



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ciencias de la Vida

**“SECUENCIACIÓN EN NANOPOROS PARA LA IDENTIFICACIÓN
MULTIPLEX DE BACTERIAS PATÓGENAS TRANSMITIDAS POR
ALIMENTOS”**

**Trabajo de titulación Previo a la obtención del Título de Magíster en Biociencias
Aplicadas con Mención en BIODESCUBRIMIENTO**

Presentado por: Byron Jesús Díaz Cárdenas

Guayaquil – Ecuador, 2024

AGRADECIMIENTO

Agradezco a mi tutor Juan Manuel Cevallos, PhD. Agradezco a mis profesores y evaluadores,

A la ESPOL, al CIBE, y la RED VLIR NETWORK ECUADOR por su contribución al proyecto de investigación

Agradezco a todas las personas especiales que me ayudaron, los llevo en mi corazón.

DEDICATORIA

*Dedico mi trabajo a Dios por ser
siempre mi fuerza y refugio,*

*A mi madre y mis hermanas, por ser
mi inspiración.*

A mi abuelita Delia, y

A ti, noble colibrí,

gracias.

TRIBUNAL DE GRADUACIÓN

Ana Tirapé Bajaña, PhD.

Presidenta del Tribunal de graduación

Juan Manuel Cevallos Cevallos, PhD.

Tutor

Eduardo Chica, PhD.

Cotutor

Luis Galarza, PhD.

Evaluator

Andrea Orellana, PhD.

Evaluator

DECLARACIÓN EXPRESA

"La responsabilidad del contenido de este trabajo de titulación, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la Escuela Superior Politécnica del Litoral".

Ing. Byron Jesús Díaz Cárdenas

RESUMEN

En Ecuador, la inocuidad alimentaria se ve amenazada por la presencia de bacterias patógenas en alimentos. El secuenciador MinION de Oxford Nanopore Technologies se perfila como una herramienta accesible para la detección de patógenos transmitidos por alimentos. Esta investigación evalúa su potencial para la detección simultánea de *Listeria monocytogenes*, *Salmonella enterica* y *Vibrio cholerae* en 12 alimentos de alto riesgo incluyendo bolones, ceviches, carne de res, encebollados, ensalada de frutas, empanadas de queso, frutas picadas, mote con choncho, pollo, quesos, salchipapas y salsas picantes.

Se desarrolló un protocolo de extracción de ADN, un método de preparación de genotecas 16S y un flujo bioinformático para la detección de bacterias usando la secuenciación de nanoporos. Se evaluó la sensibilidad del equipo minION ante bajas cantidades de patógenos (10^{-9} , UFC/ml o UFC/g) y se compararon los perfiles microbianos con la secuenciación en Illumina. El protocolo de extracción a base de hidróxido de sodio 50 mM NaOH, obtuvo rendimientos aceptables para la secuenciación de Nanopore. La preparación de genotecas 16S rARN se logró utilizando los barcodes del kit de Nanopore y ajustando las concentraciones de los amplicones equimolarmente. El flujo bioinformático NCBI+ fue una alternativa al flujo EPI2ME oficial de Nanopore.

El secuenciador Nanopore detectó *Salmonella enterica* y *Vibrio cholerae* con una abundancia relativa $< 1\%$ en la menor dilución de bacterias inoculadas (10^{-9} UFC). No se detectó *Listeria monocytogenes* en ninguna plataforma.

Se identificaron potenciales bacterias patógenas de interés económico en ambas plataformas pertenecientes a los géneros de *Bacillus*, *Campylobacter*, *Clostridium*, *Cronobacter*, *Escherichia*, *Listeria*, *Salmonella*, *Shigella*, *Staphylococcus*, *Vibrio* y *Yersinia*. La secuenciación de Nanopore obtuvo un 69% de clasificación a nivel de género, mientras que Illumina alcanzó un 92%. No se encontraron diferencias significativas en la diversidad alfa y beta entre ambas plataformas, pero si se halló una correlación del 70% entre los taxones hallados.

El secuenciador MinION es una alternativa a las plataformas de secuenciación NGS por su versatilidad, bajo costo de inversión inicial y rápida obtención de información de perfiles bacterianos en muestras de alimentos. Se requieren estudios adicionales para optimizar la detección de *Listeria monocytogenes* y aumentar la cantidad de lecturas válidas con clasificación taxonómica a nivel de género y especies.

ABSTRACT

Food safety in Ecuador is threatened by the presence of pathogenic bacteria in food. The MinION sequencer from Oxford Nanopore Technologies is emerging as an accessible tool for the detection of foodborne pathogens. This research evaluates its potential for the simultaneous detection of *Listeria monocytogenes*, *Salmonella enterica*, and *Vibrio cholerae* in 12 high-risk foods including bolones, ceviches, beef, encebollados, fruit salad, cheese empanadas, chopped fruits, mote con chanco, chicken, cheeses, salchipapas and spicy sauces.

A DNA extraction protocol, a 16S library preparation method, and a bioinformatics workflow were developed for the detection of bacteria using nanopore sequencing. The sensitivity of the minION equipment to low amounts of pathogens (10^{-9} , CFU/ml or CFU/g) was evaluated and the microbial profiles were compared with Illumina sequencing. The extraction protocol based on 50 mM NaOH sodium hydroxide, obtained acceptable yields for Nanopore sequencing. The preparation of 16S rRNA libraries was achieved using the barcodes of the Nanopore kit and adjusting the concentrations of the amplicons equimolarly. The NCBI+ bioinformatics workflow was an alternative to the official Nanopore EPI2ME workflow.

The Nanopore sequencer detected *Salmonella enterica* and *Vibrio cholerae* with a relative abundance $< 1\%$ in the lowest dilution of bacteria inoculated (10^{-9} CFU). *Listeria monocytogenes* was not detected on any platform.

Potential pathogenic bacteria of economic interest were identified on both platforms belonging to the genera *Bacillus*, *Campylobacter*, *Clostridium*, *Cronobacter*, *Escherichia*, *Listeria*, *Salmonella*, *Shigella*, *Staphylococcus*, *Vibrio* and *Yersinia*. Nanopore sequencing obtained 69% classification at the genus level, while Illumina reached 92%. No significant differences were found in alpha and beta diversity between both platforms, but a 70% correlation was found between the taxa found.

The MinION sequencer is an alternative to NGS sequencing platforms due to its versatility, low initial investment cost and rapid information acquisition of bacterial profiles in food samples. Further studies are needed to optimize the detection of *Listeria monocytogenes* and increase the number of valid reads with taxonomic classification at the genus and species level.

ÍNDICE GENERAL

AGRADECIMIENTO	II
DEDICATORIA	III
TRIBUNAL DE GRADUACIÓN	IV
DECLARACIÓN EXPRESA	V
RESUMEN.....	VI
ABSTRACT.....	VII
ÍNDICE GENERAL	VIII
ABREVIATURAS.....	X
SIMBOLOGÍA.....	XII
ÍNDICE DE FIGURAS.....	XIII
ÍNDICE DE TABLAS	XVI
CAPÍTULO 1.....	1
INTRODUCCIÓN Y OBJETIVOS	1
1.1 Antecedentes y justificación	1
1.2. Objetivo General.....	3
1.3. Objetivos Específicos	3
CAPÍTULO 2.....	4
MARCO TEÓRICO	4
2.1. Vigilancia de la inocuidad alimentaria	4
2.2. Impacto de los patógenos transmitidos por alimentos	4
2.3. Métodos para la detección de patógenos transmitidos por los alimentos.....	5
2.4 Secuenciación en nanoporos para la detección de bacterias transmitidas por alimentos.....	6
CAPÍTULO 3.....	8
METODOLOGÍA	8
3.1 Muestreo de alimentos.....	9
3.3 Inoculación artificial de alimentos.....	10

3.4. Extracción de ADN.....	11
3.5. Preparación de librerías y secuenciación en MinION.	12
3.6. Secuenciación 16S en Illumina.....	14
3.7. Análisis de datos y bioinformática	14
CAPÍTULO 4	19
RESULTADOS Y DISCUSIÓN.....	19
4.1 Extracción de ADN de las muestras de alimentos.....	19
4.2 Librerías genómicas 16S rRNA para el multiplexeado e identificación de <i>S. enterica</i> , <i>L. monocytogenes</i> y <i>V. cholerae</i>	23
4.3. Evaluación de flujos bioinformáticos para la identificación de <i>S. enterica</i> , <i>L. monocytogenes</i> y <i>V. cholerae</i>	25
4.4. Comparativa entre la secuenciación de Nanopore e Illumina, detección y diversidad de bacterias patógenas transmitidas por alimentos.	36
CAPÍTULO 5	46
CONCLUSIONES Y RECOMENDACIONES	46
BIBLIOGRAFÍA	48
ANEXOS	56
Figuras Suplementarias	56
Tablas suplementarias	59

ABREVIATURAS

ARN	Ácido RiboNucleico
ADN	Ácido Desoxirribonucleico
AFNOR	Association Française de Normalization
AGROCA LIDAD	Agencia de Regulación y Control Fito y Zoonosanitario
AOAC	Association of Analytical Communities
BPA	Buenas Prácticas Agropecuarias
BPTA	Bacterias patógenas transmitidas por alimentos
CEDIA	Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia
CMD	Símbolo del Sistema (Command Prompt, en inglés)
ETA	Enfermedades transmitidas por alimentos
FAO	Organización de las Naciones Unidas para la Alimentación y la Agricultura (Food and Agriculture Organization, en inglés)
FDA	Administración de Alimentos y Medicamentos (Food and Drug Administration, en inglés)
HACCP	Análisis de Peligros y Puntos Críticos de Control (Hazard Analysis and Critical Control Points, en inglés).
MAG	Ministerio de Agricultura y Ganadería
NCBI	Centro Nacional para la Información Biotecnológica (National Center for Biotechnology Information, en inglés)
NGS	Secuenciación de siguiente generación (Next generation sequencing, en inglés)
OMS	Organización Mundial de la Salud (World Health Organization, en inglés)
ONT	Oxford Nanopore Technologies
OTU	Unidades Taxonómicas Operativas

PBS	Solución Salina Tamponada con Fosfato (Phosphate-Buffered Saline, en inglés)
PCR	Reacción en Cadena de la Polimerasa (Polymerase Chain Reaction, en inglés)
SSD	Disco de Estado Sólido (Solid State Disk, en inglés)
TAE	Tris-Acetato-EDTA
UFC	Unidad Formadora de Colonias
USB	Bus Universal en Serie (Universal Serial Bus, en inglés)
WGS	Secuenciación del Genoma Completo (Whole Genome Sequencing, en inglés)

SIMBOLOGÍA

Símbolo	Significado
%	Porcentaje
°C	Grado centígrado
μl	microlitro
g	gramo
GB	gigabytes
h	hora
kg	kilogramo
l	litro
mg	miligramo
min	minutos
ml	mililitro
mM	milimolar
ng	nanogramos
p/v	peso/volumen
pH	Potencial de hidrógeno
seg	segundos
UFC/ml	Unidades formadoras de colonias/ mililitro

ÍNDICE DE FIGURAS

Figura 1. Esquema general del diseño de estudio	8
Figura 2. Comparación de los efectos del método de extracción sobre la calidad del ADN obtenido de las muestras de frutas picadas. El eje de las ordenadas representa la medida de calidad obtenida para cada método de extracción. El eje de las abscisas Muestra los tres métodos de extracción que se están comparando: el método de Chen et al. (2019) (Método A), Purelink (Método B) y el método de Dilhari et al. (2017) (Método C). Cada caja representa la distribución de los datos de calidad para cada método. La línea en el medio de la caja es la mediana. La parte superior e inferior de la caja representan el tercer y primer cuartil respectivamente, lo que da el rango intercuartil (IQR). Las letras ‘a’ y ‘b’ indican grupos estadísticamente diferentes. Los puntos representan valores atípicos.....	20
Figura 3. ADN genómico de 12 muestras de alimentos utilizando el protocolo de Dilhari et al. (2017) modificado. Las líneas rojas indican el tamaño de los fragmentos del marcador de 1500 pares de bases (PB). Las iniciales indican B: bolones, C: ceviches, CM: carne de res, E: encebollados, EF: ensalada de frutas, EM: empanadas de queso, F: frutas picadas, MT: mote con chanco, P: pollo, Q: quesos, SP: salchipapas, S: salsas picantes.	22
Figura 4. Representación gráfica del conteo de lecturas de ADN en 2 horas de secuenciación y el puntaje de calidad asignado, considerando 7 como el valor mínimo para ser considerado una lectura válida. La moda de las lecturas obtenidas fue de 15.55.	26
Figura 5. Distribución de las lecturas de secuenciación de Nanopore en los diferentes flujos bioinformáticos evaluados. En el eje de las ordenadas se representa las lecturas halladas y en el de las abscisas los flujos bioinformáticos. El literal A indica el total de lecturas que se lograron clasificar, el B indica las lecturas asignadas a nivel de género, el literal C muestra las lecturas asignadas a nivel de especie. El literal D, E y F indican las lecturas asignadas a Salmonella, Listeria y Vibrio, respectivamente. Los flujos son P1: EPI2ME, P2: EPI2ME_LABS, P3: Spaguetty_modificado, P4: NCBI+ y P5: Situseq. Desde A, B y C, se observan las clasificaciones de lecturas clasificadas y asignadas a un taxón. D, E y F son las lecturas asignadas por los flujos a las bacterias patógenas.....	29

Figura 6. Abundancia relativa de las bacterias patógenas inoculadas detectadas según el flujo bioinformático utilizado. Las legendas del gráfico se describen P1: EPI2ME, P2: EPI2ME_LABS, P3: Spaguetty_modificado, P4: NCBI+ y P5: Situseq. 30

Figura 7. Abundancia relativa de las especies patógenas transmitidas por alimentos detectadas en las lecturas de secuenciación de las muestras de alimentos inoculadas artificialmente según el flujo bioinformático. Los códigos por tipos de alimentos son: Queso (Q), Control positivo (Ctrl_pos), Frutas (F), salsas (S) y control negativo (Ctrl_neg). De los flujos bioinformáticos, EPI2ME (P1), EPI2ME_LABS (P2), Spaguetty_modificado (P3), NCBI+ (P4) y Situseq (P5). 32

Figura 9. Diversidad de Shannon entre los 5 flujos bioinformáticos (A) y alimentos evaluados (B). Las letras sobre cada conjunto de datos indican la diferencia significativa del grupo de TukeyHSD. Los cuadros indican el rango intercuartílico de los datos. Las barras de error se extienden hasta el valor más extremo dentro de 1,5 veces el rango intercuartil y los puntos representan valores atípicos más allá de ese rango. Los códigos por tipos de alimentos son: Queso (Q), Control positivo (Ctrl_pos), Frutas (F), salsas (S) y control negativo (Ctrl_neg). 34

Figura 9. El análisis de coordenadas principales (PcoA) revela la agrupación de los flujos P1, P2 y P4 utilizando la distancia de Bray Curtis de muestras de alimentos por ciudad de muestreo. Mientras que se observa la diferencia del microbioma en los flujos P3 y P5. Los códigos por tipos de alimentos son: Queso (Q), Control positivo (Ctrl_pos), Frutas (F), salsas (S) y control negativo (Ctrl_neg). De los flujos bioinformáticos, EPI2ME (P1), EPI2ME_LABS (P2), Spaguetty_modificado (P3), NCBI+ (P4) y Situseq (P5). 35

Figura 10. Prueba de abundancia diferencia utilizando el método Random forest y el índice MeanDecreaseGini para la detección de las lecturas de los géneros más significativos que influyen en la AR del Flujo P3 en comparación con los demás Flujos. Los códigos de los flujos bioinformáticos, EPI2ME (P1), EPI2ME_LABS (P2), Spaguetty_modificado (P3), NCBI+ (P4) y Situseq (P5). 36

Figura 11. Alfa diversidad de las muestras de alimentos usando el índice de Shannon, se muestra la comparación de la diversidad hallada entre ambas tecnologías de secuenciación Illumina vs Nanopore. 37

Figura 12. Gráfico de correlaciones que muestran las similitudes de agrupamiento jerárquico (algoritmo de Ward) entre las muestras de alimentos secuenciadas con Illumina y Nanopore. La distancia entre grupos se midió utilizando el índice de Bray-Curtis. Las líneas de colores indican nodos compartidos entre ambos dendrogramas y

las reglas horizontales indican las alturas de agrupación. Los códigos de los alimentos se indican: Bolones (B), Ceviches (C), Carne de res (CM), Encebollados (E), Ensalada de frutas (EF), Empanadas de queso (EM), Frutas picadas (F), Mote con chanco (MT), Pollo (P), Quesos (Q), Salchipapas (SP) y Salsas picantes (S)..... 38

Figura 13. Abundancias relativas de los filos hallados en las muestras de alimentos según la tecnología de secuenciación Illumina y Nanopore. Las abundancias de los filos menores al 1% se agruparon en la categoría “Otros”. Los códigos de los alimentos se indican: Bolones (B), Ceviches (C), Carne de res (CM), Encebollados (E), Ensalada de frutas (EF), Empanadas de queso (EM), Frutas picadas (F), Mote con chanco (MT), Pollo (P), Quesos (Q), Salchipapas (SP) y Salsas picantes (S)..... 40

Figura 14. Correlación entre la abundancia de géneros secuenciados con la tecnología Nanopore (eje y) e Illumina (eje x) (Pearson $r = 0,767$). Los géneros abundantes están coloreados, "Unclassified at genus level" representa las secuencias no identificadas a nivel de género y "Other" representa los filos menos abundantes. La línea discontinua negra muestra la relación lineal entre las abundancias de Nanopore e Illumina, y la línea roja muestra una proporción de 1:1. Los ejes se han transformado en raíz cuadrada. 42

Figura 15. La relación entre la abundancia de Illumina y la abundancia de nanoporos de los géneros más abundantes seleccionados. Las proporciones de muestras individuales se superponen en los diagramas de caja. La línea roja muestra una proporción de 1:1. 43

Figura 16. Gráfico NMDS con los conjuntos de datos combinados de Nanopore e Illumina. La disimilitud de Bray-Curtis en los géneros se muestran con puntos coloreados y dimensionados según las plataformas de secuenciación..... 44

Figura 18. Abundancias de las potenciales bacterias patógenas transmitidas por alimentos en ambas tecnologías de secuenciación, Illumina y Nanopore. Los códigos de los alimentos se indican: Bolones (B), Ceviches (C), Carne de res (CM), Encebollados (E), Ensalada de frutas (EF), Empanadas de queso (EM), Frutas picadas (F), Mote con chanco (MT), Pollo (P), Quesos (Q), Salchipapas (SP) y Salsas picantes (S)..... 45

ÍNDICE DE TABLAS

Tabla I. Programa de amplificación PCR con el ciclaje modificado para el kit 16S (SQK-16S024).	13
Tabla II. Flujos bioinformáticos utilizados para el análisis y detección de patógenos alimentarios.	15
Tabla III. Géneros y especies de las bacterias patógenas transmitidas por alimentos de mayor importancia económica analizadas en el presente estudio.	16
Tabla IV. Métodos de extracción evaluados, tiempo de ejecución del protocolo con una muestra de frutas y 3 réplicas, se muestra la concentración de ADN en ng/μl y la calidad en el rango 260/280. Los valores se expresaron como la media más la desviación estándar.	19
Tabla V. Extracción de ADN de 12 muestras de alimentos con el protocolo de Dilhari et al. (2017) modificados, se muestra la cantidad y la calidad medidos en Qubit y Nanodrop, respectivamente.	23
Tabla VI. Cantidad de ADN obtenida de la extracción de las muestras inoculadas luego de 24 horas de incubación en caldo de cultivo. Se anotó la cantidad de amplicones obtenidos de cada muestra usando 35 ciclos de amplificación y la cantidad de lecturas para <i>L. monocytogenes</i> , <i>S. enterica</i> y <i>V. cholerae</i> . Se expresan la media de cada muestra ± DE.	25
Tabla VII. Evaluación de los 5 flujos bioinformáticos de los programas para la detección de bacterias patógenas transmitidas por alimentos, se muestran las lecturas obtenidas, el número de géneros y especies identificados.	28
Tabla VIII. Comparación del número de secuencias de ADN obtenidas según la secuenciación de Nanopore e Illumina, se analizó el número de lecturas que se clasificaron a nivel de género (%) según los flujos de los programas EPI2ME (Nanopore) y 16S Metagenomics (Illumina).	39

CAPÍTULO 1

INTRODUCCIÓN Y OBJETIVOS

1.1 Antecedentes y justificación

Hoy en día las enfermedades transmitidas por alimentos (ETA) constituyen una grave amenaza para la seguridad alimentaria mundial, registrándose 600 millones de casos de ETA y confirmándose más de 250 patologías con alrededor de 420,000 muertes cada año ocasionadas por el consumo de alimentos contaminados (Mangal et al., 2016; World Health Organization, 2015). Cerca del 90% de los casos reportados y muertes causadas por ETA se han atribuido a bacterias con alta virulencia y patogenicidad, siendo las más comunes *Escherichia coli*, *Salmonella spp.*, *Listeria monocytogenes* y *Vibrio cholerae* (Baraketi et al., 2018; Giacometti & Josic, 2013; Palomino-Camargo & González-Muñoz, 2014; U.S. Food & Drug Administration, 2018).

En países en vías de desarrollo, las ETA son más comunes y han llegado a causar pérdidas anuales por 110,000 billones de dólares anuales en Salud Pública (The World Bank, 2018). En Ecuador se reportó 24,082 casos durante el 2018, y hasta mediados del año 2019 se registró 11,411 casos más, lo que significó millonarias pérdidas para el país (Fung et al., 2018; Ministerio de Salud Pública, 2019).

La Organización Mundial de la Salud (World Health Organization, 2015) señala que las alarmantes estadísticas relacionadas a ETA resaltan la importancia de detectar e identificar los agentes patógenos en los alimentos antes de que los productos alimenticios se distribuyan y vendan a los consumidores, sugiriendo la creación y utilización de métodos rápidos, económicos y confiables que faciliten la actuación de las autoridades en el caso de brotes epidemiológicos.

Los métodos microbiológicos convencionales basados en el cultivo y el recuento de colonias aún son utilizados en Ecuador y varios países en Latinoamérica como norma estándar para la detección de bacterias patógenas transmitidas por alimentos (BPTA), sin embargo, estas no se ajustan a los requerimientos actuales de tiempo, eficiencia y confiabilidad (Vidic et al., 2019).

En la última década, el campo de la sanidad alimentaria se ha enfocado en el desarrollo de tecnologías con la información presente en los ácidos nucleicos para la detección e identificación de BPTA con una alta fidelidad (Van Hoorde & Butler, 2018a). De esta manera, se ha implementado varias herramientas ómicas a través de las nuevas tecnologías de secuenciación, siendo la de Illumina, la más común. Sin embargo, sus altos costos de análisis, tiempo y la dificultad para determinar genomas a niveles

taxonómicos más específicos, han impedido su uso rutinario (Mitsuhashi et al., 2017a; Yang et al., 2019a). La tercera generación de secuenciadores, basados en la secuenciación de una sola molécula de lectura larga (hasta 70k pb), han superado algunas de estas deficiencias (Goldstein et al., 2019a; Kovac, 2019a; Taylor et al., 2019a).

