesias, PhD**

**Rubén Alexander Maldonado Orbe, MSc.**

Después de revisar el trabajo presentado, lo han calificado como apto para su defensa oral ante el tribunal examinador.

---

**José Rubén Ramirez  
Iglesias, PhD**

---

**Rubén Alexander  
Maldonado Orbe, MSc.**

Quito, 26 de marzo del 2025

## **DECLARACIÓN DE AUTORÍA DEL ESTUDIANTE**

Declaro que este trabajo es original, de mi autoría, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales que protegen los derechos de autor vigentes.

---

**Brigitte Aracelly Garzón Proaño**

C.I. 1723269393

## **DEDICATORIA**

A Dios, fuente inagotable de amor, sabiduría y fortaleza, y a la Virgen María cuya intercesión y protección me han acompañado a lo largo de mi vida, especialmente en esta etapa.

A mis amados padres y hermano, cuyo amor, apoyo incondicional y enseñanzas han sido la base de mi formación. Su ejemplo de esfuerzo y dedicación me ha guiado en cada paso de este camino. Este logro también es suyo.

## **AGRADECIMIENTO**

A Dios, cuya infinita bondad y sabiduría han sido mi luz y guía en este camino. Su fortaleza me ha sostenido en los desafíos, y su amor me ha dado la confianza para seguir adelante. A Él dedico cada esfuerzo y cada.

Agradezco profundamente a mis padres, Marlene y Fabián, por su amor incondicional, su paciencia y su guía, que me han permitido crecer y avanzar. A mi hermano, Santiago, mi compañero de vida, por su apoyo, sabios consejos y por cuidarme siempre. Gracias a ustedes, por ser mi fortaleza.

A Cristian, por su constante aliento, comprensión y por estar a mi lado en cada etapa de este camino. Tu apoyo me ha dado fuerzas para seguir adelante, y tu compañía ha sido un motor fundamental en mi crecimiento.

A todos mis docentes de la maestría en Biomedicina por los valiosos conocimientos impartidos y por su dedicación a la formación académica. En especial, quiero expresar mi gratitud a Andrés Herrera, por ser un tutor excepcional, cuya orientación y apoyo han sido fundamentales en mi crecimiento profesional a lo largo de este proceso.

A mis compañeros de la maestría, María José, Mylena, Juan y Anthony, por la amistad que construimos durante este tiempo. Su apoyo, compañerismo y los momentos compartidos fueron esenciales no solo para mi crecimiento académico, sino también para enriquecer esta experiencia con nuevas perspectivas y recuerdos que siempre llevaré conmigo.

A mis amigos de BioMol Ecuador por su apoyo y comprensión durante el tiempo que duró mi maestría. Su motivación y colaboración han sido fundamentales para equilibrar mis responsabilidades laborales y educativas.

Artículo de tesis

# Estandarización de pipelines para la identificación de genes de resistencia a los antimicrobianos a partir de metagenomas con lecturas obtenidas por NGS

Brigitte Garzón<sup>1</sup>, Andrés Herrera (Tutor)<sup>2</sup>

<sup>1</sup> Maestría en Biomedicina, Fac Cs Salud, UISEK; brigritte.garzon@uisek.edu.ec

<sup>2</sup> Health Sciences Faculty, Universidad Internacional SEK (UISEK), Quito, Ecuador; manuel.herrera@uisek.edu.ec

**Resumen:** El incremento mundial de resistencia a antimicrobianos exige diagnósticos más rápidos y precisos. Actualmente, la Secuenciación de Siguiete Generación (NGS) ha permitido la detección integral de genes de resistencia a antimicrobianos (ARGs), superando las limitaciones de la PCR, pero se ha demostrado que requiere complementarla con bioinformática. El presente estudio permite estandarizar pipelines bioinformáticos para determinar el método más eficiente para mejorar la sensibilidad y precisión en la identificación de ARGs a partir de metagenomas y genomas con lecturas obtenidas por NGS. Se usaron datos de metagenomas y genomas secuenciados con tecnologías Oxford Nanopore Technologies (ONT) e Illumina desde NCBI SRA, y se analizaron en Galaxy Trkr. En el caso de ONT se realizó un análisis de calidad con NanoPlot y Porechop, ensamblaje con MEGAHIT y clasificación con MetaBAT2. En ILLUMINA, se usaron las herramientas Trimmomatic, ensamblaje con metaSPAdes y binning con MetaBAT2. La identificación de ARGs se realizó con la herramienta ABRICATE y la base de datos ResFinder, para la clasificación taxonómica se empleó Kraken2 y se visualizó por Krona Pie Chart. También se evaluó la calidad del ensamblaje mediante Quast. Se detectaron ARGs en todos los experimentos, evidenciando variación en la cantidad y tipos según el origen clínico, destacando genes de importancia clínica como *mecA*, *blaTEM*, *blaOXA* y *fosA*, asociado a antibióticos de primera línea. De igual manera, el análisis Quast mostró diferencias en la cobertura y fragmentación del ensamblaje entre ambas tecnologías de secuenciación. En conclusión, los pipelines estandarizados permitieron la identificación precisa y eficiente de ARGs en metagenomas y genomas de NGS, evidenciando la diversidad genética y taxonómica de muestras clínicas. Los datos de los experimentos de ILLUMINA fueron más robustos en términos de ensamblaje, mientras que ONT, a pesar de ser útil para lecturas largas, presentaron problemas en el ensamblaje y cobertura.

**Palabras clave:** ARGs, NGS, Bioinformática, Pipeline, Ensamblaje, Metagenomas, Lecturas, ILLUMINA, ONT

**Abstract:** The global rise in antimicrobial resistance demands faster and more accurate diagnostic approaches. Next-Generation Sequencing (NGS) has enabled comprehensive detection of antimicrobial resistance genes (ARGs), surpassing the limitations of PCR; however, it has been demonstrated that bioinformatics is essential to complement its analysis. This study aims to standardize bioinformatics pipelines to determine the most efficient method for improving the sensitivity and accuracy of ARG identification from metagenomes and genomes with sequencing reads obtained through NGS. Metagenomic and genomic data sequenced using Oxford Nanopore Technologies (ONT) and Illumina were retrieved from the NCBI SRA and analyzed in Galaxy Trkr. For ONT data, quality assessment was performed using NanoPlot and Porechop, assembly was conducted with MEGAHIT, and binning was carried out with MetaBAT2. For Illumina data, Trimmomatic was used for quality control, metaSPAdes for assembly, and MetaBAT2 for binning. ARG identification was conducted using ABRICATE with the ResFinder database, while taxonomic classification was performed with Kraken2 and visualized using Krona Pie Chart. Assembly quality was further evaluated using Quast. ARGs were detected across all experiments, revealing variations in the quantity and types of ARGs depending on the clinical origin, with clinically significant genes such as *\*mecA\**, *\*blaTEM\**, *\*blaOXA\**, and *\*fosA\** associated with first-line antibiotics. Additionally, the Quast analysis indicated differences in assembly coverage and fragmentation between sequencing technologies. In conclusion, the standardized pipelines enabled accurate and efficient identification of ARGs in metagenomes and NGS-derived genomes, highlighting the genetic and taxonomic diversity of clinical samples. Illumina data provided more robust assemblies, whereas ONT, despite its advantages in long-read sequencing, exhibited challenges in assembly and coverage.

**Keywords:** ARGs, NGS, Bioinformatics, Pipeline, Assembly, Metagenomes, Reads, ILLUMINA, ONT.

## 1. Introducción

La resistencia microbiana a los antibióticos ocurre cuando las bacterias desarrollan la capacidad de sobrevivir a los medicamentos diseñados para erradicarlas. Las principales causas son: la mutación genética de los microorganismos y el uso inadecuado e irracional de medicamentos antimicrobianos (A. Gupta et al., 2019). Los mecanismos principales mediante los cuales las bacterias desarrollan resistencia incluyen: la producción de enzimas capaces de degradar el antibiótico, la modificación del sitio de unión y blancos moleculares del antimicrobiano, la expulsión activa del fármaco mediante bombas de eflujo antes de que se acumulen concentraciones letales para la célula, cambios en la permeabilidad de la membrana, el desarrollo de vías metabólicas alternativas y la modificación intracelular del antibiótico para su inactivación (Abushaheen et al., 2020; Balderrama-González et al., 2021; Hasan & Al-Harmoosh, 2020; Y. Liu et al., 2019; Schaenzer & Wright, 2020; Varela et al., 2021).

Acorde a cifras recolectadas, en 76 países, el 42% de *Escherichia coli* desarrollaron resistencia a cefalosporinas de tercera generación y el 35% de *Staphylococcus aureus* a la meticilina. De igual manera, en el 2020 se informó que 1 de cada 5 pacientes con enfermedades del tracto urinario causadas por *Escherichia coli* mostraron susceptibilidad reducida a antibióticos estándar como ampicilina, cotrimoxazol y fluoroquinolonas. Además, se ha reportado que la bacteria *Klebsiella pneumoniae* posee un nivel superior de resistencia y ha conducido a un mayor uso de antibióticos de último recurso como los carbapenémicos, lo cual ha comprometido su eficacia y aumentado la tasa de infecciones sin tratamiento adecuado (World Health Organization, 2023). En base a lo anterior, acorde a la Organización para la Cooperación y el Desarrollo Económicos (OECD), para el 2035 se estima el doble de la resistencia a los antibióticos de último recurso (World Health Organization, 2022). En relación al Ecuador, el Instituto Nacional de Investigación en Salud Pública (INSPI) reportó que *E. coli* aislado de muestras de pacientes con infecciones en vías urinarias han presentado resistencia a ampicilina (78.8%), ciprofloxacina (62.3%) y gentamicina (33.5%). De igual manera, en regiones rurales del país, de 92 aislados bacterianos de pacientes con heridas infectadas con *S. aureus* resistentes a meticilina (MRSA), el 44.7% se identificaron como oxacilino-resistentes (Ross et al., 2020).

En base a lo anterior existe la necesidad de mejorar la precisión, sensibilidad y velocidad del diagnóstico de las infecciones microbianas. Se conoce que, por cada hora de demora en brindar un tratamiento apropiado de la sepsis relacionada a una infección bacteriana que se propaga a través del torrente sanguíneo disminuye la tasa de supervivencia del paciente en un 7.6% (Sheka et al., 2021a). Además, el diagnóstico clínico erróneo para infecciones bacterianas ha provocado que el 42% de los pacientes presenten los síntomas por largos periodos de tiempo lo cual ha comprometido su calidad de vida (Sheka et al., 2021b). En ambientes hospitalarios, residencias de ancianos y centros de atención para personas inmunocomprometidas donde las infecciones relacionadas pueden ser letales, se ha informado la prevalencia de bacterias multirresistentes (MDR) debido a que presentan resistencia a al menos tres clases de antimicrobianos. Los géneros de principal preocupación son *Acinetobacter*, *Pseudomonas* y *Enterobacteriaceae* (incluidos *Klebsiella pneumoniae*, *Escherichia coli*, *Enterobacter spp.*, *Serratia spp.*, *Proteus spp.*, *Providencia spp.* y *Morganella spp.*) (Williams et al., 2022).

En los laboratorios clínicos, la PCR es una técnica rutinaria de diagnóstico para infecciones microbianas que se caracteriza por su velocidad para identificar patógenos, baja complejidad y costos (Wenger et al., 2019). Sin embargo, se ha reportado que mediante esta técnica molecular se puede determinar si está presente un mecanismo particular de resistencia, pero no es posible establecer la existencia de otros mecanismos en el mismo antibiótico. De igual forma, este método molecular se puede usar únicamente para mutaciones y mecanismos de resistencia previamente detectados y caracterizados. Además, esta prueba molecular puede sobreestimar la presencia de resistencia, ya que no analizan la expresión del gen relevante (Tan, 2003). También se ha reportado que no permite tener información específica sobre la cepa y la resistencia a antibióticos de la bacteria (Wenger et al., 2019). Respecto a las muestras clínicas de alta complejidad se ha reportado que tienen una carga variable de patógenos, flora bacteriana comensal y alta proporción de ácidos nucleicos huésped: patógeno, lo cual impide un diagnóstico oportuno de enfermedades infecciosas y compromete la efectividad de tratamientos médicos (Pendleton et al., 2017).

Un estudio prospectivo ha demostrado que una alternativa viable para eludir las limitaciones en el diagnóstico de bacterias resistentes a antibióticos es el análisis de secuenciación metagenómica (mNGS) (Tang et al., 2021). Es una técnica que permite secuenciar y analizar directamente el material genético de una comunidad microbiana completa extraída de una muestra compleja, proporcionando información exhaustiva sobre su composición y diversidad. En comparación con los métodos tradicionales, mNGS se ha caracterizado porque permiten un diagnóstico rápido, no dependen del cultivo de comunidades microbianas, y proporcionan información sobre todo el contenido del genoma microbiano (Yonkus et al., 2022). De igual forma, mNGS permite la caracterización de resistomas bacterianos y la identificación tanto de genes nuevos y conocidos con resistencia a antibióticos presentes en distintas bacterias (Cheng et al., 2022).

Actualmente existen diversas plataformas de Secuenciación de Siguiete Generación con distintas tecnologías y que han demostrado brindar resultados eficientes para mNGS. Mediante el flujo de trabajo de Oxford Nanopore Technology (ONT) se ha demostrado que a partir de muestras clínicas complejas de bacterias es posible secuenciar genomas de una forma consistente, en tiempo real y empleando lecturas largas. Sin embargo, la amplificación de librerías de ONT presenta un rendimiento bajo, dificulta identificar patógenos con pocas copias, y disminuye la sensibilidad en la detección de patógenos y genes con resistencia antimicrobiana (Sheka et al., 2021a). Así también, Illumina se caracteriza por el uso de tecnología por síntesis, sin embargo, su limitación es el uso de lecturas cortas lo

cual dificulta asociar genes de resistencia antimicrobiana con especies bacterianas individuales, además requiere un proceso de amplificación de la biblioteca y sus costos de implementación son elevados (Tausch et al., 2018).

Dentro del área de salud, es importante reportar datos de secuenciación de alta calidad ya que hace más eficiente la identificación de cepas patógenas y el análisis de información genética sobre genes como la resistencia a antibióticos. Sin embargo, el procesamiento de los datos depende de herramientas bioinformáticas difíciles de usar, y por lo tanto no se pueden aprovechar al máximo para analizar todas las lecturas obtenidas a partir de NGS (Oakeson et al., 2017b). Illumina ha desarrollado la solución BaseSpace Sequence Hub que permite a los usuarios monitorear el progreso de secuenciación, compartir datos y almacenarlos, y realizar análisis primarios, secundarios y terciarios de datos. Sin embargo, esta herramienta es gratuita por treinta días y a partir de ese tiempo hay que pagar una licencia, además que únicamente es compatible con secuenciadores Illumina (Oakeson et al., 2017a). Caso contrario, Galaxy es una plataforma de código abierto para la salud pública que permite realizar una amplia gama de análisis bioinformáticos. Es especialmente conocida por su flexibilidad y facilidad de uso, permitiendo a los investigadores sin experiencia en programación llevar a cabo análisis complejos. Algunas de las aplicaciones que destacan son control de la calidad de secuencia, ensamblajes de genomas completos, análisis de datos multi-ómicos, seguimiento y reproducción de análisis, e identificación de genes de resistencia (Afgan et al., 2016).

