



UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE CIENCIA E INGENIERÍA EN
ALIMENTOS Y BIOTECNOLOGÍA



CARRERA DE BIOTECNOLOGÍA

Análisis de secuencias de genoma completo de *Escherichia coli* aisladas de vegetales y comida callejera de Ecuador

Informe Final de Integración Curricular, Modalidad Proyecto de Investigación, previo a la obtención de título de Ingeniera Biotecnóloga, otorgado por la Universidad Técnica de Ambato, a través de la Facultad de Ciencias e Ingeniería en Alimentos y Biotecnología.

AUTORA: Lissette Katherine Sánchez Gavilanes

TUTOR: PhD. William Ricardo Calero Cáceres

Ambato – Ecuador

Marzo - 2023

APROBACIÓN DEL TUTOR

PhD. William Ricardo Calero Cáceres.

CERTIFICA:

Que el presente Informe Final de Integración Curricular ha sido prolijamente revisado. Por lo tanto, autorizo la presentación de este Informe Final de Integración Curricular bajo la modalidad de Proyecto de Investigación, el mismo que responde a las normas establecidas en el Reglamento de Titulación y Grados de la Facultad de Ciencia e Ingeniería en Alimentos y Biotecnología.

Ambato, 10 de febrero del 2023

PhD. William Ricardo Calero Cáceres.

C.I. 171434885-9

TUTOR

DECLARACIÓN DE AUTENTICIDAD

Yo, Lissette Katherine Sánchez Gavilanes, manifiesto que los resultados obtenidos en el presente Informe Final de Integración Curricular, modalidad Proyecto de Investigación, previo a la obtención del título de Ingeniera Biotecnóloga, son absolutamente originales, auténticos y personales; a excepción de las citas bibliográficas.



Lissette Katherine Sánchez Gavilanes

C.I. 180438450-9

AUTORA

APROBACIÓN DE LOS MIEMBROS DEL TRIBUNAL DE GRADO

Los suscritos Profesores Calificadores, aprueban el presente Informe Final de Integración Curricular, modalidad Proyecto de Investigación, el mismo que ha sido elaborado de conformidad con las disposiciones emitidas por la Facultad de Ciencia e Ingeniería en Alimentos y Biotecnología de la Universidad Técnica de Ambato.

Por constancia firman:

Presidente del Tribunal

B.F. Irvin Ricardo Tubón Usca, PhD.

C.I. 060425035-7

Blga. Helena Maritza De la Torre Olvera, PhD

C.I. 130965199-8

Ambato, 08 de marzo del 2023

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que haga de este presente Informe Final de la Integración Curricular o parte de él un documento disponible para su lectura, consulta y procesos de investigación, según las normas de la institución.

Cedo los derechos en línea patrimoniales de mi Informa Final de la Integración Curricular, con fines de difusión pública, además apruebo la reproducción de este, dentro de las regulaciones de la Universidad, siempre y cuando esta reproducción no suponga una ganancia económica y se realice respetando mis derechos de autor.



Lissette Katherine Sánchez Gavilanes

C.I. 180438450-9

AUTORA

DEDICATORIA

“Todo el mundo tiene estrellas, pero no significan lo mismo para todos. Para algunos, que son viajeros, las estrellas sirven de guía. Para otros no son más que lucecitas en el cielo. Para los sabios, son problemas. Pero todas esas están en silencio. Tú, y solo tú tendrás estrellas como nadie más las tiene”

El Principito

A Dios y la Virgen María, por guiar cada paso que he dado, por cuidarme y nunca dejarme caer.

*Lo más importante en mi mundo es mi **Familia**.*

*A mi **Familia***

Mis padres Elsa y Wilson por ser la luz en mi camino y fortaleza y sabiduría en momentos de fragilidad. Son el vivo reflejo de amor incondicional, les viviré eternamente agradecida.

Mis hermanitas Pao y Andre, el mejor regalo que me han dado mis padres.

Por su ternura y complicidad, las amo con todo mi corazón.

Mis abuelitos, tíos y primos por sus enseñanzas y cariño. A mi ángel en el cielo, porque guardo su sonrisa en mis ojos, su fortaleza en mis huesos y su alegría en mi alma.

Lissette Sánchez

AGRADECIMIENTO

Agradezco a Dios y a la Virgen María por su bendición, por la vida, la salud y la fuerza para enfrentar cada obstáculo y disfrutar cada alegría.

A mis padres, por ser los autores principales en mi vida, por su amor y esfuerzo. A mis hermanitas por los momentos de felicidad y llanto, por ser mi motivación y alegría. A mi hermanito Moisés por siempre estar para mí.

A mi tutor, PhD. William Ricardo Calero Cáceres por ser mi mentor y guía en esta travesía, por su enseñanzas, consejos y apoyo incondicional. Muchísimas gracias por todo.

A mis amigos por su cariño, amistad y apoyo. Por más momentos juntos.

*A mis cómplices en esta travesía Katty, Jenny, Gaby J, Gabby L, Pauli, Chris, Joyce y Anabell.
Que la vida nos vuelva juntar.*

A cada persona que ha ayudado en mi crecimiento profesional, mi más sincero agradecimiento.

A la Universidad Técnica de Ambato y Facultad de Ciencia e Ingeniería en Alimentos y Biotecnología por la formación impartida y la experiencia que me han permitido vivir. A mis docentes por brindarme sus conocimientos en mi formación profesional.

ÍNDICE DE CONTENIDO

<i>APROBACIÓN DEL TUTOR</i>	<i>ii</i>
<i>DECLARACIÓN DE AUTENTICIDAD</i>	<i>iii</i>
<i>APROBACIÓN DE LOS MIEMBROS DEL TRIBUNAL DE GRADO</i>	<i>iv</i>
<i>DERECHOS DE AUTOR</i>	<i>v</i>
<i>DEDICATORIA</i>	<i>vi</i>
<i>AGRADECIMIENTO</i>	<i>vii</i>
<i>RESUMEN</i>	<i>xii</i>
<i>ABSTRACT</i>	<i>xiii</i>
<i>CAPÍTULO 1</i>	<i>1</i>
<i>MARCO TEÓRICO</i>	<i>1</i>
1.1. Antecedentes investigativos	1
1.1.1. Emergencia de la resistencia antimicrobiana	1
1.1.2. Patógenos alimentarios.....	3
1.1.3. <i>E. coli</i> : mecanismos de resistencia y producción de toxinas.	6
1.1.3.1. <i>E. coli</i> : mecanismos de resistencia.....	6
1.1.3.2. <i>E. coli</i> : productora de toxina Shiga (STEC).....	9
1.1.4. Investigación epidemiológica de <i>E. coli</i>	10
1.1.4.1. Técnicas de tipificación bacteriana	10
1.1.4.2. Secuenciación de próxima generación (NGS)	11
1.1.4.3. Aplicación de la secuenciación de próxima generación: Secuenciación de Genoma Completo (WGS)	12
1.1.5. Análisis de secuencias de genoma completo. Herramientas bioinformáticas empleadas en el análisis de secuencias de genoma completo	13
1.2. Objetivos	19
1.2.1. Objetivo General	19
1.2.2. Objetivos Específicos.....	19

<i>CAPITULO 2</i>	20
<i>METODOLOGÍA</i>	20
2.1. Materiales	20
2.2. Métodos	22
2.2.1. Análisis de la calidad del genoma de las secuencias ensambladas	23
2.2.2. Caracterización de secuencias	24
2.2.3. Determinación de profagos y elaboración de gráficos circulares	27
2.2.4. Comparación de secuencias de genoma completo de diversos orígenes geográficos y elaboración de árboles filogenéticos	28
<i>CAPÍTULO 3</i>	31
<i>RESULTADOS Y DISCUSIÓN</i>	31
3.1. <i>Análisis y discusión de los resultados</i>	31
3.1.1. Análisis de secuencias a través de plataformas bioinformáticas	31
3.1.1.1. Análisis de secuencias de genoma completo	31
3.1.1.2. Caracterización de secuencias de genoma completo (WGS) de <i>E. coli</i> portadores de la toxina Shiga.	33
3.1.1.3. Caracterización genotípica de secuencias de genoma completo de <i>E. coli</i> portadoras de betalactamasas de espectro extendido.	41
3.1.2. Evaluación de la distancia filogenética entre aislamientos de <i>E. coli</i> de origen ecuatoriano con aislamientos de diverso origen geográfico	51
3.1.2.1. Estructura de población basada en la estrategia cgMLST	51
3.1.2.2. Caracterización de secuencias de diverso origen geográfico y análisis comparativo con secuencias de origen local	57
<i>CAPÍTULO 4</i>	62
<i>CONCLUSIONES Y RECOMENDACIONES</i>	62
4.1. Conclusiones	62
4.2. Recomendaciones	64
<i>MATERIALES DE REFERENCIA</i>	65

<i>Referencia Bibliográficas</i>	65
<i>Anexos</i>	86

ÍNDICE DE TABLAS

Tabla 1 Toxinas bacterianas, fuentes asociadas y toxicidad que generan en células eucariotas.....	5
Tabla 2 Recursos informáticos	20
Tabla 3 Equipos	22
Tabla 4 Identificación de las secuenciaciones del genoma completo de 2 cepas de <i>E. coli</i> aisladas de vegetales y comida callejera de Ecuador portadoras de genes toxina Shiga y ARGs.....	34
Tabla 5 Caracterización basada en la WGS de 2 cepas de <i>E. coli</i> aisladas de vegetales y comida callejera de Ecuador portadoras de genes toxina Shiga y ARGs	35
Tabla 6 Identificación de las secuenciaciones del genoma completo de 3 cepas de <i>E. coli</i> aisladas de vegetales y comida callejera de Ecuador portadoras de ARGs y genes de virulencia	41
Tabla 7 Caracterización basada en la secuenciación del genoma completo de 3 cepas de <i>E. coli</i> aisladas de vegetales y comida callejera de Ecuador portadoras de ARGs y genes de virulencia	43

ÍNDICE DE FIGURAS

Figura 1 Línea de tiempo. Relación entre la introducción de nuevos antibióticos e identificación de resistencia antimicrobiana	2
Figura 2 Transferencia genética horizontal entre bacterias	8
Figura 3 Mecanismo de secuenciación de próxima generación.....	12
Figura 4 Mapa circular de la secuencia genómica completa de la cepa de <i>E. coli</i> C5b1 aislada de col blanca (<i>B. oleracea</i>)	36

Figura 5 Mapa circular de la secuencia genómica completa de la cepa de <i>E. coli</i> T5bX aislada de tomate (<i>S. lycopersicum</i>)	37
Figura 6 Mapa circular de la secuencia genómica completa de la cepa de <i>E. coli</i> N1-1 aislada de comida callejera (Ensalada).	44
Figura 7 Mapa circular de la secuencia genómica completa de la cepa de <i>E. coli</i> J5_2C2 aislada de juego de caña.....	45
Figura 8 Mapa circular de la secuencia genómica completa de la cepa de <i>E. coli</i> L6by aislada de lechuga (<i>L. sativa</i>).	49
Figura 9 Árbol de expansión mínima (Minimum-spanning tree) para 50 aislados basados en el cgMLST de <i>E. coli</i>	52
Figura 10 Árbol de filogenia rectangular comparativo y predicción de genes de virulencia para 50 aislados de <i>E. coli</i> de diverso origen geográfico	58
Figura 11 Árbol de filogenia rectangular comparativo y predicción de ARGs para 50 aislados de <i>E. coli</i> de diverso origen geográfico.....	59
Figura 12 Árbol de filogenia rectangular comparativo y predicción de plásmidos para 50 aislados de <i>E. coli</i> de diverso origen geográfico.....	60

ÍNDICE ANEXOS

Anexo 1 Tabla de identificación de las secuenciaciones del genoma completo de 13 cepas de <i>E. coli</i> aisladas de vegetales y comida callejera de Ecuador	86
Anexo 2 Tabla de información relacionada a la calidad del genoma para las 13 secuencias de genoma completo aisladas de vegetales y comida callejera de Ecuador	88
Anexo 3 Tabla de información compilada de secuencias de genoma completo de diverso origen geográfico y de origen local	89
Anexo 4 Códigos de identificación (CI) y tipo de complejo clonal (ST Complex) ..	94

RESUMEN

La secuenciación de genoma completo (WGS) es una herramienta fundamental para el análisis y detección de genes de resistencia antimicrobiana (ARGs), factores de virulencia y elementos genéticos móviles (MGE) en genomas bacterianos. No obstante, la aplicación de WGS en Ecuador se desarrolla a través de la academia, mas no en la adopción de estas tecnologías como mecanismos de vigilancia y control epidemiológico en los sistemas de salud. El presente estudio se enfocó en la caracterización *in silico* de secuencias de genoma completo de *Escherichia coli* aisladas de vegetales y comida callejera procedentes de mercados de las ciudades Ambato y Riobamba, siguiendo como línea de investigación Microbiología y Biotecnología. Se emplearon secuencias de longitudes de lectura de extremo emparejado de 300 pb. La caracterización *in silico* se realizó a través de herramientas bioinformáticas empleadas para la detección de ARGs, virulencia y MGE. Los resultados de patogenicidad indicaron que 2 cepas portaban los genes *stx2* y diversos factores de virulencia. Adicionalmente, se identificó que 3 cepas portaban los genes *blaTEM* y *blaCMY* y diferentes determinantes de RAM que confieren resistencia a aminoglucósidos, quinolonas, tetraciclinas y sulfametoxazol. Los análisis de filogenia indicaron que las interacciones evolutivas de ARGs, factores de virulencia y tipos de plásmidos no se correlacionaron con el origen geográfico y fuente de aislamiento, mostrando una elevada diversidad. Estos hallazgos destacan la necesidad de emplear WGS en la vigilancia y control de *E. coli* patogénica en matrices alimentarias, teniendo en cuenta una perspectiva “One Health”.

Palabras clave: Secuenciación de genoma completo, *Escherichia coli*, resistencia, virulencia, plásmidos, fagos, filogenia.

ABSTRACT

Whole genome sequencing (WGS) is a fundamental tool for the analysis and detection of antimicrobial resistance genes (ARGs), virulence factors, and mobile genetic elements (MGEs) in bacterial genomes. Despite their advantages for epidemiological analysis, applying WGS in Ecuador is scarce and limited to Academia. The present study focused on in silico characterization of WGS of *Escherichia coli* isolated from vegetables and street food from markets in Ambato and Riobamba. The research line of this research is Microbiology and Biotechnology. For this study, paired-end read length sequences of 300 bp were used. In silico characterization was carried out through bioinformatics tools to detect ARGs, virulence, and MGE genes. Two isolates harbor the *stx2* genes and various virulence factors. It was also identified that three strains carried the *bla*_{TEM} and *bla*_{CMY} genes, and different ADR determinants confer resistance to aminoglycosides, quinolones, tetracyclines, and sulfamethoxazole. Phylogenetic analysis indicated that the evolutionary interactions of resistance genes, virulence factors, and plasmid types did not correlate with the geographic origin and source of isolation, showing high diversity. These findings highlight the need to use WGS as a surveillance tool to analyze pathogenic *E. coli* in food matrices, taking into account a "One Health" perspective.

Keywords: Whole genome sequencing, resistance, virulence, plasmids, phages, phylogeny.

CAPÍTULO 1

MARCO TEÓRICO

1.1. Antecedentes investigativos

1.1.1. Emergencia de la resistencia antimicrobiana

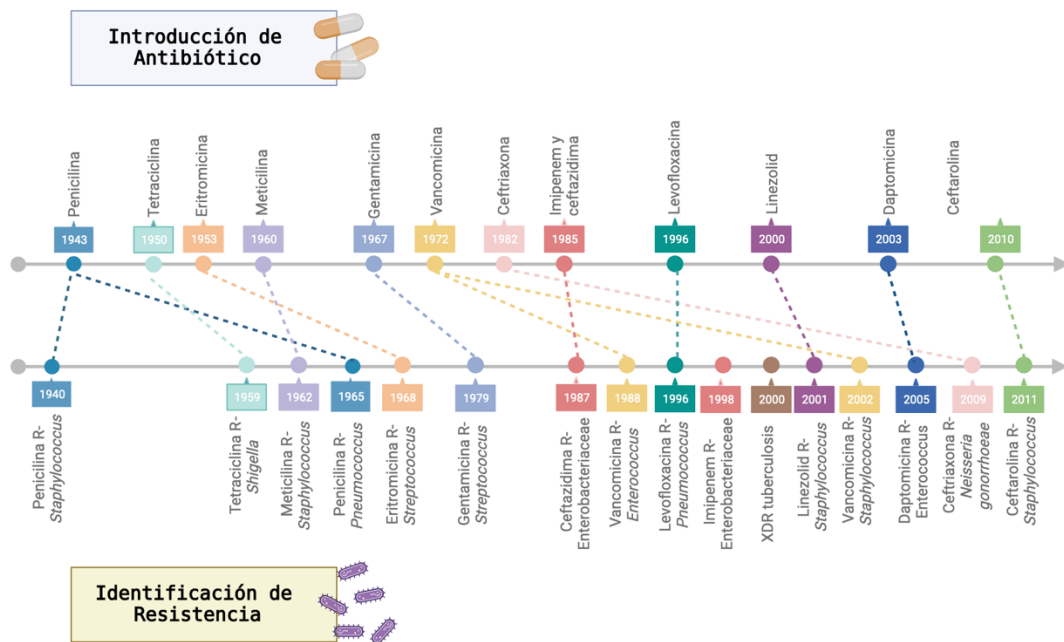
A lo largo de los últimos años, la resistencia a los antimicrobianos (RAM) ha sido considerada una de las principales amenazas mundiales para la salud pública (**World Health Organization, 2015**). Desde principios del siglo XX, cuando se descubrieron los primeros antibióticos modernos, los antimicrobianos han salvado millones de vidas en el mundo, siendo desde entonces una de las bases centrales en la medicina (**The World Bank & World Health Organization, 2022**). Sin embargo, el uso persistente de los antibióticos a nivel mundial ha originado que muchos antimicrobianos hayan perdido su eficacia, dada la presión selectiva que estos agentes antagonistas ejercen sobre las bacterias, derivando en una aceleración del ritmo evolutivo, generándose episodios de mutaciones y el desarrollo de nuevos mecanismos de resistencia (**Buelow et al., 2021; Rodgers et al., 2018**).

En este sentido, estadísticas presentadas en el año 2019, exhibieron que 4.95 millones de muertes fueron provocadas por enfermedades asociadas a RAM bacteriana, y estimaciones para el año 2050 prevé que esta cifra haya incrementado a 20 millones, dejando pérdidas económicas superiores a los 2,9 billones de dólares (**The World Bank & World Health Organization, 2022; Uddin et al., 2021**). Bajo este contexto, RAM se considera una amenaza de alto impacto para la salud de los seres humanos, animales, medioambiente y economía mundial (**Pan American Health Organization**

& World Health Organization, 2021). Alexander Fleming en su discurso de aceptación del premio Nobel en 1945 mencionó “Existe el peligro de que un hombre pueda aplicar una dosis insuficiente de antibiótico, y al exponer a este microorganismo a una cantidad no letal del mismo, los haga resistentes” (Fleming, 1945). Desafortunadamente, en la actualidad este vaticinio es una realidad. La presión selectiva antimicrobiana, permite que las bacterias adquieran ARGs (del inglés, *antibiotic resistance genes*) y elementos genéticos móviles (MGE, del inglés *mobile genetic elements*) que pueden transferirse a otras bacterias por diferentes mecanismos de transferencia genética horizontal (HGT, del inglés *horizontal gene transfer*). Cabe recalcar que las bacterias al adquirir esta RAM también adquieren una mayor capacidad de proliferar en animales, humanos y entorno natural (Velazquez-Meza et al., 2022).

Figura 1

Línea de tiempo. Relación entre la introducción de nuevos antibióticos e identificación de resistencia antimicrobiana



Adaptado de (CDC, 2019)

En retrospectiva, las primeras bacterias resistentes identificadas fueron de origen clínico, tal es el caso de *Staphylococcus aureus*, una bacteria resistente a la penicilina identificada como resistente pocos años después de la inserción del antimicrobiano en hospitales de Londres en el año de 1940 (**Levy & Bonnie, 2004**). Años posteriores, la introducción de nuevos medicamentos semisintéticos empleados para combatir bacterias resistentes y el sobreuso de los mismos, acrecentó el número de bacterias resistentes y multirresistentes, por ejemplo, las *Enterobacterias* a saber: *E. coli*, *Shigela* y *Salmonella* fueron identificadas como resistentes a múltiples fármacos en la década de 1950 y 1960 (**Watanabe, 1963**). RAM ha sido considerada una creciente fuente de preocupación a través de los años, puesto que la implementación de un antibiótico no tarda en establecer resistencia, y en adición, se ha identificado que más del 70% de las bacterias presenta resistencia al menos a un antibiótico (**Economou & Gousia, 2015**).

Ante una posible crisis, la Organización Mundial de la Salud (OMS), emitió una acción global acerca de RAM en 2015, adoptando un enfoque “One Health” (**World Health Organization, 2015**). One Health describe un enfoque multidisciplinario conjunto para proporcionar soluciones direccionadas a una salud óptima en personas, animales y medio ambiente (**Velazquez-Meza et al., 2022**). De esta forma, se han desarrollado sistemas de vigilancia e investigaciones frente al uso y consumo de antimicrobianos en la población humana, animal y sus interconexiones, personas, animales, plantas y entorno compartido, a fin de abordar la problemática a nivel regional, nacional y mundial (**FAO et al., 2022**).

1.1.2. Patógenos alimentarios

A lo largo de la historia, los seres humanos han empleado microorganismos y sus subproductos para producir y conservar alimentos mediante la fermentación. No obstante, algunos de estos microorganismos deterioran significativamente los alimentos y provocan enfermedades transmitidas por lo alimentos (**Gourama, 2020**).

Las enfermedades transmitidas por los alimentos (ETAs) son un gran desafío para los sistemas de salud pública a nivel mundial puesto que estas provocan una morbilidad, mortalidad y costos económicos considerables, en especial en países en vías de desarrollo (FAO & WHO, 2021).

La frecuencia e importancia de las enfermedades transmitidas por los alimentos depende directamente de la interacción entre la matriz alimentaria, el patógeno de transmisión por los alimentos, el huésped y el medio ambiente (Mousavi Khaneghah et al., 2020). Se estima que cada año 1 de cada 10 personas en el mundo padecen de enfermedades después de consumir alimentos contaminados por diversos agentes (bacterias, virus, parásitos, toxinas y productos químicos), de estas, 420,000 personas mueren y causan una pérdida de 33 millones de DALYS (años de vida saludable) a nivel global (World Health Organization, 2015). En Ecuador, de acuerdo a la Gaceta Epidemiológica del Ministerio de Salud, se reporta un promedio de 12,203 casos de enfermedades bacterianas transmitidas por alimentos y agua para el año 2019, y 6,728 casos en el año 2021 (Ministerio de Salud Pública, 2022).

Dado el mecanismo por el cual estos patógenos transmitidos por los alimentos causan enfermedades, las enfermedades transmitidas por medio de matrices alimentarias se pueden clasificar en tres tipos: infecciones, infecciones tóxicas e intoxicaciones transmitidas por los alimentos (Gourama, 2020). Algunos de los patógenos responsables de la mayoría de los brotes reportados de ETAs son Norovirus, *Campylobacter spp.*, *Salmonella spp.*, *Listeria monocytogenes* y *E. coli* productora de toxinas Shiga (Curtis et al., 2014). Las infecciones tóxicas son provocadas por toxinas producidas por el patógeno ingerido dentro del huésped (CDC, 2017; Zuñiga & Caro, 2017). Los microorganismos producen dos tipos de toxinas, endotoxinas: son toxinas de lipopolisacáridos que forman parte de la membrana externa o de la pared celular de bacterias Gram negativas; y exotoxinas: son proteínas tóxicas sintetizadas dentro de las células que se excretan por bacterias, hongos, algas y protozoos Gram positivos y Gram negativos (Madigan et al., 2016).

Las enterotoxinas son exotoxinas liberadas por el patógeno en el intestino, que afecta principalmente la permeabilidad de las células epiteliales de la pared intestinal (**Martin, 2012**). Algunos ejemplos de estos microorganismos son *E. coli*, *Clostridium perfringens*, *Vibrio cholera* y *S. aureus*, con frecuencia estos microorganismos generan toxinas capaces de formar poros en las membranas celulares, induciendo la muerte celular del hospedador (**Madigan et al., 2016**). En la **Tabla 1** se identifican toxinas bacterianas, sus fuentes asociadas y la toxicidad que estas generan en células eucariotas.

Tabla 1

Toxinas bacterianas, fuentes asociadas y toxicidad que generan en células eucariotas

Organismo	Toxina	Fuente	Toxicidad
<i>C. perfringens</i>	Alfa, Beta, Épsilon y Iota Enterotoxina de <i>C. perfringens</i> (CPE).	Suelo, agua, e intestinos de mamíferos.	Intoxicación alimentaria, enterocolitis, gangrena gaseosa.
<i>C. botulinum</i>	Toxina C2.	Suelo, agua, productos cárnicos, miel, intestinos de mamíferos.	Intoxicación alimentaria, botulismo infantil y adulto.
<i>B. anthracis</i>	Antígeno protector, toxina del edema, toxina letal.	Tierra, productos animales contaminados (huesos o pieles).	Edema y necrosis cutánea, shock, insuficiencia cardíaca, disfunción respiratoria grave.
<i>E. coli</i>	Toxina Shiga	Alimentos, animales, agua, personas y medio ambiente contaminados.	Diarrea leve sin sangre a grave, síndrome hemolítico urémico, insuficiencia renal.

Adaptado de (Cetin, 2020).

1.1.3. *E. coli*: mecanismos de resistencia y producción de toxinas.

E. coli es un bacilo anaerobio facultativo Gram negativo perteneciente a la familia *Enterobacteriaceae* (**Gonzalez-Escalona et al., 2019**). En seres humanos, *E. coli* es comensal del tracto intestinal; no obstante, es capaz de sobrevivir fuera del tracto intestinal como en el suelo, agua y alimentos (**Vila et al., 2016**). Algunas cepas de *E. coli* no suelen ser patógenas, mientras que otras pueden causar una variedad de enfermedades en humanos y animales. En la actualidad la amenaza identificada sobre dicho patógeno radica a la creciente prevalencia de *E. coli* RAM y a la capacidad que tiene estos para producir toxinas.

1.1.3.1. *E. coli*: mecanismos de resistencia

En general, las bacterias resisten la acción inhibitoria de los antibióticos empleando dos mecanismos: la resistencia intrínseca y la adquirida. La resistencia intrínseca es la capacidad natural que tienen los microorganismos para resistir a un determinado antimicrobiano empleando sus características estructurales y funcionales (**Cox & Wright, 2013**). Los mecanismos de resistencia adquirida se desarrollan a partir de un cambio en el ADN bacteriano que desemboca un nuevo rasgo fenotípico.

Los mecanismos intrínsecos normalmente se encuentran codificados por cromosomas, e incluyen bombas de expulsión activa o de eflujo (mecanismo capaz de modular o eliminar la actividad de varios tipos o familias de antibióticos), enzimas que inactivan antibióticos o mecanismos que sirven como barreras de permeabilidad (**Fajardo et al., 2008; Nikaido & Takatsuka, 2009**). Las bacterias Gram negativas, al poseer una segunda membrana lipídica posterior a la capa de peptidoglicano, permite generar mayor resistencia intrínseca, ocasionando reducir la permeabilidad al encontrarse expuesta a moléculas hidrófobas (**Andersson & Hughes, 2017**).