El secuenciador USB MinION de *Oxford Nanopore Technologies* (ONT), un dispositivo portátil de tercera generación, ha empezado a llamar la atención de la comunidad científica en el campo alimentario puesto que varios estudios microbiológicos lo han instaurado como un método prometedor para la detección rutinaria de BPTA, ya que ofrece un enfoque metagenómico de las muestras en estudio, proporciona análisis de los genomas completos en tiempo real y es accesible económicamente (Loit, Kaire ; Adamson, Kalev ; Bahram, Mohammad; Puusepp, Rasmus; Anslan, Sten; Kiiker, Riinu; Drenkhan, Rein; Tedersoo, 2019a; Runtuwene et al., 2019a; Tyler et al., 2018a; Yang et al., 2019a). Hoy en día, las tasas de error y los perfiles de sesgo del MinION van mejorando y actualizándose, desde su flujo de celdas hasta los diferentes programas informáticos (Kuchly et al., 2022; Taylor et al., 2019a; Tyler et al., 2018b).

En Ecuador, varias instituciones gubernamentales relacionadas a la salud pública han planteado mejorar y garantizar la inocuidad alimentaria mediante la caracterización de los contaminantes microbianos en los alimentos, buscando la transición de los métodos tradicionales basados en cultivos a pruebas de diagnóstico que utilicen plataformas modernas de secuenciación de ADN, perfilando al MinION (ONT) como una herramienta accesible a implementarse para dichos fines (Carleton et al., 2019; CEDIA, 2019; Loman et al., 2015a; World Health Organization, 2018).

En este contexto, la presente investigación tuvo como objetivo evaluar la secuenciación metagenética en nanoporos del minION. Este método se presenta como una alternativa eficiente para la detección simultánea de *Listeria monocytogenes*, *Salmonella enterica* y *Vibrio cholerae* en alimentos de alto riesgo y consumo popular. De esta manera, promover el uso del MinION como secuenciador portátil y de bajo costo de instalación en los programas de sanidad alimentaria, estableciendo alternativas de prevención a brotes epidemiológicos de bacterias patógenas transmitidas por alimentos.

1.2. Objetivo General

Evaluar la secuenciación metagenética en nanoporos del dispositivo MinION (*Oxford Nanopore Technologies*) como método rápido, exacto y preciso para la detección simultánea de los patógenos de transmisión por alimentos *Salmonella enterica*, *Listeria monocytogenes* y *Vibrio cholerae*.

1.3. Objetivos Específicos

1. Determinar un protocolo de extracción de ADN bacteriano con la cantidad y calidad para la secuenciación en nanoporos, a partir de muestras de alimentos.
2. Establecer un método de preparación de librerías genómicas 16S rRNA para el multiplexeado e identificación de *S. enterica*, *L. monocytogenes* y *V. cholerae* en secuenciación por nanoporos.
3. Estructurar un flujo bioinformático con los datos de secuenciación obtenidos Nanopore para la identificación de *S. enterica*, *L. monocytogenes* y *V. cholerae*.
4. Comparar las metodologías de secuenciación de Illumina frente a la de nanoporos para determinar las diferencias en detección y diversidad de las bacterias en estudio.

Al determinar la capacidad de secuenciación metagenética del secuenciador minION en la detección múltiple de bacterias patógenas en alimentos se busca promover el uso de este equipo en programas de vigilancia sanitaria en Ecuador.

CAPÍTULO 2

MARCO TEÓRICO

2.1. Vigilancia de la inocuidad alimentaria

Un sistema de vigilancia ante patógenos transmitidos por los alimentos tiene como función la recopilación, el análisis y la interpretación sistemática de datos esenciales para la planificación, implementación y evaluación de prácticas de salud pública, y la difusión oportuna de esta información para la acción entidades gubernamentales responsables (World Health Organization, 2008). Estos programas tienen como objetivos principales, identificar, controlar y prevenir los brotes de enfermedades transmitidas por los alimentos, determinar las causas de las enfermedades transmitidas por los alimentos y monitorear las tendencias en la aparición de enfermedades transmitidas por los alimentos (Centers for Disease Control and Prevention, 2022). Como lo menciona Rantsiou et al. (2018) en su publicación, la información sobre la epidemiología de los patógenos transmitidos por los alimentos es muy valiosa para la evaluación de riesgos y el desarrollo e implementación de un plan HACCP.

En Ecuador, el Ministerio de Agricultura y Ganadería (MAG) a través la Agencia de Regulación y Control Fito y Zoosanitario (Agrocalidad) posee las competencias para vigilar y controlar la inocuidad en la producción primaria de los alimentos y así evitar las ETA en el país. Agrocalidad lleva a cabo distintos programas e iniciativas tendientes a fortalecer la inocuidad, a través de proyectos en la producción en campo, certificación de centros de faenamiento, vigilancia y control de residuos de contaminantes en los alimentos y la resistencia antimicrobiana. De esta manera, se ha implementado un proyecto de certificación en Buenas Prácticas Agropecuarias (BPA), que tiene como ejes principales, la consecución de un alimento seguro, la protección de la salud del trabajador del campo, el cuidado del ambiente y el bienestar animal. Además, esta institución participa activamente como líder del Comité Coordinador FAO/OMS para América Latina y El Caribe, del Codex Alimentarius (AGROCALIDAD, 2020).

2.2. Impacto de los patógenos transmitidos por alimentos

Entre los patógenos identificados como causantes de enfermedades transmitidas por alimentos, *Salmonella*, *Campylobacter*, *Staphylococcus aureus*, *Listeria monocytogenes*, *Clostridium perfringens* y *Escherichia coli* O157:H7 han sido incriminados en la gran mayoría de enfermedades, hospitalizaciones y muertes (Foddai & Grant, 2020). De hecho, *Salmonella* spp., *L. monocytogenes*, *E. coli* 157:H7 y *S.*

aureus encabezan la lista de la mayor cantidad de brotes, casos y muertes alrededor del mundo (FDA, 2020; U.S. Food & Drug Administration, 2018).

Además, aunque la Listeriosis no es frecuente, porque *L. monocytogenes* tiene una menor prevalencia en los alimentos en comparación con *Campylobacter* y *Salmonella*, es una ETA con graves efectos en la salud, con una tasa de mortalidad asociada que alcanza el 13% (European Food Safety Authority, 2021).

Algunas de las ETA son causadas por bacterias productoras de toxinas, y se ha encontrado evidencia de su aumento en los últimos años, considerándolas como un peligro emergente (European Food Safety Authority, 2021). Un factor importante en este grupo de microorganismos es que sus genes que codifican toxinas pueden transferirse entre bacterias. Los brotes de origen alimentario causados por bacterias toxigénicas a menudo comparten una sintomatología común, lo que dificulta las investigaciones, por esto, es excepcionalmente importante detectarlas porque las toxinas pueden permanecer dentro o sobre el alimento y ser ingeridas mientras la bacteria se elimina o ya no está presente. Además, se carece de métodos adecuados para la detección de toxinas bacterianas (Vidic et al., 2019).

2.3. Métodos para la detección de patógenos transmitidos por los alimentos.

La inocuidad microbiológica de los alimentos ha sido monitoreada tradicionalmente a través de métodos fenotípicos que se basan en medios específicos de cultivo para la detección, caracterización e identificación de patógenos transmitidos por los alimentos. Estos microorganismos, se han estudiado principalmente a nivel de especie y fuera del contexto de las comunidades microbianas en las que residen. La regulación actual de niveles cero en la detección en alimentos, para ciertas especies de patógenos refleja tales nociones (Kovac, 2019). Estas regulaciones suelen propiciar un mayor desperdicio de alimentos y costos para la industria alimentaria y los consumidores, sin alcanzar el resultado de salud pública buscado. Esto, debido a que los subtipos de especies de patógenos varían en su capacidad para causar enfermedades en diferentes huéspedes (Kovac et al., 2017).

Para superar las limitaciones de los métodos convencionales, se han desarrollado varios métodos rápidos que están comercialmente disponibles en la industria alimentaria, que además, deben ser validados por una organización reconocida en el área de bioseguridad alimentaria como por ejemplo la Association Française de Normalization (AFNOR) en la Unión Europea o la Association of Analytical Communities (AOAC International) en los Estados Unidos, la mayoría de los kits de detección están validados según su sensibilidad y especificidad (Baraketi et al., 2018).

Idealmente para aplicaciones industriales, los métodos rápidos deben caracterizarse por su especificidad, alta sensibilidad y rápido rendimiento. Hoy en día, los métodos rápidos actuales pueden detectar patógenos en alimentos crudos y procesados en cantidades bajas para evitar el riesgo de infección, que son más eficientes en tiempo, ahorran trabajo y evitan errores humanos. Sin embargo, la sensibilidad y la especificidad aún deben mejorarse para analizar muestras de alimentos sin necesidad de enriquecerlas previamente antes del análisis (Vidic et al., 2019; Wei & Zhao, 2021).

Los métodos de detección rápida se pueden clasificar en biosensores, métodos inmunológicos y métodos basados en ácidos nucleicos. En los últimos años, las plataformas de secuenciación de próxima generación (*next generation sequencing* [NGS]), como Illumina, han surgido como una nueva estrategia tecnológica en el campo de la seguridad alimentaria. Sin embargo, sus aplicaciones son muy limitadas, puesto que estas plataformas generan lecturas de secuencia en paralelo y no en serie, siendo el análisis de datos una carga significativa al tiempo de respuesta total, además de su costo excesivo en equipos (Mitsubishi et al., 2017b; Yang et al., 2019b).

2.4 Secuenciación en nanoporos para la detección de bacterias transmitidas por alimentos.

En microbiología de alimentos, NGS se usa predominantemente para determinar la secuencia del genoma completo de un microorganismo aislado y cultivado que comúnmente se conoce como "*whole genome sequencing*" (WGS). En metagenómica, NGS se aplica a una muestra biológica generando secuencias de múltiples microorganismos (Jagadeesan et al., 2019; Kovac et al., 2017; Van Hoorde & Butler, 2018b).

En la actualidad, el WGS es un procedimiento de rutina en la investigación epidemiológica y la vigilancia de patógenos transmitidos por los alimentos. Sin embargo, la tecnología de secuenciación NGS de lectura corta (hasta 600 pb) presenta desafíos en la resolución de genomas en regiones repetitivas que conducen a ensamblajes *de novo* incompletos. La tercera generación de secuenciadores, basados en la secuenciación de una sola molécula de lectura larga (hasta 70 kpb), han superado algunas de las deficiencias (Goldstein et al., 2019b; Kovac, 2019b; Taylor et al., 2019b).

El secuenciador USB MinION de Oxford *Nanopore Technologies*, es un dispositivo de bolsillo de tercera generación, que se ha instaurado como un método prometedor para la detección de patógenos alimentarios basado en sus flujos de trabajo de secuenciación de ARN y ADN rápidos, rentables, portátiles y de alto rendimiento. Sus

aplicaciones son de gran alcance, e incluyen análisis de comunidades metagenómicas bacterianas en tiempo real, subtipado y ensamblaje de lectura larga para la secuenciación completa del genoma de organismos (Loit, Kaire ; Adamson, Kalev ; Bahram, Mohammad; Puusepp, Rasmus; Anslan, Sten; Kiiker, Riinu; Drenkhan, Rein; Tedersoo, 2019b; Runtuwene et al., 2019b; Tyler et al., 2018c; Yang et al., 2019b).

El MinION es capaz de captar directamente los fragmentos de ADN individuales nativos sin necesidad de un sistema de replicación enzimática. La secuenciación de nucleótidos se mide en los cambios de la corriente que una sola molécula de ADN ejerce al pasar a través de un nanoporo de proteínas, lo que resulta en lecturas de secuencias extremadamente largas (> 1 Mbp) manteniéndose la calidad de la secuencia (Loman et al., 2015b; Payne et al., 2019).

La identificación de nucleótidos (*basecalling*) es rápida, se produce en tiempo real y la información de la secuencia de ADN se puede recopilar a los pocos minutos de comenzar la ejecución del instrumento (Benítez-Páez et al., 2016; Parker et al., 2018). Sin embargo, la calidad de lectura de MinION es baja en comparación con las tecnologías de secuenciación de lectura corta (Jain et al., 2017).

Hoy en día, las tasas de error y los perfiles de sesgo del MinION van mejorando y actualizándose, desde su flujo de celdas hasta los diferentes programas informáticos. En varios estudios, se ha demostrado una mayor precisión en la secuenciación con nanoporos de genomas bacterianos llegando a alcanzar hasta un 99 % de precisión en la identificación, las secuencias obtenidas se compararon con secuencias obtenidas de Illumina y ajustando varios flujos bioinformáticos para mejorar la interpretación de los datos obtenidos. Algunos de estos microorganismos son patógenos transmitidos por alimentos, tales como *Cronobacter sakazakii*, *Escherichia coli*, *Salmonella enterica*, *Enterobacter cloacae*, y *Listeria monocytogenes* (Goldstein et al., 2019b; González-Escalona et al., 2019; Hyeon et al., 2017; Peritz et al., 2018; Quick et al., 2015; Schmidt et al., 2017; Senol Cali et al., 2018; Taylor et al., 2019b; Tyler et al., 2018c; Van Hoorde & Butler, 2018b; Yang et al., 2019b).

CAPÍTULO 3

METODOLOGÍA

En el contexto del actual proyecto de maestría para el diseño de estudio se llevó a cabo una investigación experimental, analítica y prospectiva con el objetivo de evaluar la capacidad de detección de perfiles bacterianos patógenos en muestras de alimentos utilizando el secuenciador de nanoporos MinION. Los aspectos clave del proyecto incluyeron un diseño de estudio con intervenciones controladas, análisis estadísticos para evaluar precisión y sensibilidad, y una perspectiva prospectiva que consideró factores como el tiempo de procesamiento y la escalabilidad de la tecnología. Además, se exploró la correlación entre los resultados obtenidos por MinION e Illumina para determinar su validez como alternativa al estándar actual. El flujo de trabajo de estudio se representa de manera gráfica en la Figura 1.

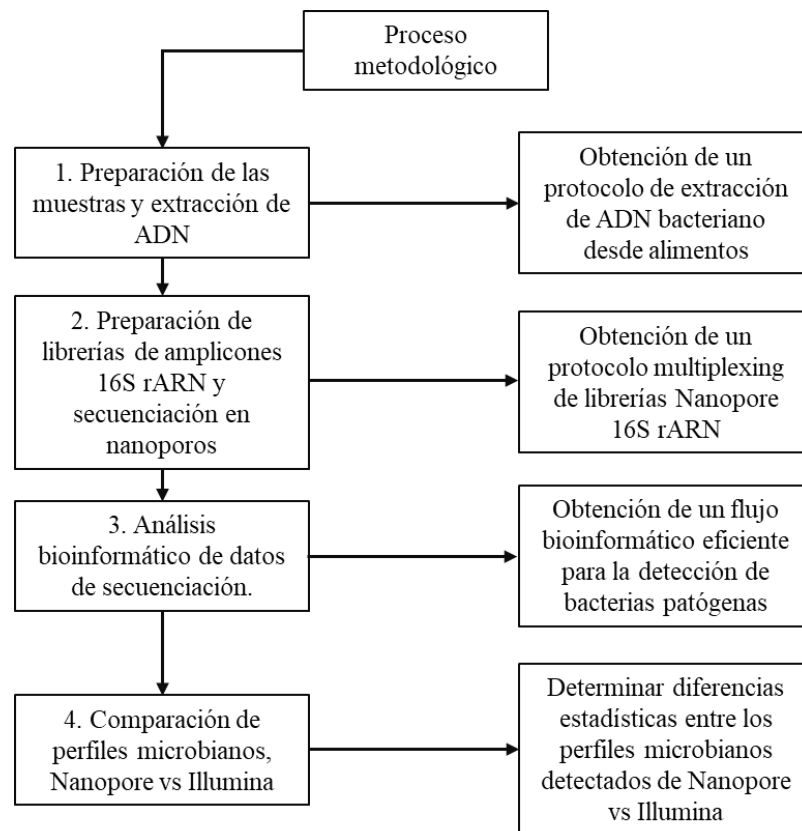


Figura 1. Esquema general del diseño de estudio

3.1 Muestreo de alimentos

Las muestras de alimentos que se utilizaron en esta investigación fueron escogidos del estudio de inocuidad alimentaria desarrollado por Cevallos-Cevallos, Morales, & Salazar-Llorente (2019) para mantener la continuidad de un proyecto macro de Seguridad alimentaria en Ecuador. En total se seleccionaron 12 alimentos, bolones (B), Ceviches (C), Carne de res cruda (CM), Encebollados (E), Ensalada de frutas (EF), Empanadas de queso (EM), Frutas picadas (F), Mote con chanco (MT), Pollo (P), Quesos (Q), Salchipapas (SP), Salsas picantes (S) cuyas características se indican en la Tabla suplementaria 1 y que forman parte de los alimentos considerados de alto riesgo para el consumidor y son los más populares para el consumo por su fácil acceso (U.S. Food & Drug Administration, 2020).

En la primera parte del estudio, se seleccionaron 6 alimentos (encebollados, frutas listas para comer, salsas a base de huevo, quesos frescos, pasteles de carne o pollo, y ensaladas crudas) para obtener ADN y validar el protocolo de extracción. Se recolectaron 3 muestras de cada alimento en sus envases originales y se transportaron a 4°C. El procesamiento se realizó dentro de las 3 horas posteriores a la compra siguiendo las recomendaciones del Codex Alimentarius (CXG 50-2004) (Organización de las Naciones Unidas para la Alimentación y la Agricultura & Organización Mundial de la Salud, 2004).

En la segunda parte del estudio, para el estudio comparativo de plataformas de secuenciación se tomaron 2 muestras de cada alimento. Las muestras se recolectaron en Cuenca, se pesaron y se procesaron dentro de las 24 horas en el laboratorio de la maestría de Biociencias.

3.2. Preparación de las muestras y aislamiento bacteriano

Luego de recolectar las muestras, en el laboratorio y bajo condiciones asépticas, para la obtención de la suspensión bacteriana de estos alimentos, se pesó 25 g de cada uno en una bolsa de muestreo estéril Whirl-Pak® (Nasco, EE. UU.) añadiendo 225 mL de caldo Mueller Hinton (Sigma-Aldrich, EE. UU.) y mezclando de manera uniforme por 1 minuto para obtener una unidad analítica de 375 g en muestras sólidas o 375 ml en muestras líquidas de cada uno de los alimentos según el protocolo señalado en el Manual Analítico Bacteriológico (U.S. Food & Drug Administration, 2003).

A continuación, para la concentración de las células bacterianas se utilizó el método de (Gu et al., 2018). Una de las limitaciones halladas fue la presencia excesiva de grasas residuales de las muestras, debido a esto, previo a la concentración final del

pellet mediante centrifugado, estas muestras fueron filtradas utilizando gasa estéril. El pellet obtenido se congeló a $-80\text{ }^{\circ}\text{C}$ hasta la extracción del ADN.

Adicionalmente, para garantizar la detección de los patógenos objetivos dentro de la muestra de alimentos y evitar resultados falsos negativos, se realizó el enriquecimiento de las muestras en proporción 1: 9 con el caldo de cultivo Mueller Hinton (MH; Sigma-Aldrich). Las muestras se incubaron por 24 horas, manteniendo $37\text{ }^{\circ}\text{C}$ como la temperatura estándar (Hyeon et al., 2017; Kobayashi et al., 2009; D. Liu, 2009; Y. Liu et al., 2019; Sun et al., 2014; Villamizar-Rodríguez et al., 2015; Vinayaka et al., 2019). Luego de la incubación, la concentración de las células bacterianas se realizó con 10 mL del sobrenadante de cada bolsa con alimento según el protocolo de Gu et al. (2018). Al final, el pellet se resuspendió en tampón TAE 1X y fue congelado a $-20\text{ }^{\circ}\text{C}$ (Tatsika et al., 2019).

3.3 Inoculación artificial de alimentos

Para evaluar la sensibilidad del secuenciador en la detección de las bacterias patógenas objetivo, se diseñó un experimento para la inoculación artificial de las bacterias en quesos, salsas y frutas picadas. Se buscó que las muestras adquiridas tuvieran el sello de registro sanitario ecuatoriano que garantiza la inocuidad de las muestras previo la inoculación de los patógenos. Se utilizaron dos cepas certificadas provenientes del *American Type Culture Collection* cuya identificación es ATCC 13932 y ATCC 14028, de las bacterias patógenas *L. monocytogenes* y *S. enterica* serovar Typhimurium respectivamente. Las cepas fueron enriquecidas por 24 horas. Posteriormente se preparó diluciones desde 10 hasta 10^{-9} UFC/ml de cada bacteria.

El experimento se llevó a cabo dividiendo cada muestra de alimento en tres réplicas, similar a la metodología de Hyeon et al. (2017) y Liu et al. (2019), 25 g de cada alimento se colocó en fundas plásticas estériles con sello ziploc, se agregó 2 ml de las diferentes diluciones de cada bacteria utilizando una micropipeta en cada grupo de alimentos. Para los controles positivo y negativo se utilizó una muestra compuesta por los 3 alimentos, el control positivo se utilizó con la dilución 10^3 UFC/ml, mientras que para el control negativo se utilizó muestras de alimentos estériles inoculadas con 2 ml de agua peptonada (Elizaquível & Aznar, 2008). Las muestras fueron enriquecidas por 24 horas a $37\text{ }^{\circ}\text{C}$ en agua peptonada o caldo Fraser, según el inóculo. Un esquema del experimento se muestra en la Figura suplementaria 1.

Para la identificación y conteo de los microorganismos viables presentes en las muestras, se utilizó el método de vertido en placa utilizando los agares XLD y Oxford Listeria agar. La dilución líquida de menor concentración 10^9 UFC/ml se utilizó para

la obtención del pellet bacteriano que se congeló el pellet a -80 °C hasta la extracción del ADN.

3.4. Extracción de ADN

Con el fin de optimizar recursos, el protocolo de extracción se probó en una de las muestras de alimentos, seleccionándose la ensalada de frutas. El pellet bacteriano obtenido fue sometido a 3 métodos de extracción de ADN. El método A, fue el método de ebullición del agua por Chen et al. (2019) que consiste en incubar el pellet en agua desionizada estéril a 100 °C durante 10 minutos. A continuación, para el método B se incluyó la extracción con el kit *PureLink® Genomic DNA Mini Kit* (PL; Invitrogen™, EE. UU.), probándose 3 tiempos de incubación durante la lisis enzimática en el rango de 30 min a 2 horas, indicado en el protocolo. Finalmente, el método C probó el protocolo de Dilhari et al. modificado (2017), con una extracción alcalina 50 mM NaOH y lisis en horno microondas a 700 *Watts* en 2 intervalos de 30 segundos.

Debido a la naturaleza compleja de las muestras de alimentos, volumen de muestras y para alcanzar la cantidad y calidad de ADN requeridos (> 10 ng y 1.80 en A260/280, respectivamente), se ajustó y modificó el protocolo de extracción de ADN de Dilhari et al. (2017).

Para la purificación del ADN extraído, se probaron tres variantes según el método, en el de Chen et al. (2019) se realizó una centrifugación a 12000 x g por 2 minutos, el kit *PureLink®* empleó una purificación con mini columnas de sílica y el protocolo de Dilhari et al. modificado (2017) utilizando una solución de precipitación de proteínas con acetato de potasio 5M, seguida de una purificación con isopropanol al 0.7X volumen. El ADN obtenido se almacenó a -20 °C para los siguientes procesos.

El ADN obtenido en la validación de protocolos, de cada una de las 12 muestras de los alimentos homogeneizados fueron procesados por triplicado y agrupadas en un solo tubo por alimento al momento de la elución, se cuantificó y evaluó la calidad del ADN para la posterior distribución en tubos individuales según la aplicación, Illumina y MinION.

