



Universidad Nacional Autónoma de México

Facultad de Estudios Superiores Zaragoza

ESTANDARIZACIÓN DEL PROCESO DE ELABORACIÓN
DE GENOTECAS BACTERIANAS PARA LA PLATAFORMA
DE SECUENCIACIÓN MASIVA ION TORRENT PGM

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
QUÍMICA FARMACÉUTICA BIÓLOGA

P R E S E N T A:

VERONICA AIDE OLVERA AVILA

DIRECTOR DE TESIS:

M. en C. ABRAHAM IZCOATL ACATZI SILVA

ASESOR INTERNO:

M. en C. ÁNGEL GARCÍA SÁNCHEZ



MÉXICO, D.F. NOVIEMBRE 2015



Universidad Nacional
Autónoma de México



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.

El presente trabajo se realizó en el Centro Nacional de Referencia en Detección de Organismos Genéticamente Modificados (CNRDOGM) en la Subdirección de Secuenciación y Bioinformática, del Servicio Nacional de Sanidad Inocuidad y Calidad Agroalimentaria (SENASICA).

ÍNDICE	
Abreviaturas	4
Resumen	6
I. INTRODUCCIÓN	7
II. MARCO TEÓRICO	9
2.1 Enfermedades transmitidas por alimentos (ETA).....	9
2.2 Identificación bacteriana.....	13
2.3 Historia de la Secuenciación de Nueva Generación.....	17
2.4 Aplicaciones de la Secuenciación Nueva Generación SNG.....	19
2.5 Plataforma de Secuenciación Masiva Ion Torrent.....	20
2.6 Flujo de trabajo de secuenciación en la plataforma Ion Torrent.....	21
2.7 Tipos de bibliotecas o genotecas en la Secuenciación Masiva.....	22
2.8 Métodos de fragmentación del ADN.....	25
2.9 Características generales de las bacterias utilizadas en el presente estudio	26
2.10 Antecedentes particulares del presente estudio.....	29
III. PLANTEAMIENTO DEL PROBLEMA	
3.1 Planteamiento del Problema.....	31
IV. HIPÓTESIS	32
V. OBJETIVOS	
5.1 Objetivos generales.....	33
5.2 Objetivos particulares.....	33
VI MATERIALES Y MÉTODOS	

6.1 Estrategia experimental.....	34
6.2 Diagrama de flujo.....	35
6.3 Selección de muestras.....	36
6.4 Extracción de ADN.....	37
6.5 Fragmentación de ADN.....	38
6.5.1 Fragmentación enzimática.....	38
6.5.2 Fragmentación por Sonicación.....	39
6.6 Purificación de fragmentos de ADN.....	40
6.7 Ligación de adaptadores y código de barras	40
6.8 Purificación de genotecas bacterianas.....	41
6.9 Selección de tamaño de genotecas bacterianas.....	41
6.10 Purificación de genotecas bacterianas después de la selección tamaño	42
6.11 Amplificación de la genoteca bacteriana.....	43
6.12 Purificación de la genoteca amplificada.....	43
6.13 Evaluación del tamaño de genotecas bacterianas.....	44
6.14 Cuantificación de las genotecas por qPCR.....	45
VII. RESULTADOS.....	46
VIII. DISCUSIÓN.....	63
IX. CONCLUSIONES.....	68
REFERENCIAS.....	69

ABREVIATURAS

ADN	Ácido Desoxirribonucleico
ARN r	ácido ribonucleico ribosomal
CDC	Centros para el Control y Prevención de Enfermedades
CNDROGM	Centro Nacional en Detección de Referencia de Organismos Genéticamente Modificados
CV%	Coefficiente de variación
COFEPRIS	la Comisión Federal para la Protección contra Riesgos Sanitarios
DGE	Dirección General de Epidemiología de la Secretaría de Salud
DGIAAP	Dirección General de Inocuidad Agroalimentaria, Acuícola y Pesquera
ETA	Enfermedades transmitidas por alimentos
H	Horas
qPCR	Reacción en Cadena de Polimerasa en tiempo real
MPC	Concentrador de Perlas Magnéticas
Min	Minutos
MID	Molecular identifier
MLEE	Multilocus Enzyme Electrophoresis
MLST	Multilocus sequence type
μL	Microlitos
SNG	Next Generation Sequencing
LCDDOP	Laboratorio de Diagnóstico para la Detección de Organismos Patógenos

Ng	Nano gramos
OMS	Organización Mundial de la Salud
pb	Pares de bases
SAGARPA	Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación
SENASICA	Servicio Nacional de Sanidad Inocuidad y Calidad Agroalimentaria
SSA	Secretaría de Salud
SNP	Single Nucleotide Polymorphism
SUAVE	Sistema Único Automatizado de Vigilancia Epidemiológica
SUIVE	Sistema Único de Información para la Vigilancia Epidemiológica

RESUMEN

En los últimos años el uso de la Secuenciación Nueva Generación (SNG) ha aumentado debido al gran número de aplicaciones, entre las que destaca la capacidad de secuenciar el genoma completo de diferentes organismos en menor tiempo a bajo costo; en comparación con tecnologías de primera generación, la secuenciación de ADN es la forma más completa de obtener la información sobre el genoma de cualquier organismo.