Algunos ejemplos de los mecanismos intrínsecos son: la bomba de eflujo AcrAB-TolC de *E. coli* capaz de expulsar diferentes antibióticos dada la amplia especificidad de sustrato, entre ellos los β -lactámicos dianónicos (Chowdhury et al., 2019); la barrera de permeabilidad impuesta por la membrana externa en *E. coli* resistente a la vancomicina permite describir otro mecanismo de resistencia intrínseca (Krishnamoorthy et al., 2016).

Los mecanismos de resistencia adquiridos pueden ser mediados por dos vías: mutaciones o con mayor frecuencia por HGT (Saha & Sarkar, 2021). La resistencia adquirida por mutaciones son procesos espontáneos que no necesariamente se generan tras la exposición a antibióticos (Sekyere & Asante, 2018). Las mutaciones pueden generar resistencia al emplear los siguientes mecanismos: 1) modificación o eliminación de la proteína a la cual el agente antimicrobiano se adhiere; 2) regulación en la producción de enzimas inactivantes del agente antimicrobiano; 3) regulación o modificación de las proteínas de membrana externa necesarias para permitir el ingreso de los antibióticos a las células; 4) regulación de las bombas de eflujo (Calero Cáceres, 2017; Morrison & Zembower, 2020).

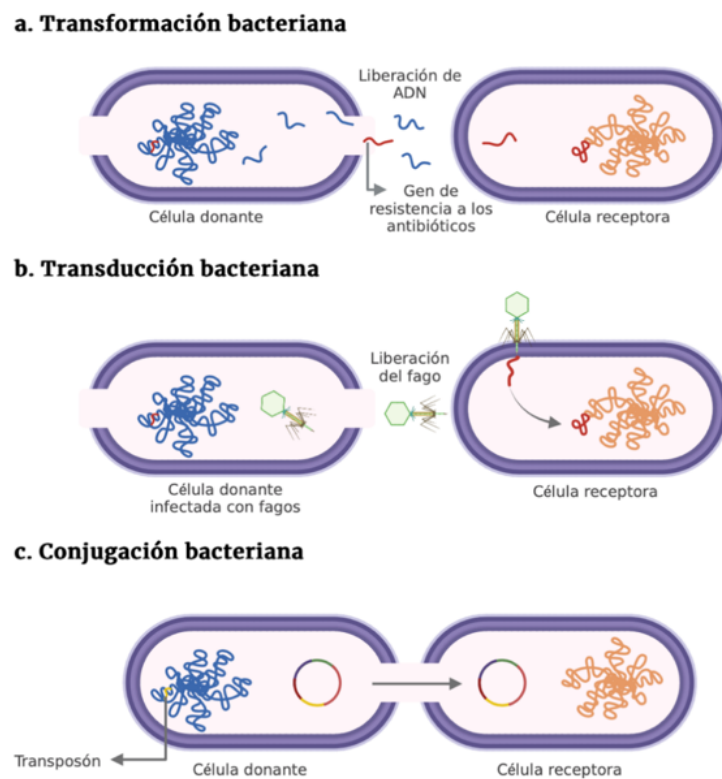
Por otro lado, la HGT es un fenómeno de intercambio genético ya sean entre bacterias u otros linajes de organismos (Burmeister, 2015). Los mediadores de HGT son los MGE. Los MGE son segmentos de ADN que codifican enzimas y proteínas capaces de cambiar de lugar entre cromosomas bacterianos (Dobrindt et al., 2015). Existen diferentes clases de HGT, describiéndose: transposones, plásmidos, bacteriófagos, agentes de transferencia genética (GTAs, del inglés *gene transfer agents*), integrones, elementos de integración conjugativa (ICEs, del inglés *integrative conjugative elements*), islas genómicas (Gis, del inglés *genomic islands*) (Frost et al., 2005; Partridge et al., 2018).

En la **Figura 2**, se identifican los mecanismos de intercambio genético; la transferencia de genes determinantes en la resistencia a los antibióticos en *E. coli* se desarrolla

mediante mecanismos de intercambio genético que involucran la transformación con ADN libre, transducción por bacteriófagos (virus que infectan bacterias) o conjugación que involucra plásmidos (moléculas de ADN bacteriano que coexiste de manera independiente al cromosoma de la célula, y son los principales transportadores de genes entre bacterias) (Clark et al., 2019; Emamalipour et al., 2020; Partridge et al., 2018).

Figura 2

Transferencia genética horizontal entre bacterias



Adaptado de (Guo, 2020)

E. coli es considerada un indicador para la resistencia antimicrobiana de las bacterias Gram negativas. Dentro de los mecanismos de resistencia de *E. coli* destaca la resistencia a β -lactámicos, dada la codificación de enzimas como las β -lactamasas de espectro extendido (ESBL, del inglés *extended-spectrum beta-lactamase*) encargadas de degradar estos antimicrobianos (Krishnamoorthy et al., 2016). Algunos de los ARGs localizados en *E. coli* de prevalencia que codifican la resistencia a β -lactámicos son: *bla*_{CTX}, *bla*_{TEM} y *bla*_{SHV} (Khalifa et al., 2021).

1.1.3.2. *E. coli*: productora de toxina Shiga (STEC)

STEC es un grupo de *E. coli* patógenas, antiguamente identificadas como *E. coli* enterohemorrágicas (EHEC). El primer brote reportado por STEC ocurrió en 1982, en el cual el serotipo *E. coli* 0157:H7 fue causante de dos brotes de colitis hemorrágica; en la actualidad, este serotipo es el causante de la mayoría de los brotes asociados a STEC. STEC es productora de toxinas reconocidas como verotoxinas o Stxs (**Zhang et al., 2021**).

Las Stx son toxinas proteicas bacterianas codificadas por los genes *stx*. Se encuentran formadas por dos subunidades principales: subunidad A, encargada de lesionar el ribosoma eucariota y detener la síntesis de proteínas en células diana y; pentámero de subunidad B, responsable de mediar la unión de toxinas al receptor presente en células endoteliales Gb3 (globotriaosilceramida) (**Fraser et al., 2004**). Las Stx se clasifican en Stx1 y Stx2. El mecanismo de acción de estas toxinas es el mismo, pero genéticamente e inmunológicamente se diferencian entre sí, exteriorizando un 55-60% de identidad genética y de aminoácidos (**Lee et al., 2007; Zhang et al., 2021**).

Las infecciones provocadas por STEC son reconocidas como un problema de salud pública importante alrededor del mundo. Según datos recopilados en el año 2010, han permitido identificar que STEC transmitida por los alimentos desencadenó más de 1 millón de enfermedades humanas, 128 muertes y una pérdida de 13 000 DALYS (**World Health Organization & Food and Agriculture Organization of the United Nations, 2018**).

1.1.4. Investigación epidemiológica de *E. coli*

Las investigaciones de brotes generalmente vinculan las etiologías a alimentos específicos, entre ellos alimentos crudos de origen animal, derivados de animales sin procesar, frutas, verduras, entre otros (CDC, 2022; Luna-Guevara et al., 2019); este precedente ha permitido que los sistemas de salud pública, las agencias reguladoras y la industria alimentaria investiguen y desarrollen programas epidemiológicos en muchos países (Dewey-Mattia et al., 2019). Sin embargo, en Ecuador no se realizan estudios epidemiológicos enfocados en la detección de reservorios de RAM o de vías de diseminación de RAM en la cadena alimentaria, con excepción de ciertos estudios realizados por la academia (Barragán-Fonseca et al., 2022; Tubón et al., 2022; Zurita, Yáñez, Sevillano, Ortega-Paredes, et al., 2020).

En 2015, la Organización Mundial de la Salud (OMS) lanzó el Sistema mundial de vigilancia de la resistencia y el uso de antimicrobianos (GLASS) para estandarizar las metodologías de vigilancia de RAM, incluidas las directrices para la recopilación, el análisis, la interpretación y el intercambio de datos de AMR de patógenos prioritarios por países a nivel clínico (World Health Organization, 2015). La iniciativa GLASS sugiere utilizar enfoques de secuenciación del genoma completo (WGS) para complementar la caracterización fenotípica de patógenos prioritarios, con la expectativa de que esto conduzca a una mejor comprensión de la transmisión y la interrelación de las poblaciones microbianas (World Health Organization, 2020).

1.1.4.1. Técnicas de tipificación bacteriana

La tipificación bacteriana es la capacidad de diferenciación de bacterias a nivel de especies o subespecies. La tipificación bacteriana se puede dar empleando métodos fenotípicos (caracterización basada en atributos expresados visualmente) y genotípicos (caracterización de contenido genético) (Harris & Okoro, 2014). La caracterización

bacteriana a nivel genotípico provee una descripción de contenido genético presente en la bacteria, principalmente la identificación de ADN plasmídico en el cual se pueden localizar ARGs, virulencia o cualquier tipo de gen de interés (**Espitia-Navarro et al., 2019**). Existen diferentes niveles de caracterización genética: caracterización de todo el contenido genético bacteriano, caracterización del ambiente cromosómico, tipificación de vectores plasmídicos o caracterización de una pequeña parte de ADN bacteriano de interés (**Ann-Luna, 2016**). La tipificación emplea dos métodos de genotipado: 1) método basado en bandas; 2) método basado en secuencias (**Oyarzabal & Kathariou, 2014**).

Los métodos de genotipado basados en tecnologías de secuenciación han incrementado la disponibilidad de datos de secuencias bacterianas, convirtiendo a la secuenciación en una herramienta factible para la vigilancia de RAM y patógenos potenciales (**Boolchandani et al., 2019**). Bajo esta línea, la secuenciación permite identificar la secuencia concreta de nucleótidos que conforman el ADN o ARN (**Wood et al., 2020**). Empleando la secuenciación se puede analizar los polimorfismos de un solo nucleótido (SNPs, del inglés *single nucleotide polymorphism*) o analizar secuencias buscando caracteres repetidos. La tipificación multilocus de secuencias (MLST, del inglés *multi-locus sequence typing*), secuenciación de genes específicos, secuenciación del genoma completo (WGS, del inglés *whole genome sequencing*), son algunos de los métodos de secuenciación emergente (**Satpathy et al., 2019**).

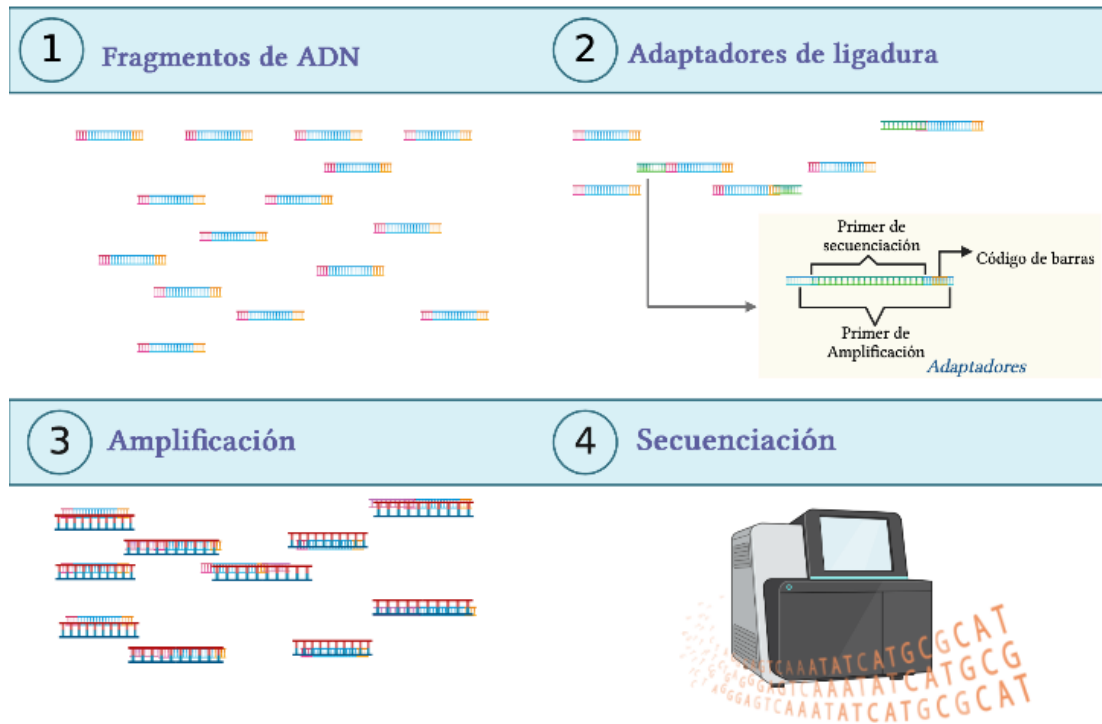
1.1.4.2. Secuenciación de próxima generación (NGS)

La secuenciación de próxima generación (NGS, del inglés *next generation sequencing*) engloba un conjunto de tecnologías destinadas a analizar la secuenciación masiva y paralela de millones de secuencias de una sola cadena individual de ADN analizada por separado, pero de forma simultánea (**Fletcher, 2021**). Esta secuenciación se diferencia del resto (secuenciación primera, segunda y tercera generación) dado que se amplifican los fragmentos de ADN antes de empezar a secuenciar. Algunas de los

equipamientos o plataformas empleadas en NGS son: Roche, 454 Life Sciences (GS FLX titanium), Life Technologies (Ion Torrent & Ion Proton) e Illumina (HiSeq, MiSeq, GenomeAnalyzer), Nanopore, PacBio (Pervez et al., 2022; Udaondo et al., 2021; Y. Wang et al., 2021).

Figura 3

Mecanismo de secuenciación de próxima generación



1.1.4.3. Aplicación de la secuenciación de próxima generación: Secuenciación de Genoma Completo (WGS)

WGS es un método integral de NGS utilizado para analizar genomas completos; es decir, permite determinar el orden exacto de las bases nitrogenadas (A, T, C y G) en el genoma de un organismo (CDC, 2022; Yin et al., 2019). WGS en la actualidad se emplea como un mecanismo clínico de rutina para el control y estudio de infecciones, así como para la caracterización de patógenos y ARGs (Oniciuc et al., 2018). Existe una gran gama de herramientas bioinformáticas para predecir RAM a partir de datos de WGS en tiempo real, dada la disponibilidad creciente de programas que permiten

ensamblar, tipificar e identificar genotipos específicos de interés. De este modo, debido a las grandes mejoras tecnológicas de secuenciación, WGS se ha convertido en una de las herramientas base para la ejecución de investigaciones epidemiológicas (Köser et al., 2014).

1.1.5. Análisis de secuencias de genoma completo. Herramientas bioinformáticas empleadas en el análisis de secuencias de genoma completo

Las herramientas bioinformáticas permiten ensamblar o mapear la lectura, al genoma de referencia (Nodehi et al., 2021). La gran cantidad de softwares o herramientas para realizar ensamblajes o mapear las lecturas están en sistema operativo Unix o Linux de forma que se emplean líneas de comandos; solo una pequeña proporción se encuentran en Sistema Operativo de Disco (DOS, del inglés *disk operating system*) utilizable en Windows con una interfaz menos compleja (Joppich & Zimmer, 2019). Pese a esto, en la actualidad ya existen diferentes herramientas bioinformáticas basadas en la web que permiten subir los datos de secuencias completas y analizar los mismos sin el empleo de líneas de comandos.

Herramientas bioinformáticas dependientes de CGE

El centro de Epidemiológica Genómica (CGE, del inglés *Center for Genomic Epidemiology*), proporciona una base de datos bioinformáticos diseñados con la finalidad de analizar e interpretar datos de secuencias completas empleando diferentes servicios (Larsen et al., 2017). Algunos de los servicios bioinformáticos webs desarrollados por CGE, se describen a continuación:

- **MLST 2.0**

Representa uno de los métodos más utilizados para la tipificación de microorganismos patógenos. Consiste en el análisis de las secuencias de fragmentos de aproximadamente 450 pb de por lo menos 7 genes *housekeeping*, cuyas variaciones permiten diferenciar las distintas cepas de la misma especie (Díaz et al., 2018). El programa MLST versión 2.0 es una herramienta bioinformática que analiza los perfiles de MLST para 66 especies bacterianas utilizando como datos de entrada WGS que puede ser ingresado en formato FASTA y FASTQ (Larsen et al., 2012).

- **ResFinder 4.1.**

ResFinder es una de las herramientas bioinformáticas de la plataforma CGE de libre acceso, que permite identificar ARGs completos, y/o mutaciones cromosómicas específicas que median la resistencia a antimicrobianos en las secuencias parciales o totales del ADN bacteriano (lecturas sin procesar y genomas ensamblados) (Florensa et al., 2022). Esta herramienta bioinformática, emplea bases de datos seleccionadas, bases de datos públicas y de artículos ResFinder puede aceptar datos en forma de genomas ensamblados en formato FASTA y en forma de lecturas crudas, en formato FASTQ (Carattoli et al., 2014).

Contiene cuatro bases de datos, la primera enfocada en identificar ARGs, la segunda direccionada en identificar mutaciones cromosómicas puntuales, la tercera base de datos se encuentra ligada a la traducción del genotipo a fenotipo y la última base de datos, es un panel específico para simular pruebas de susceptibilidad (Bortolaia et al., 2020). Esta última base de datos implementa antibióticos *in silico* que simulan pruebas de susceptibilidad a antimicrobianos, semejante al mecanismo empleado en las pruebas de susceptibilidad fenotípicas (Clausen et al., 2018).

- **PlasmidFinder 2.1.**

Según **Carattoli et al., (2014)**, PlasmidFinder 2.1. es una herramienta web que contiene una base de datos de consenso que integra replicones de 559 plásmidos completamente secuenciados. Se construyó a partir de 126 secuencias diseñadas en plásmidos completamente secuenciados disponibles en la base de datos de nucleótidos del NCBI en 2014 y se ha actualizado continuamente para incluir nuevos replicones de plásmidos secuenciados asociados con la familia *Enterobacteriaceae*. PlasmidFinder se puede utilizar para el análisis de secuencias de replicones de datos de secuenciación en bruto (fastq) y ensamblados (fasta) (**Carattoli et al., 2014**).

- **VirulenceFinder 2.0.**

VirulenceFinder 2.0 es un servidor que permite la identificación de genes de virulencia conocidos en secuencias de nucleótidos (**Besser et al., 2018**). Por otra parte, esta herramienta es empleada para examinar secuencias de genoma de fagos, profagos, detección de ARGs y virulencia (**Kleinheinz et al., 2014**).

- **CSI Phylogeny 1.4**

Herramienta bioinformática empleada en la clasificación filogenética basada en la detección de polimorfismos de nucleótido único (SNPs), infiriendo una filogenia justificada en la alineación concatenada de SNP de alta calidad. Por otro lado, CSI Phylogeny 1.4 facilita la generación de árboles filogenéticos desarrollados en base a SNP localizados entre los aislados ingresados (**Ahrenfeldt et al., 2017**).

Herramientas bioinformáticas independientes de CGE

- **Centro de Recursos de Bioinformática Bacteriana y Viral (BV-BRC)**

Sistema de información diseñado con la finalidad de desarrollar investigaciones sobre enfermedades infecciosas bacterianas y virales. Esta plataforma alberga datos sobre la estructura y función de proteínas, estudios clínicos, resistencia a medicamentos, epidemiología, entre otros; por otra parte, proporciona herramientas de código abierto para el análisis de datos y anotaciones genómicas (**Olson et al., 2023**).

- **PHASTER**

Servidor web desarrollado para la identificación y anotación de secuencias de profagos dentro de genomas y plásmidos bacterianos (**Arndt et al., 2016**).

- **Proksee**

Herramienta bioinformática web empleada en el ensamblaje, generación de anotación y visualización del genoma (**Proksee, 2022**).

- **BacWGSTdb 2.0**

BacWGSTdb 2.0 es una base de datos de libre acceso empleada en la tipificación de secuencias de genoma completos bacterias. La gran diversidad de recursos para los datos de secuenciación genoma bacterianos y metadatos recuperados de NCBI GenBank y BioSample con los que cuenta esta base de datos, permiten la vigilancia epidemiológica genómica de los patógenos bacterianos (**Feng et al., 2021**).

- **Galaxy Sciensano**

Es un servidor Galaxy que cuenta con herramientas bioinformáticas, flujos de trabajo y bases de datos empleadas en WGS para la tipificación y caracterización de patógenos. Algunas de las aplicaciones se destaca el ensamblaje, mapeo de lectura, alineación de secuencias, entre otros. En sus actualizaciones recientes permite el análisis de datos provenientes de Illumina, permitiendo tener una alta compatibilidad con esta plataforma (**Galaxy, 2023**).

- **PHYLOViZ**

PHYLOViZ es un software que permite el análisis de métodos de tipificación basados en secuencias que incluyen el análisis de SNP y el genoma completo (WGS). Es una herramienta Java independiente que permite la deducción filogenética y visualización de conjuntos de datos con enfoques epidemiológicos moleculares (datos referentes a SNP o cg/wgMLST) (**Nascimento et al., 2017**).

- **iTOL**

iTOL es una herramienta empleada en el diseño, visualización, y generación de anotaciones de árboles filogenéticos (**Letunic & Bork, 2021**).

- **Tipificación de secuencias multilocus del genoma central (cgMLST)**

Los esquemas MLST de genoma central se basan en un conjunto fijo de genes que se encuentran conservados en todo el genoma, los cuales suelen ser específicos de la especie (**Kimura, 2018**). Esta clase de esquemas son herramientas que permiten el genotipado, mismas en el área epidemiológica permiten potenciales investigaciones y vigilancia, en adición la clasificación de cepas en las diferentes muestras (**Ragupathy**

et al., 2022). El cgMLST tiene como objetivo relacionar el poder discriminatorio del MLST clásico con la amplia cantidad de datos genéticos derivados de WGS (**Neumann et al., 2019**).

1.2. Objetivos

1.2.1. Objetivo General

Analizar secuencias de genoma completo de *Escherichia coli* aisladas de vegetales y comida callejera de Ecuador.

1.2.2. Objetivos Específicos

- Caracterizar *in silico* secuencias de genoma completo de *E. coli* portadoras de la toxina Shiga aisladas en vegetales de Ecuador.

- Caracterizar *in silico* secuencias de genoma completo de *E. coli* portadoras de betalactamasas de espectro extendido aisladas en vegetales de Ecuador.

- Evaluar la distancia filogenética y las características genómicas entre las secuencias de genoma completo de *E. coli* caracterizadas en este estudio con secuencias representativas de diferente origen geográfico y tipo de muestra.

CAPITULO 2

METODOLOGÍA

2.1.Materiales

Tabla 2

Recursos informáticos

Recursos	Desarrollador
Base de datos	
NCBI	Biblioteca Nacional de Medicina, Departamento de Salud y Servicios Humanos EE. UU, Gobierno Federal de los EE.UU.
Herramientas bioinformáticas:	
Plataforma CGE	Grupo de Investigación para la Epidemiología Genómica del Instituto Nacional de Alimentos, Universidad Técnica de Dinamarca (DTU).
KBase	Departamento de Biología de Sistemas Energéticos de Estados Unidos.
iTOL	Ivica Letunic y Peer Bork (Letunic & Bork, 2007)
BacWGSTdb 2.0	Instituto de Medicina Traslacional, Universidad de Zhejiang, China. Zhi Ruan y Ye Feng (Ruan & Feng, 2016)
Galaxy Sciensano	Laboratorio de Nekrutenko en el Centro de Genómica Comparada y Bioinformática de Penn State, Laboratorio

Recursos	Desarrollador
	Taylor de la Universidad Johns Hopkins y el Laboratorio Goecks de la Universidad de Salud y Ciencias de Oregón.
Microreact	Centro de Vigilancia de Patógenos Genómicas (CGPS).
Proksee	Genome Canadá y Genome Alberta.
Pathogenwatch	Centro de Vigilancia de Patógenos Genómicas (CGPS). Instituto de Grandes Datos, Universidad de Oxford.
Phaster	Institutos Canadiense de Investigación en Salud (CIHR), Genome Alberta, Universidad de Alberta. David Arndt, Jason Grant, Ana Marcu, Tanvir Sajed, Allison Pon, Yongjie Liang, David Wishart.
BV-BRC	Instituto Nacional de Alergias y Enfermedades Infecciosas (NIAID).
Phyloviz	Bruno Goncalves y Joao Andre Carrico
Secuencias de genoma completo de <i>E. coli</i> obtenidas de aislamientos de vegetales y comida callejera de Tungurahua y Chimborazo en los años 2020 – 2021.	

Tabla 3

Equipos

Recursos	Cantidad
Computadora portátil	1
Router para internet	1

2.2.Métodos

La presente investigación planteó realizar un análisis *in silico* de secuencias de genoma completo de *E. coli*. Se utilizaron secuencias de genoma completo de *E. coli* obtenidas de aislamientos de vegetales frescos y comida callejera originarios de Ecuador, portadoras de toxina Shiga (STEC) y resistencia antimicrobiana proveniente de estudios preliminares que no incluyeron el análisis de genoma completo (**Barragán-Fonseca et al., 2022; Tubón et al., 2022**).

Previamente, se identificaron 182 aislados como pertenecientes a la especie *E. coli*. Se seleccionaron aislados que presentaron perfiles de resistencia antimicrobiana fenotípica y genotípica y, portaban los genes de virulencia *stx2*, independientemente. Los aislados fueron recuperados de vegetales frescos provenientes de mercados de la ciudad de Riobamba y la comida preparada fue obtenida en calles de la ciudad de Ambato. Las secuencias crudas de genoma completo fueron obtenidas en base a una colaboración con el Departamento de Agricultura y Mercados del Estado de Nueva York (Estados Unidos), la Universidad McGill de Montreal (Canadá) y la Universidad Central del Ecuador. En resumen, los cultivos de *E. coli* se realizaron en caldo LB e incubaron a 37°C durante 24 horas. La biomasa generada se empleó para extraer ADN genómico utilizando el kit de purificación de ADN genómico Wizard (Promega,

Madison, WI, EE. UU) siguiendo el instructivo del fabricante; el ADN fue cuantificado utilizando el fluorómetro Qubit 3.0 (Invitrogen, CA, EE.UU.).

Las bibliotecas genómicas fueron preparadas utilizando el kit de preparación de bibliotecas Nextera XT (Illumina, CA, EE.UU.), y la secuenciación de todo el genoma (WGS) se realizó empleando longitudes de lectura de extremo emparejado de 300 pb en la plataforma MiSeq (Illumina). Las lecturas en bruto se ensamblaron de *novó* a través de SKESA v2.2 haciendo uso del Pipeline de Anotación del Genoma Procariota NCBI (Souvorov et al., 2018). Las secuencias de genoma completo ensambladas fueron depositadas en la base de datos del Centro Nacional de Información Biotecnológica (NCBI) bajo el BioProject PRJNA415804, los ensamblajes (Assembly) y biomuestras (BioSample) individuales se enumeran en *Anexo 1*. Los ensamblajes fueron descargados de la base de datos GenBank de NCBI en formato FASTA.

2.2.1. Análisis de la calidad del genoma de las secuencias ensambladas

El análisis de la calidad de los datos de WGS son esenciales antes de realizar cualquier análisis bioinformático, para evaluar si estos han alcanzado un estándar adecuado. A partir de las secuencias de genoma completo previamente ensambladas y descargadas, se desarrolló el análisis de la calidad del genoma empleando la herramienta bioinformática BV-BRC (<https://www.bv-brc.org>). Una vez iniciado sesión en BV-BRC, se procedió a subir cada una de las secuencias en la carpeta principal. Posterior a ello, en el menú principal utilizando las opciones Herramientas y Servicios en la categoría Genómica, se seleccionó la opción Análisis del genoma integral. Este metaservicio de análisis integral del genoma permite realizar un análisis profundo de las secuencias que incluyen ensamblajes, anotaciones, análisis comparativos básicos que permiten distinguir el genoma de sus vecinos más cercanos.

