



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

---

**FACULTAD DE QUÍMICA**

**IDENTIFICACIÓN DE PLÁSMIDOS PRESENTES EN  
BACTERIAS PATÓGENAS DE INTERÉS AGROALIMENTARIO  
MEDIANTE TÉCNICAS DE SECUENCIACIÓN MASIVA Y  
ANÁLISIS BIOINFORMÁTICO.**

**TESIS**

**QUE PARA OBTENER EL TÍTULO DE  
QUÍMICO DE ALIMENTOS**

**PRESENTA:**

**Abraham Joshué Rosales García**

**Ciudad de México, 2017**





Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**JURADO ASIGNADO:**

**PRESIDENTE:** PROFESORA: MIREYA RODRÍGUEZ PENAGOS  
**VOCAL:** PROFESOR: JOSÉ ADELFO ESCALANTE LOZADA  
**SECRETARIO:** PROFESOR: ROBERTO DASSAEVT CARVENTE GARCÍA  
**1° SUPLENTE:** PROFESORA: BEATRIZ RUÍZ VILLAFÁN  
**2° SUPLENTE:** PROFESORA: NORMA ANGÉLICA CAMACHO DE LA ROSA

**LUGAR EN DONDE SE DESARROLLÓ EL TEMA:**

**CENTRO NACIONAL DE REFERENCIA EN DETECCIÓN DE ORGANISMOS  
GENÉTICAMENTE MODIFICADOS  
SUBDIRECCIÓN DE SECUENCIACIÓN Y BIOINFORMÁTICA  
SENASICA, TECÁMAC.**

**DIRECTOR DE TESIS:**

**SUPERVISOR TÉCNICO:**

---

**M. EN C. ROBERTO DASSAEVT  
CARVENTE GARCÍA**

---

**M. EN C. SANDRA IVETTE PÉREZ  
AGÜEROS**

**SUSTENTANTE:**

---

**ABRAHAM JOSHUÉ ROSALES GARCÍA**

*If opportunity doesn't knock, build a door.*

**Milton Berle**

ÍNDICE GENERAL

**ÍNDICE GENERAL** ..... I  
**ÍNDICE DE FIGURAS** ..... V  
**ÍNDICE DE TABLAS** ..... VII  
**ABREVIATURAS** ..... IX  
**RESUMEN** ..... 2  
**1.0 INTRODUCCIÓN** ..... 5  
    1.1 Enfermedades transmitidas por alimentos (ETA) ..... 5  
    1.2 Bacterias patógenas de interés agroalimentario ..... 6  
        1.2.1 *Listeria monocytogenes* ..... 7  
        1.2.2 *Escherichia coli* ..... 8  
        1.2.3 Genero *Salmonella* ..... 9  
    1.3 ADN bacteriano ..... 12  
        1.3.1 Cromosoma y Gen ribosomal 16S ..... 12  
        1.3.2 Generalidades de los plásmidos ..... 13  
    1.4 Extracción de ADN ..... 15  
        1.4.1 Extracción de plásmidos ..... 16  
    1.5 Secuenciación de ácidos nucleicos ..... 17  
        1.5.1 Secuenciación de primera generación ..... 17  
        1.5.2 Secuenciación de segunda generación ..... 20  
        1.5.3 Tecnología Ion Torrent ..... 22  
    1.6 Bioinformática ..... 24  
**2.0 JUSTIFICACIÓN** ..... 26

<b>3.0 HIPÓTESIS .....</b>	<b>29</b>
<b>4.0 OBJETIVOS.....</b>	<b>31</b>
4.1 Objetivo general.....	31
4.2 Objetivos particulares .....	31
<b>5.0 METODOLOGÍA .....</b>	<b>33</b>
5.1 Diagrama de flujo.....	33
5.2 Materiales .....	34
5.2.1 Material biológico .....	34
5.2.2 Material de laboratorio.....	34
5.2.3 Reactivos .....	35
5.2.4 Equipos .....	36
5.2.5 Software .....	36
5.3 Métodos.....	37
5.3.1 Selección de muestras .....	37
5.3.2 Extracción de plásmidos. ....	37
5.3.2.1 Zyppy™ Plasmid Miniprep Kit (Protocolo A).....	38
5.3.2.2 ZR Plasmid Miniprep™-Classic (Protocolo B).....	38
5.3.2.3 GeneJET Plasmid Miniprep Kit (Protocolo C).....	39
5.3.2.4 Metodología ajustada (Protocolo D).....	39
5.3.3 Evaluación de la calidad y cuantificación del ADN extraído .....	40
5.3.4 Purificación del ADN plasmídico .....	41
5.3.4.1 Búsqueda de contaminantes cromosomales.....	42
5.3.4.2 Remoción de ADN genómico .....	43
5.3.4.3 Búsqueda de ADN plasmídico.....	45
5.3.4.3.1 Diseño de iniciadores específicos .....	45

5.3.4.3.2 Amplificación de regiones únicas en plásmidos .....	46
5.3.5 Amplificación de los plásmidos purificados .....	47
5.3.6 Secuenciación .....	48
5.3.6.1 Generación de Bibliotecas Genómicas .....	48
5.3.6.1.1 Fragmentación del material genético .....	49
5.3.6.1.2 Purificación de material genético fragmentado.....	51
5.3.6.1.3 Ligación de adaptadores. ....	51
5.3.6.1.4 Selección de tamaños .....	53
5.3.6.1.5 Amplificación de la biblioteca genómica .....	53
5.3.6.1.6 Cuantificación de las bibliotecas mediante PCR tiempo real.....	54
5.3.6.2 Amplificación clonal por em-PCR .....	56
5.3.6.2.1 Enriquecimiento.....	58
5.3.6.3 Corrida de secuenciación .....	61
5.3.6.4 Análisis bioinformático.....	63
<b>6.0 RESULTADOS.....</b>	<b>65</b>
6.1 Selección de muestras.....	65
6.2 Extracción de plásmidos .....	65
6.3 Cuantificación y evaluación de la calidad del ADN extraído .....	66
6.4 Purificación del ADN plasmídico. ....	67
6.4.1 Búsqueda de contaminantes cromosomales.....	67
6.4.2 Remoción de ADN genómico. ....	67
6.4.3 Búsqueda de ADN plasmídico.....	68
6.5 Amplificación de los plásmidos purificados.....	73
6.6 Generación de bibliotecas genómicas .....	73
6.6.1 Amplificación de las Bibliotecas Genómicas .....	75

6.6.2 Cuantificación de las bibliotecas mediante PCR tiempo real. ....	76
6.7 Amplificación clonal por em-PCR y enriquecimiento.....	77
6.8 Secuenciación .....	77
6.9 Análisis bioinformático .....	79
6.9.1 Análisis comparativos.....	89
<b>7 .0 DISCUSIÓN .....</b>	<b>98</b>
7.1 Elección del mejor protocolo de extracción para la purificación de plásmidos .....	98
7.2 Secuenciación masiva .....	101
7.3 Análisis bioinformático .....	102
<b>8.0 CONCLUSIONES.....</b>	<b>110</b>
<b>9.0 PROSPECTIVAS.....</b>	<b>112</b>
<b>10.0 ANEXOS .....</b>	<b>114</b>
<b>Anexo 1.</b> Cargado y revelado de chip High Sensitivity DNA y 7500 DNA en el equipo Bioanalyzer 2100 Agilent .....	114
<b>Anexo 2.</b> Score de los iniciadores para validar especificidad.....	116
<b>Anexo 3.</b> Interpretación del porcentaje de recuperación de ISPs. ....	118
<b>Anexo 4.</b> Descripción del Ion 316 v2™ Chip. ....	119
<b>Anexo 5.</b> Líneas de comando para ejecutar CSA. ....	120
<b>11.0 Bibliografía.....</b>	<b>122</b>

ÍNDICE DE FIGURAS

<b>Figura 1.</b> Secuenciación tipo Sanger.....	19
<b>Figura 2.</b> Pirosecuenciación utilizada por el secuenciador 454 .....	21
<b>Figura 3.</b> Secuenciación por semi-conducción Ion Torrent.....	23
<b>Figura 4.</b> Condiciones para la amplificación del gen ribosomal 16S.....	43
<b>Figura 5.</b> Digestión enzimática para la eliminación de ADN cromosomal .....	44
<b>Figura 6.</b> Condiciones para la amplificación de regiones únicas en plásmidos ....	46
<b>Figura 7.</b> Amplificación con hexámeros al azar .....	47
<b>Figura 8.</b> Metodología de amplificación del material genético con Illustra™ GenomiPhi DNA Amplification Kit.....	48
<b>Figura 9.</b> Programa para la ligación de adaptadores .....	52
<b>Figura 10.</b> Condiciones de tiempo y temperatura para la amplificación de la librería .....	54
<b>Figura 11.</b> Diagrama de preparación de diluciones para cada librería.....	55
<b>Figura 12.</b> Programa para la cuantificación por PCR tiempo real .....	56
<b>Figura 13.</b> Termociclador de emPCR Ion One Touch 2.....	58
<b>Figura 14.</b> Sistema de enriquecimiento Ion One Touch.....	60
<b>Figura 15.</b> Tira de 8 pozos para el enriquecimiento de las librerías amplificadas.	60
<b>Figura 16.</b> Cargado del Ion Chip 316 v2 .....	62
<b>Figura 17.</b> Extracción del material genético plasmídico con las cuatro metodologías utilizadas .....	65
<b>Figura 18.</b> Evaluación de contaminantes cromosomales en los extractos de cada metodología.....	67
<b>Figura 19.</b> Ensayo de amplificación para el gen Ribosomal 16S posterior a la digestión enzimática .....	68
<b>Figura 20.</b> Identificación de regiones únicas en plásmidos.....	72
<b>Figura 21.</b> Fragmentación del material genético para la elaboración de librerías.	74
<b>Figura 22.</b> Bibliotecas amplificadas .....	75
<b>Figura 23.</b> Cuantificación de librerías mediante qPCR .....	76

<b>Figura 24.</b> Reporte de la corrida de secuenciación para plásmidos con la plataforma Ion Torrent PGM.....	78
<b>Figura 25.</b> Efecto del recorte de lecturas con baja calidad .....	80
<b>Figura 26.</b> Comparativo entre las longitudes de las lecturas antes y después del recorte de adaptadores .....	81
<b>Figura 27.</b> Análisis en One Codex para la corrida de plásmidos .....	83
<b>Figura 28.</b> Ensamble De Novo. Sobrelape de fragmentos con extremos similares para ensamblar una secuencia consenso .....	84
<b>Figura 29.</b> Mapeo a referencia en el software CLC .....	86
<b>Figura 30.</b> Secuencias procesadas con CSA .....	90
<b>Figura 31.</b> Alineamientos de las referencias reorganizadas y sin ordenar.....	91
<b>Figura 32.</b> Alineamiento circular del mismo par de plásmidos.....	93
<b>Figura 33.</b> Plásmido de menor tamaño, contenido en uno con longitud mayor ....	95
<b>Figura 34.</b> Identificación de inserciones dentro de una secuencia conservada..	107
<b>Figura 35.</b> Pozos de cargado de las muestras en el Chip 7500 DNA.....	115
<b>Figura 36.</b> Pozos de cargado de las muestras en el Chip High Sensitivity DNA	115
<b>Figura 37.</b> Valores obtenidos del alineamiento BLASn.....	116
<b>Figura 38.</b> Ion 316™ Chip.....	119

ÍNDICE DE TABLAS

**Tabla 1.** Cuadro comparativo de las características generales de las bacterias en el estudio. .... 11

**Tabla 2.** Cuadro comparativo de los parámetros entre las diferentes plataformas de NGS..... 22

**Tabla 3.** Cuadro comparativo para cada metodología. .... 37

**Tabla 4.** Iniciadores universales para gen ribosomal 16S..... 42

**Tabla 5.** Mezcla de reacción para la amplificación del gen Ribosomal 16S..... 42

**Tabla 6.** Mezcla por reacción para la digestión enzimática..... 44

**Tabla 7.** Criterios para el diseño de iniciadores ..... 45

**Tabla 8.** Mezcla de reacción para la amplificación las regiones únicas en plásmidos. .... 46

**Tabla 9.** Condiciones de ultrasonificación para la fragmentación. .... 49

**Tabla 10.** Mezcla de reacción para la reparación de extremos. .... 50

**Tabla 11.** Mezcla de reacción para la fragmentación enzimática..... 50

**Tabla 12.** Mezcla de reacción para la ligación de adaptadores. .... 52

**Tabla 13.** Mezcla de reacción para la amplificación de la biblioteca genómica. ... 53

**Tabla 14.** Concentración y calidad de los extractos de ADN plasmídico para cada protocolo..... 66

**Tabla 15.** Secuencias de referencia plasmídica descargadas para cada bacteria en estudio..... 70

**Tabla 16.** Iniciadores diseñados para cada plásmido según su especie y/o serotipo..... 71

**Tabla 17.** Cuantificación del material genético obtenido mediante la amplificación con hexámeros al azar. .... 73

**Tabla 18.** Cuantificación de librerías por qPCR. .... 77

**Tabla 19.** Resultados del ensayo en BLAST para el contig más largo de cada muestra. .... 85

<b>Tabla 20.</b> Resultados obtenidos de los mapeos a cada referencia. Se muestra la cobertura de manera decreciente y la profundidad que se obtuvo para cada registro en el software CLCBio.....	87
<b>Tabla 21.</b> Primera depuración y persistencia de registros altamente similares.....	92
<b>Tabla 22.</b> Resumen final de plásmidos identificados para cada bacteria dentro del estudio. Mostrados en orden decreciente.....	94
<b>Tabla 23.</b> Regiones de la secuencia transponible y enzimas para las que codifican el plásmido pLM1-2bUG1 de <i>L. monocytogenes</i> (NC_014495.1). .....	96
<b>Tabla 24.</b> Resultados de la purificación para cada método y muestra. ....	101
<b>Tabla 25.</b> Valores de identidad, cobertura y Valor de expectación para cada iniciador sintetizado.....	117
<b>Tabla 26.</b> Interpretación de los valores del % de recuperación. Parámetros implementados dentro de los ensayos de secuenciación del CNRDOGM. ....	118

ABREVIATURAS

<b>% GC</b>	Porcentaje de guanina y citosina	<b>Mb</b>	Mega base
<b>°C</b>	Grados Celsius	<b>mL</b>	Mililitro
<b>ADEC</b>	<i>E. coli</i> adherencia difusa	<b>mM</b>	Mili molar
<b>ADN</b>	Acido desoxirribonucleico	<b>NCBI</b>	National Center for Biotechnology Information
<b>ADN CCC</b>	ADN cíclico covalentemente cerrado	<b>ng</b>	Nano Gramo
<b>ARN</b>	Ácido ribonucleico	<b>NGS/SNG</b>	Secuenciación de Nueva Generación
<b>Aw</b>	Actividad acuosa	<b>nm</b>	Nanómetro
<b>BLAST</b>	Basic Local Alignment Search Tool (Herramienta de búsqueda básica de alineamiento local)	<b>pb</b>	Pares de bases
<b>CNRDOGM</b>	Referencia en Detección de Organismos Genéticamente Modificados	<b>PCR</b>	Reacción en Cadena de la Polimerasa
<b>CSA</b>	Circular Sequences Alignment	<b>PGM</b>	Personal Genome Machine
<b>ddNTP</b>	Dideoxinucleótido trifosfato	<b>pH</b>	Potencial de hidrogeno
<b>dNTP</b>	Desoxinucleótido trifosfato	<b>PPI</b>	Pirofosfato inorgánico
<b>DTT</b>	Ditiotreitol	<b>qPCR</b>	Reacción en Cadena de la Polimerasa de cuantificación
<b>EAEC</b>	<i>E. coli</i> entero agregativa	<b>SDS</b>	Dodecil sulfato sódico
<b>EHEC</b>	<i>E. coli</i> entero hemorrágica	<b>SENASICA</b>	Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria
<b>EIEC</b>	<i>E. coli</i> entero invasiva	<b>SHU</b>	Síndrome Hemolítico Urémico
<b>em-PCR</b>	Reacción en cadena de la polimerasa por emulsión	<b>SINAVE</b>	Sistema Nacional de Vigilancia Epidemiológica
<b>EPEC</b>	<i>E. coli</i> enteropatógena	<b>STEC</b>	<i>E. coli</i> sigatoxigenica
<b>ETA</b>	Enfermedades Transmitidas por Alimentos	<b>tARN<sup>Ala</sup></b>	ARN de transporte de alanina
<b>ETEC</b>	<i>E. coli</i> enterotoxigénica	<b>Tm</b>	Temperatura de fusión
<b>EtOH</b>	Etanol	<b>UV-vis</b>	Luz ultravioleta visible
<b>Gb</b>	Giga base	<b>WHO/OMS</b>	Organización Mundial de la Salud
<b>ISFET</b>	Transistor de efecto campo sensible a iones	<b>µL</b>	Micro litro
<b>ISP's</b>	Ion Sphere Particles	<b>µM</b>	Micro molar
<b>Kb</b>	Kilo base		
<b>LDDOP</b>	Laboratorio de Diagnóstico y Detección de Organismos Patógenos		
<b>M</b>	Molar		

# *Resumen*

**RESUMEN**

En México, el SENASICA desarrolla trabajos y actividades para asegurar la inocuidad de los alimentos, verificando las vías de distribución e implementando programas informativos a fin de realizar un buen manejo de los mismos. Por su parte, debido a que los padecimientos causados por la ingesta de alimentos contaminados con agentes patógenos representan un problema de salud pública para la población mexicana, la Subdirección de Secuenciación y Bioinformática del CNRDOGM ha centrado parte de sus actividades al estudio genómico de bacterias patógenas de interés agroalimentario con mayor incidencia, entre las que se encuentran *Salmonella* spp., *Listeria monocytogenes*, y *Escherichia coli*.

En los últimos años se ha encontrado que las características de patogenicidad y virulencia, así como diversos factores responsables de daños y la transmisión de enfermedades, se encuentran generalmente codificados en elementos extracromosomales, por lo cual en el presente proyecto se diseñó y estandarizó una estrategia de extracción y purificación de plásmidos de tres de las bacterias con mayor incidencia relacionadas con ETA: *Salmonella enterica* serovar Enteritidis, *Listeria monocytogenes* y *Escherichia coli*, para la obtención de material genético plasmídico que pueda ser secuenciado mediante NGS y analizado bioinformáticamente.

Para ello, se estandarizó la extracción de plásmidos a partir de cuatro diferentes metodologías, de las cuales se eligió la que presentó mejores rendimientos y calidades en los extractos de ADN, posteriormente se eliminó el material genético cromosomal mediante digestión enzimática y se evaluó la presencia de los plásmidos mediante PCR.

El material genético restante fue enriquecido empleando una amplificación por desplazamiento mediante PCR isotermal y posteriormente secuenciado en la plataforma Ion Torrent PGM™.

Finalmente, al llevar a cabo el análisis bioinformático se logró la identificación, a partir de las secuencias obtenidas, de las tres bacterias en estudio. Además de la identificación de tres plásmidos con una cobertura del 100 % para la bacteria *L. monocytogenes*; tres plásmidos para la bacteria *E. coli* con coberturas del 77.0, 85.0 y 99.0 % respectivamente y un plásmido para la bacteria *S. Enteritidis* con una cobertura del 99.0 %. Adicional a esto, se identificaron dos referencias plasmídicas, en las cuales se observó que su contenido nucleotídico fue altamente conservado, sin embargo, una de ellas poseía una región transponible dentro de la cual se identificaron genes codificantes a distintos tipos de trasnposasas y uno más para una integrasa.

# *Capítulo 1*

## **INTRODUCCIÓN**

### 1.0 INTRODUCCIÓN

A nivel mundial se ha considerado a las enfermedades causadas por el consumo de agua o alimentos contaminados con agentes nocivos como un importante problema de salud. Se identifican dentro de estos agentes nocivos a elementos de origen químico, como residuos de plaguicidas o metales pesados acarreados por el riego, además de elementos biológicos como virus, parásitos y bacterias <sup>[48]</sup>. El consumo de alimentos con agentes nocivos representa una amenaza latente para la salud a nivel mundial, poniendo en peligro la vida de lactantes, niños, mujeres embarazadas, adultos mayores y personas inmunocomprometidas, convirtiéndolos en grupos particularmente vulnerables <sup>[51]</sup>.

#### 1.1 Enfermedades Transmitidas por Alimentos (ETA)

Las ETA son consecuencia de la ingesta de alimentos y agua contaminados con microorganismos o sustancias químicas. Se conoce la existencia de aproximadamente 31 agentes causantes de 32 enfermedades, de los cuales 12 corresponden a bacterias patógenas <sup>[48]</sup>.

La OMS estima que cada año alrededor del mundo, cerca de 600 millones de personas se ven afectadas a causa de las ETA, de las cuales cerca de 420 000 casos resultan en muerte. Las infecciones relacionadas con el desarrollo de diarrea infecciosa representan más de la mitad de los casos reportados de ETA, la cual se desarrolla por la ingestión de carne, huevos y productos lácteos con un tratamiento térmico no adecuado, por el consumo de frutas y verduras mal lavadas, así como de agua contaminada <sup>[52]</sup>.

La aparición de síntomas gastrointestinales es el reflejo más común de la ingesta de alimentos contaminados, sin embargo, esta puede provocar incluso síntomas neurológicos, ginecológicos e inmunológicos, entre otros <sup>[50]</sup>.

Las ETA se clasifican en dos tipos según el origen de la enfermedad:

- **Intoxicaciones transmitidas por alimentos.** Causadas por la ingesta de alimentos los cuales portan toxinas producidas por la proliferación de las bacterias en el alimento. La presencia de la bacteria no resulta necesaria para el desarrollo de la enfermedad.
- **Infecciones transmitidas por alimentos.** Causadas por la ingestión de alimentos portadores de bacterias vivas, la cuales se instalan dentro del hospedero y desarrollan la enfermedad.

### 1.2 Bacterias patógenas de interés agroalimentario

A pesar de los esfuerzos realizados para generar campañas donde se pretende cuidar la inocuidad y calidad de los alimentos, anualmente se siguen registrando miles de casos de infecciones a causa del consumo de alimentos portadores de patógenos alimentarios. Las bacterias y los virus son la causa más común de ETA presentando una sintomatología variable según el agente patógeno consumido, llegando incluso a causar la muerte <sup>[15]</sup>.

Dentro de los 12 patógenos más recurrentemente identificados en los casos de diarrea infecciosa por la Organización Mundial de la Salud se encuentran *Salmonella* no tifoidea y *Escherichia coli* patógena <sup>[53]</sup>, ambos con una tasa de morbilidad alta. Por otra parte, una de las enfermedades transmitidas por alimentos que tiene uno de sus grupos vulnerables en las mujeres embarazadas es la listeriosis, causada por la bacteria *Listeria monocytogenes*, microorganismo ubicuo en el ambiente, el cual se presenta de modo inesperado con una baja incidencia, pero un alto índice de letalidad <sup>[30]</sup>.

Tan solo en 2014 se reportaron en los Estados Unidos de América 19,542 casos de infecciones causadas por el consumo de alimentos contaminados, de los cuales, 4,445 requirió de hospitalización y 71 de ellas terminaron en muerte. Dentro los casos de hospitalización se logró la identificación, entre otras bacterias, de las tres bacterias involucradas en este estudio, reportando los siguientes

porcentajes para cada una: *E. coli* en 23.0 %, *Listeria* spp con 11.0 % y *Salmonella* spp con un 6.0 % <sup>[7]</sup>.

Una vía de transmisión potencial es el consumo de germinados de semillas de hortalizas, el cual se ha ido popularizando a lo largo de la última década debido a la demanda de los consumidores por probar nuevos alimentos y sabores. Sin embargo, el número de brotes de ETA asociados al consumo de estos alimentos también ha ido en aumento, presentando una mayor relación con las cepas patógenas de *Salmonella* spp y *E. coli*. En América del Norte se asocian principalmente al consumo de germinados de alfalfa y soya, en tanto que, en Japón, en el año de 1996 se registró un brote de *E. coli* O157:H7 debido al consumo de germinados de rábano, el cual desató 9,451 casos de enfermedades con un total de 12 muertes. En México no se encuentran registros de casos reportados de brotes asociados al consumo de estos alimentos, sin embargo, esto no significa que no existan casos relacionados con bacterias patógenas, ya que el consumo de vegetales es común, un ejemplo de ello son los germinados de alfalfa y soya los cuales se consumen como una práctica con alta incidencia dentro de la población mexicana, con lo cual se incrementan los riesgos de desarrollar una infección relacionada con el consumo de patógenos acarreados por los mismos alimentos <sup>[45]</sup>.

### 1.2.1 *Listeria monocytogenes*

*Listeria* es un bacilo Gram-positivo, no ramificado y anaerobio facultativo, capaz de desarrollarse en un rango muy amplio de temperaturas (1 °C – 45 °C) y tolerar altas concentraciones de NaCl. Móvil a temperatura ambiente pero inmóvil a 37 °C, temperatura en que sus flagelos se inactivan. Debido a que *L. monocytogenes* puede crecer a bajas temperaturas, los alimentos contaminados con un número pequeño de microorganismos almacenados en refrigeración por periodos prologados de tiempo pueden presentar un creciente desarrollo de esta bacteria <sup>[28]</sup>.

El género *Listeria* está conformado por 10 especies, de las cuales únicamente *L. monocytogenes* es reconocida como patógena y aunque es poco frecuente, su letalidad se encuentra limitada a poblaciones bien definidas como recién nacidos, ancianos, mujeres embarazadas y pacientes con un sistema inmunológico disminuido <sup>[28]</sup>. Así mismo, se considera una de las principales causas de muertes por ETA en los Estados Unidos de América <sup>[28]</sup>.

Además, como se ha mencionado, *L. monocytogenes* es un microorganismo ubicuo localizado en fuentes ambientales, así como en las heces de mamíferos, aves y otros animales. Su principal fuente de transmisión es mediante el consumo de alimentos contaminados como carne cruda o con mal procesamiento térmico, lácteos no pasteurizados y vegetales crudos mal lavados, sin embargo, puede transmitirse de individuo a individuo, principalmente de madre a hijo al momento del nacimiento <sup>[28]</sup>.

La infección en mujeres embarazadas por *L. monocytogenes* también es responsable de muertes neonatales, nacimientos prematuros y abortos espontáneos; en adultos sanos puede causar gastroenteritis, diarrea acuosa, fiebre, náuseas, cefalea, mialgias y artralgias <sup>[28]</sup>, además de ser capaz de producir meningitis, presentando una letalidad del 20 al 50 % y secuelas neurológicas importantes en los sobrevivientes <sup>[28]</sup>.

### **1.2.2 *Escherichia coli***

Bacilo Gram-negativo, oxidasa negativa, anaerobio facultativo, que puede o no poseer flagelos que le brinden movilidad. Perteneciente a la familia de las enterobacterias <sup>[8]</sup> crece en un intervalo de los 7 a los 50 °C con un óptimo en los 37 °C <sup>[54]</sup>. Se ubica principalmente como comensal en el tracto digestivo de humanos y animales de sangre caliente que viven en asociaciones mutuamente beneficiosas y rara vez causa enfermedad <sup>[49]</sup>; sin embargo, existen algunas cepas productoras de toxinas capaces de provocar cuadros gastrointestinales graves al ser humano, de las cuales se identifican seis variedades agrupadas según sus mecanismos de patogenicidad y síndromes clínicos: *E. coli* Shigatoxigénica (STEC), *E. coli* enterohemorrágica (EHEC), *E. coli* enterotoxigénica (ETEC), *E. coli*

enteroinvasiva (EIEC), *E. coli* enteropatógena (EPEC), *E. coli* enteroagregativa (EAEC), y *E. coli* de adherencia difusa (ADEC) [14].

*E. coli* es transmitida de hospedero a hospedero mediante la ruta fecal-oral o través de contactos con superficies contaminadas. Un vehículo de transmisión muy importante es la ingesta de alimentos y bebidas contaminadas con materia fecal, la cual se encuentra estrechamente relacionada con los brotes epidemiológicos en adultos [8].

La OMS destaca a la especie de *E. coli* productora de toxina Shiga (STEC) debido a su agresividad y complicaciones producidas al consumidor, capaz de producir desde malestares estomacales y diarrea acuosa, hasta diarreas sanguinolentas o colitis hemorrágica, fiebre y vómitos, presenta un periodo de incubación promedio de cuatro días y la mayoría de los infectados logra sobreponerse a lo largo de los primeros 10 días, sin embargo, se estima que cerca del 10% de los casos de infección desarrollan Síndrome Hemolítico Urémico (SHU), el cual es acompañado de insuficiencia renal aguda, anemia hemolítica y trombocitopenia, con una tasa de letalidad del 3 al 5 %. En el 25 % de los casos que desarrollan SHU se pueden presentar complicaciones neurológicas como convulsiones, accidente cerebrovascular y coma, así como secuelas renales crónicas [54].

### 1.2.3 Genero *Salmonella*

Bacilo Gram-negativo perteneciente a la familia de las enterobacterias, anaerobio facultativo, móviles mediante flagelos peritricos, organismos no encapsulados y no esporulados, con una longitud de entre las 2 y 5 micras [13], crece en un intervalo de temperatura de 5 a 46 °C con un óptimo en los 37 °C [10].

*Salmonella* spp. habita generalmente en el tracto intestinal de animales sanos, principalmente aves de corral, ganado vacuno y porcino, así como animales domésticos. Al ser un habitante de la zona gastrointestinal, puede ser desechada en las heces de los animales, en donde llega a sobrevivir por un periodo prolongado de tiempo gracias a su resistencia a una Aw baja [10].

El género *Salmonella* se organiza en dos grandes grupos: *S. bongori*, la cual no presenta ningún riesgo para el ser humano y *S. enterica*, que a su vez se subdivide en seis subespecies y de estas se tienen caracterizados más de 2500 serotipos <sup>[26]</sup>, mismos que en su gran mayoría pueden provocar gastroenteritis. Entre estos serotipos destacan *S. Enteritidis* y *S. Typhimurium* como los dos serotipos más importantes causantes de salmonelosis transmitida desde animales a seres humanos en la mayor parte del mundo <sup>[48]</sup>.

La principal fuente de transmisión de *Salmonella* spp. se presenta vía fecal-oral por malos hábitos de higiene, así como por el consumo de alimentos contaminados con materia fecal. La salmonelosis es una enfermedad que generalmente se caracteriza por presentar fiebre alta, dolor abdominal, diarrea, náuseas y en ocasiones vómitos, los síntomas de la enfermedad se presentan entre las 6 y 72 horas posterior al consumo de la bacteria y esta puede durar de 2 a 7 días <sup>[48]</sup>.

El SINAVE en su informe hasta la semana 28 de 2017, presenta un acumulado de 43,488 casos de enfermedades intestinales identificados como salmonelosis <sup>[38]</sup>, sin embargo, se estima que entre el 60 y el 80 % de los casos no se reconocen como parte de un brote identificado y se clasifican como casos esporádicos, o ni siquiera se diagnostican <sup>[48]</sup>.

En la Tabla 1 se muestra, a modo de resumen, las principales características de cada bacteria en el estudio.

Tabla 1. Cuadro comparativo de las características generales de las bacterias en el estudio.

Bacteria	Variedades patógenas a humanos	Características	Daños a la salud	Vector de transferencia	% de mortalidad
<i>L. monocytogenes</i>	<i>L. monocytogenes</i>	Bacilo Gram-positivo, Catalasa +, No ramificado, Flagelos peritricos, Móvil 30 °C o menos, No esporulado, Halotolerante, Anaerobio facultativo.	Listeriosis, Nacimientos prematuros, Muertes neonatales, Abortos espontáneos, Meningitis.	Carne cruda o mal cocida, Vegetales crudos, Vía madre a hijo al momento de nacer.	20 – 50 %
<i>E. coli</i>	STEC EHEC ETEC EIEC EPEC EAEC ADEC	Bacilo Gram-negativo, Oxidasa negativa, Puede o no poseer flagelos, Móvil, No esporulado, Anaerobio facultativo.	Diarrea acuosa, Diarreas sanguinolentas, Colitis hemorrágica, SHU, Insuficiencia renal aguda, Anemia hemolítica, Trombocitopenia.	Ruta fecal – oral de hospedero a hospedero, Ingesta de alimentos contaminados con restos fecales.	3 – 5 %
<i>S. enterica</i>	Tifoideas y no tifoideas	Bacilo Gram-negativo, Generalmente lactosa -, Flagelos peritricos, Móvil, No esporulado, Anaerobio facultativo.	Salmonelosis, Fiebre, Dolor abdominal, Diarrea, Náuseas, Vómito.	Consumo de alimentos contaminados con material fecal Carne, huevos y leche sin procesamiento térmico, Contacto fecal oral	No determinado*

\*Los casos de letalidad son poco frecuentes y no hay reporte de ellos <sup>[48]</sup>.

### 1.3 ADN bacteriano

#### 1.3.1 Cromosoma y Gen ribosomal 16S

El cromosoma en bacterias es una molécula circular de ADN de doble cadena en donde se aloja toda la información hereditaria y de regulación de una bacteria<sup>[43]</sup>. Esta macromolécula contiene genes constitutivos esenciales, necesarios para el mantenimiento de las funciones celulares básicas como reproducción, metabolismo y funciones específicas, tales como procesos de replicación, segregación, transcripción y traducción. Estos genes regulan su nivel de expresión de acuerdo a la presión selectiva que el medio y las condiciones de crecimiento impongan sobre la célula. Los genes constitutivos representan unas de las regiones mayormente conservadas dentro del cromosoma bacteriano al presentar un nivel de evolución muy bajo sobre sus secuencias codificadas<sup>[11] [56]</sup>. El almacenamiento de la información contenida en el cromosoma plantea un enorme reto en la distribución espacial dentro de la célula, ante esto, el cromosoma se encuentra perfectamente organizado y empaquetado en una estructura anclada al citoplasma denominada nucleoide<sup>[41]</sup>, fijada mediante asociación de proteínas favoreciendo así el concentrado de ADN en estructuras compactas<sup>[34]</sup>. Esta compactación del ADN se mantiene gracias al balance de fuerzas, incluyendo el súper enrollamiento, compactación por proteínas, transcripción y las interacciones ARN-ADN<sup>[44]</sup>.