En ese mismo sentido el Centro Nacional en Detección de Referencia de Organismos Genéticamente Modificados (CNDROGM), ha empleado el uso de la técnica de Secuenciación de Nueva Generación para la identificación y tipificación de bacterias de interés agroalimentario. Sin embargo, para la secuenciación de estos microorganismos, una de las etapas cruciales es la elaboración de genotecas, donde se asegura su tamaño y concentración, para evitar los errores de cobertura y longitud de los fragmentos de ADN.

Es por ello que en este trabajo se establecieron las condiciones de fragmentación de ADN, por el método de sonicación que fue de 300 ciclos y 80 s y para el método enzimático el tiempo óptimo fue 7.5 min y para la selección de tamaño de ADN se estableció un “target” 470 pb una vez establecidos los parámetros se elaboraron diez genotecas a partir del ADN de la bacteria *Salmonella Muenchen*, a los resultados se le aplicó estadística descriptiva paramétrica el coeficiente de variación que se obtuvo fue de 0.43 % menor de 15 % por tanto se concluyó que el proceso de elaboración de genotecas es repetible y reproducible.

I. INTRODUCCIÓN

El Servicio Nacional de Sanidad Inocuidad y Calidad Agroalimentaria (SENASICA) es un órgano desconcentrado de la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA), que se ocupa de condiciones y prácticas que preservan la calidad de los alimentos en México para prevenir la contaminación y enfermedades transmitidas por el consumo de alimentos, para reducir los riesgos inherentes en materia agrícola, pecuaria, acuícola y pesquera, en beneficio de los productores y consumidores. Entre las actividades de la Dirección General de Inocuidad Agroalimentaria, Acuícola y Pesquera (DGIAAP) de SENASICA, está el identificar la presencia de contaminantes en los alimentos de origen agrícola, pecuario, acuícola y pesquero.

Las enfermedades transmitidas por alimentos (ETA) constituyen el 12% de los casos que registra el Sistema Nacional de Vigilancia Epidemiológica, lo que las convierte en la segunda causa de notificación de enfermedades en México. Es por ello que se debe trabajar en la identificación de dichos agentes contaminantes.

En este mismo sentido, el Centro Nacional de Referencia en Detección de Organismos Genéticamente Modificados (CNRDOGM) del SENASICA ha implementado la técnica de secuenciación masiva para la identificación y tipificación de microorganismos patógenos de interés agroalimentario que son causantes de ETA, las cuales son un problema importante de salud pública, con el fin de coadyuvar a la inocuidad alimentaria.

La secuenciación masiva consta de tres etapas: elaboración de la genoteca, emulsión en PCR y corrida de secuenciación. Dentro de esta técnica hay dos químicas, la de 200 pares de bases y la de reciente introducción que es más amplia, de 400 pares de bases (pb). En la plataforma de secuenciación de Ion

Torrent PGM con química de 400 pb, no se cuenta con los parámetros óptimos para el proceso de la elaboración de la genotecas bacterianas, por su reciente introducción en el Centro Nacional de Referencia en Detección de Organismos Genéticamente Modificados.

Por lo anterior, en este trabajo se realizó una estandarización de la elaboración de genotecas, para la implementación de esta nueva metodología de secuenciación Ion Torrent PGM con la química de 400 pb con el fin de realizar la identificación de bacterias patógenas en el campo agroalimentario, y con ello coadyuvar a garantizar la inocuidad alimentaria en la producción primaria con el fin de prevenir enfermedades transmitidas por los alimentos.

II MARCO TEÓRICO

2.1 Enfermedades transmitidas por Alimentos

La Organización Mundial de la Salud (OMS) ha definido a las Enfermedades Transmitidas por Alimentos (ETA), como enfermedades de carácter infeccioso o tóxico que son causadas por bacterias, virus, parásitos o sustancias químicas que penetran en el organismo a través del agua o los alimentos contaminados ⁽¹⁾.

Las ETA afectan principalmente al tracto gastrointestinal, sus síntomas más comunes son diarreas y vómitos, pero también se pueden presentar otros como choque séptico, hepatitis, cefalea, fiebre, visión doble, todo depende del microorganismo causante de la ETA ^(2, 3, 4, 5).

Las enfermedades transmitidas por alimentos pueden ser de dos tipos: infecciones alimentarias o intoxicaciones alimentarias. En las infecciosas, la enfermedad está causada por microorganismos patógenos que se reproducen en el interior del organismo, como virus, bacterias o parásitos; mientras que en el caso de intoxicaciones, son provocadas por la ingesta de toxinas producidas por microorganismos o sustancias químicas, presentes de forma natural en el alimento o añadidas de manera artificial en cualquier momento, desde la producción del alimento hasta su consumo ^(6, 7). En el Cuadro 1 se muestran algunas de las principales bacterias causantes de ETA.

Cuadro 2.1 Bacterias causantes de Enfermedades Transmitidas por Alimentos ^(5, 6).