En base a lo anterior, se plantea que un flujo de trabajo automatizado aumentará la precisión y reducirá los tiempos de detección de genes de resistencia a antibióticos en microbiomas. Motivo por el cual, el objetivo del presente estudio, busca determinar el método más óptimo para incrementar la sensibilidad y eficiencia en la identificación de genes con resistencia a antibióticos a partir de lecturas obtenidas de metagenomas con tecnologías ONT e Illumina.

## 2. Materiales y Métodos

- **Obtención de Datos**

En la base de datos de Archivos de Lectura de Secuencias (SRA) de la plataforma *National Center for Biotechnology Information* (NCBI) se realizó la búsqueda de los números de acceso de experimentos (SRX) de metagenomas y genomas con lecturas obtenidas mediante las tecnologías de secuenciación *Oxford Nanopore Technology* (ONT) e Illumina. Los metadatos correspondientes a cada SRX, se presentan detalladamente en la Tabla No. 1.

**Tabla No.1** Números de Acceso SRA y Metadatos de las Lecturas de Metagenomas y Genomas obtenidos por ONT e Illumina

Tecnología de Secuenciación	NCBI							
	SRA	Estrategia	Organismo	Recurso	Entorno	Huésped	Fuente de Aislamiento	RAM
ONT	SRX24373965	WGS	<i>Staphylococcus aureus</i>	Genómico	Hospital Militar, Vietnam	Humano	Muestra respiratoria	Sí
	SRX20109754	WGS	<i>Enterobacter cloacae</i>	Genómico	Hospital General, Bangladesh	Humano	Cultivo celular bacteriano	Sí
	SRX20109753	WGS	<i>Enterobacter hormaechei</i>	Genómico	Hospital General, Bangladesh	Humano	Cultivo celular bacteriano	Sí
	SRX14661891	WGS	NA	Metagenómico	Hospital General, China	Humano	Lavado bronqueoalveolar	Sí
	SRX14661871	WGS	NA	Metagenómico	Mock	Humano	No aplica	Sí
	SRX17062755	WGS	NA	Metagenómico	Hospital General, China	Humano	Lavado bronqueoalveolar	Sí
	SRX17062754	WGS	NA	Metagenómico	Hospital General, China	Humano	Lavado bronqueoalveolar	Sí
	SRX14661895	WGS	NA	Metagenómico	Hospital General, China	Humano	Lavado bronqueoalveolar	Sí
	SRX17062753	WGS	NA	Metagenómico	Hospital General, China	Humano	Lavado bronqueoalveolar	Sí
	SRX17062752	WGS	NA	Metagenómico	Hospital General, China	Humano	Lavado bronqueoalveolar	Sí
	SRX20504811	WGS	<i>Enterococcus faecium</i>	Genómico	Hospital General, China	Humano	Cultivo de sangre	Sí
SRX7585726	WGS	<i>Klebsiella pneumoniae</i>	Genómico	Hospital General, US	Humano	No hay Información	Sí	
SRX23055596	WGS	<i>Escherichia coli</i>	Genómico	Clínica veterinaria, Corea del Sur	Humano	No hay Información	Sí	
ILLUMINA	SRX24530459	WGS	<i>Pseudomonas aeruginosa</i>	Genómico	UCI, US	Humano	Heces y cultivos de sangre	Sí
	SRX24530457	WGS	<i>Pseudomonas aeruginosa</i>	Genómico	UCI, US	Humano	Heces y cultivos de sangre	Sí
	SRX24530456	WGS	<i>Pseudomonas aeruginosa</i>	Genómico	UCI, US	Humano	Heces y cultivos de sangre	Sí
	SRX24530454	WGS	<i>Pseudomonas aeruginosa</i>	Genómico	UCI, US	Humano	Heces y cultivos de sangre	Sí
	SRX24373962	WGS	<i>Staphylococcus aureus</i>	Genómico	Hospital Militar, Vietnam	Humano	Muestra respiratoria	Sí
	SRX24108112	WGS	<i>Klebsiella pneumoniae</i>	Genómico	Hospital Universitario, China	Humano	No hay Información	Sí
	DRX402992	WGS	<i>Acinetobacter baumannii</i>	Genómico	Hospital Militar, Vietnam	Humano	No hay Información	Sí
	DRX521347	WGS	<i>Salmonella enterica</i>	Genómico	No hay Información	Humano	Heces	Sí
	SRX22064762	WGS	<i>Acinetobacter baumannii</i>	Genómico	No hay Información	Humano	No hay Información	Sí
	SRX1033968	WGS	<i>Klebsiella pneumoniae</i>	Genómico	Brigham and Women's Hospital, US	Humano	Muestra de Sangre	Sí
	DRX374641	WGS	<i>Klebsiella pneumoniae</i>	Genómico	No hay Información	Humano	Muestra de esputo	Sí
	SRX20954232	WGS	<i>Acinetobacter baumannii</i>	Genómico	Hospital General, China	Humano	Cultivos bacterianos	Sí

	SRX1033998	WGS	<i>Klebsiella pneumoniae</i>	Genómico	Brigham and Women's Hospital, US	Humano	Fluido de abscesos	Sí
	SRX23041734	WGS	<i>Morganella morganii</i>	Genómico	No hay Información	Humano	UTI	Sí

Posteriormente, se realizó el análisis de datos genómicos para cada SRX utilizando GalaxyTrakr (Versión 3.0.3+galaxy0). Se extrajeron las lecturas en formato FASTQ a partir de la herramienta *Download and Extract Reads in FASTQ format from NCBI SRA*. El tipo de input seleccionado fue *SRR accession* y se seleccionó el formato de *output gzip compressed fastq*. Cada experimento se procesó en un historial diferente y se generó al finalizar dos archivos: *Paired-end data (fastq-fump)* y *Single-end data (fastq-dump)* (Leinonen et al., 2011).

- **Pipeline para Experimentos con Tecnología ONT**

- *Control de Calidad y procesamiento inicial de datos*

En GalaxyTrakr (Versión 3.0.3+galaxy0) se usó FASTQC (Versión 0.11.9) para evaluar la calidad de los datos de secuenciación crudos a partir del archivo *fastqsanger.gz* contenido en *Single-end data (fastq-dump)*, obtenido previamente (Andrews, 2024). Después, se eliminaron los adaptadores de los extremos de las lecturas a partir de la herramienta Porechop (Versión 0.2.4). Se usó como input el mismo archivo (Chan et al., 2020).

- *Ensamblaje y clasificación de metagenomas*

El ensamblaje de genomas se realizó con MEGAHIT (Versión 1.2.9) usando el archivo output de Porechop (C.-M. Liu et al., 2014). Posteriormente, las secuencias de metagenoma fueron clasificadas en *bins* mediante la herramienta Binning\_Refiner (Version 1.4.3) (Song & Thomas, 2017). De igual manera, los metagenomas se agruparon mediante MetaBAT2 (Version 2.15), el tamaño mínimo de *bins* para realizar el agrupamiento seleccionado fue 200000 pb. Esta herramienta generó una colección de *bins* para la identificación de genes de resistencia (Kang et al., 2019a).

- *Detección de genes de resistencia*

Se corrió la herramienta ABRicate (Galaxy Version 1.0.1), la cual permitió la detección de genes específicos en secuencias de ADN, incluyendo genes de resistencia a antimicrobianos (ARGs). Para ello, se comparó con la base de datos ResFinder, se usó un mínimo porcentaje de identidad del ADN de 80. Esto permitió la identificación y clasificación de los ARGs presentes en los experimentos analizados (Berwa & Caspar, 2024).

- **Pipeline para Experimentos con Tecnología Illumina**

- *Control de Calidad y procesamiento inicial de datos*

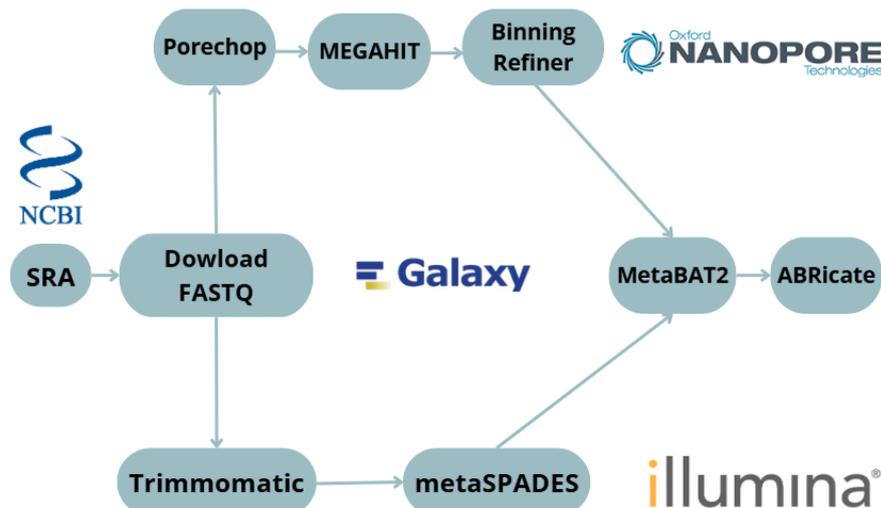
En Galaxy Australia (Galaxy Version 3.1.1+galaxy0), se extrajeron las lecturas en formato FASTQ a partir de los SRX correspondientes (Andrews, 2024). Posteriormente, se realizó el control de calidad y pre-procesamiento de lecturas mediante la herramienta Trimmomatic, con la opción *Paired-end (as collection)* y un Phred Score de 20. Esto permitió eliminar secuencias de baja calidad y adaptadores. Se generaron archivos procesados en formato *fastqsanger.gz* (Bolger et al., 2014).

- *Ensamblaje y clasificación de metagenomas*

El ensamblaje de metagenomas se realizó con metaSPAdes, se seleccionó la opción *Paired-end: list of dataset pairs* y se usó como input los datos procesados con *Trimmomatic on collection 1: paired* (Antipov et al., 2016). A continuación, los archivos generados previamente correspondientes a *MetaSPAdes on data: Scaffolds* fueron procesados usando MetaBAT2, y permitió recuperar bins representativos en cada experimento. El tamaño mínimo de bins para realizar el agrupamiento fue de 200000 pb. (Kang et al., 2019b).

- *Detección de genes de resistencia*

Para detectar genes con resistencia a antimicrobianos se corrió ABRicate, comparando los metagenomas con la base de datos Resfinder. Se determinó un 80% de identidad y cobertura para identificar secuencias. Los datos obtenidos en *MetaBAT2 on data: Bin Sequences* fueron analizados y permitió identificar genes de resistencia en los bins previamente generados (Berwa & Caspar, 2024).



**Figura No. 1 Pipeline para la identificación de genes de resistencia a antimicrobianos (ARGs) en metagenomas y genomas, a partir de lecturas NGS obtenidas ONT e Illumina.**

En la Figura No.1 se puede observar el flujo de trabajo usado para identificar genes de resistencia en datos genómicos y metagenómicos. Los pipelines estandarizados en Galaxy para cada tecnología de secuenciación incluyen la obtención de datos en la plataforma NCBI, el control de calidad y preprocesamiento de lecturas, el ensamblaje y clasificación de metagenomas y genomas, e identificación de genes de resistencia mediante ABRicate, empleando la base de datos ResFinder.

- **Mapeo y Clasificación Taxonómica**

- **Experimentos de ONT**

- *Clasificación taxonómica de lecturas*

Mediante Kraken2 se realizó la clasificación taxonómica y se empleó como input el archivo preprocesado posterior al trimado. Se seleccionó la opción *single-end* y la base de datos MiniKraken2 v2 (Wood & Salzberg, 2014).

- *Conversión de resultados taxonómicos*

Los resultados generados por Kraken2 fueron convertidos a formatos de representación taxonómica mediante Convert Kraken, se empleó la base de datos taxonómicos NCBI-2015-10-05 (Wood & Salzberg, 2014).

- *Visualización del Perfil Taxonómico*

Mediante Krona Pie Chart se generaron los gráficos metagenómicos para la interpretación del perfil taxonómico de cada experimento (Ondov et al., 2011).

- **Experimentos de Illumina**

- *Mapeo de Lecturas*

Mediante la herramienta Bowtie2 se realizó el mapeo de las lecturas, para lo cual se seleccionó la opción *Paired end*. Se seleccionó como input los archivos preprocesados posterior al trimado. Se habilitó la opción *Write unaligned reads (in fast format) to separate files*, se generó archivos FASTQ con las secuencias que no fueron alineadas acorde a los parámetros establecidos. El genoma de referencia empleado fue Humano (*Homo sapiens*) (b38): hg38 Canonical (Langmead et al., 2009).

- *Clasificación taxonómica de lecturas no alineadas*

A las lecturas no fueron alineadas previamente en Bowtie2 se les asignó una etiqueta taxonómica mediante Kraken2. Se usó la base de datos Minikraken2, y los resultados generados fueron convertidos a formatos de representación taxonómica usando Convert Kraken, con la base de datos de NCBI-2015-10-05 (Wood & Salzberg, 2014).

- *Visualización del Perfil Taxonómico*

Mediante Krona Pie Chart se generaron los gráficos metagenómicos para la interpretación del perfil taxonómico de cada experimento (Ondov et al., 2011).

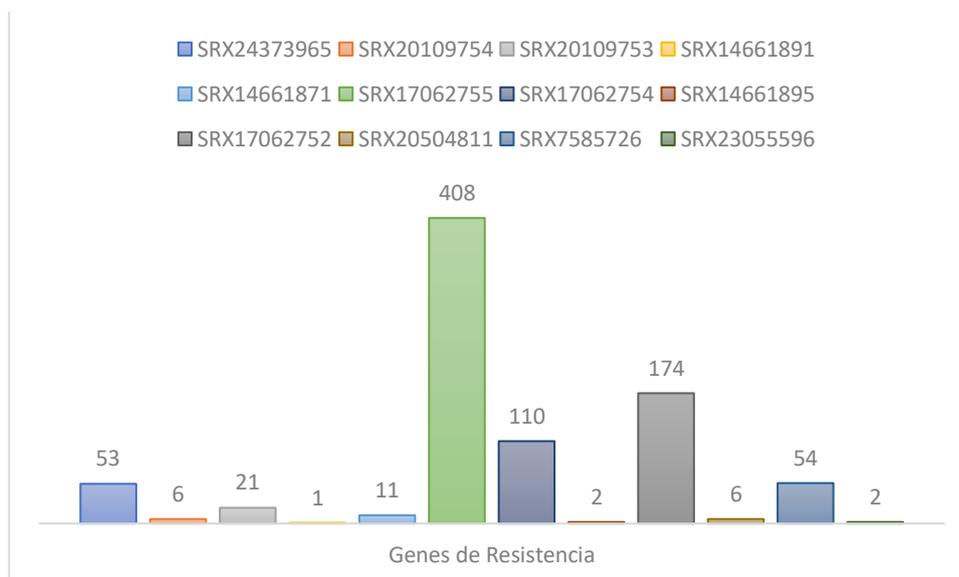
- **Evaluación de la Calidad de Ensamblaje**

A partir de los archivos *metaSPAdes on data x and data y: Scaffolds* generados previamente por metaSPAdes para la tecnología Illumina y los archivos *Assembly with MEGAHIT on data1* para la tecnología ONT, se procedió a realizar la evaluación del ensamblaje usando Quast. Para el presente análisis, se seleccionó el ensamblaje para metagenomas y se obtuvo un reporte con métricas para evaluar la calidad del ensamblaje. Entre los indicadores se encuentra *NGA50* y *LGA50* que permitieron conocer sobre la precisión del ensamblaje (Mikheenko et al., 2018).