Entre los genes constitutivos ubicados en el cromosoma bacteriano, se encuentra el gen ribosomal 16S, el cual, tiene un tamaño de aproximadamente 1500 pb que a pesar de poseer secuencias altamente conservadas, también contiene regiones hipervariables, las cuales, se emplean para el estudio de filogenia y taxonomía bacteriana, al punto de ser considerado como un cronómetro molecular propuesto por Carl Woese en la década de los 70, bajo el postulado de Zuckerkandl y Pauling en 1962, donde establecen que los cambios en los monómeros de las proteínas (aminoácidos) se producían al azar y aumentaban de manera lineal en dos organismos vivos. En 1968, Allan Wilson y Vincent Sarich observaron que los errores en la replicación de ADN inducían la aparición de mutaciones dirigiendo a

éste a la evolución molecular, determinando que con la acumulación de las diferencias evolutivas entre organismos se podría determinar su historia evolutiva. [32] [36] [57].

Con los estudios comparativos de las secuencias del gen ribosomal 16S, Woese logró determinar la división de los protistas en dos reinos: *Eubacteria* y *Archaeobacteria* para posteriormente introducir el término dominio en sustitución de reino como categoría taxonómica superior realizando la distribución de los organismos celulares en tres dominios: *Bacteria*, *Archaea* y *Eukarya* [32], en la actualidad es posible emplear al gen ribosomal 16S con fines de caracterización e identificación.

### 1.3.2 Generalidades de los plásmidos

Los plásmidos son replicones extracromosomales de doble cadena en su mayoría circulares denominados ADN CCC, aunque también se presentan en su forma lineal en algunas especies bacterianas y fúngicas [37] [46]. Estos elementos se ubican en el citoplasma de la célula y son portadores de información genética adicional a la contenida en cromosoma, capaces de facilitar la propagación de genes de selección positiva como resistencia a antimicrobianos, detoxificación de metales pesados, factores de virulencia, factores metabólicos, entre otras. [43] [46].

Son elementos con replicación autónoma y auto-regulada que utilizan la maquinaria de replicación del hospedero y son considerados como material genético de expresión eventual, la cual, generalmente se ve regulada en respuesta a presión selectiva ante la presencia de antimicrobianos o frente a condiciones ambientales a las que se vea expuesto cada microorganismo.

Entre los plásmidos mayormente estudiados se encuentran los plásmidos circulares con súper enrollamiento negativo encontrados generalmente en *E. coli* y otras bacterias Gram-negativas, con una longitud que oscila entre 1 – 100 Kb los cuales pueden representar hasta el 10 % del material genético celular. Se sabe que los plásmidos de gran tamaño se limitan a pocas copias dentro de una misma célula (plásmidos exigentes), a diferencia de los plásmidos de menor tamaño, que

se han encontrado hasta en miles de copias por célula (plásmidos tolerantes), estos inician un número indiscriminado de copias por ciclo, los cuales, al alcanzar cierta concentración, activan un sistema de autorregulación que inhibe la replicación plasmídica <sup>[43]</sup><sup>[46]</sup>.

Los plásmidos, como elementos móviles, son uno de los principales vehículos para el transporte de información genética <sup>[39]</sup>. De esta manera se cree que la transferencia de información adicional obtenida mediante la incorporación de elementos génicos ajenos a su genoma es el principal promotor de la rápida evolución y adquisición de habilidades para la adaptación a diferentes entornos en los microorganismos <sup>[37]</sup>.

Es posible realizar una clasificación de los plásmidos con base en las características que estos poseen:

- **Grupos de incompatibilidad:** según el sistema de replicación y segregación de copias <sup>[46]</sup>.
- **Por el fenotipo codificado:** plásmidos de resistencia (antimicrobianos y metales pesados), plásmidos de virulencia (codificantes de diversas toxinas), plásmidos bacteriocinogénicos (llevan a cabo la síntesis de bacteriocinas).
- **Por su capacidad de auto transferencia:** conjugativos (mediante transferencia horizontal), movilizables (en presencia de un plásmido conjugativo) y no movilizables (propios del organismo en el cual residen).
- **Número de copias:** Desde una o algunas pocas copias, hasta algunos miles de ellas. Ejemplo de estos, se tiene al plásmido *p15A* presente en cepas de *E. coli 15T* en un intervalo de 3 a 5 copias, así como el plásmido *pTZ*, utilizado como vector de clonación y en aplicaciones de ingeniería genética, capaz de generar arriba de 1000 copias de su secuencia por cada célula.

### 1.4 Extracción de ADN

Los métodos de aislamiento y purificación de ADN se fundamentan en el uso de sus características fisicoquímicas ya que el ADN está constituido por una doble hélice de nucleótidos unidas entre sí, integradas por un azúcar (desoxirribosa), un grupo fosfato y una base nitrogenada (adenina, guanina, timina o citosina). La unión de los nucleótidos ocurre entre el grupo fosfato y el azúcar, mediante enlaces fosfodiéster, dando lugar al esqueleto de la molécula. La unión de las bases complementarias se lleva a cabo mediante puentes de hidrógeno los cuales mantienen estable la estructura helicoidal <sup>[47]</sup>.

Los grupos fosfato están cargados negativamente y son polares, lo que le confiere al ADN una carga neta negativa y lo hace altamente polar, condición que permite disolverlo en soluciones acuosas y formar una capa hidratante alrededor de la molécula. Sin embargo, en presencia de alcoholes como el etanol o isopropanol, se rompe la capa hidratante y quedan expuestos los grupos fosfato, condición que favorece la interacción con cationes como el  $\text{Na}^+$ , el cual al reducir las fuerzas repulsivas entre las cadenas de nucleótidos permite la precipitación del ADN. A su vez, la carga neta negativa del ADN le permite unirse a moléculas y matrices inorgánicas cargadas positivamente permitiendo el aislamiento de la molécula mediante las interacciones electrostáticas con membranas de sílice <sup>[47]</sup>.

A lo largo del tiempo, se han diseñado distintos protocolos con la finalidad de obtener una cantidad y calidad de ADN adecuados, así como garantizar la eliminación de inhibidores potenciales que dificulten el tratamiento posterior de la molécula. En general, los protocolos tradicionales consisten en cinco etapas principales: homogenización del cultivo (para el caso de bacterias), lisis celular, separación de proteínas, lípidos y RNA, precipitación y redisolución del ADN <sup>[47]</sup>.

A partir de los años 90, se introdujeron al mercado kits de extracción que utilizan matrices inorgánicas (columnas de purificación) compactas cargadas positivamente capaces de retener varios microgramos de ADN y separarlos del resto de las biomoléculas presentes, permitiendo obtener un extracto libre de inhibidores. Durante la extracción, el ADN cargado negativamente se adsorbe o

une a la matriz selectiva de manera reversible manteniéndose unido a ésta durante la remoción de ARN, lípidos, proteínas y diversos metabolitos y posteriormente liberado de la matriz <sup>[47]</sup>, consiguiendo de esta forma su purificación.

### 1.4.1 Extracción de plásmidos

La minipreparación de plásmidos (minipreps), se consideran una técnica fundamental en biología molecular para el aislamiento del ADN plasmídico, los cuales pueden ser clasificados según los métodos de lisis celular y precipitación de ADN utilizados, teniendo los siguientes: métodos alcalinos que utilizan NaOH con SDS, así como NaOH con detergentes zwitteriónicos o no iónicos; los métodos con detergentes, que utilizan sólo SDS o en combinación con pronasa o lisozima; los procedimientos orgánicos, que utilizan fenol o mezclas con fenol/cloroformo solas o en combinación con Tritón X-100 o lisozima; y los tratamientos físicos que incluyen llevar a ebullición el medio en donde se encuentran las células en presencia de soluciones de lisozima con Tritón X-100 o Tween 20.

Dichos procedimientos han sido diseñados para llevar a cabo la ruptura de la célula y exponer todo el material biológico, así como desnaturalizar y precipitar el ADN tanto genómico como plasmídico. Una vez sedimentado el material genético, se agrega una solución de neutralización, la cual favorece la renaturalización y solubilización del ADN plasmídico, en tanto que los restos cromosomales se quedan en el sedimento. Finalmente, luego de una centrifugación se retira el sobrenadante con ADN plasmídico contenido en él.

Durante las últimas cuatro décadas los métodos más utilizados y sobre los cuales se basan los protocolos y kits comerciales hacen uso de un medio alcalino y detergentes aniónicos como el SDS, adicionalmente, muchos emplean columnas de afinidad con una composición molecular a base de sales de silicio como un paso extra para el aislamiento e incremento en el rendimiento y selectividad <sup>[24]</sup>.

### 1.5 Secuenciación de ácidos nucleicos

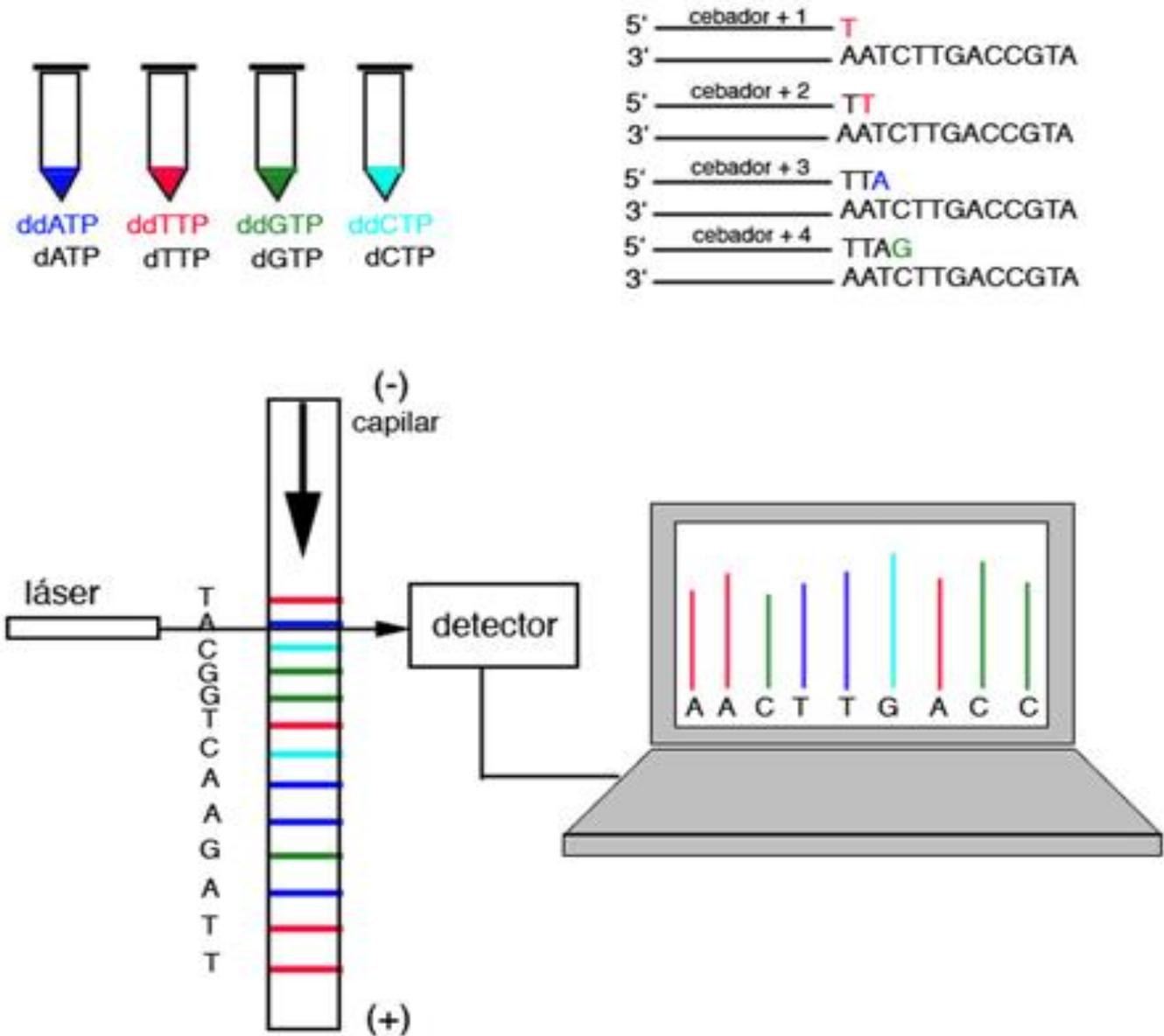
De manera determinante la Secuenciación de nueva generación, tiene origen en el desarrollo de las primeras metodologías de secuenciación de ADN, las cuales en principio se plantearon ser mucho más sencillas que la ya desarrollada secuenciación de proteínas, ya que se pensaba que encontrar la secuencia ordenada de cuatro nucleótidos sería mucho más sencillo que descifrar el posible orden de una proteína conformada por 20 aminoácidos diferentes. Años más tarde con el desarrollo de nuevas técnicas y la secuenciación de fragmentos de mayor tamaño esta idea cambió [3].

#### 1.5.1 Secuenciación de primera generación

El primer material genético secuenciado fue el tARN<sup>Ala</sup> (polinucleótido de 76 bases) por el equipo de Holley, en un proyecto que duró 7 años, el cual se publicó oficialmente en el año de 1965 [21]. La secuenciación se obtuvo al fragmentar la molécula de ADN vía enzimática, para después hacer uso de métodos cromatográficos y la posterior identificación de los extremos 5' y 3' para ordenar la secuencia [3].

Años más tarde, Frederick Sanger y su grupo de colaboradores dirigieron sus estudios al diseño de nuevas metodologías para la fragmentación y secuenciación, inicialmente de ARN y en una etapa posterior de ADN. A partir de esto, en el año de 1977 se reportaron los dos primeros protocolos que describían una gran aproximación a lo que actualmente es la secuenciación de ADN, Sanger con su método enzimático de terminación de cadenas (método de los didesoxinucleótidos – ddNTP's) y Gilbert & Maxam con el desarrollo del método por fragmentación química. Aunque ambas metodologías permitían un análisis completo de los fragmentos de ADN y la detección de modificaciones en su secuencia, la metodología de Sanger obtuvo mayor difusión y fue mediante ésta que un año más tarde se reportaría la secuencia completa del virus øX174 [35].

Posteriormente, en la década de los 90, con el desarrollo de nuevos proyectos de secuenciación, se modificó una parte fundamental de la metodología de Sanger, adaptándose a métodos menos nocivos, eliminando el uso de reactivos radioactivos con la implementación de ddNTP's marcados con fluoróforos (Figura 1) que se analizaban mediante electroforesis capilar, obteniendo un electroferograma que se traducía en el orden nucleotídico de los fragmentos secuenciados <sup>[3]</sup><sup>[27]</sup>.



**Figura 1. Secuenciación tipo Sanger.** La amplificación se detiene al incorporar a la síntesis un nucleótido modificado marcado con un fluoróforo único. Los fragmentos sintetizados se separan mediante electroforesis capilar durante la cual el detector realiza la lectura de cada nucleótido marcado y mediante la interpretación del ordenador, ésta es traducida a la secuencia del fragmento en cuestión <sup>[19]</sup>.

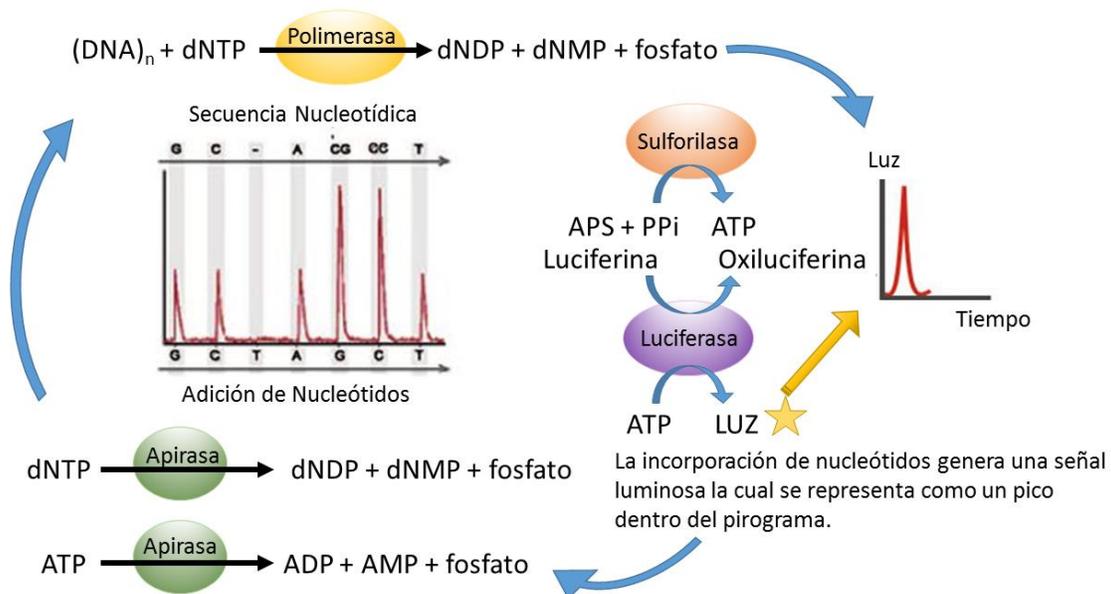
Los avances en el desarrollo de estas técnicas permitieron el diseño de nuevas tecnologías automatizadas para la secuenciación de ADN, como los equipos ABI Prism y CEQ-serie con capacidad de procesar hasta 96 muestras en pocas horas y con secuencias entre 500 y 1000 bases. En este punto quedó marcado lo que se consideró la primera generación de secuenciadores, abriendo paso a la secuenciación de segunda generación con la metodología *Whole Genome Shotgun Sequencing* la cual, aceleraba y facilitaba el análisis de los datos obtenidos <sup>[30]</sup>.

### 1.5.2 Secuenciación de segunda generación

Los avances en la ciencia de la secuenciación dieron pie en 1990 al desarrollo del “Proyecto Genoma Humano” y en 1995 se reportó el primer genoma secuenciado de un organismo de vida libre, *Haemophilus influenzae* <sup>[27]</sup>. En el año 2001 se publicó el primer borrador del genoma humano, el cual implicó un gasto cercano a los 3,000 millones de dólares, lo cual orilló a los científicos a la búsqueda de nuevas tecnologías que abarataran los costos sin afectar el rendimiento y obtención de resultados, llegando así a la creación de los denominados secuenciadores de segunda generación, capaces de generar miles de reacciones generadas en paralelo (secuenciadores de alto rendimiento [*high-throughput*]).

Fue hasta 2005 que surgió el primer desarrollador capaz de aportar una aproximación significativa y de buenos rendimientos a la secuenciación masiva con el nombre 454 FLX de Life Science que posteriormente sería absorbida por Roche, implementando la tecnología de pirosecuenciación (Figura 2), que aplica una técnica no fluorescente en la que participa el pirofosfato, producto de la reacción de polimerización del ADN, el cual, mediante una serie de reacciones enzimáticas acopladas, es transformado dando lugar a una liberación de luz. De esta forma, con la adición de cada nucleótido, es producida una imagen que se analiza para proporcionar flujogramas que una vez interpretados por el ordenador, devuelven las secuencias de nucleótidos <sup>[27]</sup>. A poco más de una década en el mercado, la pirosecuenciación en el secuenciador 454 FLX se ha visto desplazada por el desarrollo de tecnologías capaces de generar volúmenes mayores de datos

en un tiempo considerablemente menor, disminuyendo los costos e incrementando la cantidad de resultados obtenidos. En la Tabla 2 se observa un comparativo entre las plataformas más representativas de la NGS disponibles en la actualidad. Se observa que la evolución de las tecnologías ha desarrollado técnicas de secuenciación las cuales generan una cantidad mayor de información en tiempos reducidos. Es por ello que en 2017, la tecnología 454 de Roche ha salido de circulación, en búsqueda del desarrollo de técnicas competitivas con las actualmente vigentes en el mercado.



**Figura 2. Pirosecuenciación utilizada por el secuenciador 454.** Al incorporar un nucleótido conocido a la cadena que se está sintetizando, se libera un pirofosfato (PPi), el cual es transformado en ATP por la enzima ATP sulfurilasa, el ATP generado es entonces empleado junto con una luciferina por la enzima luciferasa, produciendo oxiluciferina y una señal luminosa. Finalmente, una apirasa se encarga de degradar aquellos nucleótidos que no han sido incorporados con el fin de que al añadir un nuevo dNTP, no existan restos de reacciones anteriores. El sensor del equipo captura esta señal lumínica y lo reproduce como un pico en un pirograma, con lo que se puede realizar el llamado de bases y una correcta traducción de los fragmentos secuenciados <sup>[33]</sup>.

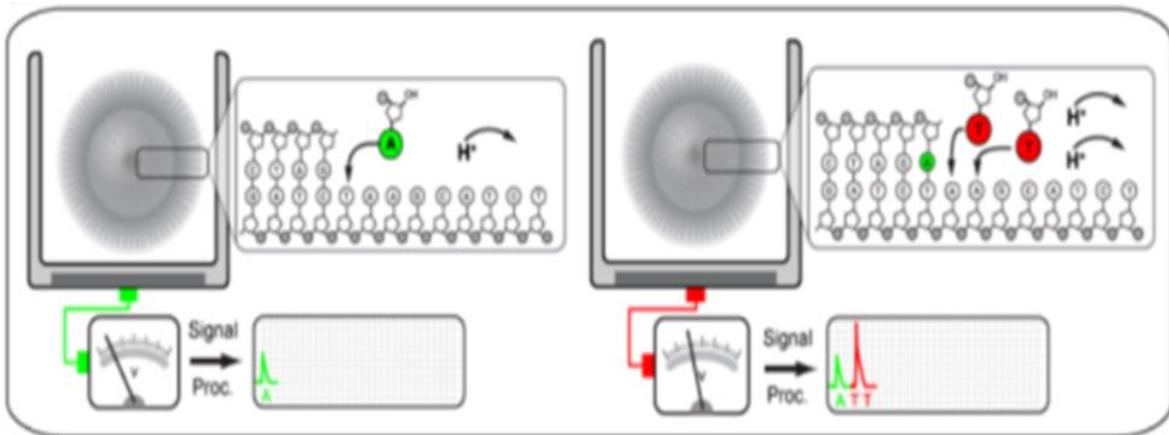
**Tabla 2. Cuadro comparativo de los parámetros entre las diferentes plataformas de NGS.**

<b>Plataforma</b> <b>Parámetro</b>	<b>454</b>	<b>Ion Torrent</b>	<b>Illumina</b>
<b>ADN para librerías</b>	1.0 µg	100.0 ng	1.0 ng
<b>Fragmentación</b>	- Sonicación - Nebulización	- Sonicación - Enzimática	Tagmentación
<b>Unión de adaptadores</b>	Ligación	Ligación	
<b>Identificadores</b>	MID	Barcodes	Índices
<b>Selección de tamaños</b>	Sizing solution	Electroforesis en agarosa	Perlas magnéticas AMPure XP: 0.4X y 1.8X
<b>Amplificación</b>	emPCR	- Amplificación de librería - emPCR	- Amplificación de librería - Clusterización
<b>Secuenciación</b>	Pirosecuenciación	Semiconducción	Fluorescencia
<b>Rendimiento</b>	Junior: 30 Mb FLX: 500 Mb	PGM: 2 Gb Proton: 10 Gb S5: 10 Gb	MiSeq: 15 Gb NextSeq: 120 Gb
<b>Tamaño de lecturas</b>	800 pb 1,600 pb	200 pb 400 pb	75 pb 150 pb 300 pb
<b>Dirección de las lecturas</b>	Single end	Single end	Pair end
<b>Tiempo de ensayo</b>	29 h	5 h	- 29 h - 56 h

### 1.5.3 Tecnología Ion Torrent

El Ion PGM™, es una plataforma de segunda generación y fue la elegida para el desarrollo de este trabajo. Este equipo, fue lanzado al mercado en el año 2010 y en la actualidad permanece en uso gracias a su versatilidad para ser aplicada en diversos trabajos de secuenciación, como la secuenciación dirigida, análisis de transcriptoma, exoma y genoma completo, así como la capacidad de escalar proyectos y migrarlos a otras plataformas. Ion Torrent implementa el principio de la semiconducción, realizando la precisa detección de los protones liberados como subproducto de la síntesis de ADN, lo cual genera un ligero cambio de voltaje en el medio. De este modo y mediante un flujo controlado y conocido de nucleótidos,

es posible determinar la secuencia de cada fragmento al detectar la adición de cada base a lo largo de la síntesis y la interpretación mediante el ordenador integrado al equipo (Figura 3). En el CNRDOGM se ha establecido para el Ion PGM™ un flujo de trabajo capaz de brindar resultados correspondientes a: la preparación de librerías en un periodo aproximado de 8 horas, la amplificación por emulsión en un estimado de 6 horas, la corrida de secuenciación con un total de 8 horas más y finalmente un par de horas adicionales para el análisis de datos, es decir que en un periodo de tres días es posible obtener resultados sobre todas las muestras secuenciadas. Además, el rendimiento obtenido por el equipo está relacionado con los soportes sólidos en donde se llevan a cabo las reacciones de síntesis, los cuales están hechos de una matriz semiconductor que, a su vez, dependerán de los elementos que se pretende estudiar. Se distribuyen en tres tipos: Ion 314 Chip con un estimado de 20 Mb secuenciadas en un periodo de 2.4 h; Ion 316 Chip con un estimado de 200 Mb en un periodo de 3.1 h e Ion 318 Chip con un estimado de más de 1 Gb en un periodo de 4.5 h. El tamaño de los fragmentos secuenciados varía según la metodología empleada entre las 200 y 400 pb secuenciadas [25].



**Figura 3. Secuenciación por semi-conducción Ion Torrent.** En cada flujo el micro-chip es cubierto con una solución de uno de los nucleótidos. Si el nucleótido se incorpora en la cadena de ADN, se libera un protón, el cual es detectado por la base del pozo, la cual actúa como un mini-pHmetro (transistor de efecto campo sensible a iones ISFET por sus siglas en inglés). Gracias a que cada lámina del microchip es semiconductor, se pueden detectar estos pequeños cambios iónicos. Este proceso se repite con una nueva solución de cada nucleótido, por lo que el micropocillo debe lavarse adecuadamente antes de iniciar el siguiente ciclo [4].

### 1.6 Bioinformática

Desde principios de los 90s, los trabajos de secuenciación se han enfocado al análisis de genomas completos de diversos organismos como protozoarios, parásitos, plantas, virus, bacterias, hongos, levaduras y mamíferos. Mucha de esta información ha sido comparada y analizada para encontrar las variaciones genéticas que se llevan a cabo con el tiempo o aquellas que se presenten de individuo a individuo; sin embargo, la cantidad de información obtenida es de tal magnitud que no resulta conveniente realizar el análisis de manera manual y es aquí en donde la bioinformática entra para facilitar y encontrar soluciones a dichos problemas <sup>[55]</sup>.

La bioinformática es un campo de la ciencia en el cual confluyen varias disciplinas tales como: biología, computación y tecnologías de la información; cuya tarea es facilitar el análisis de datos y desarrollo de nuevas ideas biológicas, así como crear perspectivas globales a partir de las cuales se puedan discernir principios unificadores en biología <sup>[55]</sup>.

El proceso de analizar e interpretar los datos es conocido como biocomputación en donde se engloban otras disciplinas como las matemáticas, la estadística y el desarrollo de nuevos softwares para el diseño de algoritmos enfocados al procesamiento de datos biológicos de una manera más específica y efectiva <sup>[55]</sup>.

Los trabajos más frecuentes en la bioinformática se enfocan en el alineamiento de secuencias, predicción de genes, ensamblado del genoma, alineamiento estructural de proteínas, predicción de estructura de proteínas, predicción de la expresión génica, interacciones proteína-proteína y modelado de la evolución; herramientas que han brindado grandes avances en los diagnósticos genéticos, predicción y tratamiento de enfermedades, rastreabilidad histórica y ubicación geográfica de especies así como una extensa exploración de proteínas y ácidos nucleicos que esta disciplina permite <sup>[55]</sup>.

# *Capítulo 2*

## **JUSTIFICACIÓN**

## 2.0 JUSTIFICACIÓN

Las ETA representan un problema de orden público debido a su ocurrencia que va de la mano con el acelerado incremento en la población y la aparición de grupos vulnerables <sup>[16]</sup>. Una de las estrategias para el combate y control de las ETA, es el uso de antimicrobianos que inhiben la acción de los microorganismos patógenos dentro del consumidor para evitar daños en la salud del mismo. Sin embargo, a lo largo del tiempo las bacterias han logrado desarrollar mecanismos de resistencia a éstos, disminuyendo su tasa de efectividad y abriendo una búsqueda de nuevas estrategias de control.

Se sabe que los genes responsables del combate a antibióticos no son vitales para una bacteria y sólo se observa su expresión bajo determinadas condiciones, es por ello que en su mayoría se encuentran codificados dentro de los plásmidos, los cuales contienen genes propiamente adaptativos que le conferirán ventajas sobre el medio con la finalidad de preservar su especie <sup>[2]</sup>.

Derivado de lo anterior, la Subdirección de Secuenciación y Bioinformática ha dirigido una parte de sus actividades al estudio del contenido genético de bacterias de interés agroalimentario mediante NGS y análisis bioinformáticos de genomas bacterianos. Una vez que se ha realizado una corrida de secuenciación se realiza su análisis para la identificación de cada organismo secuenciado, mediante ensamblés *de novo* y mapeos a un genoma de referencia. Por otra parte, se han encontrado secuencias que no corresponden a ninguna región del cromosoma bacteriano, generando dos situaciones, interferencia en los análisis bioinformáticos y la incógnita del origen de estas secuencias, lo cual, ha llevado a pensar que dichas secuencias pudieran pertenecer a elementos extracromosomales, y de este modo se ha buscado *in silico*, si el origen de esos datos pudiera estar relacionado con algún plásmido ya reportado. Aunado a esto, no existe una estrategia estandarizada para la secuenciación y análisis bioinformático enfocado al estudio exclusivo de plásmidos, que pudiera brindarnos con más certeza el origen de las

secuencias. En el presente trabajo, se diseñó y estandarizó una metodología para el eficiente aislamiento de plásmidos, los cuales pueden alojar dentro de su secuencia nucleotídica, genes responsables de conferir características de patogenicidad, virulencia, resistencia a antibióticos, y técnicas de supervivencia para prevalecer en el medio <sup>[2]</sup>.

Por lo anterior, el ampliar el conocimiento sobre el contenido genético alojado en plásmidos, a través de su purificación y secuenciación, facilitará la comprensión de los mecanismos de adaptación utilizados por cada microorganismo, mismos que proporcionaran información de utilidad para el combate y control epidemiológico al identificar las posibles ventajas adaptativas de cada patógeno, las cuales lleven a su correcto monitoreo y control que funja como apoyo a la inocuidad alimentaria del país.

# *Capítulo 3*

## **HIPÓTESIS**

### **3.0 HIPÓTESIS**

La estandarización de un método de extracción y purificación selectivo para plásmidos, proporcionara material genético adecuado para secuenciarlo, eliminará contaminaciones y sus secuencias podrán ser comparadas con bases de datos para su identificación.

# *Capítulo 4*

## **OBJETIVOS**

### 4.0 OBJETIVOS

#### 4.1 Objetivo general

Proponer y estandarizar una metodología de extracción y purificación de plásmidos de bacterias de interés agroalimentario, así como realizar la secuenciación de dicho material genético para sentar las bases de su identificación mediante el uso de secuenciación masiva y análisis bioinformático.

#### 4.2 Objetivos particulares

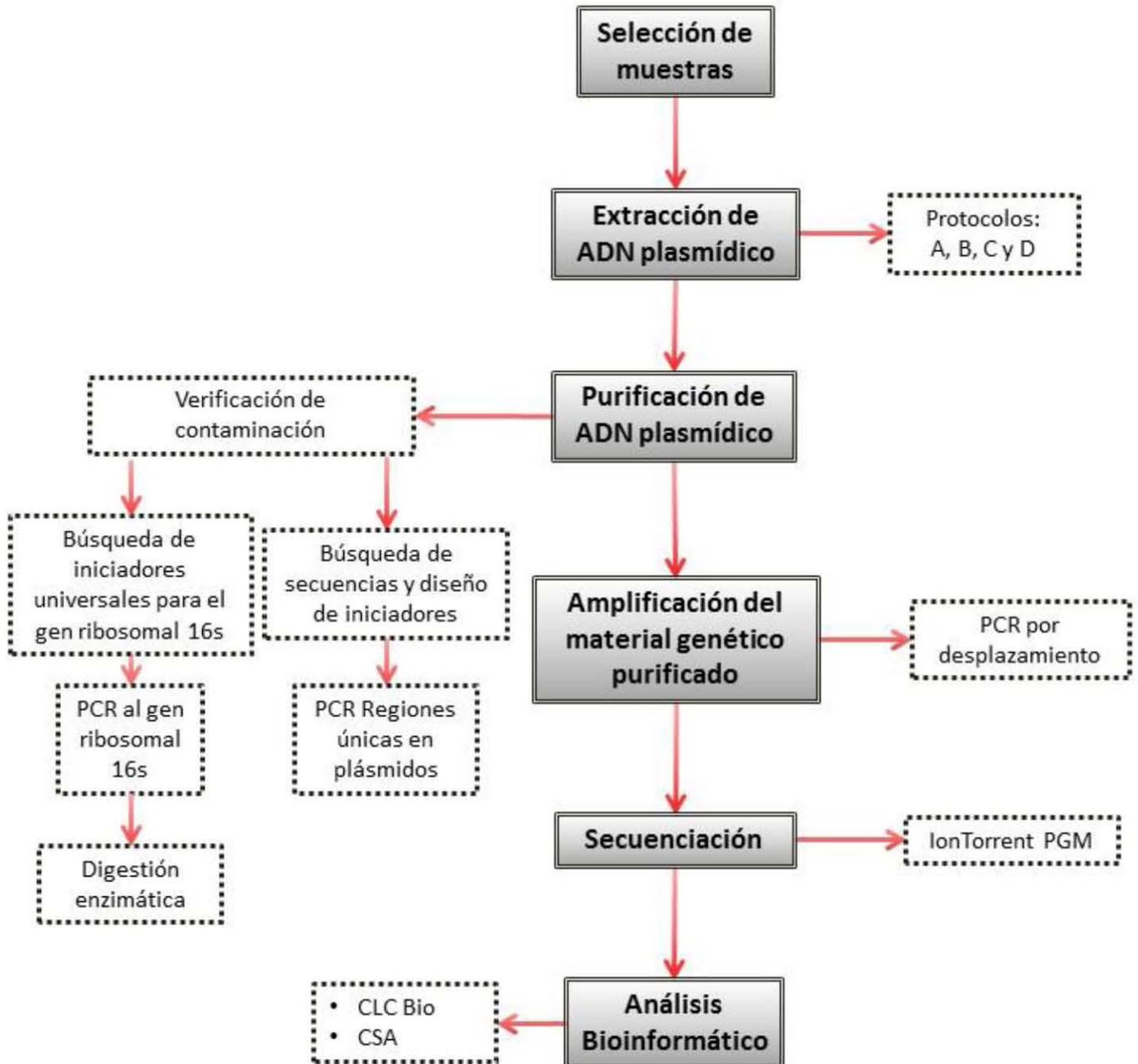
- Realizar una búsqueda bibliográfica de los plásmidos reportados para las bacterias utilizadas en este estudio.
- Estandarizar una técnica de extracción y purificación de ADN plasmídico.
- Diseñar iniciadores para amplificar regiones específicas de ADN plasmídico, así como para el gen ribosomal 16S que sirvan como controles.
- Amplificar el material genético purificado mediante una PCR por desplazamiento utilizando hexámeros al azar.
- Secuenciar los plásmidos aislados mediante la plataforma Ion Torrent PGM e identificarlos mediante análisis bioinformático.