Microorganismo	Periodo de incubación	Fuente	Síntomas
<i>Vibrio chólera</i>	6 horas a 5 días	Pescados y mariscos crudos o que no estén bien cocidos, u otros alimentos y agua contaminados.	Diarrea acuosa inicial, seguida de disentería, fiebre y vómito.
<i>E. coli enterotoxigénica</i>	6 a 56 horas	Ensalada, queso, carnes, agua	Fiebre (raro), diarrea acuosa, cólico bajo, náusea, vómito.
<i>Salmonella sp</i>	6-72 horas varía hasta 10 días.	Carne de vaca, pollo, huevos, lácteos.	Fiebre, diarrea acuosa o disentería, cefalea, dolor abdominal, vómito.
<i>Shigella sp</i>	1-4 días.	Ensalada de patatas o huevo, lechuga, verduras crudas.	Fiebre, cefalea, malestar general, anorexia, vómito(raro), diarrea acuosa (al inicio), disentería, cólico.
<i>Campylobacter jejuni</i>	36 a 120 horas	Alimentos (leche no pasteurizada, carne agua contaminada),	Inicio abrupto, cólico abdominal, diarrea acuosa o disenteriforme.
<i>Staphylococcus aureus</i>	1 a 6 horas	Jamón, pollo, ensalada de patatas	Náuseas, calambres estomacales, vómitos y diarrea.
<i>Vibrio parahemolyticus</i>	4 a 96 horas	Moluscos, Crustáceos	Diarrea, calambres estomacales, náuseas, vómitos, dolor de cabeza, fiebre y escalofríos.
<i>Clostridium perfringens</i>	8 a 16 horas	Carne de vaca, pollo, legumbres, salsas	Cólicos abdominales, diarrea
<i>Yersinia enterocolítica</i>	3-7 días	Carne de res, pescados y mariscos crudos, productos lácteos, productos frescos y agua no tratada	Fiebre, dolor abdominal, diarrea acuosa, sangre en heces, vómito, faringitis exudativa, adenitis cervical.
<i>Listeria monocytogenes</i>	9 a 48 horas	Alimentos refrigerados, listos para consumir (carne de res, pollo, pescados y mariscos, y lácteos.	Fiebre, dolor de cabeza, cansancio, dolores musculares, náuseas, vómitos, diarrea, meningitis y abortos espontáneos

En México, según la Dirección General de Epidemiología de la Secretaría de Salud, la intoxicación alimentaria de origen bacteriano ocupa el cuarto lugar de las enfermedades diarreicas ^(5, 8, 9, 10). Además de las bacterias, microorganismos como virus y parásitos también son causales de ETA. Las infecciones gastrointestinales por rotavirus son una de las causas principales para la hospitalización en los menores de cinco años de edad relacionados con las gastroenteritis agudas.

En el Cuadro 2.2 se muestran el virus y los parásitos que causan ETA. Los alimentos que contienen bacterias, virus, parásitos o sustancias químicas nocivas causan en total, más de 200 enfermedades que van desde la diarrea hasta el cáncer.¹¹

Cuadro 2.2 Etiología de Enfermedades Transmitidas por Alimentos ⁽¹¹⁾.

Patógeno	Mecanismo de transmisión	Síntomas
<i>Rotavirus</i>	Está extendido por todo el mundo y la transmisión es por contaminación fecal-oral o posiblemente por gotículas de transmisión aérea.	Inicio abrupto, diarrea líquida con vómitos y fiebre baja (menos de 38°C). Puede ser asintomática o cursar con diarrea aguda y deshidratación que puede ser mortal.
<i>Giardia duodenales</i>	Alimentos contaminados, agua y transmisión fecal –oral.	Diarrea aguda o persistente, a veces mala absorción con esteatorrea, dolor abdominal y meteorismo. Muchas de las veces es asintomática.
<i>Entamoeba histolytica</i>	Alimentos y agua contaminada con quistes del protozooario y transmisión fecal–oral.	Diarrea persistente hasta la disentería fulminante con absceso hepático.

En el año 2011 en México se reportaron 44 467 casos por intoxicación alimentaria bacteriana, que lo ubicó en el decimoctavo lugar de las veinte principales causas de morbilidad en el país. En la Figura 2.1 se muestra un mapa de la incidencia de intoxicación alimentaria bacteriana tomado de SUAVE/DGE/SALUD.SUIVE/INEGI 2010 ⁽¹²⁾.