### 3. Resultados

Para evaluar la efectividad de los pipelines estandarizados para la identificación de Genes de Resistencia a Antimicrobianos (ARGs) a partir de metagenomas y genomas obtenidos por NGS, se desarrolló un flujo de trabajo automatizado usando lecturas de experimentos de ONT e Illumina. Como se puede observar en la Figura No. 1, el control de calidad de las lecturas crudas se realizó mediante la herramienta FastQC, seguido por el trimado empleando Porechop para los datos ONT, y Trimmomatic para Illumina, lo cual garantizó la calidad de las secuencias. Los datos procesados fueron ensamblados con MEGAHIT y metaSPADES, y se clasificaron en bins usando Binning\_Refiner (Versión 1.4.3) y MetaBAT2 (Versión 2.15), lo cual permitió la agrupación eficiente de secuencias. En ambos casos mediante la herramienta ABRicate fue posible identificar los genes de resistencia a antimicrobianos al compararlos con la base de datos ResFinder.

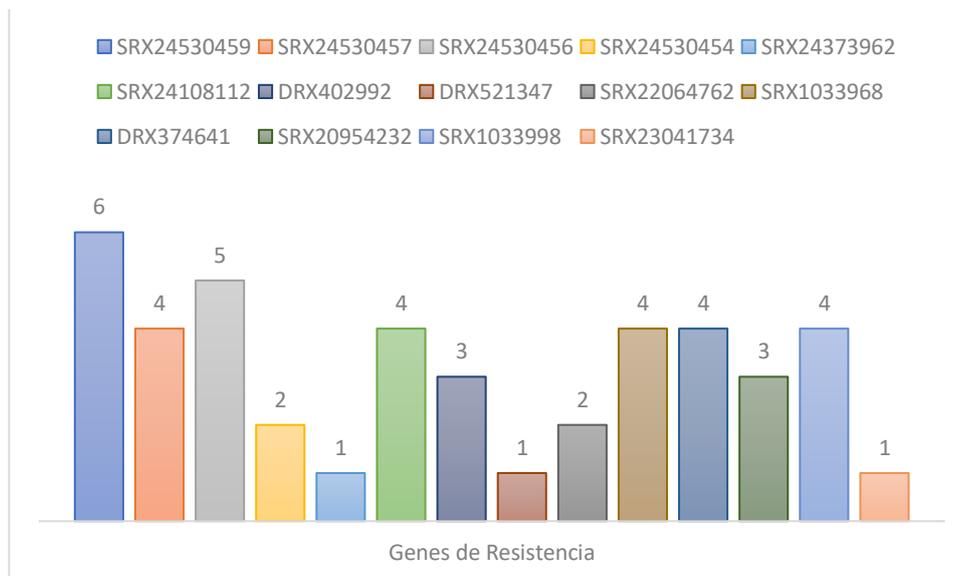
Como se puede observar en el Gráfico No. 1, para todos los experimentos de las lecturas con tecnología ONT, se logró agrupar las lecturas en bins, lo que indica que el binning funcionó correctamente, permitiendo un análisis más organizado y preciso. De igual forma, dentro de los distintos bins se identificaron ARGs. En el caso del experimento con SRA SRX14661891, se identificó un gen de resistencia. Caso contrario, los experimentos con mayor cantidad de genes de resistencia fueron SRX17062755 y SRX17062752, con 408 y 174 ARGs identificados, respectivamente. Es importante remarcar que, como se puede observar en la Tabla No. 1, para los tres casos se realizó un análisis de WGS a partir de muestras de lavado bronqueoalveolar de pacientes en un hospital de China.



**Gráfico No. 1 Cantidad de Genes de resistencia identificados en cada bin a partir del pipeline estandarizado para tecnología ONT**

En el Gráfico No.1 se muestra la distribución cuantitativa de los genes de resistencia a antimicrobianos identificados en cada bin obtenido a partir del ensamblaje y clasificación metagenómica del pipeline usando metagenomas y genomas con lecturas de ONT. Se puede observar que el experimento con mayor cantidad de genes de resistencia corresponde a SRX17062755, un análisis de WGS realizado a partir de un lavado bronqueoalveolar en un paciente de un hospital de China.

De acuerdo con lo que se presenta en el Gráfico No. 2, para todos los experimentos con tecnología ILLUMINA, también se logró agrupar las lecturas en bins y corrobora que el proceso de binning se llevó a cabo correctamente. De igual forma, dentro de los distintos bins se identificaron ARGs. En el caso de los experimentos con SRA SRX24373962, DRX521347 y SRX23041734 se identificó únicamente un gen de resistencia para cada caso. Como se puede observar en la Tabla No. 1, el SRX24373962 se realizó a partir de un análisis de WGS a partir de una muestra respiratoria de un paciente del Hospital Militar de Vietnam. Por otro lado, DRX521347 corresponde a un análisis de WGS a partir de una muestra de heces y DRX521347 se generó a partir de muestra de orina de un paciente con infección del tracto urinario (UTI). Caso contrario, los experimentos con mayor cantidad de genes de resistencia fueron SRX24530459 y SRX24530456, con 6 y 5 ARGs identificados, respectivamente. En la Tabla No. 1 se detalla que para ambos casos se realizó un análisis de WGS a partir de muestras de heces y cultivos de sangre de pacientes en unidades de cuidados intensivos (UCI) de Estados Unidos. Estos hallazgos sugieren que, en ambientes clínicos, en especial en aquellos críticos, la presencia de una considerable carga de resistencia se da principalmente por la presión selectiva provocada por el uso de antibióticos que facilita la propagación y mantenimiento del índice de ARGs en el microbiota hospitalario (Wu et al., 2022).



**Gráfico No. 2 Cantidad de Genes de resistencia identificados en cada bin a partir del pipeline estandarizado para tecnología ILLUMINA**

En el Gráfico No.2 se muestra la distribución cuantitativa de los genes de resistencia a antimicrobianos identificados en cada bin obtenido a partir del ensamblaje y clasificación metagenómica del pipeline usando metagenomas y genomas con lecturas de Illumina. Se puede observar que el experimento con mayor cantidad de genes de resistencia corresponde a SRX24530459, un análisis de WGS realizado a partir de muestras de heces y cultivo de sangre a un paciente perteneciente a la Unidad de Cuidados Intensivos en un Hospital de Estados Unidos

Mediante los pipelines estandarizados fue posible conocer la distribución cualitativa y determinar las subfamilias y variantes de ARGs presentes en cada bin de cada experimento, así como el antibiótico para el cual presentan la resistencia de las lecturas obtenidas a partir de las tecnologías ONT e ILLUMINA. Es importante mencionar que los pipelines estandarizados permitieron identificar los ARGs de mayor impacto y riesgo para la salud pública, los cuales confieren resistencias a antibióticos críticos. Entre estos ARGs se encuentran mecA, blaTEM, blaACT-7, blaZ, vanRA, oqxA, qnrB4 y fosA (Arturo et al., 2023; Nasrabadi et al., 2024; Vieira et al., 2023).

En la Tabla No. 2, se describen las variantes de ARGs generados por los flujos de trabajo automatizados y los antibióticos a los cuales generan resistencias. Sin embargo, es importante mencionar que los pipelines también permitieron identificar otros genes de resistencia que, a pesar de que no se describen detalladamente, son significativos y de preocupación en el área de salud (Arturo et al., 2023).

**Tabla No. 2 Variantes de Genes de Resistencia y Antibióticos a los cuales Presentan Resistencia Generados mediante Pipelines Estandarizados para las Tecnologías ONT e ILLUMINA**

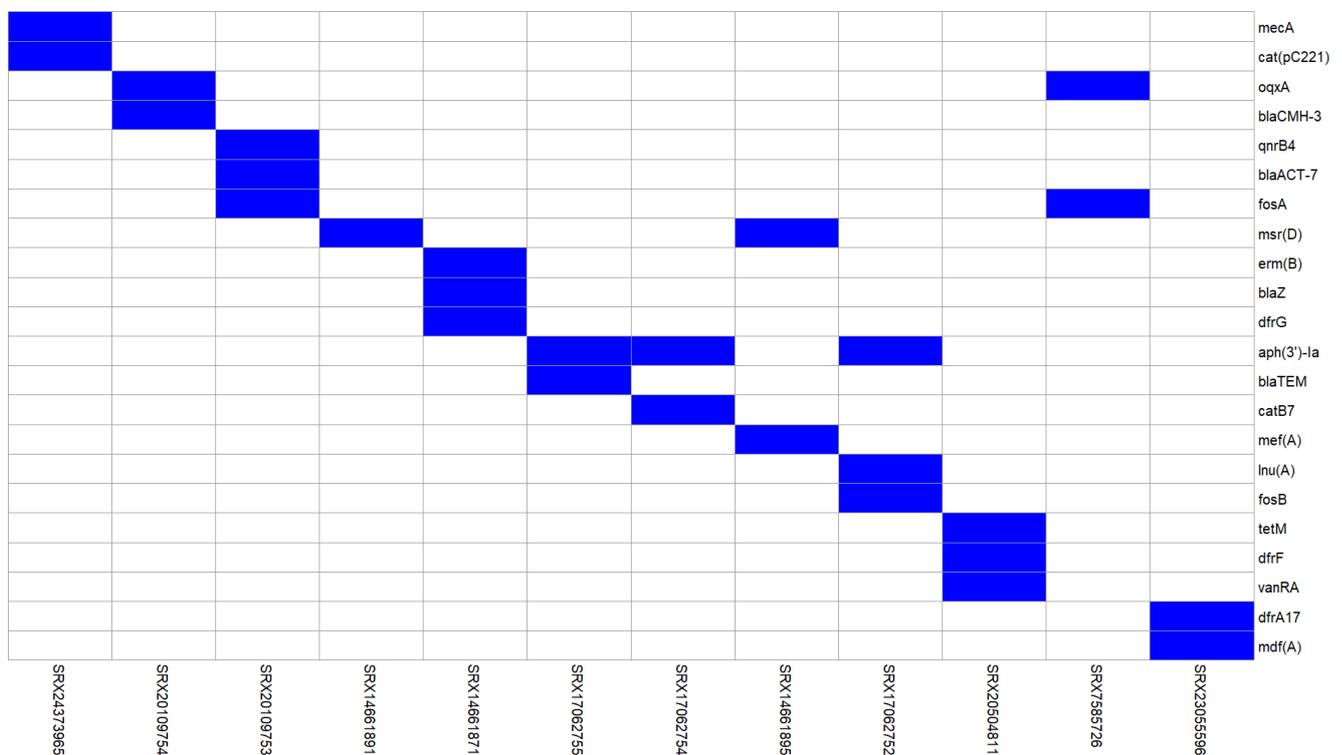
Tecnología de Secuenciación	SRA	Variante de Gen con Resistencia	Antibiótico
ONT	SRX24373965	mecA	Amoxicilina; Amoxicilina y Ácido Clavulánico; Ampicilina; Ampicilina Ácido Clavulánico; Cefepime; Cefixima; Cefotaxima; Cefoxitina; Ceftazidima; Ertapenem; Imipenem; Meropenem
		cat(pC221)_1	Cloranfenicol
	SRX20109754	oqxA_1	Ácido nalidíxico, Ciprofloxacino.
		blaCMH-3_1	β-lactámicos
	SRX20109753	qnrB4_1	Ciprofloxacino
		blaACT-7_1	β-lactámicos
		fosA_7	Fosfomicina
		qnrB4_1	Ciprofloxacino
SRX14661891	msr(D)_2	Erythro micina;Azithro micina;Telithro micina;Quinupristin;Pristina micina IA;Virginia micina S	
SRX14661871	erm(B)_12	Erythro micina;Linco micina;Clinda micina;Quinupristin;Pristina micina IA;Virginia micina S	
	blaZ_138	Amoxicilina; Ampicilina;Penicilina;Piperacilina	
	dfgG_1	Trimetoprim	
SRX17062755	aph(3')-Ia_1	Aminoglucósidos	
	blaTEM	Amoxicilina	

	SRX17062754	blaTEM-112_1 catB7_1	Amoxicilina; Ampicilina; Aztreonam; Cefepime; Cefotaxima; Ceftazidima; Ceftriaxona; Piperacilina; Ticarcilina Cloramfenicol
	SRX14661895	mef(A)_2 msr(D)_2	Eritromicina; Azitromicina. Eritromicina; Azitromicina; Telitromicina; Quinupristina; Pristinamicina IA; Virginiamicina S
	SRX17062752	blaTEM lnu(A)_1 fosB_3	Amoxicilina; Ampicilina; Aztreonam; Cefepime; Cefotaxima; Ceftazidima; Ceftriaxona; Piperacilina; Ticarcilina. Lincomicina Fosfomicina
	SRX20504811	tetM dfrF vanRA	Tetraciclina Diaminopirimidina Glicopéptido
	SRX7585726	oqxA_1 fosA6_1	Ácido nalidíxico; Ciprofloxacino Fosfomicina
	SRX23055596	dfrA17_1 mdf(A)_1	Trimetoprim Macrólidos; Lincosamidas; Estreograminas
ILLUMINA	SRX24530459	crpP_1 catB7_1 fosA_4 blaOXA-396_1 aph(3')-IIb_2 blaPAO_2	Ciprofloxacino Cloramfenicol Fosfomicina Amoxicilina; Ampicilina; Meropenem Aminoglucósidos β-lactámicos
	SRX24530457	blaPAO_1 catB7_1 fosA_4 blaOXA-396_1	β-lactámicos Cloramfenicol Fosfomicina Amoxicilina; Ampicilina; Meropenem
	SRX24530456	crpP_1 blaPAO_2 fosA_4 blaOXA-396_1 catB7_1	Ciprofloxacino Penicilina y Cefalosporina Fosfomicina Amoxicilina; Ampicilina; Meropenem Cloramfenicol
	SRX24530454	blaL1_3 aph(3')-IIc_1	β-lactámicos Aminoglucósidos
	SRX24373962	mecA_6	Amoxicilina; Amoxicilina Ácido Clavulánico; Ampicilina; Ampicilina Ácido Clavulánico; Cefepime; Cefixima; Cefotaxima; Cefoxitina; Ceftazidima; Ertapenem; Imipenem; Meropenem; Piperacilina; Piperacilina Tazobactam
	SRX24108112	fosA6 oqxB oqxA blaCTX-M-65 blaSHV	Fosfomicina Ácido nalidíxico; Ciprofloxacino. Ácido nalidíxico; Ciprofloxacino. Amoxicilina; Ampicilina; Aztreonam; Cefepime; Cefotaxima; Ceftazidima; Ceftriaxona; Piperacilina; Ticarcilina Amoxicilina; Ampicilina; Aztreonam; Cefepima; Cefotaxima; Ceftazidima; Ceftriaxona; Piperacilina; Ticarcilina.
	DRX402992	blaADC-25_1 blaOXA-23	β-lactámicos Imipenem; Meropenem
	DRX521347	aac(6)-Iaa	Amikacina; Tobramicina
	SRX22064762	blaOXA-23	Imipenem; Meropenem
	SRX1033968	fosA6 blaSHV-106 oqxA	Fosfomicina Amoxicilina; Ampicilina; Aztreonam; Cefepime; Cefotaxima; Ceftazidima; Ceftriaxona; Piperacilina; Ticarcilina Ácido nalidíxico; Ciprofloxacino.

		oqxB	Ácido nalidíxico; Ciprofloxacino.
DRX374641		oqxB	Ácido nalidíxico; Ciprofloxacino.
		oqxA	Ácido nalidíxico; Ciprofloxacino.
		fosA	Fosfomicina
		blaSHV-1	Amoxicilina; Ampicilina; Cefalotina; Piperacilina; Ticarcilina
SRX20954232		blaADC-25	β-lactámicos
		blaOXA-23	Imipenem; Meropenem
SRX1033998		fosA6	Fosfomicina
		oqxB	Ácido nalidíxico; Ciprofloxacino.
		oqxA	Ácido nalidíxico; Ciprofloxacino.
		blaSHV-12	Amoxicilina; Ampicilina; Aztreonam; Cefepime; Cefotaxime; Ceftazidime; Ceftriaxone; Piperacilina; Ticarcilina
SRX23041734		blaDHA-18	Amoxicilina; Amoxicilina Ácido Clavulánico; Ampicilina; Ampicilina Ácido Clavulánico; Cefotaxime; Cefoxitina; Ceftazidime; Piperacilina; Piperacilina Tazobactam; Ticarcilina; Ticarcilina Ácido Clavulánico.