# *Capítulo 5*

## **METODOLOGÍA**

## 5.0 METODOLOGÍA

### 5.1 Diagrama de flujo



## **5.2 Materiales**

### **5.2.1 Material biológico**

Se trabajó con tres cepas bacterianas correspondientes a:

- *Listeria monocytogenes*
- *Escherichia coli*
- *Salmonella entérica* serovar Enteritidis

### **5.2.2 Material de laboratorio**

- Guantes de látex
- Sanitas
- Micro tubos de 2.0, 1.6, 0.6 y 0.2 mL
- Tubos para Qubit
- Tubos cónicos de 15 y 50 mL
- Micro pipetas Rainin de 10, 20, 100, 200, 1000, 2000 y 5000  $\mu$ L
- Puntas con filtro Rainin de 10, 20, 100, 200, 1000, 2000 y 5000  $\mu$ L
- Concentrador magnético de partículas
- Placas de 96 pozos
- Cubiertas adheribles para placas
- Reaction Filter - Ion One Touch 2
- Placa de amplificación - Ion One Touch 2
- Micro tubos para recuperar la emulsión - Ion One Touch 2
- Placas de 8 pozos- Ion One Touch ES
- Punta con filtro para enriquecimiento - Ion One Touch ES
- Ion 316™ Chip
- 7500 DNA chip
- High Sensitivity DNA chip
- E-Gel® Size Select™

**5.2.3 Reactivos**

- Agua grado biología molecular
- Isopropanol grado biología molecular
- Buffer de Lisis (3 % SDS, 50 mM Tris pH 12.6)
- Fenol / cloroformo / alcohol isoamílico (25:24:1 vol./vol./vol.)
- Cloroformo grado biología molecular
- Acetato de Sodio 3.0 M (pH 5.2)
- Etanol al 70 %
- Zyppy™ Plasmid Miniprep Kit
- ZR Plasmid Miniprep™-Classic Kit
- GeneJET Plasmid Miniprep Kit
- Platinum Taq DNA polymerase
- Plasmid-safe ATP-Dependent DNase
- Agilent DNA 7500 Kit
- High Sensitivity DNA Kit
- Illustra™ GenomiPhi DNA Amplification Kit
- Ion Shear Plus Kit
- Ion Plus Fragment Library Kit
- KAPA Library Quantification Kits for Ion Torrent™ platform
- Ion PGM Template OT2 400 Kit
- Ion PGM™ Enrichment Beads
- Ion PGM™ Sequencing 400 Kit

### **5.2.4 Equipos**

- Gabinete de seguridad biológica clase 2
- Sonicador S220 Covaris
- Agitador Vortex-2 Genie
- Microcentrífuga science Med D1008
- Centrífuga eppendorf 5424
- Fluorómetro Qubit 2.0
- Espectrofotómetro UV-visible NanoDrop 2000
- Termociclador punto final GeneAmp PCR System
- Termociclador de Gradiente Veriti
- 2100 Bioanalyzer Agilent
- Vortex mixer - Bioanalizador
- Bloque de Calentamiento eppendorf
- Ion One Touch 2
- Ion One Touch ES
- Ion Torrent PGM

### **5.2.5 Software**

- OligoAnalyzer 3.1
- Primer3Plus
- BLAST v2.4.0
- *In silico* PCR Amplification
- COVARIS SonoLab 7
- 2100 expert
- CLCBio v8.5.1
- CSA v1.11

### 5.3 Métodos

#### 5.3.1 Selección de muestras

Se seleccionaron tres cepas (*L. monocytogenes*, *E. coli* y *S. Enteritidis*), todos por su relevancia a nivel agroalimentario, así como sus implicaciones en la salud del consumidor. El LDDOP de SENASICA realizó la siembra, enriquecimiento y resuspensión de cada bacteria, por lo que se recibieron alícuotas con cultivo bacteriano en microtubos de 2 mL con medio de enriquecimiento, listas para la extracción del material genético.

#### 5.3.2 Extracción de plásmidos.

Se realizó la extracción del material genético plasmídico mediante cuatro metodologías diferentes para evaluar la selectividad y el rendimiento de cada una (En adelante referidos como Protocolo **A**, **B**, **C** y **D**). Los protocolos empleados corresponden a los Kits comerciales: **(A)** Zyppy™ Plasmid Miniprep Kit, **(B)** ZR Plasmid Miniprep™-Classic y **(C)** GeneJET Plasmid Miniprep Kit, así mismo se empleó una metodología ajustada **(D)** de Aya Brown, Itai, & Itzhak, (2013) <sup>[1]</sup>. En la tabla 3 se describe de manera general las características de cada metodología.

**Tabla 3 Cuadro comparativo para cada metodología.**

	<b>Protocolo A</b>	<b>Protocolo B</b>	<b>Protocolo C</b>	<b>Protocolo D</b>
<b>Tipo de muestra</b>	Cultivo enriquecido	Cultivo enriquecido	Cultivo enriquecido	Cultivo enriquecido
<b>ADN blanco</b>	Plásmidos	Plásmidos	Plásmidos	Total
<b>Método de extracción</b>	Lisis alcalina	Lisis alcalina	Lisis alcalina	Lisis alcalina
<b>Método de selección plasmídica</b>	Membranas de afinidad	Membranas de afinidad	Membranas de afinidad	Precipitación
<b>Tiempo de extracción (teóricos)</b>	20 – 25 minutos	15 – 20 minutos	10 – 15 minutos	18 horas

**5.3.2.1 Zyppy™ Plasmid Miniprep Kit (Protocolo A).**

A un volumen de 600 µL de suspensión bacteriana, se agregaron 100 µL de *Buffer de Lisis 7X* y se mezcló por inversión 6 veces, inmediatamente después se agregaron 350 µL de *Buffer de Neutralización* y se invirtió 3 veces para asegurar la neutralización, identificándose mediante un cambio en el medio a una tonalidad amarilla. Posteriormente, la muestra se centrifugó a 16,000 x g por 4 minutos y se transfirió el sobrenadante (cuidando no tocar el botón celular) a la columna *Zymo-Spin™ IIN*, para centrifugar nuevamente a 16,000 x g por 15 segundos. El volumen colectado se decantó y la columna se colocó nuevamente dentro del mismo tubo colector. Para el lavado de la columna se agregaron 400 µL de *Zyppy™ Wash Buffer* y se centrifugó a 16,000 x g por 1 minuto. Se colocó la columna dentro de un tubo limpio de 1.5 mL y se agregaron 30 µL de *Zyppy™ Elution Buffer* directamente a la matriz, se incubó por 1 minuto a temperatura ambiente y se centrifugó a 16,000 x g por 30 segundos para eluir el ADN plasmídico.

**5.3.2.2 ZR Plasmid Miniprep™-Classic (Protocolo B).**

Se trabajó con un volumen de 5 mL de suspensión bacteriana, el cual se centrifugó a 16,000 x g por 5 minutos, el sobrenadante se descartó. Al botón celular se le agregaron 200 µL del *P1 Buffer* y se resuspendió por completo con ayuda del vórtex, una vez resuspendido se agregaron 200 µL del *P2 Buffer (Lisis)* y se mezcló por inversión hasta que la solución tornó a una apariencia purpura y viscosa. Posteriormente se agregaron 400 µL de *P3 Buffer (Neutralización)* y se mezcló suavemente para asegurar una neutralización completa, la cual se pudo evidenciar con un cambio en la tonalidad del medio a color amarillo. Al observar el cambio de tonalidad, se incubó el lisado a temperatura ambiente durante 2 minutos y se centrifugó a 16,000 x g por 2 minutos. El sobrenadante se transfirió a una columna *Zymo-Spin™ IIN* (cuidando de no tomar el botón formado en la etapa anterior) y la columna se colocó dentro de un tubo colector, se centrifugó a 16,000 x g por 30 segundos, se agregaron 200 µL de *Endo-Wash Buffer* y se centrifugó a 16,000 x g por 30 segundos, se adicionaron 400 µL de *Plasmid Wash Buffer* y

nuevamente se centrifugó a 16,000 x g por 1 minuto. Finalmente se transfirió la columna a un tubo limpio de 1.5 mL y se eluyó el contenido con 30 µL de *DNA Elution Buffer*, centrifugando a 16,000 x g por 1 minuto.

### **5.3.2.3 GeneJET Plasmid Miniprep Kit (Protocolo C).**

Se siguió la metodología de este protocolo partiendo de un botón celular proveniente de un volumen de 5 mL de suspensión. El botón se resuspendió con vórtex en 250 µL de *Resuspension Solution*, se agregaron 250 µL de *Lysis Solution* y se mezcló por inversión hasta que la muestra se tornara viscosa y ligeramente clara, se dejó incubar por 2 minutos a temperatura ambiente y se agregaron 350 µL de *Neutralization Solution*, mezclando inmediata y completamente por inversión. Posterior a este paso, se centrifugó a 16,000 x g por 5 minutos. Se transfirió el sobrenadante a una columna *GeneJET spin* y se agregaron 500 µL de *Wash Solution*, se centrifugó a 16,000 x g por 1 minuto, se descartó el volumen colectado y se repitió el paso anterior una vez más. Finalmente se adicionaron 50 µL de *Elution Solution*, se incubó por 2 minutos a temperatura ambiente y se centrifugó a 16,000 x g por 2 minutos para colectar el material genético.

### **5.3.2.4 Metodología ajustada (Protocolo D).**

Se siguió la metodología de Aya Brown, Itai, & Itzhak (2013), con un ajuste en los volúmenes para la cantidad de células empleadas, a continuación, se describe la metodología:

A partir de 5 mL de suspensión bacteriana se obtuvo un botón celular al centrifugar a 16,000 x g por 5 minutos, se decantó el sobrenadante y se agregaron 500 µL de Buffer de Lisis (3 % SDS, 50 mM Tris pH 12.6). Se incubó por 1 hora a 60 °C mezclando el contenido del tubo cada 15 minutos. Se agregaron 500 µL de una mezcla fenol/cloroformo/alcohol isoamílico (25:24:1 vol. /vol. /vol.), se mezcló por completo y se centrifugo a 16,000 x g por 15 minutos, se decantó la fase acuosa y se repitió el paso anterior una vez más. Se extrajo la fase acuosa, a la cual se le agregaron 500 µL de cloroformo y se mezcló completamente, se

centrifugó a 16,000 x g por 15 minutos y se repitió el paso anterior una vez más. La fase acuosa fue transferida a un tubo nuevo de 1.6 mL y se agregaron 5 µL de acetato de sodio 3 M (pH 5.2) y 30 µL de isopropanol. Se incubó por 16 horas a 4 °C. Posterior a la incubación, se centrifugó la muestra a 16,000 x g por 30 minutos y se removió el sobrenadante. Se agregaron 500 µL etanol al 70 %, se mezcló y centrifugó a 16,000 x g por 15 minutos. Finalmente se removió el sobrenadante y se permitió que el botón se secase para resuspenderlo en 30 µL de agua libre de nucleasas.

### **5.3.3 Evaluación de la calidad y cuantificación del ADN extraído**

Los extractos de ADN, deben contar con la calidad adecuada para los procesos de secuenciación, así como estar libres de la mayor cantidad de contaminantes posibles, para asegurar resultados confiables. Por otra parte, la cantidad de ADN aislado también resulta importante, ya que se requiere de una cantidad mínima de 100 ng totales para dar inicio al ensayo.

#### **5.3.3.1 Evaluación de la calidad mediante espectrofotometría**

La calidad de los extractos se evaluó con el equipo NanoDrop 2000 de la marca Thermo Fisher Scientific™, el cual realiza una serie de lecturas en un intervalo de 210-310 nm y registra la absorbancia de la muestra a estas diferentes longitudes de onda, siendo de especial interés la relación establecida por el cociente de Absorbancias 260nm y 280nm (A260/A280). De esta forma se determina la dominancia de ADN en el medio por sobre los posibles contaminantes acarreados durante la extracción.

Para realizar la medición, se realizaron los siguientes pasos:

1. Se inició el software, seleccionando la opción de cuantificación de ADN.
2. Se utilizó como blanco el buffer de elución según la metodología de extracción de ADN utilizada, colocando 1.0 µL de buffer en el puerto de cargado.
3. La muestra se leyó colocando 1.0 µL en el puerto de cargado.

4. Se evaluó el cociente 260/280 nm el cual debe estar en el rango de 1.8 a 2.1, para asegurar una calidad y pureza de ADN aceptable <sup>[23]</sup>.

### **5.3.3.2 Cuantificación mediante fluorimetría**

Se cuantificaron los extractos de ADN utilizando el equipo *Qubit® 2.0 Fluorometer* bajo la modalidad *Broad range*, el cual tiene la sensibilidad para cuantificar cantidades de 2-1000 ng de ADN. Para la cuantificación, el equipo requiere establecer métricas de referencia, las cuales quedan determinadas con la lectura de dos estándares, utilizados para generar una curva de calibración y así, interpolar los valores de fluorescencia obtenidos en cada muestra y de este modo estimar la concentración del ADN extraído.

Para realizar la medición, se siguieron los siguientes pasos:

1. Se preparó un volumen de solución de trabajo agregando 199  $\mu\text{L}$  de Qubit Buffer y 1  $\mu\text{L}$  de Reactivo DNA Br.
2. Se prepararon dos estándares agregando 190  $\mu\text{L}$  solución de trabajo y 10  $\mu\text{L}$  de estándar correspondiente, para generar en el equipo una curva de calibración.
3. Se prepararon las muestras agregando a un tubo Assay Qubit® 199  $\mu\text{L}$  solución de trabajo y 1  $\mu\text{L}$  de muestra.
4. Muestras y estándares se mezclaron e incubaron durante 2 minutos a temperatura ambiente.
5. Pasado el tiempo de incubación, se realizó la cuantificación indicando en el equipo el volumen de muestra utilizado para obtener el valor de cuantificación correcto.

### **5.3.4 Purificación del ADN plasmídico**

La purificación tuvo como objetivo dirigir el trabajo sólo al asilamiento del ADN plasmídico, por lo cual resultó de particular importancia remover los fragmentos contaminantes de origen cromosomal.

**5.3.4.1 Búsqueda de contaminantes cromosomales**

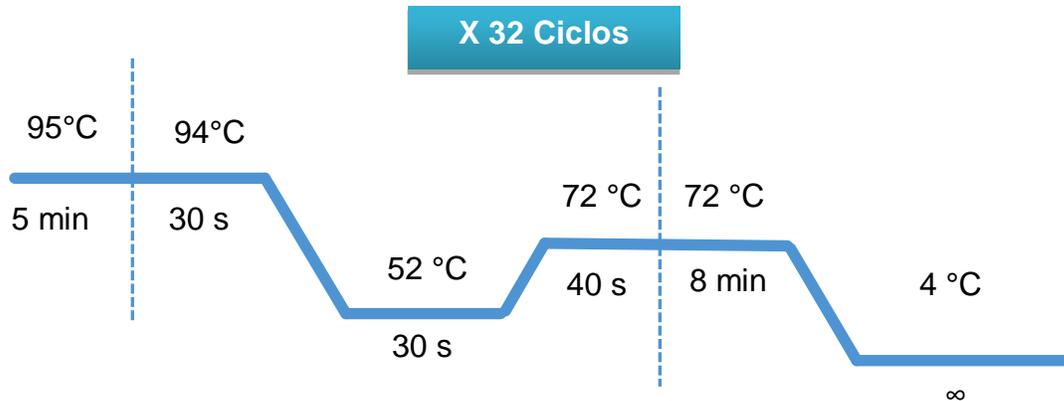
Para verificar la presencia o ausencia de ADN cromosomal, se diseñó una estrategia experimental mediante amplificación por PCR dirigida al gen ribosomal 16S contenido en el cromosoma bacteriano. Para esto, se emplearon los iniciadores universales mostrados en la Tabla 4 [22], así como la mezcla de reacción y condiciones de PCR mostradas en la Tabla 5 y Figura 4 respectivamente.

**Tabla 4. Iniciadores universales para gen ribosomal 16S.**

Oligonucleótido		Secuencia	Tamaño de fragmento esperado
27F	Sentido	5'- AGA GTT TGA TCM TGG CTC AG -3'	~1500 pb
1492R	Antisentido	5'- CGG TTA CCT TGT TAC GAC TT -3'	

**Tabla 5. Mezcla de reacción para la amplificación del gen Ribosomal 16S.**

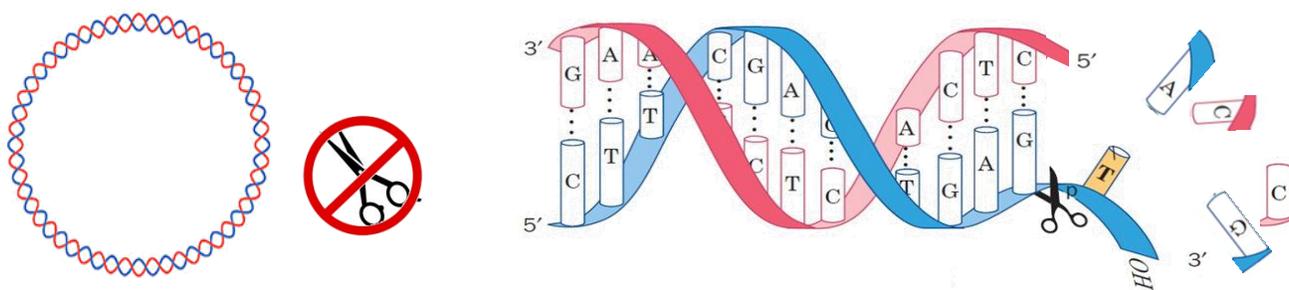
REACTIVO	VOLUMEN $\mu$ L
Buffer PCR 10X	5.0
DNTP's (10mM)	1.0
Iniciador sentido (10 $\mu$ M)	1.0
Iniciador antisentido (10 $\mu$ M)	1.0
Platinum® Taq DNA Polymerase High Fidelity	0.2
ADN (1.0 ng/ $\mu$ L)	3.0
Agua libre de nucleasas	38.8
Volumen final	50.0



**Figura 4. Condiciones para la amplificación del gen ribosomal 16S.** Se describe el proceso de amplificación detallando tiempos y temperaturas. Apertura inicial de cadenas a 95°C por cinco minutos; 32 ciclos de amplificación con una apertura de cadenas a 94°C por 30 segundos, alineamiento de los iniciadores a las cadenas molde a 55°C por 30 segundos y una elongación de las cadenas 72°C por 40 segundos; elongación final a 72°C por ocho minutos para completar la amplificación de todas las cadenas.

#### 5.3.4.2 Remoción de ADN genómico

En los casos en los que se observó amplificación en la reacción anterior por PCR, se procedió a eliminar el ADN cromosomal residual, mediante una reacción por digestión enzimática. A cada reacción se le agregó 5.0  $\mu\text{L}$  de DTT (10 mM) como agente reductor inhibiendo la formación de puentes disulfuro, evitando las modificaciones conformacionales dentro de la enzima y asegurando su funcionalidad <sup>[5]</sup>, se incubó a 97 °C por 2 minutos para relajar las cadenas y facilitar el acoplamiento de la enzima sobre las cadenas de ADN. La reacción de digestión se realizó con el kit *Plasmid-safe ATP-Dependent DNAsa*, el cual utiliza una exonucleasa que degrada los fragmentos de doble cadena con extremos libres presentes en el medio (Figura 5). Adicional al protocolo del kit, se adicionó la enzima *Exo I* para digerir todos los fragmentos de cadena sencilla eliminando los iniciadores residuales de la primera reacción de PCR.



**Figura 5. Digestión enzimática para la eliminación de ADN cromosomal.** Degradación de todos los fragmentos con extremos libres, dejando intactos en el medio a los elementos circulares o con extremos protegidos

La mezcla de reacción empleada se describe en la Tabla 6, misma que se efectuó durante un tiempo de 30 minutos a 37 °C.

**Tabla 6. Mezcla por reacción para la digestión enzimática.**

<b>REACTIVO</b>	<b>VOLUMEN <math>\mu</math>L</b>
Buffer de reacción	5.0
Plasmid-safe DNAsa 100U/mL	5.0
Exo I	3.0
DTT (10mM)	5.0
ATP (25 mM)	2.0
ADN	5.0
Agua libre de nucleasas	25.0
Volumen total	50.0

Posterior a la digestión enzimática se realizó una segunda PCR dirigida al gen ribosomal 16S, para verificar la eliminación de contaminantes cromosomales. La reacción se realizó con la mezcla de reacción y condiciones descritas en la Tabla 5 y Figura 4 respectivamente.

**5.3.4.3 Búsqueda de ADN plasmídico**

Una vez que se verificó la ausencia del amplicón correspondiente al gen ribosomal 16S, se planteó una PCR punto final a regiones únicas en los plásmidos de cada bacteria en estudio, para comprobar su presencia en el medio. Para esto, se realizó la búsqueda de las secuencias de plásmidos reportadas en la base de datos de referencia Ref-Seq del NCBI y posteriormente se diseñaron sus respectivos iniciadores.

**5.3.4.3.1 Diseño de iniciadores específicos**

Obtenidas las secuencias de referencia de plásmidos, se diseñaron iniciadores para la amplificación de regiones específicas de cada uno de ellos. Se utilizó el software en línea *Primer3Plus*, en el cual se establecieron los criterios necesarios para el diseño de cebadores (tamaño, temperatura de alineamiento, % de GC, así como longitud del amplicón), Tabla 7.

**TABLA 7. Criterios para el diseño de iniciadores.**

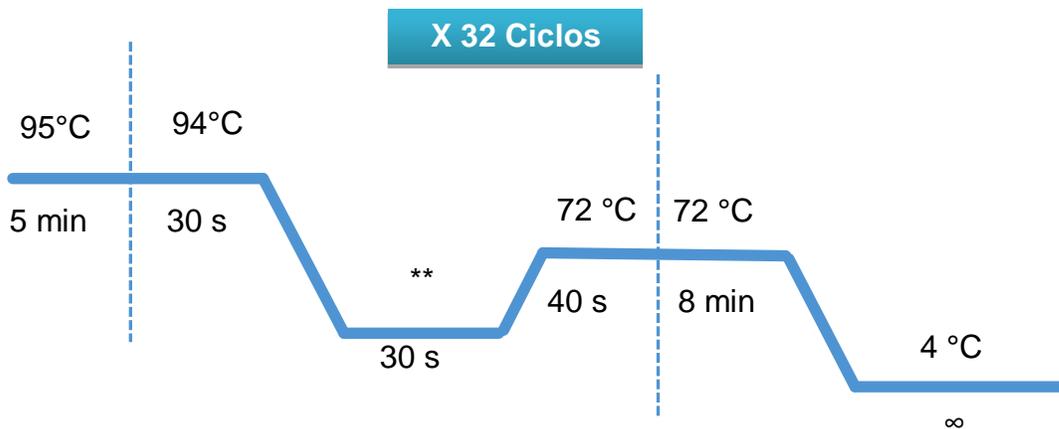
CRITERIO	VALOR
LONGITUD DEL INICIADOR	<b>20-25 PB</b>
TEMPERATURA DE ALINEAMIENTO	<b>50-55 °C CON UNA VARIACIÓN MENOR A 2 °C ENTRE CADA PAR DE INICIADORES</b>
% GC	<b>40-60 %</b>
<i>SIN FORMACIÓN DE DÍMEROS NI ESTRUCTURAS SECUNDARIAS ENTRE CADA INICIADOR</i>	

**5.3.4.3.2 Amplificación de regiones únicas en plásmidos**

Una vez diseñados los iniciadores, se prepararon las reacciones de amplificación, para lo cual se utilizó la mezcla de reacción, así como las condiciones descritas en la Tabla 8 y Figura 6 respectivamente.

**Tabla 8. Mezcla de reacción para la amplificación las regiones únicas en plásmidos.**

REACTIVO	VOLUMEN $\mu\text{L}$
Buffer PCR 10X	5.0
DNTP's (10mM)	1.0
Iniciador forward (10 $\mu\text{M}$ )	1.0
Iniciador reverse (10 $\mu\text{M}$ )	1.0
Platinum® Taq DNA Polymerase High Fidelity	0.2
ADN (1.0 ng/ $\mu\text{L}$ )	3.0
Agua libre de nucleasas	38.8
Volumen final	50.0



**Figura 6. Condiciones para la amplificación de regiones únicas en plásmidos.**

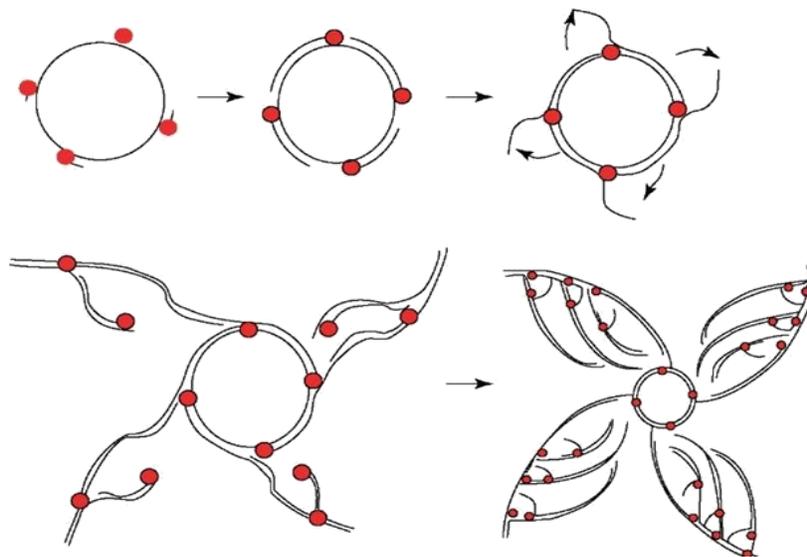
Apertura de cadenas inicial a 95°C por cinco minutos; 32 ciclos de amplificación con una apertura de cadenas a 94°C por 30 segundos, alineamiento de los iniciadores a las cadenas molde a diferentes temperaturas por 30 segundos y una elongación de las cadenas 72°C por 40 segundos; elongación final a 72°C por ocho minutos para completar la amplificación de todas las cadenas.

\*\* Se emplearon dos Tm diferentes para cada bacteria: *L. monocytogenes* a) 61.6 °C y b) 61.6 °C; *E. coli* a) 55.2 °C y b) 54.3 °C; *S. Enteritidis* a) 62.3 °C y b) 61.9 °C.

**5.3.5 Amplificación de los plásmidos purificados**

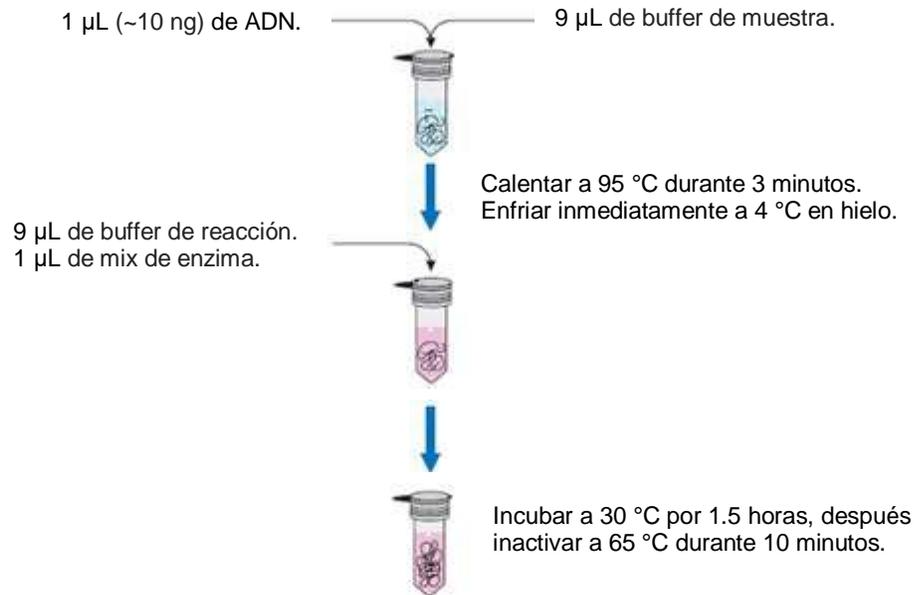
Después de verificar que el amplicón de gen ribosomal 16S se encontraba ausente y la amplificación de las regiones únicas en plásmidos resultara positiva, se procedió a amplificar todo el material genético presente en cada muestra, para contar con una cantidad suficiente de ADN y así realizar una corrida de secuenciación por la plataforma Ion Torrent PGM™.

Para realizar la amplificación se utilizó *Illustra™ GenomiPhi DNA Amplification Kit*, el cual actúa mediante una PCR isotermal por desplazamiento. En esta metodología se utilizan hexámeros al azar (Random primers) con un tamaño de seis pares de bases, lo que incrementa la posibilidad de hibridar con un mayor número de sitios sobre la secuencia del plásmido. En la Figura 7 se muestra el mecanismo de acción mediante el cual se desarrolla la amplificación [18].



**Figura 7. Amplificación con hexámeros al azar.** Los hexámeros al azar (representados como puntos rojos) hibridan en más de un sitio sobre la secuencia molde del plásmido, comienza la extensión y al encontrarse con un nuevo punto de unión, desplaza la cadena ya amplificada para continuar con la amplificación. Al desprender la cadena complementaria se tiene un nuevo fragmento disponible para que los hexámeros vuelvan a hibridar y se pueda repetir el proceso.

Para la reacción se ocuparon los reactivos y cantidades ilustrados en la Figura 8.



**Figura 8. Metodología de amplificación del material genético con Illustra™ GenomiPhi DNA Amplification Kit.**

Procesada la muestra, se realizó la cuantificación del producto amplificado con el equipo *Qubit® 2.0 Fluorometer* (Ver sección 5.3.3.2), utilizando los reactivos del kit *High sensitivity*, el cual determina cantidades entre 0.2-100 ng de ADN. Así mismo, se evaluó la calidad del material genético utilizando el equipo *NanoDrop 2000 UV-Vis Spectrophotometer* (Ver sección 5.3.3.1). Las muestras que obtuvieron un valor en la relación 260/280 entre 1.8 y 2.19 fueron utilizadas para preparar la corrida de secuenciación.

### 5.3.6 Secuenciación

#### 5.3.6.1 Generación de Bibliotecas Genómicas

El proceso de secuenciación da inicio con la elaboración de las bibliotecas genómicas, en las cuales se ajustan los tamaños de los fragmentos según la metodología a emplear mediante una fragmentación al azar, se acondicionan estos fragmentos con la ligación de adaptadores, selección de tamaños a

fragmentos entre 350-450 pb, amplificación y enriquecimiento de las mismas para generar una buena cantidad de librería con una calidad adecuada.

### **5.3.6.1.1 Fragmentación del material genético**

Para cada bacteria se emplearon dos metodologías de fragmentación. La primera se realizó por emisión de ondas ultrasónicas, la cual se llevó a cabo con el *Sonicador S220 de la marca COVARIS™*, en tanto que la segunda se hizo vía enzimática con el *Kit Ion Shear Plus* de la marca Thermofisher la cual es la recomendada por la plataforma Ion Torrent. Ambas metodologías se muestran a continuación:

#### **a) Fragmentación por ultrasonicación**

En un microtubo de Qubit® se colocaron 100 ng de material genético purificado y amplificado en un volumen de 50 µL de agua libre de nucleasas. El ensayo se llevó a cabo de acuerdo al protocolo del equipo S220 COVARIS™, el cual se detalla en seguida:

1. La preparación con los 100 ng de ADN se colocó en la estación de cargado del equipo y se introdujo en el tanque con agua destilada, se establecieron las condiciones en el software COVARIS SonoLab 7, las cuales se muestran en la Tabla 9.
2. Al terminar la sonicación, se analizó una alícuota de 1 µL del ADN fragmentado, para confirmar un intervalo de tamaño de fragmentos en donde la mayoría de ellos se encuentre entre 350-450 pb, en el equipo Bioanalyzer 2100 Agilent en un chip High Sensitivity DNA (Véase Anexo 1).

**Tabla 9. Condiciones de ultrasonicación para la fragmentación.**

<b>PARÁMETRO</b>	<b>CONDICIONES</b>
<b>PIP</b>	200 Watts
<b>DutyFactor</b>	30 %
<b>Ciclos</b>	200
<b>Tiempo</b>	460 s

Debido al daño mecánico aleatorio provocado durante la sonicación sobre las cadenas de ADN, fue necesario realizar una reparación para obtener extremos romos, sobre los cuales posteriormente se ligaron adaptadores. La reparación se realizó agregando al tubo de fragmentación los reactivos descritos en la Tabla 10, dejando incubar a temperatura ambiente durante 20 minutos.