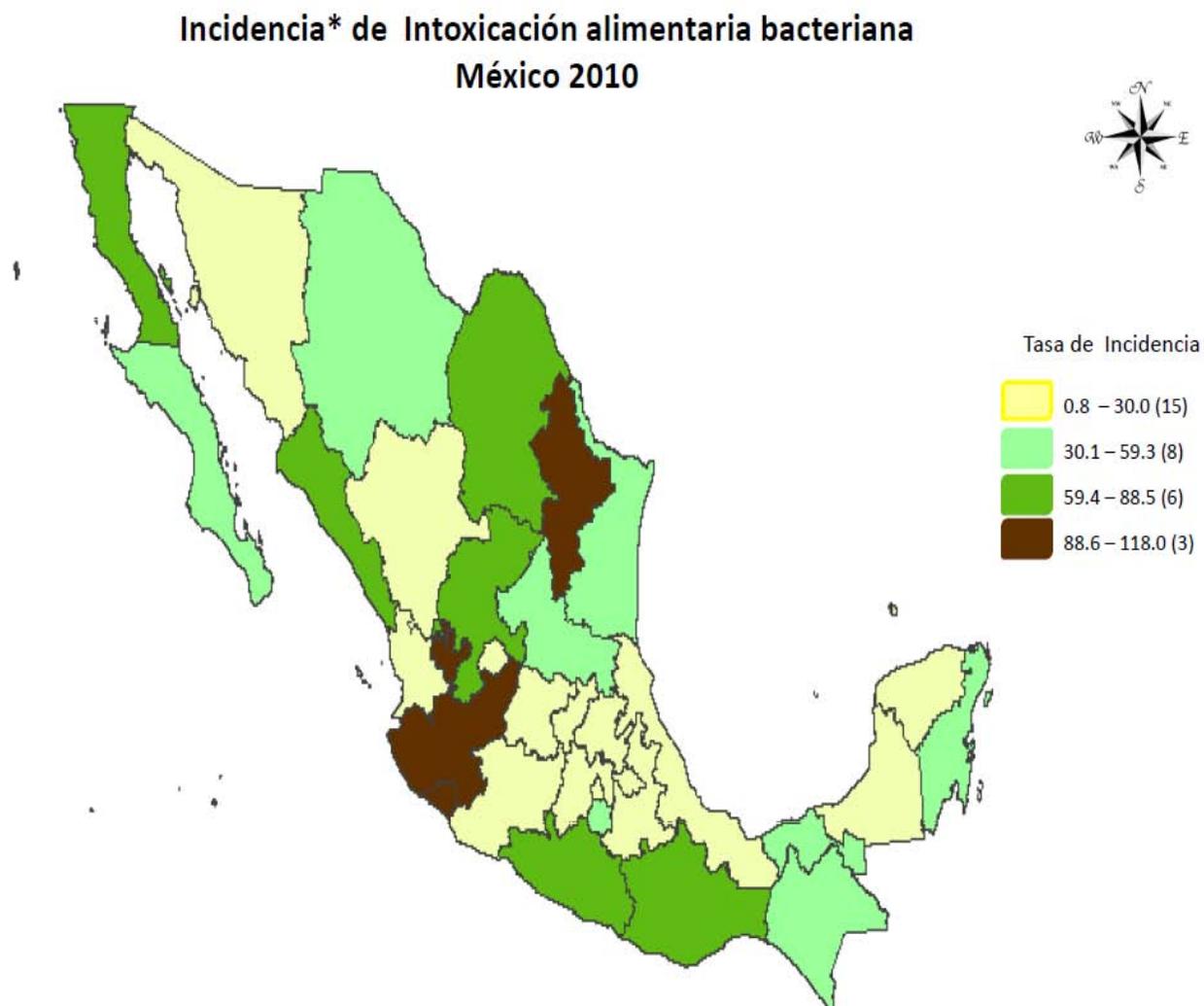


Figura 2.1. Incidencia de Intoxicación alimentaria bacteriana ⁽¹³⁾.

En México el interés por el tema de inocuidad de los alimentos ha crecido de manera importante, en particular por el impacto que tiene sobre la salud de la población. Por ende, en el país existen dos agencias encargadas de la inocuidad de los alimentos frescos y procesados. Dichas agencias son: la Secretaría de Salud (SSA), a través de la Comisión Federal para la Protección contra Riesgos Sanitarios (COFEPRIS) y la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA) que se encarga de los aspectos de inocuidad a través del Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria (SENASICA) ^(9,10).

El gobierno mexicano, al comprender la necesidad de prevenir la contaminación de los alimentos, crea dentro del Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria (SENASICA) la Dirección General de Inocuidad Agroalimentaria, Acuícola y Pesquera, que desarrolla y ejecuta esquemas de aplicación voluntaria sobre temas de inocuidad para la parte primaria de la industria, que promueve la aplicación y certificación de los sistemas de reducción de riesgos de contaminación de los alimentos^(9,10).

2.2 Identificación de las bacterias

La identificación de las bacterias realizada por medio de métodos convencionales se basa en las características fenotípicas de las bacterias, como su morfología, propiedades bioquímicas y metabólicas; sin embargo estos métodos tienen algunas limitaciones, porque no todas las cepas de una misma especie muestran

una característica específica, por ejemplo, una misma cepa puede generar diferentes resultados en ensayos repetidos, por lo cual se pueden utilizar los métodos genotípicos de identificación bacteriana como procedimientos alternativos (14, 15).

Recientemente se han desarrollado nuevas herramientas moleculares como la PCR, PCR tipo multiplex y la secuenciación, para la identificación de bacterias, ya que con los métodos anteriores se presentaban dificultades para diferenciar fenotípicamente las bacterias, además de problemas en el aislamiento, crecimiento lento o uso de medios de cultivo *in vitro* complejos, la baja actividad en las pruebas bioquímicas, ausencia o baja especificidad de técnicas serológicas.¹⁴ De forma adicional han surgido plataformas de identificación de patógenos que modifican o sustituyen la secuenciación tradicional, como sucede con la secuenciación de genoma completo o secuenciación de nueva generación⁽¹⁴⁾.

Los métodos genotípicos se han desarrollado para la tipificación de las cepas bacterianas de difícil identificación fenotípica, así como para determinar la fuente de la infección y reconocer las cepas particularmente virulentas en un determinado brote. Estos métodos se basan en técnicas de identificación molecular de bacterias, mediante el análisis del ARN ribosomal 16S (ARNr 16S) que es un polirribunucleotido de aproximadamente 1500 nucleótidos (nt), por medio de la amplificación y secuenciación de éste o fragmentos de genes^(14, 15, 16, 17).