Como se puede observar en el Gráfico No. 3, todos los experimentos presentan ARGs. Los SRA SRX20109753, SRX14661871, SRX17062752 y SRX20504811 identificaron tres subfamilias distintas de ARGs, mientras que para SRX14661891 sólo se identificaron una subfamilia.

En relación al SRX20109753 se identificaron qnrB4, relacionado a la resistencia a ciprofloxacino; blaACT-7, que confiere resistencia a β-lactámicos; y fosA, relacionado con resistencia a fosfomicina. De igual manera, respecto a SRX17062752 se detectaron los siguientes genes blaTEM, lnu(A) y fosB, que confieren resistencia a los antibióticos a la amoxicilina, lincomicina y fosfomicina respectivamente. Así también, en el SRX20504811 se identificaron los genes tetM, asociado a la resistencia a las tetraciclinas; dfrF, relacionado a resistencia a diaminopirimidina; y vanRA a glucopéptidos. Por otro lado, en SRX14661891, solo se detectó el gen de resistencia msr(D) que confiere resistencia a la eritromicina.



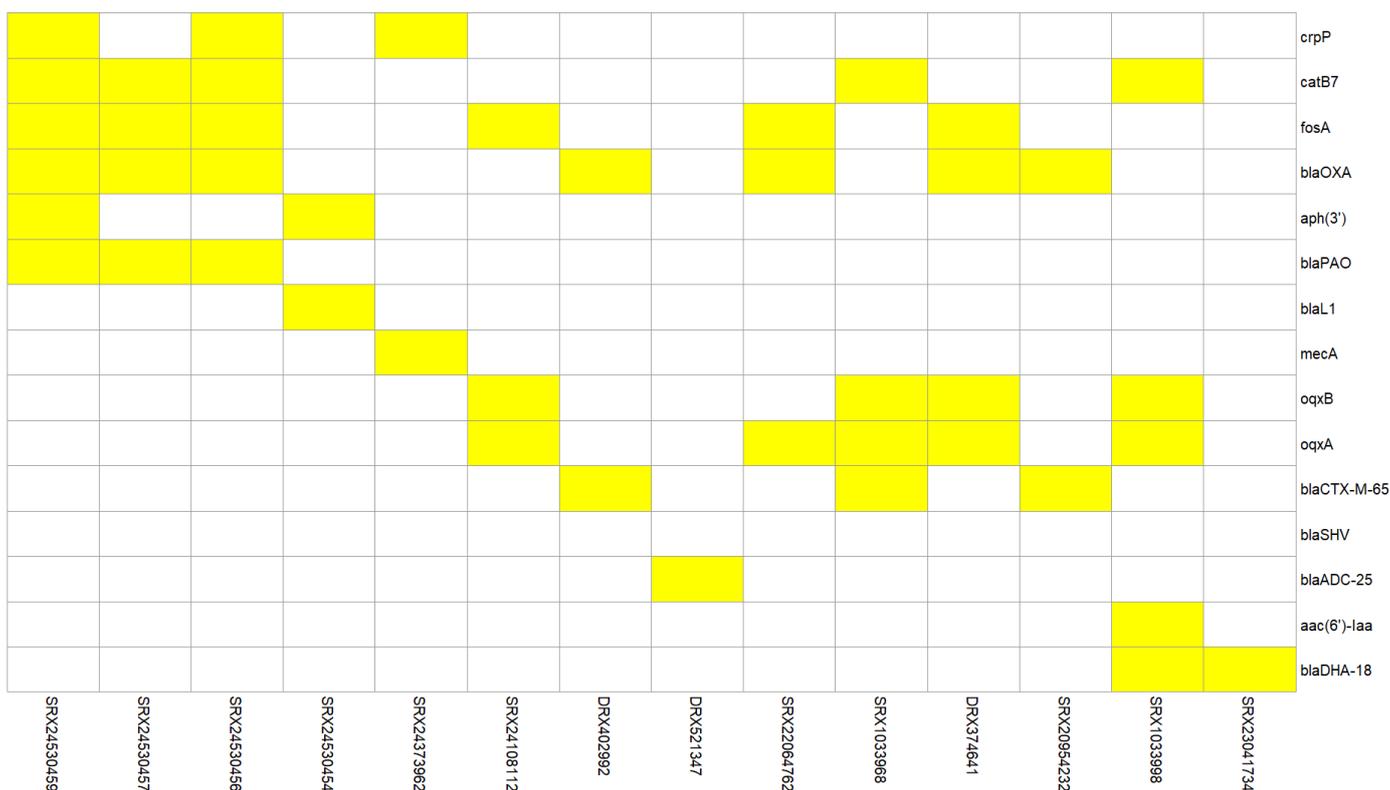
**Gráfico No. 3 Mapa de Calor de la Distribución de Genes de Resistencia a Antimicrobianos Identificados en Cada Bin a partir del Pipeline Estandarizado para Tecnología ONT**

En el Gráfico No.3 se detalla la distribución cualitativa de los genes de resistencia a antimicrobianos identificados en cada bin, obtenidos a partir del ensamblaje y clasificación metagenómica usando un pipeline estandarizado para el análisis de metagenomas y genomas a partir de lecturas de ONT. Se puede observar que el experimento con mayor cantidad de genes de resistencia corresponde a SRX4661871, un análisis de WGS realizado a partir de una muestra Mock. De manera similar, el experimento SRX20504811, que corresponde a un análisis de WGS realizado a partir de un cultivo de sangre de un paciente de un hospital de China, revela tres genes de resistencia.

Como se aprecia en el Gráfico No. 4, todos los experimentos para la tecnología ILLUMINA presentan ARGs. En relación a SRX24530459 presentó seis subfamilias diferentes, entre ellos crpP, asociado a la resistencia al ciprofloxacino; catB7 que confiere resistencia a cloramfenicol; fosA, relacionado a fosfomicina; blaOXA implicado en la inactivación de la amoxicilina; aph(3') asociado a aminoglucósidos; y blaPAO perteneciente a la familia de β-lactamasas.

Por su parte, SRX24530456 presentó cinco subfamilias distintas, incluyendo crpP, asociado a la resistencia al ciprofloxacino; blaPAO perteneciente a la familia de β-lactamasas; fosA, relacionado con la resistencia a fosfomicina; blaOXA a amoxicilina; y catB7 al cloramfenicol. De manera similar, en SRX1033998 se identificaron ARGs como fosA que confiere resistencia a la fosfomicina; oqxB y oqxA asociados a ácido nalidíxico y ciprofloxacino; y blaSHV-106 a la amoxicilina.

Caso contrario, DRX521347 y SRX23041734 mostraron una menor diversidad de ARGs. En DRX521347 se identificó el gen aac(6)-Iaa que confiere resistencia a aminoglucósidos como amikacina y trobamicina. Finalmente, en SRX23041734 se identificó blaDHA-18, que confiere resistencia a la amoxicilina.

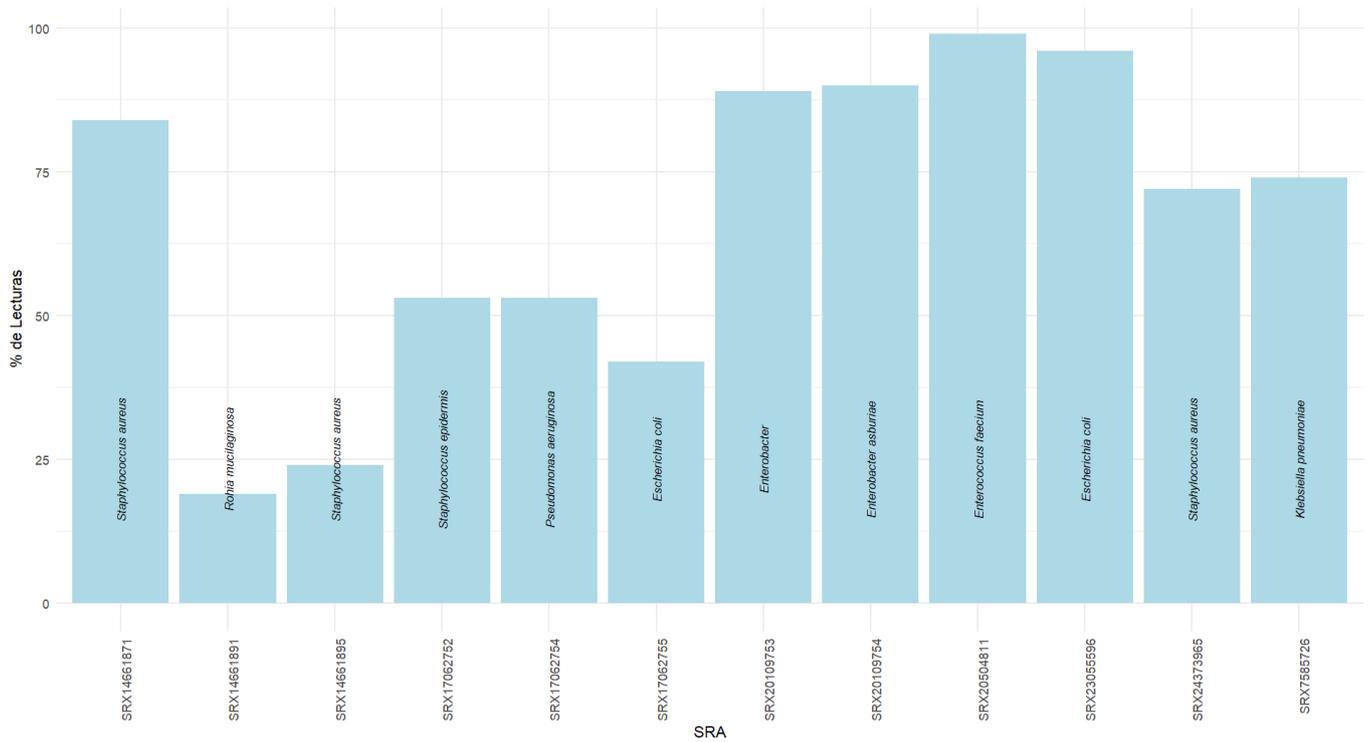


**Gráfico No. 4 Mapa de Calor de la Distribución de Genes de Resistencia a Antimicrobianos Identificados en Cada Bin a partir del Pipeline Estandarizado para Tecnología ILLUMINA**

En el Gráfico No.4 se detalla la distribución cualitativa de los genes de resistencia a antimicrobianos identificados en cada bin, obtenidos a partir del ensamblaje y clasificación metagenómica usando un pipeline estandarizado para el análisis de metagenomas y genomas a partir de lecturas de ILLUMINA. Se puede observar que el experimento con mayor cantidad de genes de resistencia corresponde a SRX24530459, un análisis de WGS realizado a partir de muestras de heces y cultivo de sangre a un paciente perteneciente a la Unidad de Cuidados Intensivos en un Hospital de Estados Unidos.

Respecto al mapeo y clasificación taxonómica para conocer la efectividad de los pipelines estandarizados en la identificación de ARGs, se realizó el análisis con la herramienta Krona Pie Chart. En los Gráficos No. 5 y 6 se detalla la presencia del microorganismo predominante en cada experimento analizado tanto para la tecnología ONT e ILLUMINA.

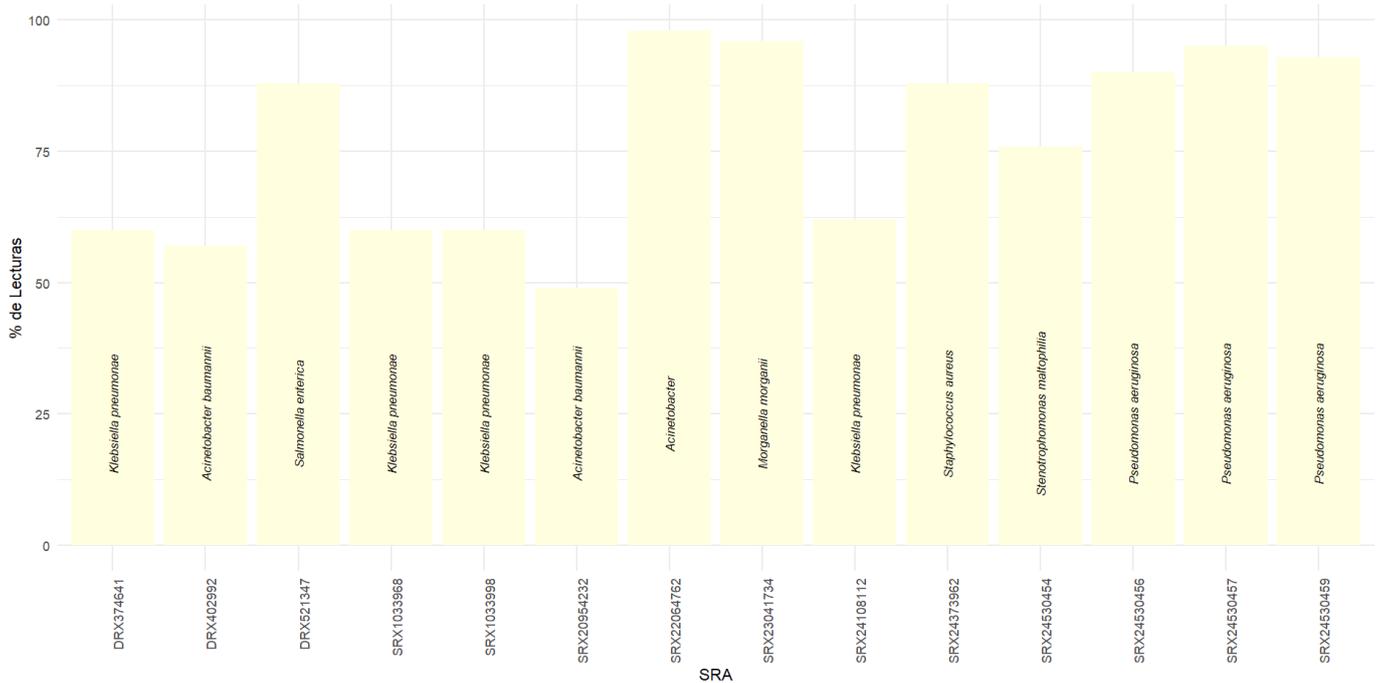
En el Gráfico No. 5 se puede observar que para el SRA SRX20504811, el 99% de las lecturas correspondió a *Enterococcus faecium*, lo cual sugiere una alta dominancia de esta bacteria en la composición microbiana de la muestra. De igual forma, respecto al SRX23055596, el 96% de las lecturas se relacionaron con *Escherichia coli*. Así también, el 90% de las lecturas del experimento SRX20109754 se identificaron para la bacteria *Enterobacter asburiae*. Caso contrario, el SRA SRX14661891 con el 19% de lecturas se vinculó con *Rohia mucilaginosa*. Mientras que en SRX14661895, el 26% de las lecturas corresponden a *Staphylococcus aureus*. Estos resultados indicaron variabilidad en la composición microbiana, e indica diferencias en las características de las muestras o en los ambientes de donde provienen (Pal et al., 2015).