La cantidad del ADN obtenido en cada método evaluado se determinó utilizando el kit *Qubit® dsDNA HS Assay* en un fluorómetro *Qubit® 4.0* (Invitrogen™), mientras que la calidad se evaluó en un espectrofotómetro *NanoDrop™ 2000/c* (Thermo Fisher Scientific). De modo complementario, para la valoración visual del ADN obtenido, se realizó electroforesis en geles de agarosa al 1 % (p/v) (Invitrogen™) teñidos con *SYBR*

Safe (Invitrogen™) y visualizados en un transiluminador de luz azul *Safe Imager™* 2.0 (Invitrogen™).

3.5. Preparación de librerías y secuenciación en MinION.

Para la preparación de genotecas 16S con el kit de Nanopore, se realizaron 2 secuenciaciones.

La primera, luego de la extracción del ADN en el ensayo de inoculación artificial de patógenos, se realizó la preparación de la librería genómica según el protocolo del kit de secuenciación rápida de amplicones 16S (SQK-16S024, Oxford Nanopore Technologies, Reino Unido). Este kit contiene cebadores 27 F/1492 R para la amplificación del gen rRNA 16S de longitud completa y tiene 24 pares de cebadores con código de barras (*Barcodes*), lo que permite la secuenciación simultánea de 24 muestras.

Se usaron 10 ng de ADN por muestra como molde para la reacción en cadena de la polimerasa (PCR), excepto en muestras con concentraciones bajas de ADN extraído (<1 ng/ul, determinado usando un fluorómetro Qubit, ThermoFisher, EE. UU.) donde se utilizó al menos 3 ng de ADN, se asignó los códigos de barras del kit SQK-16S024 desde el número 10 al 20. Sin embargo, luego de realizar la amplificación de las muestras con los cebadores/*barcodes*, se cuantificó usando el Qubit a las 5 muestras y se obtuvieron valores menores a 0.1 ng/μl.

Debido a esto, el número de ciclos se extendió de 25 a 35 en las condiciones de PCR (Tabla I), además, se mantuvo el mismo método de dilución para garantizar la cantidad suficiente de amplicones a secuenciar en las muestras con baja cantidad de ADN inicial. El programa se ejecutó en el termociclador (Techne Prime, USA). En este ensayo, se realizó una extracción de ADN del control negativo paralela a las extracciones de ADN de las muestras. Las concentraciones de ADN fueron demasiado bajas, sin embargo, se incluyeron en la preparación de las librerías.

Tabla I. Programa de amplificación PCR con el ciclaje modificado para el kit 16S (SQK-16S024).

Fases	Temperatura	Tiempo	Ciclos
Desnaturalización inicial	95°C	1 MIN	
Desnaturalización	95°C	20 SEG	
Anillamiento	55°C	30 SEG	35 veces
Extensión	72°C	2 MIN	
Extensión final	72°C	5 MIN	
Mantenimiento	12°C	∞	

Los productos PCR obtenidos de las muestras de alimentos se purificaron con perlas AMPure XP (Beckman Coulter, EE. UU.), de acuerdo con las instrucciones del kit. Después, los amplicones purificados se cuantificaron en el Qubit antes de agruparse y ser normalizados para una secuenciación equimolar.

Para poder realizar el *multiplexing* de manera equimolar de las distintas muestras de amplicones y evitar el sesgo de lecturas al momento de la secuenciación considerando que las muestras de queso tenían mayor cantidad de ADN y de amplicones en 10 µl de eluido, se realizó la dilución de dichas muestras a 10 fmol usando la siguiente fórmula:

$$\text{fmol} = \frac{\text{ng} \times 10^6}{\text{MW}}$$

dónde: ng es la cantidad de ADN en nanogramos. MW es el peso molecular del ADN en g/mol. Para el ADN de doble cadena, el peso molecular se puede estimar como el número de pares de bases multiplicado por 660 g/mol, para el gen 16S rRNA (~ 1500 pb) (Johnson et al., 2019).

De esta manera, las muestras de queso se diluyeron a 10 ng en 10 µl (1 ng/µl), mientras que las demás muestras se cargaron en la misma cantidad sin dilución. Después de la ligación de los códigos de barras en la PCR, y la normalización equimolar, el agrupamiento (*pooling*) de las muestras se realizó de modo que se obtuvo 110 ng (~ 110 fmol) de amplicones, incluyendo nueve muestras, el control positivo y el control negativo para la ejecución de la secuenciación. Finalmente se realizó la ligación del adaptador de secuenciación tomando solo 10 µl del *pooling* y se cargó en el minION.

La segunda secuenciación se realizó para el ensayo de comparación de tecnologías entre Nanopore e Illumina, el ADN obtenido de las extracciones de las 12 muestras de

alimentos se procesó con la misma metodología del kit 16S de Nanopore según el protocolo del fabricante sin ninguna modificación para evitar diferencias y variaciones en la comparación de tecnologías. El *pooling* de las muestras se realizó de modo que se cargó 60 ng de amplicones, el negativo no mostró amplificación por lo que no fue utilizado en la secuenciación.

Las secuenciaciones se realizaron con una celda de flujo (R9.4.1) en un minION Mk1B (Oxford Nanopore Technologies, Reino Unido) conectado a una computadora portátil Aorus Gygabyte core i7 con 32 GB de RAM, 1 TB SSD y una tarjeta de video portátil integrada de 16 GB vRAM que se utilizó con el programa MinKNOW Standalone Release for Mk1B 23.04.3 (Oxford Nanopore Technologies, Reino Unido) para configurar los parámetros de secuenciación.

La secuenciación de las muestras para la evaluación de la sensibilidad del minION se realizó por 2 horas, mientras que la secuenciación para la comparación de plataformas Illumina vs Nanopore con 12 muestras de alimentos se realizó por 3 horas.

3.6. Secuenciación 16S en Illumina

El mismo ADN extraído de las 12 muestras de alimentos que se usó para la secuenciación de Nanopore se usó para la secuenciación de Illumina. Las muestras se prepararon según indicaciones del centro de secuenciación. La secuenciación de Illumina se realizó siguiendo el protocolo de Nextera XT de la marca Illumina, los amplicones utilizados en el presente estudio son de la región 341F-805R (Op De Beeck et al 2014), correspondientes a gen 16S. La secuenciación se realizó en una flowcell V3 de 600 ciclos en la plataforma MiSeq System de Illumina. Los archivos Fastq fueron generados mediante el demultiplexing a través del programa MiSeq Reporter (MSR) y se realizó el control de la corrida con el software Sequencing Analysis Viewer (SAV). Finalmente, el análisis y clasificación taxonómica primario de los archivos Fastq generados se realizó en la aplicación 16S *Metagenomics* alojada en la plataforma Bioinformática BaseSpace Hub de Illumina.

3.7. Análisis de datos y bioinformática

En el experimento para establecer el protocolo de extracción de ADN, se analizaron los datos de validación y las secuencias obtenidas de las muestras de alimentos. Se empleó un análisis de varianza unidireccional (ANOVA) para comparar las medias de las variables dependientes (cantidad y calidad de ADN), así como los tiempos de cada método evaluado. Para identificar diferencias significativas entre las medias de los grupos, se utilizó la prueba post hoc de Tukey.

Todos los archivos FAST5 obtenidos de la secuenciación Nanopore se transformaron en archivos FASTQ utilizando el programa traductor “*basecaller*” Guppy v6.0.1 en la configuración de alta precisión “*Super high accuracy*” que permite la clasificación de lecturas de ADN con una calidad y puntaje >10.

Para la identificación de *S. enterica*, *L. monocytogenes* y *V. cholerae*, los archivos FASTQ que se obtuvieron del *basecalling* de las muestras de alimentos inoculadas fueron procesadas dentro del flujo standard de Oxford Nanopore con el programa de acceso web EPI2ME (v3.6.2, <https://epi2me.nanoporetech.com/downloads>) y la versión de uso local EPI2ME-LABS (v4.1.2, <https://labs.epi2me.io/downloads/>). Además, se utilizaron 3 diferentes flujos bioinformáticos que incluyeron Spaghetti (Latorre-Pérez et al., 2021), BLAST+ en su versión local (<ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>) y SituSeq (Zorz et al., 2023). Dentro de los flujos se utilizaron 2 bases de datos como referencia con secuencias del gen 16S ARNr incluyendo la del Centro Nacional de Información Biotecnológica (NCBI) y una base de datos preformateada de SILVA v. 138 (McLaren & Callahan, 2021). Los flujos bioinformáticos se muestran en la Tabla II.

Tabla II. Flujos bioinformáticos utilizados para el análisis y detección de patógenos alimentarios.

#	Programas	Ejecución	Base de referencia	Alineación/ mapeo	Interfaz
P1	EPI2ME	ONLINE	16S-NCBI	MINIMAP2	Gráfica
P2	EPI2ME-LABS	ONLINE/LOCAL	16S-NCBI	KRAKEN2	Gráfica
P3	Spaguetty-modificado	LOCAL	SILVA 138	MINIMAP2	Línea de comandos-Semiautomático-Rstudio
P4	NCBI +	LOCAL	16S-NCBI	BLASTn	Línea de comandos-CMD
P5	SituSeq	LOCAL	SILVA 138.1	Método de Clasificador Bayesiano NAÏVE	Línea de comandos-Rstudio

Posterior a la clasificación taxonómica efectuada por cada uno de los flujos bioinformáticos, se procedió a evaluar su capacidad para detectar y clasificar taxonómicamente las secuencias de ADN provenientes tanto de patógenos inoculados como de aquellos no inoculados. Durante este proceso de clasificación taxonómica de las Unidades Taxonómicas Operativas (OTUs), se observó la presencia de patógenos alimentarios de relevancia económica.

Con el objetivo de profundizar en este hallazgo, se realizó una revisión bibliográfica exhaustiva en fuentes internacionales de renombre, como la Administración de Alimentos y Medicamentos (FDA), la Organización Panamericana de la Salud (OPS) y la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO). Esta revisión se centró en la búsqueda de OTUs correspondientes a los géneros patogénicos más comúnmente encontrados en alimentos. Entre estos se incluyen *Bacillus*, *Campylobacter*, *Clostridium*, *Cronobacter*, *Escherichia*, *Shigella*, *Staphylococcus*, *Vibrio* y *Yersinia* (FDA, 2023; Doyle et al., 2013; Schmidt, 2017). En la Tabla III se muestran las especies más comunes transmitidas por alimentos correspondientes a los géneros mencionados que se consideraron en el análisis.

Tabla III. Géneros y especies de las bacterias patógenas transmitidas por alimentos de mayor importancia económica analizadas en el presente estudio.

Género	Especies
<i>Bacillus</i>	<i>Bacillus cereus</i> , <i>Bacillus anthracis</i>
<i>Campylobacter</i>	<i>Campylobacter jejuni</i> , <i>Campylobacter coli</i> , <i>Campylobacter upsaliensis</i>
<i>Clostridium</i>	<i>Clostridium botulinum</i> , <i>Clostridium perfringens</i>
<i>Cronobacter</i>	<i>Cronobacter sakazakii</i> , <i>Cronobacter malonaticus</i> , <i>Cronobacter turicensis</i>
<i>Escherichia</i>	<i>Escherichia coli</i>
<i>Listeria</i>	<i>Listeria monocytogenes</i> *, <i>Listeria ivanovii</i>
<i>Salmonella</i>	<i>Salmonella enterica</i> *
<i>Shigella</i>	<i>Shigella sonnei</i> , <i>Shigella flexneri</i> , <i>Shigella boydii</i> , <i>Shigella dysenteriae</i>
<i>Staphylococcus</i>	<i>Staphylococcus aureus</i>
<i>Vibrio</i>	<i>Vibrio parahaemolyticus</i> , <i>Vibrio vulnificus</i> , <i>Vibrio cholerae</i>
<i>Yersinia</i>	<i>Yersinia enterocolitica</i> , <i>Yersinia pseudotuberculosis</i>

* Cepas bacterianas inoculadas artificialmente en las muestras de alimentos.

Además de la detección de patógenos, se evaluaron otros parámetros clave. Entre ellos, tiempo de ejecución, la cantidad de lecturas que fueron clasificadas taxonómicamente, lo que proporcionó una medida de la diversidad microbiana en cada flujo de trabajo.

Los 5 objetos phyloseq obtenidos del análisis de cada flujo de los programas se unieron en un solo objeto phyloseq para los análisis de abundancia relativa, diversidad alfa y beta, así como, los análisis de abundancia diferencial. Estos se llevaron a cabo usando los programas y funciones de R (v4.3.2) incluyendo phyloseq (v1.44.0) (McMurdie & Holmes, 2013), microeco (v1.4.0) (C. Liu et al., 2021) y fantaxtic (v0.2.0) (Teunisse, 2022).

Los archivos FASTQ de las 12 muestras de alimentos secuenciadas por Nanopore fueron clasificadas taxonómicamente utilizando el flujo del programa EPI2ME con el fin de mantener un nivel de comparación similar a la clasificación taxonómica que se realizó con los datos FASTQ de Illumina, ya que se usó el software análogo 16S del sitio web de Illumina. Para la comparación de los perfiles bacterianos y la diversidad hallada, se obtuvieron las OTUs de cada tecnología y también se procesaron en R (v4.3.2) utilizando el paquete *Vegan* (v2.6.4) (Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Sólymos, M. Henry H. Stevens, 2022) para los análisis diversidad.

La alfa-diversidad para ambos conjuntos de datos se calculó utilizando el método de Shannon, las matrices de distancia utilizadas para el agrupamiento jerárquico se obtuvieron con el método de Bray-Curtis. La clusterización se realizó utilizando el algoritmo de Ward. Se calculó la correlación cofenética entre dendrogramas con la función `corr.dendlist()` del paquete *dendextend* (v1.17.1) (Galili, 2015). Se realizó el análisis de varianza multivariante permutacional (PERMANOVA) para evaluar si la composición del microbioma difería según los métodos de secuenciación evaluados.

Las clasificaciones taxonómicas y abundancias relativas a nivel de filo y género de las secuencias 16S rRNA obtenidas de los 12 alimentos fueron utilizadas para comparar la secuenciación de Illumina y Nanopore.

Para visualizar las relaciones entre las plataformas de secuenciación, se realizó las ordenaciones de escalamiento multidimensional no métrico (NMDS, por sus siglas en inglés) y las pruebas de análisis de similitudes (ANOSIM, por sus siglas en inglés). La disimilitud de Bray-Curtis se utilizó como medida de disimilitud para las ordenaciones NMDS, así como la prueba ANOSIM. Para determinar la abundancia relativa de

taxones en los conjuntos de datos de Illumina y Nanopore se utilizó el método de correlación de Pearson.

En la evaluación de los cinco flujos bioinformáticos para obtener perfiles bacterianos, se sometieron las lecturas válidas obtenidas y la cantidad de géneros clasificados taxonómicamente a un MANOVA. Se consideraron significativas las diferencias con un valor de p menor a 0.05. Además, se realizó un ANOVA multifactorial de dos niveles para comparar las medias de diversidad de Shannon entre las plataformas de secuenciación. La prueba post hoc de Tukey se utilizó para explorar las diferencias entre las medias de las plataformas evaluadas.

CAPÍTULO 4

RESULTADOS Y DISCUSIÓN

4.1 Extracción de ADN de las muestras de alimentos

Luego de las extracciones realizadas con los 3 métodos evaluados, las comparaciones entre medias se realizaron tomando como factor cada método y los tiempos como variable independiente, las variables de respuesta fueron la concentración y la calidad de ADN de cada método (Tabla IV). No se observaron diferencias estadísticas al comparar las concentraciones obtenidas de ADN ($P = 0.145$), o el tiempo de duración del protocolo ($P = 0.102$) con los métodos evaluados. Si se observó una diferencia significativa al comparar el método de extracción y la calidad de las muestras de ADN ($P < 0.001$).

Tabla IV. Métodos de extracción evaluados, tiempo de ejecución del protocolo con una muestra de frutas y 3 réplicas, se muestra la concentración de ADN en ng/μl y la calidad en el rango 260/280. Los valores se expresaron como la media más la desviación estándar.

Métodos	Tiempo de ejecución (minutos)	Concentración de ADN (ng/μl)	Calidad 260/280
Chen et al. (2019)	20	10,03 ± 7,60	1,21 ± 0,081
Purelink	70	1,54 ± 1,55	1,77 ± 0,075
Dilhari et al. (2017)	50	9,77 ± 1,04	1,78 ± 0,032

La comparación entre medias para diferenciar el efecto del método sobre la calidad del ADN se realizó con la prueba de Tukey. Los resultados de la prueba de Diferencia Significativa Honesta (HSD) de Tukey mostraron que la media del protocolo de *Purelink* fue significativamente mayor que la media del protocolo de Chen en aproximadamente 0.563 unidades. El valor p fue altamente significativo $P < 0.001$, lo que confirmó la importancia de esta diferencia.

Al igual que el método de *Purelink*, la media del método de Dilhari fue significativamente mayor que la media de Chen en aproximadamente 0.573 unidades. El valor p ajustado fue $P < 0.001$. La diferencia entre las medias del método de Dilhari y el de *Purelink* no fue significativa. La diferencia media fue de aproximadamente 0.010 unidades, pero el intervalo de confianza del 95% para esta diferencia fue de -0.156 a 0.176 unidades, lo que indica que no hay una diferencia significativa. En la Figura 2 se observa que los métodos de *Purelink* y Dilhari tienen puntuaciones de calidad significativamente más altas que el de Chen, pero no existe una diferencia significativa entre los dos métodos restantes.

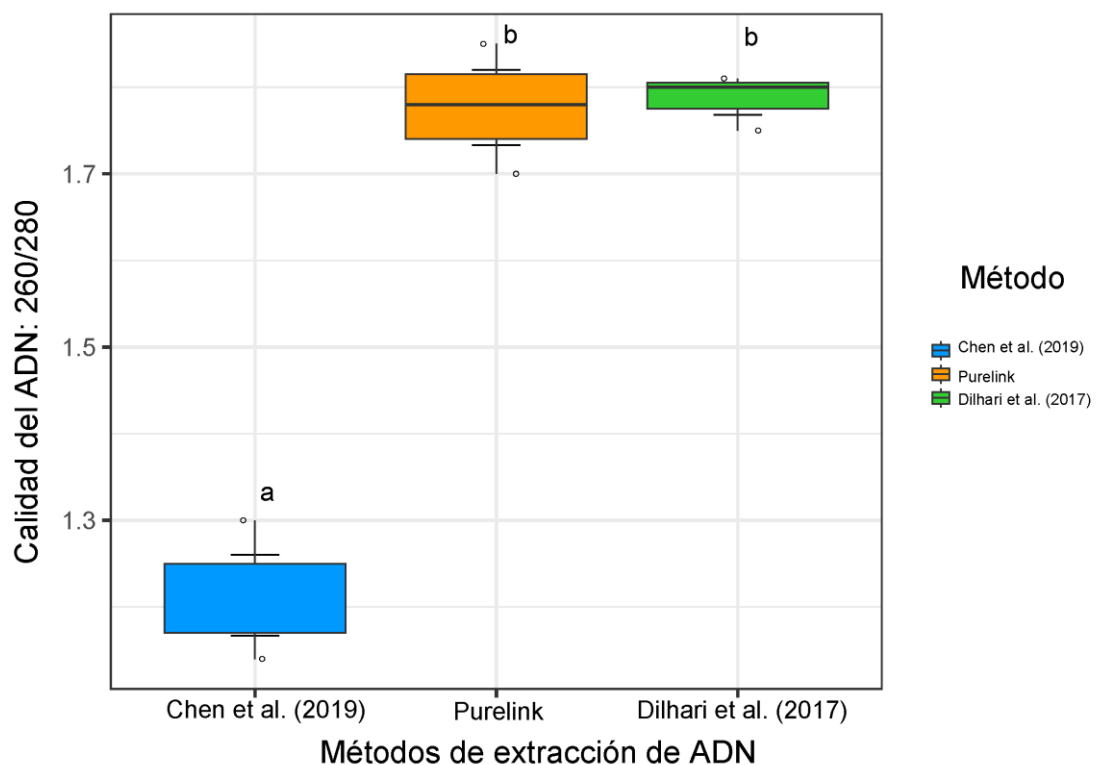


Figura 2. Comparación de los efectos del método de extracción sobre la calidad del ADN obtenido de las muestras de frutas picadas. El eje de las ordenadas representa la medida de calidad obtenida para cada método de extracción. El eje de las abscisas muestra los tres métodos de extracción que se están comparando: el método de Chen et al. (2019) (Método A), Purelink (Método B) y el método de Dilhari et al. (2017) (Método C). Cada caja representa la distribución de los datos de calidad para cada método. La línea en el medio de la caja es la mediana. La parte superior e inferior de la caja representan el tercer y primer cuartil respectivamente, lo que da el rango intercuartil (IQR). Las letras 'a' y 'b' indican grupos estadísticamente diferentes. Los puntos representan valores atípicos.

La cantidad de ADN que se obtuvo con estos recursos fue partiendo de 10 ml de muestra inicial. Estos resultados son esperados en este tipo de muestras, puesto que al ser una mezcla de frutas existen varios tipos de partículas que interfieren en el proceso de extracción del ADN, estas partículas pueden ser proteínas, lípidos, azúcares que pueden influir en la calidad y cantidad del ADN extraído (Barbosa et al., 2016; Maksimov et al., 2017). Al igual que la investigación realizada por McGaughey et al. (2018), donde obtuvieron cantidades promedio de 20 a 50 ng/ μ l de matrices fecales con calidades promedio 1.84 similares a las obtenidas en el rango 260/280 de las extracciones realizadas en esta investigación, el uso de kits comerciales de extracción no siempre garantiza los mismos resultados que se obtienen en las pruebas del kit puesto que existen muchos factores que pueden influenciar estos parámetros de calidad, como por ejemplo, el tiempo de incubación, la cantidad de enzimas utilizadas, entre otras.

La purificación del ácido desoxirribonucleico (ADN), llevada a cabo mediante la utilización de acetato de potasio a una concentración de 5M, junto con la concentración del ADN extraído a través del uso de isopropanol, se demostraron como métodos altamente efectivos para la obtención de ADN de calidad superior y en volúmenes adecuados. Estos procedimientos resultaron ser más eficientes en términos de tiempo y desde el punto de vista económico.

En contraste con la extracción utilizando el kit PureLink®, que requiere aproximadamente 2 horas por muestra, las modificaciones propuestas permitieron la obtención del ADN en un lapso significativamente reducido, aproximadamente 50 minutos, el mejor protocolo es que permite la extracción de la mayor cantidad de muestras posible en una sola vez con el fin de evitar sesgos en los análisis subsiguientes (Elie et al., 2023; Heller et al., 2003; Xin & Chen, 2012).

Posteriormente, se utilizó la metodología de Dilhari et al. (2017) para el resto de las 12 muestras de alimentos, obteniéndose las cantidades de ADN y calidad requeridas para las actividades de secuenciación (Tabla V). El volumen de elución final de ADN que se obtuvo de cada muestra fue de 200 μ l, para cada tecnología de secuenciación se utilizó el mismo ADN dividiendo 100 μ l para cada una.

Es importante destacar que la calidad de las muestras de ADN no alcanzaron la requerida por el protocolo de secuenciación de amplicones de Oxford Nanopore (206/280: > 1.80- 2.00), sin embargo, la relación entre la calidad y la cantidad no influyeron en la capacidad de amplificación de los cebadores de Nanopore en la PCR, debido principalmente a la cantidad utilizada para la preparación del coctel PCR (2 μ l)

alcanzando diluciones de 5 a 10 ng de ADN plantilla, esto concuerda con estudios similares donde se validaron protocolos de extracción de ADN de diferentes matrices alimentarias (Piskata et al., 2019).