La secuenciación del ARNr 16S es utilizada frecuentemente como un marcador universal para identificar una bacteria patógena, sin embargo su nivel de resolución es bajo, es decir, solo permite identificar género y especie⁽¹⁸⁾. Otra

metodología de identificación bacterias es el *Multilocus Enzyme Electrophoresis* (MLEE) que es un método genético indirecto que analiza la movilidad electroforética de los genes en geles de almidón o poliacrilamida. Las variaciones observadas en estas movilidades corresponden con variaciones en el *locus* o gen codificante. Por lo tanto, cada variante se define como “variante alélica” y los diferentes alelos de cada uno de los genes conforman el perfil alélico que a su vez define el tipo electroforético, aun en el caso de encontrar tan sólo cuatro alelos por *locus*, o lo que es lo mismo, cuatro variantes en cada gen analizado ⁽¹⁹⁾.

Por otro lado, el acceso a la secuenciación del ADN permitió el desarrollo de la técnica *Multilocus Sequence Typing* (MLST) que se utiliza para la caracterización de aislados de especies bacterianas con uso de las secuencias de fragmentos internos de siete genes de expresión constitutiva o también llamados *house-keeping*. Esta técnica utiliza fragmentos de aproximadamente 450-500 pares de bases (pb) de los *house-keeping* de cada uno de los genes secuenciados dentro de una especie bacteriana, que se asignan como alelos distintos para cada aislamiento, los alelos en cada uno de los siete *loci* definen el perfil alélico o tipo de secuencia (ST) ^(19,20).

Una de las limitaciones de los análisis de tipo MLST es que existen algunas especies de bacterias donde los genes *house-keeping* o constitutivos presentan una muy baja diversidad nucleotídica (especies conocidas como monomórficas, como por ejemplo *Salmonella* y *Bacillus anthracis*), o incluso existen otras bacterias en las que los genes constitutivos o *house-keeping* son tan uniformes en todos los aislados, es decir tienen la misma secuencia y parecen ser

el mismo tipo de bacteria.

Además la técnica MLST presenta el problema de baja cobertura del genoma completo, es decir, solo se secuencian pequeñas fracciones del genoma, lo cual no permite conocer la diversidad nucleotídica del genoma completo. Por lo anterior la secuenciación de genomas completos se utiliza cada vez con mayor frecuencia para la identificación y caracterización de las bacterias patógenas.

En la Figura 2.2 se muestra el esquema de diferentes niveles de resolución dentro de una población bacteriana por los diversos métodos de identificación antes mencionados, donde se nota una tendencia actual a la identificación de las bacterias por medio de polimorfismos de un solo nucleótido (SNPs), cuya resolución puede ser obtenida por medio de la secuenciación masiva⁽¹⁸⁾. La secuenciación de ADN, por lo tanto hoy se utiliza como una técnica para identificación de bacterias que permite obtener una serie ordenada de nucleótidos de un gen o genoma.

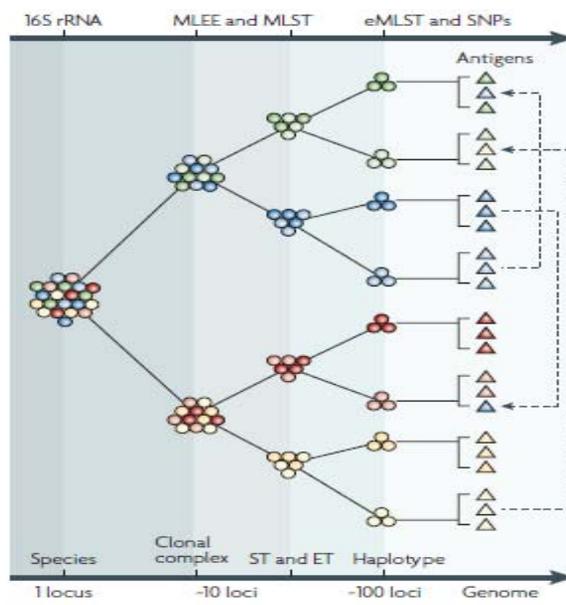


Figura 2.2 Representación esquemática de diferentes niveles de resolución, para la identificación bacteriana⁽¹⁸⁾.

2.3 Historia de la Secuenciación de Nueva Generación

Las primeras técnicas de secuenciación surgen en 1977 con Allan Maxam y Walter Gilbert, y Frederick Sanger, que se basan en la separación por electroforesis capilar de una sola reacción de PCR con nucleótidos dideoxi marcados con fluoróforos ⁽²¹⁾.

Aunque el método automatizado de Sanger había dado lugar a una serie de importantes logros, que incluían la finalización de la secuenciación del genoma humano, así como genomas animales y vegetales, las limitaciones del método mostraron una necesidad en la búsqueda de nuevas y mejores tecnologías para la secuenciación de un gran número de genomas.

A finales del siglo XX y principios del XXI, se han realizado esfuerzos hacia el desarrollo de nuevos métodos para sustituir la secuenciación tipo Sanger automatizado, que se considera como una Tecnología de “primera generación” ²³,

²⁴ La secuenciación masiva es una de esas innovaciones, que se conoce como tecnología de segunda generación o secuenciación de nueva generación o SNG de sus siglas en Inglés “*Next Generation Sequencing*” ^(22, 23).