**Gráfico No. 5 Abundancia Relativa del Microorganismo con Mayor Porcentaje de Lecturas en el Análisis de Metagenomas y Genomas a partir del Pipeline Estandarizado para Tecnología ONT**

En el Gráfico No.5 se presenta el análisis de los microorganismos predominantes según perfil taxonómico realizado con la herramienta Krona Pie Chart. Se puede observar que en el experimento SRX20504811, el 99% de lecturas identificadas corresponden a *Enterococcus faecium*.

Tal como se muestra en el Gráfico No. 6, los análisis de los siguientes SRA presentaron que más del 90% de las lecturas se relacionaron con los siguientes microorganismos: para los SRA SRX24530459, SRX24530457 y SRX24530456, la mayor parte de las lecturas se asociaron con *Pseudomonas aeruginosa*, mientras que el experimento SRX22064762 estuvo predominantemente relacionado a *Acinetobacter* y SRX23041734 mostró una alta abundancia de *Morganella morganii*. Caso contrario, SRX20954232 presentó una distinta distribución, con sólo el 49% relacionadas con *Acinetobacter baumannii*, lo cual indica una composición microbiana más variada.



**Gráfico No. 6 Abundancia Relativa del Microorganismo con Mayor Porcentaje de Lecturas en el Análisis de Metagenomas y Genomas a partir del Pipeline Estandarizado para Tecnología ILLUMINA**

En el Gráfico No.6 se presenta el análisis de los microorganismos predominantes según perfil taxonómico realizado con la herramienta Krona Pie Chart. Se puede observar que en el experimento SRX22064762, el 98% de lecturas identificadas correspondieron a *Acinetobacter sp.*

El análisis de la calidad genómica reveló diferencias en la cobertura y continuidad del ensamblaje. Mediante la herramienta Quast de Galaxy fue posible determinar la integridad de los ensamblajes generados en los pipelines estandarizados para la identificación de ARGs a partir de lecturas ONT e Illumina (Mikheenko et al., 2018). Se analizaron parámetros importantes como la longitud total alineada, NGA50 y LGA50, lo cual permitió estimar de forma precisa la calidad estructural en relación con los genomas de referencia. El NGA50 indica la longitud mínima de los fragmentos del ensamblaje alineados con la referencia, de tal forma que al sumarlos alcance al menos el 50% del genoma de referencia. Esta métrica permite conocer la continuidad y exactitud del ensamblaje, proporcionando una estimación más confiable de su calidad en comparación con el genoma de referencia. Un valor alto indica un ensamblaje menos fragmentado y una mayor proporción de secuencias correctamente ensambladas (Hiltemann et al., 2023). En contraste, el parámetro LGA50 indica el número mínimo de loci necesarios para cubrir al menos el 50% del genoma de referencia y refleja la fragmentación del ensamblaje en relación con la estructura del genoma real. Mientras más bajo sea este valor, indica un ensamblaje continuo y de mejor calidad (Batut et al., 2018).

Como se puede observar en la Tabla No.3, se identificó una variedad de microorganismos con distinta cobertura y continuidad del ensamblaje. Respecto a ONT, aquellos experimentos que presentaron un buen equilibrio entre la longitud total alineada y la calidad del ensamblaje (NGA50 y LGA50) fueron: en SRX24373965, 4235336 bases se alinearon con el microorganismo *Staphylococcus aureus*, lo cual sugiere una buena cobertura. Además, se obtuvo un NGA50 de 2930, el cual indicó que el 50% de las lecturas se agruparon en contigs de esa longitud o mayor, lo cual también refleja un ensamblaje bueno en relación a la longitud de los fragmentos alineados. De igual forma, LGA50 fue 297 el cual indicó la longitud del contig en donde se encuentra la mitad de la longitud total alineada, este valor sugiere que la longitud de los fragmentos más largos no es muy alta en contraste con otras especies.

El SRA SRX20109753 evidenció que 92709748 bases se alinearon con *Enterobacter hormaechei ATCC 49162*, con un NGA50 de 6039 que sugiere un ensamblaje de buena calidad y LGA50 de 265 que indica un valor relativamente alto lo cual permitiría inferir que hay fragmentación en el ensamblaje.

En relación a SRX17062754, 4217185 de bases se alinearon con *Pseudomonas aeruginosa PAO1*, con un NGA50 de 944 que refleja un ensamblaje aceptable y LGA50 de 1598 que indica que la mitad de la longitud total alineada proviene de contigs de esa longitud o mayor, e indican una calidad buena.

El experimento SRX1466187, presentó una longitud total alineada de 14840186 bases y se relacionó con el microorganismo *Enterococcus faecalis*, un NGA50 de 5104 y sugiere que el ensamblaje posee contigs relativamente largos, y por ende las lecturas se agrupan de manera eficaz en fragmentos grandes y un ensamblaje de buena calidad. Sin embargo, el LGA50 es bajo, 230, lo cual indica que hay fragmentos más cortos en el ensamblaje que pueden afectar la continuidad del genoma reconstruido.

De igual forma, se evidenciaron casos en donde se reportaron los valores de Longitud Total Alineada pero no se generó un valor de NGA50 ni LGA50, esto se pudo deber a la presencia significativa de contigs pequeños que no contribuyen de manera significativa a la continuidad del genoma o bien se puede atribuir a una baja cobertura.

En relación al pipeline con lecturas generadas con tecnología ILLUMINA, 3725176 bases se alinearon con *Morganella morganii*, lo cual sugiere un ensamblaje completo y representativo. Se obtuvo un NGA50 de 802420 que indicó una menor fragmentación y que el 50% del genoma alineado se encuentra contenido en contigs de al menos esta longitud. El parámetro LGA50 fue 99329, lo cual indica un ensamblaje con fragmentos largos.

El SRA SRX24108112 se alineó con *Klebsiella pneumoniae\_UHKPC33* mediante una longitud total de 5251116 pb, un NGA50 de 115897 y LGA50 de 15 que indicó que los alineamientos se distribuyeron en pocos fragmentos grandes. El experimento DRX402992 obtuvo 3650015 pb alineados con *Acinetobacter baumannii\_TYTH-1*, se obtuvo un NGA50 de 103150 que indicó un buen ensamblaje y LGA50 de 14 que sugirió que los alineamientos se encontraban concentrados en pocos fragmentos largos.

De igual forma, el SRX1033998 se alineó con *Klebsiella pneumoniae* con una longitud total de 4861635, un NGA50 de 82951 y LGA50 de 23. Dichos valores indicaron un ensamblaje con fragmentos relativamente largos y estructurados.

En las lecturas con SRA DRX521347, presentó una longitud total alineada correspondiente a 4372297 pb asociado con *Salmonella enterica*. Presentó un NGA50 de 63263 y LGA50 de 25 los cuales reflejan un ensamblaje de calidad y continuidad moderada y los contigs se distribuyeron de forma homogénea.

**Tabla No. 3** Evaluación de la Calidad del Ensamblaje del Genoma: Microorganismos Alineados y Métricas de Cobertura en Pipelines Estandarizados a partir de Lecturas de ONT e ILLUMINA

Tecnología de Secuenciación	SRA	Microorganismos Alineados	Longitud Total Alineada	NGA50	LGA50
ONT	SRX24373965	<i>Staphylococcus aureus</i>	4235336	2930	297
		<i>Staphylococcus aureus M0077</i>	6503291	4728	196
		<i>Staphylococcus aureus subsp. aureus DSM 20231</i>	4175372	2953	287
		<i>Staphylococcus aureus subsp. aureus RN4220</i>	4229591	3151	260
	SRX20109754	<i>Enterobacter asburiae</i>	8843610	960	1160
		<i>Enterobacter kobei</i>	7241783	864	1424
		<i>Enterobacter ludwigii</i>	6573666	780	1769
		<i>Enterobacter sp. ODB01</i>	6931462	868	1380
		<i>Leclercia adecarboxylata</i>	3217786	516	2909
	SRX20109753	<i>Serratia marcescens</i>	189582	539	147
		<i>Enterobacter cloacae</i>	156595	838	25
		<i>Enterobacter hormaechei</i>	92709748	6039	265
	SRX14661891	<i>Enterobacter hormaechei ATCC 49162</i>	10262751	1220	1203
		<i>Anaeroglobus geminatus</i>	1586488	1227	414
		<i>Capnocytophaga sputigena</i>	1659277	613	1011
		<i>Eikenella corrodens ATCC 23834</i>	1838731	1293	333
		<i>Eikenella sp. NML130454</i>	557029	0	0
		<i>Prevotella pallens ATCC 700821</i>	2196647	836	949
	SRX14661871	<i>Streptococcus sp. oral taxon 064</i>	1084575	506	781
		<i>Veillonella parvula DSM 2008</i>	1254500	676	661
<i>Enterococcus faecalis</i>		14840186	5104	230	
<i>Enterococcus faecalis ATCC 19433</i>		14523943	5440	187	
<i>Enterococcus faecalis ATCC 35038</i>		13679573	4882	215	
<i>Serratia marcescens</i>		543502	2647	26	
<i>Shigella dysenteriae</i>		3679201	961	1246	
<i>Shigella flexneri 2002017</i>		3343999	810	1418	
<i>Shigella sonnei</i>		8753463	2920	533	
<i>Staphylococcus aureus</i>		21676899	5468	178	
<i>Staphylococcus aureus M0077</i>	30225806	7045	143		
<i>Staphylococcus aureus subsp. aureus JH9</i>	30467527	7222	140		

	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> RN4220	21406359	5645	161
	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> ST772-MRSA-V	22724091	5945	161
	<i>Staphylococcus phage phiN315</i>	146646	2516	7
	synthetic <i>Escherichia coli</i> C321.deltaA	10409732	3618	418
SRX17062755	<i>Escherichia coli</i>	2298353	-	-
	<i>Shigella flexneri</i> 1a	1837290	-	-
	<i>Shigella sonnei</i>	3131846	568	2282
SRX17062754	<i>Escherichia coli</i>	559807	-	-
	<i>Pseudomonas aeruginosa</i>	4152324	701	2239
	<i>Pseudomonas aeruginosa</i> DHS01	4304599	787	2061
	<i>Pseudomonas aeruginosa</i> DSM 50071 NBRC 12689	4013569	849	1741
	<i>Pseudomonas aeruginosa</i> PAO1	4217185	944	1598
	<i>Pseudomonas aeruginosa</i> PAO1H2O	3998451	847	1755
	<i>Pseudomonas aeruginosa</i> UCBPP-PA14	3374718	538	2387
	<i>Pseudomonas</i> sp. AK6U	4079633	731	2108
	<i>Pseudomonas</i> sp. P179	4170623	769	2044
	<i>Shigella</i> sp. PAMC 28760	934069	-	-
SRX14661895	<i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i> 3 1 27	365050	-	-
	<i>Fusobacterium periodonticum</i>	429429	-	-
	<i>Gemella sanguinis</i>	494334	-	-
	<i>Moraxella catarrhalis</i>	1193674	774	359
	<i>Porphyromonas pasteri</i>	423260	-	-
	<i>Rothia mucilaginosa</i>	491647	-	-
	<i>Rothia mucilaginosa</i> ATCC 25296	584916	-	-
	<i>Staphylococcus aureus</i>	1503099	465	1107
	<i>Streptococcus constellatus</i> subsp. <i>pharyngis</i> C818	1224993	619	759
	<i>Streptococcus parasanguinis</i> ATCC 15912	849801	-	-
SRX17062752	<i>Escherichia coli</i> K-12	1488401	-	-
	<i>Staphylococcus epidermidis</i>	5559636	2456	380
	<i>Staphylococcus epidermidis</i> ATCC 12228	6347368	2608	372
SRX20504811	<i>Enterococcus faecium</i>	4245929	6506	153
SRX7585726	<i>Klebsiella pneumoniae</i>	4877867	80686	200
SRX23055596	<i>Escherichia coli</i>	221681256	1287	592
SRX24530459	<i>Pseudomonas aeruginosa</i>	5845593	6590	279
SRX24530457	<i>Pseudomonas aeruginosa</i>	5797336	6632	276
SRX24530456	<i>Pseudomonas aeruginosa</i>	5896796	5542	315
SRX24530454	<i>Xanthomonas phaseoli</i> pv. <i>dieffenbachiae</i>	71763	-	-
SRX24373962	<i>Staphylococcus argenteus</i>	762519	-	-
	<i>Staphylococcus aureus</i>	2595813	78122	10
SRX24108112	<i>Klebsiella pneumoniae</i> UHKPC33	5251116	115897	15
DRX402992	<i>Acinetobacter baumannii</i> TYTH-1	3650015	103150	14
DRX521347	<i>Salmonella enterica</i>	4372297	63263	25
SRX22064762	<i>Acinetobacter baumannii</i> TYTH-1	3684817	17416	65
SRX1033968	<i>Klebsiella pneumoniae</i>	4890744	65279	25
DRX374641	<i>Klebsiella pneumoniae</i>	4877667	80686	20
SRX20954232	<i>Acinetobacter baumannii</i> TYTH-1	3676354	60620	19

	SRX1033998	<i>Klebsiella pneumoniae</i>	4861635	82951	23
	SRX23041734	<i>Morganella morganii</i>	3725176	802420	99329

#### 4. Discusión

La resistencia a antibióticos es considerada una amenaza a nivel global en constante crecimiento que ha comprometido la salud humana, en especial los tratamientos clínicos de las enfermedades. Actualmente se han reportado ARGs en todos los entornos, incluso hábitats naturales, diseñados y clínicos (Allen et al., 2010). Los principales impulsores de la diseminación de ARGs son las actividades antropogénicas y el uso clínico de antibióticos (Zhang et al., 2022). Una alternativa viable para poder contrarrestar el impacto de los ARGs es el uso de tecnologías de última generación, como NGS, las cuales en conjunto con análisis bioinformáticos han permitido la identificación de dichos genes (Florensa et al., 2022).