Una vez abierto el formulario de entrada de Análisis del genoma integral, se seleccionó los datos de entrada “Contigs ensamblados”; los archivos de entrada fueron seleccionados uno por uno de manera independiente de la carpeta principal “Espacios de trabajo”. El parámetro seleccionado de dominio del organismo objetivo fue Bacterias o Arqueas; por otro parte, el nombre de la taxonomía insertada fue “*Escherichia coli*”. La identificación de la taxonomía se rellenó automáticamente después de introducir el nombre de la taxonomía (ID *Escherichia coli*: 562). Los nombres de salida fueron insertados en base a la especie microbiana e identificación del aislamiento (Ejm. *Escherichia coli*_JC5_2C2). Los datos de salida fueron depositados en una carpeta denominada “Grupos de genoma”.

Los análisis efectuados se visualizaron en el espacio de trabajo privado en la categoría “Mis genomas”. Cada secuencia analizada fue seleccionada una por una y visualizada mediante la opción “Vista del Genoma”. Los resultados provenientes del análisis desarrollado en esta sección son expuestos en el apartado de anexos, **Anexo I**.

2.2.2. Caracterización de secuencias

La caracterización de las secuencias se desarrolló empleando herramientas disponibles en el sitio web del Centro de Epidemiología Genómica (CGE, del inglés Center for Genomic Epidemiology) (<http://www.genomicepidemiology.org/services/>). La caracterización *in silico* fue efectuada a partir de datos de WGS de *E. coli*. Los resultados obtenidos fueron cribados en base a las combinaciones y coincidencias perfectas para un gen determinado, evaluándose la longitud del mejor gen de resistencia de coincidencia en la base de datos y la longitud de alineación entre el mejor gen de interés (gen de resistencia, gen de virulencia, identificación del replicón de plásmido encontrado, entre otros) y la secuencia correspondiente en el genoma.

- **Tipificación de secuencias multilocus (MLST)**

Una vez analizada la calidad de los ensamblajes, se procedió a tipificar los aislamientos empleando la herramienta MLST 2.0. Para lo cual, se seleccionó el servicio de herramientas bioinformáticas para tipificación con el nombre “MLST”. Seguidamente se seleccionó el tipo de microorganismos al cual pertenecen las muestras “*Escherichia coli* #1”. El parámetro subsiguiente de profundidad mínima fue deshabilitado dado que las entradas de lecturas eran contigs ensamblados, siendo este parámetro relevante en entradas de lectura en bruto. Se seleccionó el tipo de entrada de datos FASTA y se cargaron los archivos uno a uno. Los resultados referentes al perfil MLST, organismo al que pertenecen los aislados, tipo de secuencia del genoma de entrada (ST, del inglés *Sequence Type*), se detallan en el apartado resultados (CGE, 2022).

- **Identificación de ARGs y mutaciones puntuales**

Empleando la herramienta ResFinder 4.1 se identificó ARGs completos y mutaciones puntuales. De forma que, se seleccionó el servicio de herramientas bioinformáticas para fenotipado con el nombre “ResFinder”. Los parámetros seleccionados fueron “Mutaciones puntuales y/o genes adquiridos”, en este caso se evaluaron tanto las mutaciones en puntos cromosómicos como los ARGs adquiridos; para cada ítem, se seleccionaron las opciones predeterminadas. Posteriormente se escogió la especie a la cual pertenecen las muestras “*Escherichia coli*” y el tipo de lectura seleccionado fue “Genoma/Contigs ensamblados”. Los archivos se cargaron uno a uno y los resultados obtenidos referente al antimicrobiano evaluado, clases del antibiótico o antimicrobiano, fenotipo predicho de secuenciación de genoma completo (resistencia o no resistencia), trasfondo genético, es decir, los genes o mutaciones puntuales responsables de la resistencia genética (en el caso de haber resistencia), se presentan en el capítulo “Resultados y Discusión” (CGE, 2022).

- **Identificación de los genes de virulencia adquiridos**

La identificación de los genes de virulencia adquiridos se realizó empleando el servicio bioinformático VirulenceFinder 2.0. Inicialmente, se eligió el servicio de herramientas bioinformáticas para fenotipado con el nombre “VirulenceFinder”. Posterior a ello, se seleccionó la base de datos del microorganismo correspondiente al análisis (*Escherichia coli*). Los parámetros referentes al umbral y longitud mínima fueron seleccionados por defecto 90% y 60%, respectivamente. El umbral para %ID, permite identificar el porcentaje mínimo de nucleótidos idénticos entre el gen de virulencia adecuado en la base de datos y la secuencia muestra; por otro parte, la longitud mínima hace referencia al número de nucleótidos que una secuencia debe superponerse al gen de virulencia, representándolo en forma de porcentaje de longitud total del gen de virulencia. El tipo de lecturas seleccionado fue “Genoma/ Contig ensamblado”; y los archivos se cargaron uno a uno, de manera independiente.

Los resultados obtenidos concerniente al factor de virulencia, porcentaje de identidad, longitud del mejor gen de virulencia coincidente en la base de datos y longitud de los mejores segmentos de alta puntuación (HSP), fenotipo previsto, número de adhesión, son presentados en el capítulo “Resultados y Discusión”.

- **Identificación del potencial de patogenicidad**

La predicción del potencial de patogenicidad de los aislamientos se llevó a cabo a través del servicio bioinformático PathogenFinder 1.1. Se seleccionó el servicio de herramientas bioinformáticas para fenotipado con el nombre “PathogenFinder”. El parámetro referente al filo o clase del microorganismo interés fue seleccionado en base al modelo automático; por otra parte, el tipo de entrada seleccionado fue “Genoma/ Contig ensamblado”. Los archivos fueron cargados uno por uno de manera independiente y los resultados referentes a la probabilidad de que los aislamientos sean un patógeno humano son expuestos en el apartado “Resultados y Discusión”.

- **Identificación de plásmidos**

Se empleó la herramienta PlasmidFinder 2.1 para identificar la presencia de plásmidos en los aislados de secuencias completas de *E. coli*. Para ejecutar la predicción de plásmidos utilizando el servidor web PlasmidFinder, se procedió a seleccionar la base de datos “*Enterobacteriaceae*”. El umbral para el porcentaje mínimo de identidad predeterminado fue 95%; por otro lado, el umbral de cobertura mínimo predeterminado fue 60%. El tipo de lectura seleccionado fue “Genoma/ Contigs ensamblado o borrador (fasta)”. Los archivos fueron cargados uno por uno de manera independiente. Los resultados de identificación del replicón de plásmido encontrado con mayor identidad se muestran en el apartado “Resultados y Discusión”.

2.2.3. Determinación de profagos y elaboración de gráficos circulares

El servidor web PHASTER (<http://phaster.ca>) se empleó para identificar regiones de bacteriófagos y profagos (estado del genoma de bacteriófagos integrados en el genoma huésped). Los archivos de secuencias de nucleótidos fueron cargados de manera

individual en la plataforma PHASTER. Se seleccionó el ítem “Mi entrada consiste en varios contigs separados (solo formato FASTA)””; por otra parte, el ítem consecuente se dejó como predeterminado. Los datos fueron enviados y una vez completada la carga del archivo, se identificó las regiones de profagos intactas (puntuación >90, este parámetro es calculado mediante 3 métodos para puntuar las regiones del fago como intactas, cuestionables o incompletas). Las regiones de ADN de fagos intactas fueron recopiladas en archivos de texto con extensión “. fasta” y el contenido de G+C se presenta en la sección “Resultados y Discusión”.

Se desarrollaron gráficas circulares de las secuencias de genoma completo de forma individual, empleando la herramienta bioinformática web Proksee (<https://proksee.ca>), en el cual se insertó las regiones de profagos intactas e identificaron los ARGs y virulencia previamente analizados. Las WGS previamente descargadas fueron cargadas de manera independiente través del Browser. Una vez creado el mapa circular, se procedió a insertar el archivo “. fasta” de las regiones de profagos intactas para cada secuencia empleando la herramienta BLAST localizada en la categoría “Comparación de Secuencias”. En la sección de Trabajo, se seleccionó el archivo que contenía las regiones de profagos intactas cargado, y se añadió los resultados en el mapa. Los ARGs localizados con anterioridad por ResFinder fueron insertados empleando la herramienta “CARD Resistance Gene Identifier” localizada en la categoría “Anotación del Genoma”; los genes de resistencia insertados sin previo análisis fueron eliminados del gráfico circular.

2.2.4. Comparación de secuencias de genoma completo de diversos orígenes geográficos y elaboración de árboles filogenéticos

Para la elaboración de los árboles filogenéticos, se tomaron 50 WGS de aislamientos de *E. coli* procedentes de diferentes países del mundo; con la finalidad de: identificar las diferencias existentes entre los genomas de *E. coli* disponibles en las bases de datos y evaluar la relación existente con aislamientos de origen ecuatoriano. Las secuencias

empleadas fueron secuencias referenciadas en artículos y revistas científicas en diferentes años y secuencias aisladas de diferentes matrices (Humana, Animal, Alimentaria y Ambiental).

De igual forma, para la selección de los aislamientos de orígenes diversos, se empleó el concepto de referencia expuesto en el estudio de (Nagy et al., 2020), el cual detalla tres apartados:

- i. Los aislamientos de un solo país se preferirán si provienen de una misma publicación.
- ii. Cada país estará representado por un máximo de 10 secuencias de diverso origen (alimentario, ambiental o clínico).
- iii. Se seleccionarán las secuencias provenientes de artículos publicados con una preferencia de los 5 últimos años.

Las secuencias seleccionadas fueron descargadas empleando la base de datos NCBI, de forma que, la descarga de las WGS ensambladas provenientes de diversos orígenes fue desarrollada a través de la base de datos GenBank de NCBI en formato FASTA y guardado con el número de acceso ID Strain_localización_año (Ejem: RICC00000000.1_C1-067_USA_2006). La información referente al país de origen, matriz, tipo de muestra, fecha de aislamiento, código de acceso, y referencia son expuestos en el *Anexo 3*.

WGS descargadas de diversos orígenes fueron caracterizadas siguiendo el mismo protocolo de análisis de WGS mencionado con anterioridad; de esta forma, se utilizaron las herramientas de CGE para identificar los ARGs, genes de virulencia, plásmidos, perfiles de MLST, subtipificación de las cepas, MGE y regiones de profago (PHASTER). De igual forma, se empleó la plataforma BV-BRC para conocer la calidad del genoma ensamblado, el tamaño total del genoma, número de contigs, N_{50} y contenido G+C.

Los árboles filogenéticos se realizaron empleando las secuencias de diferentes ubicaciones geográficas previamente caracterizadas y las secuencias descritas en la presente investigación. Se empleó la tipificación de la secuencia multilocus del genoma central (cgMLST), la cual es una extensión del esquema MLST convencional de siete locus que amplía el rango de genes diana al nivel del genoma completo y, a menudo, se usa como una solución para proporcionar una relación filogenética muy detallada (Feng et al., 2021).

Para realizar el análisis de tipificación cgMLST se empleó la herramienta bioinformática Galaxy Sciensano. Se ingresaron los archivos FASTA previamente extraídos de NCBI a la plataforma Galaxy Sciensano. Posterior a la construcción de una lista de conjunto de datos, se empleó la herramienta “Multi Locus Sequence Typing (MLST-BLAST) para analizar las cgMLST individualmente. Seguidamente se utilizó la herramienta “MLST-Phylogeny” de la cual se extrajo la matriz alélica para la generación de los árboles de filogenia basado en la tipificación cgMLST. La construcción del árbol de filogenia (Árbol de Expansión Mínima) se realizó a partir de los datos obtenidos en Galaxy Sciensano a través de la herramienta bioinformática Phyloviz.

La elaboración de los árboles de filogenia rectangular se realizó empleando iTOL. Las anotaciones referentes a las secuencias de genoma completo, origen de aislamiento, ARGs, virulencia y tipo de plásmidos fueron insertadas a través del editor de iTOL. Los análisis complementarios fueron realizados empleando las plataformas Microreact y Pathogenwatch. Los resultados son presentados en el capítulo “Resultados y Discusión”.

CAPÍTULO 3

RESULTADOS Y DISCUSIÓN

3.1. Análisis y discusión de los resultados

3.1.1. Análisis de secuencias a través de plataformas bioinformáticas

La información recopilada mediante las diversas bases de datos y plataformas bioinformáticas NCBI, Enterobase, EMBL-EBI, KBase, BV-BRC, CGE, BacWGSTdb, iTol, Microreact, Pathogenwatch y Phyloviz, empleados en la caracterización de aislados de *E. coli*, se presentan a continuación.

3.1.1.1. Análisis de secuencias de genoma completo

En la presente investigación se analizaron 5 secuencias de genoma completo de *E. coli* aisladas en estudios anteriores de vegetales y comida callejera de Ecuador, de las cuales dos aislamientos son portadores de genes productores de la toxina Shiga, y tres aislamientos codifican ARGs betalactámicos. Adicionalmente, se analizaron 8 aislamientos de *E. coli* los cuales no portan los mecanismos de resistencia y de virulencia caracterizados en los objetivos de la presente investigación, sin embargo, se añadió su análisis preliminar en la sección de anexos. La calidad del genoma se encuentra predicha en base a datos estadísticos en el cual se incluye el número total de lecturas, contenido de guanina-citosina (%GC), número de contigs, contigs N₅₀ y L₅₀, puntuación de integridad y contaminación CheckM del genoma (Parks et al., 2015; Parrello et al., 2019). En el *Anexo 1 y 2*, se detalla la información general

correspondiente a las secuencias de genoma completo analizadas en el presente estudio.

En base a la información general para las secuencias de genoma completo, se identificó que el número total de lecturas varió de 4.0 Mb a 5.3 Mb, expresando rendimientos idóneos con referencia a las cantidades secuenciadas. El genoma de *E. coli* suele bordear los 5.0 Mb, no obstante, los genomas de *E. coli* dependiendo la cepa aislada pueden variar aproximadamente 1 Mb, oscilando entre 4.5 y 5.5 Mb (**Hozzari et al., 2020**).

Por otra parte, el porcentaje total de contenido de GC es empleado como parámetro de control de calidad. El contenido genómico de GC de las bacterias varía en relación a las diferencias en el proceso mutacional o procesos neutrales y selectivos, estas variaciones van desde menos del 25% hasta más del 75% (**Hershberg, 2016; Raghavan et al., 2012**). Bajo literatura, el contenido de GC para la secuenciación del genoma completo es alrededor del 49-51% (**Guo et al., 2014**); en este contexto el porcentaje promedio para las secuencias analizadas fue de 50.7% G+C.

En referencia al número de contigs, estos oscilaron entre los 54 hasta 171; por lo general el número total de contigs ensamblados <100 contigs representa un valor realista en referencia a la calidad del genoma. La longitud media de todas las longitudes de contigs N_{50} se encontraron en el rango de 74,583 pb a 276,038 pb; en este sentido, medidas de $N_{50} > 15,000$ pb son indicadores de buena calidad, no obstante, se prefiere un tamaño mínimo de 30,000 pb (**Ellington et al., 2017**).

Otras de las métricas que terminaron la integridad del genoma fueron las métricas de integridad y contaminación del genoma, las cuales evidenciaron el 100% de integridad CheckM y el porcentaje de contaminación CheckM el cual osciló entre 0.1% al 1.5%,

indicando que la calidad del genoma para las secuencias evaluadas es considerada buena.

3.1.1.2. Caracterización de secuencias de genoma completo (WGS) de *E. coli* portadores de la toxina Shiga.

El enfoque principal de esta sección se delimitó en el análisis de dos secuencias de genoma completo de aislamientos de *E. coli* de origen vegetal portadores de la toxina Shiga (aislamientos C5b1 y T5bX). Como parte de la caracterización *in silico*, se determinó los tipos de plásmidos que poseen, ARGs, análisis de la presencia de profagos, serovar, genes de virulencia complementarios, y tipo de secuencia MLST.

En ambos aislamientos, se observó la presencia de genes productores de enterotoxinas como *stx* y *astA*. Los dos aislamientos fueron *stx2* positivos, es decir portaban genes de la toxina Shiga del subtipo A y B (*stx2A* y *stx2B*), la información general de la caracterización *in silico* se describe en la **Tabla 4**. A la par se identificó el tipo de secuencia MLST, tipo de plásmido y regiones de profagos intactas para estas dos cepas de *E. coli* C5b1 y T5bX, esta información se detalla en la **Tabla 5**. El tipo de secuencia (ST) identificada para estas dos cepas fue ST677, el tipo de plásmido detectado fue *IncFII* y se encontraron 5 y 6 regiones de profagos intactas.

Las evaluaciones fenotípicas realizadas en etapas previas por **Barragán & Calero, (2022)** para las cepas de *E. coli* C5b1 y T5bX, indicaron la presencia de patotipos diarreogénicos, puesto que, posterior a la etapa de PCR se identificó la amplificación del gen de virulencia *stx2*. Estos resultados fueron corroborados en base al análisis de WGS realizado en el presente estudio.

Tabla 4

Identificación de las secuenciaciones del genoma completo de 2 cepas de E. coli aisladas de vegetales y comida callejera de Ecuador portadoras de genes toxina Shiga y ARGs

Aislamiento	C5b1	T5bX
Año de aislamiento	2021	2021
Tipo de muestra	Col Blanca (<i>Brassica oleracea</i>)	Tomate (<i>Solanum lycopersicum</i>)
Ubicación geográfica	Chimborazo	Chimborazo
Número de lecturas totales	5,184,284	5,181,902
Porcentaje de G+C (%)	50.71	50.71
N50(bp)	217,893	229,587
L50(bp)	8	7
No. de contigs	87	85
No. de CDSs	5,242	5,229
No. de tRNA	82	82
No. de rRNA	10	8
Compleitud (CheckM)	100	100
Contaminación (CheckM)	0.1	0.1
No. Acceso a BioSample	SAMN28681467	SAMN28681464
No. Acceso a SRA	SRS13189686	SRS13189707
No. Acceso a Assembly	GCA_023561385.1	GCA_023561575.1

Tabla 5

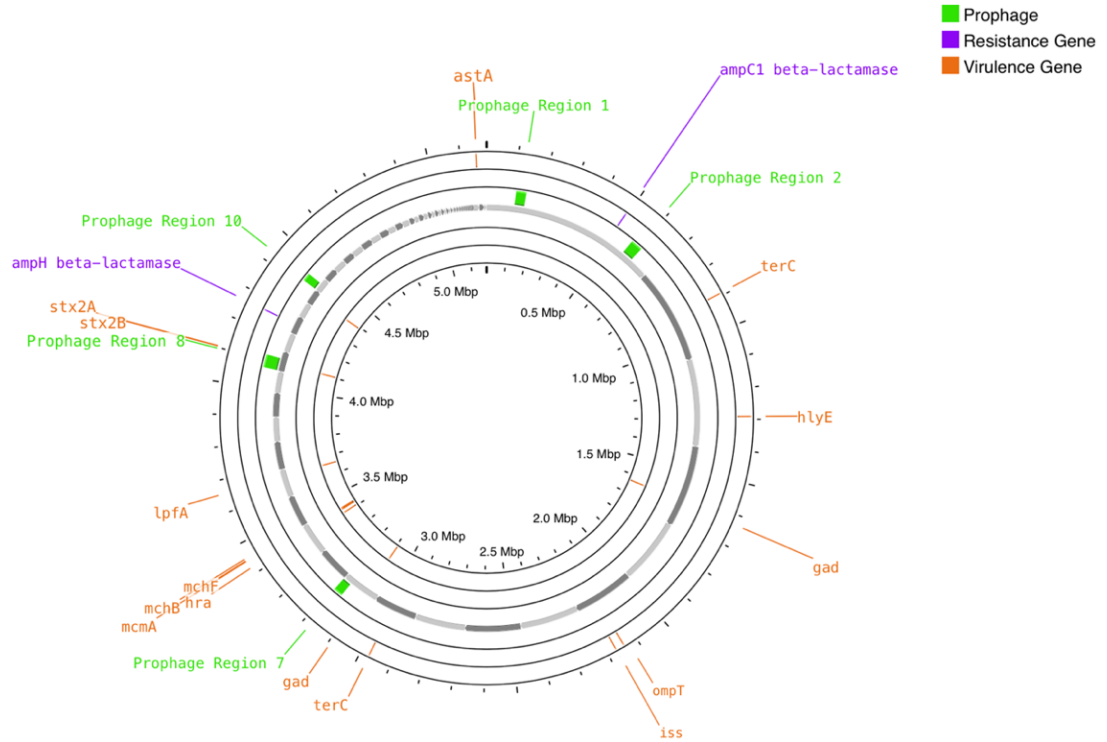
Caracterización basada en la WGS de 2 cepas de *E. coli* aisladas de vegetales y comida callejera de Ecuador portadoras de genes toxina Shiga y ARGs

Cepa	MLST	Tipo de Plásmido	Profagos (Regiones Intactas)	Genes de Virulencia	Genes de Resistencia
C5b1	ST677	IncFII(pHN7A8)	5	<i>astA</i> ; <i>gad</i> ; <i>hlyE</i> ; <i>hra</i> ; <i>iss</i> ; <i>lpfA</i> ; <i>mcmA</i> ; <i>mchB</i> ; <i>mchF</i> ; <i>ompT</i> ; <i>stx2A</i> ; <i>stx2B</i> ; <i>terC</i> ; <i>traT</i>	<i>ampC₁</i> <i>ampH</i>
T5bX	ST677	IncFII(pHN7A8)	6	<i>astA</i> ; <i>gad</i> ; <i>hlyE</i> ; <i>hra</i> ; <i>iss</i> ; <i>lpfA</i> ; <i>mcmA</i> ; <i>mchB</i> ; <i>mchF</i> ; <i>ompT</i> ; <i>stx2A</i> ; <i>stx2B</i> ; <i>terC</i> ; <i>traT</i>	<i>ampC₁</i> <i>ampH</i>

Los factores de virulencia como *astA* encargado de codificar la enterotoxina estable térmica, *gad* (glutamato descarboxilasa), *hra*, gen que codifica la aglutinina resistente al calor, *lpfA* (fimbriae polar larga) factor adhesivo de STEC, y *ompT*, factor asociado a la patogenicidad extraintestinal pues codifica la proteasa de la membrana externa de *E. coli*, se encontraron presentes en los aislamientos portadores de genes toxina Shiga, como se observa en la **Figura 4 y 5** (Bidet et al., 2022; da Silva et al., 2022; Hammad et al., (2020)., Zhang et al., 2021).

Figura 4

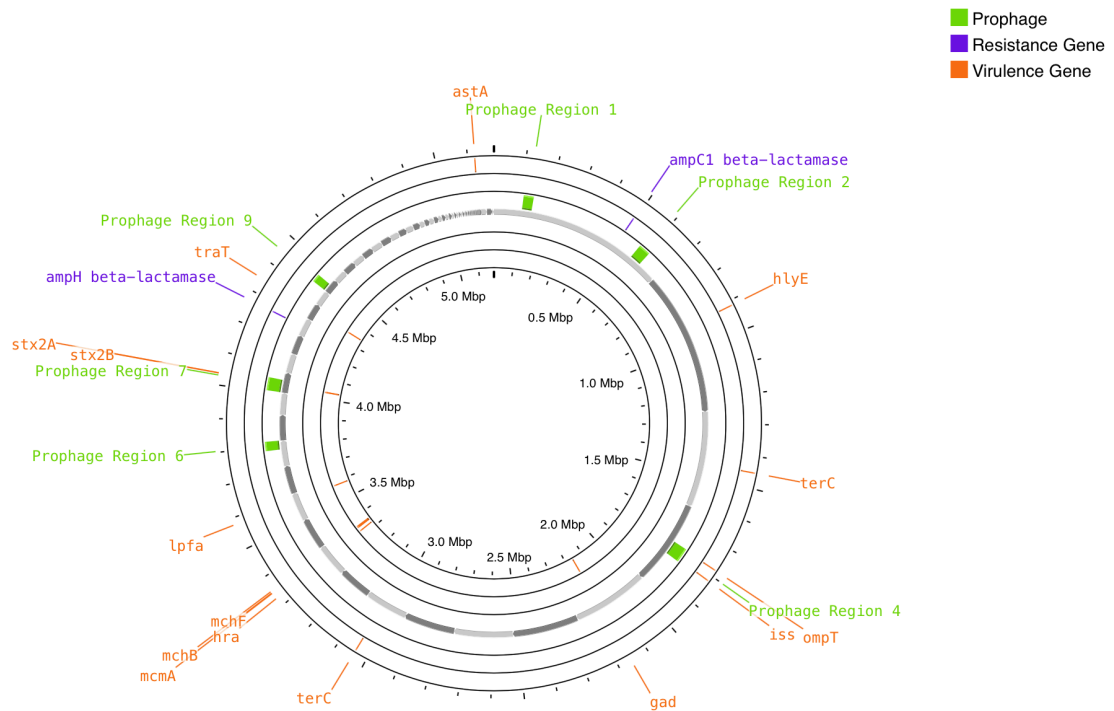
Mapa circular de la secuencia genómica completa de la cepa de *E. coli* C5b1 aislada de col blanca (*B. oleracea*)



Nota: Los genes de resistencia, virulencia y presencia de regiones intactas de profagos están codificados por colores de la siguiente forma: verde (regiones intactas de profagos), morado (ARGs), y anaranjado (genes de virulencia)

Figura 5

Mapa circular de la secuencia genómica completa de la cepa de *E. coli* T5bX aislada de tomate (*S. lycopersicum*)



Nota: Los genes de resistencia, virulencia y presencia de regiones intactas de profagos están codificados por colores de la siguiente forma: verde (regiones intactas de profagos), morado (ARGs), y anaranjado (genes de virulencia)

Los aislados de *E. coli* positivos para *stx2* curiosamente fueron de origen vegetal. En este contexto, a lo largo de los años se ha identificado que el principal reservorio de STEC son animales (**European Centre for Disease Prevention and Control, 2022**). No obstante, estudios epidemiológicos actuales atribuyen el incremento de hospitalizaciones al consumo de verduras frescas contaminadas con STEC, infiriendo un mayor alcance en la diseminación de STEC al medio ambiente (**Leonard et al., 2016**). En el caso de los aislamientos C5b1 y T5bX, ambos pertenecen al serovar O174:H21, el cual no ha sido previamente reportado en Ecuador, pero ha sido detectado en aislamientos STEC origen humano, animal y alimentario en Chile y Argentina, en donde se observa una diseminación emergente de este serovar en los últimos años (**Galarce et al., 2021; Torres et al., 2018**). El uso de estiércol animal sin tratar en los campos de cultivo, así como el uso de agua de riego contaminada, favorece

la diseminación de microorganismos potencialmente patógenos hacia el suelo y los alimentos (**Pushpakanth et al., 2019; Ramos et al., 2021**). La detección de STEC en alimentos vegetales cuyo consumo no implica un tratamiento térmico representa una alerta adicional, considerando que puede facilitarse su adquisición por los consumidores y provocar brotes de infecciones transmitidas por los alimentos (**Waltenburg et al., 2022**).