**Tabla 10. Mezcla de reacción para la reparación de extremos.**

<b>REACTIVO</b>	<b>VOLUMEN <math>\mu</math>L</b>
Agua grado biología molecular	29.0
End repair buffer	20.0
End repair enzyme	1.0
ADN fragmentado	50.0

**b) Fragmentación Enzimática**

Para la fragmentación enzimática, se colocaron en un microtubo de 1.6 mL, 100 ng de material genético en un volumen final de 35  $\mu$ L de agua libre de nucleasas. Posteriormente se empleó el *Ion Shear Plus Kit*, el cual utiliza una mezcla de endonucleasas para generar tamaños entre 350-450 pb. Para esto se colocó la mezcla de reacción de la Tabla 11 en el tubo con los 100ng de ADN y se incubó durante 8.5 min a 37°C, transcurrido el tiempo de incubación se agregaron 5.0  $\mu$ L de Ion Shear Stop Buffer para detener la reacción.

**Tabla 11. Mezcla de reacción para la fragmentación enzimática.**

<b>REACTIVO</b>	<b>VOLUMEN <math>\mu</math>L</b>
Ion shear plus 10X reaction buffer	5.0
Ion shear plus enzyme mix II	10.0
ADN	35.0

### **5.3.6.1.2 Purificación de material genético fragmentado**

Una vez fragmentado el material genético, se realizó una purificación con perlas magnéticas, para eliminar los componentes extras de la reacción y de esta manera trabajar únicamente con el ADN contenido en cada muestra. La purificación se realizó con perlas *Agencourt AMPure* como se muestra a continuación:

1. Se atemperaron las perlas *Agencourt AMPure* durante 10 minutos.
2. Se homogenizó por 1 minuto el contenido del frasco mediante vórtex.
3. Se alicuotaron las perlas *Agencourt AMPure* en un tubo de 1.6 mL (99  $\mu$ L para el método enzimático y 180  $\mu$ L para el método de sonicación).
4. Se agregó a las perlas el contenido de la reacción de fragmentación.
5. Se mezcló el tubo con vórtex y se centrifugó con un spin rápido.
6. Se Incubó a temperatura ambiente por 5 min.
7. Se colocó el tubo en un rack magnético durante 3 min.
8. Se descartó el sobrenadante.
9. Se realizó un lavado con 500  $\mu$ L de EtOH al 70% recién preparado. En un tiempo no mayor a 30 segundos se dieron dos giros de 180° a cada tubo para realizar un correcto lavado de las perlas y se eliminó el EtOH.
10. Se repitió el paso anterior.
11. Se dejó el tubo destapado durante 5 minutos para que el EtOH residual se evaporara.
12. Se eluyó con 25  $\mu$ L de Buffer Low TE, dejando interactuar al buffer con las perlas por 1 min y colocándolo nuevamente en él magneto.
13. Finalmente se recuperó el material genético contenido en el sobrenadante.

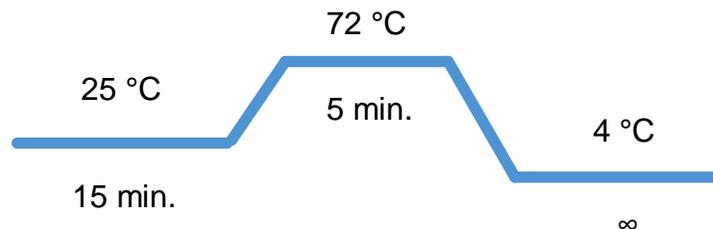
### **5.3.6.1.3 Ligación de adaptadores.**

Una vez obtenidas las secuencias fragmentadas y con extremos romos, se realizó la ligación de los adaptadores mediante una reacción de ligación. Para este paso se utilizó *Ion Plus Fragment Library Kit* y adaptadores con identificadores (Barcode) moleculares dentro de su secuencia. Se preparó el mix para la ligación de los adaptadores como se muestra en la Tabla 12:

**Tabla 12. Mezcla de reacción para la ligación de adaptadores.**

<b>REACTIVO</b>	<b>VOLUMEN <math>\mu\text{L}</math></b>
ADN	25.0
Buffer de ligasa 10X	10.0
Barcode	2.0
Adaptador P1	2.0
DNTP's Mix	2.0
DNA ligasa	2.0
Polimerasa nick repair	8.0
Agua libre de nucleasas	49.0

La mezcla de reacción para cada muestra se colocó dentro del termociclador y se ejecutó el programa mostrado en la Figura 9.



**Figura 9. Programa para la ligación de adaptadores.** Se lleva a cabo el proceso de ligación uniendo las secuencias de los adaptadores a los fragmentos de ADN plasmídico.

Al igual que en la fragmentación, la librería se purificó con perlas magnéticas *Agencourt AMPure*, siguiendo la metodología antes descrita en el paso 5.3.6.1.2 modificando el volumen inicial de perlas a 120  $\mu\text{L}$  y eluyendo con 20  $\mu\text{L}$  de Low TE.

**5.3.6.1.4 Selección de tamaños**

La selección de tamaños se realizó con el sistema E-Gel® Size Select™, empleando un gel de agarosa al 2 %, para separar sólo los fragmentos con longitud aproximada de 480 pb, tamaño estimado de las bibliotecas genómicas (ADN + adaptadores) de acuerdo a los siguientes pasos:

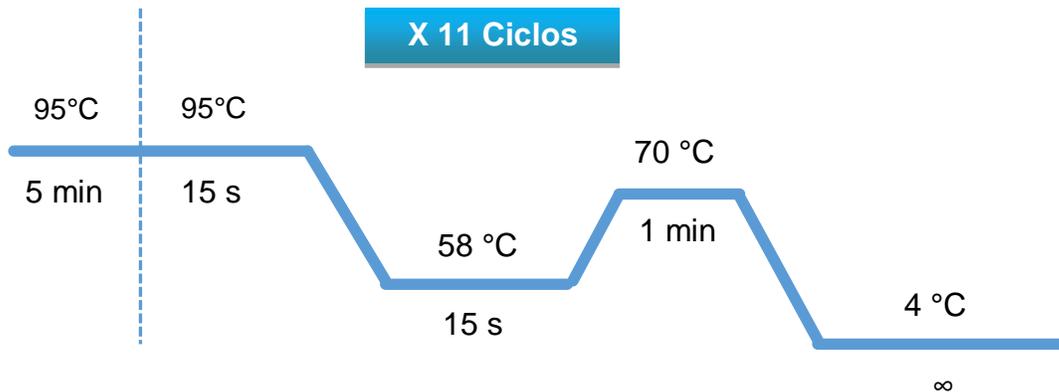
1. Se colocó el gel de agarosa al 2% en la cámara de electroforesis y se seleccionó el programa Size Select 2%.
2. Se cargaron los 20 µL de la muestra previamente purificada en los pozos superiores, y se agregaron 10 µL de marcador de peso molecular en el pozo central del cassette.
3. Se adicionaron 25 µL de agua libre de nucleasas en todos los pozos vacíos.
4. Se corrió programa Size Select 2%. Monitoreando con el transiluminador la posición de las bandas para su posterior recuperación.
5. Cuando la banda de 500 pb se encontraba próxima al pozo de recuperación, se detuvo la corrida y se recuperaron las librerías del pozo colector, agregando 10 µL de agua libre de nucleasas por duplicado.

**5.3.6.1.5 Amplificación de la biblioteca genómica**

Una vez seleccionados los fragmentos del tamaño deseado, se realizó una amplificación del material genético seleccionado con la finalidad de tener al menos 100 ng de librería [6]. Para esto, se preparó la mezcla de reacción descrita en la Tabla 13 para su posterior amplificación bajo las condiciones mostradas en la Figura 10.

**Tabla 13. Mezcla de reacción para la amplificación de la biblioteca genómica.**

<b>REACTIVO</b>	<b>VOLUMEN µL</b>
Platinum PCR SuperMix High Fidelity	100.0
Library Amplification Primer Mix	5.0
Librería sin amplificar	25.0

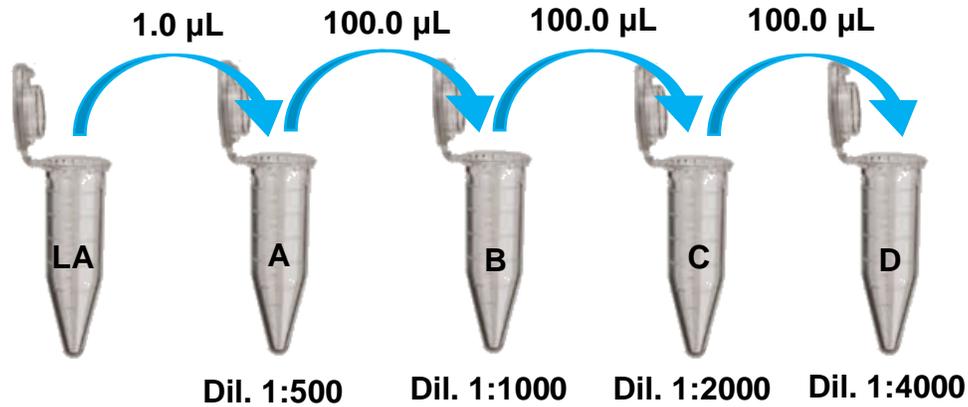


**Figura 10. Condiciones de tiempo y temperatura para la amplificación de la librería.** Se lleva a cabo una PCR convencional sin importar la variación entre el contenido nucleotídico de cada fragmento, ya que en la ligación los adaptadores se adicionan secuencias universales para las librerías genómicas de IonTorrent, en donde se simula una secuencia conservada para todos los fragmentos sobre la cual se alinean los iniciadores y se permite realizar la amplificación de todas los fragmentos involucrados en la reacción.

#### 5.3.6.1.6 Cuantificación de las bibliotecas mediante PCR tiempo real.

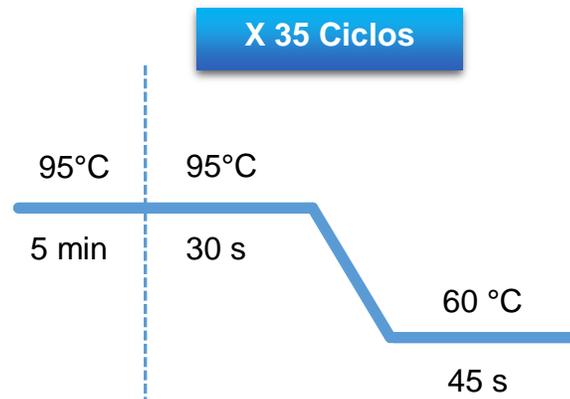
Posterior a la amplificación se realizó la cuantificación de los productos mediante PCR tiempo real. Para realizar este ensayo se utilizaron los reactivos *KAPA Library Quantification Kits for Ion Torrent™ platform* y el equipo *LightCycler® 480* de Roche, realizando el siguiente procedimiento:

1. Se preparó un tubo (A) con 499.0  $\mu\text{L}$  y tres tubos (B, C y D) con 100  $\mu\text{L}$  de agua libre de nucleasas para cada dilución de librería amplificada (LA). Posteriormente se realizaron las diluciones siguiendo el diagrama mostrado en la Figura 11.



**Figura 11. Diagrama de preparación de diluciones para cada librería.** Se tomó 1.0 µL de librería amplificada (LA) y se diluyó en 499.0 µL de agua libre de nucleasas obteniendo la dilución (A) 1:500. A partir de esta, se realizaron diluciones en serie tomando 100.0 µL de la dilución previa y diluyendo en 100.0 µL más de agua libre de nucleasas para obtener las diluciones: (B) 1:1000; (C) 1:2000 y (D) 1:4000.

2. Se preparó un tubo por librería diluida con 18 µL de KAPA SYBR® FAST qPCR Master Mix y 12 µL de templado para realizar lecturas por triplicado con reacciones de 10 µL cada una.
3. Adicional a cada librería, se preparó una curva de calibración con seis estándares de concentración conocida agregando 18 µL de KAPA SYBR® FAST qPCR Master Mix y 12 µL de cada estándar, para realizar lecturas por triplicado con reacciones de 10 µL cada una.
4. Se realizó la cuantificación siguiendo el programa descrito en la Figura 12.



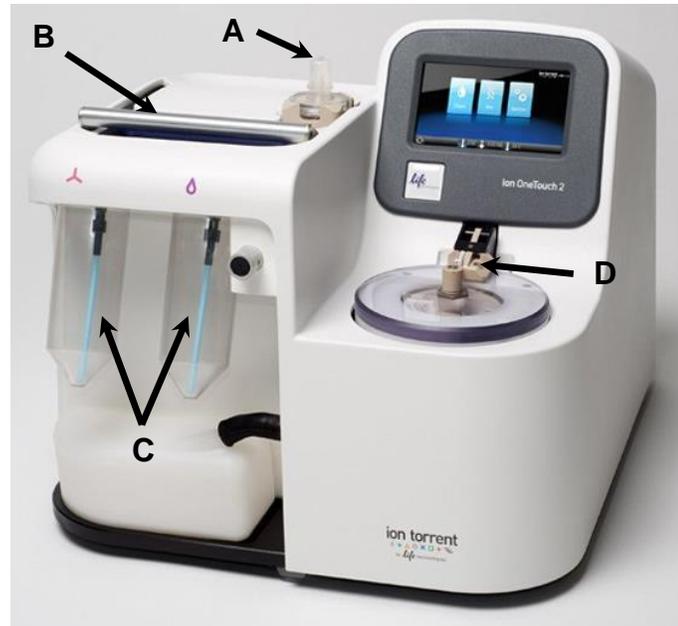
**Figura 12. Programa para la cuantificación por PCR tiempo real.** Apertura de cadenas a 95°C por cinco minutos; 35 ciclos de amplificación con una apertura de cadenas a 95°C por 30 segundos y alineamiento de iniciadores, amplificación y cuantificación a 60 °C por 45 segundos.

Una vez conocida la cantidad de ADN contenido en cada tubo de amplificación, se calculó el factor de dilución para las librerías amplificadas y así obtener una concentración de  $1.5 \times 10^7$  moléculas/ $\mu\text{L}$ . De cada librería diluida a la concentración deseada, se tomó una alícuota de 25  $\mu\text{L}$ , para obtener una mezcla con todas las genotecas a la misma concentración, de esta mezcla se tomaron 25  $\mu\text{L}$  para continuar con el proceso de em-PCR.

### 5.3.6.2 Amplificación clonal por em-PCR

Una vez obtenida la cantidad adecuada de bibliotecas genómicas, estas fueron ligadas a perlas acarreadoras para generar la amplificación clonal en un sistema de emulsión agua en aceite, de modo que cada micela funge como un micro reactor en donde se lleva a cabo la amplificación. De esta forma se favorece que cada perla contenga un único fragmento, el cual se amplifica mediante PCR. El ensayo de amplificación se realizó con el *Ion PGM Template OT2 400 Kit* y el sistema automatizado *Ion One Touch 2* (Figura 13) siguiendo las siguientes especificaciones:

1. En un microtubo de 1.5 mL se colocaron los siguientes reactivos, respetando el orden indicado y se mezcló mediante pipeteo.
  - Ion PGM™ Template OT2 400 Reagent Mix 500 µL
  - Ion PGM™ Template OT2 400 PCR Reagent B 285 µL
  - Ion PGM™ Template OT2 400 Enzyme Mix 50 µL
  - Ion PGM™ Template OT2 400 Reagent X 40 µL
  - Pool de Librerías ( $1.5 \times 10^7$  moléculas/µL) 25 µL
2. Se dio vórtex a la solución preparada en el paso anterior a velocidad máxima durante 5 segundos, después se centrifugó la solución durante dos segundos y se agregaron 100 µL de Ion PGM Template OT2 400 Ion Sphere Particles (ISP´s).
3. Se mezcló perfectamente a máxima velocidad en vórtex 5 segundos más.
4. Se colocaron los reactivos e insumos necesarios en el equipo Ion One Touch.
5. Se cargó la librería en el filtro de reacción y añadieron 1000 µL del aceite de reacción.
6. Se colocó el filtro en el puerto correspondiente y se inició la amplificación clonal en emulsión (el proceso tarda aproximadamente 5.5 h).
7. Terminada la emulsión, se centrifugó dentro del equipo y al finalizar se descartó el exceso de la solución, dejando únicamente 100 µL de cada tubo de recuperación para conservar las perlas con la librería adherida.



**Figura 13. Termociclador de emPCR Ion One Touch 2.** (A) Filtro de reacción, contiene todos los reactivos para la emPCR. (B) Ranura para colocar la placa de amplificación en el termobloque. (C) Ion OneTouch™ Oil  y Ion PGM™ OT2 Recovery Solution . (D) Centrifuga incorporada para instalar los tubos de recuperación y el router de recuperación.

#### 5.3.6.2.1 Enriquecimiento

Finalizada la amplificación, se preparó el proceso de enriquecimiento para recuperar únicamente las perlas vivas (perlas unidas a librería) y descartar aquellas que permanecieron vacías durante el proceso. Para esto se empleó el equipo Ion One Touch ES (Figura 14) y los reactivos *Ion PGM™ Template OT2 Solutions 400 Kit*, así como las perlas *Ion PGM™ Enrichment Beads*, para el cual se acondicionó la muestra de la siguiente manera:

1. Con una punta nueva se resuspendió el botón formado en los 100  $\mu$ L de librería amplificada de ambos tubos con muestra, hasta dispersar la totalidad de perlas en el medio.
2. Se agregaron 500  $\mu$ L de Ion OneTouch™ Wash Solution a cada tubo de recuperación y se resuspendieron las ISP's mezclando por pipeteo.
3. Se transfirió el contenido de ambos tubos a un tubo nuevo de 1.5  $\mu$ L debidamente etiquetado.

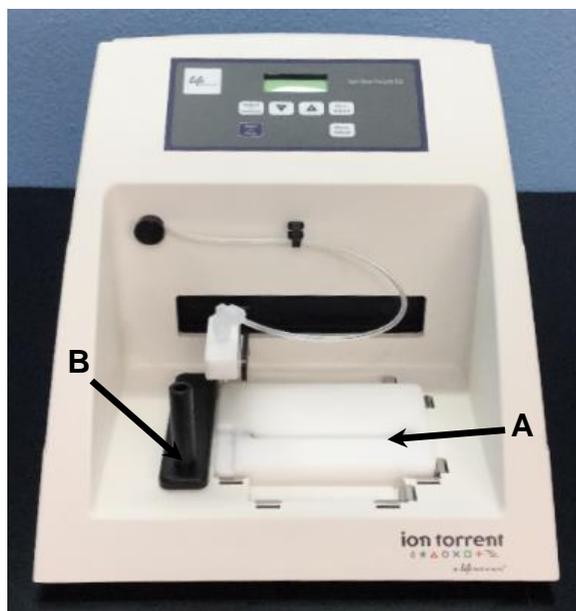
4. Se calentaron las ISP's a 50 °C por 2 minutos.
5. Se centrifugaron las ISP's por 2.5 minutos a 15,500.
6. Con una pipeta, se retiró todo el volumen excepto 100 µL. Se retiró sobrenadante desde la superficie y desde la pared opuesta a la formación del botón.

Una vez acondicionada la muestra se realizó el enriquecimiento mediante el equipo Ion OneTouch™ ES con las especificaciones del fabricante:

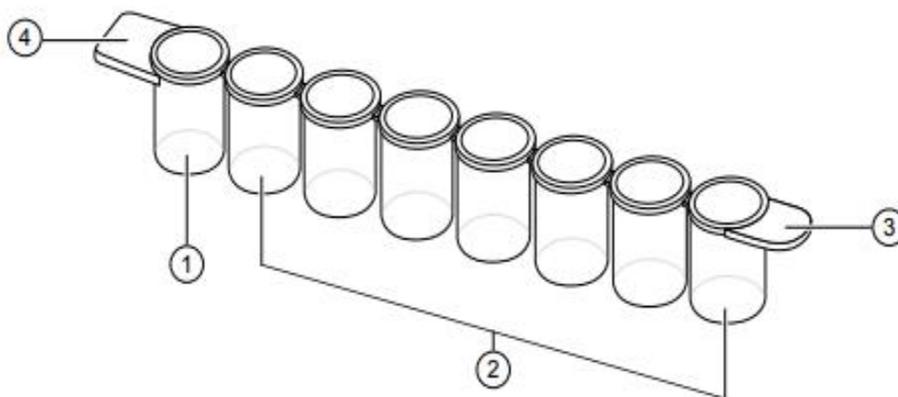
1. Se preparó la solución Melt-Off añadiendo 280 µL de Tween® Solution y 40 µL de NaOH 1 M.
2. Se resuspendieron 13 µL de Dynabeads MyOne Streptavidin beads en 130 µL de MyOne Beads Wash Solution.
3. Se cargó la tira de 8 pozos de la siguiente manera y se aseguró colocar el extremo con la pestaña cuadrada del lado derecho: (Figura 15).

1. Librería amplificada	100 µL
2. <i>My One Dynabeads</i>	130 µL
3. <i>Ion OneTouch Wash Solution</i>	300 µL
4. <i>Ion OneTouch Wash Solution</i>	300 µL
5. <i>Ion OneTouch Wash Solution</i>	300 µL
6. Vacío	
7. Solución <i>Melt-off</i>	300 µL
8. Vacío	

4. Se colocó la tira de 8 pozos en el equipo Ion One Touch ES y se comenzó el enriquecimiento.
5. Se agregaron 10 µL de solución de neutralización en un tubo para PCR de 0.2 mL.
6. Se colocó el tubo de PCR abierto con la solución de neutralización dentro del orificio colocado en la base de la punta del equipo, el cual servirá como tubo colector de la solución enriquecida contenedora de las perlas acarreadoras (mono y policlonales) con fragmentos de ADN adheridos a su superficie.



**Figura 14. Sistema de enriquecimiento Ion One Touch.** (A) Ranura base para colocar la tira de 8 pozos para el enriquecimiento. (B) Orificio del tubo colector para la solución enriquecida.



- ① Pozo 1
- ② Pozo 2-8
- ③ Pestaña redonda
- ④ Pestaña cuadrada

**Figura 15. Tira de 8 pozos para el enriquecimiento de las librerías amplificadas.** Se muestra la posición en que habrá que colocar la tira y el orden en el cual se deberán agregar los reactivos descritos en el párrafo superior.

### 5.3.6.3 Corrida de secuenciación

La corrida de secuenciación se realizó en el equipo Ion Torrent PGM™ para el cual se llevó a cabo un prelavado de las vías internas del equipo, se prepararon los reactivos necesarios para el secuenciador utilizando el estuche de secuenciación *Ion PGM™ Sequencing 400 Kit*, y posteriormente se preparó la muestra a secuenciar agregando los iniciadores de secuenciación y la polimerasa para su cargado en un chip tipo 316 v2, a continuación, se describe el proceso:

Prelavado de líneas internas:

1. Se preparó una solución clorada empleando agua Milli-Q 18 MΩ.
2. Se realizó la inicialización del equipo colocando y preparando los insumos y reactivos necesarios para el ensayo (buffers, recipientes, dNTP's).

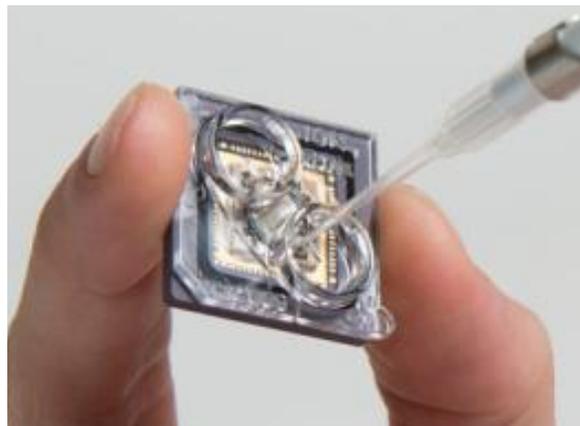
Preparación de la muestra:

1. Se añadieron 5 µL de Control Ion Sphere Particles (librería control) al total de las perlas recuperadas en el enriquecimiento.
2. Se centrifugó por 2 min a 15500 × g.
3. Se eliminó el sobrenadante sin perturbar las perlas y se dejaron aproximadamente 15 µL en el tubo.
4. Se añadieron 12 µL del primer de secuenciación y se alineó a la mezcla anterior, colocando la reacción en termociclador bajo las siguientes condiciones:

95 °C	2 min.
37 °C	2 min.
5. Se añadieron 3 µL de la Ion PGM™ Sequencing 400 Polymerase para tener un volumen final de 30 µL y se incubó por 5 minutos a temperatura ambiente.
6. Se realizó el chip-check en el equipo Ion Torrent PGM™.

Finalizado el chip check, se procedió a realizar el cargado de la librería en el chip 316 v2 (Figura 16), enfatizando en realizar todo este procedimiento sin guantes y sobre el soporte plástico de su centrifuga para evitar cualquier daño a causa de las posibles cargas electrostáticas, siguiendo los pasos que a continuación se enlistan:

1. Se sostuvo el chip a 45° de inclinación con el puerto de cargado hacia abajo y se retiró el líquido que pudiese contener.
2. Se colocó el chip en la mesa de trabajo, y con una pipeta de 200  $\mu\text{L}$  a 90° (perpendicular a la mesa) se cargaron los 30  $\mu\text{L}$  totales previamente preparados.
3. Se centrifugó el chip con un spin de 30 segundos con el puerto de cargado hacia el centro.
4. Nuevamente se sostuvo el chip a 45° y se mezcló la librería subiendo y bajando 3 veces tratando de no generar burbujas, se centrifugó con un spin de 30 s.
5. Se repitió el paso anterior una vez más.
6. Por última vez se sostuvo el chip a 45° y se retiró la mayor cantidad posible de líquido contenido.
7. Finalmente se colocó el chip en el equipo Ion Torrent PGM y comenzó la corrida de secuenciación.



**Figura 16. Cargado del Ion Chip 316 v2.** Procedimiento sin guantes para evitar daños a causa de cargas electrostáticas.

#### **5.3.6.4 Análisis bioinformático**

Concluida la corrida de secuenciación, se extrajeron las lecturas (secuencias) obtenidas y se analizaron con herramientas bioinformáticas especializadas, a continuación se menciona brevemente el proceso bioinformático desarrollado:

1. Evaluación de la calidad de las lecturas crudas con el software FastQC (software libre, plataforma Linux).
2. Recorte de adaptadores, eliminando las primeras 50 bases de cada lectura y supresión de lecturas de baja calidad, recortando cada lectura a partir de la posición que comience a caer por debajo del valor de Q20 CLCBio (Requiere Licencia, Multiplataforma).
3. Segunda evaluación de las lecturas y visualización del efecto del recorte en la calidad de las mismas con el software FastQC.
4. Cargado de los datos limpios en OneCodex (Plataforma en línea, Acceso limitado libre) y visualización de resultados preliminares.
5. Ensamblés *De Novo* con el software CLCBio, estableciendo como parámetros una identidad y cobertura de las lecturas de al menos el 90 % para dejar cabida a la captura de posibles polimorfismos.
6. Extracción de los contigs más largos (secuencias consenso formadas por un conjunto de lecturas) y alineamiento en BLASTn (alineamiento de los nucleótidos en las lecturas con la base de datos de nucleótidos de NCBI) utilizado los parámetros predeterminados para la verificación del origen e identidad de cada lectura.
7. Mapeos de las lecturas a las referencias descargadas y/o a las identidades obtenidas mediante el alineamiento en BLAST con el software CLCBio, estableciendo como parámetros de cobertura e identidad, por lo menos el 90 %.
8. Búsqueda de referencias redundantes con los softwares CSA (programa libre, plataforma Linux) y CLCBio.
9. Reporte de resultados depurados.

# *Capítulo 6*

## **RESULTADOS**

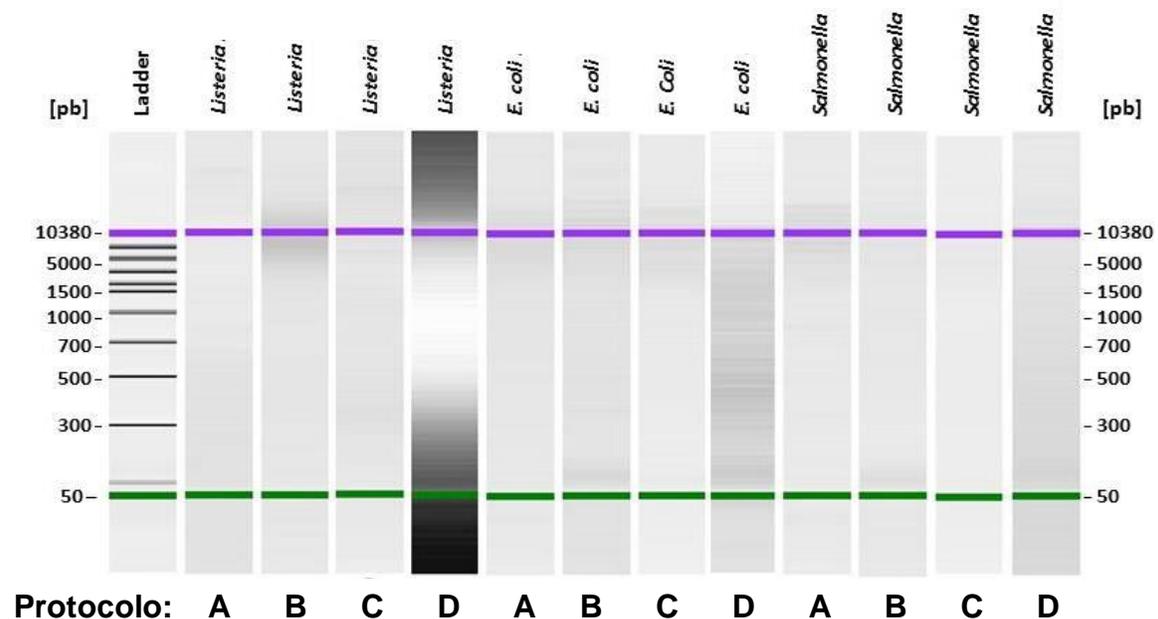
## 6.0 RESULTADOS

### 6.1 Selección de muestras

Como se describió en el punto 5.3.1 las muestras para el desarrollo del presente trabajo, fueron tres cepas aisladas e identificadas mediante pruebas bioquímicas por el LDDOP, correspondientes a *L. monocytogenes*, *E. coli* y *S. Enteritidis*.

### 6.2 Extracción de plásmidos

Se extrajo el material genético plasmídico de cada muestra mediante las cuatro metodologías ya descritas e identificadas como: **(A)** Zyppy™ Plasmid Miniprep Kit, **(B)** ZR Plasmid Miniprep™-Classic, **(C)** GeneJET Plasmid Miniprep Kit y **(D)** metodología ajustada de Aya Brown, Itai, & Itzhak, (2013) [1]. Finalizadas las extracciones, se observaron en el equipo Bioanalizador 2100 empleando un Chip 7500 (Véase Anexo 1), para visualizar la integridad y contenido del ADN extraído (Figura 17).



**Figura 17.** Extracción del material genético plasmídico con las cuatro metodologías utilizadas. Gel digital obtenido mediante electroforesis capilar. Se observan pequeñas cantidades de material genético a lo largo de todos los carriles, con una mayor cantidad en las muestras extraídas bajo el protocolo D.

### 6.3 Cuantificación y evaluación de la calidad del ADN extraído

Se evaluarán los extractos obtenidos mediante fluorimetría y espectrofotometría, utilizando los equipos *Qubit* y *NanoDrop* respectivamente. Con el equipo *Qubit* se cuantificó la cantidad de ADN de doble cadena presente en cada extracto, en tanto que la calidad de los mismos se evaluó mediante la relación de absorbancias a 260 y 280 nm con el equipo *NanoDrop* Tabla 14.

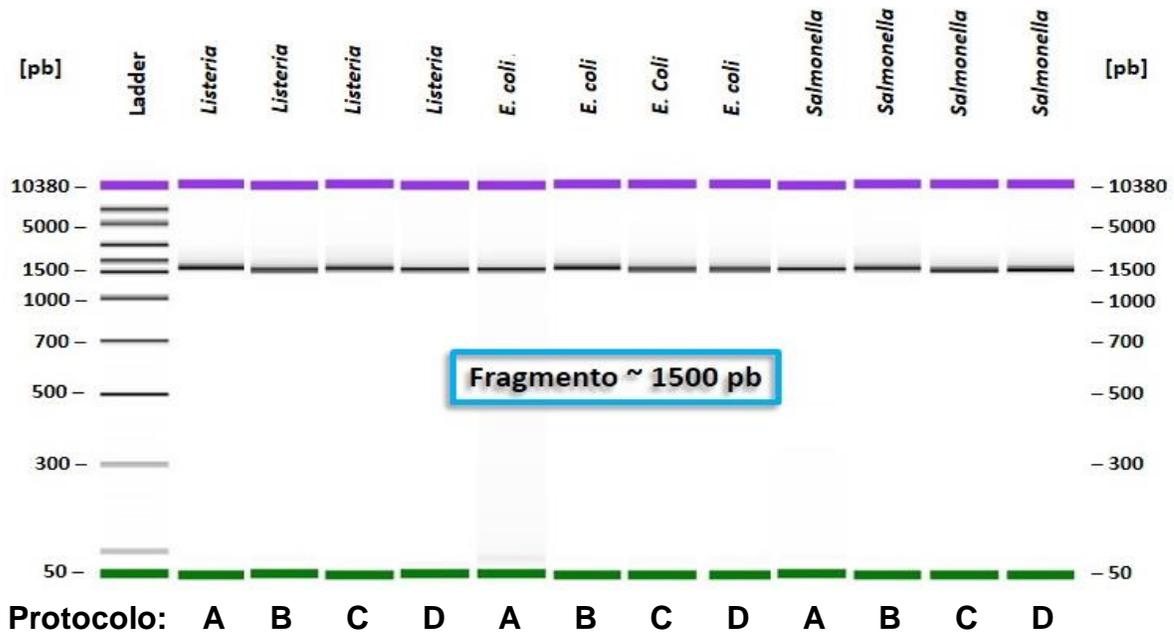
**Tabla 14. Concentración y calidad de los extractos de ADN plasmídico para cada protocolo.** En negritas se muestran los valores de calidad fuera del intervalo de aceptación para el CNRDOGM (1.8-2.19). Cada muestra se trabajó por duplicado.