La historia de la Secuenciación de Nueva Generación comenzó en el año 2000, cuando la empresa Lynx Therapeutics (EE.UU.) puso en marcha la primera plataforma de secuenciación, la cual posteriormente fue comprada por Illumina, que también es una compañía estadounidense creada en 1998.

Posteriormente la empresa *454 Life Sciences* comercializó la versión de su secuenciador masivo, basado en la pirosecuenciación y en los años 2005–2006 fue adquirida por Roche. Esta plataforma revolucionó la secuenciación del

ADN, por su capacidad de producir 20 Mpb (20 millones de bases) en corto tiempo, y en el año 2007 fue sustituido por el modelo GS FLX que tenía la capacidad de producir más de 100 M pb en cuatro horas⁽²²⁾.

Por otro lado la compañía Solexa, lanzó en el año 2005 el equipo Genome Analyzer (GA) que se basaba en secuenciación por síntesis; que fue comprada por Illumina, y para el año 2007 la plataforma GA II x era capaz de generar fragmentos de 50 pb⁽²²⁾.

Posteriormente, en el año de 2011 la empresa Life Technologies® lanzó el secuenciador Ion Torrent basado en tecnologías de semiconducción utiliza chips semiconductores. Actualmente está en proceso el desarrollo de tecnologías de tercera generación, las cuales, pretenden secuenciar una sola molécula de ADN muy larga, sin necesidad de realizar una amplificación clonal, como se realiza en las plataformas de segunda generación. En el Cuadro 2.3 se muestran algunas de las características de las principales plataformas de Secuenciación de Nueva Generación SNG⁽²²⁾.

Cuadro 2.3. Principales características de las plataformas de Secuenciación NuevaGeneración SNG ^(22, 24).

Plataforma de Secuenciación	Método de amplificación	Longitud de lectura (pb)	Tiempo de Corrida	Ventajas	Desventajas
454 (Roche)	Emulsión PCR	700	23 horas	Lecturas largas	Errores homopolimeros, caro
Illumina MiSeq	Puentes de PCR	150	27 horas	Fácil uso	Caro por base
Ion Torrent(Life Technologies)	Emulsión PCR	>100	2 horas	Tiempo de corrida corto, bajo costo de reactivos	Rendimiento medio
PacBio(Pacific Biosciences)	No hay amplificación	>1000	90 min	Lecturas largas, corto tiempo de corrida, secuencia una sola molécula	Tasa de error alta ,rendimiento bajo

h= hora, min= minutos, PCR= Reacción en cadena de la polimerasa

2.4 Aplicaciones de la Secuenciación Nueva Generación SNG

Algunos ejemplos de las principales aplicaciones de la SNG son: la resecuenciación, que se lleva a cabo cuando se dispone de un genoma de referencia con el cual se puede comparar al momento de realizar los análisis bioinformáticos; la secuenciación de ChiP “*ChiP secuencing*” que permite observar interacciones ADN-proteína; otra aplicación es la secuenciación de metilación “*Methyl-secuencing*” que se utiliza para medir el nivel de metilación en los genomas; una aplicación más es la, la secuenciación *De Novo*, que se aplica cuando no se conoce un genoma de referencia, y por último la metagenómica, que

es el estudio del contenido de los genomas microbianos en una muestra ambiental. Estas son algunas de las aplicaciones de SNG, sin embargo no son las únicas, sólo se enunciaron las de mayor relevancia ^(22, 25, 26, 27, 28).

La secuenciación masiva tiene el potencial de detectar todos los tipos de variación genómica en un único experimento, que incluye las variantes de nucleótido único o mutaciones puntuales, pequeñas inserciones y deleciones ^(23, 24). Uno de los beneficios que ofrece la secuenciación en SNG es la capacidad de producir una enorme cantidad de datos, con lo que se ha logrado una reducción importante en los costos, así como un aumento en el rendimiento y precisión ^(21, 22, 23).

2.5 Plataforma de Secuenciación Masiva Ion Torrent

La plataforma de secuenciación Ion Personal Genome Machine (PGM), utiliza tecnología semiconductor que detecta un cambio de voltaje, el cual se basa en la detección en tiempo real de la concentración de iones de hidrógeno, lanzados como un subproducto, cuando un nucleótido se incorpora en una cadena de ADN por la acción de la polimerasa. El Ion Torrent o PGM utiliza una matriz de alta densidad de pozos micro-mecanizada, para realizar el proceso de secuenciación de manera paralela y masiva.

Cada pocillo del chip semiconductor tiene una cadena de ADN única de cada biblioteca o genoteca y debajo de los pozos, hay un sensor de iones para detectar el cambio en la concentración de iones hidrógeno debido a la incorporación de nucleótidos durante la secuenciación. En la Figura 2.3 se muestra el fundamento de la secuenciación masiva por Ion Torrent ^(29, 30).