Las toxinas Shiga (Stx) en las STEC generalmente se codifican en el genoma de los bacteriófagos lambdoides, que pasan la mayor parte del tiempo de su ciclo de vida integrados como profagos en sitios específicos del cromosoma bacteriano (**Rodríguez-Rubio et al., 2021**). En la **Figura 4 y 5**, se puede identificar que el gen *stx2* subtipo A y B se encontraban ubicados en regiones identificadas como intactas de profagos de los aislamientos C5b1 y T5bX (Regiones de Profagos 8 y 7, respectivamente). Según **Koutsoumanis et al., (2020)** los bacteriófagos convertidores de toxinas Shiga tienen la capacidad de movilizar genes *stx* y lisogenizar cepas bacterianas no patógenas convirtiéndolas en STEC. La inducción de los bacteriófagos Stx lisogénicos puede darse a partir de cultivos de la célula hospedadora mediante diferentes agentes inductores como factores abióticos (uso de antibióticos como mitomicina C, factores de huésped, luz ultravioleta) (**Grande et al., 2014; Rodríguez-Rubio & Muniesa, 2021**). Por lo tanto, una continuación al presente estudio sería la inducción del ciclo lítico de los profagos presentes en C5b1 y T5bX, y su posterior caracterización con la finalidad de conocer si siguen una vía lisogénica o lítica, y si son capaces de transferir los genes Stx a una célula receptora por medio de transconjugación.

IncFII (pHN7A8) es un tipo de plásmido identificado en la presente investigación en las cepas positivas para *stx2*, evidenciándose este análisis *in silico* en la **Tabla 5**. Aunque en el presente estudio no se ha identificado la presencia de ARGs y factores de virulencia en plásmidos en los STEC evaluados, es importante recalcar que, los plásmidos de tipo *IncFII* (pHN7A8) regularmente están involucrados en la difusión de los genes *bla_{CTX-M}* que codifican las cefotaximasas, las cuales son enzimas encargadas de conferir la resistencia a cefalosporinas (**Eskandari-Nasab et al., 2018; He et al.,**

2013; Zhou et al., 2021). Lastimosamente, la secuenciación de fragmentos cortos de Illumina no permite ensamblar diferencialmente a plásmidos y al cromosoma, motivo por el cual una investigación complementaria a la presente podría enfocarse en WGS por medio de fragmentos largos a través de la tecnología Nanopore o PacBio, y posteriormente ejecutar un ensamblaje híbrido (lecturas cortas y largas) para obtener una secuencia consenso de alta calidad que diferencie a plásmidos y cromosomas (**De Maio et al., 2019**).

La tipificación MLST reveló que los dos aislamientos portadores de genes *stx2* compartían un mismo ST (ST677). Estudios de vigilancia han identificado con frecuencia que las cepas de *E. coli* ST677 como cepas STEC de serovar O174:H21 que circulan en matrices alimentarias, animales, ambientales y humanas distribuidas con frecuencia en América, Europa, Asia y África (**EnteroBase, 2023; Galarce et al., 2021**). En el caso de aislamientos de origen alimentario, la diseminación de *E. coli* ST677 se evidencia en productos alimenticios (sin especificar el tipo de alimento, ni el origen) por parte de la vigilancia epidemiológica de la Agencia Canadiense de Inspección de Alimentos (**EnteroBase, 2023**). Adicionalmente, estudios han revelado que las cepas STEC de ST677 con frecuencia portan múltiples ARGs tal es el caso del gen *ampC* (**Mahindroo & Taneja, 2021**). En este contexto las cepas de *E. coli* C5b1 y T5bX analizadas, presentaron también genes de resistencia a β -lactámicos, tal es el caso de los genes *ampC₁* y *ampH*. Los genes *ampC₁* y *ampH* son encargados de codificar β -lactamasas de clase C (cefalosporinasas AmpC) (**Huang et al., 2021**). Estos reportes sugieren que *E. coli* ST677 se encuentra ampliamente diseminada a nivel de matriz de aislamiento y distribución geográfica, motivo por el cual se justifica la necesidad de realizar monitoreo y vigilancia genómica, lo cual permitiría prevenir brotes de ETAs ligados a STEC resistentes a antimicrobianos.

En la caracterización *in silico* efectuada mediante ResFinder, se detectaron “mutaciones no definidas” en los genes *ampC*, *parC* y *pmrB* de los aislamientos *E. coli* C5b1 y T5bX. En estos dos aislamientos se identificaron dos mutaciones del gen *ampC*, una mutación en el gen *parC* y dos mutaciones en el gen *pmrB*. En este

contexto, las mutaciones no definidas son mecanismos de defensa que algunas bacterias han desarrollado ante la presión selectiva generada por los antimicrobianos (Hager et al., 2017). Los mecanismos de resistencia mutacional incluyen, modificaciones en los centros de unión bacteriano a antibióticos, sobreexpresión de proteínas o enzimas que inhiben la actividad antimicrobiana, mutaciones que dan como resultado la sobreexpresión de numerosas bombas de eflujo intrínsecas (Cho & Misra, 2021; Coolen et al., 2021; Khalifa et al., 2021). En este sentido, se ha mencionado que las mutaciones significativas de aminoácidos en el gen *parC* (codifica la topoisomerasa IV) desencadena mecanismos de resistencia a las fluoroquinolonas (Johnning et al., 2015). Un estudio de 58 aislamientos clínicos de *E. coli* resistentes a fluoroquinolonas realizado en hospitales de EE. UU, encontraron que el 85% de los aislamientos presentaban sustituciones de aminoácidos en *parC* (Morgan-Linnell et al., 2009). Por otra parte, se ha descrito que las mutaciones cromosómicas puntuales en el gen *pmrB* (codifica la proteína cinasa sensora PmrB) desencadenan mecanismos de resistencia a colistina y polimixina B en *E. coli* (Cannatelli et al., 2017; Phan et al., 2017).

Con respecto a la resistencia fenotípica, en una caracterización preliminar se identificó que los aislamientos *E. coli* C5b1 y T5bX poseen resistencia a gentamicina (CN), ampicilina (AM), cefotaxima (CTX), cefalotina (KF) (Barragán-Fonseca et al., 2022). La identificación fenotípica de la resistencia a aminoglucósidos, cefalosporinas y penicilinas viene determinada por la presencia del gen cromosómico *ampC* (Kohlmann & Gatermann, 2020; Singh et al., 2020). Según Tamma et al., (2019), las enzimas AmpC son inducibles y pueden sobreexpresarse en altos niveles por mutaciones; esta sobreexpresión confiere a la bacteria resistencia a cefalosporinas de amplio espectro tales como cefotaxima, ceftazidima y ceftriaxona. Los análisis *in silico* realizados a partir de la WGS representan un enfoque potencial para proporcionar datos completos referentes a genes de resistencia o mutaciones relacionadas, que podrían ser usados para comparar las evaluaciones fenotípicas tradicionales con resistencias antimicrobianas inferidas genotípicamente.

3.1.1.3. Caracterización genotípica de secuencias de genoma completo de *E. coli* portadoras de betalactamasas de espectro extendido.

Posterior al análisis de secuencias de genoma completo para la identificación de ARGs en las secuencias evaluadas, se detectó que 3 secuencias portaban genes de resistencia a β -lactámicos de espectro extendido (BLEE). Las cepas de *E. coli* portadores de BLEEs fueron aisladas de comida callejera (ensalada y jugo de caña) y vegetales (lechuga), como se presenta en la **Tabla 6**.

El análisis de WGS confirmó la presencia de los genes *bla*_{TEM-1B} y *bla*_{CMY-132} en tres cepas de *E. coli* (N1_1, Jc5_2C2 y L6by). Los genes que portan BLEEs comúnmente detectados fue *bla*_{TEM-1B}, dado que fue identificado en los tres aislamientos; por otra parte, el gen *bla*_{CMY-132} fue detectado en el aislamiento de *E. coli* Jc5_2C2. Adicionalmente, se detectaron otros 16 ARGs adquiridos relacionados a la resistencia a diversas clases de antibióticos. Los 3 aislamientos presentaban al menos un gen de resistencia adicional, y la mayoría presentó por lo mínimo 8 ARGs. En adición, todas las cepas portaban las variantes de los genes *tet* (resistencia a tetraciclina y doxiciclina) y *sul* (resistencia al sulfametoxazol).

Tabla 6

Identificación de las secuenciaciones del genoma completo de 3 cepas de E. coli aisladas de vegetales y comida callejera de Ecuador portadoras de ARGs y genes de virulencia

Cepa	N1_1	Jc5_2C2	L6bY
Año de aislamiento	2020	2021	2021
Tipo de muestra	Ensalada	Jugo de caña	Lechuga (<i>Lactuca sativa</i>)
Ubicación geográfica	Tungurahua	Tungurahua	Chimborazo

Cepa	N1_1	Jc5_2C2	L6bY
Número de lecturas	5,058,102	5,125,220	4,007,122
totales			
Porcentaje de G+C	50.54	50.53	50.68
(%)			
N50(bp)	119,031	276,038	210,300
L50(bp)	15	7	6
No. de contigs	115	61	56
No. de CDSs	5,024	5,020	4,607
No. de tRNA	80	82	78
No. de rRNA	7	5	6
Completitud	100	100	100
(CheckM)			
Contaminación	1.5	0.1	0.1
(CheckM)			
No. Acceso a BioSample	SAMN28681473	SAMN28681469	SAMN28681465
No. Acceso a SRA	SRS13189681	SRS13189654	SRS13189668
No. Acceso a Assembly	GCA_023561435.1	GCA_023561735.1	GCA_023561375.1

Algunos de los ARGs que codifican enzimas modificadoras de aminoglucósidos localizados en los aislamientos fueron *aac (6')-Ib-cr*, *ant(3'')-IIa*; *aph(6)-Id*; *aph (3'')-Ib*, *ant (3'')-Ib* y *aadA2* (Eftekhar & Seyedpour, 2015; Zhang et al., 2017). Dos de los aislamientos portaban el gen *dfrA14* (codifica una dihidrofolato reductasa resistente a trimetoprima). Uno de los aislamientos portaba los genes *mphA* (confiere resistencia a macrólidos) y *qnrB19* (otorga resistencia a las quinolonas) (Alcock et al., 2023; Moreno-Switt et al., 2019).

Adicional a la identificación de ARGs, se realizó el análisis de MLST, determinando que las cepas portadoras de BLEEs (N1_1, Jc5_2C2 y L6by) pertenecían a una amplia

selección de tipo de secuencias, siendo estos ST101, ST12373 y ST1286, respectivamente. Por otra parte, el buscador de plásmidos (PlasmidFinder) predijo varios replicones de plásmidos en todas las cepas portadoras de BLEEs. Los tipos de plásmidos localizados fueron IncI1-I (Alpha), IncFIC (FII); IncFIB (AP001918); IncFIA, IncFII (pHN7A8) y Col (pHAD28). Adicional, se identificaron las regiones de profagos intactas, determinando que las cepas de *E. coli* N1_1, Jc5_2C2 y L6by presentaban una región de profago intacta en la cual no se localizaron ARGs y virulencia; la información detallada se describe en la **Tabla 7**.

Tabla 7

Caracterización basada en la secuenciación del genoma completo de 3 cepas de E. coli aisladas de vegetales y comida callejera de Ecuador portadoras de ARGs y genes de virulencia

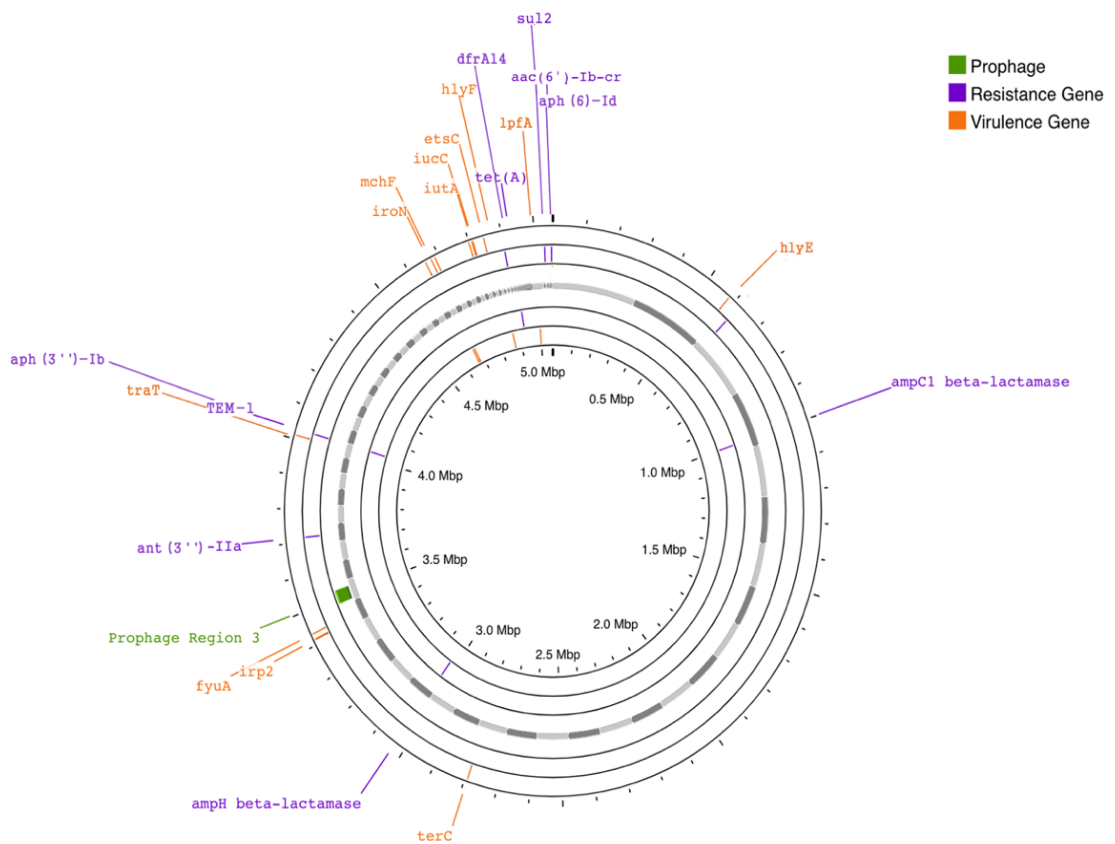
Cepa	MLST	Tipo de Plásmido	Profagos (Regiones Intactas)	Genes de Resistencia	Genes de Virulencia
N1_1	ST101	IncI1-I(Alpha); IncFIC(FII); IncFIB(AP001918); IncFIA	1	<i>aac</i> (6')- <i>Ib-cr</i> ; <i>ant</i> (3'')- <i>Ila</i> ; <i>aph</i> (3'')- <i>Ib</i> ; <i>ampH</i> ; <i>bla</i> _{TEM-1B} ; <i>sul2</i> ; <i>tetA</i>	<i>etsC</i> ; <i>fyuA</i> ; <i>hlyE</i> ; <i>hlyF</i> ; <i>iucC</i> ; <i>iroN</i> ; <i>irp2</i> ; <i>iutA</i> ; <i>lpfA</i> ; <i>mchF</i> ; <i>terC</i> ; <i>traT</i>
Jc5_2c2	ST12373	IncFIA; IncFIB(AP001918); IncFIC(FII); IncFII(pHN7A8)	1	<i>aph</i> (6)- <i>Id</i> ; <i>aph</i> ((3'')- <i>Ib</i>); <i>aadA2</i> ; <i>ampH</i> ; <i>bla</i> _{CMY-132} ; <i>bla</i> _{TEM-1B} ; <i>sul2</i> ; <i>tetB</i>	<i>astA</i> ; <i>chuA</i> ; <i>eilA</i> ; <i>gad</i> ; <i>hlyE</i> ; <i>kpsE</i> ; <i>sitA</i> ; <i>terC</i> ; <i>traT</i>
L6by	ST1286	Col(pHAD28)	1	<i>aadA2</i> ; <i>ant</i> (3'')- <i>Ib</i> ; <i>ampH</i> ; <i>dfrA12</i> ; <i>qnrB19</i> ; <i>sul1</i> ; <i>tetA</i>	<i>terC</i>

En las **Figuras 6, 7 y 8** se presentan los mapas circulares de las secuencias de genoma completo de las cepas de *E. coli* que portaban BLEEs. Se ha identificado que los aislamientos de *E. coli* obtenidos a partir de comida callejera (ensalada y jugo de caña) adicional a portar varios ARGs, portaban una gran variedad de genes de virulencia

como *astA*, *chuA*, *eilA*, *etsC*, *fyuA*, *gad*, *hlyE*, *hlyF*, *iucC*, *ironN*, *irp2*, *iutA*, *kpsE*, *lpfA*, *mchF*, *sitA*, *terC*, *traT*. Por otra parte, el aislamiento de *E. coli* obtenido a partir de lechuga únicamente portaba el gen de virulencia *terC*.

Figura 6

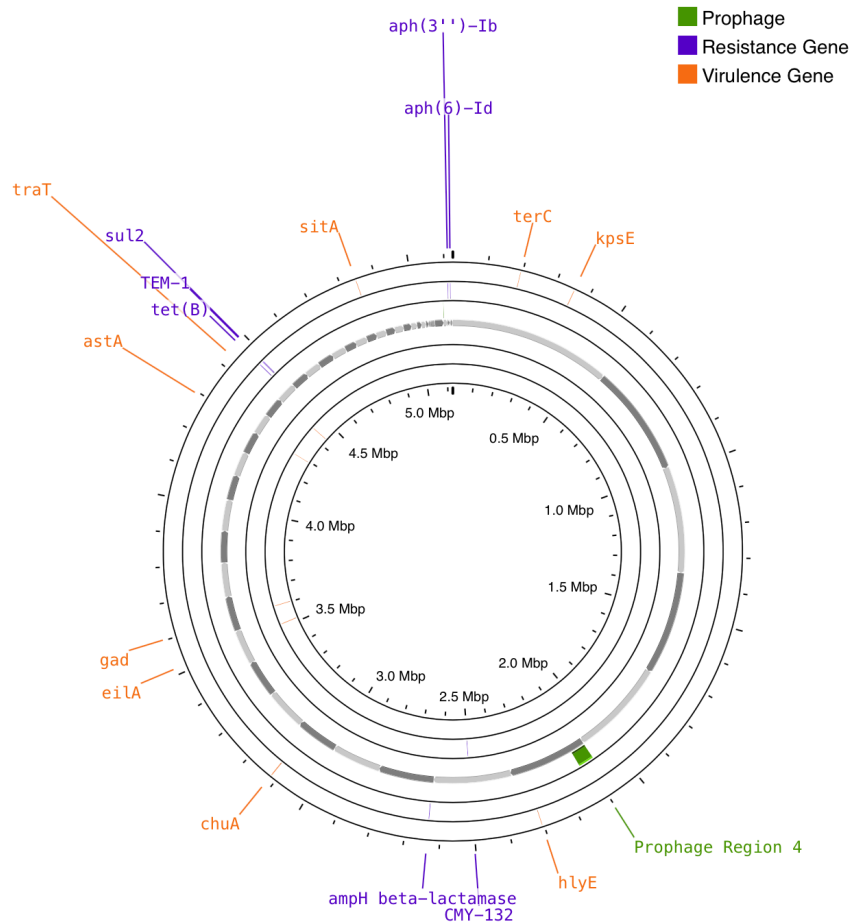
Mapa circular de la secuencia genómica completa de la cepa de *E. coli* N1-1 aislada de comida callejera (Ensalada).



Nota: Los genes de resistencia, virulencia y presencia de regiones intactas de profagos están codificados por colores de la siguiente forma: verde (regiones intactas de profagos), morado (ARGs), y anaranjado (genes de virulencia)

Figura 7

Mapa circular de la secuencia genómica completa de la cepa de *E. coli* J5_2C2 aislada de juego de caña.



Nota: Los genes de resistencia, virulencia y presencia de regiones intactas de profagos están codificados por colores de la siguiente forma: verde (regiones intactas de profagos), morado (ARGs), y anaranjado (genes de virulencia)

Los alimentos callejeros listos para el consumo humano son potenciales vehículos de transmisión de microorganismos patógenos, debido a que las limitadas prácticas de seguridad alimentaria ejecutadas por los vendedores de alimentos callejeros incrementan el número de enfermedades bacterianas transmitidas por los alimentos (Tenea et al., 2023). Por su parte, la contaminación de los alimentos con microorganismos resistentes a antimicrobianos representa un problema de salud en

aumento en todo el mundo. Bajo este contexto, *E. coli* resistente a antimicrobianos se ha convertido en un patógeno peligroso implicado en epidemias de gastroenteritis en América, Europa, Asia y África localizado con frecuencia en alimentos poco cocidos, carne contaminada, jugos de frutas y leche sin pasteurizar, y verduras frescas (**Giri et al., 2021; Pilamala et al., 2023**). En consecuencia, la OMS recomienda la vigilancia de las bacterias entéricas productoras de BLEEs (WHO, 2014). Por consiguiente, el control y vigilancia de la seguridad de los alimentos vendidos en la calle es un desafío para muchos países. En Ecuador, a pesar de contar con un Reglamento para el Control Sanitario de Alimentos que se expenden en vías públicas (Acuerdo No. 14381), se ha identificado que la comida callejera lista para comer no está regularizada y es una tarea compleja para los vendedores ambulantes el control de los puntos de riesgo de contaminación (**Reglamento de Control Sanitario de Alimentos, 1992; Zurita et al., 2020**). De esta forma, la detección de *E. coli* portadora de BLEEs es indispensable para comprender la diversidad y las rutas de diseminación de la resistencia a los antimicrobianos en matrices alimentarias, en especial alimentos de alto consumo local.

Los aislamientos de comida callejera empleados en el presente análisis de WGS se derivaron de muestras que contenían integrados una serie de ingredientes, tal es el caso del aislado de ensalada, este contenía ingredientes como mote, choclo, habas, ensalada fresca y cerdo; por lo cual no se puede vincular la contaminación de dichos alimentos a un solo ingrediente principal. Por su parte, se ha descrito que los productos alimenticios derivados de frutas frescas son una fuente de preocupación constante, puesto que estos productos frescos son con frecuencia más sensibles a la contaminación, dados sus mecanismos de cultivo y recolección inadecuados, suministros de agua de regadío contaminada, limitada limpieza antes de la distribución, y almacenamiento inapropiado de frutas (**Tenea et al., 2023**). Según **Azanaw et al., (2022)** las principales fuentes que aportan a la contaminación microbiana en alimentos son temperaturas de mantenimiento inadecuadas, manipulación excesiva de alimentos por parte de los vendedores, utensilios y equipos sucios para cocinar y servir, materias primas contaminadas, almacenamiento de materia prima inadecuado y falta de higiene personal de los vendedores.

Se ha identificado la presencia de los genes *bla*_{TEM-1B} y *bla*_{CMY-132} en cepas de *E. coli* aisladas de comida callejera, como se detalla en la **Gráfica 6 y 7**. El gen *bla*_{TEM-1B} es una de las variantes del gen *bla*_{TEM} que se localiza con mayor frecuencia en bacterias Gram negativas, por lo cual se encuentra mayormente distribuida a nivel mundial (**Delmani et al., 2017**). Las enzimas que codifican estos genes son responsables de conferir la resistencia a penicilina, ampicilina y cefalosporinas. En base a estudios realizados en Ecuador por **Calva et al., (2016)** se ha identificado que el gen *bla*_{TEM} presente en *E. coli* es predominante en el país localizándose en aislamientos de *E. coli* de la flora comensal de humanos y animales; lo que sugiere la posible incidencia de la mala higiene personal de los manipuladores de los alimentos expendidos en las calles. Por otra parte, aunque no existió la presencia del gen *bla*_{CTX-M} en las secuencias analizadas, se ha detectado la alta prevalencia del gen *bla*_{CTX-M} en muestras aisladas de alimentos principalmente de cevichochos y salsa de ají en Ecuador, elucidando el papel de fuentes no humanas como reservorios para la resistencia a BLEEs (**Zurita et al., 2020**). Por otra parte, la variante *bla*_{CMY-132} del gen *bla*_{CMY} que codifica la resistencia a las penicilinas, cefalosporinas y cefamicina (**Seiffert et al., 2017**); se encontró únicamente en la cepa de *E. coli* J5_2c2 aislada de jugo de caña. Investigaciones realizadas en Ecuador han identificado la incidencia del gen *bla*_{CMY} en aislamientos de origen humano y animal, por lo que se sugiere que la diseminación de estos genes a matrices alimentarias puede darse debido a la materia prima contaminada, mala higiene personal y pobres mecanismos de agua segura en el empleo de la preparación de los alimentos (**Álvarez & Vera, 2020; Garcés & Calero, 2020; Vinueza & Dávila, 2018**).

La diseminación de estos genes a diferentes especies microbianas puede atribuirse al transporte de los genes *bla*_{CMY} por plásmidos. Estudios asociados a las diferentes variantes del gen *bla*_{CMY} han indicado que existe una alta probabilidad que la propagación de estas variantes sea mediada por la HGT de plásmidos de tipo Inc11 (**Fang et al., 2018; Sidjabat et al., 2014**). De igual forma, se ha descrito que el tipo de plásmido IncFIA (localizado en los aislamientos N1_1 y J5_2c2) se encarga de movilizar y diseminar los genes *bla*_{TEM-1B} (**Rodríguez, 2020**). Por otra parte, el análisis de WGS identificó la presencia de otros ARGs tales como los genes *tet* y *sul*. Un

estudio realizado en Reino Unido ha identificado que los genes portadores de BLEEs se asocian con variantes genéticas *tet* y *sul* para una misma cepa de *E. coli*, describiendo que la co-resistencia antimicrobiana es probablemente impulsada por la ubicación de las variantes de genes *sul* y *tet* en plásmidos de tipo IncI1 (**Duggett et al., 2020**). En el presente estudio, a pesar de que se identificó diferentes tipos de plásmidos (IncFI1-1, IncFIA, IncFIB, IncFIC), se detectó que ningún gen de resistencia se ubicó en los plásmidos localizados. En este contexto, un punto de inflexión de WGS de lectura corta es la difícil reconstrucción de los plásmidos, por lo cual el empleo de ensamblajes híbridos (lecturas cortas y largas) es un mecanismo efectivo para reconstruir genomas contiguos precisos que podrían mejorar la reconstrucción y caracterización de plásmidos que acarrean ARGs (**Berbers et al., 2020**).

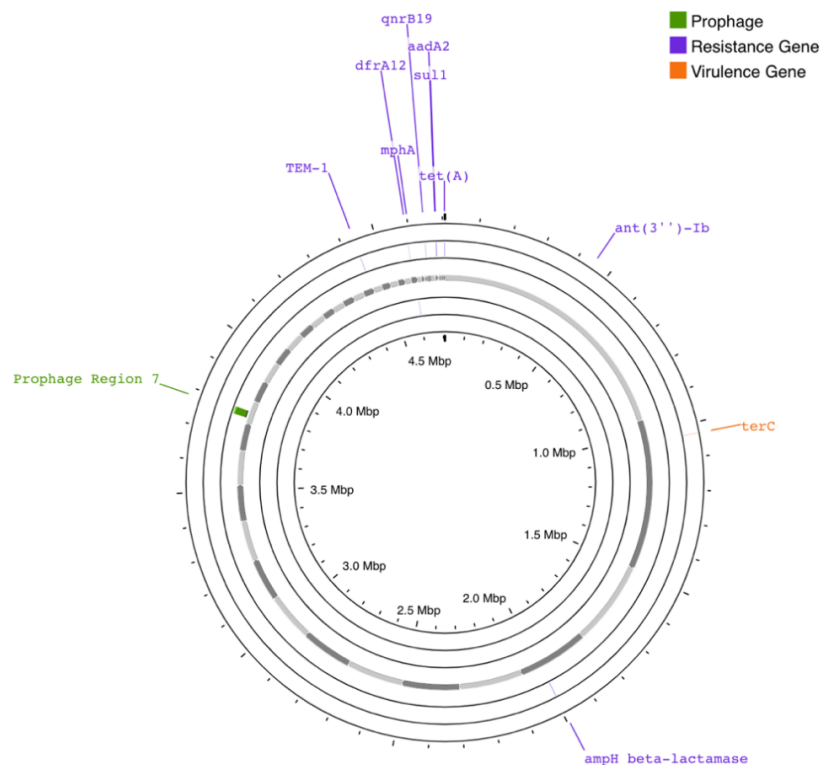
La tipificación MLST reveló que las cepas de *E. coli* N1_1 y Jc5_2C2 aisladas de comida callejera presentaban los tipos de secuencias ST101 y ST12373, respectivamente. *E. coli* ST101 ha sido ampliamente detectada en aislamientos de diverso origen geográfico y diversas fuentes de aislamiento; no obstante, los aislamientos derivados de fuentes alimentarias han sido identificadas con frecuencia en países como Alemania, Bangladesh, Camboya, China, Canadá, EE. UU, México, Países Bajos, Polonia, Reino Unido, Tailandia y Tanzania (**EnteroBase, 2023**). A pesar de la baja incidencia de *E. coli* ST101 aislada de comida en países de América del Sur, se ha detectado este tipo de secuencia en aislamientos de origen animal, ambiental y humano en Brasil, Colombia, Chile y Perú, siendo la cepa de *E. coli* N1_1 la primera evidencia de ST101 en matrices alimentarias, lo que sugiere que este clon juega un papel relevante en la diseminación en matrices humana, animal, medio ambiente y en la actualidad en matrices alimentarias (**Galarce et al., 2021**). Por otra parte, *E. coli* ST12373 solo ha sido descrito en aislamientos de matrices humanas en Australia, siendo un linaje inusualmente detectado (**EnteroBase, 2023**).