En el presente estudio, se estandarizaron flujos de trabajo automatizados para un análisis eficaz a partir de metagenomas y genomas de lecturas obtenidas por ONT e Illumina. Estos pipelines garantizaron un control de calidad riguroso, un ensamblaje optimizado y un binning preciso de las secuencias, lo cual permitió mejorar la reproducibilidad y la consistencia de los resultados. La combinación de herramientas dentro de la aplicación Galaxy demostraron ser eficientes y efectivas para la caracterización de comunidades microbianas y su diversidad genética (Mehta et al., 2023). Como se puede observar en los Gráficos No. 1 y No. 2, la cantidad de bins generados variaron significativamente, en experimentos como SRX17062755 (408 bins) y SRX17062752 (174 bins), los cuales evidencian una complejidad metagenómica mayor en contraste con muestras como SRX14661891 (1 bin). Estas diferencias se pueden atribuir tanto a la riqueza como complejidad microbiana de cada muestra. Sin embargo, también se pudo deber a factores técnicos, incluyendo la calidad de las lecturas o la profundidad de secuenciación (Schiml et al., 2023). Además, el uso de la base de datos ResFinder con la herramienta ABRicate permitió la identificación precisa de ARGs con parámetros de identidad y cobertura bien definidos, lo cual optimizó la precisión de los análisis (Wee & Yap, 2021a). En conjunto, estos hallazgos sugieren también una considerable carga de resistencia en los ambientes clínicos analizados, lo cual resalta la importancia de optimizar enfoques bioinformáticos para la vigilancia de la resistencia antimicrobiana (Cheng et al., 2022). El uso de herramientas bioinformáticas para la identificación precisa de ARGs ha permitido optimizar los análisis metagenómicos y conocer las implicaciones prácticas significativas en el ámbito clínico y epidemiológico (Chiu & Miller, 2019). Se ha reportado que para el desarrollo de estrategias de control es importante la caracterización de ARGs en muestras metagenómicas ya que permiten la comprensión de la distribución y abundancia de determinantes genéticos con resistencia en distintos entornos (Zhuang et al., 2021). De igual forma, desde el ámbito terapéutico, la detección de ARGs favorece a la selección informada de tratamientos antimicrobianos, lo cual reduce el uso inapropiado de antibióticos y reduciendo la propagación de la resistencia (Muteeb et al., 2023). Así también, dentro de la vigilancia epidemiológica, el monitoreo sistemático de ARGs permite determinar tendencias en la resistencia antimicrobiana y detecta posibles brotes asociados a patógenos resistentes, lo cual es importante para mitigar los riesgos de la transferencia horizontal de genes en la salud pública (Djordjevic et al., 2023).

Como se puede observar en los Gráficos No. 3 y 4 existe una diversidad de genes de resistencia en función de las distintas muestras y tecnología usada. En el SRA SRX4661871, el pipeline generó en la muestra Mock una mayor cantidad de ARGs, lo cual se puede atribuir a la naturaleza controlada de la muestra, en contraste con las muestras clínicas, donde se ha reportado una composición microbiana más diversa y dinámica (Nicholls et al., 2019). La cantidad variable de ARGs identificados indica cuán heterogéneas son las muestras y sus fuentes (Ibekwe et al., 2011). Por ejemplo, el pipeline para el experimento SRX24530456, relacionado a muestras clínicas de un paciente en UCI de un hospital de los US, permitió identificar cinco subfamilias de ARGs, con genes asociados a la resistencia de distintos antimicrobianos como ciprofloxacino,  $\beta$ -lactamasas, fosfomicina, amoxicilina y cloramfenicol. Así también, mediante los pipelines estandarizados fue posible identificar ARGs de preocupación crítica para la salud mundial, en especial aquellos genes relacionados con la resistencia a antibióticos de primera línea y los de amplio espectro (Nasrabadí et al., 2024). El experimento SRX24373965 presentó el gen *mecA*, responsable de la resistencia a  $\beta$ -lactámicos como la amoxicilina, meropenem y cefalosporinas. Se ha reportado que *mecA*, codifica la proteína de unión a penicilinas de baja afinidad PBP2, está asociada con la resistencia a infecciones complejas en ambientes hospitalarios, y facilitan la propagación de *S. Aureus* resistente a la meticilina (MRSA) (Wielders et al., 2002). Así también, el SRA SRX24530459, permitió identificar varios genes como *crpP* (resistencia a ciprofloxacino), *blaOXA* (resistencia a amoxicilina, ampicilina y meropenem), y *fosA* (resistencia a fosfomicina). El gen *blaOXA* se ha asociado con una amplia resistencia a carbapenémicos, los cuales son antibióticos de última línea usados en el tratamiento de infecciones graves, incluidas las del tracto urinario causadas por bacterias multirresistentes, lo cual es una amenaza directa contra la salud pública (Sivarajan et al., 2025). En relación al SRX20109753, identificó el gen *qnrB4* que confiere la resistencia a ciprofloxacino y *blaACT-7* relacionado a los  $\beta$ -lactámicos (Ku et al., 2018; H. Y. Yang et al., 2014). En el caso de los  $\beta$ -lactámicos, se ha reportado que la resistencia a estos antibióticos compromete a brindar un tratamiento a diversas infecciones, desde urinarias complicadas y neumonía bacteriana para el caso del

ciprofloxacino, hasta neumonía adquirida en la comunidad y hospitalaria, así como sepsis y bacteriemias (Hawkey et al., 2018).

Para corroborar que los pipelines estandarizados funcionaron correctamente se realizó un análisis taxonómico con Krona Pie Chart. La identificación correcta de los microorganismos predominantes en cada SRA respalda la implementación de cada paso de los flujos de trabajo, desde el control de calidad hasta la clasificación taxonómica y detección de ARGs (Cuccuru et al., 2014). Como se puede observar en el Gráfico No.5, se obtuvo una predominancia alta de *Enterococcus faecium* en el experimento SRX20504811 (99% de lecturas), *Acetivobacter sp.* en SRX22064762 (98% de lecturas) y *Morganella morganii* en el SRX23041734 (98% de lecturas), esto sugiere una posible fuerte influencia por la fuente de origen o por factores asociados al procesamiento de datos (Rauer et al., 2025). En ambientes clínicos, es común una alta presencia de *Enterococcus faecium* y *Acetivobacter sp.*, ya que se relacionan ampliamente con infecciones nosocomiales, y se ha reportado que pueden colonizar superficies hospitalarias e incluso el microbioma de pacientes inmunocomprometidos (Sangiorgio et al., 2024; Suleyman et al., 2018). Por otro lado, la composición taxonómica observada se pudo ver influenciada por el procesamiento de datos bioinformáticos (Rauer et al., 2025). Los factores que podrían afectar la identificación de especies podrían ser la calidad inicial de las lecturas, la eficiencia en la eliminación de adaptadores y la precisión del ensamblaje (Lamas et al., 2023). Es por ello que el uso de herramientas en la etapa del control de calidad y preprocesamiento tales como FastQC, Porechop y Trimmomatic, fueron fundamentales para garantizar que únicamente se usen lecturas de alta calidad en los posteriores análisis. En el caso de que estos pasos no se configuren correctamente se podría dar lugar a sesgos que favorezcan la detección de ciertos microorganismos sobre otros, afectando la abundancia relativa de las muestras (Tzec-Interián et al., 2025). De igual manera, la capacidad de ensamblaje y binning influyeron en los resultados obtenidos. Mediante la herramienta MEGAHIT usadas para el pipeline de los experimentos de ONT se garantizó que se recuperen contigs de calidad y fue esencial para la adecuada reconstrucción de los metagenomas y posterior identificación de perfiles taxonómicos y ARGs (Kayani et al., 2021). En los SRA donde se obtuvo sólo una especie predominante como el caso de SRX20504811 con *Enterococcus faecium*, el ensamblaje fue más preciso y se generó bins bien definidos. Caso contrario, en muestras donde se evidencia una mayor diversidad, tal como SRX20954232, se puede ver afecta la recuperación de especies menos abundantes, lo cual genera una distribución taxonómica más heterogénea (D. Gupta & Sardar, 2021).

Por otro lado, como se puede observar en el Gráfico No. 6, se evidenció la presencia predominante de *Pseudomonas aeruginosa* en SRX24530459, SRX24530457 y SRX24530456 (más del 90% de lecturas), los cuales correspondieron a muestras de pacientes en UCI en un hospital de US, e indicaron la presencia de una baja diversidad de comunidades microbianas, o a su vez también se puede atribuir a ciertos sesgos durante el muestreo y procesamiento que favorecieron a que se identifique a esta bacteria (Y. Yang et al., 2022). Se ha reportado que *Pseudomonas aeruginosa* es un patógeno oportunista prevalente, en especial en dispositivos médicos contaminados, heridas, o infecciones respiratorias. Se caracterizan porque forman biopelículas y resisten condiciones adversas por lo cual sobreviven en ambientes donde la competencia con otras especies es limitada. Esto explicaría el motivo por el cual el 90% de las lecturas se asociaron con este patógeno (Diggle & Whiteley, 2020; Madalina Mihai et al., 2015). Sin embargo, la constancia de los resultados obtenidos en múltiples experimentos sugirió que el ensamblaje con MetaSPADES y posterior clasificación con MetaBAT2 permitieron una adecuada reconstrucción de genomas y recuperación de bins representativos (Kang et al., 2019a; Maguire et al., 2020). Mediante el pipeline se corrobora sobre la capacidad de la herramienta metaSPADES para manejar datos obtenidos con tecnología Illumina y generar ensamblajes de alta calidad que permitieron tener contigs bien definidos (Antipov et al., 2016). Como se puede observar en los Gráficos No. 5 y 6, MetaBAT2 permitió una segmentación taxonómica eficiente y permitió identificar el patógeno de mayor predominancia con alta confianza en cada experimento (Kang et al., 2019a). La combinación de los ensamblajes generados por MEGAHIT y MetaSPADES aseguró una clasificación robusta de los metagenomas.

De igual manera, para evaluar detalladamente la calidad estructural de los genomas reconstruidos a partir de lecturas ONT e Illumina se usó la herramienta QUAST. Los resultados generados evidenciaron diferencias significativas en la continuidad y cobertura de los ensamblajes, lo cual indica las particularidades de cada tecnología de secuenciación. Los parámetros evaluados, como la Longitud Total Alineada, NGA50 y LGA50, fueron importantes para estimar la calidad estructural de los ensamblajes y su relación con genomas de referencia (Gurevich et al., 2013).

En el caso de los ensamblajes generados a partir de los experimentos de las lecturas ONT, se observó que la calidad del ensamblaje mostró variación en relación al microorganismo predominante en cada experimento. Como se puede observar en la Tabla No. 3, el experimento SRX24373965 tuvo una Longitud Total Alineada 4235336 bases con *S. aureus*. Esta información coincide con la información de la muestra subida a la base de datos de NCBI SRA y detallada a la Tabla No. 1. De igual manera, los indicadores NGA50 fueron de 2930 y LGA50 de 297, estos valores sugirieron una cobertura adecuada y segmentación estructural aceptable, aunque el valor bajo de LGA50 indicaron la presencia de fragmentos pequeños en el ensamblaje que podían afectar la continuidad (Mikheenko et al., 2018). Caso contrario, SRX20109753 se alineó con *Enterobacter hormaechei ATCC 49162* (92709748 bases), pero con el NGA50 de 6039 y LGA50 de 265, lo cual indicó un ensamblaje de buena calidad, pero con posibles fragmentaciones que podrían impactar la precisión en la identificación de genes de resistencia (Mäkinen et al., 2012). En la Tabla No.1 se puede observar que los investigadores reportaron y secuenciaron el mismo microorganismo a partir de una muestra de cultivo celular bacteriano de una muestra de un paciente del Hospital de Bangladesh. Los experimentos con *Pseudomonas aeruginosa* PAO1 (SRX17062754) y *Enterococcus faecalis* (SRX1466187) mostraron valores de NGA50 de 944 y

5104, respectivamente. Sin embargo, el LGA50 en *E.faecalis* fue de 230, lo cual indica que a pesar de que se obtuvieron fragmentos largos, la presencia de fragmentos cortos podría afectar la continuidad del genoma ensamblado (Gurevich et al., 2013). Esto podría indicar la variabilidad en la eficiencia del ensamblaje dependiendo de la complejidad del genoma bacteriano y las características innatas de las lecturas generadas por ONT (Mikheenko et al., 2018).

En relación a los experimentos con la tecnología ILLUMINA, se generaron ensamblajes con una mejor fragmentación y continuidad, lo cual se respalda por valores superiores de NGA50 y LGA50. El SRA SRX24108112, se alineó con *Klebsiella pneumoniae UHKPC33* (Longitud Total Alineada 5251116 bases), con un NGA50 DE 115897 Y LGA50 de 15, lo cual nos indicó un ensamblaje de alta calidad con grandes fragmentos contiguos, lo cual favorece la identificación de regiones genómicas completas (Wee & Yap, 2021b). Como se puede observar en la Tabla No. 1, este microorganismo coincide con el reportado en NCBI a partir de una muestra de un paciente de un Hospital de China. De igual forma, el experimento DRX402992, tuvo un valor de Longitud Total Alineada de 3650015 con el microorganismo *Acinetobacter baumannii TYTH-1*, presentó un NGA50 de 103150 y LGA50 de 14, lo cual corrobora que la Tecnología Illumina permite obtener ensamblajes más completos y estructurados en comparación con ONT (Afgan et al., 2018). De igual forma, en la Tabla No.1 se confirma que el microorganismo generado mediante el pipeline coincide con el reportado en NCBI SRA.

Mediante el análisis de valores de NGA50 y LGA50 indicó que, a pesar de que la tecnología ONT brinda lecturas más largas, la fragmentación del ensamblaje puede ser mayor por la presencia de errores propias de esta plataforma, como tasas altas de error de inserción y deleción (Mäkinen et al., 2012). Se ha reportado que estos errores afectan la reconstrucción precisa del genoma, lo cual puede indicar el motivo por el cual los ensamblajes presentan menores valores de NGA50 y LGA50. Caso contrario, ILLUMINA generó ensamblajes más robustos con una mayor continuidad y menor fragmentación, lo cual nos permite inferir que es una mejor opción para estudios de genómica comparativa reconstrucción de genomas completos (Gurevich et al., 2013).

De igual forma, como se observa en la Tabla No. 3, se generaron casos en los que no se obtuvieron valores de NGA50 y LGA50 lo cual se podría atribuir por la presencia significativa de contigs pequeños que no contribuyeron de forma significativa a la continuidad del genoma o con una baja cobertura de secuenciación que no permitió un ensamblaje efectivo de la estructura genómica (Gurevich et al., 2013). Esto nos indicó sobre la importancia de una buena profundidad de secuenciación y estrategias de ensamblaje optimizadas para cada tecnología. Para que exista una mejor precisión y continuidad del ensamblaje en futuros análisis bioinformáticos, sería recomendable evaluar la combinación de enfoques que usen la longitud de lectura de ONT y la precisión de ILLUMINA, lo cual puede dar lugar a ensamblajes más completos y correctos estructuralmente (Afgan et al., 2018).