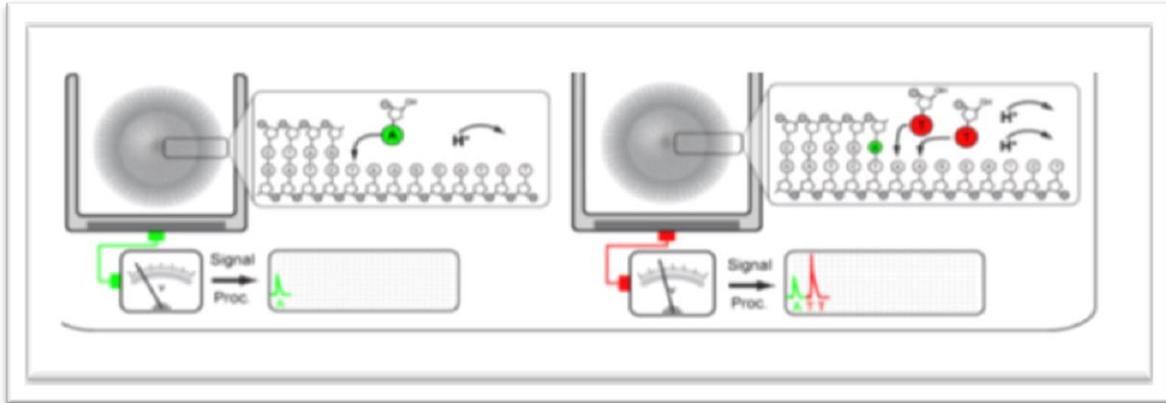


Figura 2.3 Fundamento de secuenciación masiva por Ion Torrent. Se muestra la incorporación de un nucleótido y la liberación del protón hidrógeno como un subproducto de la reacción, al detectar el cambio en el pH, PGM ⁽³¹⁾.

2.6 Flujo de trabajo de secuenciación en la plataforma Ion Torrent

El flujo de trabajo consta de tres etapas; la primera etapa de secuenciación Ion Torrent comienza con la preparación de la genoteca –que se produce mediante la fragmentación del ADN, a estos fragmentos se le ligan adaptadores de secuenciación, una vez ligados los adaptadores se convierten en genotecas- que posteriormente se amplifican por medio de una PCR de emulsión, es decir una reacción de amplificación por PCR dentro de una emulsión (PCR), donde una genoteca se une a una esfera y se amplifica millones de veces. Una vez concluido ese proceso, el siguiente paso es la corrida de secuenciación masiva ⁽³⁰⁾.

2.7 Tipos de bibliotecas o Genotecas en la Secuenciación Masiva

La genoteca de ADN es una colección aleatoria de fragmentos de ADN que en conjunto representan una parte o el genoma completo de un organismo ⁽³²⁾. El primer paso para la preparación de una Genoteca es la fragmentación del ADN, que se puede realizar de forma mecánica por medio de sonicación o con una nebulización con nitrógeno gaseoso y de forma enzimática, con endonucleasas.

Los fragmentos de ADN se convierten en genoteca mediante la ligación de adaptadores, que son secuencias de nucleótidos ya conocidas, que contienen un *primer* de secuenciación y un *primer* de amplificación y un “MID” (*Molecular identifier*) ó código de barras, que se ligan en cada uno de los extremos de los fragmentos del ADN.

Posteriormente, se realiza la selección de tamaño del material genético, por medio de una electroforesis en gel automatizada, y después se amplifica por medio de una PCR punto final, una vez terminada la amplificación, se cuantifica la genoteca por una PCR Tiempo real (qPCR) o por fluorimetría, en la Figura 2.4 se muestra el proceso para la elaboración de genotecas ⁽³⁰⁾.

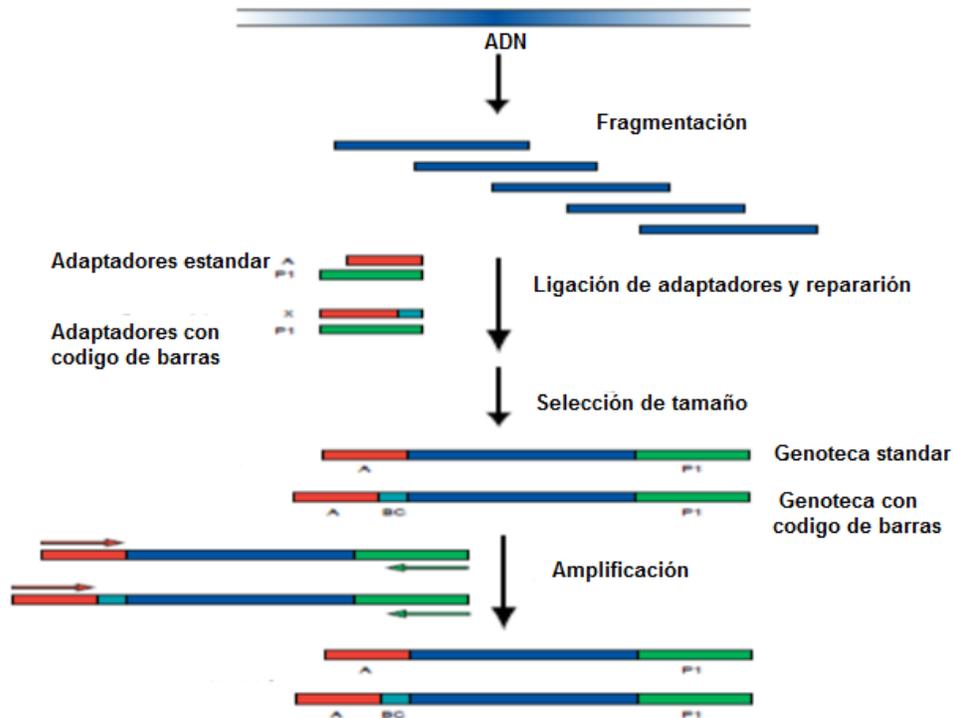


Figura 2. 4 Proceso de elaboración de genotecas. El primer paso es la fragmentación posteriormente la ligación de adaptadores, selección de tamaño y finalmente la amplificación ⁽³³⁾.