Para las cepas de *E. coli* aisladas de comida callejera se identificó la patogenicidad detectándose genes de virulencia (*etsC*, *gad*, *hlyE* *hlyF*, *iroN*, *iss*, *ompT* y *traT*)

relacionados con patotipos de origen extraintestinal de *E. coli* (Founou et al., 2022). Adicionalmente, se identificaron genes de enterotoxinas tales como *astA*. La alta detección de genes de virulencia localizados sugiere que las cepas aisladas de comida callejera tenían el potencial de ser patógenas para los seres humanos y los animales. De esta forma se ha identificado que las cepas portadoras de BLEEs, no solo albergan ARGs sino también factores de virulencia, lo cuales pueden ser transmitidos a través de HGT a otras bacterias diseminándose con amplio espectro a diversas matrices; motivo por el cual, se enfatiza en la necesidad de garantizar las medidas adecuadas de seguridad alimentaria y medidas acertadas de prevención y control de infecciones en los hospitales.

Figura 8

Mapa circular de la secuencia genómica completa de la cepa de *E. coli* L6by aislada de lechuga (*L. sativa*).



Nota: Los genes de resistencia, virulencia y presencia de regiones intactas de profagos están codificados por colores de la siguiente forma: verde (regiones intactas de profagos), morado (ARGs), y anaranjado (genes de virulencia)

En la **Figura 8**, se visualiza un mapa circular de la secuencia de genoma completo de la cepa *E. coli* L6by aislada de lechuga (*L. sativa*). Se ha mencionado que los productos de origen agrícola suelen estar contaminados por *E. coli* RAM proveniente de fertilizantes elaborados a base de estiércol, agua de riego, suelo, animales salvajes e insectos (**Reid et al., 2020**). En dicho aislamiento se han identificado gran variedad de ARGs destacando los genes que confieren resistencia β -lactámicos (*bla*_{TEM-1}, *ampH*), tetraciclinas (*tetA*), quinolonas (*qnrB19*), macrólidos (*mphA*), aminoglucósidos (*aadA2*, *ant (3'')-Ib*) y antagonista de la vía del folato (*sulI*, *dfrA12*). La presencia de *E. coli* portadora de una amplia gama de ARGs representa un motivo de preocupación, por lo tanto, es indispensable aportar hacia el desarrollo de estrategias que garanticen la seguridad alimentaria de vegetales sin previo tratamiento térmico distribuidos ampliamente en mercados agrícolas, de tal forma que se prevenga y reduzca la propagación de infecciones resistentes a antibióticos entre la población.

Por otra parte, el tipo de secuencia identificado para la cepa *E. coli* L6by fue ST1286, este tipo de secuencia pertenece al complejo clonal ST10 Cplx distribuido ampliamente en América, África, Asia, Europa y Oceanía. En adición, se ha identificado este tipo de secuencia en aislamientos de matrices ambientales, humanas y animales en Ecuador, lo que insinúa que este tipo de secuencia se encuentra ampliamente diseminado en diferentes matrices (**EnteroBase, 2023**). Por otra parte, el plásmido identificado en la secuencia de la cepa *E. coli* L6by fue Col(pHAD28). En el presente estudio se identificó que el gen *qnrB19* y el plásmido Col(pHAD28) se encontraban ubicados en un mismo contig (ABGJUB010000055.1). Según **Jibril et al., (2021)**, los plásmidos Col(pHAD28) habitualmente albergan genes *qnrB*; adicionalmente, el plásmido Col(pHAD28) caracterizado en diversos estudios ha sido identificado en cepas de *Salmonella* (**Jibril et al., 2021**); infiriendo la alta incidencia de diseminación de ARGs a diversos géneros y especies microbianas mediante mecanismos de HGT.

3.1.2. Evaluación de la distancia filogenética entre aislamientos de *E. coli* de origen ecuatoriano con aislamientos de diverso origen geográfico

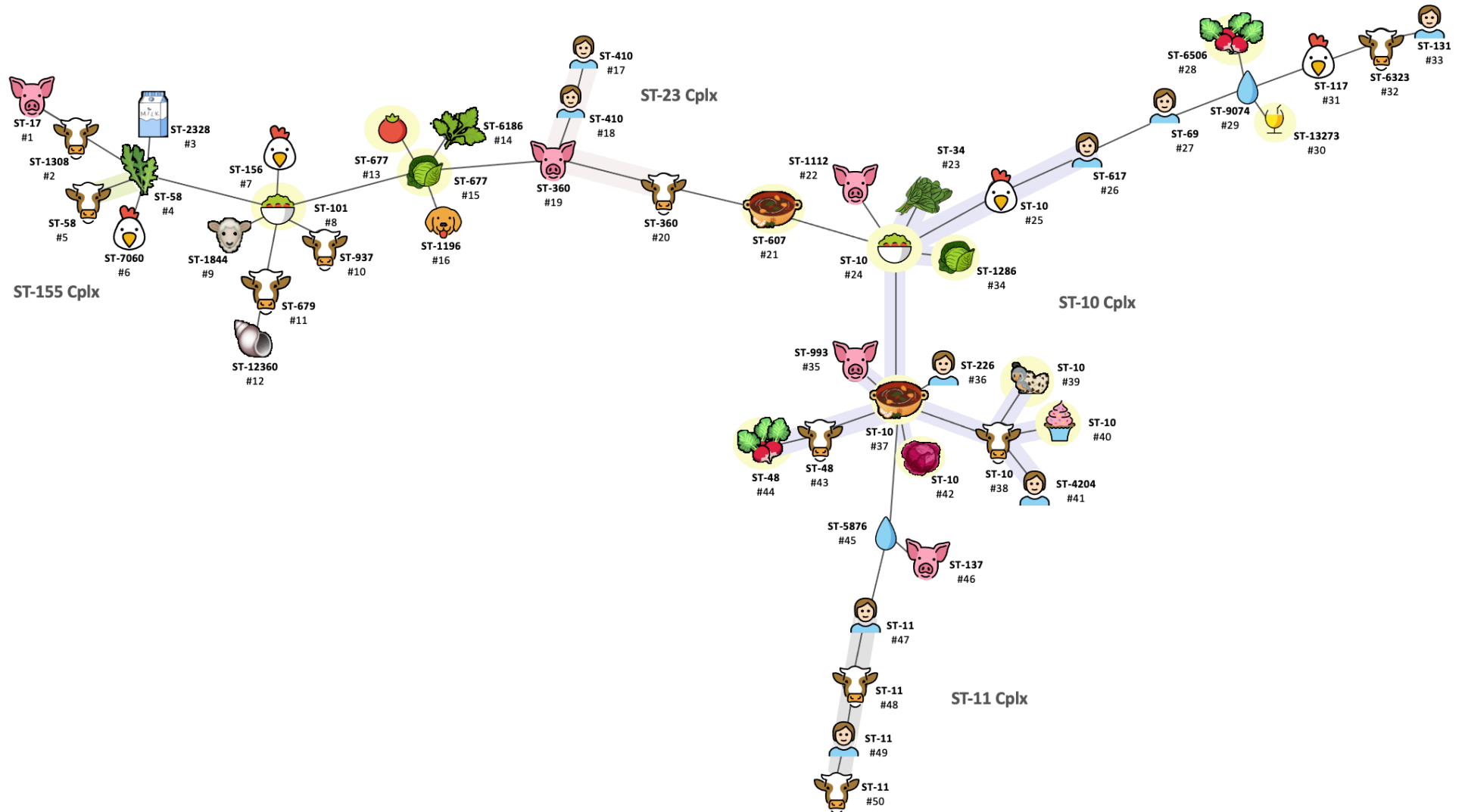
3.1.2.1. Estructura de población basada en la estrategia cgMLST

Se realizó una comparativa de los aislamientos de *E. coli* caracterizados en el presente estudio con aislamientos de diverso origen geográfico a fin de identificar las relaciones y diferencias existentes entre los genomas. Los análisis se realizaron empleando 50 secuencias de genoma completo de *E. coli* aisladas de diversos orígenes y países. Dentro de las 50 secuencias de genoma completo se incluían las secuencias de genoma completo de origen local caracterizadas en el presente estudio; se desarrolló una comparación filogenómica cgMLST basado en las secuencias de genoma completo (2485 alelos). Posteriormente, se obtuvo un árbol de expansión mínimo construido mediante PHYLOViZ 2.0.

En la **Figura 9**, se identifica un esquema general de la filogenia de las 50 cepas de *E. coli* pertenecientes a 20 países tales como Alemania, Argentina, Australia, Chile, Ecuador, Egipto, España, Estados Unidos, Francia, India, Indonesia, Italia, Japón, Malasia, México, Países Bajos, Sudáfrica, Suecia, Suiza y Tailandia. La fuente de donde fueron obtenidas las cepas derivaba de muestras de origen animal, ambiental, alimentario y humano.

Figura 9

Árbol de expansión mínima (Minimum-spanning tree) para 50 aislados basados en el cgMLST de *E. coli*



Nota: Las imágenes representan el origen de las muestras. Las etiquetas corresponden al: 1) perfil alélico (ST, del inglés sequence type) y 2) codificación de identificación (CI). Los aislamientos sombreados corresponden a STs asociados o relacionados (ST complex) y los aislamientos encerrados en círculos de color amarillo representan los aislamientos analizados en el presente estudio

El análisis cgMLST mostró que los genomas se agrupaban según el serotipo y tipo de secuencia, indicando que los aislados de *E. coli* pertenecían a diferentes linajes; de forma que, de los 50 genomas, 24 se agruparon en 4 grupos (Complex “Cplx”), mientras que los genomas restantes no se agruparon. El primer grupo de aislados identificado perteneció al ST-10 Cplx (ST10, ST34, ST48, ST617, ST993, ST1286, ST4204), el segundo grupo abarcó el ST-11 Cplx (ST11), el tercer grupo comprendió el ST-23 Cplx (ST360, ST410) y el cuarto grupo incluyó el ST-155 Cplx (ST58).

En esta dirección, en el árbol de expansión mínima se exhibe que siete de las trece cepas de *E. coli* de origen local analizadas en el presente estudio (cepas Ch5_1a1, L6bY, M1_1, C2_1c1, E2_2b, CM3aX, R6c2), pertenecían a ST-10 Cplx. De igual forma, se logra evidenciar la alta incidencia de este complejo clonal en diversas matrices, detallándose aislamientos de origen alimentario, ambiental, animal y humano en países como Alemania, Australia, Chile, Egipto, Italia, India y México. Estudios han reportado que, *E. coli* ST-10 Cplx es uno de los complejos clonales mayoritariamente distribuidos, puesto que *E. coli* del complejo clonal 10 es altamente eficaz colonizando humanos, animales, productos frescos y nichos ambientales (**Liu et al., 2016; Manges et al., 2015**). Adicionalmente se ha identificado que *E. coli* ST-10 Cplx con frecuencia son resistentes a antimicrobianos y potenciales patógenos extraintestinales para humanos y animales (**Reid et al., 2019**). En este punto se puede inferir que es probable que los aislamientos designados a los mismos ST o complejos clonales ST-Complex tengan un antepasado reciente común.

En la misma línea se logra identificar que el aislamiento de CI#24 de origen ecuatoriano se encontraba en un punto central, detectándose que existían entre 326 hasta 667 diferencias alélicas entre los aislados CI #25, #37, #38, #39, #40 y #42 con referencia al aislamiento #24. Estos resultados evidencian que los aislamientos de ST10 se encuentran relativamente más cercanos a diferencia de los ST variables que conforman el complejo clonal ST-10Cplx, lo que sugiere que existe una amplia diversidad de STs dentro de dicho complejo clonal. Por otra parte, también se ha descrito que, la cercanía o lejanía de los linajes depende de la serotipificación basada

en la presencia de antígenos O (diferenciación basada en la composición química) y H (diferenciación basada en el contenido proteico de los flagelos) (CDC, 2022). Es de esta forma que se detectó que las cepas NAEC1115 (CI#25) y 10270 (CI#26) portaban el serovar O101:H9 y las cepas de origen ecuatoriano C2_1c1 (CI#39) y E2_2b (CI#40) portaban el serovar O51:H4. En referencia a estudios realizados en China, se describe que las cepas que pertenecen al serogrupo 0:51 de *E. coli* son clasificadas como *E. coli* enteropatógena atípica (aEPEC), pues estos serotipos han sido identificados y aislados con frecuencia en pacientes con gastroenteritis; sin embargo, con poca incidencia han sido también aislados de matrices animales y carne cruda (Xu et al., 2017). Para las cepas *E. coli* C2_1c1 (#39) y E2_2b (#40) no se encontraron diferencias alélicas puesto que estos aislamientos presentaban variaciones nulas en las secuencias elucidando la presencia de aislamientos homólogos que indican que provienen de una ascendencia común.

Por otro lado, empleando SerotypeFinder 2.0 se identificó que las cepas SAKAI (Japón CI#47), MC2 (Francia CI#48), STEC 343 (Países Bajos CI#49), C1-057 (Estados Unidos CI#50) pertenecientes a *E. coli* ST-11 Cplx presentaron un mismo serotipo, O157:H7. Se ha descrito que el serotipo *E. coli* O157:H7 es uno de los serotipos ampliamente distribuidos en patógenos que provocan ETAs, ha sido predominante en países como Argentina, EE. UU, Gran Bretaña y Japón; sin embargo, dicho tipo de secuencia y serotipo no fue localizado en aislamientos de origen ecuatoriano. Curiosamente, las cepas de *E. coli* O157:H7 fueron aisladas de heces de animales (ganado) y humanos. Según Mead & Griffin, (1998), *E. coli* O157:H7 regularmente se encuentra en las heces de ganado sano y se disemina mediante alimentos y agua contaminados. De igual forma, las diferencias alélicas localizadas en estos aislamientos variaron de 174 hasta 256, especificando que las cepas STEC 343 (Países Bajos CI#49) y C1-057 (Estados Unidos CI#50) se localizaban mucho más cerca dado el reducido número de diferencias alélicas, detectándose 174 diferencias alélicas. La amplia distribución geográfica de este tipo de secuencia ST11 es un indicativo de la preservación de este linaje y diseminación con el paso de los años, puesto que la cepa *E. coli* SAKAI fue identificada en 1996 en Japón (Hayashi et al., 2001). Por su parte, este *E. coli* ST11 ha sido ampliamente detectado en América del Norte en EE. UU y

Canadá, y con muy baja prevalencia en países de América del Sur identificándose en Argentina y Paraguay.

Por otra parte, el complejo clonal ST-23Cplx ha sido detectado en las cepas de *E. coli* C2117, 53037, STEC 778g y STEC908e2 identificadas en Tailandia, Países Bajos y Australia. Aunque se ha descrito la presencia de dichas cepas en Asia, Europa y Oceanía, se ha identificado que este tipo de complejo clonal se encuentra también ampliamente diseminado en América y África, identificándose en América del Sur en países como Brasil, Ecuador y Perú; los aislamientos ecuatorianos han sido en su mayoría detectados en muestras de origen humano en estudios de vigilancia desarrollados por la Universidad de California-Berkley, el Instituto de Tecnología de Georgia y la Unidad de Investigación de Enfermedades Transmitidas por Alimentos y Resistencia a los Antimicrobianos (UNIETAR) (**EnteroBase, 2023**).

El último complejo clonal *E. coli* ST-155 Cplx identificado en matrices alimentarias y animales, unifica el tipo de secuencia ST58 proveniente de Alemania (RE2 CI#4) y Sudáfrica (CFG_335 CI#5). Se ha descrito que *E. coli* ST58 del complejo clonal ST-155 Cplx es un grupo clonal ExPEC causante de infecciones del torrente sanguíneo; los aislamientos de *E. coli* ST58 han sido identificados en animales productores de alimentos, estiércol, aguas alledañas a granjas, cereales (cebada y avena) (**Reid et al., 2022**). A pesar de la nula incidencia de este tipo de secuencia en las secuencias analizadas en el presente estudio, en Ecuador se ha detectado la presencia en muestras de origen ambiental, animal y humano; sin embargo, no se ha reportado la presencia en muestras de origen alimentario, según datos de vigilancia realizados por la Universidad de California-Berkley, el Instituto de Tecnología de Georgia y el Centro de Medicina Veterinaria (CVM)-FDA (**EnteroBase, 2023**).

Por su parte, las cepas T5bX y C5b1 de ST677 identificadas como secuencias portadoras de STEC, se encontraron filogenéticamente cercanas, pues que se determinó que presentaban una diferencia alélica de 2, por lo cual se infiere que estos

aislamientos provenían de un mismo ancestro en común. Por otra parte, se logra identificar que existen aislamientos de diverso origen geográfico relacionados filogenéticamente a los aislamientos ecuatorianos mencionados tales como la cepa de *E. coli* EK 5.12 (Alemania); no obstante, las diferencias alélicas identificadas para dichos aislamientos con respecto a las secuencias de origen ecuatoriano eran considerables, identificándose diferencias alélicas de 1590 con respecto a las cepas T5bX y C5b1, respectivamente, lo que se infiere que estos aislamientos pueden no compartir un ancestro en común. Por otra parte, los aislamientos de *E. coli* N1_1 (CI#8), Jc5_2c2 (CI#30) y L6by (CI#34) portadores de BLEEs presentaban diferencias alélicas significativas por lo cual no se puede inferir que dichos aislamientos no comparten un ancestro en común. En cuanto al serovar identificado para dichas cepas fue H40, O12:H48 y O16:H40 para las cepas N1_1, Jc5_2c2 y L6by, respectivamente.

En el caso de la cepa de *E. coli* N1_1 se identifica gráficamente la relación filogenética con las cepas 71CH (CI#7), MEZEC10 (CI#9), STEC1828-1 (CI#10), STEC816b (CI#11) y EPEC919f (CI#12); sin embargo, se identifican grandes diferencias alélicas, determinándose diferencias de 1584, 1647, 1494, 1598 y 1776, respectivamente para cada aislamiento con referencia al aislamiento ecuatoriano. Los serovares detectados en la presente eran variables determinándose los serotipos H28, O8:H49, O43:H2, O100:H30 y O26:H11. Con respecto a la cepa Jc5_2c2 se encontró una diferencia alélica de 2249 con referencia a su aislamiento más cercano siendo este la cepa *E. coli* EC14 aislada en Malasia. De esta forma se infiere que el análisis MLST de genoma central mostró una distribución aleatoria para las secuencias portadoras de ARGs, y no se identificó una asociación con un linaje clonal específico, salvo la cepa de *E. coli* L6by de complejo clonal ST-10 Cplx.

De forma general, se logra evidenciar que los diferentes tipos de secuencias y serovares se encuentran ampliamente diseminados en diversas matrices y origen geográfico desde el punto de vista filogenético; por lo cual, se puede inferir que los aislamientos que no pertenecen a un complejo clonal, no exhiben relaciones filogenéticas cercanas. Por su parte se demuestra que los aislamientos filogenéticamente cercanos eran de origen animal y humano, y por su parte los aislamientos provenientes de matrices

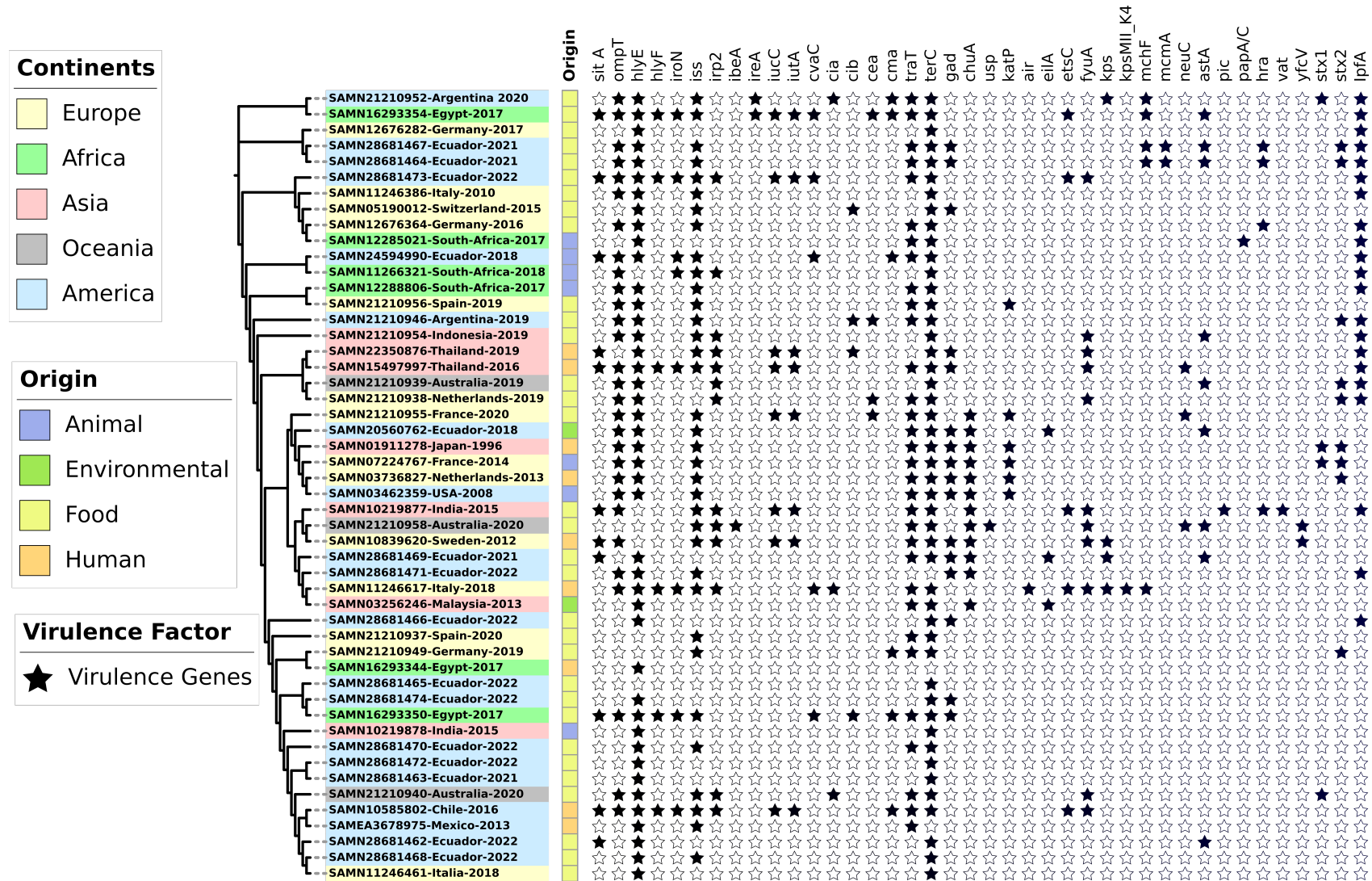
alimentarias eran filogenéticamente cercanas a aislamientos de origen ambiental y animal.

3.1.2.2. Caracterización de secuencias de diverso origen geográfico y análisis comparativo con secuencias de origen local

Se realizó la caracterización de los factores de virulencia, ARGs y presencia de tipos de plásmidos para las secuencias de diverso origen geográfico empleando la plataforma VirulenceFinder, ResFinder y PlasmidFinder del CGE; dichas caracterizaciones fueron realizadas a fin de identificar la distribución de los factores de virulencia, los ARGs y la presencia de plásmidos en aislamientos de diverso origen geográfico y fuente de aislamiento. Se identificaron 40 factores de virulencia, diversos genes para 9 familias de antibióticos y 18 tipos de plásmidos. Cada árbol de filogenia rectangular cuenta con una leyenda que indica el continente del cual es proveniente el aislamiento, el origen o matriz, y la representación de presencia de los genes y tipo de plásmido; la presencia de los genes de virulencia, resistencia y tipos de plásmidos se identifica con la figura pintada de negro o llena como se muestra en la **Figura 10, 11 y 12**. La información detallada de los aislamientos se presenta en el apartado de anexos bajo la identificación de **Anexo 3**.