Así también es importante la validación y optimización de los pipelines bioinformáticos empleados para la identificación de ARGs en distintas muestras clínicas como, fluidos corporales, tejidos infectados, y muestras obtenidas de pacientes con infecciones nosocomiales (Sanabria et al., 2021). Mediante dichas validaciones será posible ajustar los parámetros de los flujos de trabajo para optimizar la sensibilidad y especificidad en la detección de ARGs en entornos clínicos de alta complejidad microbiológica (Kaprou et al., 2021). De igual forma, la integración de ideas adicionales tales como modelos basados en machine learning, podría ofrecer una mayor capacidad para analizar patrones complejos de resistencia y anticipar su evolución en infecciones bacterianas (Anahtar et al., 2021). Otra área prometedora, sería la creación de flujos de trabajo integrados que combinen datos genómicos y clínicos, lo que permitiría una identificación más rápida y precisa de ARGs, lo cual podría favorecer en las decisiones terapéuticas personalizadas y optimizar la eficacia de los tratamientos con antimicrobianos (Branda & Scarpa, 2024). Finalmente, la implementación de bases de datos más actualizadas y específicas, que abarquen una mayor diversidad de ARGs de distintas regiones geográficas y fuentes, podría ampliar el alcance de los análisis y mejorar la identificación de resistencias emergentes en poblaciones microbianas poco exploradas (Chen et al., 2025).

## 5. Conclusiones

Los resultados obtenidos demostraron que la estandarización de pipelines automatizados permitió la identificación precisa de genes de resistencia a los antimicrobianos a partir de metagenomas con lecturas obtenidas por NGS de una forma eficiente y reproducible. La combinación de herramientas de la plataforma Galaxy para el control de calidad, ensamblaje, binning y detección de ARGs permitió la caracterización de la diversidad genética y taxonómica de las muestras, demostrando la presencia de ARGs relevantes y una variación en la composición microbiana que refleja la complejidad de los ambientes clínicos analizados.

De igual forma, que a pesar de que ONT tiene como ventaja generar lecturas largas, sus ensamblajes evidenciaron una mayor fragmentación, y en determinados casos, incluso valores bajos o ausentes de métricas tales como NGA50 y LGA50. Esto se puede atribuir a la presencia de contigs pequeños y a una probable baja cobertura de secuenciación, lo cual afectó la continuidad del genoma reconstruido. En contraste, los datos generados a partir del pipeline de ILLUMINA mostraron ensamblajes más robustos, con valores altos de NGA50 y LGA50, lo cual indica que esta plataforma es la más adecuada para la generación de contigs largos y bien organizados, lo que ayudó en la identificación de regiones genómicas completas.

Finalmente, el presente estudio destaca la importancia de implementar enfoques bioinformáticos estandarizados e integrarlos con tecnologías de secuenciación. La combinación de las lecturas largas de ONT con la precisión de ILLUMINA podría representar una estrategia prometedora para superar las limitaciones inherentes de cada plataforma, permitiendo el ensamblaje de más completos y precisos. Estos avances permiten mejorar la vigilancia y el control de ARGs en entornos clínicos, y también contribuyen a un avance de flujos de trabajo más eficientes para el análisis de microbiomas de distintos recursos.

### Referencias citadas

- Abushaheen, M. A., Muzahed, Fatani, A. J., Alosaimi, M., Mansy, W., George, M., Acharya, S., Rathod, S., Divakar, D. D., Jhugroo, C., Vellappally, S., Khan, A. A., Shaik, J., & Jhugroo, P. (2020). Antimicrobial resistance, mechanisms and its clinical significance. *Disease-a-Month*, *66*(6), 100971. <https://doi.org/10.1016/J.DISAMONTH.2020.100971>
- Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Ech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B. A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., & Blankenberg, D. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, *46*(W1), W537–W544. <https://doi.org/10.1093/NAR/GKY379>
- Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Grüning, B., Guerler, A., Hillman-Jackson, J., Kuster, G. Von, Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., & Goecks, J. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research*, *44*(W1), W3–W10. <https://doi.org/10.1093/NAR/GKW343>
- Allen, H. K., Donato, J., Wang, H. H., Cloud-Hansen, K. A., Davies, J., & Handelsman, J. (2010). Call of the wild: antibiotic resistance genes in natural environments. *Nature Reviews Microbiology* *2010* *8*:4, *8*(4), 251–259. <https://doi.org/10.1038/nrmicro2312>
- Anahar, M. N., Yang, J. H., & Kanjilal, S. (2021). Applications of Machine Learning to the Problem of Antimicrobial Resistance: an Emerging Model for Translational Research. *Journal of Clinical Microbiology*, *59*(7). <https://doi.org/10.1128/JCM.01260-20/ASSET/D6BD3115-0DFD-4335-BC91-2E4D9AC64511/ASSETS/IMAGES/LARGE/JCM.01260-20-F0002.JPG>
- Antipov, D., Korobeynikov, A., McLean, J. S., & Pevzner, P. A. (2016). hybridSPAdes: an algorithm for hybrid assembly of short and long reads. *Bioinformatics*, *32*(7), 1009–1015. <https://doi.org/10.1093/BIOINFORMATICS/BTV688>
- Arturo, L., Silvas, C., & Autor, ). (2023). Resistencia bacteriana, una crisis actual. *Revista Española de Salud Pública*, *97*, e202302013. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10541255/>
- Balderrama-González, A. S., Piñón-Castillo, H. A., Ramírez-Valdespino, C. A., Landeros-Martínez, L. L., Orrantia-Borunda, E., & Esparza-Ponce, H. E. (2021). Antimicrobial Resistance and Inorganic Nanoparticles. *International Journal of Molecular Sciences* *2021*, Vol. *22*, Page *12890*, *22*(23), 12890. <https://doi.org/10.3390/IJMS222312890>
- Batut, B., Hiltemann, S., Bagnacani, A., Baker, D., Bhardwaj, V., Blank, C., Bretaudeau, A., Brillet-Guéguen, L., Čech, M., Chilton, J., Clements, D., Doppelt-Azeroual, O., Erxleben, A., Freeberg, M. A., Gladman, S., Hoogstrate, Y., Hotz, H. R., Houwaart, T., Jagtap, P., ... Grüning, B. (2018). Community-Driven Data Analysis Training for Biology. *Cell Systems*, *6*(6), 752-758.e1. <https://doi.org/10.1016/J.CELS.2018.05.012>
- Berwa, A., & Caspar, Y. (2024). Easy analysis of bacterial whole-genome sequencing data for clinical microbiologists using open-source Galaxy platform: Characterization of ESBL-producing Enterobacterales from bloodstream infections. *Journal of Global Antimicrobial Resistance*, *39*, 153–158. <https://doi.org/10.1016/J.JGAR.2024.08.012>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. <https://doi.org/10.1093/BIOINFORMATICS/BTU170>
- Branda, F., & Scarpa, F. (2024). Implications of Artificial Intelligence in Addressing Antimicrobial Resistance: Innovations, Global Challenges, and Healthcare's Future. *Antibiotics* *2024*, Vol. *13*, Page *502*, *13*(6), 502. <https://doi.org/10.3390/ANTIBIOTICS13060502>
- Chan, W. S., Au, C. H., Chung, Y., Leung, H. C. M., Ho, D. N., Wong, E. Y. L., Lam, T. W., Chan, T. L., Ma, E. S. K., & Tang, B. S. F. (2020). Rapid and economical drug resistance profiling with Nanopore MinION for clinical specimens with low bacillary burden of Mycobacterium tuberculosis. *BMC Research Notes*, *13*(1), 1–7. <https://doi.org/10.1186/S13104-020-05287-9/TABLES/1>
- Chen, X., Yin, X., Xu, X., & Zhang, T. (2025). Species-resolved profiling of antibiotic resistance genes in complex metagenomes through long-read overlapping with Argo. *Nature Communications* *2025* *16*:1, *16*(1), 1–14. <https://doi.org/10.1038/s41467-025-57088-y>
- Cheng, H., Sun, Y., Yang, Q., Deng, M., Yu, Z., Zhu, G., Qu, J., Liu, L., Yang, L., & Xia, Y. (2022). A rapid bacterial pathogen and antimicrobial resistance diagnosis workflow using Oxford nanopore adaptive sequencing method. *Briefings in Bioinformatics*, *23*(6). <https://doi.org/10.1093/BIB/BBAC453>

- Chiu, C. Y., & Miller, S. A. (2019). Clinical metagenomics. *Nature Reviews Genetics* 2019 20:6, 20(6), 341–355. <https://doi.org/10.1038/s41576-019-0113-7>
- 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*, 30(13), 1928–1929. <https://doi.org/10.1093/BIOINFORMATICS/BTU135>
- Diggle, S. P., & Whiteley, M. (2020). Microbe profile: *Pseudomonas aeruginosa*: Opportunistic pathogen and lab rat. *Microbiology (United Kingdom)*, 166(1), 30–33. <https://doi.org/10.1099/MIC.0.000860/CITE/REFWORKS>
- Djordjevic, S. P., Jarocki, V. M., Seemann, T., Cummins, M. L., Watt, A. E., Drigo, B., Wyrsch, E. R., Reid, C. J., Donner, E., & Howden, B. P. (2023). Genomic surveillance for antimicrobial resistance — a One Health perspective. *Nature Reviews Genetics* 2023 25:2, 25(2), 142–157. <https://doi.org/10.1038/s41576-023-00649-y>
- Florensa, A. F., Kaas, R. S., Clausen, P. T. L. C., Aytan-Aktug, D., & Aarestrup, F. M. (2022). ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microbial Genomics*, 8(1), 000748. <https://doi.org/10.1099/MGEN.0.000748/CITE/REFWORKS>
- Gupta, A., Mumtaz, S., Li, C. H., Hussain, I., & Rotello, V. M. (2019). Combatting antibiotic-resistant bacteria using nanomaterials. *Chemical Society Reviews*, 48(2), 415–427. <https://doi.org/10.1039/C7CS00748E>
- Gupta, D., & Sardar, R. (2021). Bioinformatics of genome annotation. *Bioinformatics and Human Genomics Research*, 7–30. <https://doi.org/10.1201/9781003005926-2/BIOINFORMATICS-GENOME-ANNOTATION-DINESH-GUPTA-RAHILA-SARDAR>
- Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072–1075. <https://doi.org/10.1093/BIOINFORMATICS/BTT086>
- Hasan, T. H., & Al-Harmoosh, R. A. (2020). Mechanisms of antibiotics resistance in bacteria. *Systematic Reviews in Pharmacy*, 11(6), 817–823. <https://doi.org/10.31838/srp.2020.6.118>
- Hawkey, P. M., Warren, R. E., Livermore, D. M., McNulty, C. A. M., Enoch, D. A., Otter, J. A., & Wilson, A. P. R. (2018). Treatment of infections caused by multidrug-resistant Gram-negative bacteria: report of the British Society for Antimicrobial Chemotherapy/Healthcare Infection Society/British Infection Association Joint Working Party. *Journal of Antimicrobial Chemotherapy*, 73(suppl\_3), iii2–iii78. <https://doi.org/10.1093/JAC/DKY027>
- Hiltemann, S., Rasche, H., Gladman, S., Hotz, H. R., Larivière, D., Blankenberg, D., Jagtap, P. D., Wollmann, T., Bretaudeau, A., Goué, N., Griffin, T. J., Royaux, C., Bras, Y. Le, Mehta, S., Syme, A., Coppens, F., Droesbeke, B., Soranzo, N., Bacon, W., ... Batut, B. (2023). Galaxy Training: A powerful framework for teaching! *PLoS Computational Biology*, 19(1). <https://doi.org/10.1371/JOURNAL.PCBI.1010752>
- Ibekwe, A. M., Murinda, S. E., & Graves, A. K. (2011). Genetic Diversity and Antimicrobial Resistance of *Escherichia coli* from Human and Animal Sources Uncovers Multiple Resistances from Human Sources. *PLOS ONE*, 6(6), e20819. <https://doi.org/10.1371/JOURNAL.PONE.0020819>
- Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019a). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, 2019(7), e7359. <https://doi.org/10.7717/PEERJ.7359/SUPP-3>
- Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019b). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, 2019(7), e7359. <https://doi.org/10.7717/PEERJ.7359/SUPP-3>
- Kaprou, G. D., Bergšpica, I., Alexa, E. A., Alvarez-Ordóñez, A., & Prieto, M. (2021). Rapid Methods for Antimicrobial Resistance Diagnostics. *Antibiotics* 2021, Vol. 10, Page 209, 10(2), 209. <https://doi.org/10.3390/ANTIBIOTICS10020209>
- Kayani, M. U. R., Huang, W., Feng, R., & Chen, L. (2021). Genome-resolved metagenomics using environmental and clinical samples. *Briefings in Bioinformatics*, 22(5), 1–20. <https://doi.org/10.1093/BIB/BBAB030>
- Ku, Y. H., Lee, M. F., Chuang, Y. C., & Yu, W. L. (2018). Detection of Plasmid-Mediated  $\beta$ -Lactamase Genes and Emergence of a Novel AmpC (CMH-1) in *Enterobacter cloacae* at a Medical Center in Southern Taiwan. *Journal of Clinical Medicine* 2019, Vol. 8, Page 8, 8(1), 8. <https://doi.org/10.3390/JCM8010008>
- Lamas, A., Garrido-Maestu, A., Prieto, A., Cepeda, A., & Franco, C. M. (2023). Whole genome sequencing in the palm of your hand: how to implement a MinION Galaxy-based workflow in a food safety laboratory for rapid *Salmonella* spp. serotyping, virulence, and antimicrobial resistance gene identification. *Frontiers in Microbiology*, 14, 1254692. <https://doi.org/10.3389/FMICB.2023.1254692/BIBTEX>
- Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10(3), 1–10. <https://doi.org/10.1186/GB-2009-10-3-R25/TABLES/5>
- Leinonen, R., Sugawara, H., Shumway, M., & Collaboration, on behalf of the I. N. S. D. (2011). The Sequence Read Archive. *Nucleic Acids Research*, 39(suppl\_1), D19–D21. <https://doi.org/10.1093/NAR/GKQ1019>