La extracción alcalina con NaOH (Figura 3) probó ser un método eficaz y de bajo costo en las muestras analizadas en esta investigación, ajustando la purificación con soluciones de bajo costo, estos resultados probaron ser concluyentes en su implementación para el estudio de perfiles microbianos metagenéticos al igual que se ha utilizado en estudios similares de seguridad alimentaria (Shwani et al., 2023).



Figura 3. ADN genómico de 12 muestras de alimentos utilizando el protocolo de Dilhari et al. (2017) modificado. Las líneas rojas indican el tamaño de los fragmentos del marcador de 1500 pares de bases (PB). Las iniciales indican B: bolones, C: ceviches, CM: carne de res, E: encebollados, EF: ensalada de frutas, EM: empanadas de queso, F: frutas picadas, MT: mote con choncho, P: pollo, Q: quesos, SP: salchipapas, S: salsas picantes.

Estos resultados concuerdan con estudios preliminares relacionados al análisis del gen 16S ARNr en perfiles microbianos de interés como las bacterias patógenas transmitidas por alimentos de Salazar-Llorente et al. (2020) y Witte et al. (2018), hallando que los métodos de extracción de formulación abierta o caseros funcionan eficientemente para el procesamiento de muestras complejas como lo son los diferentes tipos de alimentos, realizando pequeñas modificaciones según la complejidad y característica intrínseca de la muestra.

Tabla V. Extracción de ADN de 12 muestras de alimentos con el protocolo de Dilhari et al. (2017) modificados, se muestra la cantidad y la calidad medidos en Qubit y Nanodrop, respectivamente.

#	Alimentos	ID	Cantidad ADN (ng/μl)	Calidad ADN (260/280)
1	Bolones	B	39.65	1.88
2	Ceviches	C	71.2	1.79
3	Carne de res	C	36.1	1.84
		M		
4	Encebollados	E	26.35	1.80
5	Ensalada de frutas	EF	76	1.85
6	Empanadas de queso	E M	26.35	1.60
7	Frutas picadas	F	33.5	1.79
8	Mote con chancho	M T	54	1.80
9	Pollo	P	39.4	1.88
10	Quesos	Q	50.7	1.92
11	Salchipapas	SP	23.85	1.77
12	Salsas picantes	S	15.91	1.80

En la investigación de Markusková et al. (2021) donde se investigó los perfiles microbianos en muestras de queso, que son muestras con altos contenidos de grasas y azúcares, también se evaluó y comparó protocolos de extracción con kits comerciales y caseros, hallando valores desde 20 a 28 ng/μl en las concentraciones de ADN cuando utilizó protocolos caseros permitiendo la cuantificación de algunas especies mediante PCR en tiempo real (qPCR).

4.2 Librerías genómicas 16S rRNA para el multiplexeado e identificación de *S. enterica*, *L. monocytogenes* y *V. cholerae*.

En el experimento de sensibilidad del secuenciador minION donde se realizaron inoculaciones artificiales de las bacterias objetivo, las cantidades de ADN para la

preparación de librerías 16S rRNA que se obtuvieron fueron muy bajas ($< 0.01 \text{ ng}/\mu\text{l}$) debido a que las colonias que se obtuvieron de la dilución 10^{-9} después de 24 horas de incubación fueron apenas visibles (Figura suplementaria 2). Debido a estas variables se optó en la modificación del método de dilución y aumentar la cantidad de ciclos en la PCR. Estos resultados son consistentes con los hallazgos de Villette et al. (2021), quienes emplearon protocolos de extracción complejos para obtener suficiente ADN y analizar perfiles microbianos en muestras con concentraciones bacterianas inferiores a 10^{-6} .

Como lo mencionó Teng et al. (2018) la variabilidad de las matrices y las diferencias en los tipos de bacterias que se hallan influye directamente en la cantidad de ADN que se obtiene y se logra amplificar mediante PCR, sin embargo, el ajuste de los protocolos a cada grupo de muestras en cuanto al método de lisis, beneficia y amortigua el rendimiento de los protocolos de formula abierta.

De esta manera, para las muestras de queso la cantidad que se determinó para la librería 16S fue de $\sim 1 \mu\text{l}$ de la muestra amplificada (PCR) y $9 \mu\text{l}$ de agua bidestilada (ddH_2O). El resto de las muestras no se diluyeron y se unieron tomando $10 \mu\text{l}$ directamente de cada muestra. En total el volumen del *pooling* fue de $110 \mu\text{l}$, de este se tomó solo $10 \mu\text{l}$ para la ligación del adaptador de secuenciación.

Estas modificaciones dieron como resultados una secuenciación satisfactoria que permitió obtener datos de secuenciación de *L. monocytogenes*, *S. enterica* y *V. cholerae* utilizando el flujo oficial de Nanopore, EPI2ME.

En la Tabla VI Se observa la cantidad de ADN obtenida de la extracción de las muestras inoculadas luego de 24 horas de incubación en caldo de cultivo, además se anotó la cantidad de amplicones obtenidos de cada muestra usando 35 ciclos de amplificación y las lecturas obtenidas para los patógenos en estudio.

Tabla VI. Cantidad de ADN obtenida de la extracción de las muestras inoculadas luego de 24 horas de incubación en caldo de cultivo. Se anotó la cantidad de amplicones obtenidos de cada muestra usando 35 ciclos de amplificación y la cantidad de lecturas para *L. monocytogenes*, *S. enterica* y *V. cholerae*. Se expresan la media de cada muestra \pm DE.

Muestra	Cantidad de ADN (ng/ul)	Cantidad de amplicones (ng/ul)	<i>Salmonella Enterica</i> (Taxones/ Lecturas)	<i>Listeria</i> (Taxones/ Lecturas)	<i>Vibrio cholerae</i> (Taxones/ Lecturas)
C. positivo	2.25	1.06	1856	0	3
Quesos	1.66 \pm 0.97	70.60 \pm 5.30	4	9	3
Salsas	5.34 \pm 2.66	0.85 \pm 0.16	132	0	34
Frutas	0.32 \pm 0.04	0.89 \pm 0.17	0	0	2
C. negativo	0.653	0.926	0	0	0

4.3. Evaluación de flujos bioinformáticos para la identificación de *S. enterica*, *L. monocytogenes* y *V. cholerae*.

Tras la preparación de la genoteca 16S ARNr utilizando las muestras que fueron inoculadas artificialmente, se procedió a su secuenciación en el equipo MinION y posterior análisis con los 5 flujos bioinformáticos establecidos.

Culminada las 2 horas de secuenciación en el minION de estas muestras inoculadas, se procesó en total 169 108 mil secuencias individuales de ADN, además se obtuvo un rendimiento de 240.4 millones de bases (Mbases), una longitud media de secuencias de 1421 pb (N50). La calidad de las lecturas obtuvo un puntaje promedio de 14.72, siendo 7 el puntaje mínimo para la validación de una lectura de ADN, la moda de estos datos se ubicó en 15.55 como se observa en la Figura 4.

En estudios similares de detección de bacterias patógenas usando la plataforma de Nanopore se utilizaron tiempos similares de secuenciación, de entre 2 a 6 horas, alcanzado resultados satisfactorios cuando se evaluaron librerías de amplicones 16S rARN para la detección de *Salmonella* (Quick et al., 2015; Xu et al., 2020). Al utilizar estos tiempos cortos de secuenciación en la detección de bacterias patógenas en diferentes matrices de importancia económica los programas de seguridad alimentaria se benefician y adquieren una mayor eficiencia para prevenir brotes epidémicos de enfermedades transmitidas por bacterias patógenas (Rantsiou et al., 2018b).

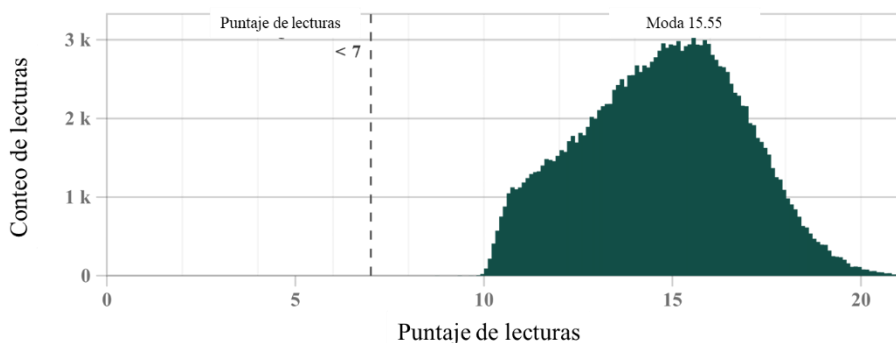


Figura 4. Representación gráfica del conteo de lecturas de ADN en 2 horas de secuenciación y el puntaje de calidad asignado, considerando 7 como el valor mínimo para ser considerado una lectura válida. La moda de las lecturas obtenidas fue de 15.55.

La cantidad de archivos FAST5 (Secuencias en bruto “raw”) que se obtuvieron fue de 4.49 gigabytes (Gb). A continuación, se realizó el basecalling de estos datos convirtiendo los archivos FAST5 en archivos FASTQ obteniéndose un total de 513 megabytes (Mb). La cantidad de datos de secuenciación obtenidos con estos parámetros de filtrado de calidad de las lecturas (Qscore) son similares a los obtenidos por Forghani et al. (2020) que al utilizar los parámetros de secuenciación de cortos periodos de tiempo (1 hora) en experimentos de inoculación artificial de alimentos para la detección de especies de *Salmonella* y *Escherichia*, obtuvieron cerca de 15 K de lecturas con un tamaño promedio de 3 kb en sus lecturas.

En el presente estudio, se evaluó la sensibilidad del secuenciador para detectar cantidades mínimas de ADN derivado de células microbianas. Se emplearon concentraciones de ADN tan bajas ($< 1 \text{ ng}/\mu\text{l}$) desde una concentración microbiana de 10E^{-9} , siendo estos parámetros similares a las que se encuentran en condiciones ambientales naturales (Lobiuc et al., 2023).

Los flujos de trabajo bioinformáticos empleados en este estudio se ejecutaron en un equipo portátil cuyas especificaciones técnicas se detallaron previamente. Los parámetros clave de rendimiento, incluyendo los tiempos de ejecución de cada programa y el consumo de memoria RAM, se registraron y analizaron. Estos datos proporcionaron una medida cuantitativa del rendimiento y la eficiencia de cada flujo de trabajo en el contexto de este estudio específico. Los resultados de este análisis se presentan en anexos de la Tabla Suplementaria I.

Similar a las investigaciones de de Siqueira et al. (2021) y Zorz et al. (2023) la capacidad para realizar estos análisis metagenéticos de perfiles microbianos en estudios ambientales o en condiciones que implican el escaso acceso a recursos de un laboratorio es esencial minimizar los requerimientos de software y hardware así como contar con la oportunidad de realizar análisis sin acceso a internet, los flujos probados mostraron versatilidad y facilidad de manejo al emplear bases de datos 16S rARN curadas y de bajo peso en megabytes (< de 100 mb), lo que permite complementar la portabilidad del secuenciador Nanopore (Santos et al., 2020).

La Tabla VII presenta los resultados obtenidos de la detección de microorganismos patógenos y el recuento total de lecturas clasificadas en OTUs clasificadas en cada uno de los flujos bioinformáticos evaluados. Es notable que, en los flujos 1 y 4, correspondientes a EPI2ME y NCBI+ respectivamente, se logró identificar a las tres especies propuestas en este estudio, a pesar de que únicamente se inoculó en los alimentos dos especies certificadas, *Salmonella enterica* y *Listeria monocytogenes*. Es importante destacar que *Listeria monocytogenes* no fue identificada a nivel de especie, sino únicamente a nivel de género. Por otro lado, las especies pertenecientes al género *Vibrio* fueron detectadas principalmente en la muestra de queso.

En investigaciones similares utilizando la secuenciación de Nanopore para la detección de bacterias patógenas realizando inoculación artificial de bacterias patógenas como *Salmonella enterica*, *E. coli*, entre otros, los flujos de EPI2ME y NCBI+ obtuvieron resultados similares en cuanto al número de OTUs clasificadas puesto que comparten la misma base taxonómica de clasificación y su método de depuración y selección de lecturas válidas no es riguroso y requiere de menos procesos de filtrado (Forghani et al., 2020; Voorhuijzen-Harink et al., 2019; Wu et al., 2023).

La ausencia en la detección de la especie inoculada de *L. monocytogenes*, tuvo relación con el método de extracción utilizado, la dilución empleada y las características de pertenecer al grupo de las bacterias gram positivas (Markusková et al., 2021; Teng et al., 2018). Como lo declaran Lakicevic et al. (2023) y Wagner et al. (2021), ante la dificultad que implica la detección de *L. monocytogenes* en estudios de secuenciación de siguiente generación, tanto de Illumina como Nanopore, el medio de cultivo para el enriquecimiento de las muestras, el método de extracción y el tiempo de secuenciación favorecen la aparición y detección de lecturas con tamaños más grandes en Millones de bases que permiten una mejor clasificación al momento de realizar el contraste con la base de datos utilizada, puesto que además, es importante recalcar que a nivel de género la detección de esta especie fue lograda.

En relación a la detección de *V. cholerae*, aunque no se esperaba su detección debido a la condición de inocuidad de la muestra adquirida, su detección ha sido común en diferentes estudios de inocuidad alimentaria, especialmente en productos derivados de lácteos y con cantidades importantes de cloruro de sodio, ya que son condiciones microbiológicas ideales para el desarrollo de estas bacterias (Díaz Cárdenas et al., 2024; Dugat-Bony et al., 2020; El-Zamkan et al., 2023; Salazar-Llorente et al., 2020).

Tabla VII. Evaluación de los 5 flujos bioinformáticos de los programas para la detección de bacterias patógenas transmitidas por alimentos, se muestran las lecturas obtenidas, el número de géneros y especies identificados.

Flujo	Lecturas clasificadas	Taxones asignados a géneros	Taxones asignados a especies	<i>Salmonella enterica</i> (Taxones/Lecturas)	<i>Listeria</i> (Taxones/Lecturas)	<i>Vibrio</i> (Taxones/Lecturas)
P1	166528	224	795	6 (1992)	6 (9)	14(42) *6 lecturas de <i>Vibrio cholerae</i>
P2	166528	82	253	1 (2018)	0 (0)	1(1)
P3	134977	58	184	1 (1240)	0 (0)	1(1)
P4	167534	238	888	6 (1985)	10 (14)	15(37) *2 lecturas de <i>Vibrio cholerae</i>
P5	8780	21	82	1(1)	0 (0)	1(1)

Para evaluar la relación y diferencias entre estas variables dependientes que implican las lecturas clasificadas, así como su relación con los flujos utilizados se implementó el análisis de varianza multivariado (MANOVA).

Entre las variables de respuesta, se observó una diferencia notable en la distribución de datos de las lecturas clasificadas taxonómicamente, las que se asignaron a nivel de género y las que se asignaron a nivel de especies en comparación con las lecturas evaluadas de los patógenos detectados (Figura 5).

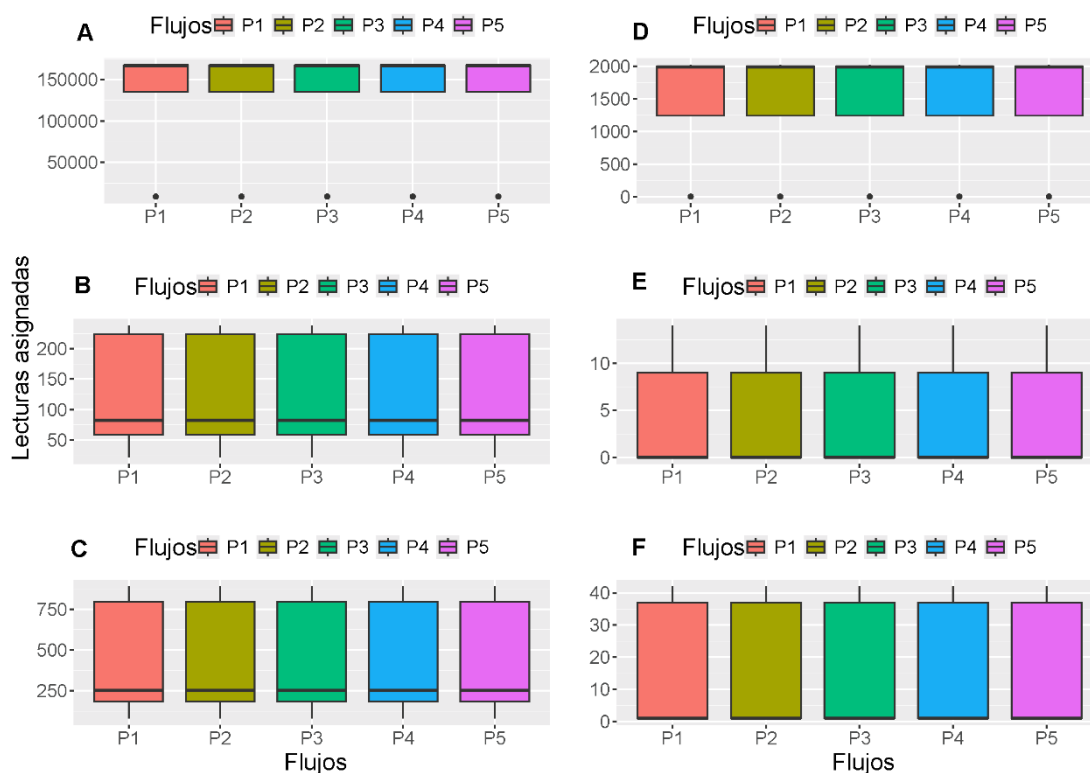


Figura 5. Distribución de las lecturas de secuenciación de Nanopore en los diferentes flujos bioinformáticos evaluados. En el eje de las ordenadas se representa las lecturas halladas y en el de las abscisas los flujos bioinformáticos. El literal A indica el total de lecturas que se lograron clasificar, el B indica las lecturas asignadas a nivel de género, el literal C muestra las lecturas asignadas a nivel de especie. El literal D, E y F indican las lecturas asignadas a Salmonella, Listeria y Vibrio, respectivamente. Los flujos son P1: EPI2ME, P2: EPI2ME_LABS, P3: Spaguetty_modificado, P4: NCBI+ y P5: Situseq. Desde A, B y C, se observan las clasificaciones de lecturas clasificadas y asignadas a un taxón. D, E y F son las lecturas asignadas por los flujos a las bacterias patógenas.

La investigación incluyó dos pruebas de MANOVA para analizar el impacto de diferentes flujos bioinformáticos en la cantidad de lecturas procesadas y asignadas a bacterias patógenas. Sin embargo, no se encontraron diferencias estadísticas significativas en ninguno de los dos casos. Los resultados sugieren una independencia en la obtención de datos entre los flujos de programas, ya que cada uno tiene sus propios parámetros y base de datos para la clasificación (Hiergeist et al., 2023).

En cuanto a los flujos de trabajo específicos, el programa EPI2ME demostró una notable capacidad para detectar lecturas asociadas a patógenos alimentarios, identificando especies de *S. enterica*, *Listeria* y *V. cholerae*. El flujo EPI2ME-LABS, en su versión local, también detectó varios géneros bacterianos relacionados con los patógenos objetivos. Spaguetty y NCBI+ también lograron detectar especies de interés, aunque con variaciones en la cantidad y tipo de lecturas identificadas. Por último, el programa Situseq detectó una menor cantidad de lecturas relacionadas con los patógenos objetivos.

Estos hallazgos sugieren que la elección del flujo bioinformático puede influir en la detección de bacterias potencialmente patógenas, pero no necesariamente en la cantidad de lecturas procesadas. Los resultados detallados de las detecciones de patógenos según el flujo utilizado se presentan en la Figura 6.

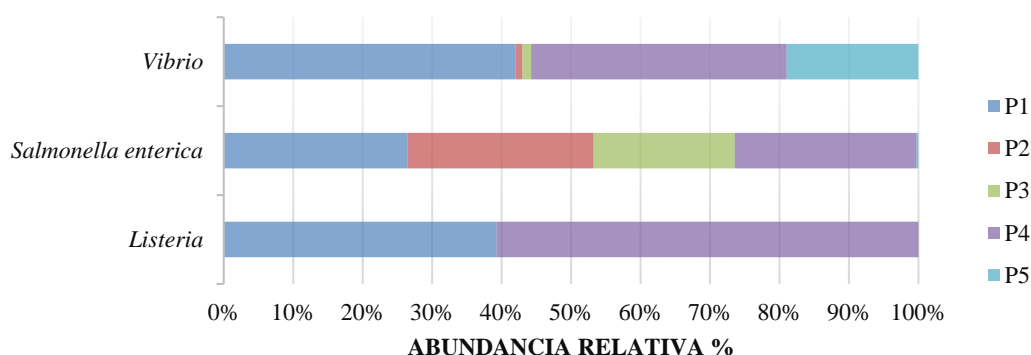


Figura 6. Abundancia relativa de las bacterias patógenas inoculadas detectadas según el flujo bioinformático utilizado. Las legendas del gráfico se describen P1: EPI2ME, P2: EPI2ME_LABS, P3: Spaguetty_modificado, P4: NCBI+ y P5: Situseq.

En la evaluación realizada con los diferentes flujos de programas bioinformáticos para la identificación de bacterias patógenas en muestras de alimentos se utilizaron principalmente dos bases de datos: SILVA 138 y NCBI, cada una con sus propias ventajas y desventajas.

La base SILVA 138 ofrece una cobertura integral de las lecturas y bases de ADN, lo que favorece la identificación precisa de las OTUs. En contraste, la base del NCBI integra múltiples grupos de bases, como GenBank, ENA y DDBJ, lo que puede generar

diferencias en la clasificación debido al tamaño de su base y la variedad de grupos de bases que maneja (Camilla Ceccarani, 2023; Hiergeist et al., 2023; Molano et al., 2024).

Los resultados, detallados en las figuras y en las tablas suplementarias 3 y 4, mostraron que los flujos de trabajo variaron en su capacidad para detectar bacterias potencialmente patógenas de transmisión alimentaria. Por ejemplo, el flujo EPI2ME identificó un alto número de lecturas asociadas a patógenos transmitidos por alimentos (23 especies), mientras que el flujo de Spaguetty modificado no detectó ninguna de las especies buscadas. El flujo de NCBI+ también logró identificar varias especies de bacterias potencialmente patógenas (19 especies), aunque en menor medida que el flujo EPI2ME. Por otro lado, el flujo de Situseq tuvo un rendimiento notablemente inferior en la detección de estas bacterias pues solo detectó 2 especies.

En resumen, los resultados resaltan la importancia de seleccionar el flujo bioinformático adecuado para la detección precisa de bacterias patógenas en muestras de alimentos, considerando las características y cobertura de las bases de datos utilizadas en cada flujo.

Adicional a la detección de las bacterias patógenas inoculadas, en cada flujo se detectaron varias de las especies patógenas indicadas en la Tabla III. Se observó en la Figura 7 las abundancias relativas de todas las bacterias patógenas detectadas en cada alimento según el flujo bioinformático evaluado.