Figura 10

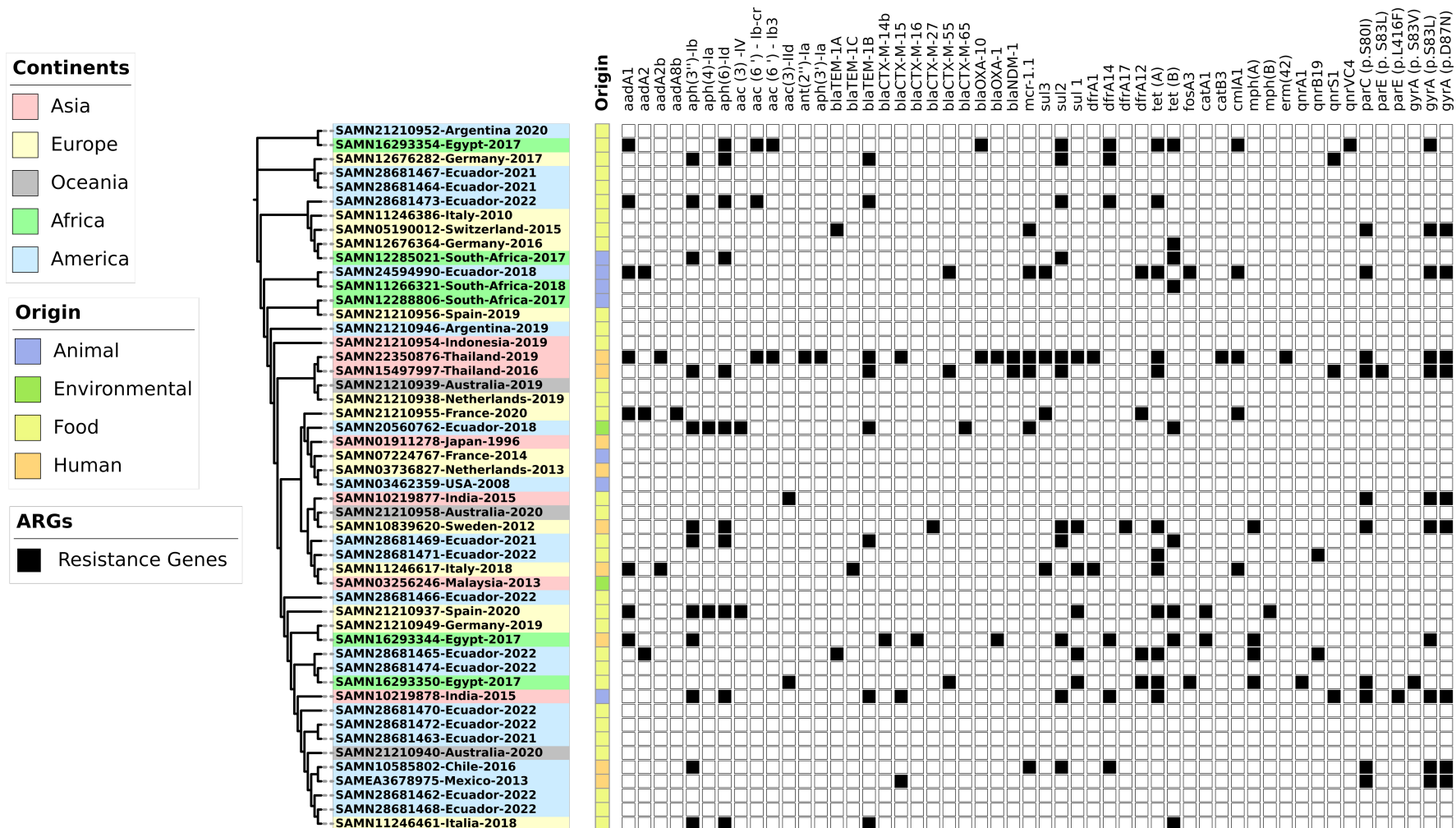
Árbol de filogenia rectangular comparativo y predicción de genes de virulencia para 50 aislados de *E. coli* de diverso origen geográfico



Nota: Cada columna identificada en la parte superior corresponde al factor de virulencia predicho en cada secuencia. La detección del factor de virulencia se identifica con la figura llena

Figura 11

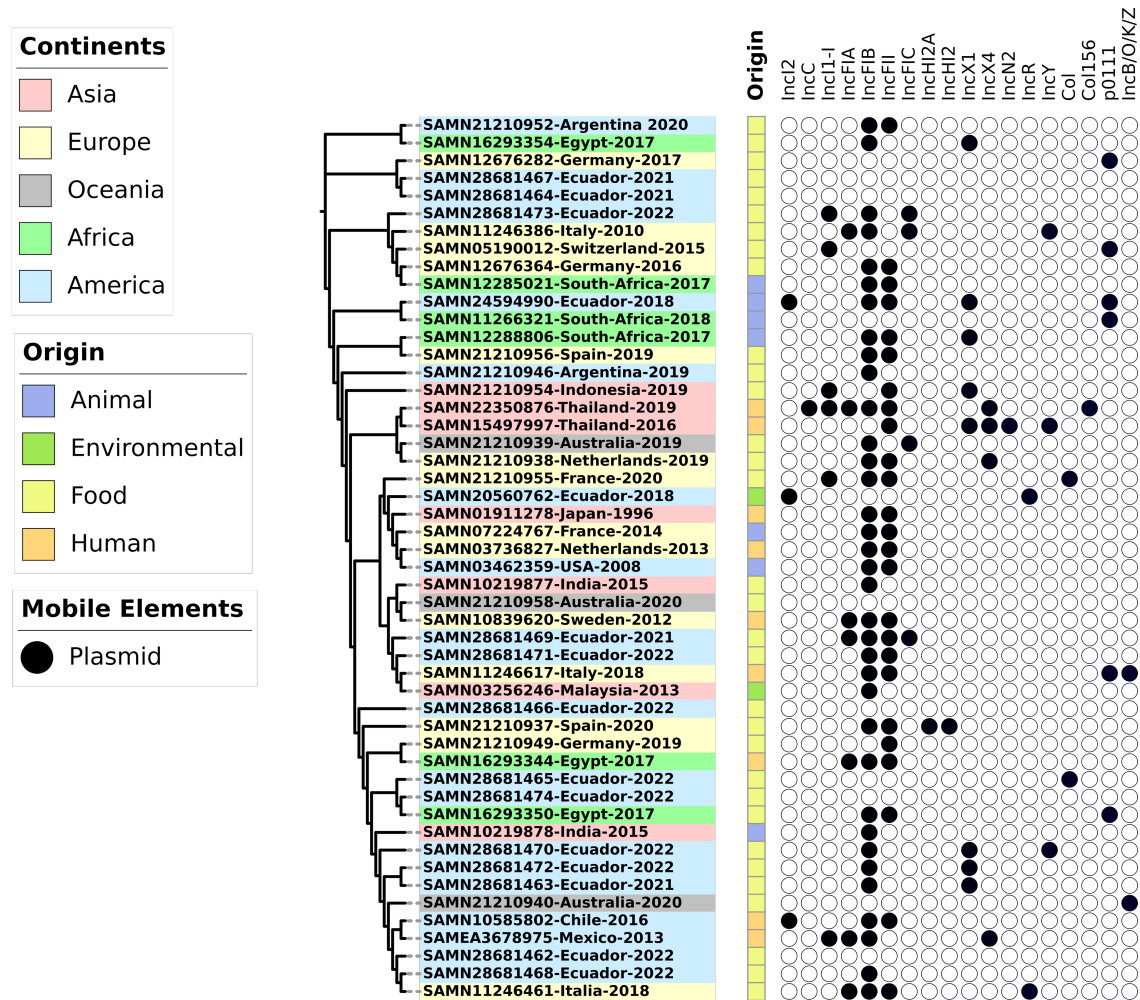
Árbol de filogenia rectangular comparativo y predicción de ARGs para 50 aislados de *E. coli* de diverso origen geográfico



Nota: Cada columna identificada en la parte superior corresponde al ARGs predichos en cada secuencia. La detección del gen que confiere la resistencia se identifica con la figura llena

Figura 12

Árbol de filogenia rectangular comparativo y predicción de plásmidos para 50 aislados de E. coli de diverso origen geográfico.



Nota: Cada columna identificada en la parte superior corresponde a los tipos de plásmidos predichos en cada secuencia. La localización de los diferentes plásmidos se identifica con la figura llena

De forma general, los aislamientos de origen alimentario de origen ecuatoriano caracterizados en el presente estudio presentan un bajo porcentaje de ARGs, destacando principalmente los genes de resistencia a aminoglicósidos como *aadAI*, *aph(3'')-Ib* y *aph(6)-Id*, los cuales se encuentran rondando el 20% de las secuencias analizadas y cuyos porcentajes son similares a los detectados en aislamientos del resto de América, Europa, Asia y África, en donde han sido detectados en un rango entre el 14% y el 33%. Con respecto a los mecanismos de virulencia, aquellos que presentan un porcentaje de presencia superior al 50% corresponden a los genes *terC* (confiere resistencia al telurio), *hlyE* (codifica para una proteína formadora de poros), *iss* (codifica para la proteína de supervivencia sérica), *traT* (resistencia serica de *E. coli*), *ompT* (factor asociado a la patogenicidad extraintestinal) (Sarowska et al., 2019); los cuales guardan concordancia con los porcentajes de presencia de tales genes en los aislamientos de América, Europa, Asia y África evaluados en el presente estudio, así como con respecto a su origen. Con respecto a los plásmidos, aquellos que han sido detectados en mayor frecuencia son los de tipo IncFIB e InFII, los cuales están presentes en más del 50% de los aislamientos de origen alimentario de Ecuador, y cuyo porcentaje se asemeja a los detectados en aislamientos de origen humano y animal a nivel global.

Los plásmidos representan uno de los principales MGE. Desafortunadamente, los datos de las secuencias se basaron principalmente en la secuenciación de lectura corta (Illumina), motivo por el cual, la identificación de la ubicación específica de los ARGs y de virulencia no está clara, porque el ensamblaje *de novo* de las lecturas del ADN genómico total no permite la separación de los ensamblajes según su ubicación original (plásmidos o cromosoma) (Wang et al., 2021). Por otra parte, las comparaciones topológicas del árbol filogenético rectangular indica que las relaciones evolutivas de los ARGs, genes de virulencia y tipos de plásmidos que albergan *E. coli* no se correlacionaron completamente con el origen de los aislamientos (animal, ambiental, humano y alimentos) y su localización geográfica.

CAPÍTULO 4

CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

- Se caracterizó *in silico* secuencias de genoma completo de *E. coli* portadoras de la toxina Shiga mediante la plataforma VirulenceFinder 2.0 de CGE, identificándose que las cepas de *E. coli* C5b1 y T5bX provenientes de aislamientos de origen vegetal portaban genes *stx2* que codifican hacia la producción de toxina Shiga; empleando la herramienta PHASTER y Proksee se detectó que los genes *stx2* se localizaban en regiones de profagos. Adicionalmente se encontraron genes de resistencia que codifican las β -lactamasas de clase C y mutaciones no definidas que se ha descrito que confieren la resistencia a las cefalosporinas de amplio espectro, polimixinas y fluoroquinolonas. Simultáneamente, la tipificación MLST obtenida mediante la plataforma MLST 2.0 de CGE reveló que el tipo de secuencia para estos 2 aislamientos fue ST677.
- Se caracterizó *in silico* secuencias de genoma completo de *E. coli* portadoras de β -lactamasas de espectro extendido mediante la plataforma ResFinder 4.1 de CGE, detectándose que las 3 cepas de *E. coli* N1_1 (ST101), Jc5_2C2 (ST12373) aisladas a partir de comida callejera y la cepa L6by (ST1286) aislada de vegetales portaban los genes *bla*_{TEM} y *bla*_{CMY} que codifica para BLEEs; adicionalmente se identificaron 18 genes de resistencia adquiridos destacando que todas las cepas portaban las variantes de los genes *tet* y *sul* que confieren resistencia a la tetraciclina y al sulfametoxazol. Los tipos de plásmidos identificados en dichos aislamientos fueron IncFI1-1, IncFIA, IncFIB, IncFIC y Col; sin embargo, se identificó que únicamente el gen de

resistencia *qnrB19* que confiere resistencia a quinolonas se ubicó en la misma región del plásmido Col(pHAD28).

- La distancia filogenética entre aislamientos de origen local y de diverso origen geográfico basados en la tipificación cgMLST, detectó que 6 aislamientos locales pertenecían a un mismo complejo clonal (ST-10 Cplx) y las diferencias alélicas identificadas entre aislamientos de India y Australia fueron de 326 hasta 667 con referencia a los aislamientos ecuatorianos, lo que se infiere que estas cepas provenían de un antepasado común. Las cepas de origen ecuatoriano T5bX y C5b1 identificadas como STEC, se localizaron filogenéticamente cercanas, pues se determinó que presentaban 2 diferencias alélicas. Los aislamientos que no fueron clasificados dentro de un mismo complejo clonal y serovar no exhibieron relaciones filogenéticas cercanas. De igual forma, se determinó que las comparaciones topológicas basadas en las interacciones evolutivas de ARGs, factores de virulencia y tipos de plásmidos de *E. coli* no se correlacionaron con el origen geográfico y tipo de fuente de aislamiento.
- Se analizó secuencias de genoma completo de *E. coli* aisladas a partir de vegetales y comida callejera en función de herramientas bioinformáticas empleadas en la caracterización de la calidad del genoma de las secuencias ensambladas, determinación de ARGs y factores virulencia, tipos de plásmidos, regiones de profagos, tipos de secuencias MLST y cgMLST. En general, las secuencias analizadas presentaron altos estándares de calidad dado su bajo índice de contaminación el cual se encontró entre 0.1 % y 1.5 % y porcentaje de completitud del 100%. De igual manera se detectó una mayor frecuencia de detección de factores de virulencia, ARGs, tipos de plásmidos y cantidades superiores de regiones de profagos en secuencias aisladas a partir de comida callejera. Se logró detectar la presencia de STEC y *E. coli* portador de BLEEs en matrices alimentarias en función de la tecnología de WGS. Los hallazgos identificados en el presente estudio proporcionan una base para generar una mayor evaluación del riesgo relacionado con la RAM y la

diseminación de microorganismos potencialmente patógenos; por lo cual, a pesar que la aplicación de WGS en el país todavía se centra en situaciones académicas se propone la adopción de estas tecnologías como mecanismos de vigilancia y control epidemiológico para comprender los orígenes y diseminación de RAM.

4.2. Recomendaciones

- El empleo de WGS de lectura corta no permite ensamblar diferencialmente plásmidos y cromosoma, por lo cual el empleo de ensamblajes híbridos que implementen la WGS mediante fragmentos largos y cortos facilitarían la generación de secuencias consenso de alta calidad que diferencie plásmidos y cromosomas.
- Evaluar los mecanismos de inducción a la expresión de bacteriófagos portadores de genes *stx2*.
- Se plantea secuenciar genomas completos de *E. coli* aislados de diversas fuentes a fin de comprender la diseminación y relaciones evolutivas entre aislamientos de origen ecuatoriano.
- La importancia de la vigilancia y monitoreo epidemiológico considerando WGS contribuirá a la implementación de estrategias que permitan comprender la evolución y diseminación de RAM y patógenos potenciales en todo el mundo.

MATERIALES DE REFERENCIA

Referencia Bibliográficas

- Ahrenfeldt, J., Skaarup, C., Hasman, H., Pedersen, A. G., Aarestrup, F. M., & Lund, O. (2017). Bacterial whole genome-based phylogeny: construction of a new benchmarking dataset and assessment of some existing methods. *BMC Genomics*, *18*(1). <https://doi.org/10.1186/S12864-016-3407-6>
- Alcock, B., Huynh, W., Chalil, R., Smith, K., Raphenya, A., Wlodarski, M., Edalatmand, A., Petkau, A., Syed, S., Tsang, K., Baker, S., Dave, M., McCarthy, M., Mukiri, K., Nasir, J., Golbon, B., Imtiaz, H., Jiang, X., Kaur, K., McArthur, A. (2023). CARD 2023: expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database. *Nucleic Acids Research*, *51*(D1). <https://doi.org/10.1093/NAR/GKAC920>
- Álvarez, F., & Vera, G. (2020). *Ttipos de betalactamasas de espectro extendido en aislados de Escherichia coli (Escherich, 1885), obtenidos de muestras de heces de pacientes con gastroenteritis en Quito-Ecuador* [Universidad Central del Ecuador]. <http://www.dspace.uce.edu.ec/handle/25000/21350>
- Andersson, D., & Hughes, D. (2017). Selection and transmission of antibiotic-resistant bacteria. *Microbiology Spectrum*, *5*(4). <https://doi.org/10.1128/MICROBIOLSPEC.MTBP-0013-2016/ASSET/BD8A417B-C5AC-486F-A84A-79A70BC879EA/ASSETS/GRAPHIC/MTBP-0013-2016-FIG4.GIF>
- Ann-Luna, R. (2016). Molecular typing instruments and methods. In *Manual of Commercial Methods in Clinical Microbiology*. <https://doi.org/10.1002/9781119021872.ch18>
- Arndt, D., Grant, J., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, *44*(Web Server issue), W16. <https://doi.org/10.1093/NAR/GKW387>
- Azanaw, J., Engdaw, G. T., Dejene, H., Bogale, S., & Degu, S. (2022). Food hygiene knowledge, and practices and their associated factors of street food vendors in

- Gondar city, Northwest Ethiopia, 2021: A cross-sectional study. *Heliyon*, 8(11).
<https://doi.org/10.1016/J.HELIYON.2022.E11707>
- Barragán, G., & Calero, W. (2022). *Evaluación de la resistencia antimicrobiana de Enterobacterias aisladas a partir de vegetales de la ciudad de Riobamba* [Universidad Técnica de Ambato].
<https://repositorio.uta.edu.ec:8443/jspui/handle/123456789/35439>
- Barragán-Fonseca, G., Tubón, J., & Calero-Cáceres, W. (2022). Data on antibiogram and resistance genes of *Enterobacterales* isolated from fresh vegetables in Ecuador. *Data in Brief*, 42(108249), 0–5.
<https://doi.org/10.1016/j.dib.2022.108249>
- Berbers, B., Saltykova, A., Garcia-Graells, C., Philipp, P., Arella, F., Marchal, K., Winand, R., Vanneste, K., Roosens, N., & de Keersmaecker, S. (2020). Combining short and long read sequencing to characterize antimicrobial resistance genes on plasmids applied to an unauthorized genetically modified *Bacillus*. *Scientific Reports* 2020 10:1, 10(1), 1–13.
<https://doi.org/10.1038/s41598-020-61158-0>
- Besser, J., Carleton, H., Gerner-Smidt, P., Lindsey, R., & Trees, E. (2018). Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clinical Microbiology and Infection*, 24(4), 335–341.
<https://doi.org/10.1016/J.CMI.2017.10.013>
- Bidet, P., Birgy, A., Ouldali, N., Bechet, S., Levy, C., Madhi, F., Sobral, E., Cohen, R., & Bonacorsi, S. (2022). Comparative genomic analysis of ESBL-producing *Escherichia coli* from faecal carriage and febrile urinary tract infection in children: a prospective multicentre study. *JAC-Antimicrobial Resistance*, 4(3).
<https://doi.org/10.1093/JACAMR/DLAC056>
- Boolchandani, M., D'Souza, A., & Dantas, G. (2019). Sequencing-based methods and resources to study antimicrobial resistance. *Nature Reviews. Genetics*, 20(6), 356.
<https://doi.org/10.1038/S41576-019-0108-4>
- Bortolaia, V., Kaas, R., Ruppe, E., Roberts, M., Schwarz, S., Cattoir, V., Philippon, A., Allesoe, R., Rebelo, A., Florensa, A., Fagelhauer, L., Chakraborty, T., Neumann, B., Werner, G., Bender, J., Stingl, K., Nguyen, M., Coppens, J., Xavier, B., Aarestrup, F. (2020). ResFinder 4.0 for predictions of phenotypes

- from genotypes. *Journal of Antimicrobial Chemotherapy*, 75(12), 3491. <https://doi.org/10.1093/JAC/DKAA345>
- Buelow, E., Ploy, M., & Dagot, C. (2021). Role of pollution on the selection of antibiotic resistance and bacterial pathogens in the environment. *Current Opinion in Microbiology*, 64, 117–124. <https://doi.org/10.1016/J.MIB.2021.10.005>
- Burmeister, A. (2015). Horizontal gene transfer. *Evolution, Medicine, and Public Health*, 2015(1), 193. <https://doi.org/10.1093/EMPH/EOV018>
- Calero Cáceres, W. (2017). Evaluación de reservorios ambientales de partículas fágicas portadoras de genes resistencia a antibióticos [Universitat de Barcelona]. In *Universitat de Barcelona*. <http://diposit.ub.edu/dspace/handle/2445/105534>
- Calva, D., Toledo, Z., Ochoa, S., Arévalo, A., & Ausilia, A. (2016). Detection and molecular characterization of β -lactamase genes in clinical isolates of Gram-negative bacteria in Southern Ecuador. *The Brazilian Journal of Infectious Diseases*, 20(6), 627. <https://doi.org/10.1016/J.BJID.2016.07.001>
- Cannatelli, A., Giani, T., Aiezza, N., Di Pilato, V., Principe, L., Luzzaro, F., Galeotti, C. & Rossolini, G. (2017). An allelic variant of the PmrB sensor kinase responsible for colistin resistance in an *Escherichia coli* strain of clinical origin. *Scientific Reports 2017 7:1*, 7(1), 1–6. <https://doi.org/10.1038/s41598-017-05167-6>
- Carattoli, A., Zankari, E., Garcíá-Fernández, A., Larsen, M. V., Lund, O., Villa, L., Aarestrup, F., & Hasman, H. (2014). *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial Agents and Chemotherapy*, 58(7), 3895–3903. <https://doi.org/10.1128/AAC.02412-14>
- CDC. (2017). *Toxinas: Programa Nacional de Biovigilancia*. CDC. <https://www.cdc.gov/biomonitoring/toxins.html>
- CDC. (2019). *How antimicrobial resistance happens*. <https://www.cdc.gov/drugresistance/pdf/threats-report/Select-Germs-Develop-Resistance-Over-Time.pdf>
- CDC. (2022). *Serotypes and the importance of serotyping Salmonella*. Centers for Disease Control and Prevention. <https://www.cdc.gov/salmonella/reportspubs/salmonella-atlas/serotyping-importance.html>

- CDC. (2022). *Whole Genome Sequencing*.
<https://www.cdc.gov/pulsenet/pathogens/wgs.html>
- CDC. (2022). *Foods that can cause food poisoning*.
<https://www.cdc.gov/foodsafety/foods-linked-illness.html#:~:text=Raw%20foods%20of%20animal%20origin,vegetables%20also%20may%20get%20contaminated.>
- Cetin, Y. (2020). Microbial toxins. In A. Demirci, H. Feng, & K. Krishnamurthy (Eds.), *Food Engineering Series* (pp. 51–83). Springer.
https://doi.org/10.1007/978-3-030-42660-6_3
- CGE. (2022). *MLST 2.0*. <https://cge.food.dtu.dk/services/MLST/>
- CGE. (2022). *ResFinder 4.1*. <https://cge.food.dtu.dk/services/ResFinder/>
- Cho, H., & Misra, R. (2021). Mutational activation of antibiotic-resistant mechanisms in the absence of major drug efflux systems of *Escherichia coli*. *Journal of Bacteriology*, 203(14). <https://doi.org/10.1128/JB.00109-21>
- Chowdhury, N., Suhani, S., Purkaystha, A., Begum, M., Raihan, T., Alam, M., Islam, K., & Azad, A. (2019). Identification of AcrAB-TolC efflux pump genes and detection of mutation in efflux repressor AcrR from omeprazole responsive multidrug-resistant *Escherichia coli* isolates causing urinary tract infections. *Microbiology Insights*, 12, 117863611988962.
<https://doi.org/10.1177/1178636119889629>
- Clark, D., Pazdernik, N., & McGehee, M. (2019). Plasmids. In *Molecular Biology* (pp. 712–748). Academic Cell. <https://doi.org/10.1016/B978-0-12-813288-3.00023-9>
- Clausen, P., Aarestrup, F., & Lund, O. (2018). Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinformatics*, 19(1). <https://doi.org/10.1186/S12859-018-2336-6>
- Coolen, J., den Drijver, E., Verweij, J., Schildkraut, J., Neveling, K., Melchers, W., Kolwijck, E., Wertheim, H., Kluytmans, J., & Huynen, M. (2021). Genome-wide analysis in *Escherichia coli* unravels a high level of genetic homoplasmy associated with cefotaxime resistance. *Microbial Genomics*, 7(4). <https://doi.org/10.1099/MGEN.0.000556>
- Cox, G., & Wright, G. (2013). Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *International Journal of Medical Microbiology*, 303(6–7), 287–292. <https://doi.org/10.1016/J.IJMM.2013.02.009>

- Curtis, D., Hill, A., Wilcock, A., & Charlebois, S. (2014). Foodborne and waterborne pathogenic bacteria in selected Organization for Economic Cooperation and Development (OECD) Countries. *Journal of Food Science*, 79(10). <https://doi.org/10.1111/1750-3841.12646>
- da Silva, C., do Valle-Barroso, M., Gozi, K., Fontana, H., Nogueira, M., Lincopan, N., & Casella, T. (2022). Genomic analysis of *Escherichia coli* circulating in the Brazilian poultry sector. *Brazilian Journal of Microbiology*, 53(4), 2121–2131. <https://doi.org/10.1007/S42770-022-00799-X/METRICS>
- De Maio, N., Shaw, L., Hubbard, A., George, S., Sanderson, N., Swann, J., Wick, R., Oun, M., Stubberfield, E., Hoosdally, S., Crook, D., Peto, T., Sheppard, A., Bailey, M., Read, D., Anjum, M., Sarah-Walker, A., & Stoesser, N. (2019). Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. *Microbial Genomics*, 5(9). <https://doi.org/10.1099/mgen.0.000294>
- Delmani, F., Jaran, A., Tarazi, Y., Masaadeh, H., Delmani, F., al Tarazi, Y., & Zaki, O. (2017). Characterization of ampicillin resistant gene (*bla*_{TEM-1}) isolated from *E. coli* in northern Jordan. Characterization of Ampicillin Resistant Gene (*bla*_{TEM-1}) isolated from *E. coli* in Northern Jordan. *Asian Journal of Biomedical and Pharmaceutical Sciences*. <https://www.researchgate.net/publication/316221061>
- Dewey-Mattia, D., Manikonda, K., Hall, A., Wise, M., & Crowe, S. (2019). Surveillance for foodborne disease outbreaks — United States, 2009–2015. *MMWR. Surveillance Summaries*, 67(10), 1–11. <https://doi.org/10.15585/MMWR.SS6710A1>
- Dobrindt, U., Tjaden, S., Shah, S., & Hacker, J. (2015). Mobile genetic elements and pathogenicity islands encoding bacterial toxins. In *The Comprehensive Sourcebook of Bacterial Protein Toxins* (pp. 40–76). Academic Press. <https://doi.org/10.1016/B978-0-12-800188-2.00002-1>
- Duggett, N., AbuOun, M., Randall, L., Horton, R., Lemma, F., Rogers, J., Crook, D., Teale, C., & Anjum, M. F. (2020). The importance of using whole genome sequencing and extended spectrum β -lactamase selective media when monitoring antimicrobial resistance. *Scientific Reports*, 10(1). <https://doi.org/10.1038/S41598-020-76877-7>

- Economou, V., & Gousia, P. (2015). Agriculture and food animals as a source of antimicrobial-resistant bacteria. *Infection and Drug Resistance*, 8, 49–61. <https://doi.org/10.2147/IDR.S55778>
- Eftekhar, F., & Seyedpour, S. (2015). Prevalence of *qnr* and *aac(6')-Ib-cr* genes in clinical isolates of *Klebsiella pneumoniae* from Imam Hussein Hospital in Tehran. *Iranian Journal of Medical Sciences*, 40(6), 515. [/pmc/articles/PMC4628142/](https://doi.org/10.1186/s13047-015-0142-1)
- Ellington, M., Ekelund, O., Aarestrup, F., Canton, R., Doumith, M., Giske, C., Grundman, H., Hasman, H., Holden, M., Hopkins, K., Iredell, J., Kahlmeter, G., Köser, C., MacGowan, A., Mevius, D., Mulvey, M., Naas, T., Peto, T., Rolain, J. M., & Woodford, N. (2017). The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. *Clinical Microbiology and Infection*, 23(1), 2–22. <https://doi.org/10.1016/j.cmi.2016.11.012>
- Emamalipour, M., Seidi, K., Zununi-Vahed, S., Jahanban-Esfahlan, A., Jaymand, M., Majdi, H., Amoozgar, Z., Chitkushev, L., Javaheri, T., Jahanban-Esfahlan, R., & Zare, P. (2020). Horizontal gene transfer: from evolutionary flexibility to disease progression. *Frontiers in Cell and Developmental Biology*, 8, 229. <https://doi.org/10.3389/fcell.2020.00229>/BIBTEX
- Enterobase. (2023). *Enterobase - Search E. coli*. https://enterobase.warwick.ac.uk/species/ecoli/search_strains?query=st_search
- Enterobase. (2023). *Enterobase - Search E. coli ST23*. https://enterobase.warwick.ac.uk/species/ecoli/search_strains?query=st_search
- Enterobase. (2023). *Enterobase - Search E. coli ST101*. https://enterobase.warwick.ac.uk/species/ecoli/search_strains?query=st_search
- Enterobase. (2023). *Enterobase - Search E. coli ST12373*. https://enterobase.warwick.ac.uk/species/ecoli/search_strains?query=st_search
- Eskandari-Nasab, E., Moghadampour, M., & Tahmasebi, A. (2018). Prevalence of *bla*_{CTX-M} gene among extended-spectrum β -lactamases producing *Klebsiella pneumoniae* clinical isolates in Iran: A Meta-Analysis. *Iranian Journal of Medical Sciences*, 43(4), 347. [/pmc/articles/PMC6055219/](https://doi.org/10.1186/s13047-018-0119-1)