- Liu, C.-M., Luo, R., & Lam, T.-W. (2014). GPU-Accelerated BWT Construction for Large Collection of Short Reads. *Bioinformatics*, *31*(10), 1674–1676. <http://arxiv.org/abs/1401.7457>
- Liu, Y., Li, R., Xiao, X., & Wang, Z. (2019). Antibiotic adjuvants: an alternative approach to overcome multi-drug resistant Gram-negative bacteria. *Critical Reviews in Microbiology*, *45*(3), 301–314. <https://doi.org/10.1080/1040841X.2019.1599813>
- Madalina Mihai, M., Maria Holban, A., Giurcaneanu, C., Gabriela Popa, L., Mihaela Oanea, R., Lazar, V., Carmen Chifiriuc, M., Popa, M., & Ioan Popa, M. (2015). Microbial Biofilms: Impact on the Pathogenesis of Periodontitis, Cystic Fibrosis, Chronic Wounds and Medical Device-Related Infections. *Current Topics in Medicinal Chemistry*, *15*(16), 1552–1576. <https://doi.org/10.2174/1568026615666150414123800>
- Maguire, F., Jia, B., Gray, K. L., Lau, W. Y. V., Beiko, R. G., & Brinkman, F. S. L. (2020). Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic islands. *Microbial Genomics*, *6*(10), 1–12. <https://doi.org/10.1099/MGEN.0.000436/CITE/REFWORKS>
- Mäkinen, V., Salmela, L., & Ylinen, J. (2012). Normalized N50 assembly metric using gap-restricted co-linear chaining. *BMC Bioinformatics*, *13*(1), 1–5. <https://doi.org/10.1186/1471-2105-13-255/FIGURES/1>
- Mehta, S., Bernt, M., Chambers, M., Fahrner, M., Föll, M. C., Gruening, B., Horro, C., Johnson, J. E., Loux, V., Rajczewski, A. T., Schilling, O., Vandenbrouck, Y., Gustafsson, O. J. R., Thang, W. C. M., Hyde, C., Price, G., Jagtap, P. D., & Griffin, T. J. (2023). A Galaxy of informatics resources for MS-based proteomics. *Expert Review of Proteomics*, *20*(11), 251–266. <https://doi.org/10.1080/14789450.2023.2265062>
- Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., & Gurevich, A. (2018). Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*, *34*(13), i142–i150. <https://doi.org/10.1093/BIOINFORMATICS/BTY266>
- Muteeb, G., Rehman, M. T., Shahwan, M., & Aatif, M. (2023). Origin of Antibiotics and Antibiotic Resistance, and Their Impacts on Drug Development: A Narrative Review. *Pharmaceuticals* 2023, Vol. 16, Page 1615, *16*(11), 1615. <https://doi.org/10.3390/PH16111615>
- Nasrabadi, K. S., Ahmadi, Y., Ghasemi, S. M., & Shokri, D. (2024). Identification of blaOXA-51-23-58, blaVIM, blaNDM, and blaIMP carbapenemase genes in Acinetobacter baumannii isolates from hospitalized patients. *BMC Research Notes*, *17*(1). <https://doi.org/10.1186/S13104-024-07047-5>
- Nicholls, S. M., Quick, J. C., Tang, S., & Loman, N. J. (2019). Ultra-deep, long-read nanopore sequencing of mock microbial community standards. *GigaScience*, *8*(5), 1–9. <https://doi.org/10.1093/GIGASCIENCE/GIZ043>
- Oakeson, K. F., Wagner, J. M., Mendenhall, M., Rohrwasser, A., & Atkinson-Dunn, R. (2017a). Bioinformatic Analyses of Whole-Genome Sequence Data in a Public Health Laboratory. *Emerging Infectious Diseases*, *23*(9), 1441. <https://doi.org/10.3201/EID2309.170416>
- Oakeson, K. F., Wagner, J. M., Mendenhall, M., Rohrwasser, A., & Atkinson-Dunn, R. (2017b). Bioinformatic Analyses of Whole-Genome Sequence Data in a Public Health Laboratory - Volume 23, Number 9—September 2017 - Emerging Infectious Diseases journal - CDC. *Emerging Infectious Diseases*, *23*(9), 1441–1445. <https://doi.org/10.3201/EID2309.170416>
- Ondov, B. D., Bergman, N. H., & Phillippy, A. M. (2011). Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics*, *12*(1), 1–10. <https://doi.org/10.1186/1471-2105-12-385/FIGURES/4>
- Pal, C., Bengtsson-Palme, J., Kristiansson, E., & Larsson, D. G. J. (2015). Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics*, *16*(1). <https://doi.org/10.1186/S12864-015-2153-5>
- Pendleton, K. M., Erb-Downward, J. R., Bao, Y., Branton, W. R., Falkowski, N. R., Newton, D. W., Huffnagle, G. B., & Dickson, R. P. (2017). Rapid pathogen identification in bacterial pneumonia using real-time metagenomics. *American Journal of Respiratory and Critical Care Medicine*, *196*(12), 1610–1612. [https://doi.org/10.1164/RCCM.201703-0537LE/SUPPL\\_FILE/DISCLOSURES.PDF](https://doi.org/10.1164/RCCM.201703-0537LE/SUPPL_FILE/DISCLOSURES.PDF)
- Rauer, L., De Tomassi, A., Müller, C. L., Hülpiusch, C., Traidl-Hoffmann, C., Reiger, M., & Neumann, A. U. (2025). De-biasing microbiome sequencing data: bacterial morphology-based correction of extraction bias and correlates of chimera formation. *Microbiome* 2025 *13:1*, *13*(1), 1–17. <https://doi.org/10.1186/S40168-024-01998-4>
- Ross, J., Larco, D., Colon, O., Coalson, J., Gaus, D., Taylor, K., & Lee, S. (2020). Índices de resistencia a los antibióticos en aislamientos clínicos en Santo Domingo, Ecuador. *Práctica Familiar Rural*, *5*(1). <https://doi.org/10.23936/pfr.v5i1.144>
- Sanabria, A. M., Janice, J., Hjerde, E., Simonsen, G. S., & Hanssen, A. M. (2021). Shotgun-metagenomics based prediction of antibiotic resistance and virulence determinants in Staphylococcus aureus from periprosthetic tissue on blood culture bottles. *Scientific Reports* 2021 *11:1*, *11*(1), 1–13. <https://doi.org/10.1038/s41598-021-00383-7>
- Sangiorgio, G., Calvo, M., Migliorisi, G., Campanile, F., & Stefani, S. (2024). The Impact of Enterococcus spp. in the Immunocompromised Host: A Comprehensive Review. *Pathogens* 2024, Vol. 13, Page 409, *13*(5), 409. <https://doi.org/10.3390/PATHOGENS13050409>
- Schaenzer, A. J., & Wright, G. D. (2020). Antibiotic Resistance by Enzymatic Modification of Antibiotic Targets. *Trends in Molecular Medicine*, *26*(8), 768–782. <https://doi.org/10.1016/J.MOLMED.2020.05.001>

- Schimpl, V. C., Delogu, F., Kumar, P., Kunath, B., Batut, B., Mehta, S., Johnson, J. E., Grüning, B., Pope, P. B., Jagtap, P. D., Griffin, T. J., & Arntzen, M. (2023). Integrative meta-omics in Galaxy and beyond. *Environmental Microbiome*, 18(1), 1–13. <https://doi.org/10.1186/S40793-023-00514-9/FIGURES/4>
- Sheka, D., Alabi, N., & Gordon, P. M. K. (2021a). Oxford nanopore sequencing in clinical microbiology and infection diagnostics. *Briefings in Bioinformatics*, 22(5), 1–17. <https://doi.org/10.1093/BIB/BBAA403>
- Sheka, D., Alabi, N., & Gordon, P. M. K. (2021b). Oxford nanopore sequencing in clinical microbiology and infection diagnostics. *Briefings in Bioinformatics*, 22(5), 1–17. <https://doi.org/10.1093/BIB/BBAA403>
- Sivarajan, V., Ganesh, A. V., Subramani, P., Ganesapandi, P., Sivanandan, R. N., Prakash, S., Manikandan, N., Dharmarajan, A., Arfuso, F., Warriar, S., Raj, M., & Perumal, K. (2025). Prevalence and genomic insights of carbapenem resistant and ESBL producing Multidrug resistant *Escherichia coli* in urinary tract infections. *Scientific Reports 2025 15:1*, 15(1), 1–13. <https://doi.org/10.1038/s41598-024-84754-w>
- Song, W. Z., & Thomas, T. (2017). Binning\_refiner: improving genome bins through the combination of different binning programs. *Bioinformatics*, 33(12), 1873–1875. <https://doi.org/10.1093/BIOINFORMATICS/BTX086>
- Suleyman, G., Alangaden, G., & Bardossy, A. C. (2018). The Role of Environmental Contamination in the Transmission of Nosocomial Pathogens and Healthcare-Associated Infections. *Current Infectious Disease Reports*, 20(6), 1–11. <https://doi.org/10.1007/S11908-018-0620-2/TABLES/1>
- Tan, T. Y. (2003). Use of molecular techniques for the detection of antibiotic resistance in bacteria. *Expert Review of Molecular Diagnostics*, 3(1), 93–103. <https://doi.org/10.1586/14737159.3.1.93>
- Tang, W., Zhang, Y., Luo, C., Zhou, L., Zhang, Z., Tang, X., Zhao, X., & An, Y. (2021). Clinical Application of Metagenomic Next-Generation Sequencing for Suspected Infections in Patients With Primary Immunodeficiency Disease. *Frontiers in Immunology*, 12. <https://doi.org/10.3389/fimmu.2021.696403>
- Tausch, S. H., Strauch, B., Andrusch, A., Loka, T. P., Lindner, M. S., Nitsche, A., & Renard, B. Y. (2018). LiveKraken—real-time metagenomic classification of illumina data. *Bioinformatics*, 34(21), 3750–3752. <https://doi.org/10.1093/BIOINFORMATICS/BTY433>
- Tzec-Interián, J. A., González-Padilla, D., & Góngora-Castillo, E. B. (2025). Bioinformatics perspectives on transcriptomics: A comprehensive review of bulk and single-cell RNA sequencing analyses. *Quantitative Biology*, 13(2), e78. <https://doi.org/10.1002/QUB2.78>
- Varela, M. F., Stephen, J., Lekshmi, M., Ojha, M., Wenzel, N., Sanford, L. M., Hernandez, A. J., Parvathi, A., & Kumar, S. H. (2021). Bacterial Resistance to Antimicrobial Agents. *Antibiotics 2021, Vol. 10, Page 593*, 10(5), 593. <https://doi.org/10.3390/ANTIBIOTICS10050593>
- Vieira, A. de A., Piccoli, B. C., y Castro, T. R., Casarin, B. C., Tessele, L. F., Martins, R. C. R., Schwarzbald, A. V., & Trindade, P. de A. (2023). Pipeline validation for the identification of antimicrobial-resistant genes in carbapenem-resistant *Klebsiella pneumoniae*. *Scientific Reports 2023 13:1*, 13(1), 1–11. <https://doi.org/10.1038/s41598-023-42154-6>
- Wee, S. K., & Yap, E. P. H. (2021a). GALAXY Workflow for Bacterial Next-Generation Sequencing De Novo Assembly and Annotation. *Current Protocols*, 1(9), e242. <https://doi.org/10.1002/CPZ1.242>
- Wee, S. K., & Yap, E. P. H. (2021b). GALAXY Workflow for Bacterial Next-Generation Sequencing De Novo Assembly and Annotation. *Current Protocols*, 1(9), e242. <https://doi.org/10.1002/CPZ1.242>
- Wenger, A. M., Peluso, P., Rowell, W. J., Chang, P. C., Hall, R. J., Concepcion, G. T., Ebler, J., Fungtammasan, A., Kolesnikov, A., Olson, N. D., Töpfer, A., Alonge, M., Mahmoud, M., Qian, Y., Chin, C. S., Phillippy, A. M., Schatz, M. C., Myers, G., DePristo, M. A., ... Hunkapiller, M. W. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology 2019 37:10*, 37(10), 1155–1162. <https://doi.org/10.1038/s41587-019-0217-9>
- Wielders, C. L. C., Fluit, A. C., Brisse, S., Verhoef, J., & Schmitz, F. J. (2002). *mecA* gene is widely disseminated in *Staphylococcus aureus* population. *Journal of Clinical Microbiology*, 40(11), 3970–3975. <https://doi.org/10.1128/JCM.40.11.3970-3975.2002/ASSET/832DF0CD-2AD7-4306-9D31-2B39E1202DB1/ASSETS/GRAPHIC/JM1120335003.JPEG>
- Williams, P. C. M., Qazi, S. A., Agarwal, R., Velaphi, S., Bielicki, J. A., Nambiar, S., Giaquinto, C., Bradley, J., Noel, G. J., Ellis, S., O'Brien, S., Balasegaram, M., & Sharland, M. (2022). Antibiotics needed to treat multidrug-resistant infections in neonates. *Bulletin of the World Health Organization*, 100(12), 797. <https://doi.org/10.2471/BLT.22.288623>
- Wood, D. E., & Salzberg, S. L. (2014). Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, 15(3), 1–12. <https://doi.org/10.1186/GB-2014-15-3-R46/FIGURES/5>
- World Health Organization. (2022). Global Antimicrobial Resistance and Use Surveillance System (GLASS) Report 2022. *World Health Organization*, 71.
- World Health Organization. (2023, November 21). *Antimicrobial resistance*. <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance>

- Wu, D., Jin, L., Xie, J., Liu, H., Zhao, J., Ye, D., & Li, X. (2022). Inhalable antibiotic resistomes emitted from hospitals: metagenomic insights into bacterial hosts, clinical relevance, and environmental risks. *Microbiome*, *10*(1), 1–16. <https://doi.org/10.1186/S40168-021-01197-5/FIGURES/6>
- Yang, H. Y., Nam, Y. S., & Lee, H. J. (2014). Prevalence of Plasmid-Mediated Quinolone Resistance Genes among Ciprofloxacin-Nonsusceptible *Escherichia coli* and *Klebsiella pneumoniae* Isolated from Blood Cultures in Korea. *Canadian Journal of Infectious Diseases and Medical Microbiology*, *25*(3), 163–169. <https://doi.org/10.1155/2014/329541>
- Yang, Y., Che, Y., Liu, L., Wang, C., Yin, X., Deng, Y., Yang, C., & Zhang, T. (2022). Rapid absolute quantification of pathogens and ARGs by nanopore sequencing. *Science of the Total Environment*, *809*. <https://doi.org/10.1016/J.SCITOTENV.2021.152190>
- Yonkus, J. A., Whittle, E., Alva-Ruiz, R., Abdelrahman, A. M., Horsman, S. E., Suh, G. A., Cunningham, S. A., Nelson, H., Grotz, T. E., Smoot, R. L., Cleary, S. P., Nagorney, D. M., Kendrick, M. L., Patel, R., Truty, M. J., & Chia, N. (2022). “Answers in hours”: A prospective clinical study using nanopore sequencing for bile duct cultures. *Surgery*, *171*(3), 693–702. <https://doi.org/10.1016/J.SURG.2021.09.037>
- Zhang, Z., Zhang, Q., Wang, T., Xu, N., Lu, T., Hong, W., Penuelas, J., Gillings, M., Wang, M., Gao, W., & Qian, H. (2022). Assessment of global health risk of antibiotic resistance genes. *Nature Communications* *2022 13:1*, *13*(1), 1–11. <https://doi.org/10.1038/s41467-022-29283-8>
- Zhuang, M., Achmon, Y., Cao, Y., Liang, X., Chen, L., Wang, H., Siame, B. A., & Leung, K. Y. (2021). Distribution of antibiotic resistance genes in the environment. *Environmental Pollution*, *285*, 117402. <https://doi.org/10.1016/J.ENVPOL.2021.117402>