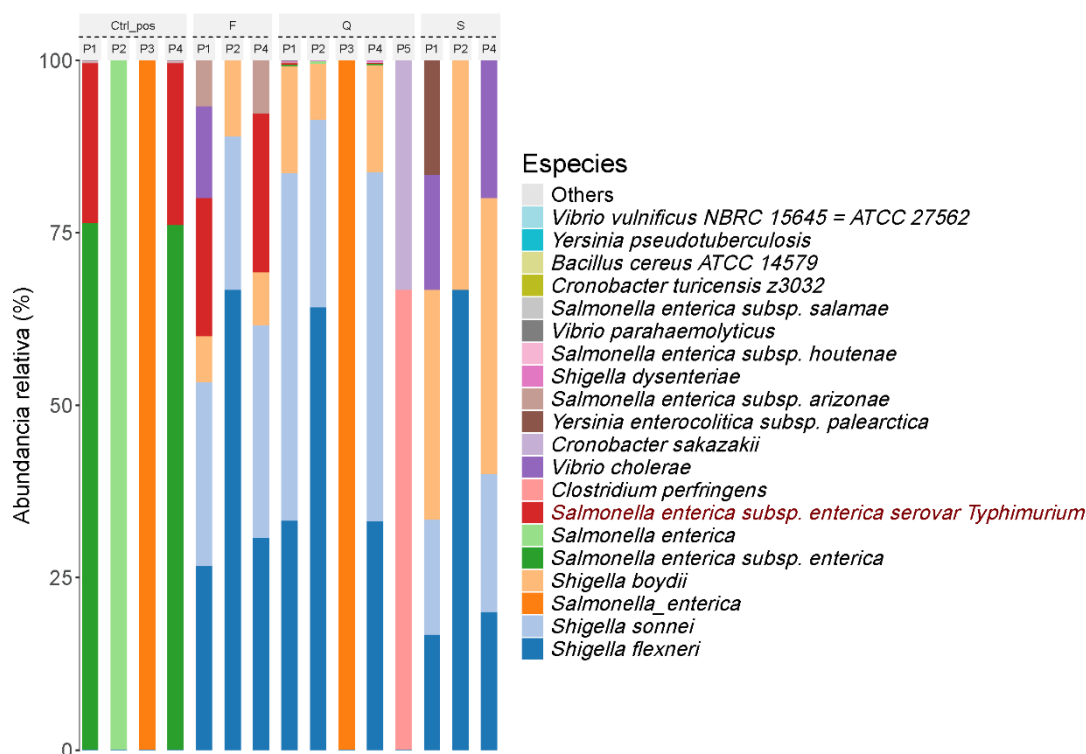


Figura 7. Abundancia relativa de las especies patógenas transmitidas por alimentos detectadas en las lecturas de secuenciación de las muestras de alimentos inoculadas artificialmente según el flujo bioinformático. Los códigos por tipos de alimentos son: Queso (Q), Control positivo (Ctrl_pos), Frutas (F), salsas (S) y control negativo (Ctrl_neg). De los flujos bioinformáticos, EPI2ME (P1), EPI2ME_LABS (P2), Spaghetty_modificado (P3), NCBI+ (P4) y Situseq (P5).

Hiergeist et al. (2023) señalan que los flujos bioinformáticos varían en la clasificación taxonómica de bacterias, con algoritmos que contrastan en parámetros según el tamaño de bases detectado. Los flujos EPI2ME y NCBI+ destacan en la identificación de lecturas. Los flujos Sphagetty y Situseq se distinguen por su proceso de validación de lecturas, que reduce la cantidad y tamaño de lecturas aprobadas hasta en un 50 % (Latorre-Pérez et al., 2021; Zorz et al., 2023).

Los datos de secuenciación de lecturas largas, como el gen 16S ARNr, permiten estudios preliminares de comunidades microbianas, incluyendo patógenos alimentarios, aunque se requieran tecnologías más precisas para la identificación. Estos datos preliminares pueden guiar futuros diseños de experimentos para estudios de vigilancia en soberanía alimentaria, como estudios metagenéticos (Grinevich et al., 2024; Lewis et al., 2020; Mira Miralles et al., 2019).

La biodiversidad microbiana en muestras de alimentos, detectada mediante diferentes flujos bioinformáticos, se evaluó mediante el índice de diversidad alfa de Shannon. El flujo P3 de Spaguetty-modificado mostró la mayor diversidad ($P_{\text{adjust}} = 0.021$). La diversidad también varió entre las muestras de alimentos, siendo el queso el más diverso $P_{\text{adjust}} = 0.024$. El análisis de la varianza multifactorial mostró una alta significancia estadística ($P < 0.0001$) en la influencia de los flujos bioinformáticos y las muestras de alimentos en la diversidad. La diversidad en el flujo de Spaghetty se atribuye a la cantidad de programas integrados y los filtros aplicados, que reducen lecturas inválidas y aseguran una clasificación precisa (Latorre-Pérez et al., 2021).

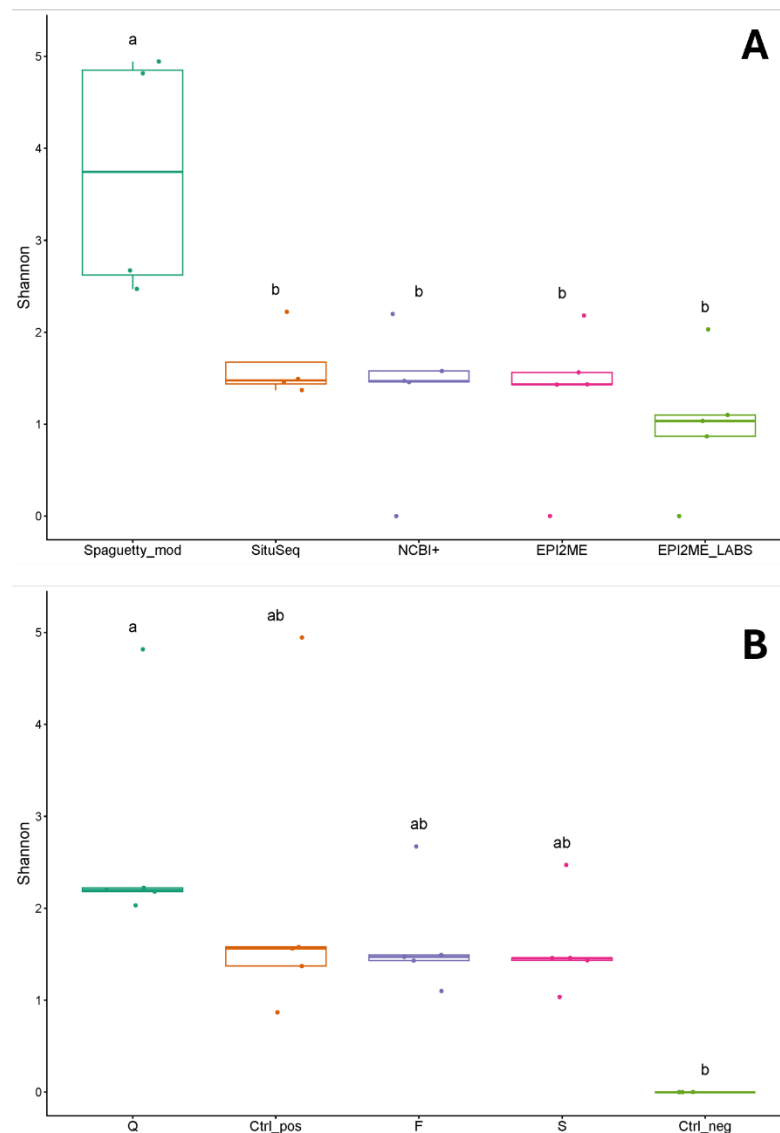


Figura 8. Diversidad de Shannon entre los 5 flujos bioinformáticos (A) y alimentos evaluados (B). Las letras sobre cada conjunto de datos indican la diferencia significativa del grupo de TukeyHSD. Los cuadros indican el rango intercuartílico de los datos. Las barras de error se extienden hasta el valor más extremo dentro de 1,5 veces el rango intercuartil y los puntos representan valores atípicos más allá de ese rango. Los códigos por tipos de alimentos son: Queso (Q), Control positivo (Ctrl_pos), Frutas (F), salsas (S) y control negativo (Ctrl_neg).

El estudio encontró que las comunidades bacterianas en las muestras de alimentos variaban significativamente dependiendo de los flujos bioinformáticos analizados,

según los resultados de la prueba de PERMANOVA (disimilitud de Bray-Curtis: $F = 2.090$, $R^2 = 0.317$, $P = 0.001$, con 999 permutaciones). La Figura 9 ilustra este hallazgo mediante un análisis de coordenadas principales (PcoA) utilizando medidas de distancias de diversidad beta de Bray. Cada punto en el gráfico representa una muestra individual. Se observó que alrededor del 9.2% al 9.9% de la variabilidad entre los flujos se explicaba por las diferencias en la diversidad encontrada entre los programas Spaguetty-modificado y Situseq, lo que concordaba con los resultados obtenidos utilizando el índice de diversidad de Shannon. Sin embargo, no se encontraron diferencias significativas entre los grupos de bacterias de los flujos utilizando las distancias de Bray-Curtis.

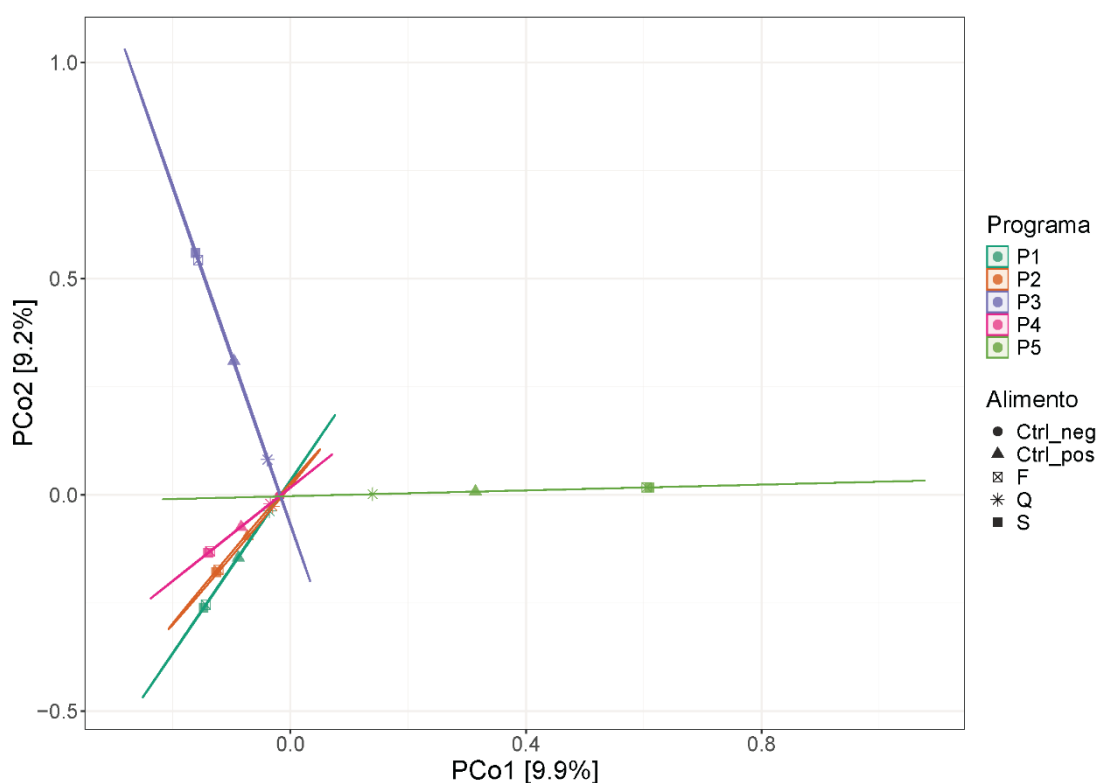


Figura 9. El análisis de coordenadas principales (PcoA) revela la agrupación de los flujos P1, P2 y P4 utilizando la distancia de Bray Curtis de muestras de alimentos por ciudad de muestreo. Mientras que se observa la diferencia del microbioma en los flujos P3 y P5. Los códigos por tipos de alimentos son: Queso (Q), Control positivo (Ctrl_pos), Frutas (F), salsas (S) y control negativo (Ctrl_neg). De los flujos bioinformáticos, EPI2ME (P1), EPI2ME_LABS (P2), Spaghetty_modificado (P3), NCBI+ (P4) y Situseq (P5).

El análisis de abundancia diferencial que determinó la diferencia en los perfiles microbianos de las muestras de alimentos según el flujo bioinformático utilizado

mostró que las lecturas de los géneros detectados en el programa de Spaghetti-modificado influyen en la variabilidad que este alcanza en comparación con los demás flujos, como lo muestra la Figura 10, el índice de Gini para *Escherichia-Shigella*, *Raoultella*, *Enterobacter*, *Lactococcus*, *Citrobacter* y *Klebsiella* muestra una similitud en el índice alcanzado > 0.4 influenciando directamente en la AR del Flujo 3.

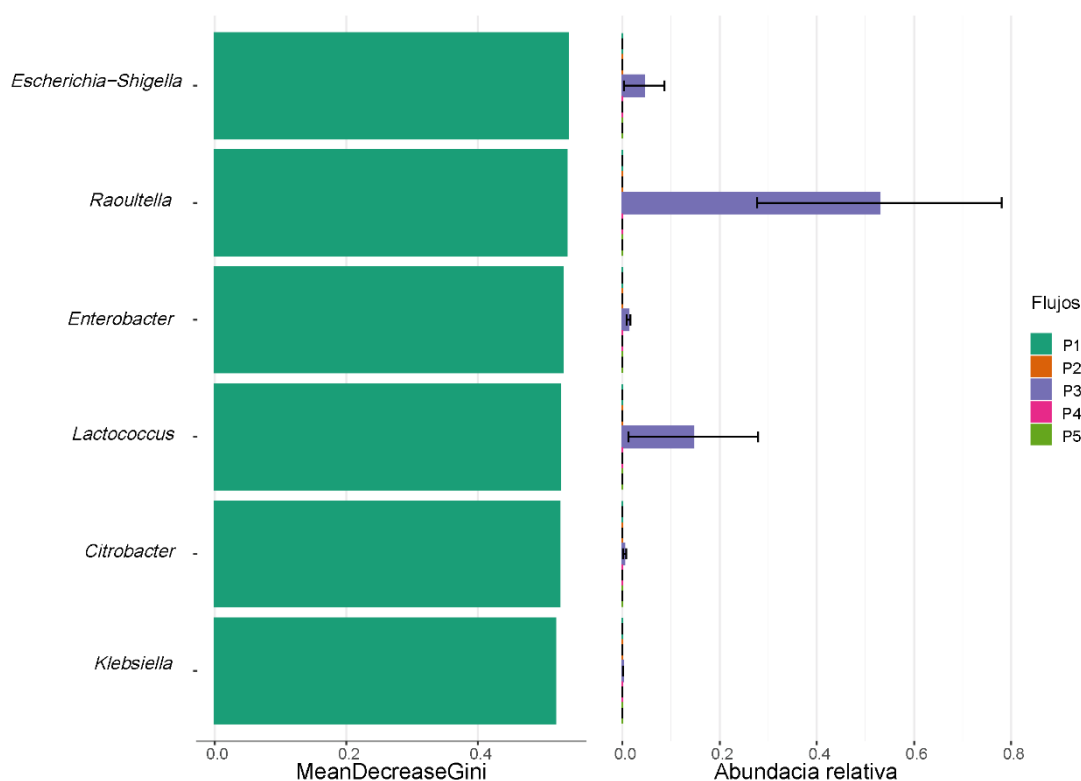


Figura 10. Prueba de abundancia diferencia utilizando el método Random forest y el índice MeanDecreaseGini para la detección de las lecturas de los géneros más significativos que influyen en la AR del Flujo P3 en comparación con los demás Flujos. Los códigos de los flujos bioinformáticos, EPI2ME (P1), EPI2ME_LABS (P2), Spaghetti_modificado (P3), NCBI+ (P4) y Situseq (P5).

4.4. Comparativa entre la secuenciación de Nanopore e Illumina, detección y diversidad de bacterias patógenas transmitidas por alimentos.

Se compararon secuenciaciones de Nanopore e Illumina utilizando 12 muestras de alimentos. Aunque se utilizaron réplicas durante la extracción de ADN y secuenciación, para el análisis se combinaron estas réplicas sumando las OTUs detectadas.

Se obtuvieron 677 45 mil lecturas después de la secuenciación en Nanopore, pero tras el Basecalling (considerando lecturas con un Qscore >10), se analizaron 554.45 mil lecturas válidas en Nanopore y 752.85 mil en Illumina. Las librerías de amplicones de Nanopore tuvieron una longitud media de secuencias de 1421 pb (N50), mientras que las de Illumina tuvieron 440 pb en promedio. La diferencia en la cantidad de datos válidos de secuenciación (Qscore >10 en Nanopore) influyó en las estimaciones de diversidad, mostrando una mayor diversidad en las muestras de Illumina que en las de Nanopore. Esto se evidenció en la Figura 11 del estudio.

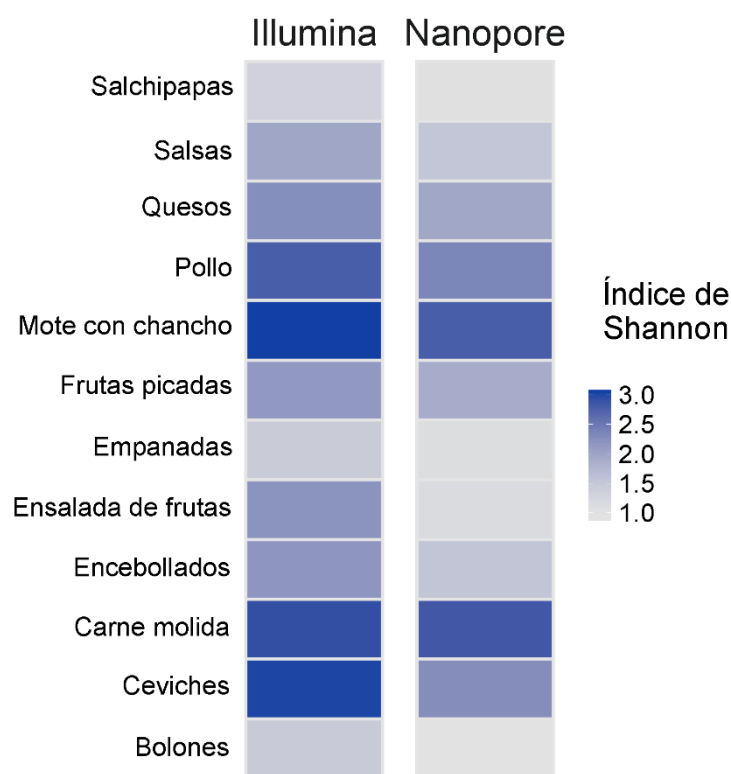


Figura 11. Alfa diversidad de las muestras de alimentos usando el índice de Shannon, se muestra la comparación de la diversidad hallada entre ambas tecnologías de secuenciación Illumina vs Nanopore.

A pesar de las diferencias en la diversidad alfa entre las tecnologías de secuenciación Illumina y Nanopore, el análisis mostró un alto grado de concordancia a nivel de género en el agrupamiento jerárquico de las muestras, con un índice de correlación del 70% (Figura 12). El tanglegrama reveló la similitud entre las muestras de alimentos según los datos genéticos secuenciados por ambas tecnologías. Las muestras de quesos,

frutas, mote con chanco y ceviches exhibieron una alta correlación y similitud de OTUs a nivel de géneros en comparación con otras muestras. El entrelazamiento de 0.07 en el tanglegrama sugiere una similitud significativa entre los árboles de agrupamiento, respaldado por una correlación cercana a 1 (Correlación: 0.7), indicando una concordancia sustancial entre los datos de secuenciación de Illumina y Nanopore. Estos hallazgos coinciden con estudios previos que compararon muestras bacterianas utilizando dendrogramas, destacando la consistencia en la detección bacteriana independientemente de la tecnología de secuenciación utilizada (Golparian et al., 2018).

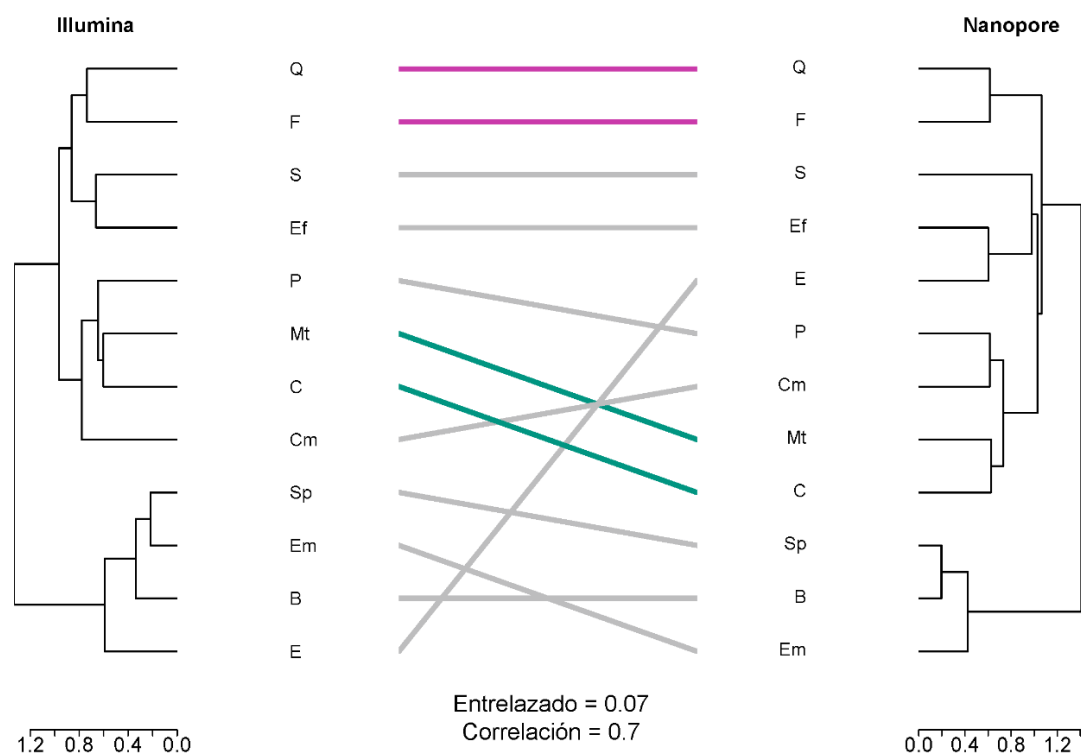


Figura 12. Gráfico de correlaciones que muestran las similitudes de agrupamiento jerárquico (algoritmo de Ward) entre las muestras de alimentos secuenciadas con Illumina y Nanopore. La distancia entre grupos se midió utilizando el índice de Bray-Curtis. Las líneas de colores indican nodos compartidos entre ambos dendrogramas y las reglas horizontales indican las alturas de agrupación. Los códigos de los alimentos se indican: Bolones (B), Ceviches (C), Carne de res (CM), Encebollados (E), Ensalada de frutas (EF), Empanadas de queso (EM), Frutas picadas (F), Mote con chanco (MT), Pollo (P), Quesos (Q), Salchipapas (SP) y Salsas picantes (S).

En promedio, la cantidad de lecturas obtenidas y asignadas taxonómicamente a nivel de género por medio de la secuenciación de Nanopore fue de un 69% mientras que las de Illumina alcanzaron un 92% de clasificación. En la Tabla VIII se observa la cantidad de lecturas de secuenciación obtenidas para cada una de las muestras de alimentos evaluadas según el método de secuenciación.