Microorganismo	NanoDrop		Qubit BR		
	A260/A280		Concentración de Ácido Nucleico (ng/μl)		
	R1	R2	R1	R2	Media
<i>L. monocytogenes (A)</i>	2,10	<b>1,58</b>	12,50	1,56	7,03
<i>E. coli (A)</i>	<b>2,24</b>	<b>1,63</b>	4,64	3,30	3,97
<i>S. Enteritidis (A)</i>	<b>1,66</b>	<b>1,65</b>	7,38	4,10	5,74
<i>L. monocytogenes (B)</i>	1,95	2,00	8,10	16,50	12,30
<i>E. coli (B)</i>	2,07	2,19	8,54	4,86	6,70
<i>S. Enteritidis (B)</i>	1,90	2,19	4,06	6,34	5,20
<i>L. monocytogenes (C)</i>	2,00	1,97	5,40	2,94	4,17
<i>E. coli (C)</i>	1,90	2,01	13,10	8,80	10,95
<i>S. Enteritidis (C)</i>	1,99	2,08	2,24	1,23	1,74
<i>L. monocytogenes (D)</i>	<b>1,59</b>	<b>1,56</b>	-	5,82	5,82
<i>E. coli (D)</i>	2,02	2,04	6,68	14,80	10,74
<i>S. Enteritidis (D)</i>	<b>1,66</b>	1,97	4,64	21,60	13,12

## 6.4 Purificación del ADN plasmídico.

### 6.4.1 Búsqueda de contaminantes cromosomales

Una vez extraído el material genético, se evaluó la presencia de residuos cromosomales mediante PCR dirigida a la amplificación del gen ribosomal 16S. Para realizar el ensayo se utilizaron iniciadores universales dirigidos al gen ribosomal 16S. Obteniendo los resultados mostrados en la Figura 18.

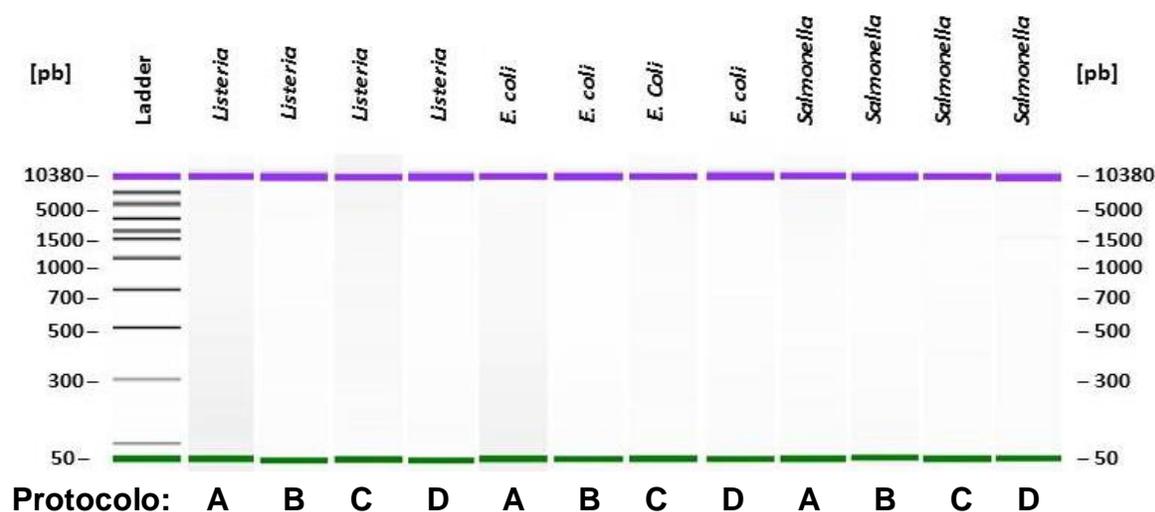


**Figura 18. Evaluación de contaminantes cromosomales en los extractos de cada metodología.** Se observa la banda correspondiente a la amplificación del gen ribosomal 16S, indicativo de la presencia del cromosoma en todas las muestras.

### 6.4.2 Remoción de ADN genómico.

Identificada la presencia de residuos cromosomales, se realizó una digestión enzimática con la enzima *Plasmid-safe DNasa*, la cual, por su actividad de exonucleasa degrada fragmentos con extremos libres, dejando en el medio sólo los elementos circulares, de este modo se pretende conservar los plásmidos purificados.

Una vez realizada la digestión, se evaluó nuevamente la presencia de residuos de material genético cromosomal por amplificación del gen ribosomal 16S. En la Figura 19 se observa la ausencia de productos de amplificación, descartando así la presencia de contaminantes cromosomales.



**Figura 19.** Ensayo de amplificación para el gen ribosomal 16S posterior a la digestión enzimática. Se observa la ausencia de bandas de amplificación provenientes de elementos cromosomales.

### 6.4.3 Búsqueda de ADN plasmídico

Purificados los extractos, se comprobó que aún se contaba con ADN plasmídico realizando una PCR dirigida a regiones únicas en los plásmidos, utilizando iniciadores específicos para cada especie y/o serotipo, capaces de discriminar entre los elementos de interés y el ADN cromosomal. Para el diseño de cada iniciador, se realizó una búsqueda en NCBI descargando los registros para cada bacteria en estudio, obteniendo 19 registros para *L. monocytogenes*, 11 para *S. Enteritidis* y se encontraron un total de 816 para *E. coli*, lo cual, al tratarse de una gran cantidad de secuencias, representó una complicación en el análisis de cada referencia. Para solventar dicho inconveniente, se realizó la búsqueda de un elemento único que estuviera presente en plásmidos de *E. coli*, se ubicó el Clúster *mccX*, cuyos genes codifican para diferentes tipos de microcinas o bacteriocinas que *E. coli* emplea para su predominancia en el medio <sup>[39]</sup> (Tabla 15).

A partir de las secuencias encontradas y haciendo uso del software *Primer3 Plus*, se hizo el diseño de los iniciadores, adicionalmente se utilizó BLAST, para asegurar que los cebadores diseñados contaran con dos características fundamentales: capacidad de diferenciar elementos plasmídicos de cromosomales e hibridar sólo con elementos génicos propios para los cuales fueron diseñados. (Anexo 2). En la Tabla 16 se detallan las características de cada iniciador diseñado, así como el tamaño del amplicón esperado.

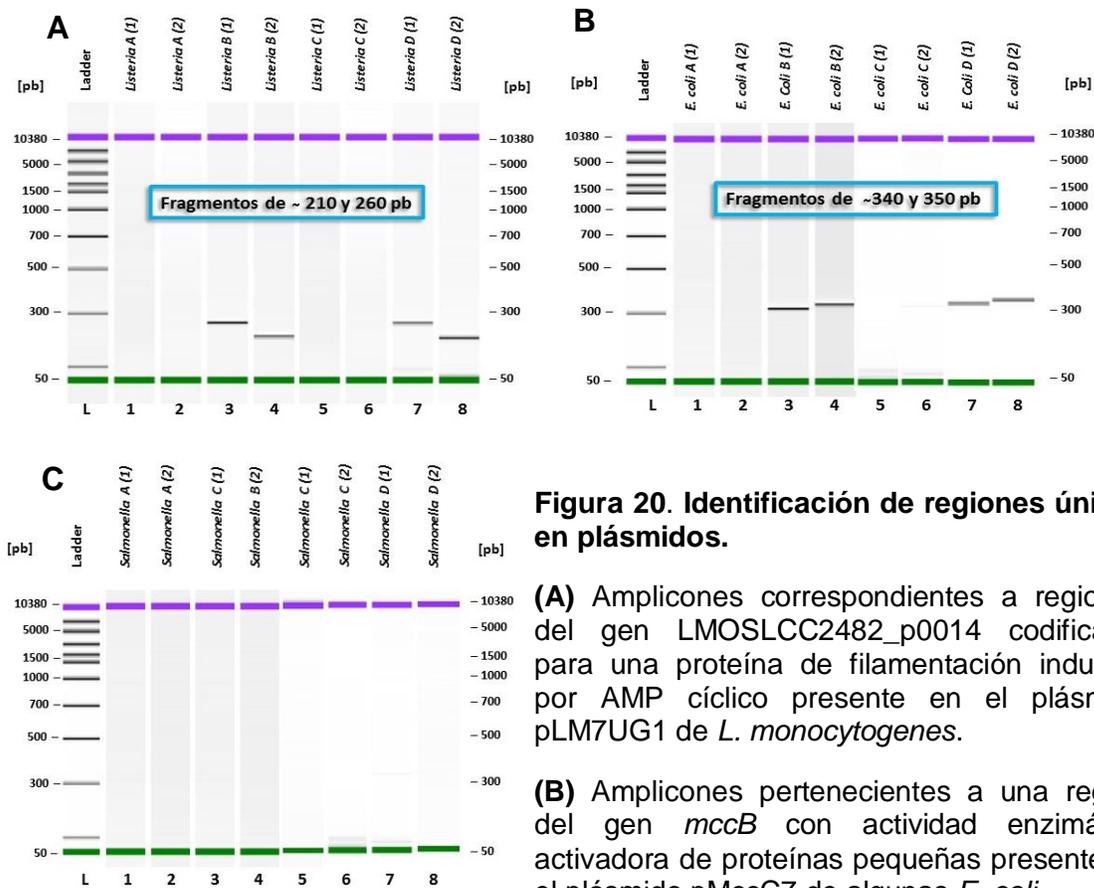
**Tabla 15. Secuencias de referencia plasmídica descargadas para cada bacteria en estudio.**

Bacteria	ID	Nombre del Plásmido	No. Acceso	Tamaño (kb)	% GC
<i>L. monocytogenes</i>	1	pLM33	NC_014255.1	32.307	36.15
	2	CFSAN023459_01	NZ_CP014253.1	12.949	36.36
	3	pLM7UG1	NC_018888.1	50.1	35.48
	4	pLM1-2cUG1	NC_018889.1	50.1	35.48
	5	CFSAN023459_02	NZ_CP014254.1	52.687	35.05
	6	pCFSAN010068_01	NZ_CP014251.1	55.521	36.7
	7	Unnamed	NC_022046.1	55.804	34.92
	8	Unnamed	NC_022047.1	55.804	34.92
	9	Unnamed	NC_022051.1	55.803	34.92
	10	Unnamed	NC_021828.1	57.557	36.03
	11	pLM1-2bUG1	NC_014495.1	57.78	36.04
	12	LM6179	NZ_HG813250.1	62.206	36.53
	13	pCFSAN008100	NZ_CP011399.1	68.224	35.67
	14	pLM5578	NC_013767.1	77.054	36.59
	15	pLMIV	NZ_CM001470.1	77.825	34.02
	16	pLmN1546	NZ_CP013725.1	86.616	36.98
	17	II	NZ_HG813248.1	86.652	36.98
	18	Unnamed	NC_022045.1	148.96	37.05
	19	pl2015TE24968	NZ_CP015985.1	57.530	36.04
<i>E. coli</i>	1	MccC7-H22	NC_010257.1	32.014	43.50
	2	B	NZ_CP010127.1	42.655	39.89
	3	unnamed_2	NZ_CP009074.1	24.185	39.22
	4	mccABCDE	AJ487788.2	5.734	34.32
	5	pRPEC180_47	JN935898.1	47.506	41.43
	6	pMccC7	X57583.1	6.458	34.42
	7	pMccC7	AY913945.1	6.360	34.66
<i>S. Enteritidis</i>	1	pCFSAN000111_01	NZ_CP007599.1	39.599	51.96
	2	pCFSAN000006	NZ_CP011395.1	59.336	51.96
	3	pSLA5	NC_019002.1	59.374	48.34
	4	Unnamed	NZ_CP007508.1	59.372	51.96
	5	plasmid00	NZ_CP007529.1	59.369	51.92
	6	Unnamed	NZ_CP008927.1	59.372	39.88
	7	pCMCC50041	NZ_CP013098.1	59.373	51.96
	8	pFORC7	NZ_CP009767.1	101.43	51.95
	9	Virulence	NZ_CP011943.1	59.37	51.96
	10	Intermediate	NZ_CP011944.1	3.372	55.10
	11	Small	NZ_CP011945.1	2.096	55.63

Tabla 16. Iniciadores diseñados para cada plásmido según su especie y/o serotipo.

Número de Acceso	Microorganismo	Oligonucleótido	Secuencia	Tamaño de fragmento esperado	T <sub>m</sub> (°C)
NC_01888 8.1	<i>L. monocytogenes</i>	pLM7UG1_F1	5'-TCG CTC GCC GAT ACG GTG TCT ATT-3'	266 pb	61.6
		pLM7UG1_R1	5'-GTG GAC TCG CCG TCT GGA AAT CTG-3'		
		pLM7UG1_F2	5'-CAG ATT TCC AGA CCG CGA GTC CAC-3'	211 pb	61.6
		pLM7UG1_R2	5'-TCA CCA TAC GTC CTG TTC GTC CGT-3'		
NZ_CP009 074.1	<i>E. coli</i>	mccB_F1	5'-GAT GAT CTG CAG AAA CAA CCC AT-3'	343 pb	55.2
		mccB_R1	5'- ATG GTG CTA ATC CTG TTC TTG TG-3'		
		mccB_F2	5'-CTA TTT CCA CAC ACA GAG CAA AC -3'	355 pb	54.3
		mccB_R2	5'-TCA GCC CTA TATA AA TGC TGG GT-3'		
NZ_CP007 599.1	<i>S. Enteritidis</i>	S_pCFSAN_F1	5'-AGA GTG GGA AAC ATT TCA GCG CAG GAC-3'	390 pb	62.3
		S_pCFSAN_R1	5'-CCC AGT CCT GGG AGT GTG AAA CAA AAC T-3'		
		S_pCFSAN_F2	5'-TCG GCG GTA GCG TAC TTG AAT GCT CT-3'	333 pb	61.9
		S_pCFSAN_R2	5'-ACT CGC AAC TCA GGA ACA ACA TAA GCG-3'		

Con cada par de cebadores se llevaron a cabo ensayos de PCR. En la Figura 20 se aprecia que las muestras de *L. monocytogenes* y *E. coli* obtuvieron la amplificación deseada comprobando la presencia de material genético plasmídico. En el caso de *S. Enteritidis* no fue posible identificar la presencia de material genético, sin embargo, se siguió trabajando con la muestra.



**Figura 20. Identificación de regiones únicas en plásmidos.**

**(A)** Amplicones correspondientes a regiones del gen LMO5LCC2482\_p0014 codificante para una proteína de filamentación inducida por AMP cíclico presente en el plásmido pLM7UG1 de *L. monocytogenes*.

**(B)** Amplicones pertenecientes a una región del gen *mccB* con actividad enzimática activadora de proteínas pequeñas presente en el plásmido pMccC7 de algunas *E. coli*.

**(C)** No se obtuvo amplificación para los plásmidos de *S. Enteritidis*. El ensayo estuvo dirigido a amplificar una región del gen SEEE1427\_22585 codificante para una proteína de transferencia por conjugación presente en el plásmido pCFSAN000111\_01 de *S. Enteritidis*.

### 6.5 Amplificación de los plásmidos purificados

Una vez purificados los plásmidos, se amplificaron mediante una PCR por desplazamiento de hebra, de acuerdo al protocolo del *Kit Illustra DNA GenomiPhi Amplification de GE Healthcare*, con lo cual se generó un aumento en la cantidad de material genético presente en los tubos de reacción; culminada esta última amplificación, se cuantificó mediante fluorometría. En la Tabla 17 se describe la cantidad inicial de ADN en cada muestra y la cantidad resultante posterior a la amplificación.

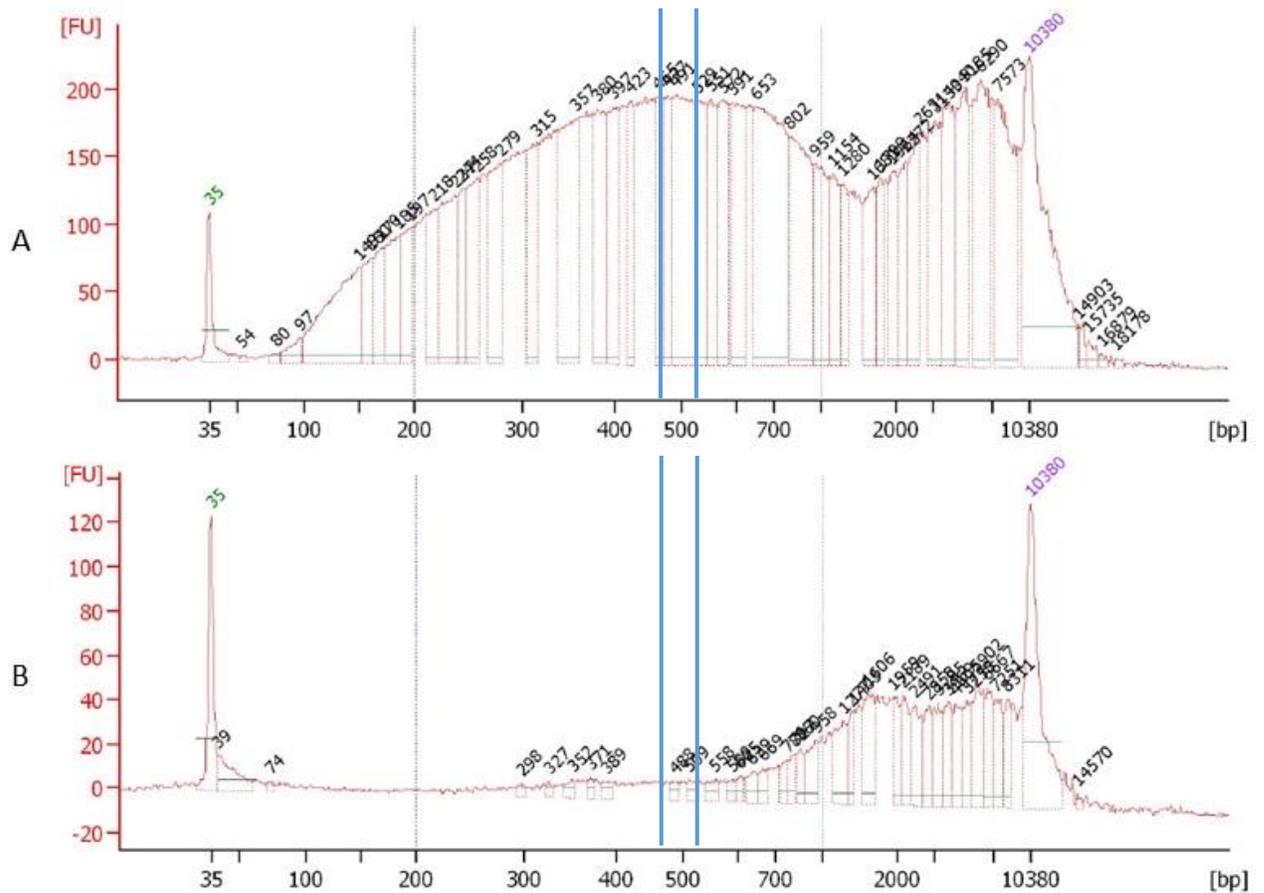
**Tabla 17. Cuantificación del material genético obtenido mediante la amplificación con hexámeros al azar.**

Muestra	Qubit High Sensitivity	
	Purificada sin amplificar [ng/μl]	Purificada amplificada [ng/μl]
<i>L. monocytogenes</i>	Too low**	0.168
<i>E. coli</i>	0.498	2.46
<i>S. Enteritidis</i>	Too low**	0.186

\*\* Concentración muy baja de ADN en doble cadena, bajo esta condición, el método fluorométrico no logra la cuantificación

### 6.6 Generación de bibliotecas genómicas

Al contar con la cantidad suficiente de material genético, se comenzó el proceso de secuenciación. El primer paso fue la elaboración de bibliotecas genómicas, para la cual fue necesaria la fragmentación. Para esto, se probaron condiciones para realizarla vía enzimática y mecánica, con el fin de obtener la mayor cantidad de fragmentos alrededor de las 400 pb. La fragmentación fue visualizada y analizada mediante un Chip High Sensitivity en el Bioanalizador 2100. En la Figura 21 se observa que mediante la metodología de fragmentación ultrasónica se logró obtener una distribución más uniforme, así como una mayor cantidad de ADN de la longitud deseada.

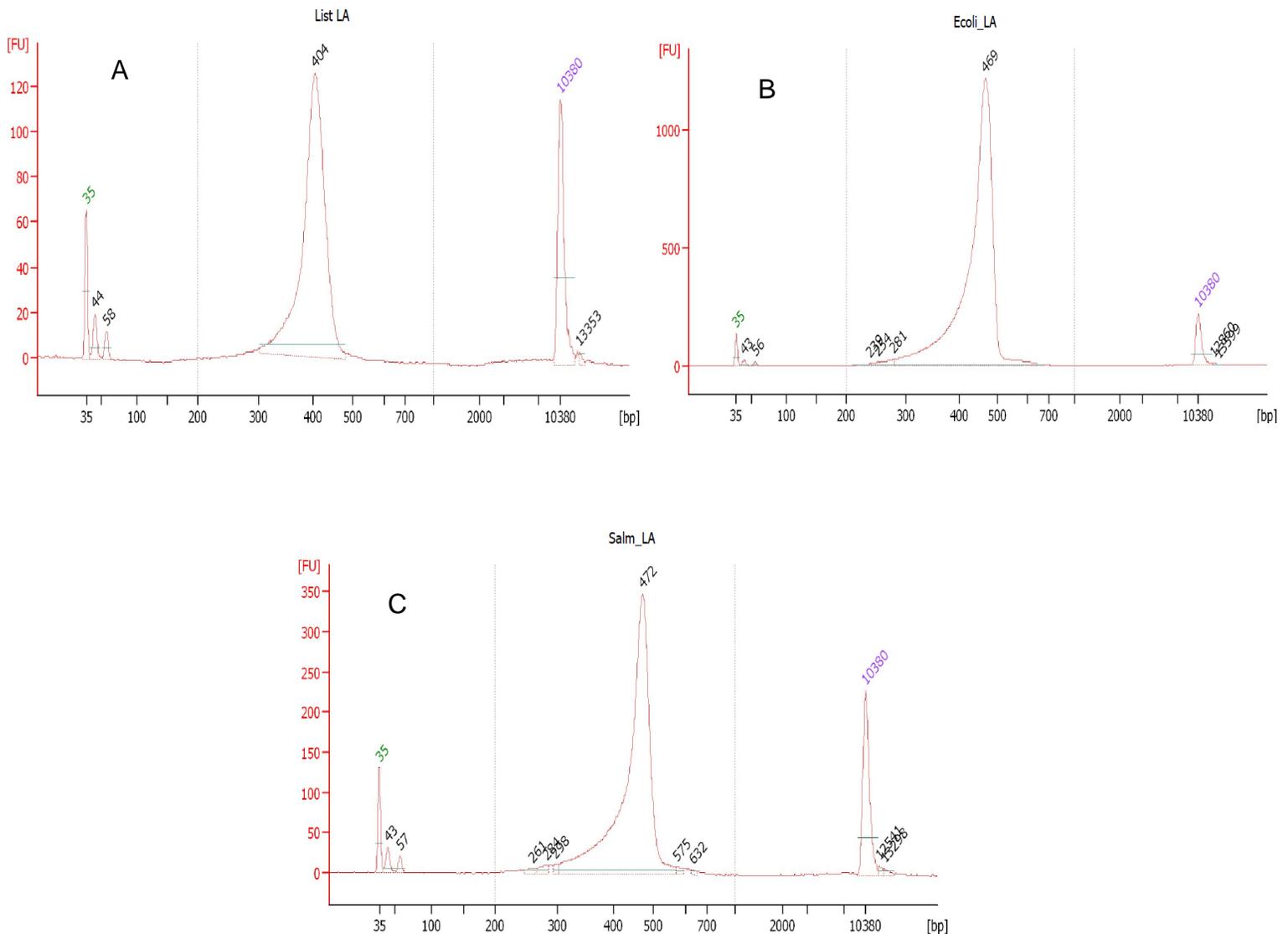


**Figura 21. Fragmentación del material genético para la elaboración de librerías. (A)** Fragmentación por sonicación, **(B)** Fragmentación enzimática. Las líneas en azul marcan el intervalo con los tamaños deseados

Los fragmentos obtenidos por sonicación fueron reparados para obtener extremos romos y posteriormente proceder a la ligación de adaptadores. Finalmente se llevó a cabo una selección de fragmentos por electroforesis en agarosa al 2%, para recuperar únicamente secuencias de aproximadamente 480 pb. (400 pb de secuencia problema + 2 adaptadores de 40 pb cada uno).

### 6.6.1 Amplificación de las Bibliotecas Genómicas

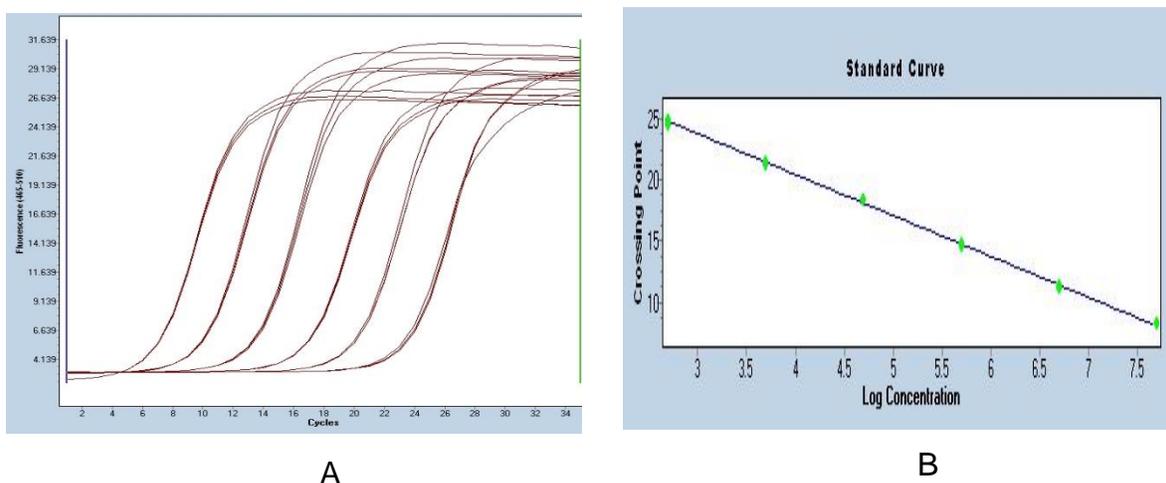
Una vez seleccionados los fragmentos del tamaño deseado, se realizó su amplificación, lo anterior para incrementar la cantidad de librería de 480 pb. Los productos de amplificación se verificaron en un Chip High Sensitivity. En la Figura 22 se puede apreciar un pico cercano a las 480pb, para cada una de las muestras.



**Figura 22. Bibliotecas amplificadas.** Amplificación de los fragmentos seleccionados para incrementar la cantidad de ADN a secuenciar. (A) Biblioteca genómica de *L. monocytogenes*, (B) Biblioteca genómica de *E. coli* y (C) Biblioteca genómica de *S. Enteritidis*.

### 6.6.2 Cuantificación de las bibliotecas mediante PCR tiempo real.

La cantidad de librería obtenida previamente se cuantificó mediante qPCR con el *KAPA Library Quantification Kit for Ion Torrent™*. Para esto, cada muestra se evaluó por triplicado y mediante una curva de calibración con seis estándares de concentración conocida, fue posible conocer la concentración de cada una de las bibliotecas genómicas. En la Figura 23 se muestran las diluciones seriadas que se realizaron para la curva de calibración, cuya linealidad obtuvo los siguientes valores:  $m = -3.33$ ,  $b = 33.8$  y  $r^2 = 0.94$ .



**Figura 23. Cuantificación de librerías mediante qPCR.** (A) Curva estándar para la cuantificación de librerías mediante qPCR. (B) Gráfica de regresión lineal de la curva de calibración.

Los resultados de la cuantificación se resumen en la Tabla 18. Las librerías cuantificadas se diluyeron hasta obtener una mezcla equimolar de todas las librerías a una concentración de  $1.5 \times 10^7$  moléculas/ $\mu\text{L}$ . Esta normalización se realizó para poder proceder a la amplificación por emulsión (em-PCR) y asegurar que la cantidad de moléculas por muestra fuera aproximadamente igual, para evitar la sobre-representación de alguna de ellas, lo cual mermaría la cantidad de información a obtener del resto de las muestras.

Tabla 18. Cuantificación de librerías por qPCR.

Librería	Concentración obtenida por qPCR (moléculas/ $\mu$ L)	Factor de dilución empleado (1:X)	Concentración obtenida
<i>L. monocytogenes</i>	$1.56 \times 10^{09}$	1:100.62	1.5x10 <sup>7</sup> moléculas/ $\mu$ L
<i>E. coli</i>	$1.11 \times 10^{10}$	1:716.20	
<i>S. enterica</i> serovar Enteritidis	$1.73 \times 10^{07}$	1:1.11	

### 6.7 Amplificación clonal por em-PCR y enriquecimiento

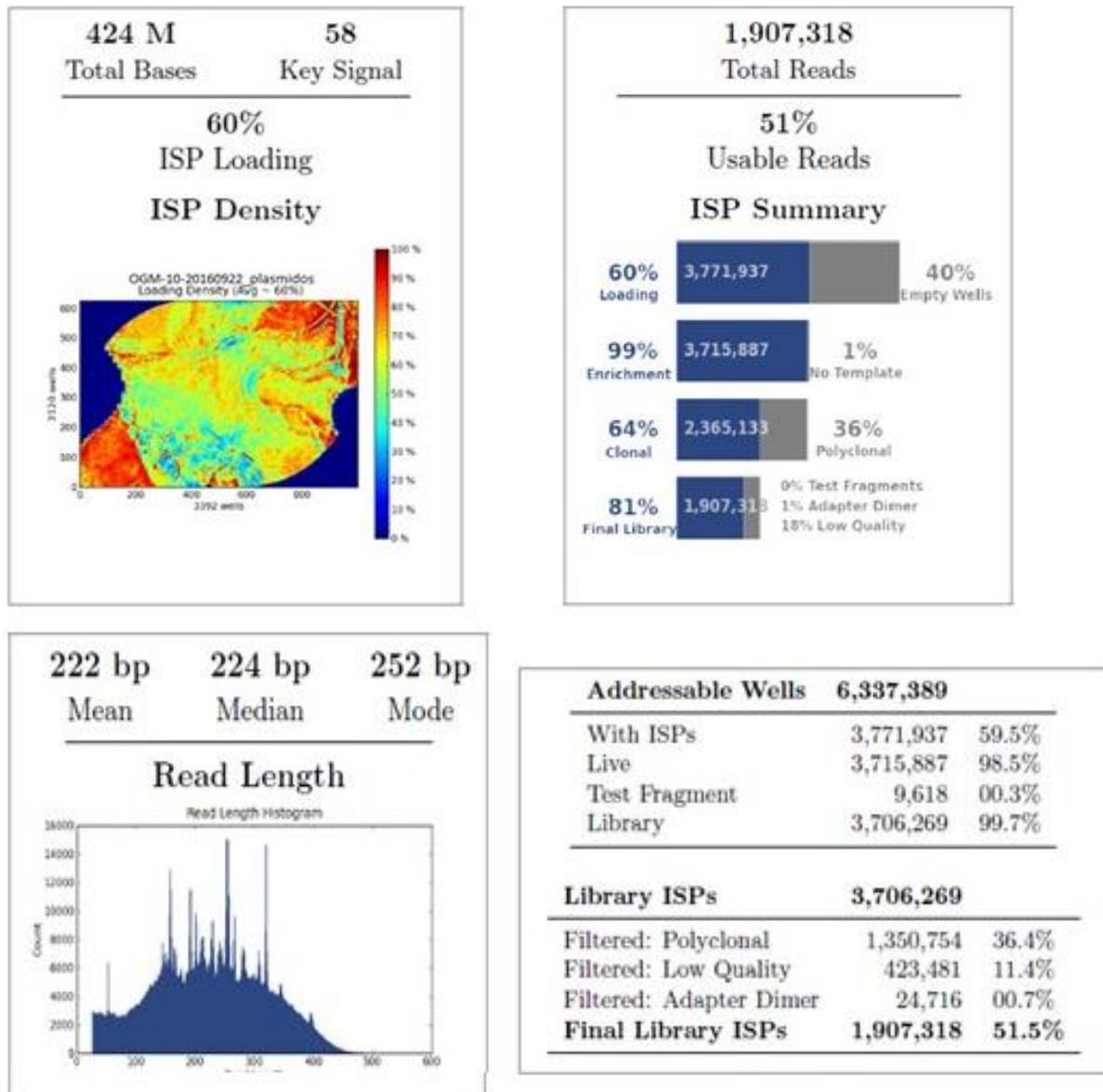
Se realizó la amplificación de los fragmentos ligados a las perlas de secuenciación y el enriquecimiento de las mismas. Previo al enriquecimiento de las perlas amplificadas, se realizó un ensayo control para evaluar el porcentaje de fragmentos amplificados y unidos a las ISP's (% de recuperación). Se recomienda un valor en el intervalo del 10-40% de perlas amplificadas (véase Anexo 3). Para este ensayo se obtuvo un porcentaje de recuperación del 33.90 %.

Posterior a esta evaluación, se procedió a realizar el enriquecimiento (véase sección 5.3.6.2.1 "Enriquecimiento"), a modo de recuperar la mayor cantidad de perlas cargadas con fragmentos de ADN y desechar aquellas perlas vacías a las cuales no se les acopló ningún fragmento.

### 6.8 Secuenciación

La corrida de secuenciación se llevó a cabo cargando un Ion 316v2™ Chip (Véase Anexo 4), siguiendo los pasos descritos en el punto 5.3.6.3 "Corrida de secuenciación".