- Espitia-Navarro, H., Rishishwar, L., Mayer, ., & King Jordan, I. (2019). Bioinformatics. *Microbial Forensics*, 267–282. <https://doi.org/10.1016/B978-0-12-815379-6.00018-0>
- European Centre for Disease Prevention and Control. (2022). *STEC infection: Annual Epidemiological Report for 2020*. <https://www.ecdc.europa.eu/sites/default/files/documents/STEC-infection-AER-2020-JD-FINAL.pdf>
- Fajardo, A., Martínez-Martín, N., Mercadillo, M., Galán, J., Ghysels, B., Matthijs, S., Cornelis, P., Wiehlmann, L., Tümmler, B., Baquero, F., & Martínez, J. L. (2008). The neglected intrinsic resistome of bacterial pathogens. *PLOS ONE*, 3(2), e1619. <https://doi.org/10.1371/JOURNAL.PONE.0001619>
- Fang, L., Li, X., Li, L., Chen, M., Wu, C., Li, L., Liao, X., Liu, Y., & Sun, J. (2018). ISEcpI-mediated transposition of chromosome-borne *bla_{CMY-2}* into an endogenous ColE1-like plasmid in *Escherichia coli*. *Infection and Drug Resistance*, 11, 995. <https://doi.org/10.2147/IDR.S159345>
- FAO & WHO. (2021). Microbiological risk assessment - Guidance for food. In *Microbiological Risk Assessment Series no. 36* (p. 288). <https://doi.org/10.4060/cb5006en>
- FAO, UNEP, WHO, & WHOH. (2022). *One Health Joint Plan of Action (2022–2026)*. FAO; UNEP; WHO; World Organization for Animal Health (WOAH) (founded as OIE); <https://doi.org/10.4060/cc2289en>
- Feng, Y., Zou, S., Chen, H., Yu, Y., & Ruan, Z. (2021). BacWGSTdb 2.0: A one-stop repository for bacterial whole-genome sequence typing and source tracking. *Nucleic Acids Research*, 49(D1), D644–D650. <https://doi.org/10.1093/nar/gkaa821>
- Fleming, A. (1945). *Penicillin: Nobel Lecture*.
- Fletcher, C. D. (2021). Application of modern techniques: Next Generation Sequencing. In *diagnostic histopathology of tumors* (5th ed., Vol. 2, pp. 2305–2320). Elsevier.
- Florensa, A., Kaas, R., Clausen, P., Aytan-Aktug, D., & Aarestrup, F. (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). <https://doi.org/10.1099/MGEN.0.000748>

- Founou, L., Founou, R., Allam, M., Ismail, A., & Essack, S. (2022). Genome Analysis of ESBL-Producing *Escherichia coli* Isolated from pigs. *Pathogens*, *11*(7), 776. <https://doi.org/10.3390/PATHOGENS11070776/S1>
- Fraser, M., Fujinaga, M., Cherney, M., Melton-Celsa, A., Twiddy, E., O'Brien, A., & James, M. (2004). Structure of Shiga toxin type 2 (Stx2) from *Escherichia coli* O157:H7. *Journal of Biological Chemistry*, *279*(26), 27511–27517. <https://doi.org/10.1074/JBC.M401939200>
- Frost, L., Leplae, R., Summers, A., & Toussaint, A. (2005). Mobile genetic elements: the agents of open-source evolution. *Nature Reviews Microbiology* *2005* *3*:9, 3(9), 722–732. <https://doi.org/10.1038/nrmicro1235>
- Galarce, N., Sánchez, F., Escobar, B., Lapierre, L., Cornejo, J., Alegría-Morán, R., Neira, V., Martínez, V., Johnson, T., Fuentes-Castillo, D., Sano, E., & Lincopan, N. (2021a). Genomic epidemiology of Shiga toxin-producing *Escherichia coli* isolated from the livestock-food-human interface in South America. *Animals*, *11*(7), 1845. <https://doi.org/10.3390/ANI11071845/S1>
- Galaxy. (2023). *Galaxy*. <https://galaxy.sciensano.be/welcome/new>
- Garcés, X., & Calero, W. (2020). *Determinación de la presencia de genes de resistencia a betalactámicos y evaluación de diversidad clonal en aislados de Escherichia coli de origen canino de la ciudad de Ambato* [Universidad Técnica de Ambato]. <https://repositorio.uta.edu.ec:8443/jspui/handle/123456789/30903>
- Giri, S., Kudva, V., Shetty, K., & Shetty, V. (2021). Prevalence and characterization of extended-spectrum β -lactamase-producing antibiotic-resistant *Escherichia coli* and *Klebsiella pneumoniae* in ready-to-eat street foods. *Antibiotics*, *10*(7), 850. <https://doi.org/10.3390/ANTIBIOTICS10070850/S1>
- Gonzalez-Escalona, N., Meng, J., & Doyle, M. P. (2019). Shiga toxin-producing *Escherichia coli*. In *Food Microbiology: Fundamentals and Frontiers* (pp. 289–315). Wiley. <https://doi.org/10.1128/9781555819972.CH11>
- Gourama, H. (2020). Foodborne Pathogens. In A. Demirci, H. Feng, & K. Krishnamurthy (Eds.), *Food Engineering Series* (pp. 25–49). Springer. https://doi.org/10.1007/978-3-030-42660-6_2/COVER
- Guo, S. (2020). Whole genome sequencing analysis of antimicrobial resistant *Escherichia coli*: from food to human [Nanyang Technological University]. In *Nanyang Technological University*. <https://doi.org/10.32657/10356/145480>

- Guo, Y., Ye, F., Sheng, Q., Clark, T., & Samuels, D. C. (2014). Three-stage quality control strategies for DNA re-sequencing data. *Briefings in Bioinformatics*, *15*(6), 879. <https://doi.org/10.1093/BIB/BBT069>
- Hager, S., Jensen, E., Johnson, T., & Mitchell, D. (2017). Adaptation of *Escherichia coli* to antibiotic cycling via single nucleotide polymorphisms. *American Journal of Undergraduate Research*, *14*(1), undefined-undefined. <https://doi.org/10.33697/AJUR.2017.004>
- Hammad, A., Gonzalez-Escalona, N., el Tahan, A., Abbas, N., Koenig, S., Allué-Guardia, A., Eppinger, M., & Hoffmann, M. (2022.). Pathogenome comparison and global phylogeny of *Escherichia coli* ST1485 strains. *Scientific Reports* *1*, *12*, 18495. <https://doi.org/10.1038/s41598-022-20342-0>
- Harris, S., & Okoro, C. (2014). Whole-genome sequencing for rapid and accurate identification of bacterial transmission pathways. In *Methods in Microbiology* (Vol. 41). <https://doi.org/10.1016/bs.mim.2014.07.003>
- Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogasawara, N., Yasunaga, T., Kuhara, S., Shinagawa, H. (2001). Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes*, *8*(1), 11–22. <https://doi.org/10.1093/DNARES/8.1.11>
- He, L., Partridge, S., Yang, X., Hou, J., Deng, Y., Yao, Q., Zeng, Z., Chen, Z., & Liu, J. (2013). Complete nucleotide sequence of pHN7A8, an F33: A-: B- type epidemic plasmid carrying *bla*_{CTX-M-65}, *fosA3* and *rmtB* from China. *The Journal of Antimicrobial Chemotherapy*, *68*(1), 46–50. <https://doi.org/10.1093/JAC/DKS369>
- Hershberg, R. (2016). Codon Usage and Translational Selection. *Encyclopedia of Evolutionary Biology*, 293–298. <https://doi.org/10.1016/B978-0-12-800049-6.00178-5>
- Hozzari, A., Behzadi, P., Kerishchi-Khiabani, P., Sholeh, M., & Sabokroo, N. (2020). Clinical cases, drug resistance, and virulence genes profiling in uropathogenic *Escherichia coli*. *Journal of Applied Genetics*, *61*(2), 265–273. <https://doi.org/10.1007/S13353-020-00542-Y>

- Huang, W., Kao, C., Mao, Y., Lai, C., Lai, K., Lai, C., Tseng, C., Huang, Y., Liu, P., Mao, Y., Lai, C., Lai, C., Tseng, C., Huang, Y., Liu, P., & Kolář, M. (2021). *Shewanella algae* and *Morganella morganii* coinfection in cobra-bite wounds: A Genomic Analysis. *Life* 2021, Vol. 11, Page 329, 11(4), 329. <https://doi.org/10.3390/LIFE11040329>
- Jay, M., Loessner, J., & Golden, D. A. (2008). Modern food Microbiology. In *Modern Food Microbiology* (Seventh Edition, Vol. 250).
- Jibril, A., Okeke, I., Dalsgaard, A., Menéndez, V., & Olsen, J. (2021). Genomic analysis of antimicrobial resistance and resistance plasmids in *Salmonella* serovars from poultry in Nigeria. *Antibiotics* 2021, Vol. 10, Page 99, 10(2), 99. <https://doi.org/10.3390/ANTIBIOTICS10020099>
- Johnning, A., Kristiansson, E., Fick, J., Weijdegård, B., & Larsson, D. (2015). Resistance Mutations in *gyrA* and *parC* are Common in *Escherichia* communities of both fluoroquinolone-polluted and uncontaminated aquatic environments. *Frontiers in Microbiology*, 6(DEC), 1355. <https://doi.org/10.3389/FMICB.2015.01355>
- Joppich, M., & Zimmer, R. (2019). From command-line bioinformatics to bioGUI. *PeerJ*, 2019(11). <https://doi.org/10.7717/PEERJ.8111/SUPP-2>
- Khalifa, S., Abd El-Aziz, A., Hassan, R., & Abdelmegeed, E. (2021). β -lactam resistance associated with β -lactamase production and porin alteration in clinical isolates of *E. coli* and *K. pneumoniae*. *PloS One*, 16(5). <https://doi.org/10.1371/JOURNAL.PONE.0251594>
- Kimura, B. (2018). Will the emergence of core genome MLST end the role of *in silico* MLST. *Food Microbiology*, 75, 28–36. <https://doi.org/10.1016/J.FM.2017.09.003>
- Kleinheinz, K., Joensen, K., & Larsen, M. (2014). Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and *E. coli* virulence genes in bacteriophage and prophage nucleotide sequences. *Bacteriophage*, 4(2), e27943. <https://doi.org/10.4161/BACT.27943>
- Kohlmann, R., & Gatermann, S. (2020). Species-specific epidemiology of AmpC production. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 71(10), 2767–2768. <https://doi.org/10.1093/CID/CIAA120>

- Köser, C., Ellington, M., & Peacock, S. (2014). Whole-genome sequencing to control antimicrobial resistance. *Trends in Genetics*, 30(9), 401–407. <https://doi.org/10.1016/J.TIG.2014.07.003>
- Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bover-Cid, S., Chemaly, M., Davies, R., de Cesare, A., Herman, L., Hilbert, F., Lindqvist, R., Nauta, M., Peixe, L., Ru, G., Simmons, M., Skandamis, P., Suffredini, E., Jenkins, C., Monteiro-Pires, S., Morabito, S., Bolton, D. (2020). Pathogenicity assessment of Shiga toxin-producing *Escherichia coli* (STEC) and the public health risk posed by contamination of food with STEC. *EFSA Journal*, 18(1), e05967. <https://doi.org/10.2903/J.EFSA.2020.5967>
- Krishnamoorthy, G., Wolloscheck, D., Weeks, J., Croft, C., Rybenkov, V., & Zgurskaya, H. (2016). Breaking the permeability barrier of *Escherichia coli* by controlled hyperporination of the outer membrane. *Antimicrobial Agents and Chemotherapy*, 60(12), 7372. <https://doi.org/10.1128/AAC.01882-16>
- Larsen, M., Joensen, K., Zankari, E., Ahrenfeldt, J., Lukjancenko, O., Kaas, R., Roer, L., Leekitcharoenphon, P., Saputra, D., Cosentino, S., Thomsen, M., Cisneros, J., Jurtz, V., Rasmussen, S., Petersen, T., Hasman, H., Sicheritz-Ponten, T., Aarestrup, F., & Lund, O. (2017). The CGE Tool Box. *Applied Genomics of Foodborne Pathogens*, 65–90. https://doi.org/10.1007/978-3-319-43751-4_5
- Lee, J., Reed, J., Shields, M., Spiegel, K., Farrell, L., & Sheridan, P. (2007). Phylogenetic analysis of Shiga toxin 1 and Shiga toxin 2 genes associated with disease outbreaks. *BMC Microbiology*, 7(1), 1–12. <https://doi.org/10.1186/1471-2180-7-109/TABLES/3>
- Leonard, S., Mammel, M., Lacher, D., & Elkins, C. (2016). Strain-level discrimination of Shiga toxin-producing *Escherichia coli* in spinach using metagenomic sequencing. *PLoS ONE*, 11(12). <https://doi.org/10.1371/JOURNAL.PONE.0167870>
- Letunic, I., & Bork, P. (2007). Interactive Tree of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, 23(1), 127–128. <https://doi.org/10.1093/BIOINFORMATICS/BTL529>
- Letunic, I., & Bork, P. (2021). Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49(W1), W293–W296. <https://doi.org/10.1093/NAR/GKAB301>

- Levy, S., & Bonnie, M. (2004). Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine*, *10*(12 Suppl), S122–S129. <https://doi.org/10.1038/NM1145>
- Liu, X., Liu, H., Li, Y., & Hao, C. (2016). High prevalence of β -lactamase and plasmid-mediated quinolone resistance genes in extended-spectrum cephalosporin-resistant *Escherichia coli* from dogs in Shaanxi, China. *Frontiers in Microbiology*, *7*(NOV), 1843. <https://doi.org/10.3389/FMICB.2016.01843/BIBTEX>
- Luna-Guevara, J., Arenas-Hernández, M., Martínez, C., Silva, J., & Luna-Guevara, M. (2019). The Role of Pathogenic *E. coli* in fresh vegetables: behavior, contamination factors, and preventive measures. *International Journal of Microbiology*, 2019. <https://doi.org/10.1155/2019/2894328>
- Madigan, M., Martinko, J., Bender, K., Buckley, D., & Stahl, D. (2016). Microbiology de Brock. In *Brock biology of microorganisms*.
- Mahindroo, J., & Taneja, N. (2021). *First report of enteroaggregative E. coli harbouring NDM-4 and NDM-7 from asymptomatic children in India*. <https://doi.org/10.21203/rs.3.rs-1022546/v1>
- Manges, A., Harel, J., Masson, L., Edens, T., Portt, A., Reid-Smith, R., Zhanel, G., Kropinski, A., & Boerlin, P. (2015). Multilocus sequence typing and virulence gene profiles associated with *Escherichia coli* from human and animal sources. *Https://Home.Liebertpub.Com/Fpd*, *12*(4), 302–310. <https://doi.org/10.1089/FPD.2014.1860>
- Martin, G. (2012). Sepsis, severe sepsis and septic shock: Changes in incidence, pathogens and outcomes. In *Expert Review of Anti-Infective Therapy* (Vol. 10, Issue 6). <https://doi.org/10.1586/eri.12.50>
- Mead, P., & Griffin, P. (1998). *Escherichia coli* O157:H7. *The Lancet*, *352*(9135), 1207–1212. [https://doi.org/10.1016/S0140-6736\(98\)01267-7](https://doi.org/10.1016/S0140-6736(98)01267-7)
- Ministerio de Salud Pública. (2022). *Enfermedades transmitidas por agua o alimentos – Ministerio de Salud Pública*. Ecuador. <https://www.salud.gob.ec/enfermedades-transmitidas-por-agua-o-alimentos-2/>
- Moreno-Switt, A., Pezoa, D., Sepúlveda, V., González, I., Rivera, D., Retamal, P., Navarrete, P., Reyes-Jara, A., & Toro, M. (2019). Transduction as a potential dissemination mechanism of a clonal *qnrB19* carrying plasmid isolated from

- Salmonella* of multiple serotypes and isolation sources. *Frontiers in Microbiology*, 10, 2503. <https://doi.org/10.3389/FMICB.2019.02503/BIBTEX>
- Morgan-Linnell, S., Boyd, L., Steffen, D., & Zechiedrich, L. (2009). Mechanisms accounting for fluoroquinolone resistance in *Escherichia coli* clinical isolates. *Antimicrobial Agents and Chemotherapy*, 53(1), 235. <https://doi.org/10.1128/AAC.00665-08>
- Morrison, L., & Zembower, T. (2020). Antimicrobial Resistance. *gastrointestinal endoscopy clinics of North America*, 30(4), 619–635. <https://doi.org/10.1016/J.GIEC.2020.06.004>
- Mousavi Khaneghah, A., Abhari, K., Eş, I., Soares, M., Oliveira, R., Hosseini, H., Rezaei, M., Balthazar, C., Silva, R., Cruz, A., Ranadheera, C., & Sant’Ana, A. (2020). Interactions between probiotics and pathogenic microorganisms in hosts and foods: A review. *Trends in Food Science & Technology*, 95, 205–218. <https://doi.org/10.1016/J.TIFS.2019.11.022>
- Reglamento de Control Sanitario de Alimentos, (1992). www.lexis.com.ec
- Nagy, T., Szmolka, A., Wilk, T., Kiss, J., Szabó, M., Pászti, J., Nagy, B., & Olasz, F. (2020). Comparative genome analysis of hungarian and global strains of *Salmonella infantis*. *Frontiers in Microbiology*, 0, 539. <https://doi.org/10.3389/FMICB.2020.00539>
- Nascimento, M., Sousa, A., Ramirez, M., Francisco, A., Carriço, J., & Vaz, C. (2017). PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics*, 33(1), 128–129. <https://doi.org/10.1093/BIOINFORMATICS/BTW582>
- Neumann, B., Prior, K., Bender, J., Harmsen, D., Klare, I., Fuchs, S., Bethe, A., Zühlke, D., Göhler, A., Schwarz, S., Schaffer, K., Riedel, K., Wieler, L., & Werner, G. (2019). A core genome multilocus sequence typing scheme for *Enterococcus faecalis*. *Journal of Clinical Microbiology*, 57(3). https://doi.org/10.1128/JCM.01686-18/SUPPL_FILE/JCM.01686-18-SD002.XLS
- Nikaido, H., & Takatsuka, Y. (2009). Mechanisms of RND multidrug efflux pumps. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1794(5), 769–781. <https://doi.org/10.1016/J.BBAPAP.2008.10.004>

- Nodehi, H., Tabatabaiefar, M., & Sehhati, M. (2021). Selection of optimal Bioinformatic tools and proper reference for reducing the alignment error in Targeted Sequencing Data. *Journal of Medical Signals and Sensors*, 11(1), 37. https://doi.org/10.4103/JMSS.JMSS_7_20
- Olson, R., Assaf, R., Brettin, T., Conrad, N., Cucinell, C., Davis, J., Dempsey, D., Dickerman, A., Dietrich, E., Kenyon, R., Kuscuoglu, M., Lefkowitz, E., Lu, J., Machi, D., Macken, C., Mao, C., Niewiadomska, A., Nguyen, M., Olsen, G., Stevens, R. (2023). Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. *Nucleic Acids Research*, 51(D1), D678–D689. <https://doi.org/10.1093/NAR/GKAC1003>
- Oniciuc, E., Likotrafiti, E., Alvarez-Molina, A., Prieto, M., Santos, J., & Alvarez-Ordóñez, A. (2018). The present and future of whole genome sequencing (WGS) and whole metagenome sequencing (WMS) for surveillance of antimicrobial resistant microorganisms and antimicrobial resistance genes across the food chain. *Genes* 2018, Vol. 9, Page 268, 9(5), 268. <https://doi.org/10.3390/GENES9050268>
- Oyarzabal, O., & Kathariou, S. (2014). DNA methods in food safety: molecular typing of foodborne and waterborne bacterial pathogens. In *DNA methods in food safety: Molecular typing of foodborne and waterborne bacterial pathogens* (Vol. 9781118278673). <https://doi.org/10.1002/9781118278666>
- Pan American Health Organization, & World Health Organization. (2021). *CE168/13, Rev. 1 - One Health: A Comprehensive Approach for Addressing Health Threats at the Human-Animal-Environment Interface - PAHO/WHO | Pan American Health Organization*. <https://www.paho.org/en/documents/ce16813-rev-1-one-health-comprehensive-approach-addressing-health-threats-human-animal>
- Parks, D., Imelfort, M., Skennerton, C., Hugenholtz, P., & Tyson, G. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, 25(7), 1043–1055. <https://doi.org/10.1101/GR.186072.114>
- Parrello, B., Butler, R., Chlenski, P., Olson, R., Overbeek, J., Pusch, G., Vonstein, V., & Overbeek, R. (2019). A machine learning-based service for estimating quality of genomes using PATRIC. *BMC Bioinformatics*, 20(1), 1–9. <https://doi.org/10.1186/S12859-019-3068-Y/FIGURES/5>

- Partridge, S., Kwong, S., Firth, N., & Jensen, S. (2018). Mobile genetic elements associated with antimicrobial resistance. *Clinical Microbiology Reviews*, 31(4). <https://doi.org/10.1128/CMR.00088-17/ASSET/808BB10D-7E85-4717-AE02-F26B28D2DBE2/ASSETS/GRAPHIC/ZCM0041826370008.JPEG>
- Pervez, M., Hasnain, M., Abbas, S., Moustafa, M., Aslam, N., & Shah, S. (2022). A comprehensive review of performance of next-generation sequencing platforms. *BioMed Research International*, 2022. <https://doi.org/10.1155/2022/3457806>
- Phan, M., Nhu, N., Achard, M., Forde, B., Hong, K., Chong, T., Yin, W., Chan, K., West, N., Walker, M., Paterson, D., Beatson, S., & Schembri, M. (2017). Modifications in the *pmrB* gene are the primary mechanism for the development of chromosomally encoded resistance to polymyxins in uropathogenic *Escherichia coli*. *Journal of Antimicrobial Chemotherapy*, 72(10), 2729–2736. <https://doi.org/10.1093/JAC/DKX204>
- Pilamala, A., Linnemann, A. R., & Luning, P. A. (2023). Food safety knowledge, self-reported hygiene practices, and street food vendors perceptions of current hygiene facilities and services - An Ecuadorian case. *Food Control*, 144, 109377. <https://doi.org/10.1016/J.FOODCONT.2022.109377>
- Proksee. (2022). *Proksee - New*. <https://proksee.ca/>
- Pushpakanth, P., John Kennedy, Z., & Balachandar, D. (2019). Source tracking of Shiga-like toxin-producing *Escherichia coli* in the fresh vegetable production system of South India. *Annals of Microbiology*, 69(9). <https://doi.org/10.1007/s13213-019-01479-2>
- Raghavan, R., Kelkar, Y. D., & Ochman, H. (2012). A selective force favoring increased G+C content in bacterial genes. *Proceedings of the National Academy of Sciences of the United States of America*, 109(36), 14504–14507. https://doi.org/10.1073/PNAS.1205683109/SUPPL_FILE/PNAS.201205683SI.PDF
- Ragupathy, R., Jolley, K., Zamuner, C., Jones, J., Redfern, J., Behlau, F., Ferreira, H., & Enright, M. (2022). Core genome MLST for epidemiological and evolutionary analyses of phytopathogenic *Xanthomonas citri*. *BioRxiv*, 2022.12.20.521341. <https://doi.org/10.1101/2022.12.20.521341>
- Ramos, T., Jay-Russell, M., Millner, P., Baron, J., Stover, J., Pagliari, P., Hutchinson, M., Lilley, J., Rowley, N., Haghani, V., Aminabadi, P., Kenney, A., Hashem, F.,

- Martínez-López, B., Bihn, E., Clements, D., Shade, J., Sciligo, A., & Pires, A. (2021). Survival and persistence of foodborne pathogens in manure-amended soils and prevalence on fresh produce in certified organic farms: A Multi-Regional Baseline Analysis. *Frontiers in Sustainable Food Systems*, 5. <https://doi.org/10.3389/fsufs.2021.674767>
- Reid, C. J., Blau, K., Jechalke, S., Smalla, K., & Djordjevic, S. (2020). Whole genome sequencing of *Escherichia coli* from store-bought produce. *Frontiers in Microbiology*, 10. <https://doi.org/10.3389/FMICB.2019.03050/FULL>
- Reid, C., Cummins, M., Börjesson, S., Brouwer, M., Hasman, H., Hammerum, A., Roer, L., Hess, S., Berendonk, T., Nešporová, K., Haenni, M., Madec, J., Bethe, A., Michael, G., Schink, A., Schwarz, S., Dolejska, M., & Djordjevic, S. (2022). A role for ColV plasmids in the evolution of pathogenic *Escherichia coli* ST58. *Nature Communications* 2022 13:1, 13(1), 1–15. <https://doi.org/10.1038/s41467-022-28342-4>
- Reid, C., de Maere, M., & Djordjevic, S. (2019). Australian porcine clonal complex 10 (CC10) *Escherichia coli* belong to multiple sublineages of a highly diverse global CC10 phylogeny. *Microbial Genomics*, 5(3), e000225. <https://doi.org/10.1099/MGEN.0.000225/CITE/REFWORKS>
- Rodgers, K., McLellan, I., Peshkur, T., Williams, R., Tonner, R., Hursthouse, A., Knapp, C., & Henriquez, F. (2018). Can the legacy of industrial pollution influence antimicrobial resistance in estuarine sediments. *Environmental Chemistry Letters* 2018 17:2, 17(2), 595–607. <https://doi.org/10.1007/S10311-018-0791-Y>
- Rodríguez, J. (2020). *Diversidad plasmídica en cepas de Escherichia coli y Klebsiella pneumoniae: Comparación entre aislados comúndales y clínicos* [Universitat Autònoma de Barcelona]. <https://www.tdx.cat/bitstream/handle/10803/671304/jrn1de1.pdf;jsessionid=AE3283AABCAE6B7B0623BF94011833C3?sequence=1>
- Rodríguez-Rubio, L., Haarmann, N., Schwidder, M., Muniesa, M., & Schmidt, H. (2021). Bacteriophages of Shiga Toxin-Producing *Escherichia coli* and their contribution to Pathogenicity. *Pathogens*, 10(4), 404. <https://doi.org/10.3390/PATHOGENS10040404>