Tabla VIII. Comparación del número de secuencias de ADN obtenidas según la secuenciación de Nanopore e Illumina, se analizó el número de lecturas que se clasificaron a nivel de género (%) según los flujos de los programas EPI2ME (Nanopore) y 16S Metagenomics (Illumina).

Alimentos	Nanopore (Secuencias ADN)	Illumina (Secuencias ADN)	Secuencias Nanopore asignadas a género	Secuencias Illumina asignadas a género
Bolones	106131	54858	92%	89%
Ceviches	52093	71165	56%	83%
Carne de res	888	62944	86%	96%
Encebollados	10931	62184	54%	94%
Ensalada de frutas	17401	54160	41%	91%
Empanadas de queso	22521	73960	81%	98%
Frutas picadas	17277	69397	56%	86%
Mote con chancho	86103	53056	75%	92%
Pollo	43483	71952	68%	97%
Quesos	95703	56065	72%	90%
Salchipapas	22116	60729	57%	88%
Salsas picantes	74896	62381	90%	99%

En el conjunto de datos de las 12 muestras de alimentos usando Illumina se identificaron 38 filos mientras que en Nanopore solo se identificaron 6 (Figura 13). En promedio los filos más abundantes identificados en los datos de Illumina fueron Firmicutes (56.66%), Proteobacteria (37.00%), Bacteroidetes (2.31%), Cyanobacteria/Chloroplast (1.78%), y Fusobacteria (1.74%). Entre los grupos más abundantes a nivel de género se obtuvieron *Bacillus* (21.75%), Unclassified at genus level (8.333), *Clostridium sensu stricto* (7.14%), *Enterobacter* (6.06%), *Weissella* (5.56%), *Lactococcus* (4.16%), *Enterococcus* (3.26%), *Pediococcus* (3.12%), *Pseudomonas* (2.69%), *Citrobacter* (2.02%), *Shewanella* (1.96%), *Clostridium XI*

(1.94%), *Lactobacillus* (1.89%), *Aeromonas* (1.85%), *Raoultella* (1.84%), *Streptophyta* (1.76%), *Pantoea* (1.71%), *Fusobacterium* (1.69%), *Klebsiella* (1.56%), *Acinetobacter* (1.55%), *Escherichia/Shigella* (1.50%), *Kurthia* (1.26%), *Myroides* (1.14%), *Hafnia* (1.13%), *Bacteroides* (1.09%), *Veillonella* (1.05%).

En Nanopore solo se identificó al filo Firmicutes (56.79%) y Proteobacteria (43.06%) como los más abundantes (Figura 13). Los grupos más abundantes identificados a nivel de género usando esta tecnología fueron *Bacillus* (33.74%), *Enterobacter* (9.96%), *Raoultella* (7.53%), *Lactococcus* (6.17%), *Enterococcus* (5.40%) y *Citrobacter* (5.01%). En menor medida, se encuentran *Clostridium* (2.89%), *Weissella* (2.72%), *Aeromona* (2.46%), *Escherichia* (2.40%), *Klebsiella* (2.28%), *Pseudomonas* (2.10%), *Proteus* (2.07%), *Hafnia-Obesumbacterium* (2.06%), *Kurthia* (2.02%), *Buttiauxella* (1.18%) y *Pantoea* (1.11%).

Estos perfiles bacterianos hallados se asemejan en gran medida a los detectados en el estudio realizado por Salazar-Llorente et al. (2020) donde también se incluyeron varias matrices alimentarias de importancia económica. En su estudio se analizaron las lecturas obtenidas mediante la plataforma de Illumina.

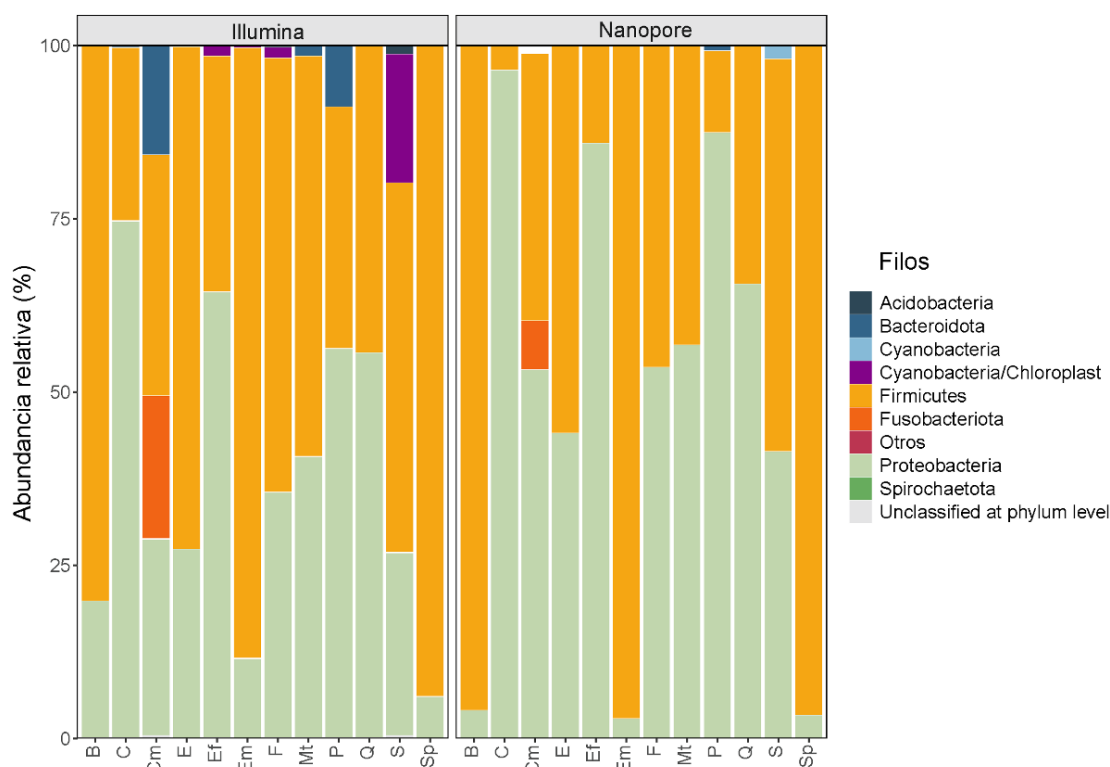


Figura 13. Abundancias relativas de los filos hallados en las muestras de alimentos según la tecnología de secuenciación Illumina y Nanopore. Las abundancias de los filos

menores al 1% se agruparon en la categoría “Otros”. Los códigos de los alimentos se indican: Bolones (B), Ceviches (C), Carne de res (CM), Encebollados (E), Ensalada de frutas (EF), Empanadas de queso (EM), Frutas picadas (F), Mote con chanco (MT), Pollo (P), Quesos (Q), Salchipapas (SP) y Salsas picantes (S).

El perfil bacteriano de las muestras de alimentos no mostró cambios significativos debido a la tecnología de secuenciación, según lo revelado por un análisis multivariado PERMANOVA utilizando el índice de disimilitud de Bray-Curtis. El método de secuenciación explicó solo el 4.3% de la variación en el modelo evaluado, con un p-valor de 0.43. Sin embargo, al considerar las muestras de alimentos como variable, se encontró un cambio significativo en el perfil bacteriano detectado (PERMANOVA: $F = 6.72$, $R^2 = 0.86$, $P = 9.99E-05$), explicando el 86% de la variación. A nivel de género, ambos métodos de secuenciación mostraron una correlación moderada ($r = 0.767$), indicando perfiles bacterianos similares (Figura 14).

La abundancia relativa de algunos géneros varió entre las plataformas, como *Bacillus* y *Citrobacter* siendo más altos en Nanopore, y *Weisella* más alto en Illumina (Figura 15), estas diferencias se atribuyen a la cobertura de las regiones alcanzadas por los cebadores y las bases de datos utilizadas, Nanopore utiliza NCBI mientras el software de Illumina utilizó una base curada de DADA2, SILVA 138 (Camilla Ceccarani, 2023; Hiergeist et al., 2023). Las diferencias de abundancias se detallan en la Tabla suplementaria V.

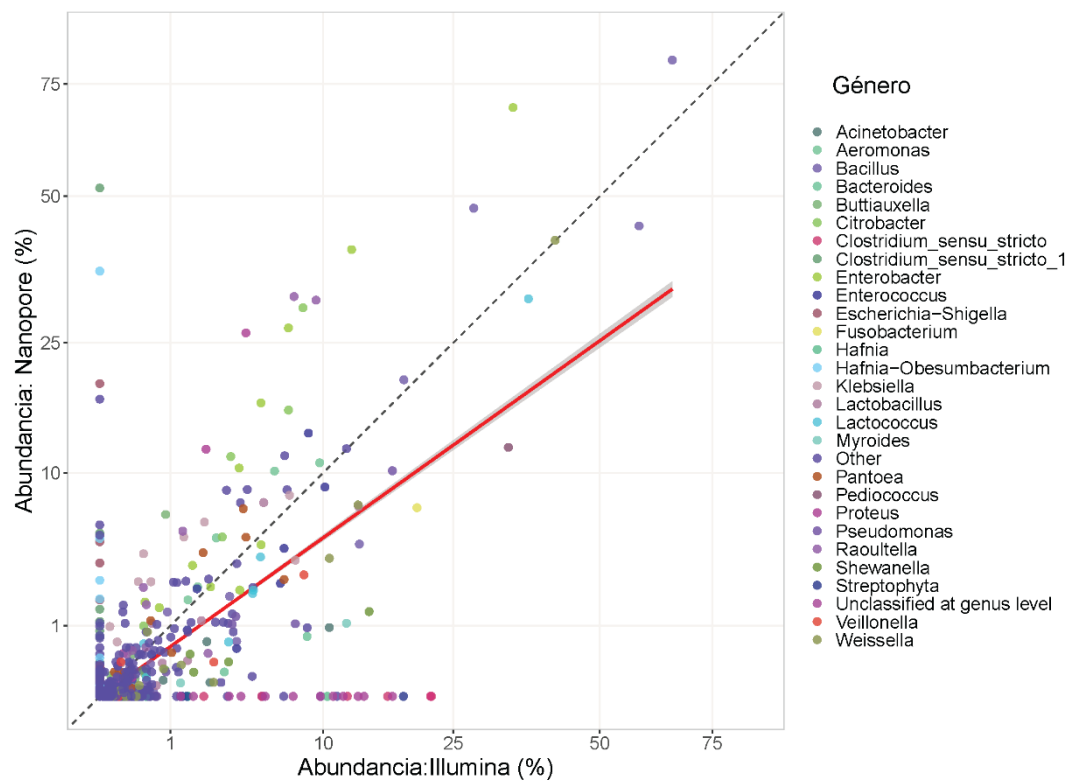


Figura 14. Correlación entre la abundancia de géneros secuenciados con la tecnología Nanopore (eje y) e Illumina (eje x) (Pearson $r = 0,767$). Los géneros abundantes están coloreados, "Unclassified at genus level" representa las secuencias no identificadas a nivel de género y "Other" representa los filos menos abundantes. La línea discontinua negra muestra la relación lineal entre las abundancias de Nanopore e Illumina, y la línea roja muestra una proporción de 1:1. Los ejes se han transformado en raíz cuadrada.

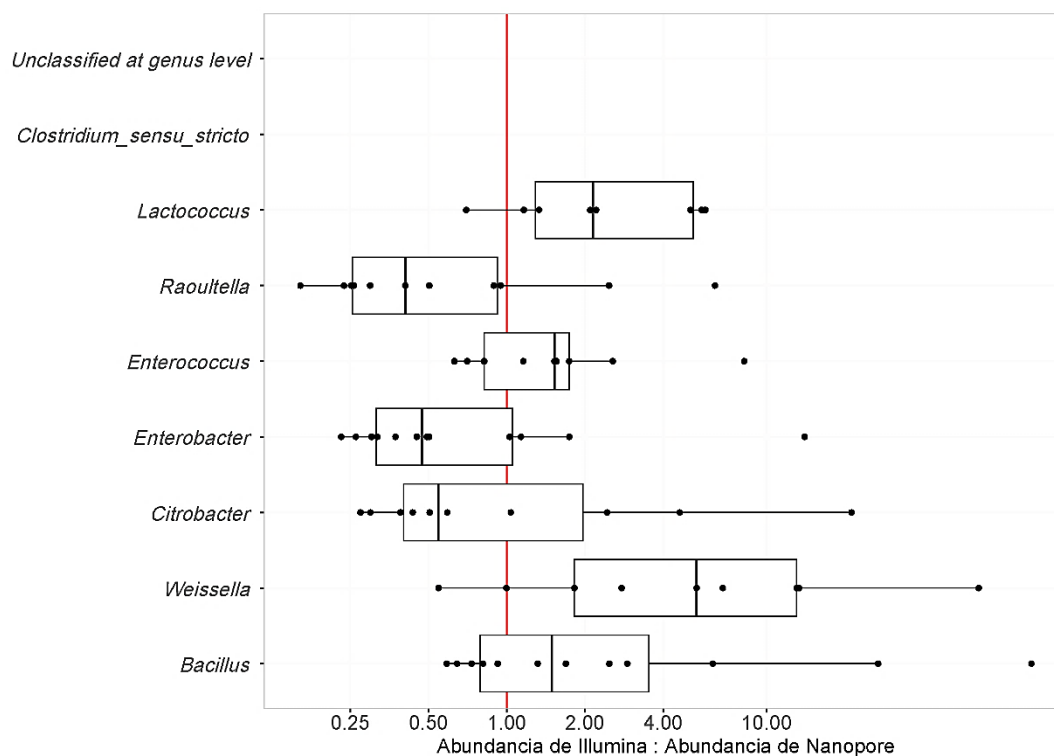


Figura 15. La relación entre la abundancia de Illumina y la abundancia de nanoporos de los géneros más abundantes seleccionados. Las proporciones de muestras individuales se superponen en los diagramas de caja. La línea roja muestra una proporción de 1:1.

Se combinaron los conjuntos de datos de amplicones generados por las tecnologías de secuenciación Nanopore e Illumina a nivel de género en R para evaluar si la elección de la tecnología afecta las conclusiones ecológicas a gran escala en un gráfico NMDS (Figura 16). Los resultados indican que la tecnología de secuenciación no tuvo un efecto estadísticamente significativo en la composición de la comunidad microbiana de las muestras (ANOSIM $p = 0.0233$). Sin embargo, el efecto de las muestras de alimentos sí influyó la variación a gran escala (ANOSIM $R = 0.89$, $p < 1e-04$). Dado que no se han encontrado estudios comparativos de ambas plataformas de secuenciación en diferentes muestras de alimentos simultáneamente, se compararon los datos obtenidos con estudios previos realizados en otras áreas. Estos estudios evidencian y concuerdan con la presente investigación, destacando el uso de la plataforma de Nanopore para la detección de microorganismos en muestras variadas de alimentos, y la falta de diferencia significativa entre ambas tecnologías en la detección de perfiles microbianos similares en las 12 muestras evaluadas. La variable que marcó la diferencia en los perfiles microbianos a nivel de género fue la de los alimentos, lo

cual coincide con hallazgos anteriores en los estudios de de Siqueira et al. (2021) y Zorz et al. (2023).

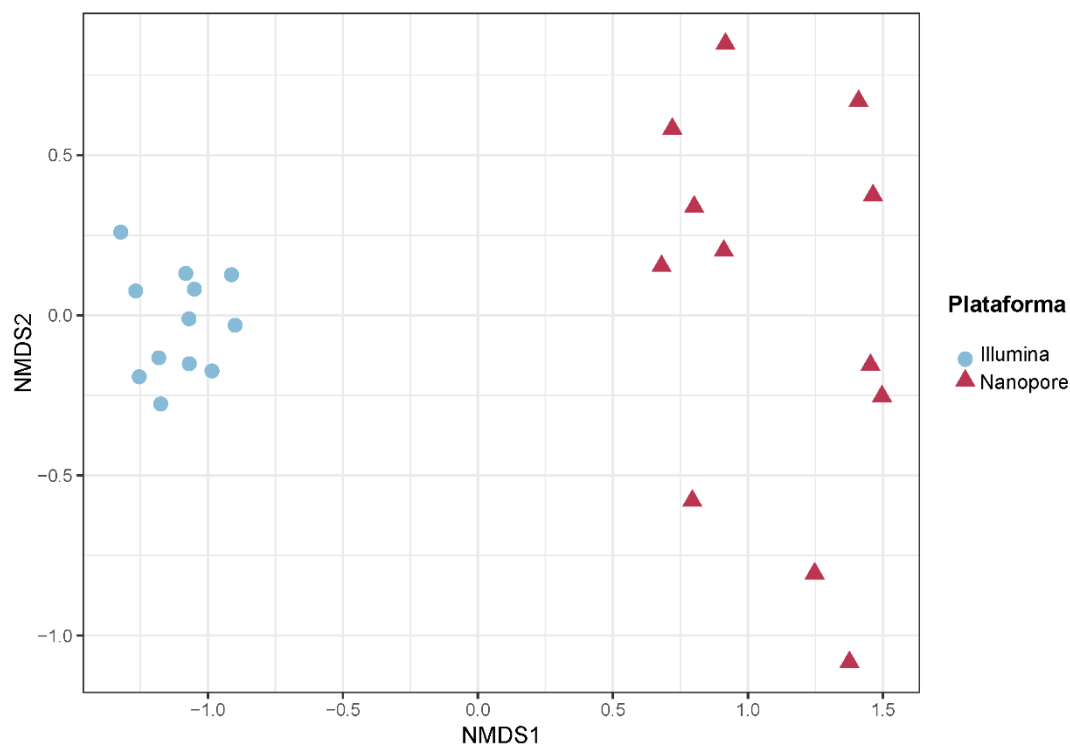


Figura 16. Gráfico NMDS con los conjuntos de datos combinados de Nanopore e Illumina. La disimilitud de Bray-Curtis en los géneros se muestran con puntos coloreados y dimensionados según las plataformas de secuenciación.

El mismo grupo de bacterias patógenas de la Tabla IX a nivel de género fueron filtrados de ambos conjuntos de datos de las plataformas de secuenciación Illumina y Nanopore. La abundancia se consideró en relación al total de lecturas clasificadas de las 12 muestras de alimentos.

Varios géneros bacterianos de este grupo de posibles patógenos fueron detectados (Figura 17) entre estos están *Bacillus* con una AR 21.75% con Illumina y 33.74% con Nanopore. *Campylobacter* y *Listeria* con una AR de 0.001% y 0.012%, respectivamente, solo se detectaron en los datos de Illumina. *Cronobacter* presentó un 0.06% de AR con Illumina y un 0.02% con Nanopore. *Salmonella* tuvo una AR de 0.17% con Illumina, pero no se detectó con Nanopore. *Staphylococcus* fue detectado con una AR de 0.73% usando la tecnología Illumina y un 0.46% usando la tecnología Nanopore. Finalmente, *Yersinia* fue identificada con valores bajos de AR 0.044% y

0.008% en ambas tecnologías. Algunos de estos patógenos también fueron detectados en un estudio similar en el que se realizó un perfil de bacterias incluyendo bacterias patógenas en la plataforma de Illumina, mostrando la consistencia de identificación de estos géneros (Díaz Cárdenas et al., 2024).

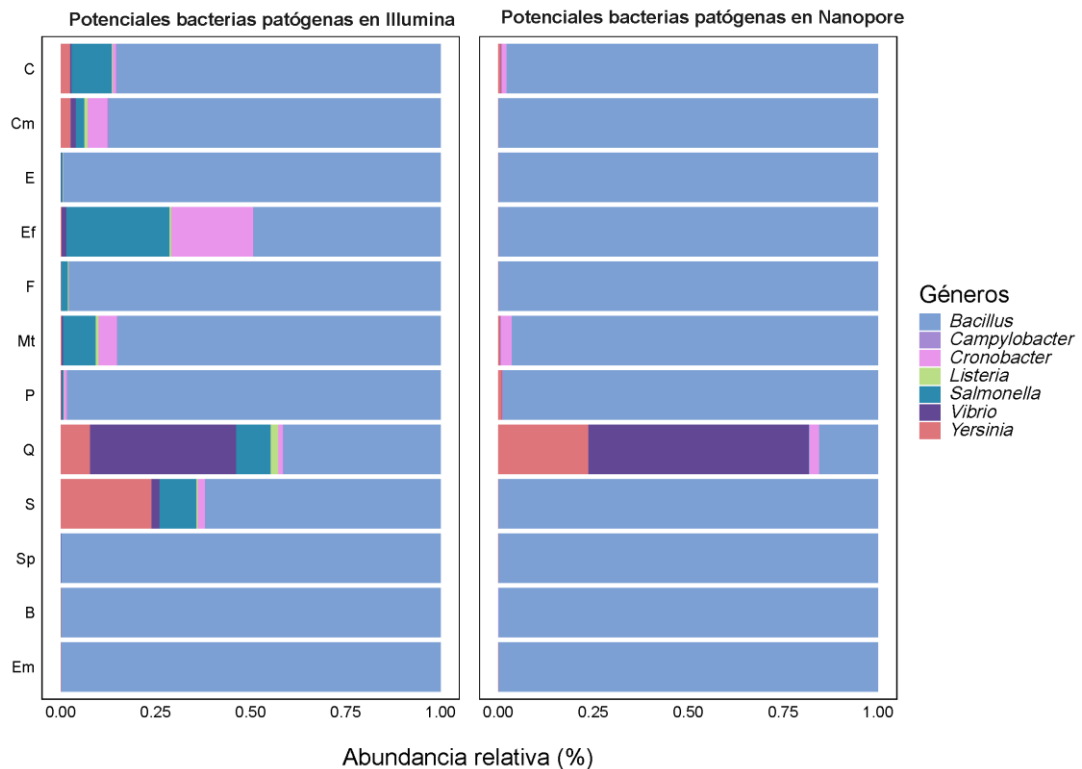


Figura 17. Abundancias de las potenciales bacterias patógenas transmitidas por alimentos en ambas tecnologías de secuenciación, Illumina y Nanopore. Los códigos de los alimentos se indican: Bolones (B), Ceviches (C), Carne de res (CM), Encebollados (E), Ensalada de frutas (EF), Empanadas de queso (EM), Frutas picadas (F), Mote con chanco (MT), Pollo (P), Quesos (Q), Salchipapas (SP) y Salsas picantes (S).

CAPÍTULO 5

CONCLUSIONES Y RECOMENDACIONES

En respuesta al objetivo 1 para la obtención del protocolo de extracción de ADN a partir de diferentes tipos de matrices alimentarias, el protocolo hallado a base de hidróxido de sodio (NaOH) y la purificación con acetato de potasio evidenció que se pudo obtener un rendimiento apropiado a partir de la cantidad inicial de muestra usada (10 ml) y la cantidad de ADN recuperada. Dado a que es un método rápido y de bajo costo, la calidad obtenida en el rango 260/280 se consideró aceptable para la secuenciación en nanoporos y de Illumina.

Para este protocolo de extracción hallado se recomienda probar concentraciones diferentes a 50 mM de NaOH haciendo los ajustes según el tipo de matriz de alimento, para evaluar posibles diferencias en la cantidad y calidad que se obtenga, así como determinar si con estas nuevas modificaciones se podría obtener más ADN de bacterias gran positivas como *L. monocytogenes* que en este estudio no logró ser detectada como especie.