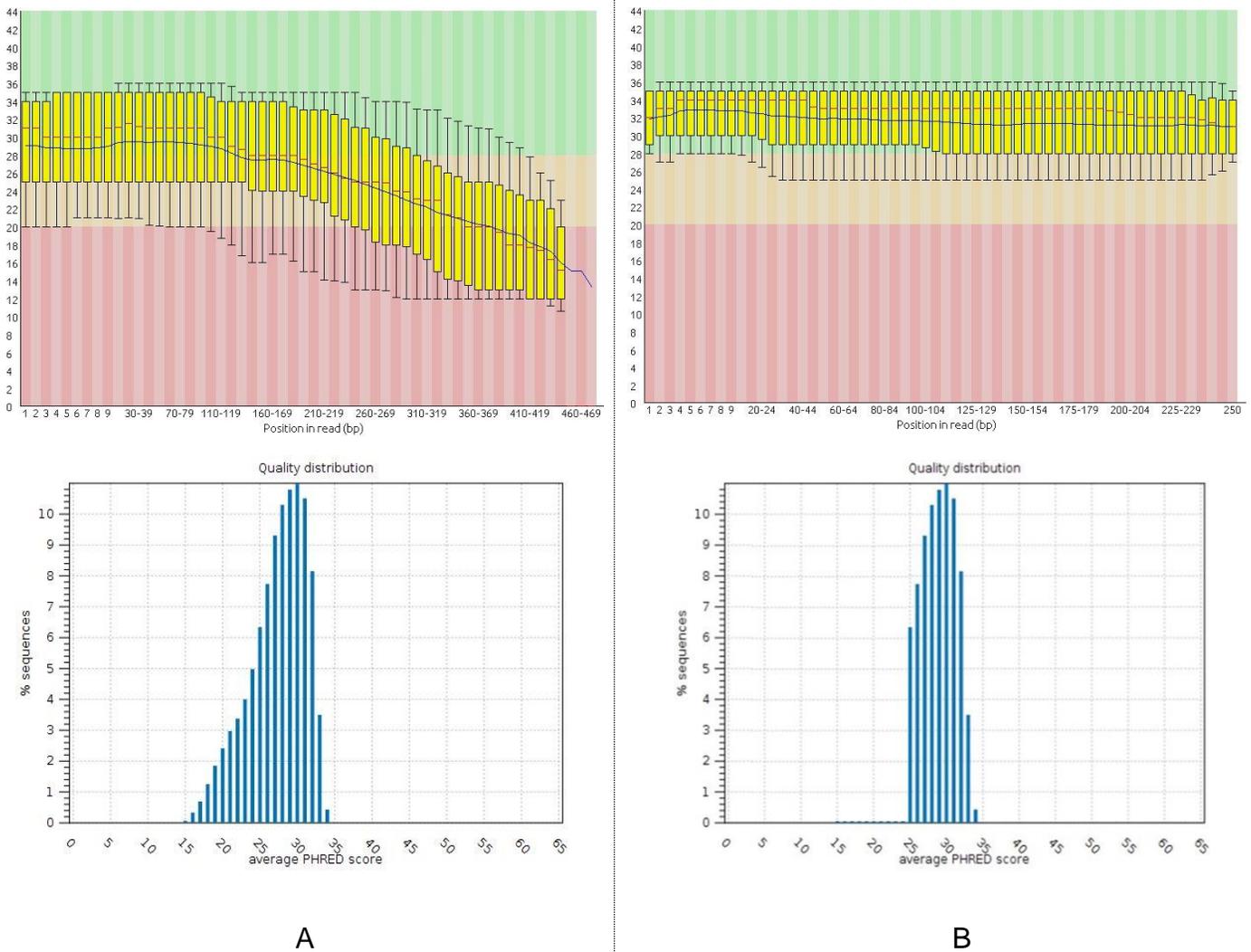
Finalizado el ensayo de secuenciación, se extrajo directamente del servidor *Torrent Server* el reporte de resultados (Figura 24)., donde se observa un total de 424 Mb secuenciadas con un 60% de cargado del chip (3, 771, 937 pozos cargados de 6, 337, 389 disponibles), del cual se registraron 1, 907, 318 lecturas útiles con un tamaño promedio de 222 pb.



**Figura 24. Reporte de la corrida de secuenciación para plásmidos con la plataforma Ion Torrent PGM. Se detalla un cargado del 60% y un resumen del filtrado de las lecturas, obteniendo un total de 1 907 318 lecturas útiles con una longitud promedio de 222 pb.**

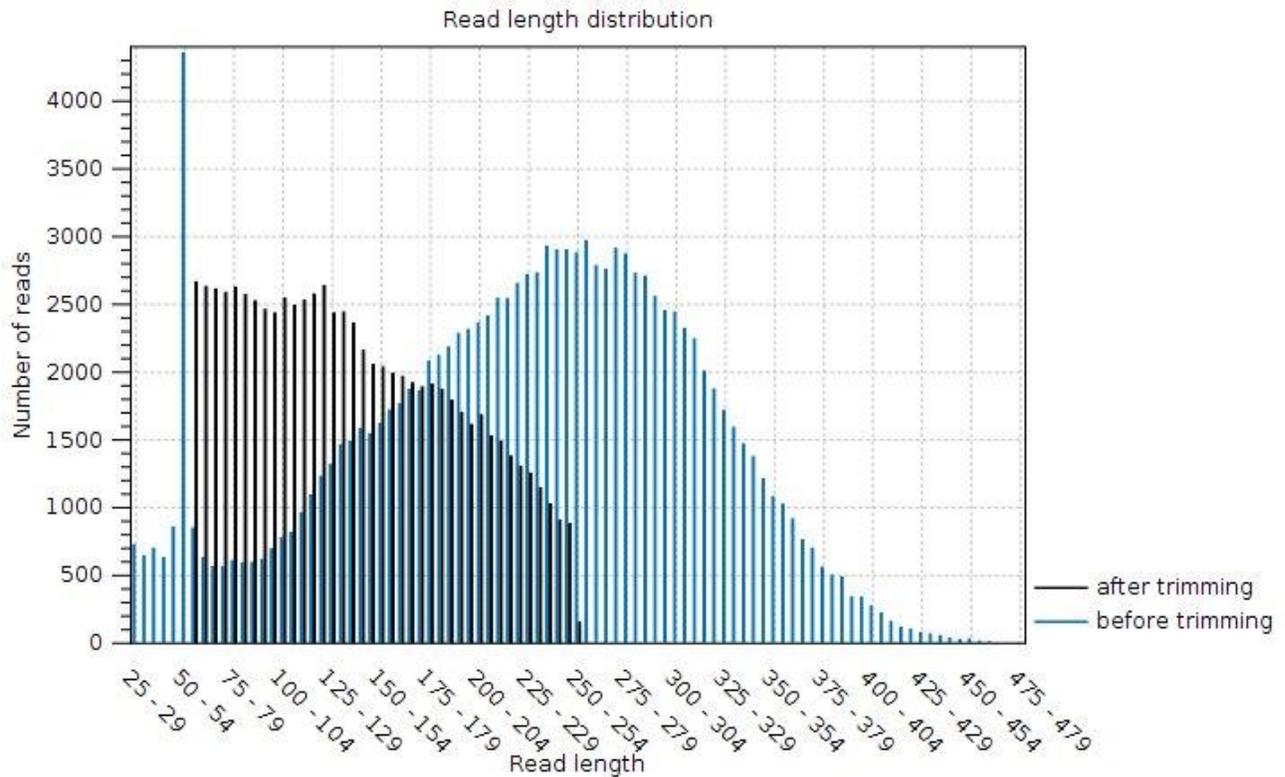
**6.9 Análisis bioinformático**

Los datos crudos obtenidos del secuenciador fueron analizados mediante CLCBio, en donde se realizó el recorte de adaptadores (secuencias ajenas al material genético de interés ubicados al inicio y final de cada lectura) y la eliminación de aquellas lecturas con un valor de calidad en la escala Phred por debajo de Q20 (1 error por cada 100 bases secuenciadas). En la Figura 25 se observa el comparativo entre las lecturas y sus calidades antes y después de este proceso. Se aprecia que previo al tratamiento, existen lecturas que se extienden hasta valores por debajo de Q20, mientras que posterior a éste, ya se observa un comportamiento regular con valores que oscilan cerca de Q30 (1 error por cada 1000 bases secuenciadas).



**Figura 25. Efecto del recorte de lecturas con baja calidad.** (A) Previo al recorte del inicio de las secuencias, así como la presencia de adaptadores y lecturas con baja calidad, se muestran lecturas con valores por debajo de Q20, posteriormente (B) se observa una distribución uniforme de las calidades de las lecturas con valores por encima de Q25.

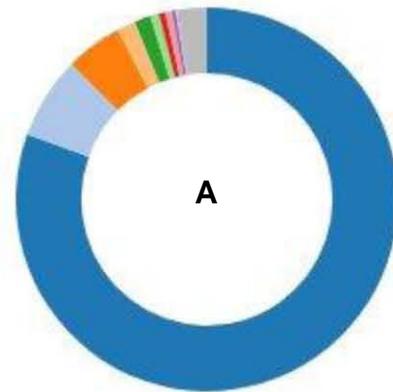
Por otra parte, la distribución de los tamaños en las lecturas, posterior al recorte de adaptadores, cambia con respecto a las obtenidas de manera cruda del secuenciador. En la Figura 26, obtenida del software CLC, se muestra un comparativo del antes y el después del tratamiento de limpieza de lecturas y el impacto que este tiene sobre el tamaño de las mismas.



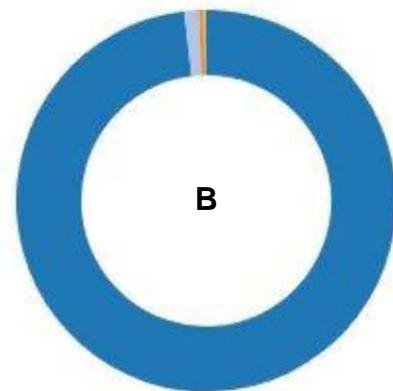
**Figura 26. Comparativo entre las longitudes de las lecturas antes y después del recorte de adaptadores.** En azul se observa la distribución de tamaños obtenidas directamente del secuenciador en un intervalo de 25–479 pb. Una vez analizada la calidad de cada base a lo largo de las lecturas obtenidas, se realizó el recorte de los adaptadores y de las secuencias cuyas bases secuenciadas contarán con una mala calidad, obteniendo una longitud final de las lecturas en un intervalo de 50–250 pb, con las cuales se realizaron los análisis bioinformáticos.

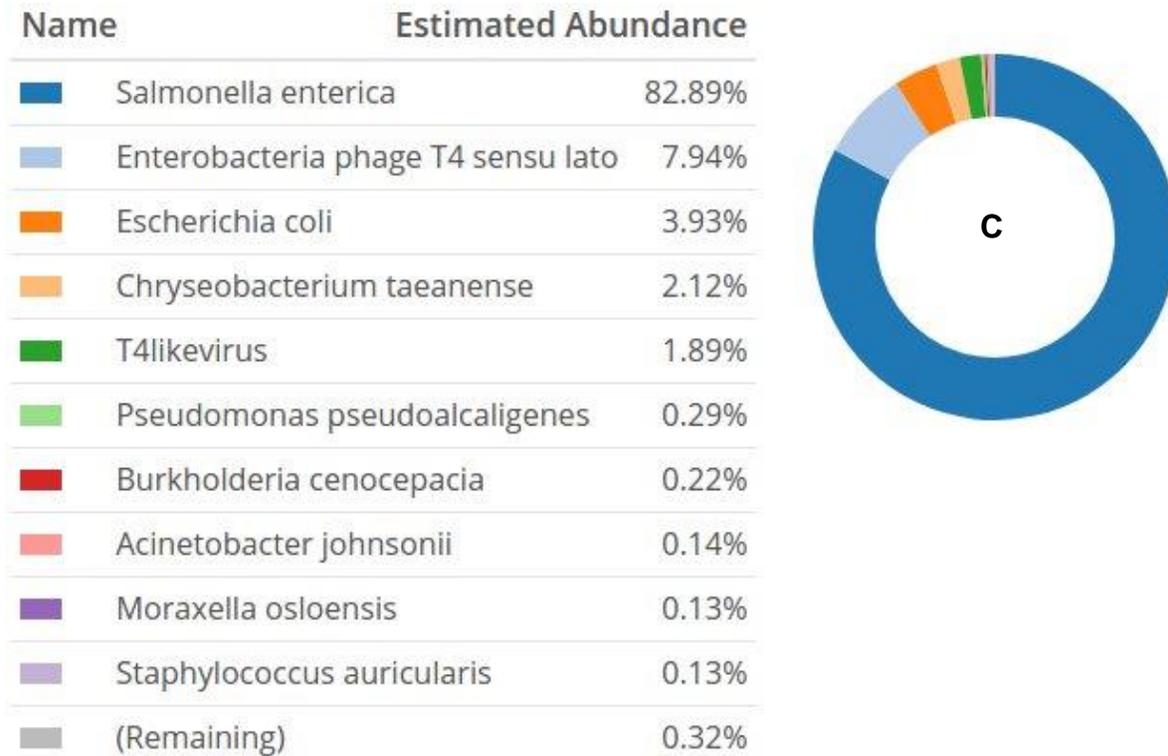
El set de lecturas limpias se visualizó en One Codex, con el objeto de caracterizar y obtener resultados presuntivos sobre el origen de las lecturas. Como se puede apreciar en la Figura 27, se obtuvo un 80.50% de las lecturas asignadas para *L. monocytogenes*, 98.15% asignadas a *E. coli* y 82.89 % asignadas a *S. Enteritidis* en sus respectivos análisis individuales por muestra, con lo cual se verificó la distribución de las muestras trabajadas.

Name	Estimated Abundance
 Listeria monocytogenes	80.50%
 Alphapapillomavirus 7	6.93%
 Propionibacterium acnes	4.80%
 Aerococcus viridans	1.58%
 Citrobacter braakii	1.38%
 Aerococcus urinaeequi	0.72%
 Staphylococcus capitis	0.56%
 Escherichia coli	0.55%
 Acinetobacter johnsonii	0.32%
 Acinetobacter lwoffii	0.17%
 (Remaining)	2.46%



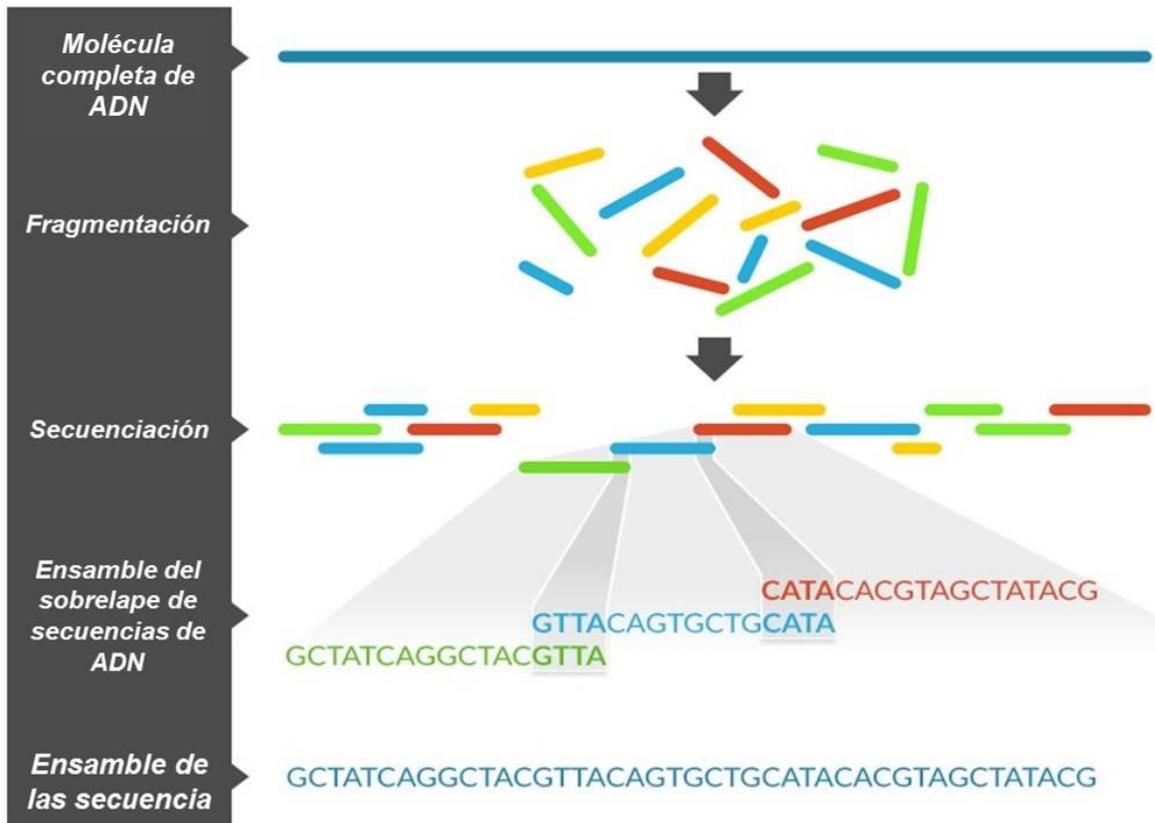
Name	Estimated Abundance
 Escherichia coli	98.15%
 Burkholderia cenocepacia	1.25%
 Burkholderia multivorans	0.36%
 Acinetobacter sp. CIP 101966	0.10%
 Herbaspirillum huttiense	0.06%
 Pseudomonas fluorescens	0.05%
 Pseudomonas sp. Ag1	0.03%





**Figura 27. Análisis en One Codex para la corrida de plásmidos.** Se observa la clasificación de las lecturas con base en el contenido nucleotídico de cada muestra, la clasificación obtenida para cada muestra llegó hasta nivel de especie. Se obtuvo un porcentaje del contenido secuenciado del 80.50 % para *L. monocytogenes* (22A), 98.15 % asignado para *E. coli* (22B) y 82.89 % asignado para *S. enterica* (22C).

Con las lecturas procesadas y teniendo un resultado presuntivo del organismo origen en One Codex, se realizaron ensambles *de novo* en CLCBio con las lecturas de cada muestra, llevando a cabo un sobrelape de los fragmentos altamente similares (Figura 28), los cuales tienen su origen en la misma región del genoma. De este modo se conectan los fragmentos individuales para convertirse en secuencias contiguas más largas obteniendo un consenso de mayor longitud de ADN. Estas secuencias consenso reciben el nombre de *contigs*.



**Figura 28. Ensamble *de novo*. Sobrelape de fragmentos con extremos similares para ensamblar una secuencia consenso.** Se muestra el alineamiento de los extremos de cada fragmento hasta conseguir una secuencia consenso de mayor tamaño.

Se extrajo el *contig* de mayor longitud, (ya que este representa una secuencia con mayor especificidad y fidelidad comparada con un *contig* de menor tamaño) de cada muestra para realizar un alineamiento en BLAST y verificar el origen de los mismos. Con los resultados obtenidos se determinó que los datos secuenciados tienen su origen en ADN plasmídico. En la Tabla 19 se muestran los resultados obtenidos de alinear en BLAST el *contig* de mayor longitud correspondiente a cada muestra.

**Tabla 19. Resultados del ensayo en BLAST para el *contig* más largo de cada muestra.**

Organismo	Contigs ensamblados	Contig más largo (pb)	Referencia plasmídica asociada	Profundidad Promedio	Query cover	Identidad
<i>L. monocytogenes</i>	617	27, 113	CP006596.1	183.15 X	100.0 %	99.0 %
<i>E. coli</i>	4	3, 040	KM085453.1	5, 311.93 X	99.0 %	99.0 %
S. Enteritidis	1 082	6, 146	CP020826.1	23.2 X	100.0 %	100.0 %

Una vez identificado el origen plasmídico de los datos, se realizó el mapeo del total de las lecturas contra los registros descargados previamente de la base de datos del NCBI (Tabla 15), así como las secuencias mostradas en la Tabla 19, obteniendo resultados con una cobertura al 100 % de las secuencias plasmídicas, mapeos incompletos y mapeos en donde sólo se capturó un porcentaje significativo de la secuencia nucleotídica para las diferentes referencias plasmídicas.

En la Figura 29 se observa el comparativo entre un mapeo con cobertura al 100 %, en contraste a uno con sólo una fracción del plásmido capturado en donde se incluyen diferencias entre el contenido nucleotídico de la referencia y el contenido secuenciado. Asimismo, en la Tabla 20 se observan a detalle los resultados obtenidos, destacando el tamaño en Kb, el % GC, la cobertura longitudinal y la profundidad obtenida para cada mapeo.



**Figura 29. Mapeo a referencia en el software CLC.** De manera horizontal, se observa la cobertura de la referencia mapeada y de modo vertical se observa la profundidad del mapeo, la cual, corresponde al número de veces que cada nucleótido fue secuenciado. Las líneas en verde representan las lecturas que se alinearon en dirección al sentido de la referencia (lecturas en forward) en tanto que las rojas representan el reverso complementario de las lecturas, necesario para realizar un correcto alineamiento (lecturas en reverse). (A) Plásmido encontrado al 100 % entre las lecturas recolectadas de la secuenciación. (B) Región con alto nivel de similitud a la referencia mapeada pero no cubierto al 100 %.

**Tabla 20. Resultados obtenidos de los mapeos a cada referencia. Se muestra la cobertura de manera decreciente y la profundidad que se obtuvo para cada registro en el software CLCBio.**

Bacteria	ID	Plásmido reportado	No. Acceso	Tamaño (kb)	% GC	Cobertura	Profundidad
<i>L. monocytogenes</i>	1	pLM33	NC_014255.1	32.307	36.15	100.00 %	16.04 X
	3	pLM7UG1	NC_018888.1	50.1	35.48	100.00 %	18.45 X
	4	pLM1-2cUG1	NC_018889.1	50.1	35.48	100.00 %	18.45 X
	10	unnamed	NC_021828.1	57.557	36.03	100.00 %	17.54 X
	11	pLM1-2bUG1	NC_014495.1	57.78	36.04	100.00 %	17.71 X
	13	pCFSAN008100	NZ_CP011399.1	68.224	35.67	68.00 %	12.56 X
	6	pCFSAN010068_01	NZ_CP014251.1	55.521	36.7	65.00 %	12.85 X
	14	pLM5578	NC_013767.1	77.054	36.59	52.00 %	9.61 X
	18	unnamed	NC_022045.1	148.959	37.05	44.00 %	6.84 X
	12	LM6179	NZ_HG813250.1	62.206	36.53	43.00 %	7.07 X
	7	unnamed	NC_022046.1	55.804	34.92	35.00 %	7.88 X
	8	unnamed	NC_022047.1	55.804	34.92	35.00 %	7.88 X
	9	unnamed	NC_022051.1	55.803	34.92	35.00 %	7.88 X
	16	pLmN1546	NZ_CP013725.1	86.616	36.98	33.00 %	6.35 X
	17	II	NZ_HG813248.1	86.652	36.98	33.00 %	6.35 X
	5	CFSAN023459_02	NZ_CP014254.1	52.687	35.05	7.00 %	0.56 X
	2	CFSAN023459_01	NZ_CP014253.1	12.949	36.36	3.00 %	15.00 X
	15	pLMIV	NZ_CM001470.1	77.825	34.02	0.01 %	0.01 X

## RESULTADOS

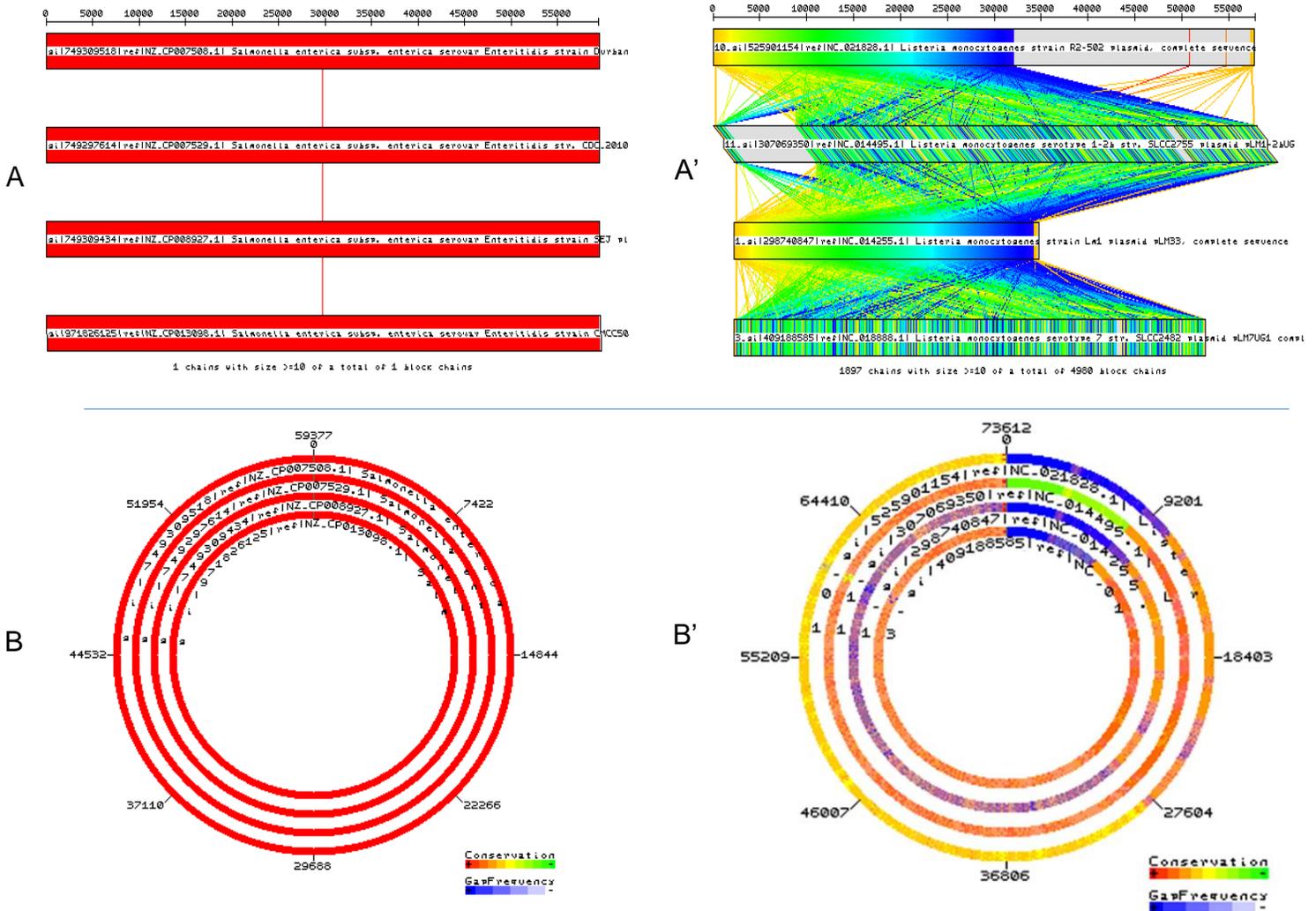
<b>E. coli</b>	9*	P010447_S 3	Km085453.1	3.173	46.90	99.00 %	14 049 X
	8*	P9705	AB040037.1	3.172	46.82	97.00 %	13 157 X
	7	pMccC7	AY913945.1	6.360	34.66	85.00 %	2.45 X
	6	pMccC7	X57583.1	6.458	34.42	83.00 %	2.41 X
	4	mccABCDE	AJ487788.2	5.734	34.32	78.00 %	2.09 X
	3	unnamed_2	NZ_CP009074.1	24.185	39.22	77.00 %	5.31 X
	1	MccC7-H22	NC_010257.1	32.014	43.50	36.00 %	3.53 X
	2	B	NZ_CP010127.1	42.655	39.89	30.00 %	2.74 X
	5	pRPEC180 _47	JN935898.1	47.506	41.43	28.00 %	2.99 X
<b>S. Enteritidis</b>	4	Unnamed	NZ_CP007508.1	59.372	51.96	99.00 %	18.24 X
	5	plasmid00	NZ_CP007529.1	59.369	51.92	99.00 %	18.24 X
	6	Unnamed	NZ_CP008927.1	59.372	39.88	99.00 %	18.24 X
	7	pCMCC500 41	NZ_CP013098.1	59.373	51.96	99.00 %	18.24 X
	2	pCFSAN00 0006	NZ_CP011395.1	59.336	51.96	98.00 %	16.72 X
	3	pSLA5	NC_019002.1	59.374	48.34	98.00 %	16.72 X
	9	Virulence	NZ_CP011943.1	59.37	51.96	98.00 %	16.72 X
	8	pFORC7	NZ_CP009767.1	101.43	51.95	52.00 %	8.71 X
	10	Intermediat e	NZ_CP011944.1	3.372	55.10	3.00 %	0.09 X
	1	pCFSAN00 0111_01	NZ_CP007599.1	39.599	51.96	0.00 %	0.00 X
	11	Small	NZ_CP011945.1	2.096	55.63	0.00 %	0.00 X

\*Plásmidos de *E. coli* encontrados mediante el ensamble *de novo* y el armado de *contigs*. No reportados para el cluster de genes Mcc.

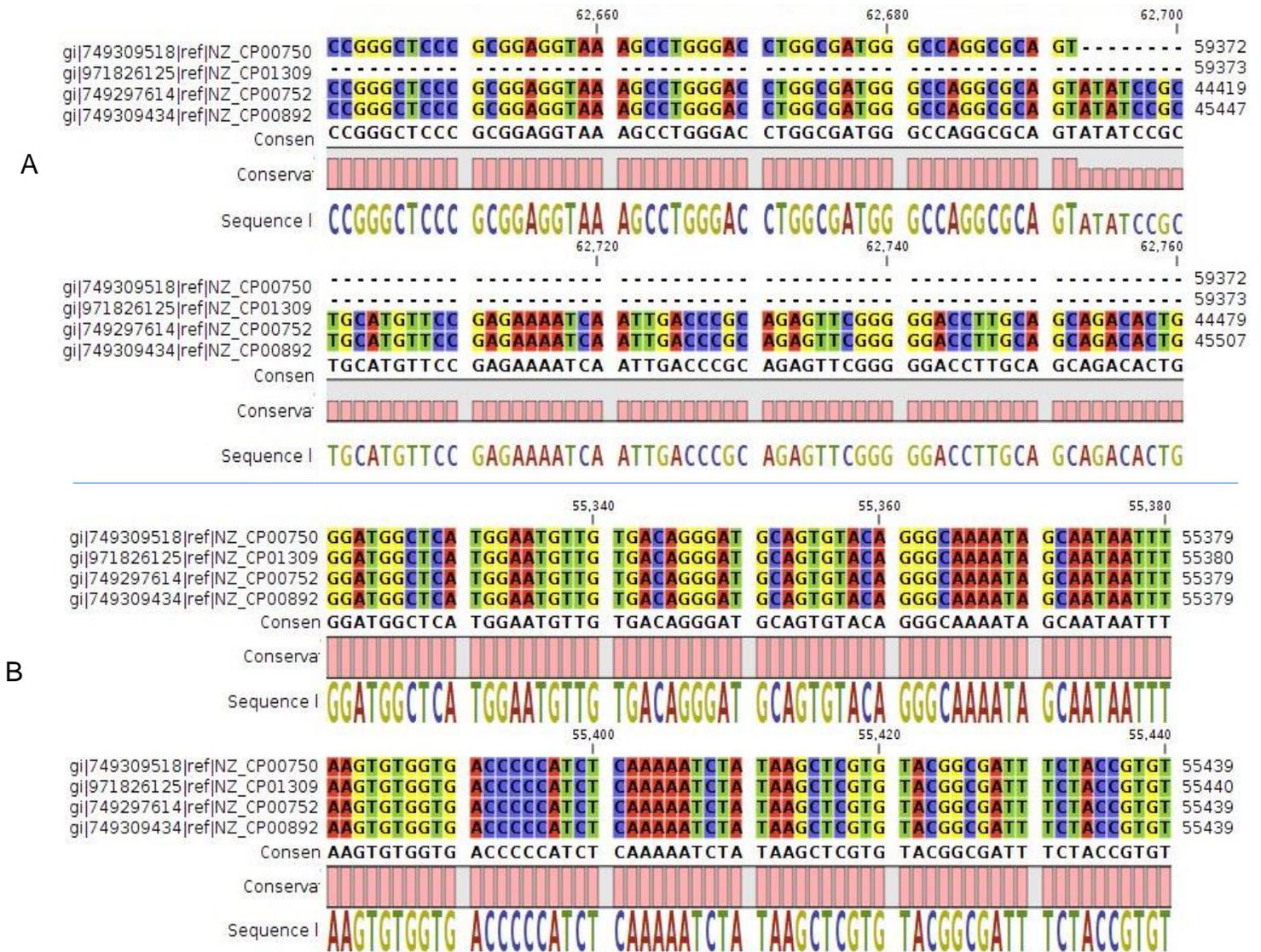
### **6.9.1 Análisis comparativos**

Los resultados descritos en la Tabla 20 se agruparon y ordenaron de manera decreciente con base el porcentaje de cobertura obtenido para cada plásmido. Como se puede observar, aquellos registros que obtuvieron el mismo porcentaje de mapeo, también comparten valores muy similares de profundidad, % GC y tamaño en pb, lo cual puede significar la presencia de referencias redundantes y por lo tanto una cantidad errónea de plásmidos identificados.

Para descartar la presencia de secuencias repetidas se utilizó el software CSA, el cual realiza el alineamiento de secuencias circulares. Este software, previo a realizar el alineamiento, establece un inicio hipotético en todas las secuencias a alinear, considerando la secuencia más larga compartida entre todos los registros, (Anexo 5) de esta forma, se obtuvieron los resultados mostrados en la Figura 30, donde las imágenes con plásmidos con similitud presuntiva, mantienen un alto grado de conservación (tonalidades rojas), pudiéndose considerar como redundantes entre sí, es decir, secuencias iguales, las cuales pueden descartarse para reducir la cantidad de referencias a analizar. Del mismo modo, aquellas que presentaron diferencias significativas tales como bajo nivel de conservación (tonalidad amarilla) así como huecos en el alineamiento (color azul-morado), se asumieron como referencias únicas. A su vez, también se realizaron alineamientos con el software CLCBio a modo de visualizar el contenido nucleótido base por base y confirmar las similitudes y diferencias que presentan los plásmidos agrupados, así como el efecto sobre los resultados para las secuencias procesadas con CSA y las que no recibieron tratamiento previo (Figura 31).