- Ruan, Z., & Feng, Y. (2016). BacWGSTdb, a database for genotyping and source tracking bacterial pathogens. *Nucleic Acids Research*, *44*(D1), D682–D687. <https://doi.org/10.1093/NAR/GKV1004>
- Saha, M., & Sarkar, A. (2021). Review on multiple facets of drug resistance: a rising challenge in the 21st century. *Journal of Xenobiotics 2021*, *Vol. 11*, Pages 197–214, *11*(4), 197–214. <https://doi.org/10.3390/JOX11040013>
- Sarowska, J., Futoma-Koloch, B., Jama-Kmiecik, A., Frej-Madrzak, M., Ksiazczyk, M., Bugla-Ploskonska, G., & Choroszy-Krol, I. (2019). Virulence factors, prevalence and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources: Recent reports. *Gut Pathogens*, *11*(1), 1–16. <https://doi.org/10.1186/S13099-019-0290-0/TABLES/5>
- Satapathy, P., Khan, K., Devi, A., Patil, A., Govindaraju, A., Gopal, S., Prasad, M., More, V., Kakarla, R., Raghu, A., Hudeda, S., More, S., & Zameer, F. (2019). Synthetic gutomics: deciphering the microbial code for futuristic diagnosis and personalized medicine. *Methods in Microbiology*, *46*, 197–225. <https://doi.org/10.1016/BS.MIM.2019.02.001>
- Seiffert, S., Carattoli, A., Schwendener, S., Collaud, A., Endimiani, A., & Perreten, V. (2017). Plasmids carrying *bla_{CMY-2/4}* in *Escherichia coli* from poultry, poultry meat, and humans belong to a novel IncK subgroup designated IncK2. *Frontiers in Microbiology*, *8*(MAR), 407. <https://doi.org/10.3389/FMICB.2017.00407/BIBTEX>
- Sekyere, J., & Asante, J. (2018). Emerging mechanisms of antimicrobial resistance in bacteria and fungi: Advances in the era of genomics. In *Future Microbiology* (Vol. 13, Issue 2). <https://doi.org/10.2217/fmb-2017-0172>
- Sidjabat, H., Seah, Y., Coleman, L., Sartor, A., Derrington, P., Heney, C., Faoagali, J., Nimmo, G., & Paterson, D. (2014). Expansive spread of IncI1 plasmids carrying *bla_{CMY-2}* amongst *Escherichia coli*. *International Journal of Antimicrobial Agents*, *44*(3), 203–208. <https://doi.org/10.1016/J.IJANTIMICAG.2014.04.016>
- Singh, T., Das, S., & Ramachandran, V. (2020). Effect of mutation on AmpC promoter in multidrug resistant isolates of diarrheagenic *Escherichia coli* in children. *International Journal of Infectious Diseases*, *101*(S1), 67. <https://doi.org/10.1016/J.IJID.2020.09.206>

- Souvorov, A., Agarwala, R., & Lipman, D. (2018). SKESA: Strategic k-mer extension for scrupulous assemblies. *Genome Biology*, *19*(1), 1–13. <https://doi.org/10.1186/S13059-018-1540-Z/TABLES/6>
- Tamma, P., Doi, Y., Bonomo, R., Johnson, J., & Simner, P. (2019). A primer on AmpC β -Lactamases: necessary knowledge for an increasingly multidrug-resistant world. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *69*(8), 1446–1455. <https://doi.org/10.1093/CID/CIZ173>
- Tenea, G., Reyes, P., Molina, D., & Ortega, C. (2023). Pathogenic microorganisms linked to fresh fruits and juices purchased at low-cost markets in Ecuador, potential carriers of antibiotic resistance. *Antibiotics 2023, Vol. 12, Page 236*, *12*(2), 236. <https://doi.org/10.3390/ANTIBIOTICS12020236>
- The World Bank, & World Health Organization. (2022). *Sustaining Action Against Antimicrobial Resistance*.
- Torres, A., Amaral, M., Bentancor, L., Galli, L., Goldstein, J., Krüger, A., & Rojas-Lopez, M. (2018). Recent advances in Shiga toxin-producing *Escherichia coli* research in Latin America. *Microorganisms*, *6*(4), 1–19. <https://doi.org/10.3390/microorganisms6040100>
- Tubón, J., Barragán-Fonseca, G., Lalaleo, L., & Calero-Cáceres, W. (2022). Data on antibiograms and resistance genes of *Enterobacteriales* isolated from ready-to-eat street food of Ambato, Ecuador. *F1000Research*, *11*(669), 1–8. <https://doi.org/10.12688/f1000research.117116.1>
- Udaondo, Z., Sittikankaew, K., Uengwetwanit, T., Wongsurawat, T., Sonthirod, C., Jenjaroenpun, P., Pootakham, W., Karoonuthaisiri, N., & Nookaew, I. (2021). Comparative analysis of PacBio and Oxford Nanopore sequencing technologies for transcriptomic landscape identification of *Penaeus monodon*. *Life*, *11*(8). <https://doi.org/10.3390/LIFE11080862/S1>
- Uddín, T., Chakraborty, A., Khusro, A., Zidan, B., Mitra, S., Emran, T., Dhama, K., Ripon, M., Gajdács, M., Sahibzada, M., Hossain, M., & Koirala, N. (2021). Antibiotic resistance in microbes: History, mechanisms, therapeutic strategies and future prospects. *Journal of Infection and Public Health*, *14*(12), 1750–1766. <https://doi.org/10.1016/J.JIPH.2021.10.020>

- Velazquez-Meza, M., Galarde-López, M., Carrillo-Quiróz, B., & Alpuche-Aranda, C. (2022). Antimicrobial resistance: One Health approach. *Veterinary World*, *15*(3), 743–749. <https://doi.org/10.14202/VETWORLD.2022.743-749>
- Vila, J., Sáez-López, E., Johnson, J., Römling, U., Dobrindt, U., Cantón, R., Giske, C., Naas, T., Carattoli, A., Martínez-Medina, M., Bosch, J., Retamar, P., Rodríguez-Banó, J., Baquero, F., & Soto, S. (2016). *Escherichia coli*: An old friend with new tidings. In *FEMS Microbiology Reviews* (Vol. 40, Issue 4). <https://doi.org/10.1093/femsre/fuw005>
- Vinueza, C., & Dávila, D. (2018). *Aislamiento e identificación de Salmonella y Escherichia coli productor de β-lactamasas de espectro extendido (BLEE) en granjas avícolas de reproductoras pesadas en las provincias de Napo y Pastaza, Ecuador* [Universidad Central del Ecuador]. <https://www.dspace.uce.edu.ec/handle/25000/17671>
- Waltenburg, M., Schwensohn, C., Madad, A., Seelman, S., Peralta, V., Koske, S., Boyle, M., Arends, K., Patel, K., Mattioli, M., Gieraltowski, L., & Neil, K. (2022). Two multistate outbreaks of a reoccurring Shiga toxin-producing *Escherichia coli* strain associated with romaine lettuce: USA, 2018-2019. *Epidemiology and Infection*, *150*. <https://doi.org/10.1017/S0950268821002703>
- Wang, X., Li, Q., Kang, J., Zhang, Z., Song, Y., Yin, D., Guo, Q., Song, J., Li, X., Wang, S., & Duan, J. (2021). Co-production of NDM-1, CTX-M-9 family and mcr-1 in a *Klebsiella pneumoniae* ST4564 strain in China. *Infection and Drug Resistance*, *14*. <https://doi.org/10.2147/IDR.S292820>
- Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology*, *39*(11), 1348. <https://doi.org/10.1038/S41587-021-01108-X>
- Watanabe, T. (1963). Infective heredity of multiple drug resistance in bacteria. *Bacteriological Reviews*, *27*(1), 87. <https://doi.org/10.1128/BR.27.1.87-115.1963>
- WHO. (2014). Antimicrobial resistance global report on surveillance. *World Health Organization*.
- Wood, R., Jensen, T., Wadsworth, C., Clement, M., Nagpal, P., & Pitt, W. (2020). Analysis of identification method for bacterial species and antibiotic resistance

- genes using optical data from DNA oligomers. *Frontiers in Microbiology*, *11*, 257. <https://doi.org/10.3389/FMICB.2020.00257/BIBTEX>
- World Health Organization. (2015). *Global action plan on antimicrobial resistance*. WHO Document Production Service.
- World Health Organization. (2015). *Global antimicrobial resistance surveillance system: Manual for Early Implementation*. In *Who*.
- World Health Organization. (2015). *Estimates of the global burden of foodborne diseases*.
- World Health Organization. (2020). *GLASS whole-genome sequencing for surveillance of antimicrobial resistance*.
- World Health Organization, & Food and Agriculture Organization of the United Nations. (2018). *Shiga toxin-producing Escherichia coli (STEC) and food: attribution, characterization, and monitoring: report*. 152.
- Xu, Y., Bai, X., Jin, Y., Hu, B., Wang, H., Sun, H., Fan, R., Fu, S., & Xiong, Y. (2017). High prevalence of virulence genes in specific genotypes of atypical enteropathogenic *Escherichia coli*. *Frontiers in Cellular and Infection Microbiology*, *7*(APR), 109. <https://doi.org/10.3389/FCIMB.2017.00109/BIBTEX>
- Yin, R., Kwok, C., & Zheng, J. (2019). Whole genome sequencing analysis. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, *1–3*, 176–183. <https://doi.org/10.1016/B978-0-12-809633-8.20095-2>
- Zhang, G., Leclercq, S., Tian, J., Wang, C., Yahara, K., Ai, G., Liu, S., & Feng, J. (2017). A new subclass of intrinsic aminoglycoside nucleotidyltransferases, ANT(3"-II), is horizontally transferred among *Acinetobacter spp.* by homologous recombination. *PLoS Genetics*, *13*(2). <https://doi.org/10.1371/JOURNAL.PGEN.1006602>
- Zhang, P., Essendoubi, S., Keenlside, J., Reuter, T., Stanford, K., King, R., Lu, P., & Yang, X. (2021). Genomic analysis of Shiga toxin-producing *Escherichia coli* O157:H7 from cattle and pork-production related environments. *NPJ Science of Food*, *5*(1). <https://doi.org/10.1038/S41538-021-00097-0>
- Zhang, Y., Liao, Y., Salvador, A., & Wu, V. (2021). Genomic characterization of two Shiga toxin-converting bacteriophages induced from environmental Shiga toxin-

- producing *Escherichia coli*. *Frontiers in Microbiology*, 12, 118. <https://doi.org/10.3389/FMICB.2021.587696/BIBTEX>
- Zhou, W., Zhang, E., Zhou, J., He, Z., Zhou, Y., Han, J., & Qu, D. (2021). Characterization and comparative genomics analysis of IncFII multi-resistance plasmids carrying *bla*_{CTX-M} and type1 integrons from *Escherichia coli*. *Frontiers in Microbiology*, 12, 3564. <https://doi.org/10.3389/FMICB.2021.753979/BIBTEX>
- Zuñiga, I., & Caro, J. (2017). Enfermedades transmitidas por los alimentos: Una mirada puntual para el personal de salud. *Enfermedades Infecciosas y Microbiología*, 37(3).
- Zurita, J., Yáñez, F., Sevillano, G., Ortega-Paredes, D., & Paz y Miño, A. (2020). Ready to eat street food: a potential source for dissemination of multidrug-resistant *Escherichia coli* epidemic clones in Quito, Ecuador. *Letters in Applied Microbiology*, 70(3), 203–209. <https://doi.org/10.1111/LAM.13263>

Anexos

Anexo 1

Tabla de identificación de las secuenciaciones del genoma completo de 13 cepas de E. coli aisladas de vegetales y comida callejera de Ecuador

Cepa	Año de aislamiento	Tipo de muestra	Ubicación geográfica	No. Acceso a BioSample	No. Acceso a SRA	No. Acceso a Assembly
M1_1	2020	Guiso	Tungurahua	SAMN28681462	SRS13189682	GCA_023561655.1
N1_1	2020	Ensalada (Mote, choclo, hablas y fritada)	Tungurahua	SAMN28681473	SRS13189681	GCA_023561435.1
C2_1c1	2021	Huevos de cordoniz cocidos	Tungurahua	SAMN28681463	SRS13189695	GCA_023561635.1
Jc5_2C2	2021	Jugo de caña (<i>Saccharum officinarum</i>)	Tungurahua	SAMN28681469	SRS13189654	GCA_023561735.1
Ch5_1a1	2021	Ensalada de chochos (<i>Lupinus mutabilis</i>)	Tungurahua	SAMN28681468	SRS13189704	GCA_023561535.1
E2_2b	2021	Merengue dulce de huevo (Espumilla)	Tungurahua	SAMN28681472	SRS13189662	GCA_023562495.1

Cepa	Año de aislamiento	Tipo de muestra	Ubicación geográfica	No. Acceso a BioSample	No. Acceso a SRA	No. Acceso a Assembly
Hmm	2021	Habas, melloco, mapahuiria	Tungurahua	SAMN28681466	SRS13189663	GCA_023561715.1
L6bY	2021	Lechuga (<i>Lactuca sativa</i>)	Chimborazo	SAMN28681465	SRS13189668	GCA_023561375.1
CM3aX	2021	Col Roja (<i>Brassica oleracea</i>)	Chimborazo	SAMN28681470	SRS13189687	GCA_024740505.1
C5b1	2021	Col Blanca (<i>Brassica oleracea</i>)	Chimborazo	SAMN28681467	SRS13189686	GCA_023561385.1
T5bX	2021	Tomate (<i>Solanum lycopersicum</i>)	Chimborazo	SAMN28681464	SRS13189707	GCA_023561575.1
R6c2	2021	Rábano (<i>Raphanus raphanistrum</i> subsp. <i>Sativus</i>)	Chimborazo	SAMN28681474	SRS13189649	GCA_023562295.1
R3cX	2021	Rábano (<i>Raphanus raphanistrum</i> subsp. <i>Sativus</i>)	Chimborazo	SAMN28681471	SRS13189648	GCA_023562535.1

Anexo 2

Tabla de información relacionada a la calidad del genoma para las 13 secuencias de genoma completo aisladas de vegetales y comida callejera de Ecuador

Cepa	No. de lecturas totales	Contenido G+C (%)	N_{50} (bp)	L_{50} (bp)	No. de contigs	No. de CDSs	No. de tRN A	No. de rRN A	Compleitud (CheckM)	Contaminación (CheckM)
M1_1	4,720,963	50.67	209,096	7	54	4,634	78	8	100	-
N1_1	5,058,102	50.54	119,031	15	115	5,024	80	7	100	1.5
C2_1c1	4,687,622	50.77	221,921	8	66	4,601	82	6	100	-
Jc5_2C2	5,125,220	50.53	276,038	7	61	5,020	82	5	100	0.1
Ch5_1a1	4,751,717	50.81	234,975	7	68	4,675	78	7	100	-
E2_2b	4,651,603	50.78	221,316	8	65	4,566	81	5	100	-
Hmm	4,830,023	50.76	228,829	8	73	4,767	74	7	100	-
L6bY	4,007,122	50.68	210,300	6	56	4,607	78	6	100	0.1
CM3aX	4,940,394	50.90	74,583	20	171	4,968	77	8	100	0.1
C5b1	5,184,284	50.71	217,893	8	87	5,242	82	10	100	0.1
T5bX	5,181,902	50.71	229,587	7	85	5,229	82	8	100	0.1
R6c2	4,616,297	50.85	83,004	14	97	4,559	76	6	100	0.2
R3cX	5,306,726	50.35	117,185	14	150	5,358	77	4	100	0.1

Anexo 3

Tabla de información compilada de secuencias de genoma completo de diverso origen geográfico y de origen local

Biosample	Bioproject	Pais	Cepa	Continente	Año	Origen	Descripción
SAMN12676282	PRJNA563564	Alemania	EK5.12	Europe	2017	Alimentario	Cilantro
SAMN12676364	PRJNA563564	Alemania	RE3	Europe	2016	Alimentario	Rúcula
SAMN21210949	PRJNA760238	Alemania	STEC 826a	Europe	2019	Alimentario	Carne de cerdo (congelada)
SAMN21210946	PRJNA760238	Argentina	STEC 816b	America	2019	Alimentario	Carne de res (congelada)
SAMN21210952	PRJNA760238	Argentina	STEC1828-1	America	2020	Alimentario	Carne de res (congelada)
SAMN21210939	PRJNA760238	Australia	STEC 908e2	Oceania	2019	Alimentario	Carne de res (congelada)
SAMN21210940	PRJNA760238	Australia	STEC 1309b	Oceania	2020	Alimentario	Carne de res (congelada)
SAMN21210958	PRJNA760238	Australia	EPEC 1668-4a	Oceania	2020	Alimentario	Carne de res (congelada)
SAMN10585802	PRJNA509640	Chile	UC-457	America	2016	Humano	Orina
SAMN20560762	PRJNA751928	Ecuador	UTA144	America	2018	Ambiental	Sedimentos de canal de riego
SAMN24594990	PRJNA751928	Ecuador	UTA069	America	2018	Animal	Heces

Biosample	Bioproject	Pais	Cepa	Continente	Año	Origen	Descripción
SAMN28681462	PRJNA415804	Ecuador	M1_1	América	2020	Alimentario	Guiso de Vaca (comida callejera)
SAMN28681463	PRJNA415804	Ecuador	C2_1c1	América	2021	Alimentario	Huevos de cordoniz (comida callejera)
SAMN28681464	PRJNA415804	Ecuador	T5bX	América	2021	Alimentario	Tomate (<i>Solanum lycopersicum</i>)
SAMN28681465	PRJNA415804	Ecuador	L6bY	América	2021	Alimentario	Lechuga (<i>Lactuca sativa</i>)
SAMN28681466	PRJNA415804	Ecuador	Hmm	América	2021	Alimentario	Frijoles hervidos, cerdo y batatas de comida callejera
SAMN28681467	PRJNA415804	Ecuador	C5b1	América	2021	Alimentario	Col Blanca (<i>Brassica oleracea</i>)
SAMN28681468	PRJNA415804	Ecuador	Ch5_1a1	América	2021	Alimentario	Ensalada <i>Lupinus mutabilis</i> (comida callejera)
SAMN28681469	PRJNA415804	Ecuador	Jc5_2C2	América	2021	Alimentario	Jugo de caña (comida callejera)

Biosample	Bioproject	Pais	Cepa	Continente	Año	Origen	Descripción
SAMN28681470	PRJNA415804	Ecuador	CM3aX	América	2021	Alimentario	Col Roja (<i>Brassica oleracea</i>)
SAMN28681471	PRJNA415804	Ecuador	R3cX	América	2021	Alimentario	Rábano (<i>Raphanus raphanistrum</i> subsp. <i>Sativus</i>)
SAMN28681472	PRJNA415804	Ecuador	E2_2b	América	2021	Alimentario	Merengue dulce de huevo (comida callejera)
SAMN28681473	PRJNA415804	Ecuador	N1_1	América	2020	Alimentario	Ensalada (comida callejera)
SAMN28681474	PRJNA415804	Ecuador	R6c2	América	2021	Alimentario	Rábano (<i>Raphanus raphanistrum</i> subsp. <i>Sativus</i>)
SAMN16293344	PRJNA666443	Egipto	23ST	Africa	2017	Humana	Heces
SAMN16293350	PRJNA666443	Egipto	21M	Africa	2017	Alimentario	Carne molida
SAMN16293354	PRJNA666443	Egipto	71CH	Africa	2017	Alimentario	Carcasas de pollo
SAMN21210937	PRJNA760238	España	STEC 10531-2	Europe	2020	Alimentario	Carne de cerdo (congelada)

Biosample	Bioproject	Pais	Cepa	Continente	Año	Origen	Descripción
SAMN21210956	PRJNA760238	España	EPEC 765e	Europe	2019	Alimentario	Carne de cerdo (congelada)
SAMN07224767	PRJNA390247	Francia	MC2 (O157:H7)	Europe	2014	Animal	Heces
SAMN21210955	PRJNA760238	Francia	EPEC 1095a	Europe	2020	Alimentario	Carne de cerdo
SAMN10219877	PRJNA495131	India	NAEC1042	Asia	2015	Alimentario	Carne de pollo cruda
SAMN10219878	PRJNA495132	India	NAEC1115	Asia	2015	Animal	Ciego de pollo
SAMN21210954	PRJNA760238	Indonesia	EPEC 919f	Asia	2019	Alimentario	Caracol de mar
SAMN11246386	PRJNA528851	Italia	EC269	Europe	2010	Alimentario	Leche
SAMN11246461	S/N	Italia	EC845	Europe	2018	Alimentario	Espinaca
SAMN11246617	PRJNA528851	Italia	EC850	Europe	2018	Humana	Heces
SAMN01911278	PRJNA226	Japón	SAKAI (EHEC) (O157:H7)	Asia	1996	Humana	Heces
SAMN03256246	PRJNA269669	Malasia	EC14	Asia	2013	Ambiental	Agua
SAMEA3678975	PRJEB11777	Mexico	10270	America	2013	Humana	Abcesos hepatico
SAMN03736827	PRJNA285020	Países Bajos	STEC 343 (O157:H7)	Europe	2013	Humana	Heces

Biosample	Bioproject	Pais	Cepa	Continente	Año	Origen	Descripción
SAMN21210938	PRJNA760238	Países Bajos	STEC 778g	Europe	2019	Alimentario	Carne de cerdo (congelada)
SAMN11266321	PRJNA293225	Sudáfrica	MEZEC10	Africa	2018	Animal	Oveja
SAMN12285021	PRJNA554852	Sudáfrica	CFG_335	Africa	2017	Animal	Heces
SAMN12288806	PRJNA555014	Sudáfrica	CFC_154	Africa	2017	Animal	Heces
SAMN10839620	PRJNA517648	Suecia	B1483-PB_2012	Europe	2012	Humana	Sangre
SAMN05190012	PRJNA323827	Suiza	S51	Europe	2015	Alimentario	Carne de ave
SAMN15497997	PRJNA525849	Tailandia	53037	Asia	2016	Humana	Orina
SAMN22350876	PRJNA525849	Tailandia	C2117	Asia	2019	Humana	Pus
SAMN03462359	PRJNA280460	USA	C1-057 (O157:H7)	America	2008	Animal	Heces rectales

Anexo 4*Códigos de identificación (CI) y tipo de complejo clonal (ST Complex)*

Biosample	País	Cepa	Origen	Descripción	Código de Identificación	ST	ST Complex
SAMN21210956	España	EPEC 765e	Alimentario	Carne de cerdo (congelada)	1	17	ST20 Cplx
SAMN12288806	Sudáfrica	CFC_154	Animal	Heces de Ganado	2	1308	ST86 Cplx
SAMN11246386	Italia	EC269	Alimentario	Leche	3	2328	S/N
SAMN12676364	Alemania	RE3	Alimentario	Rúcula	4	58	ST155 Cplx
SAMN12285021	Sudáfrica	CFG_335	Animal	Heces Ganado	5	58	ST155 Cplx
SAMN05190012	Suiza	S51	Alimentario	Carne de ave congelada (pollo)	6	7060	S/N
SAMN16293354	Egipto	71CH	Alimentario	Carcasas de pollo	7	156	ST156 Cplx
SAMN28681473	Ecuador	N1_1	Alimentario	Ensalada (comida callejera)	8	101	ST101 Cplx
SAMN11266321	Sudáfrica	MEZEC10	Animal	Oveja	9	1844	S/N
SAMN21210952	Argentina	STEC1828-1	Alimentario	Carne de res (congelada)	10	937	S/N
SAMN21210946	Argentina	STEC 816b	Alimentario	Carne de res (congelada)	11	679	ST469 Cplx

Biosample	País	Cepa	Origen	Descripción	Código de Identificación	ST	ST Complex
SAMN21210954	Indonesia	EPEC 919f	Alimentario	Caracol de mar	12	12360	ST29 Cplx
SAMN28681464	Ecuador	T5bX	Alimentario	Tomate (<i>Solanum lycopersicum</i>)	13	677	S/N
SAMN12676282	Alemania	EK5.12	Alimentario	Cilantro	14	6186	S/N
SAMN28681467	Ecuador	C5b1	Alimentario	Col Blanca (<i>Brassica oleracea</i>)	15	677	S/N
SAMN24594990	Ecuador	UTA069	Animal	Heces de <i>Canis lupus</i>	16	1196	S/N
SAMN22350876	Tailandia	C2117	Humana	Pus	17	410	ST23 Cplx
SAMN15497997	Tailandia	53037	Humana	Orina	18	410	ST23 Cplx
SAMN21210938	Países Bajos	STEC 778g	Alimentario	Carne de cerdo (congelada)	19	360	ST23 Cplx
SAMN21210939	Australia	STEC 908e2	Alimentario	Carne de res (congelada)	20	360	ST23 Cplx
SAMN28681466	Ecuador	Hmm	Alimentario	Frijoles hervidos, cerdo y batatas de (habas melloco y mapahuirá) comida callejera	21	607	S/N

Biosample	País	Cepa	Origen	Descripción	Código de Identificación	ST	ST Complex
SAMN21210937	España	STEC 10531-2	Alimentario	Carne de cerdo (congelada)	22	1112	S/N
SAMN11246461	Italia	EC845	Alimentario	Espinaca	23	34	ST10 Cplx
SAMN28681468	Ecuador	Ch5_1a1	Alimentario	Ensalada <i>Lupinus mutabilis</i>	24	10	ST10 Cplx
SAMN10219878	India	NAEC1115	Animal	Ciego de pollo	25	10	ST10 Cplx
SAMEA3678975	Mexico	10270	Humana	Abcesos hepatico	26	617	ST10 Cplx
SAMN11246617	Italia	EC850	Humana	Heces	27	69	ST69 Cplx
SAMN28681471	Ecuador	R3cX	Alimentario	Rábano (<i>Raphanus raphanistrum subsp. Sativus</i>)	28	6506	S/N
SAMN03256246	Malasia	EC14	Ambiental	Agua	29	9074	S/N
SAMN28681469	Ecuador	Jc5_2C2	Alimentario	Jugo de caña (comida callejera)	30	13273	S/N
SAMN10219877	India	NAEC1042	Alimentario	Carne de pollo cruda	31	117	S/N
SAMN21210958	Australia	EPEC 1668-4a	Alimentario	Carne de res (congelada)	32	6323	ST28 Cplx

Biosample	País	Cepa	Origen	Descripción	Código de Identificación	ST	ST Complex
SAMN10839620	Suecia	B1483-PB_2012	Humana	Sangre	33	131	ST131 Cplx
SAMN28681465	Ecuador	L6bY	Alimentario	Lechuga (<i>Lactuca sativa</i>)	34	1286	ST10 Cplx
SAMN21210949	Alemania	STEC 826a	Alimentario	Carne de cerdo (congelada)	35	993	ST10 Cplx
SAMN16293344	Egipto	23ST	Humana	Heces	36	226	ST226 Cplx
SAMN28681462	Ecuador	M1_1	Alimentario	Guiso de Vaca	37	10	ST10 Cplx
SAMN21210940	Australia	STEC 1309b	Alimentario	Carne de res (congelada)	38	10	ST10 Cplx
SAMN28681463	Ecuador	C2_1c1	Alimentario	Huevos de cordorniz	39	10	ST10 Cplx
SAMN28681472	Ecuador	E2_2b	Alimentario	Merengue dulce de huevo	40	10	ST10 Cplx
SAMN10585802	Chile	UC-457	Humano	Orina	41	4204	ST10 Cplx
SAMN28681470	Ecuador	CM3aX	Alimentario	Col Roja (<i>Brassica oleracea</i>)	42	10	ST10 Cplx
SAMN16293350	Egipto	21M	Alimentario	Carne molida	43	48	ST10 Cplx

Biosample	País	Cepa	Origen	Descripción	Código de Identificación	ST	ST Complex
SAMN28681474	Ecuador	R6c2	Alimentario	Rábano (<i>Raphanus raphanistrum</i> subsp. <i>Sativus</i>)	44	48	ST10 Cplx
SAMN20560762	Ecuador	UTA144	Ambiental	Sedimentos de canal de riego	45	5876	S/N
SAMN21210955	Francia	EPEC 1095a	Alimentario	Carne de cerdo	46	137	ST32 Cplx
SAMN01911278	Japón	SAKAI (EHEC) (O157:H7)	Humana	Heces	47	11	ST11 Cplx
SAMN07224767	Francia	MC2 (O157:H7)	Animal	Heces (<i>Bos tauro</i>)	48	11	ST11 Cplx
SAMN03736827	Países Bajos	STEC 343 (O157:H7)	Humana	Heces	49	11	ST11 Cplx
SAMN03462359	USA	C1-057 (O157:H7)	Animal	Heces rectales (ganado)	50	11	ST11 Cplx