En el caso de procesar muestras con alto contenido de grasas, es recomendable después del paso de separación de sólidos en la centrifugación inicial, filtrar la muestra por decantación a un tubo nuevo usando gasa estéril como filtro de retención de sólidos. Finalmente, durante la extracción se recomienda realizar la elución o resuspensión del pellet de ADN que se obtenga en al menos 100 ul de solución TE o agua bidestilada libre de nucleasas para diluir contaminantes y evitar la concentración innecesaria del ADN obtenido.

En respuesta al objetivo 2, para la preparación de librerías genómicas 16S rRNA que permitan el multiplexeado e identificación de múltiples patógenos de interés como lo fueron *S. enterica*, *L. monocytogenes* y *V. cholerae* en secuenciación por nanoporos. Se concluye que gracias a la dilución inicial que se realiza al momento de empezar la librería, a 10 ng, la influencia de la calidad del ADN sobre la amplificación de las muestras de ADN se ve disminuida permitiendo la obtención de amplicones en todas las muestras. Además, las purificaciones dentro del protocolo de preparación de las librerías con microperlas magnéticas, permiten mejorar la calidad de los amplicones obtenidos.

El multiplexeado y detección de *S. enterica* hasta el nivel de especie y a nivel de género con *Listeria* y *Vibrio* de las muestras se pudo lograr de manera satisfactoria debido al proceso de diluciones utilizadas en la que se consideró equiparar las cantidades de ADN

en femtomoles (Entre 50 a 100 fmol) según el tamaño de fragmento a secuenciar (16S rARN), para de esta manera secuenciar amplicones en cantidades equivalentes molarmente según la muestra y se evitó la sobre secuenciación de muestras con mayor cantidad de ADN.

Se recomienda evaluar diferentes configuraciones en cantidad inicial de ADN y reactivos en el coctel de PCR, puesto que el protocolo de Nanopore pide la preparación de reacciones de 50 ul, siendo esto innecesario debido a que en el *pooling* de las muestras que se realiza en las siguientes etapas de preparación solo se utiliza 10 µl de dicha mezcla. Además, se recomienda el ajuste del programa de PCR cuando se utiliza cantidades iniciales menores a 5 ng de ADN.

En búsqueda de la obtención de un flujo bioinformático adecuado para la detección de *S. enterica*, *L. monocytogenes* y *V. cholerae*. Se pudo evidenciar que la selección de los flujos se puede dar según las condiciones físicas del estudio y la pregunta de investigación, para estudios donde se cuente con acceso online y escasos recursos de hardware, el flujo de EPI2ME es el más adecuado. Para los casos en donde no se pueda acceder a la web y no se tenga recursos de hardware, el flujo de NCBI+ y Spaghetty modificado son los más adecuados.

Se recomienda realizar ciertos ajustes en cuanto al puntaje de lecturas válidas a los programas de Spaguetty modificado para disminuir la rigurosidad al momento de filtrar los datos y pasar las lecturas válidas. También, es deseable trabajar al momento de la secuenciación con el basecalling apagado. Adicionalmente, para la ejecución de los flujos seleccionados es recomendable que la laptop posea un núcleo corei7 o similar, memoria RAM de al menos 16GB y 1 terabyte de almacenamiento.

Finalmente, en la comparación de las plataformas de secuenciación Illumina y Nanopore con el fin de evidenciar posibles diferencias en la obtención de perfiles microbianos en muestras de alimentos, se pudo concluir que los taxones a nivel de género detectados en Nanopore no difieren significativamente de los detectados en la secuenciación de Illumina, la correlación de la diversidad bacteriana hallada es similar. Se concluye que la secuenciación de ADN bacteriano de alimentos es una tecnología válida y costo efectiva para la detección de bacterias de interés económico incluyendo bacterias patógenas transmitidas por alimentos.

Se recomienda evaluar la disminución del tiempo de enriquecimiento de las muestras previo la extracción del ADN, así como el tiempo de secuenciación para determinar si la detección de las bacterias objetivo se mantiene.

BIBLIOGRAFÍA

- Administración de Alimentos y Medicamentos (FDA). (2023). *Foodborne Pathogens. Outbreaks of Foodborne Illness*. <https://www.fda.gov/food/outbreaks-foodborne-illness/foodborne-pathogens>
- AGROCALIDAD. (2020, June 4). *LA INOCUIDAD DE ALIMENTOS, UN ASUNTO DE TODOS - AGROCALIDAD*. <https://www.agrocalidad.gob.ec/la-inocuidad-de-alimentos-un-asunto-de-todos/>
- Baraketi, A., Salmieri, S., & Lacroix, M. (2018). Foodborne Pathogens Detection: Persevering Worldwide Challenge. In *Biosensing Technologies for the Detection of Pathogens - A Prospective Way for Rapid Analysis: Vol. i* (Issue tourism, p. 13). InTech. <https://doi.org/10.5772/intechopen.74421>
- Barbosa, C., Nogueira, S., Gadanho, M., & Chaves, S. (2016). DNA extraction: Finding the most suitable method. In *Molecular Microbial Diagnostic Methods: Pathways to Implementation for the Food and Water Industries*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-416999-9.00007-1>
- Camilla Ceccarani, M. S. (2023). A comparison between Greengenes, SILVA, RDP, and NCBI reference databases in four published microbiota datasets. *BioRxiv*. <https://doi.org/https://doi.org/10.1101/2023.04.12.535864>
- Centers for Disease Control and Prevention, C. (2022). *Steps in a Foodborne Outbreak Investigation | CDC*. <https://www.cdc.gov/foodsafety/outbreaks/steps/index.html>
- Cevallos-Cevallos, J. M., Morales, M. F., & Salazar-Llorente, E. (2019). Red interuniversitaria de inocuidad alimentaria del Ecuador. Resultados de contaminación microbiana de alimentos de consumo masivo en Quito, Guayaquil y Cuenca. In J. M. Cevallos, J. Ruales, J. Ortiz, & P. Maldonado (Eds.), *Memorias del Seminario Taller: Avances en Biociencias e inocuidad alimentaria en Ecuador* (pp. 12–15). Vllir Network Ecuador. <http://drive.google.com/file/d/1CjcAFFvIOMDshpsGEoEWjJjaBTQbWua9/view?usp=sharing>
- Chen, J., Liu, X., Chen, J., Guo, Z., Wang, Y., Chen, G., Chen, X., Yan, Q., Yang, P., Li, R., Yang, G., Lan, Q., & Wang, J. (2019). Development of a Rapid Test Method for Salmonella enterica Detection Based on Fluorescence Probe-Based Recombinase Polymerase Amplification. *Food Analytical Methods*, 12(8), 1791–1798. <https://doi.org/10.1007/s12161-019-01526-3>
- de Siqueira, G. M. V., Pereira-dos-Santos, F. M., Silva-Rocha, R., & Guazzaroni, M. E. (2021). Nanopore Sequencing Provides Rapid and Reliable Insight Into Microbial Profiles of Intensive Care Units. *Frontiers in Public Health*, 9(August), 1–11. <https://doi.org/10.3389/fpubh.2021.710985>

- Díaz Cárdenas, B., Salazar Llorente, E., Gu, G., Nou, X., Ortiz, J., Maldonado, P., & Cevallos-Cevallos, J. M. (2024). Microbial Composition and Diversity of High-demand Street-vended Foods in Ecuador. *Journal of Food Protection*, 87(4), 100247. <https://doi.org/10.1016/j.jfp.2024.100247>
- Dilhari, A., Sampath, A., Gunasekara, C., Fernando, N., Weerasekera, D., Sissons, C., McBain, A., & Weerasekera, M. (2017). Evaluation of the impact of six different DNA extraction methods for the representation of the microbial community associated with human chronic wound infections using a gel-based DNA profiling method. *AMB Express*, 7(1). <https://doi.org/10.1186/s13568-017-0477-z>
- Doyle, M. P., Steenson, L. R., & Meng, J. (2013). Bacteria in Food and Beverage Production. In *The Prokaryotes* (pp. 241–256). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31331-8_27
- Dugat-Bony, E., Lossouarn, J., De Paepe, M., Sarthou, A.-S., Fedala, Y., Petit, M.-A., & Chaillou, S. (2020). Viral metagenomic analysis of the cheese surface: A comparative study of rapid procedures for extracting viral particles. *Food Microbiology*, 85, 103278. <https://doi.org/10.1016/j.fm.2019.103278>
- El-Zamkan, M. A., Ahmed, A. S., Abdelhafeez, H. H., & Mohamed, H. M. A. (2023). Molecular characterization of *Vibrio* species isolated from dairy and water samples. *Scientific Reports*, 13(1), 15368. <https://doi.org/10.1038/s41598-023-42334-4>
- Elie, C., Perret, M., Hage, H., Sentausa, E., Hesketh, A., Louis, K., Fritah-Lafont, A., Leissner, P., Vachon, C., Rostaing, H., Reynier, F., Gervasi, G., & Saliou, A. (2023). Comparison of DNA extraction methods for 16S rRNA gene sequencing in the analysis of the human gut microbiome. *Scientific Reports*, 13(1), 10279. <https://doi.org/10.1038/s41598-023-33959-6>
- Elizaquível, P., & Aznar, R. (2008). Comparison of four commercial DNA extraction kits for PCR detection of *Listeria monocytogenes*, *Salmonella*, *Escherichia coli* O157:H7, and *Staphylococcus aureus* in fresh, minimally processed vegetables. *Journal of Food Protection*, 71(10), 2110–2114. <https://doi.org/10.4315/0362-028X-71.10.2110>
- European Food Safety Authority. (2021). The European Union One Health 2020 Zoonoses Report. *EFSA Journal*, 19(12). <https://doi.org/10.2903/j.efsa.2021.6971>
- FDA. (2020). *Foodborne Pathogens* | FDA. <https://www.fda.gov/food/outbreaks-foodborne-illness/foodborne-pathogens>
- Foddai, A. C. G., & Grant, I. R. (2020). Methods for detection of viable foodborne pathogens: current state-of-art and future prospects. *Applied Microbiology and Biotechnology*, 104(10), 4281–4288. <https://doi.org/10.1007/s00253-020-10542-x>

- Forghani, F., Li, S., Zhang, S., Mann, D. A., Deng, X., den Bakker, H. C., & Diez-Gonzalez, F. (2020). Salmonella enterica and Escherichia coli in Wheat Flour: Detection and Serotyping by a Quasimetagenomic Approach Assisted by Magnetic Capture, Multiple-Displacement Amplification, and Real-Time Sequencing. *Applied and Environmental Microbiology*, 86(13). <https://doi.org/10.1128/AEM.00097-20>
- Galili, T. (2015). dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*, 31(22), 3718–3720. <https://doi.org/10.1093/BIOINFORMATICS/BTV428>
- Golparian, D., Donà, V., Sánchez-Busó, L., Foerster, S., Harris, S., Endimiani, A., Low, N., & Unemo, M. (2018). Antimicrobial resistance prediction and phylogenetic analysis of Neisseria gonorrhoeae isolates using the Oxford Nanopore MinION sequencer. *Scientific Reports*, 8(1), 17596. <https://doi.org/10.1038/s41598-018-35750-4>
- Grinevich, D., Harden, L., Thakur, S., & Callahan, B. (2024). Serovar-level identification of bacterial foodborne pathogens from full-length 16S rRNA gene sequencing. *MSystems*. <https://doi.org/10.1128/msystems.00757-23>
- Gu, G., Ottesen, A., Bolten, S., Ramachandran, P., Reed, E., Rideout, S., Luo, Y., Patel, J., Brown, E., & Nou, X. (2018). Shifts in spinach microbial communities after chlorine washing and storage at compliant and abusive temperatures. *Food Microbiology*, 73, 73–84. <https://doi.org/10.1016/j.fm.2018.01.002>
- Heller, L. C., Davis, C. R., Peak, K. K., Wingfield, D., Cannons, A. C., Amuso, P. T., & Cattani, J. (2003). Comparison of Methods for DNA Isolation from Food Samples for Detection of Shiga Toxin-Producing Escherichia coli by Real-Time PCR. *Applied and Environmental Microbiology*, 69(3), 1844–1846. <https://doi.org/10.1128/AEM.69.3.1844-1846.2003>
- Hiergeist, A., Ruelle, J., Emler, S., & Gessner, A. (2023). Reliability of species detection in 16S microbiome analysis: Comparison of five widely used pipelines and recommendations for a more standardized approach. *PLoS ONE*, 18(2 February), 1–16. <https://doi.org/10.1371/journal.pone.0280870>
- Hyeon, J., Li, S., Mann, D. A., Zhang, S., Li, Z., Chen, Y., & Deng, X. (2017). Quasimetagenomics-Based and Real-Time-Sequencing-Aided Detection and Subtyping of Salmonella enterica from Food Samples. *Applied and Environmental Microbiology*, 84(4), 1–15. <https://doi.org/10.1128/AEM.02340-17>
- Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Sóllymos, M. Henry H. Stevens, H. W. (2022). *vegan: Community Ecology Package*. <http://cran.r->

project.org/package=vegan

- Johnson, J. S., Spakowicz, D. J., Hong, B. Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., & Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, *10*(1), 1–11. <https://doi.org/10.1038/s41467-019-13036-1>
- Kobayashi, H., Kubota, J., Fujihara, K., Honjoh, K., Iio, M., Fujiki, N., Nakabe, M., Oda, S., Satoyama, T., Takasu, K., Nakanishi, H., & Miyamoto, T. (2009). Simultaneous Enrichment of Salmonella spp, Escherichia coli O157:H7, Vibrio parahaemolyticus, Staphylococcus aureus, Bacillus cereus, and Listeria monocytogenes by Single Broth and Screening of the Pathogens by Multiplex Real-time PCR. *Food Science and Technology Research*, *15*(4), 427–438. <https://doi.org/https://doi.org/10.3136/fstr.15.427>
- Kovac, J. (2019). Precision Food Safety: a Paradigm Shift in Detection and Control of Foodborne Pathogens. *MSystems*, *4*(3), 1–5. <https://doi.org/10.1128/msystems.00164-19>
- Kovac, J., Cummings, K. J., Rodriguez-Rivera, L. D., Carroll, L. M., Thachil, A., & Wiedmann, M. (2017). Temporal genomic phylogeny reconstruction indicates a geospatial transmission path of Salmonella cerro in the United States and a clade-specific loss of hydrogen sulfide production. *Frontiers in Microbiology*, *8*(MAY), 737. <https://doi.org/10.3389/FMICB.2017.00737/BIBTEX>
- Lakicevic, B., Jankovic, V., Pietzka, A., & Ruppitsch, W. (2023). Wholegenome sequencing as the gold standard approach for control of Listeria monocytogenes in the food chain. *Journal of Food Protection*, *86*(1), 100003. <https://doi.org/10.1016/j.jfp.2022.10.002>
- Latorre-Pérez, A., Gimeno-Valero, H., Tanner, K., Pascual, J., Vilanova, C., & Porcar, M. (2021). A Round Trip to the Desert: In situ Nanopore Sequencing Informs Targeted Bioprospecting. *Frontiers in Microbiology*, *12*(December). <https://doi.org/10.3389/fmicb.2021.768240>
- Lewis, E., Hudson, J. A., Cook, N., Barnes, J. D., & Haynes, E. (2020). Next-generation sequencing as a screening tool for foodborne pathogens in fresh produce. *Journal of Microbiological Methods*, *171*, 105840. <https://doi.org/10.1016/j.mimet.2020.105840>
- Liu, C., Cui, Y., Li, X., & Yao, M. (2021). Microeco: An R package for data mining in microbial community ecology. *FEMS Microbiology Ecology*, *97*(2). <https://doi.org/10.1093/femsec/fiaa255>
- Liu, D. (2009). Molecular Detection of Foodborne Pathogens. In D. Liu (Ed.), *Taylor & Francis Group (First)*. CRC Press. <https://doi.org/10.1201/9781420076448>

- Liu, Y., Cao, Y., Wang, T., Dong, Q., Li, J., & Niu, C. (2019). Detection of 12 Common Food-Borne Bacterial Pathogens by TaqMan Real-Time PCR Using a Single Set of Reaction Conditions. *Frontiers in Microbiology*, *10*(FEB), 1–9. <https://doi.org/10.3389/fmicb.2019.00222>
- Lobiuc, A., Pavál, N.-E., Dimian, M., & Covașă, M. (2023). Nanopore Sequencing Assessment of Bacterial Pathogens and Associated Antibiotic Resistance Genes in Environmental Samples. *Microorganisms*, *11*(12), 2834. <https://doi.org/10.3390/microorganisms11122834>
- Maksimov, P., Schares, G., Press, S., Fröhlich, A., Basso, W., Herzig, M., & Conraths, F. J. (2017). Comparison of different commercial DNA extraction kits and PCR protocols for the detection of *Echinococcus multilocularis* eggs in faecal samples from foxes. *Veterinary Parasitology*, *237*, 83–93. <https://doi.org/10.1016/J.VETPAR.2017.02.015>
- Markusková, B., Minarovičová, J., Véghová, A., Drahovská, H., & Kaclíková, E. (2021). Impact of DNA extraction methods on 16S rRNA-based profiling of bacterial communities in cheese. *Journal of Microbiological Methods*, *184*, 106210. <https://doi.org/10.1016/j.mimet.2021.106210>
- McGaughey, K. D., Yilmaz-Swenson, T., Elsayed, N. M., Cruz, D. A., Rodriguez, R. R., Kritzer, M. D., Peterchev, A. V., Gray, M., Lewis, S. R., Roach, J., Wetsel, W. C., & Williamson, D. E. (2018). Comparative evaluation of a new magnetic bead-based DNA extraction method from fecal samples for downstream next-generation 16S rRNA gene sequencing. *PLOS ONE*, *13*(8), e0202858. <https://doi.org/10.1371/journal.pone.0202858>
- McLaren, M. R., & Callahan, B. J. (2021). *Silva 138.1 prokaryotic SSU taxonomic training data formatted for DADA2*. <https://doi.org/10.5281/ZENODO.4587955>
- McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, *8*(4), e61217. <https://doi.org/10.1371/journal.pone.0061217>
- Mira Miralles, M., Maestre-Carballa, L., Lluesma-Gomez, M., & Martinez-Garcia, M. (2019). High-throughput 16s rRNA sequencing to assess potentially active bacteria and foodborne pathogens: a case example in ready-to-eat food. *Foods*, *8*(10), 480. <https://doi.org/10.3390/foods8100480>
- Molano, L.-A. G., Vega-Abellaneda, S., & Manichanh, C. (2024). GSR-DB: a manually curated and optimized taxonomical database for 16S rRNA amplicon analysis. *MSystems*, *9*(2). <https://doi.org/10.1128/msystems.00950-23>
- Organización de las Naciones Unidas para la Alimentación y la Agricultura, & Organización Mundial de la Salud. (2004). *Codex Alimentarius- Directrices*. <https://www.fao.org/fao-who-codexalimentarius/sh->

proxy/es/?lnk=1&url=https%253A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252Fstandards%252FCXG%2B50-2004%252FCXG_050s.pdf

- Piskata, Z., Servusova, E., Babak, V., Nesvadbova, M., & Borilova, G. (2019). The Quality of DNA Isolated from Processed Food and Feed via Different Extraction Procedures. *Molecules*, *24*(6), 1188. <https://doi.org/10.3390/molecules24061188>
- Quick, J., Ashton, P., Calus, S., Chatt, C., Gossain, S., Hawker, J., Nair, S., Neal, K., Nye, K., Peters, T., De Pinna, E., Robinson, E., Struthers, K., Webber, M., Catto, A., Dallman, T. J., Hawkey, P., & Loman, N. J. (2015). Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. *Genome Biology*, *16*(1), 1–14. <https://doi.org/10.1186/s13059-015-0677-2>
- Rantsiou, K., Kathariou, S., Winkler, A., Skandamis, P., Saint-Cyr, M. J., Rouzeau-Szynalski, K., & Amézquita, A. (2018a). Next generation microbiological risk assessment: opportunities of whole genome sequencing (WGS) for foodborne pathogen surveillance, source tracking and risk assessment. *International Journal of Food Microbiology*, *287*(March 2017), 3–9. <https://doi.org/10.1016/j.ijfoodmicro.2017.11.007>
- Rantsiou, K., Kathariou, S., Winkler, A., Skandamis, P., Saint-Cyr, M. J., Rouzeau-Szynalski, K., & Amézquita, A. (2018b). Next generation microbiological risk assessment: opportunities of whole genome sequencing (WGS) for foodborne pathogen surveillance, source tracking and risk assessment. *International Journal of Food Microbiology*, *287*, 3–9. <https://doi.org/10.1016/j.ijfoodmicro.2017.11.007>
- Salazar-Llorente, E., Morales, M., Sornoza, I., Mariduena-Zavala, M., Gu, G., Nou, X., Ortiz, J., Maldonado-Alvarado, P., & Cevallos-Cevallos, J. (2020). Microbiological quality of high-demand foods from three major cities in Ecuador. *Journal of Food Protection*. <https://doi.org/10.4315/jfp-20-271>
- Santos, A., van Aerle, R., Barrientos, L., & Martinez-Urtaza, J. (2020). Computational methods for 16S metabarcoding studies using Nanopore sequencing data. *Computational and Structural Biotechnology Journal*, *18*, 296–305. <https://doi.org/10.1016/j.csbj.2020.01.005>
- Schmidt, R. H. (2017). A Reference Guide for Foodborne Pathogens. *Fshn05-16, February*, 1–5. <http://edis.ifas.ufl.edu>
- Shwani, A., Zuo, B., Alrubaye, A., Zhao, J., & Rhoads, D. D. (2023). A Simple, Inexpensive Alkaline Method for Bacterial DNA Extraction from Environmental Samples for PCR Surveillance and Microbiome Analyses. *Applied Sciences*, *14*(1), 141. <https://doi.org/10.3390/app14010141>
- Sun, Z., Peng, Y., Zhang, M., Wang, K., Bai, J., Li, X., Ning, B., & Gao, Z. (2014). Simultaneous and highly sensitive detection of six different foodborne pathogens

- by high-throughput suspension array technology. *Food Control*, 40(1), 300–309. <https://doi.org/10.1016/j.foodcont.2013.11.028>
- Tatsika, S., Karamanoli, K., Karayanni, H., & Genitsaris, S. (2019). Metagenomic Characterization of Bacterial Communities on Ready-to-Eat Vegetables and effects of Household Washing on their Diversity and Composition. *Pathogens*, 8(1). <https://doi.org/10.3390/pathogens8010037>
- Teng, F., Darveekaran Nair, S. S., Zhu, P., Li, S., Huang, S., Li, X., Xu, J., & Yang, F. (2018). Impact of DNA extraction method and targeted 16S-rRNA hypervariable region on oral microbiota profiling. *Scientific Reports*, 8(1), 16321. <https://doi.org/10.1038/s41598-018-34294-x>
- Teunisse, G. M. (2022). *Fantaxtic - Nested Bar Plots for Phyloseq Data*. Github. <https://github.com/gmteunisse/Fantaxtic>
- U.S. Food & Drug Administration. (2003). *BAM: Food Sampling/Preparation of Sample Homogenate | FDA*. <https://www.fda.gov/food/laboratory-methods-food/bam-food-samplingpreparation-sample-homogenate>
- U.S. Food & Drug Administration. (2018). *Los 14 patógenos principales transmitidos por los alimentos de Seguridad alimentaria para futuras mamás | FDA*. <https://www.fda.gov/food/people-risk-foodborne-illness/los-14-patogenos-principales-transmitidos-por-los-alimentos-de-seguridad-alimentaria-para-futuras>
- U.S. Food & Drug Administration. (2020). *Food Safety: Importance for At-Risk Groups | FDA*. <https://www.fda.gov/food/people-risk-foodborne-illness/food-safety-importance-risk-groups>
- Vidic, J., Vizzini, P., Manzano, M., Kavanaugh, D., Ramarao, N., Zivkovic, M., Radonic, V., Knezevic, N., Giouroudi, I., & Gadjanski, I. (2019). Point-of-need DNA testing for detection of foodborne pathogenic bacteria. *Sensors (Switzerland)*, 19(5). <https://doi.org/10.3390/s19051100>
- Villamizar-Rodríguez, G., Fernández, J., Marín, L., Muñoz, J., González, I., & Lombó, F. (2015). Multiplex detection of nine food-borne pathogens by mPCR and capillary electrophoresis after using a universal pre-enrichment medium. *Frontiers in Microbiology*, 6(NOV), 1–16. <https://doi.org/10.3389/fmicb.2015.01194>
- Villette, R., Autaa, G., Hind, S., Holm, J. B., Moreno-Sabater, A., & Larsen, M. (2021). Refinement of 16S rRNA gene analysis for low biomass biospecimens. *Scientific Reports*, 11(1), 10741. <https://doi.org/10.1038/s41598-021-90226-2>
- Vinayaka, A. C., Ngo, T. A., Kant, K., Engelsmann, P., Dave, V. P., Shahbazi, M. A., Wolff, A., & Bang, D. D. (2019). Rapid detection of *Salmonella enterica* in food