**Figura 30. Secuencias procesadas con CSA.** (Sección AA') Alineamiento de las secuencias reorganizadas (con un inicio hipotético sugerido) y visualización de su contenido nucleotídico mediante bloques colineales, cada bloque y su localización física se representan con un color diferente. (Sección BB') Circularización y alineamiento de las secuencias plasmídicas, el grado de conservación (similitud) se identifica mediante el gradiente colorimétrico de rojo (+) a verde (-) así como también la presencia de huecos mediante tonos grises a azules. (Sección AB) Identificación de registros repetitivos dada su alto grado de similitud. (Sección A'B') Identificación de registros únicos dado a las diferencias entre referencias.



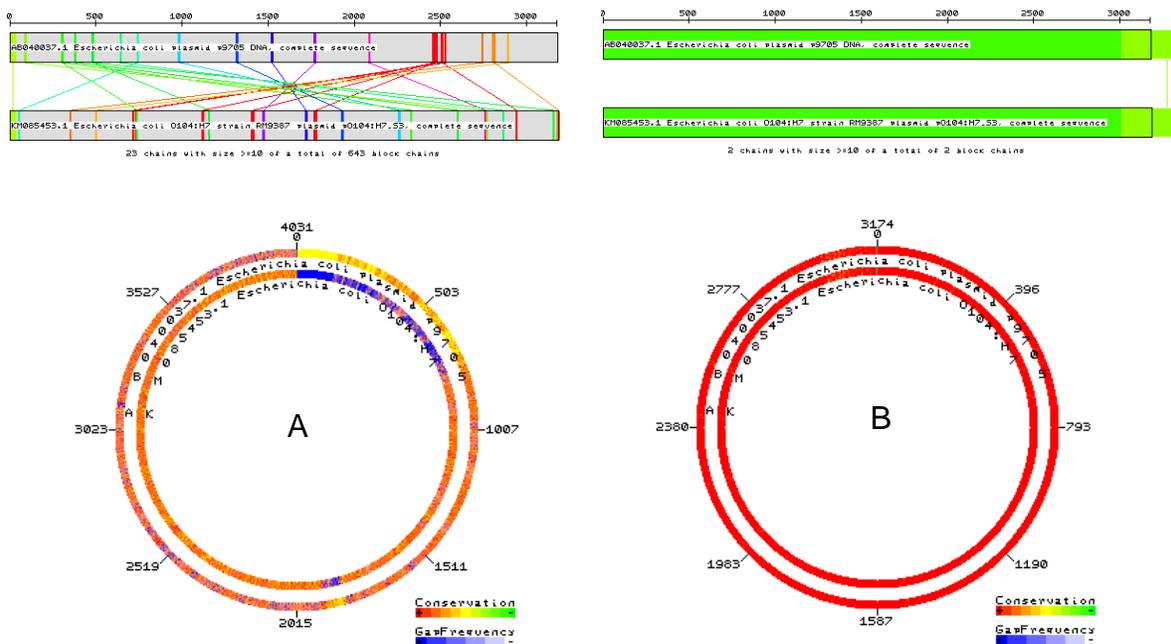
**Figura 31. Alineamientos plásmidos sin ordenar (A) y reorganizados (B).** En ambas imágenes se observa un alineamiento de las mismas cuatro referencias obtenidas de NCBI, (A) Referencias desordenadas sin reorganización. A simple vista presumen ser registros diferentes en cuanto al contenido nucleotídico, sin embargo al no poseer un orden o un inicio definido, es difícil determinar sus identidades y asumir que se tratan de secuencias distintas. (B) Referencias reorganizadas y ordenadas con el software CSA, se observan las identidades en los cuatro registros, pudiendo descartar así estas secuencias repetidas.

Eliminados los registros repetidos, se observó que hubo secuencias que continuaron presentando características en % GC, longitud, cobertura y profundidad, altamente similares, por lo que se decidió realizar un análisis extra, evaluando el sentido (5'-3' o 3'-5') en el cual se encontraban reportadas las secuencias En la Tabla 21 se agrupan en bloques de colores las referencias con persistencia en características similares.

**Tabla 21. Primera depuración y persistencia de registros altamente similares.**

Bacteria	ID	Plásmido reportado	No. Acceso	Tamaño (kb)	% GC	Cobertura	Profundidad
<i>L. monocytogenes</i>	3	pLM7UG1	NC_018888.1	50.1	35.48	100.00 %	18.45 X
	4	pLM1-2cUG1	NC_018889.1	50.1	35.48	100.00 %	18.45 X
	10	unnamed	NC_021828.1	57.557	36.03	100.00 %	17.54 X
	11	pLM1-2bUG1	NC_014495.1	57.78	36.04	100.00 %	17.71 X
	13	pCFSAN008100	NZ_CP011399.1	68.224	35.67	68.00 %	12.56 X
	6	pCFSAN010068_01	NZ_CP014251.1	55.521	36.7	65.00 %	12.85 X
	7	unnamed	NC_022046.1	55.804	34.92	35.00 %	7.88 X
	8	unnamed	NC_022047.1	55.804	34.92	35.00 %	7.88 X
	9	unnamed	NC_022051.1	55.803	34.92	35.00 %	7.88 X
	16	pLmN1546	NZ_CP013725.1	86.616	36.98	33.00 %	6.35 X
	17	II	NZ_HG813248.1	86.652	36.98	33.00 %	6.35 X
<i>E. coli</i>	9*	P010447_S3	Km085453.1	3.173	46.90	99.00 %	14 049 X
	8*	P9705	AB040037.1	3.172	46.90	99.00 %	13 157 X
	7	pMccC7	AY913945.1	6.360	34.66	85.00 %	2.45 X
	6	pMccC7	X57583.1	6.458	34.42	83.00 %	2.41 X
<i>S. Enteritidis</i>	4	unnamed	NZ_CP007508.1	59.372	51.96	99.00 %	18.24 X
	5	plasmid00	NZ_CP007529.1	59.369	51.92	99.00 %	18.24 X
	6	unnamed	NZ_CP008927.1	59.372	39.88	99.00 %	18.24 X
	7	pCMCC50041	NZ_CP013098.1	59.373	51.96	99.00 %	18.24 X
	2	pCFSAN000006	NZ_CP011395.1	59.336	51.96	98.00 %	16.72 X
	3	pSLA5	NC_019002.1	59.374	48.34	98.00 %	16.72 X
	9	Virulence	NZ_CP011943.1	59.37	51.96	98.00 %	16.72 X

Para realizar este análisis, se cargaron las secuencias similares dentro del software CLC, se obtuvo el reverso complementario de las secuencias para homologar las cadenas en una misma dirección. Homologados los datos, se realizó nuevamente el alineamiento en CSA, obteniendo resultados similares a los mostrados en la Figura 32. Se observa un comparativo entre el alineamiento de hebras complementarias (+/-) y el alineamiento ajustándolas hacia el mismo sentido (+/+).



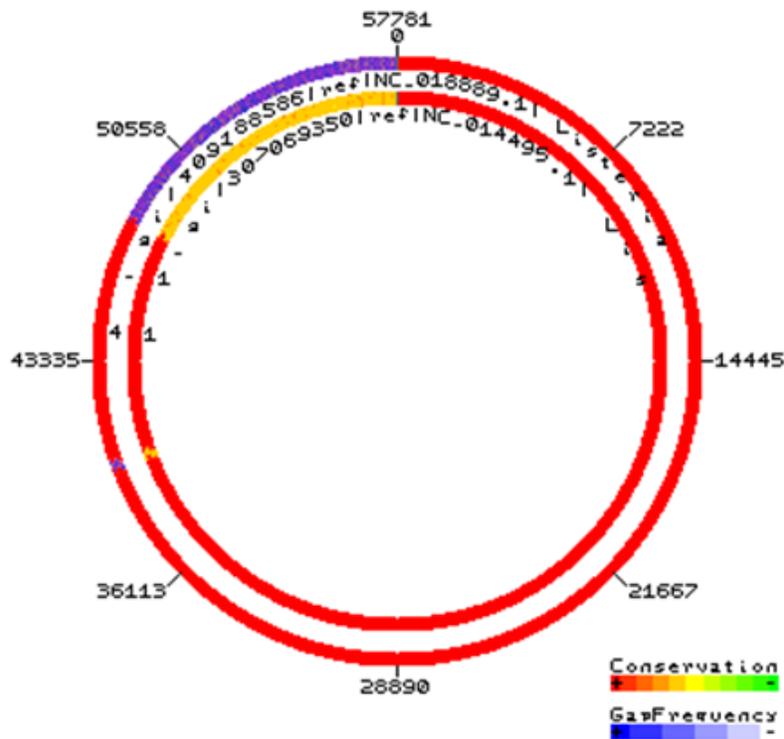
**Figura 32. Alineamiento circular del mismo par de plásmidos.** (A) Análisis en CSA con los datos crudos. Presunta identificación de registros distintos entre sí, sentido de las secuencias +/- . (B) Datos homologados al mismo sentido +/+ Identificación de plásmidos similares y redundantes entre sí.

Terminados los análisis de depuración, se realizó el reporte final de la identificación de aquellos plásmidos presentes dentro de las bacterias en estudio (Tabla 22), evitando presentar secuencias redundantes y una cantidad errada de elementos extracromosomales capturados.

**Tabla 22. Resumen final de plásmidos identificados para cada bacteria dentro del estudio. Mostrados en orden decreciente.**

Bacteria	ID	Plásmido reportado	No. Acceso	Tamaño (kb)	% GC	Cobertura	Profundidad
<i>L. monocytogenes</i>	1	pLM33	NC_014255.1	32.307	36.15	100.00 %	16.04 X
	4	pLM1-2cUG1	NC_018889.1	50.1	35.48	100.00 %	18.45 X
	11	pLM1-2bUG1	NC_014495.1	57.78	36.04	100.00 %	17.71 X
	13	pCFSAN008100	NZ_CP011399.1	68.224	35.67	68.00 %	12.56 X
	6	pCFSAN010068_01	NZ_CP014251.1	55.521	36.7	65.00 %	12.85 X
	14	pLM5578	NC_013767.1	77.054	36.59	52.00 %	9.61 X
	18	unnamed	NC_022045.1	148.959	37.05	44.00 %	6.84 X
	12	LM6179	NZ_HG813250.1	62.206	36.53	43.00 %	7.07 X
	9	unnamed	NC_022051.1	55.803	34.92	35.00 %	7.88 X
	17	II	NZ_HG813248.1	86.652	36.98	33.00 %	6.35 X
	5	CFSAN023459_02	NZ_CP014254.1	52.687	35.05	7.00 %	0.56 X
	2	CFSAN023459_01	NZ_CP014253.1	12.949	36.36	3.00 %	15 X
<i>E. coli</i>	9	P010447_S3	KM085453.1	3.173	46.90	99.00 %	14 049 X
	7	pMccC7	AY913945.1	6.360	34.66	85.00 %	2.45 X
	3	unnamed_2	NZ_CP009074.1	24.185	39.22	77.00 %	5.31 X
	1	MccC7-H22	NC_010257.1	32.014	43.50	36.00 %	3.53 X
	2	B	NZ_CP010127.1	42.655	39.89	30.00 %	2.74 X
	5	pRPEC180_47	JN935898.1	47.506	41.43	28.00 %	2.99 X
<i>S. Enteritidis</i>	9	Virulence	NZ_CP011943.1	59.373	51.96	99.00 %	18.24 X
	8	pFORC7	NZ_CP009767.1	101.43	51.95	52.00 %	8.71 X
	10	Intermediat e	NZ_CP011944.1	3.372	55.10	3.00 %	0.09 X

Adicional a la eliminación de secuencias redundantes, se encontró que algunos registros parecían estar alojados dentro de una secuencia de mayor tamaño, es decir, que el fragmento faltante no se encontraba reportado para el plásmido pequeño o bien podía pertenecer a una región móvil, la cual no se encontraba ligada al mismo al momento de determinar su secuencia (Figura 33), para verificar lo anterior se extrajo la secuencia no conservada ligada al plásmido de mayor tamaño y se realizó un BLAST a modo de evaluar el contenido de dicho fragmento. Se encontraron codificadas distintos tipos diferentes de transposasas además de una integrasa (Tabla 23), factores fundamentales para la trasposición y movilidad de elementos génicos.



**Figura 33. Plásmido de menor tamaño, contenido en uno con longitud mayor.** El plásmido alineado al centro de la imagen posee una región en tonos amarillos, representativo de un nivel intermedio de conservación, en contraste del plásmido ubicado al exterior del alineamiento, el cual ubica la misma región en tonos azules, representativo de una alta incidencia de huecos, este fenómeno se traduce en la falta de dicha región nucleotídica, es decir que a pesar de poseer una parte altamente conservada, el plásmido de mayor tamaño incluyó a su secuencia un elemento adicional el cual, una vez analizado su contenido se determinó que corresponde a una región codificante para enzimas utilizadas en la movilización de transposones e integrones.

**Tabla 23 Regiones de la secuencia transponible y enzimas para las que codifican el plásmido pLM1-2bUG1 de *L. monocytogenes* (NC\_014495.1).**

El fragmento transponible corresponde a la región delimitada por las posiciones 5,176 – 23,139 de la referencia antes mencionada, con una longitud de 17,964 nt, por lo tanto los elementos identificados dentro de dicho fragmento se identificaron en el intervalo de posiciones 1 – 17,964.

<b>Región</b>	<b>Enzima</b>
<1 – 656	<b>Transposasa IS712A</b>
9,756 – 10,436	<b>Transposasa IS1216</b>
10,467 – 10,868	
17,640 – >17,964	<b>Transposasa IS946V</b>
14,449 – 15,174	<b>Transposasa</b>
9,275 – 9,727	<b>Integrasa</b>

# *Capítulo 7*

## **DISCUSIÓN**

## **7.0 DISCUSIÓN**

### **7.1 Elección del mejor protocolo de extracción para la purificación de plásmidos**

El éxito en una corrida de secuenciación, radica en la calidad del ADN a secuenciar, es por esta razón que en el presente proyecto se trabajó con cuatro metodologías diferentes para la extracción de plásmidos, seleccionando aquella que mejor se adaptara a las necesidades de secuenciación en cuanto a calidad y cantidad de material genético.

Por lo anterior, una vez obtenidos los extractos de cada protocolo de extracción, se realizó el análisis cualitativo mediante electroforesis capilar. Se observó que todas las metodologías mostraron trazas de material genético a lo largo del gel digital mostrado por el Bioanalizador (Figura 17), por lo cual, las cuatro metodologías lograron la extracción del ADN. En el caso particular del Protocolo D, con la muestra de *L. monocytogenes*, se observó un barrido a lo largo del carril, con tamaños de alto (>10 Kb) y bajo (< 50 pb) peso molecular.

Cabe destacar que la sensibilidad del chip utilizado para visualizar el ADN de cada extracto abarca un intervalo de 50 a 10,380 pb. Por otro lado, como se puede apreciar en la Tabla 15, las referencias plasmídicas obtenidas en bases de datos poseen tamaños por encima de las 24 Kb; por lo cual, si las metodologías lograron realizar la purificación de algún elemento extracromosomal de manera íntegra, éste no podría ser visualizado dentro del intervalo representado, de tal modo que lo que se buscó fue, solo observar presencia de material genético para así asegurar la extracción, sin embargo de observar una cantidad elevada de ácidos nucleicos dentro del intervalo visible, tal como el caso del ensayo identificado como Protocolo D para la muestra *L. monocytogenes*, implicaría una prolongada degradación del ADN, convirtiendo a la muestra en inadecuada para continuar con el proceso.

Aunado a lo anterior, se evaluó la calidad de cada extracción, en la cual, de acuerdo con los parámetros establecidos en el CNRDOGM, se consideran aceptables valores de relación de absorbancias 260 y 280 nm (A260/A280) dentro del intervalo entre 1.8 – 2.1. El obtener valores por encima de 2.1 representa una contaminación de la extracción con cantidades representativas de ARN, así mismo, valores por debajo de 1.8 son indicativos de la presencia de contaminación por proteínas. En ambos casos este tipo de contaminación podría afectar el proceso de secuenciación <sup>[42]</sup>.

Por otra parte, también se realizó el análisis cuantitativo de material genético de doble cadena. Considerando que el tamaño promedio de un plásmido representa entre el 1 – 10 % del cromosoma bacteriano <sup>[31]</sup>, resulta complicado establecer un intervalo de aceptación para la cantidad de ADN a extraído, sin embargo, al realizar la cuantificación siguiendo el protocolo del equipo Qubit en su modalidad Broad Range, se lograron registrar valores de cuantificación para ADN de doble cadena (Tabla 14) para todas las muestras, exceptuando la muestra de *L. monocytogenes* extraída con el protocolo D, en la cual no se obtuvo valor de cuantificación, además de arrojar resultados de calidad por debajo de los parámetros establecidos corroborando la degradación de esta muestra (Figura 17).

Adicional a la cuantificación y cualificación de los extractos, se decidió determinar la especificidad para extraer plásmidos de las metodologías empleadas. Para esto, se evaluó la presencia de residuos cromosomales mediante amplificación por PCR del gen ribosomal 16S, gen multicopia ubicado a lo largo del cromosoma bacteriano, ya que la presencia de este amplicón sería evidencia de la existencia de restos cromosomales.

Como se aprecia en la Figura 18, todas las metodologías presentaron amplificación del gen ribosomal 16S, por lo cual ningún método resultó ser completamente selectivo a plásmidos, acarreado con ellos residuos cromosomales, mismos que fueron removidos mediante digestión enzimática, con

la enzima Plasmid Safe-DNasa, capaz de degradar material genético de cadena sencilla, gracias a su actividad de exonucleasa de doble cadena.

Cabe destacar que los plásmidos no sólo se presentan de forma circular, la estructura lineal ha sido reportada para otras bacterias por lo cual además de digerir remanentes cromosomales, también eliminarían aquellos elementos plasmídicos no circulares, sin embargo, se ha demostrado que los plásmidos lineales han desarrollado estrategias de protección en sus extremos generando secuencias complementarias al final de cada cadena, propiciando la formación de bucles e impidiendo el ataque por exonucleasas <sup>[43]</sup>. Una vez realizada esta digestión enzimática, se realizó un nuevo ensayo de PCR para el gen ribosomal 16S, el cual ya no mostró amplificación (Figura 19), asegurando así la digestión del cromosoma.

Descartados los residuos cromosomales, se realizó a cada extracto purificado un ensayo de amplificación dirigido específicamente a plásmidos, con el fin de confirmar su presencia. Se emplearon iniciadores específicos a secuencias propias de los plásmidos de cada especie y/o serovariedad, las cuales resultaron positivas para los protocolos de extracción B y D, de las muestras de *L. monocytogenes* y *E. coli* (Figura 20A y 20B respectivamente) demostrando la presencia de estos elementos para su posterior secuenciación. La muestra de *S. Enteritidis* no mostró amplificación (Figura 20C), lo cual podría relacionarse con la ausencia del plásmido para el cual se diseñaron los cebadores, sin embargo, se continuó trabajando con ella por la posible presencia de más elementos plasmídicos.

En resumen, en las cuatro metodologías se realizó un correcto aislamiento de material genético; las metodologías A y D registraron calidades fuera del intervalo de aceptación y únicamente en las metodologías B y D, se verificó la presencia de elementos plasmídicos posteriores a la digestión enzimática. De tal modo que, tomando en cuenta las calidades de cada extracto en conjunto con la efectiva

purificación de plásmidos, el protocolo B (ZR Plasmid Miniprep™-Classic) resultó ser el más adecuado para los propósitos del proyecto.

En la Tabla 24 se resume y compara el desempeño de cada protocolo ante cada prueba realizada. La metodología B cumplió de manera satisfactoria en dos de sus tres extracciones con la totalidad de las pruebas.

Tabla 24. Resultados de la purificación para cada método y muestra.

Muestra y Protocolo	Cuantificación	Calidad	Amplificación de regiones plasmídicas	Adecuado
<i>L. monocytogenes</i> (A)	✓	✗	✗	✗
<i>E. coli</i> (A)	✓	✗	✗	
<i>S. enterica</i> serovar enteritidis (A)	✓	✗	✗	
<i>L. monocytogenes</i> (B)	✓	✓	✓	✓
<i>E. coli</i> (B)	✓	✓	✓	
<i>S. enterica</i> serovar enteritidis (B)	✓	✓	✗	
<i>L. monocytogenes</i> (C)	✓	✓	✗	✗
<i>E. coli</i> (C)	✓	✓	✗	
<i>S. enterica</i> serovar enteritidis (C)	✓	✓	✗	
<i>L. monocytogenes</i> (D)	✓	✗	✓	✗
<i>E. coli</i> (D)	✓	✓	✓	
<i>S. enterica</i> serovar enteritidis (D)	✓	✗	✗	

## 7.2 Secuenciación masiva

Debido a que la cantidad de ADN aislado de una extracción plasmídica es mucho menor con respecto a una extracción de la totalidad de ácidos nucleicos presentes en la célula, fue necesaria la amplificación de todo el material genético contenido en cada extracto. Ya que no existe una secuencia en común dentro de todos los plásmidos que pueda utilizarse como región universal para dar inicio a una reacción de PCR, se empleó el *Illustra™ GenomiPhi DNA Amplification Kit* para

amplificar los plásmidos purificados. Como se mencionó en la metodología, el kit actúa mediante una PCR isotermal por desplazamiento de hebra, empleando iniciadores de seis nucleótidos. Al ser un método inespecífico, resulta imprescindible haber realizado una correcta purificación de los extractos de modo que el ADN amplificado corresponda sólo a plásmidos y no a contaminantes cromosomales provenientes de la extracción.

Amplificada la muestra, se inició la generación de librerías, se continuó con las etapas de amplificación clonal y de enriquecimiento y posteriormente se dio paso a la corrida de secuenciación.

Como resultado de la corrida, se obtuvieron un total de 424 millones de bases secuenciadas, en 1, 907, 318 lecturas con una longitud promedio de 222 pb. Como se detalló previamente, los fragmentos esperados corresponden a longitudes cercanas a las 400 bases con lo cual el tratamiento de los datos resulta más sencillo, sin embargo, a pesar de haber obtenido una longitud promedio menor a lo esperado, los datos recibieron el mismo tratamiento, sustentado en la flexibilidad que posee la plataforma Ion Torrent la cual cuenta con un protocolo estandarizado para fragmentos de longitudes menores.

Los resultados obtenidos del reporte de secuenciación nos indican haber aprovechado el chip a capacidad intermedia. El obtener un cargado de 60 % representó información suficiente para realizar los análisis comparativos y la captura de elementos plasmídicos. Al fijar como blanco la captura de plásmidos, elementos génicos relativamente pequeños en comparación a un genoma bacteriano, la cantidad de datos obtenidos resultó suficiente para desarrollar los análisis de identificación y captura de plásmidos de manera favorable.

### **7.3 Análisis bioinformático**

Culminada la corrida de secuenciación, se dio paso al análisis bioinformático, el cual inició con el análisis de las lecturas mediante el software FastQC, evaluando las calidades de las mismas al establecerles un valor Q de la escala Phred. El índice Phred se basa en una escala logarítmica descrita con el algoritmo  $q =$

$(-10)(\log(p))$  siendo “q” el valor de la escala y “p” la probabilidad de obtener un error cada determinado número “n” de nucleótidos asignados, se eligió realizar los análisis de calidad, aceptando solo aquellas lecturas que permanecieran por encima del valor de Q20 (Figura 25) dado que éste es el valor internacionalmente aceptado, de este modo se aseguró que de existir un error, la probabilidad de este sería de uno por cada 100 bases asignadas, lo cual, asegura resultados de calidad en los ensamblajes y mapeos realizados <sup>[12]</sup>.

Así mismo, se hizo uso de la base de datos de One Codex, la cual cuenta con un motor de búsqueda para la identificación de la abundancia de diversos organismos dentro de una muestra. One Codex aloja en sus registros más de 40 000 genomas completos de microorganismos, correspondientes a 11 000 especies bacterianas distintas, 4 000 especies virales y cientos de especies eucariotas y archeas, además cuenta con una versión del genoma humano para analizar las lecturas correspondientes al hospedero <sup>[17]</sup>.

One Codex permite realizar una identificación precisa de microorganismos partiendo de las lecturas obtenidas mediante secuenciación tipo “shotgun” <sup>[17]</sup>. Mediante el análisis de los datos arrojados por el secuenciador, se logró la identificación de los tres microorganismos involucrados en el estudio (*L. monocytogenes*, *E. coli*, y *S. Enteritidis*).

La identificación a nivel especie de microorganismos mediante One Codex resulta confiable; sin embargo, este software no es capaz de realizar los análisis de identificación de cada elemento génico presente en cada muestra. Por lo cual, para realizar la identificación de los plásmidos presentes, se utilizó el software CLCBio, que forma parte de un grupo de softwares de alta eficiencia, con la capacidad de analizar datos provenientes de distintas plataformas de secuenciación como Sanger, Illumina, Roche 454, Pacific Biosciences y Life Technologies.

Entre otras características, el software permite realizar trabajos con ácidos nucleicos como el mapeo de las lecturas provenientes tanto de Sanger como de secuenciación de alto rendimiento, ensambles *de novo*, detección de SNP's, análisis de transcriptómica y BLAST, por lo que resulta una herramienta muy completa <sup>[20]</sup>.

Los datos de secuenciación generados se introdujeron en el software CLCBio y se llevó a cabo el ensamble *de novo*, el cual tras realizar el solapamiento de las secuencias similares que tienen su origen en la misma región de la secuencia original, genera consensos de cadenas de ADN con una longitud mayor a las lecturas individuales, armando una pequeña región del elemento génico origen (Figura 28). Al extraer el consenso de mayor longitud (*contig* más largo), éste se analizó mediante BLAST, obteniendo valores de cobertura (Query cover) e identidad de entre 99 y 100 %, los cuales representan el porcentaje de bases alineadas y de estas bases, cuantas coincidieron de manera exacta a la referencia asociada respectivamente. Estos valores fueron suficientes para determinar el origen plasmídico de los datos y corroborar lo previamente obtenido mediante One Codex, por lo cual se confirmó que la metodología utilizada para la purificación de plásmidos resultó adecuada para secuenciar y analizar de manera dominante elementos de origen extracromosomal.

Para la búsqueda de los posibles plásmidos presentes en cada una de las bacterias en estudio, se utilizó otra herramienta del software CLCBio que también consiste en montar un genoma, pero esta vez realizando un mapeo de todas las lecturas obtenidas a un plásmido de referencia, eliminando así la búsqueda a ciegas de secuencias consenso. Los plásmidos de referencia utilizados para los mapeos fueron los que se encontraron reportados en la base de datos del NCBI (Tabla 15), incluyendo aquellos a los cuales se les encontró identidad mediante el ensamble *de novo* y el alineamiento en BLAST de los *contig* más representativos de cada muestra.

De esta forma se capturaron para *L. monocytogenes* cinco plásmidos con una cobertura del 100 %, diez plásmidos con coberturas de entre 90 – 10 % y tres más con porcentajes mínimos de cobertura de entre 3.0 – 0.0 %; para *E. coli* dos plásmidos con coberturas de 99.0 y 97.0 % y siete con coberturas de entre 90 – 10 %; en el caso de *S. Enteritidis*, se obtuvieron cuatro plásmidos con una cobertura de 99.0 %, tres con una cobertura del 98.0 %, un plásmido con cobertura del 52.0 % así como tres más con porcentajes mínimos del 3.0 – 0.0 % (Tabla 20)

Dentro de las características elementales de los elementos extracromosomales tenemos la capacidad de autorregular su replicación, mantener un número definido de copias, así como permanecer dentro de la célula siempre y cuando estos pertenezcan a distintos grupos de incompatibilidad <sup>[31]</sup>. Siguiendo con lo anterior y considerando que el número de copias y la distribución espacial dentro de la célula representan un reto para la organización del material genético, se piensa que la presencia de una gran cantidad de elementos extracromosomales distintos dentro de una bacteria resulta improbable por lo que se realizó una búsqueda de redundancias en nuestros registros descargados.

Con base en el porcentaje de cobertura obtenido para cada plásmido, se ordenaron decrecientemente las secuencias capturadas para agrupar los registros similares en una referencia única, descartando los datos con contenido repetido. La depuración se obtuvo con el software CSA que realiza un alineamiento circular, partiendo de una región altamente conservada entre los registros comparados y extendiendo el alineamiento a lo largo de las secuencias. Con esto se logró una primera eliminación de secuencias redundantes, sin embargo, se observó que aún persistían registros con características altamente similares por lo que se procedió a un segundo análisis. El software CSA al ser sólo un alineador de elementos circulares trabaja únicamente con un sentido de las secuencias, por lo cual, se determinó el antisentido de éstas con el software CLCBio y posteriormente se realizó el alineamiento con CSA, obteniendo más registros similares (Figura 32).

Finalmente, se redujo la cantidad de elementos identificados a 12 plásmidos para *L. monocytogenes*, seis plásmidos para *E. coli* y 3 para *S. enterica* serovar Enteritidis, con porcentajes de captura de entre el 3.0 al 100.0 %.

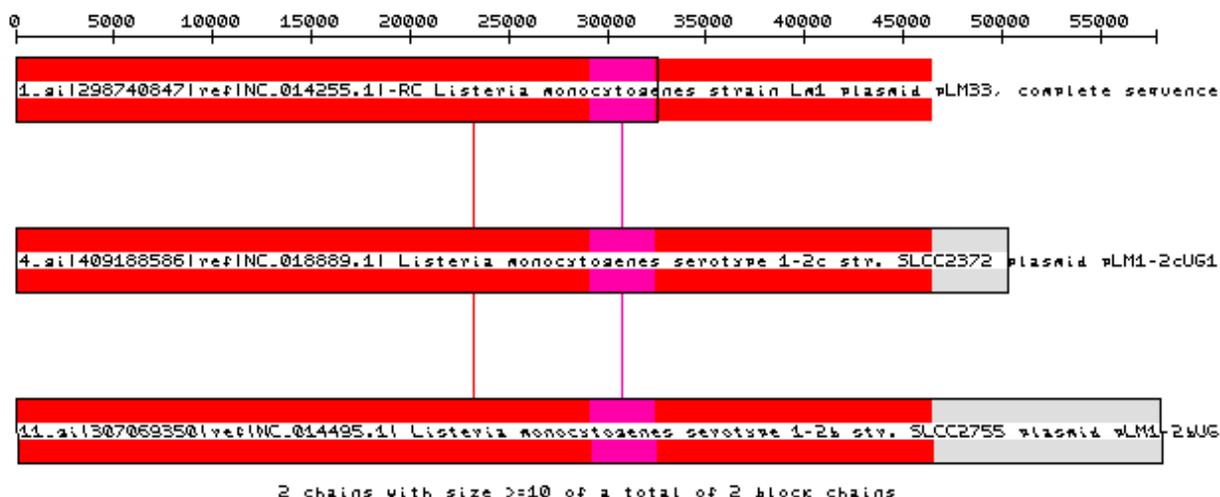
Ya que el trabajo consistió en el aislamiento y purificación de ADN de origen plasmídico y en consecuencia los datos secuenciados y analizados corresponden a elementos extracromosomales, se consideraron como plásmidos presentes a aquellos cuyo porcentaje de captura resultó superior al 70 %. Para el resto, la posibilidad de encontrarlos de manera física se vuelve improbable. Al hablar de plásmidos debemos considerar la presencia de elementos transportables, capaces de migrar de un elemento a otro en forma de transposones, integrones, cassettes de genes; compartidos a través de transferencia horizontal mediante cualquiera de sus mecanismos (transformación, conjugación o transducción), por lo que es posible haber realizado la captura de regiones móviles insertadas en las referencias utilizadas, pero éstas a su vez no forman parte estática de un plásmido en particular <sup>[31]</sup>.

Sin embargo, se han encontrado plásmidos pequeños de entre 3 y 6 Kb alojados dentro de la bacteria *E. coli*, entre los cuales se identificó la presencia parcial del bloque de genes *mccX* responsables de propiciar la inhibición de algunos patógenos entéricos mediante la síntesis de diferentes tipos de bacteriocinas <sup>[39]</sup>, así como la presencia del gen que codifica para la proteína Rop conocida como supresora del iniciador responsable de la regulación negativa en la replicación plasmídica estabilizando las interacciones entre el RNAII (cebador complementario al sitio *ori colE1*) y el RNAI (secuencia complementaria del RNAII) impidiendo la formación del complejo de síntesis, regulando la cantidad de copias presentes dentro de la bacteria <sup>[40]</sup>.

La presencia de estos elementos de corto tamaño y con funciones no esenciales pero adaptativas, suelen formar parte de acarreadores de genes trasponibles o elementos transportables que generalmente salen de la secuencia extracromosomal para insertarse dentro del cromosoma e incrementar su

replicación favoreciendo los niveles de expresión de cierta característica adaptativa [9].

Dicho fenómeno se vió reflejado en los análisis comparativos del contenido nucleotídico en busca de redundancias plasmídicas (Figura 33), en donde se encontró un caso particular para la bacteria *L. monocytogenes*, ya que los plásmidos identificados con los ID 1, 4 y 11 mostrados en la Figura 34, se observan alineados en bloques colineales, identificados con el mismo color y unidas mediante una línea recta a las regiones conservadas entre ellos. Por otra parte, los plásmidos 4 y 11 muestran en bloques grises segmentos de las secuencias que no se comparten por lo que no adquieren la representación de bloque colineal, lo cual se interpreta como secuencias adicionales dentro de cada referencia, por tanto, se puede presumir que se trata de la inserción de un elemento móvil localizado dentro del plásmido en la muestra secuenciada.



**Figura 34. Identificación de inserciones dentro de una secuencia conservada.** Se agruparon dos bloques coloniales de secuencias altamente conservadas identificados con tonos rojos y rosas alineados en la misma posición para los tres registros. A los extremos se muestran dos bloques en tono gris, que no presentan identidad ni similitud con ningún contenido presente en el resto de las referencias, se sugiere la posible inserción de elementos móviles a la secuencia plasmídica.

El aplicar técnicas de NGS provee una herramienta rápida y eficaz para la identificación y caracterización de los organismos involucrados en brotes epidemiológicos y alertas sanitarias. El ampliar el conocimiento sobre el contenido genético alojado en los plásmidos de las bacterias facilitará la comprensión de los mecanismos de adaptación utilizados por cada microorganismo, además de revelar información de utilidad para el control epidemiológico al identificar las posibles ventajas adaptativas de cada patógeno.