- samples by a novel approach with combination of sample concentration and direct PCR. *Biosensors and Bioelectronics*, 129, 224–230. <https://doi.org/10.1016/j.bios.2018.09.078>
- Voorhuijzen-Harink, M. M., Hagelaar, R., van Dijk, J. P., Prins, T. W., Kok, E. J., & Staats, M. (2019). Toward on-site food authentication using nanopore sequencing. *Food Chemistry: X*, 2, 100035. <https://doi.org/10.1016/j.fochx.2019.100035>
- Wagner, E., Fagerlund, A., Langsrud, S., Møretrø, T., Jensen, M. R., & Moen, B. (2021). Surveillance of *Listeria monocytogenes*: Early Detection, Population Dynamics, and Quasimetagenomic Sequencing during Selective Enrichment. *Applied and Environmental Microbiology*, 87(24). <https://doi.org/10.1128/AEM.01774-21>
- Wei, X., & Zhao, X. (2021). Advances in typing and identification of foodborne pathogens. *Current Opinion in Food Science*, 37, 52–57. <https://doi.org/10.1016/j.cofs.2020.09.002>
- Witte, A. K., Leeb, C., Piniør, B., Mester, P., Fister, S., Bobal, M., Schoder, D., & Rossmannith, P. (2018). Influence of sampling and DNA extraction on 16S rRNA gene amplicon sequencing - Comparison of the bacterial community between two food processing plants. *LWT*, 90, 186–192. <https://doi.org/10.1016/j.lwt.2017.12.027>
- World Health Organization. (2008). Foodborne disease outbreaks: guidelines for investigation and control. In *guidelines for investigation and control*. <https://apps.who.int/iris/handle/10665/43771>
- Wu, X., Luo, H., Ge, C., Xu, F., Deng, X., Wiedmann, M., Baker, R. C., Stevenson, A. E., Zhang, G., & Tang, S. (2023). Evaluation of multiplex nanopore sequencing for *Salmonella* serotype prediction and antimicrobial resistance gene and virulence gene detection. *Frontiers in Microbiology*, 13. <https://doi.org/10.3389/fmicb.2022.1073057>
- Xin, Z., & Chen, J. (2012). A high throughput DNA extraction method with high yield and quality. *Plant Methods*, 8(1), 26. <https://doi.org/10.1186/1746-4811-8-26>
- Xu, F., Ge, C., Luo, H., Li, S., Wiedmann, M., Deng, X., Zhang, G., Stevenson, A., Baker, R. C., & Tang, S. (2020). Evaluation of real-time nanopore sequencing for *Salmonella* serotype prediction. *Food Microbiology*, 89, 103452. <https://doi.org/10.1016/j.fm.2020.103452>
- Zorz, J., Li, C., Chakraborty, A., Gittins, D., Surcon, T., Morrison, N., Bennett, R., Macdonald, A., & Hubert, C. R. J. (2023). SituSeq: An offline protocol for rapid and remote Nanopore amplicon sequence analysis. *Geological Survey of Canada-Atlantic, March*, 1–11. <https://doi.org/10.1038/s43705-023-00239-3>

ANEXOS

Figuras Suplementarias

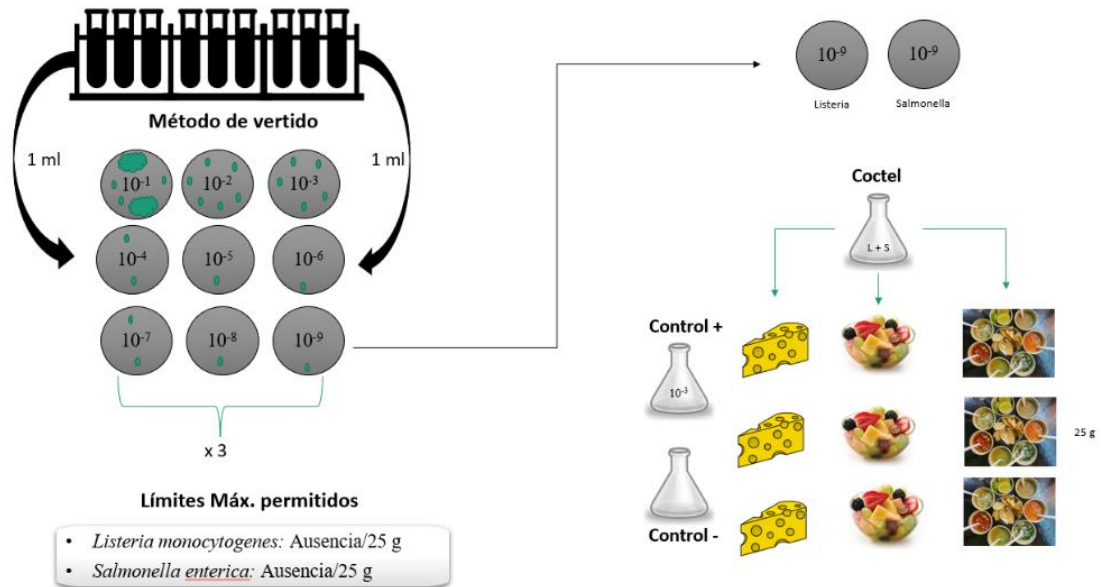


Figura Suplementaria 1. Diseño del experimento de inoculación de bacterias patógenas para las pruebas de sensibilidad de detección del secuenciador minION.

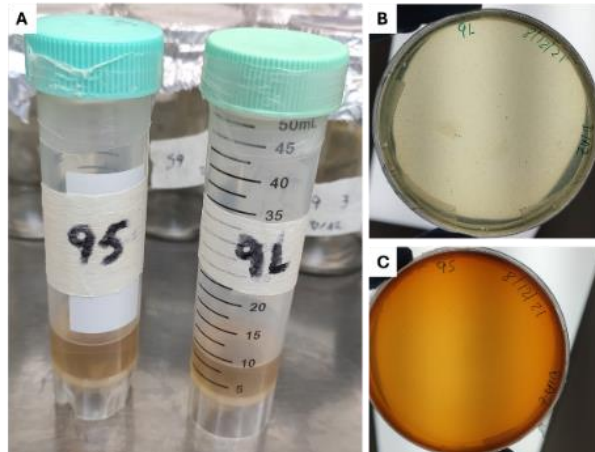


Figura Suplementaria 2. Evaluación del crecimiento bacteriano de *S. enterica* y *L. monocytogenes* luego de 24 horas de incubación en las muestras de alimentos Q, S y F. En la Figura A, se observó el caldo recolectado de cada bacteria inoculada de la muestra de salsas. En las Figuras B y C, se muestra el crecimiento en placa de *Listeria* y *Salmonella*, respectivamente.

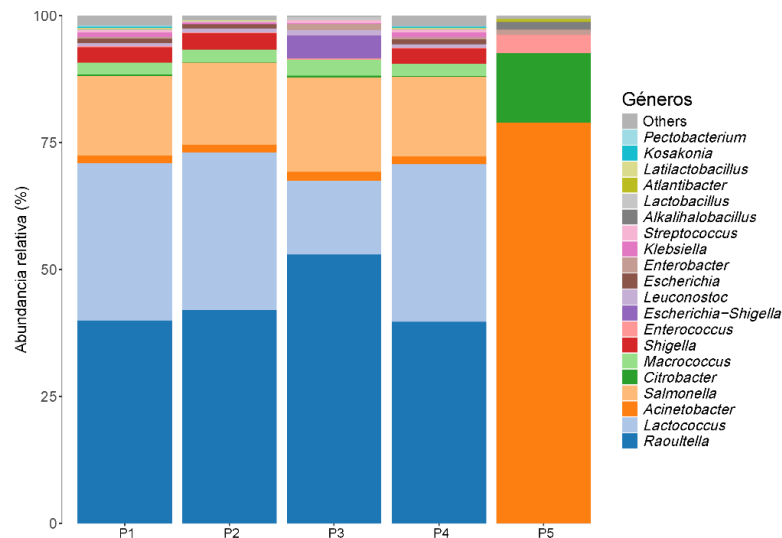


Figura Suplementaria 3. Abundancias relativas de los 20 géneros más abundantes de los distintos flujos bioinformáticos evaluados. Los códigos de los flujos bioinformáticos, EPI2ME (P1), EPI2ME_LABS (P2), Spaghetti_modificado (P3), NCBI+ (P4) y Situseq (P5).

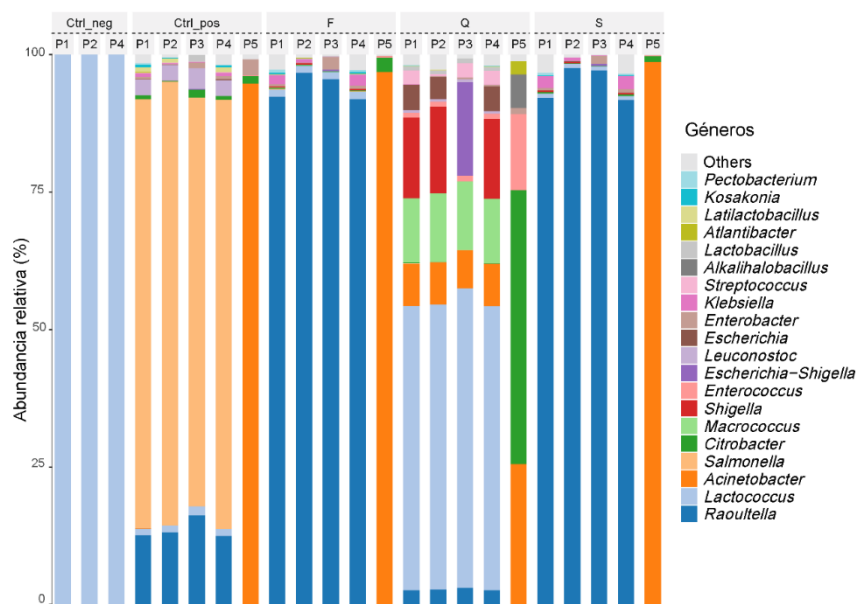


Figura Suplementaria 4. Abundancias relativas de los 20 géneros más abundantes en las muestras de alimentos evaluadas según el flujo bioinformático utilizado. Los códigos por tipos de alimentos son: Queso (Q), Control positivo (Ctrl_pos), Frutas (F), salsas (S) y control negativo (Ctrl_neg). De los flujos bioinformáticos, EPI2ME (P1), EPI2ME_LABS (P2), Spaghetti_modificado (P3), NCBI+ (P4) y Situseq (P5).

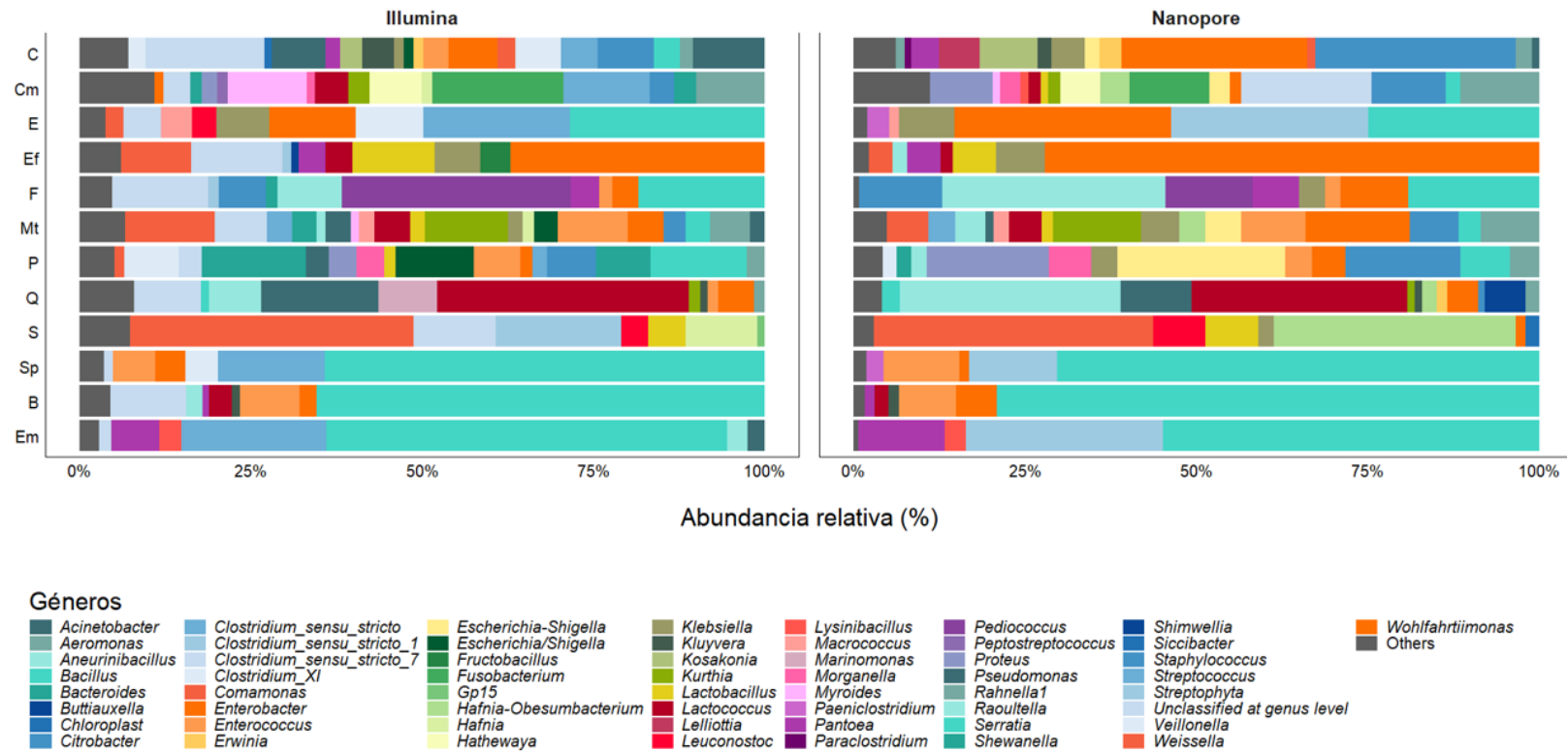


Figura Suplementaria 5. Abundancias relativas de los géneros más abundantes en las muestras de alimentos según la tecnología de secuenciación Illumina (A) y Nanopore (B). Las abundancias de los filos menores al 1% se agruparon en la categoría “Otros”. Los códigos de los alimentos se indican: Bolones (B), Ceviches (C), Carne de res (CM), Encebollados (E), Ensalada de frutas (EF), Empanadas de queso (EM), Frutas picadas (F), Mote con choncho (MT), Pollo (P), Quesos (Q), Salchipapas (SP) y Salsas picantes (S).

Tablas suplementarias

Tabla suplementaria I. Lista de alimentos y sus características utilizadas para la comparación de tecnologías de secuenciación Illumina e Nanopore.

#	Alimentos	Identificación (ID)	Descripción
1	Bolones	B	Puré de plátano verde frito de forma esférica y rellena con queso.
2	Ceviches	C	Mariscos cocinados y marinados con limón, cebolla y tomate.
3	Carne de res	CM	Carne de res cruda molida con grasa
4	Encebollados	E	Caldo de pescado con cebolla, yuca y yerbita.
5	Ensalada de frutas	EF	Trozos de fruta de guineo, sandía, papaya, sumergidos en jugo de naranja y sandía.
6	Empanadas de queso	EM	Masa de harina de trigo frita y rellena con queso
7	Frutas picadas	F	Tajadas de piña o papaya
8	Mote con chancho	MT	Grano de maíz seco cocido y pelado mezclado con chancho hornado
9	Pollo	P	Carne de pollo cruda
10	Quesos	Q	Quesos sin pasteurizar
11	Salchipapas	SP	Papas y salchichas de res fritas
12	Salsas picantes	S	Ají curtido con limón y cebolla

Tabla suplementaria II. Flujos bioinformáticos utilizados para la detección de los patógenos alimentarios, se muestra el consumo de memoria RAM en gigabytes y el tiempo de ejecución en minutos.

#	Programas	Memoria RAM utilizada (Gb)	Tiempo de ejecución (Minutos)
P1	EPI2ME	4	30
P2	EPI2ME-LABS	16	65
P3	Spaguetty-modificado	8	45
P4	NCBI +	16	60
P5	Situseq	16	60

Tabla Suplementaria III. Lecturas de las secuencias analizadas de cada muestra de alimento según el flujo utilizado. Se observa el total de cada flujo y los totales de cada alimento según el patógeno objetivo.

Total de especies según cada flujo		166528	166528	134977	167534	8780	
Patógeno	Muestras	P1	P2	P3	P4	P5	TOTAL
<i>Listeria</i>	Control positivo	0	0	0	0	0	0
	Frutas	0	0	0	0	0	0
	Queso	9	0	0	14	0	23
	Salsas	0	0	0	0	0	0
	Totales	9	0	0	14	0	23
<i>Salmonella enterica</i>	Muestras	P1	P2	P3	P4	P5	
	Control positivo	1856	1929	1193	1864	1	6843
	Frutas	4	0	0	4	0	8
	Queso	132	89	47	117	0	385
	Salsas	0	0	0	0	0	0
Totales	1992	2018	1240	1985	1	7236	
<i>Vibrio</i>	Muestras	P1	P2	P3	P4	P5	
	Control positivo	3	1	1	3	0	8
	Frutas	3	0	0	4	0	7
	Queso	34	0	0	28	1	63
	Salsas	2	0	0	2	0	4
Totales	42	1	1	37	1	82	

Tabla Suplementaria IV. Géneros de bacterias patógenas transmitidas por alimentos detectadas en las muestras de alimentos estudiadas para el estudio comparativo en la obtención del flujo bioinformático para detectar *L. monocytogenes*, *S. enterica* y *V. cholerae*.

Géneros	P1	P2	P3	P4	P5
<i>Bacillus cereus</i> ATCC 14579	7	0	0	0	0
<i>Clostridium perfringens</i>	1	1	0	1	2
<i>Cronobacter sakazakii</i>	3	5	0	2	1
<i>Cronobacter turicensis</i> z3032	7	0	0	8	0
<i>Escherichia coli</i>	133	38	0	121	0
<i>Listeria ivanovii</i>	1	0	0	0	0
<i>Salmonella enterica</i>	0	2018	0	0	0
<i>Salmonella enterica</i> subsp. <i>arizonae</i>	37	0	0	29	0
<i>Salmonella enterica</i> subsp. <i>enterica</i>	1466	0	0	1467	0
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Typhimurium</i>	484	0	0	484	0
<i>Salmonella enterica</i> subsp. <i>houtenae</i>	3	0	0	3	0
<i>Salmonella enterica</i> subsp. <i>indica</i>	1	0	0	1	0
<i>Salmonella enterica</i> subsp. <i>salamae</i>	1	0	0	1	0
<i>Shigella boydii</i>	3636	2055	0	3629	0
<i>Shigella dysenteriae</i>	48	33	0	45	0
<i>Shigella flexneri</i>	7787	16138	0	7756	0
<i>Shigella sonnei</i>	11825	6838	0	11829	0
<i>Staphylococcus aureus</i>	26	0	0	6	0
<i>Vibrio cholerae</i>	6	0	0	2	0
<i>Vibrio parahaemolyticus</i>	2	0	0	1	0
<i>Vibrio vulnificus</i> NBRC 15645 = ATCC 27562	1	0	0	1	0
<i>Yersinia enterocolitica</i> subsp. <i>palaearctica</i>	5	0	0	0	0
<i>Yersinia pseudotuberculosis</i>	1	0	0	1	0
TOTAL	25481	27126	0	25387	3
Abundancia relativa al total de lecturas	15,30%	16,29%	0,00%	15,15%	0,03%

Tabla Suplementaria V. Géneros más abundantes AR > 1% detectados en las 12 muestras de alimentos según la tecnología de secuenciación usada, Illumina vs Nanopore. Se muestra la abundancia relativa del total de lecturas del conjunto de datos.

Géneros Illumina	Géneros Nanopore	Abundancias Illumina %	Abundancias Nanopore %
<i>Acinetobacter</i>	(blank)	0.00%	1.55%
<i>Aeromonas</i>	<i>Hafnia-Obesumbacterium</i>	2.06%	1.85%
<i>Bacillus</i>	<i>Bacillus</i>	33.74%	21.75%
<i>Bacteroides</i>	(blank)	0.00%	1.10%
<i>Citrobacter</i>	<i>Escherichia-Shigella</i>	2.40%	2.03%
<i>Clostridium_sensu_stricto</i>	<i>Raoultella</i>	7.53%	7.14%
<i>Clostridium_XI</i>	<i>Pseudomonas</i>	2.10%	1.95%
<i>Enterobacter</i>	<i>Lactococcus</i>	6.17%	6.06%
<i>Enterococcus</i>	<i>Clostridium_sensu_stricto_1</i>	2.89%	3.26%
<i>Escherichia/Shigella</i>	(blank)	0.00%	1.50%
<i>Fusobacterium</i>	(blank)	0.00%	1.70%
<i>Hafnia</i>	(blank)	0.00%	1.13%
<i>Klebsiella</i>	(blank)	0.00%	1.56%
<i>Kurthia</i>	(blank)	0.00%	1.27%
<i>Lactobacillus</i>	<i>Proteus</i>	2.07%	1.89%
<i>Lactococcus</i>	<i>Citrobacter</i>	5.01%	4.16%
<i>Myroides</i>	(blank)	0.00%	1.14%
<i>Pantoea</i>	<i>Pantoea</i>	1.11%	1.71%
<i>Pediococcus</i>	<i>Weissella</i>	2.72%	3.12%
<i>Pseudomonas</i>	<i>Aeromonas</i>	2.46%	2.70%
<i>Raoultella</i>	<i>Kurthia</i>	2.02%	1.84%
<i>Shewanella</i>	<i>Klebsiella</i>	2.28%	1.96%
<i>Streptophyta</i>	<i>Buttiauxella</i>	1.18%	1.77%
<i>Unclassified at genus level</i>	<i>Enterobacter</i>	9.96%	8.33%
<i>Veillonella</i>	(blank)	0.00%	1.06%
<i>Weissella</i>	<i>Enterococcus</i>	5.40%	5.57%