# *Capítulo 8*

## **CONCLUSIONES**

### 8.0 CONCLUSIONES

Se estableció el protocolo B *ZR Plasmid Miniprep™-Classic* como la metodología de estandarización para la extracción, purificación y posterior secuenciación de ADN plasmídico para las tres bacterias en estudio, para lo cual:

A partir de un total de 27 referencias plasmídicas reportadas en la base de datos Ref-Seq para las bacterias de interés, se diseñaron seis pares de iniciadores dirigidos a regiones únicas de plásmido.

Mediante PCR con los iniciadores diseñados, así como con iniciadores universales dirigidos al gen ribosomal 16S, se verificó el aislamiento y purificación de plásmidos.

El ADN purificado fue amplificado mediante PCR isotermal por desplazamiento de hebra, estableciéndolo como un método flexible para incrementar la cantidad de ADN en una muestra cuyo contenido molecular es reducido.

Mediante NGS y su posterior análisis bioinformático, se identificó a cada una de las bacterias en estudio, asimismo, se encontraron un total de 22 presuntas secuencias plasmídicas, de las cuales se descartaron las secuencias redundantes, obteniendo un total de nueve secuencias únicas. Por otra parte, fue posible localizar secuencias altamente conservadas con diferentes inserciones de fragmentos de ADN distintos entre sí, entre los cuales se identificaron enzimas involucradas en la transposición de elementos génicos.

# *Capítulo 9*

## **PROSPECTIVAS**

## **9.0 PROSPECTIVAS**

- Dar inicio a la construcción de una base de datos propia del CNRDOGM con los registros plasmídicos presentes en las bacterias aisladas de alimentos provenientes de todo el territorio nacional.
- Profundizar en el análisis y estudio del contenido molecular de los plásmidos aislados, propiciando el desarrollo y comprensión de los métodos de adaptación, virulencia y patogenicidad presentados por las bacterias patógenas de interés agroalimentario.
- Identificar las principales resistencias a antimicrobianos contenidas en plásmidos, las cuales se encuentran en los organismos patógenos aislados de los alimentos, esto para fomentar el desarrollo del combate y prevención contra la adquisición de enfermedades causadas por el consumo de alimentos.

# *Capítulo 10*

**ANEXOS**

## 10.0 ANEXOS

### Anexo 1. Cargado y revelado de chip High Sensitivity DNA y 7500 DNA en el equipo Bioanalyzer 2100 Agilent

1. Atemperar la mezcla gel-dye durante 30 minutos antes realizar el cargado de los pozos (Figuras 35 y 36), protegiéndola de la exposición a la luz durante este periodo.
2. Colocar el chip High Sensitivity DNA o 7500 DNA en la estación de cargado correspondiente y agregar 9.0  $\mu$ L de la mezcla gel-dye en el pozo marcado con el símbolo .
3. Colocar el cronometro en 60 s (High Sensitivity DNA) o 30 s (7500 DNA) asegurando que el émbolo este posicionado en la marca de 1.0 mL y presionar el émbolo hasta escuchar “click”. Comenzar la cuenta regresiva. Transcurrido el tiempo, liberar el gatillo permitiendo que el mecanismo se relaje hasta pasar a la marca de 0.7 mL y lentamente regresar manualmente el émbolo hasta la marca de 1.0 mL.
4. Agregar 9.0  $\mu$ l de la mezcla gel-dye en cada pozo marcado con la letra G.
5. Agregar 5.0  $\mu$ l del marcador de ADN contenido en el vial con tapa verde en los pozos de muestra (12 para 7500 DNA Chip y 11 para High Sensitivity DNA Chip) incluyendo el pozo para el marcador de peso molecular identificado con el símbolo .
6. Agregar 1  $\mu$ l del marcador de peso molecular contenido en el vial con tapa amarilla en el pozo identificado con el símbolo .
7. Agregar a cada pozo de muestra 1  $\mu$ l de muestra y en caso de no contar con muestras suficientes para los pozos disponibles, agregar 1  $\mu$ l de agua grado biología molecular a los pozos sin usar.
8. Mezclar en vórtex el chip por 60 s a 2400 rpm e insertarlo en el equipo 2100 Bioanalyzer, previamente lavado.
9. Abrir el software 2100 expert y seleccionar la modalidad a trabajar (7500 o High Sensitivity).
10. Nombrar los pozos con las muestras correspondientes y seleccionar la opción START.

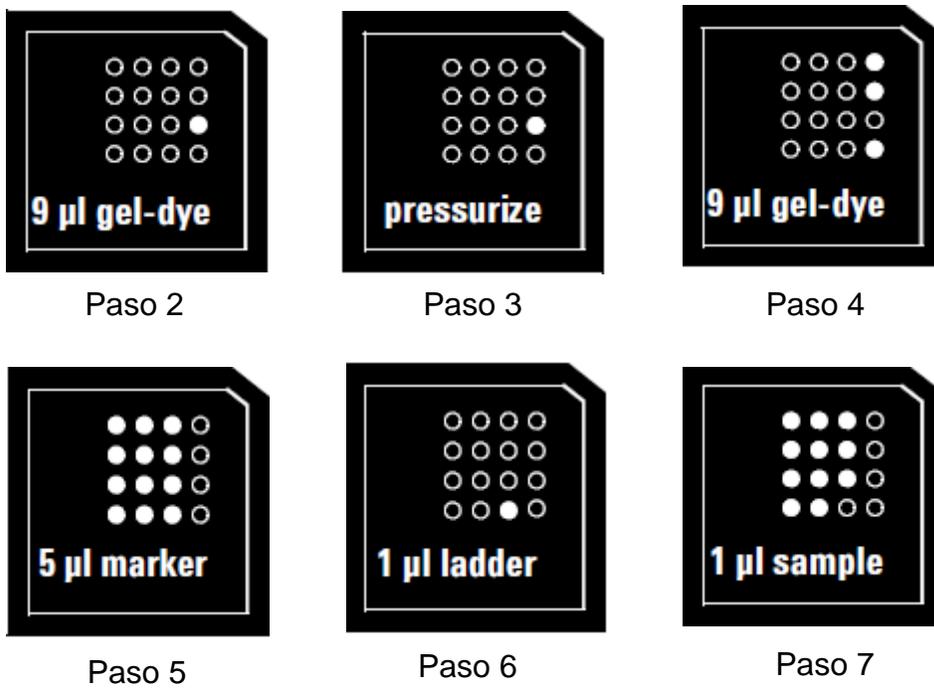


Figura 35. Pozos de cargado de las muestras en el Chip 7500 DNA

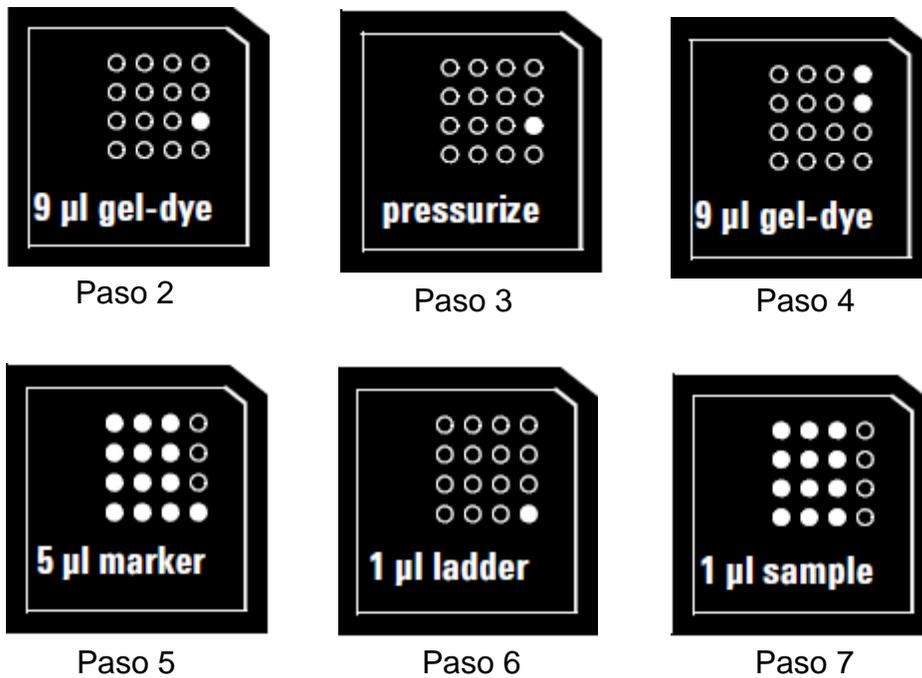
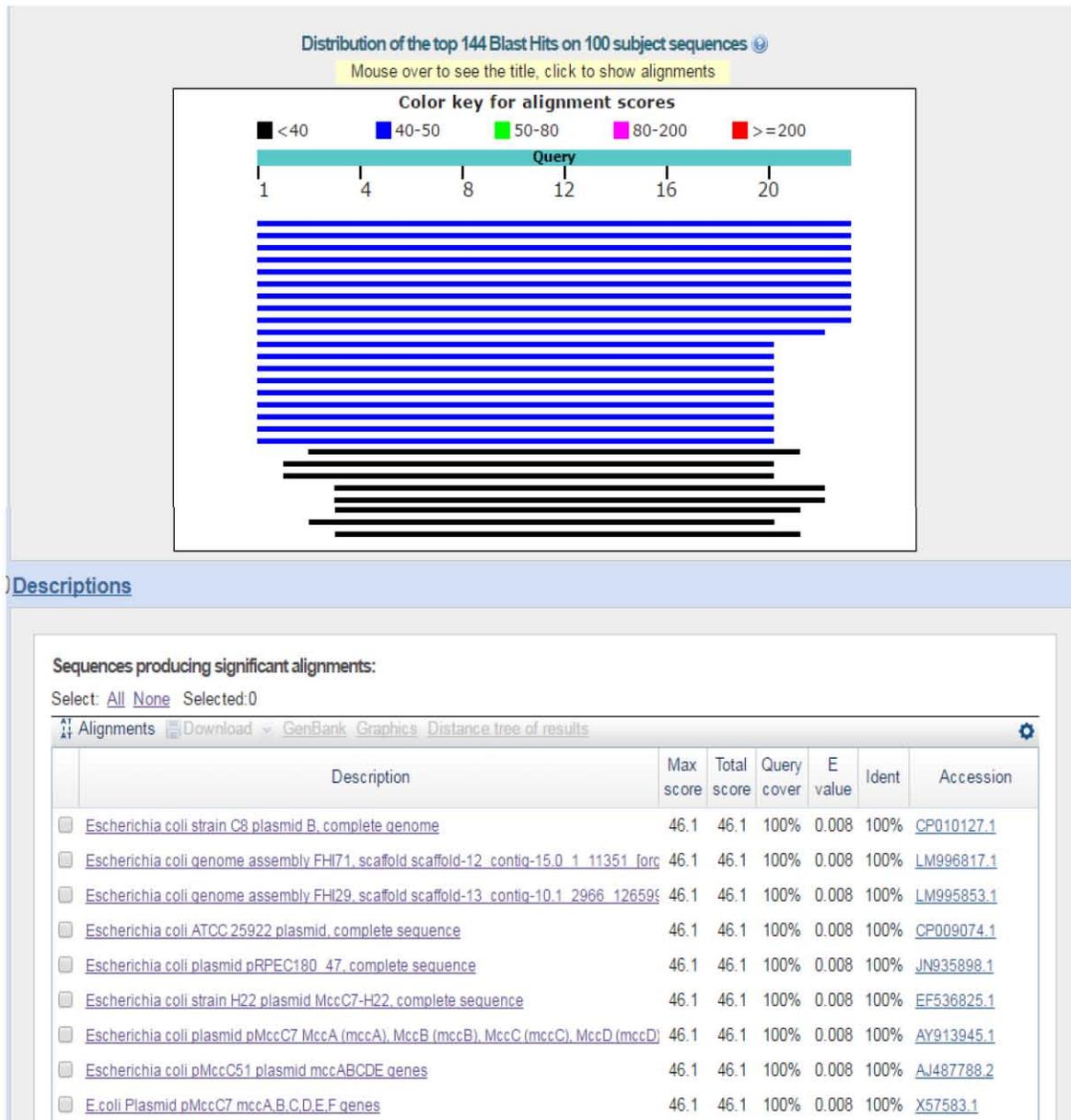


Figura 36. Pozos de cargado de las muestras en el Chip High Sensitivity DNA

## Anexo 2. Score de los iniciadores para validar especificidad

Porcentajes de identidad y cobertura del 100 % para los alineamientos realizados. Pantalla con los valores obtenidos en BLASTn (Figura 37) y resumen detallado del total de iniciadores diseñados (Tabla 25).



**Figura 37. Valores obtenidos del alineamiento BLASn.** Se observan alineamientos con coberturas e identidades del 100 % a elementos génicos de origen plasmídico, ratificando la especificidad de los iniciadores diseñados.

**Tabla 25. Valores de identidad, cobertura y Valor de expectación para cada iniciador sintetizado.**

Microorganismo	Oligonucleótido		Query cover	E value*	Identidad	No. de acceso
<i>L. monocytogenes</i>	pLM7UG1_F1	Sentido	100 %	0.002	100 %	NC_018888.1
	pLM7UG1_R1	Antisentido	100 %	0.002	100 %	
	pLM7UG1_F2	Sentido	100 %	0.002	100 %	
	pLM7UG1_R2	Antisentido	100 %	0.002	100 %	
<i>E. coli</i>	mccB_F1	Sentido	100 %	0.008	100 %	NZ_CP009074.1
	mccB_R1	Antisentido	100 %	0.008	100 %	
	mccB_F2	Sentido	100 %	0.008	100 %	
	mccB_R2	Antisentido	100 %	0.008	100 %	
S. Enteritidis	S_pCFSAN_F1	Sentido	100 %	3e-05	100 %	NZ_CP007599.1
	S_pCFSAN_R1	Antisentido	100 %	8e-06	100 %	
	S_pCFSAN_F2	Sentido	100 %	1e-04	100 %	
	S_pCFSAN_R2	Antisentido	100 %	3e-05	100 %	

\* El parámetro de expectación o E value, representa el número de hits que podemos encontrar al buscar dentro de una base de datos de tamaño determinado, cuanto más cercano a cero sea el valor del E value más significativo se vuelve el alineamiento y disminuye la aleatoriedad del alineamiento. Cabe mencionar que cuando se trabajan con secuencias cortas es muy común encontrar valores de expectación altos ya que la probabilidad de encontrar más de una coincidencia dentro de la base de datos incrementa. Se observa que los valores obtenidos de E-value de las secuencias de iniciadores diseñados presentan valores muy cercanos a cero <sup>[29]</sup>.

Los cálculos detallados del E-value se describen en el portal del NCBI *The Statistics of Sequence Similarity Scores* (<https://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html>).

### Anexo 3. Interpretación del porcentaje de recuperación de ISPs.

La cantidad óptima de biblioteca corresponde al punto de dilución de la biblioteca que da porcentajes de ISPs entre 10-40%. Las muestras que están dentro del rango recomendado generalmente producen una mayor cantidad de datos; Sin embargo, las muestras que caen fuera del rango recomendado aún pueden cumplir con las especificaciones de rendimiento en los Ion Chips. El rango óptimo recomendado no pretende ser un criterio de aprobación / rechazo. El rango proporciona orientación para la calidad de la muestra (Tabla 26).

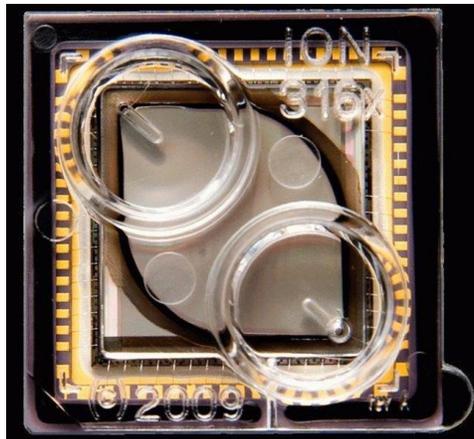
**Tabla 26. Interpretación de los valores del % de recuperación. Parámetros implementados dentro de los ensayos de secuenciación del CNRDOGM.**

Porcentaje de ISPs	Descripción
< 10 %	La muestra contiene un número insuficiente de ISPs para lograr una densidad de carga óptima en el chip de iones.
10 – 40* %	Cantidad óptima de librería.
> 40* %	La muestra proporcionará ISP policlonales (lecturas mixtas).

\* Intervalo establecido dentro de los parámetros de aceptación del CNRDOGM

**Anexo 4. Descripción del Ion 316 v2™ Chip.**

- Proporciona entre 300 MB a 1 GB de datos de secuencia con un tiempo de ciclo rápido de 3 a 5 horas con un estimado de 2 a 3 millones de lecturas (pozos).
- Compatibilidad con diversas aplicaciones, incluidas la secuenciación génica selectiva y la secuenciación microbiana con entre 2 y 3 millones de lecturas por ciclo.
- Fácil manejo y carga de partículas Ion Sphere™ con plantilla.
- Compatibilidad con todos los métodos actuales de preparación de bibliotecas.
- Versión del Chip grabada sobre la superficie. Figura 38



**Figura 38. Ion 316™ Chip.**

## Anexo 5. Líneas de comando para ejecutar CSA.

Sistema operativo Linux.

Descargar los archivos .fasta dentro del mismo directorio y colocarlo dentro del directorio "CSA".

Situarse sobre el directorio que contiene las secuencias.

Para concatenar todos los archivos .fasta en un solo archivo multifasta:

```
$cat *fasta > archivo.fasta
```

Para desplegar el menú de opciones CSA:

```
$CSA
```

```
>USAGE:
```

```

[ Rotate+Align+Image ] CSA <multi-fasta-file>
[ Rotation only      ] CSA R <multi-fasta-file>
[ Alignment only     ] CSA A <multi-fasta-file>
[ Alignment Image only ] CSA I <multi-fasta-file>

```

```
>TOOLS:
```

```

[ Clean FASTA file   ] CSA C <multi-fasta-file>
[ Get Alignment score ] CSA S <multi-fasta-file>
[ Convert FASTA to MSF ] CSA M <multi-fasta-file>

```

Para obtener la imagen del alineamiento colineal, alineamiento circular y secuencias rotadas en un solo paso:

```
$ CSA <multi-fasta-file>
```

# *Capítulo 12*

## **REFERENCIAS**

---

## 11.0 BIBLIOGRAFÍA

1. Aya Brown, K., Itai, B., & Itzhak, M. (2013). A Method For Purifying High Quality And High Yield Plasmid DNA For Metagenomic And Deep Sequencing Approaches. *Journal of Microbiological Methods*, 272-279. doi:10.1016/j.mimet.2013.09.008
2. Cabrera, C. E., Gomez, R. F., & Zuñiga, A. E. (2007). La resistencia de bacterias a antibióticos, antisépticos y desinfectantes una manifestación de los mecanismos de supervivencia y adaptación. *Colombia Médica Vol. 38(38)*, 149-158. Recuperado el 24 de febrero de 2016, de <http://132.248.9.34/hevila/Colombiamedica/2007/vol38/no2/7.pdf>
3. Campion, R. D., & Tec, J. C. (2004). Secuenciación de Ácidos Nucleicos. Recuperado el 03 de febrero de 2017, de [http://www.ibt.unam.mx/computo/pdfs/met/secuenciacion\\_acidos\\_nucleicos.pdf](http://www.ibt.unam.mx/computo/pdfs/met/secuenciacion_acidos_nucleicos.pdf)
4. Chen, J., & Xiao, P. (2012). Advance in Sequence Chemistry of High-Throughput DNA Sequencing by Synthesis. *Hans* (2), 1-6. doi:dx.doi.org/10.12677/bp.2012.21001
5. Cleland, W. W. (1963). Dithiothreitol, a New Protective Reagent for SH Groups. *Department of Biochemisfty, University of Wisconsin, Madison*, 480-482. doi:10.1021/bi00892a002
6. Corporation, L. T. (2014). Documentos, Manuales y Protocolos: User Guide: Ion Xpress Plus gDNA Fragment Library Preparation. Recuperado el 29 de agosto de 2016, de [https://tools.thermofisher.com/content/sfs/manuals/MAN0009847\\_IonXpressPlus\\_gDNA\\_FragLibraryPrep\\_UG.pdf](https://tools.thermofisher.com/content/sfs/manuals/MAN0009847_IonXpressPlus_gDNA_FragLibraryPrep_UG.pdf)
7. Crim, S., Griffin, P., Tauxe, R., Marder, E., Gillis, D., Cronquist, A., . . . Henao, O. (2015). Preliminary Incidence and Trends of Infection With Pathogens Transmitted Commonly Through Food - Foodborne Diseases Active Surveillance Network, 10 U.S.

- sites, 2006-2014. *Morb Mortal Wkly Rep.*, 64(18), 495-9. Recuperado el 21 de octubre de 2016, de <https://www.ncbi.nlm.nih.gov/pubmed/25974634>
8. Croxen, M., Law, R., Scholz, R., Keeney, K., Wlodarska, M., & Finlay, B. (2013). Recent Advances in Understanding Enteric Pathogenic *Escherichia coli*. *Clinical Microbiology Reviews*, 26(4), 822-880. doi:10.1128/CMR.00022-13
  9. de Lorenzo, V., Eltis, L., Kessler, B., & Timmis, K. N. (1993). Analysis of *Pseudomonas* Gene Products Using lacP/P<sub>trp</sub>-lac Plasmids and Transposons That Confer Conditional Phenotypes. *Elsevier Science Publishers*, 123, 17-24. doi:10.1016/0378-1119(93)90533-9
  10. Eli, E., & Levanon, E. (2003). Human Housekeeping Genes are Compact. *Trends in Genetics*, 19(7), 362-365.
  11. Elika. (2013). *Salmonella*. *Fundacion Vasca Para la Seguridad Agroalimentaria*. Recuperado el 24 de junio de 2016, de [http://www.elika.net/datos/pdfs\\_agrupados/Documento82/1.Salmonella.pdf](http://www.elika.net/datos/pdfs_agrupados/Documento82/1.Salmonella.pdf)
  12. Ewing, B., & Green, P. (1998). Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. *Genome Research*(8), 186-194. doi:10.1101/gr.8.3.186
  13. Fabrega, A., & Villa, J. (2013). *Salmonella enterica* Serovar Typhimurium Skills To Succeed in the Host: Virulence and Regulation. *CMR*, 308-341. doi:10.1128/CMR.00066-12
  14. FAO. (2011). Boletín de Enfermedades Transfronterizas de los Animales. *EMPRES*, 20-26. Recuperado el 03 de noviembre de 2016, de <http://www.fao.org/3/a-i2530s.pdf>
  15. Foodsafety. (2016). Bacteria and Viruses. Recuperado el 21 de octubre de 2016, de <https://www.foodsafety.gov/poisoning/causes/bacteriaviruses/>
  16. Gonzalez Flores, T., & Rojas Herrera, R. A. (2005). Enfermedades Transmitidas por Alimentos y PCR: Prevención y Diagnóstico. *Salud Publica de México*, 47(5), 388-390. Recuperado el 13 de febrero de 2016, de <http://www.scielosp.org/pdf/spm/v47n5/28385.pdf>

17. Greenfield, N. (2017). One Codex Database. Recuperado el 08 de junio de 2017, de <https://docs.onecodex.com/v1/docs/one-codex-database>
18. Healthcare, G. (2006). Product Web Protocol. Recuperado el 2016 de septiembre de 12, de ilustra GenomiPhi V2 DNA Amplification Kit: [https://www.gelifesciences.com/gehcls\\_images/GELS/Related%20Content/Files/1314774443672/litdoc25660030\\_20161014112135.pdf](https://www.gelifesciences.com/gehcls_images/GELS/Related%20Content/Files/1314774443672/litdoc25660030_20161014112135.pdf)
19. Herveg, J.-P., & Barcia Macay, M. (abril de 2006). Secuenciación Automática del ADN. Recuperado el 17 de 10 de 2017, de <https://ingenieriagen.weebly.com/la-secuenciacion-acuten-del-adn.html>
20. Hill, J. (2017). CLC Genomics Workbench 10. Recuperado el 22 de junio de 2017, de Release Highlights: <https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/version-10-available/>
21. Holley, R. W., Everett, G. A., Madison, J. T., & Zamir, A. (1965). Nucleotide Sequences in the Yeast Alanine Transfer Ribonucleic Acid. *The Journal of Biological Chemistry*, 240(5), 2122-2128. Recuperado el 17 de marzo de 2017, de <http://www.jbc.org/content/240/5/2122.full.pdf>
22. Hongchen , J., Hailiang, D., Gengxin , Z., Bingsong , Y., Leah R. , C., & Matthew W. , F. (2006). Microbial Diversity in Water and Sediment of Lake Chaka, an Athalassohaline Lake in Northwestern China. *Applied and Environmental Microbiology*, 3832–3845. doi:10.1128/AEM.02869-05
23. Kim, H. S., Byun, S. H., & Lee, B. M. (2005). Effects of Chemical Carcinogens and Physicochemical Factors on the UV Spectrophotometric Determination of DNA. *Journal of Toxicology and Environmental Health*, 2081-2095. doi:10.1080/15287390500182503
24. Lezin, G., Kosaka, Y., Yost, H. J., Kuehn, M. R., & Brunelli, L. (2011). A One-Step Miniprep for the Isolation of Plasmid DNA and Lambda Phage Particles. *Plos one*, 6(8), 1-9. doi:10.1371/journal.pone.0023457

25. LifeTechnologies. (2012). Specification Sheet - Ion Personal Genome Machine™ (PGM™) Sequencer. Recuperado el 03 de febrero de 2017, de [https://www3.appliedbiosystems.com/cms/groups/applied\\_markets\\_marketing/documents/generaldocuments/cms\\_094139.pdf](https://www3.appliedbiosystems.com/cms/groups/applied_markets_marketing/documents/generaldocuments/cms_094139.pdf)
26. Lin-Hui, S., & Cheng-Hsun, C. (2007). *Salmonella*: Clinical Importance and Evolution of Nomenclature. *Chang Gung Med Journal*, 30(3), 210-219. Recuperado el 21 de noviembre de 2016, de <http://memo.cgu.edu.tw/cgmj/3003/300302.pdf>
27. Moreno, R. B. (2010). Las Tres Generaciones de la Secuenciación. *¿Cómo funciona?*, 3(128), 27-28. Recuperado el 03 de febrero de 2017
28. Murray, P., Rosenthal, K., & Pfaller, M. (2014). Microbiología Médica. 216-219. Recuperado el 24 de octubre de 2016
29. National Center for Biotechnology Information, N. (2017). Frequently Asked Questions. Recuperado el 22 de junio de 2017, de What is the Expect (E) Value?: [https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE\\_TYPE=BlastDocs&DOC\\_TYPE=FAQ#expect](https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=FAQ#expect)
30. Noordhout, C. M., Devleeschauwer, B., Angulo, F. J., Verbeke, G., Hagasma, J., Kirk, M., & Havelaar, A. (2014). The Global Burden of Listeriosis: A Systematic Review and Meta-analysis. *The Lancet, Infectious Diseases*, 14(11), 1073-1082. doi:[http://dx.doi.org/10.1016/S1473-3099\(14\)70870-9](http://dx.doi.org/10.1016/S1473-3099(14)70870-9)
31. Novick, R. P. (2002). Plasmids. *Encyclopedia of Life Sciences*, 1-8. doi:10.1038/npg.els.0001490
32. Rodicio, M. d., & Mendoza, M. d. (2003). Identificación Bacteriana Mediante Secuenciación del ARNr 16S: Fundamento, Metodología y Aplicaciones en Microbiología Clínica. *Enfermedades Infecciosas Microbiología Clínica*, 22(4), 238-245. Recuperado el 23 de enero de 2017

33. Rodríguez Tarduchy, G., & Santiago Martínez, M. C. (2012). Secuenciación Automática de ADN. Recuperado el 17 de 10 de 2017, de Pirosecuenciación: <http://www2.iib.uam.es/seq/tecnicas/piro.htm>
34. Sandman, K., Pereira, S. L., & Reeve, J. N. (1998). Diversity of Prokaryotic Chromosomal Proteins and the Origin of the Nucleosome. *Cell and Molecular Life Sciences*(54), 1350–1364. Recuperado el 19 de enero de 2017
35. Sanger, F., Air, G., Barrel, B., Bro, N. B., Coulson, A., Fiddes, J., . . . Smith, M. (1977). Nucleotide sequence of bacteriophage Phi X174 DNA. *Nature*, 265, 687-695. doi:10.1038/265687a0
36. Sarich, V. M., & Wilson, A. C. (1967). Immunological Time Scale for Hominid Evolution. *Science*, 158, 1200-1203. Recuperado el 23 de enero de 2017, de <http://science.sciencemag.org/content/158/3805/1200.full.pdf+html>
37. Shintani, M., Sanchez, Z., & Kimbara, K. (2015). Genomics of Microbial Plasmids: Classification and Identification Based on Replication and Transfer Systems and Host Taxonomy. *Frontiers on Microbiology*, 1-16. doi:10.3389/fmicb.2015.00242
38. SINAVE. (2017). Boletín epidemiológico 2017. Recuperado el 27 de julio de 2017, de <https://www.gob.mx/cms/uploads/attachment/file/241912/sem28.pdf>
39. Smajs, D., Strouhal, M., Jkova, P. M., Cejkova, D., Cursino, L., Chartone-Souza, E., Nascimento, A. (2008). Complete Sequence of Low-copy-number Plasmid MccC7-H22 of Probiotic *Escherichia coli* H22 and the Prevalence of mcc Genes Among Human *E. coli*. *Science Direct* (59), 1-10. doi:10.1016/j.plasmid.2007.08.002
40. Sue Ling, C., Wen Tsuan, C., & Ten Tsao, W. (1992). High Copy Number of the pUC Plasmid Results from a Rom/Rop-suppressible Point Mutation In RNA II. *Molecular Microbiology*, 6(22), 3385-3393. doi:10.1111/j.1365-2958.1992.tb02206.x
41. Thanbichler, M., Wang, S. C., & Shapiro, a. L. (2005). The Bacterial Nucleoid: A Highly Organized and Dynamic Structure. *Journal of Cellular Biochemistry*(96), 506–521. doi:10.1002/jcb.20519

42. The University of Arizona. (2017). Interpreting Nanodrop (Spectrophotometric) Results. Recuperado el 22 de junio de 2017, de <http://www.u.arizona.edu/~gwatts/azcc/InterpretingSpec.pdf>
43. Thomas, C. M., & Summers, D. (2008). Bacterial Plasmids. *Encyclopedia of Life Sciences*, 1-9. doi:10.1002/9780470015902.a0000468.pub2
44. Toro, E., & Shapiro, L. (2010). Bacterial Chromosome Organization and Segregation. *Cold Spring Harbor Perspectives in Biology*, 1-16. doi:10.1101/cshperspect.a000349
45. Vargas, E. R., Aldapa, C. G., Vitela, M. T., Lopez, A. V., Martínez, A. G., & Rosas, J. C. (2015). Presence and Correlation of Some Enteric Indicator Bacteria, Diarrheagenic *Escherichia coli* Pathotypes, and *Salmonella* Serotypes in Alfalfa Sprouts from Local Retail Markets in Pachuca, Mexico. *Journal of Food Protection*, 609-614. Recuperado el 15 de noviembre de 2016
46. Vázquez, M. F. (2010). Entorno Genético de  $\beta$ -lactamasas de Espectro Extendido en Enterobacterias. Salamanca, España: Departamento de Medicina Preventiva, Salud Pública y Microbiología Médica. Recuperado el 02 de febrero de 2017
47. Velázquez, L. A., Martínez, M. A., & Romero, A. C. (2010). Extracción y Purificación de ADN. *Herramientas Moleculares Aplicadas a la Ecología*. Recuperado el 01 de diciembre de 2016, de <http://www2.inecc.gob.mx/publicaciones/libros/710/extraccion.pdf>
48. WHO. (2013). *Salmonella* (no Tifoidea). Recuperado el 27 de noviembre de 2016, de <http://www.who.int/mediacentre/factsheets/fs139/es/>
49. WHO. (2015). *Escherichia coli* Infections. Recuperado el 03 de noviembre de 2016, de [http://www.who.int/topics/escherichia\\_coli\\_infections/en/](http://www.who.int/topics/escherichia_coli_infections/en/)
50. WHO. (2015). Foodborne Diseases. Recuperado el 12 de octubre de 2016, de [http://www.who.int/topics/foodborne\\_diseases/en/](http://www.who.int/topics/foodborne_diseases/en/)

51. WHO. (2015). Inocuidad de los Alimentos. Recuperado el 17 de marzo de 2017, de La Inocuidad de los Alimentos: Una Prioridad de Salud Pública:  
<http://www.who.int/mediacentre/factsheets/fs399/es/>
52. WHO. (2015). WHO Estimates of the Global Burden of Foodborne Diseases. Recuperado el 19 de septiembre de 2016, de  
[http://apps.who.int/iris/bitstream/10665/199350/1/9789241565165\\_eng.pdf?ua=1](http://apps.who.int/iris/bitstream/10665/199350/1/9789241565165_eng.pdf?ua=1)
53. WHO. (2015). WHO's First Ever Global Estimates of Foodborne Diseases Find Children Under 5 Account for Almost One Third of Deaths. Recuperado el 26 de septiembre de 2016, de  
<http://www.who.int/mediacentre/news/releases/2015/foodborne-disease-estimates/en/>
54. WHO. (2016). *E. coli*. Recuperado el 07 de noviembre de 2016, de Centro de Prensa:  
<http://www.who.int/mediacentre/factsheets/fs125/en/>
55. Zepeda García, O. (2010). Bioinformatica. Recuperado el 26 de enero de 2017, de ¿Que es la bioinformatica?: <http://www.solociencia.com/biologia/bioinformatica.htm>
56. Zhu, J., He, F., Hu, S., & Yu, J. (2008). On the Nature of Human Housekeeping Genes. *Trends in Genetics*, 24(10), 481-484.
57. Zuckerkandl, E., & Pauling, L. (1965). Molecules as Documents of Evolutionary History. *J. Theoret. Biol.*, 8, 357-366. Recuperado el 23 de enero de 2017