Existen tres tipos de genotecas de ADN que se pueden generar para la secuenciación de ADN, las cuales son: La genoteca *“single-end”*, en donde por cada fragmento de ADN se realiza una sola lectura, es decir, la secuenciación de esta genoteca se realiza en un sólo sentido, este tipo de genotecas son ideales para el análisis de sustituciones de un solo nucleótido o variaciones, y para identificar que son reales estas variaciones se necesita una cobertura mayor a 30x. El diseño de estas genotecas se muestra en la Figura 2.5 ^(25, 26).

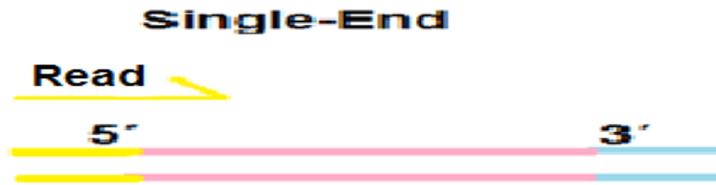


Figura 2.5 Genoteca “*single-end*”. La secuenciación de la genoteca es del extremo 5' al extremo 3' obteniéndose una sola secuencia, por genoteca.

La genoteca “*paired-end*”, que realiza las lecturas en ambas direcciones: sentido y antisentido de un fragmento de ADN o genoteca, en el cual se obtienen dos lecturas por cada segmento, lo que facilita la detección de reordenamientos genómicos, elementos repetitivos, así como la fusión de genes. En la Figura 6 se muestra la genoteca *paired-end* ⁽³⁴⁾.

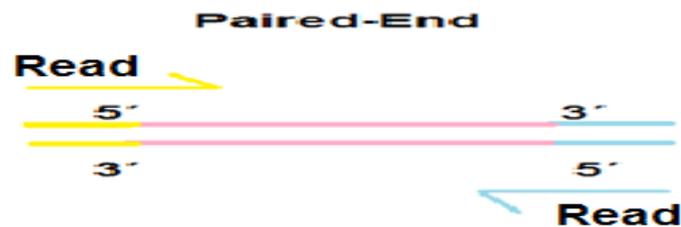


Figura 2.6 Genoteca “*paired-end*”. A partir de una genoteca “*paired-end*” se obtienen dos secuencias, sentido y antisentido.

La genoteca “*mate-pair*”, que es construida por fragmentos genómicos de ADN mayores a 1000 pb, lo cual significa un exceso en la longitud de las lecturas y alto costo en la secuenciación, por ello los segmentos largos son circularizados vía ligación con un adaptador que promueve la recircularización, posteriormente estos son fragmentados para generar dos subfragmentos más cortos. Este tipo de genoteca provee información acerca de inversiones, sustituciones de un sólo

nucleótido y variación en la estructura del genoma. En la Figura 2.7 se muestra la genoteca *mate pair* ⁽³⁴⁾.

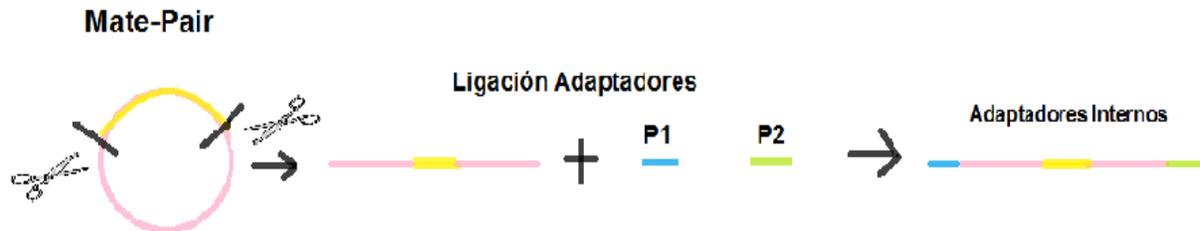


Figura 2.7 Genoteca “*mate pair*”, En la primera etapa de la construcción de estas genotecas el ADN es recircularizado por medio de dos adaptadores.

2.8. Métodos de fragmentación de ADN

La fragmentación de ADN es una etapa crucial para la construcción de las genotecas, para los cuales existen dos métodos distintos para la ruptura de forma mecánica de ADN que son la sonicación y nebulización. El otro método empleado es por medio de digestión enzimática con endonucleasas las cuales rompen los enlaces fosfodiéster del material genético.

En este trabajo se utilizaron dos métodos para fragmentar el material: el enzimático y el método mecánico por medio de sonicación; este último realiza la ruptura de ADN por medio de ondas de energía acústica de alta frecuencia, el proceso se realiza en un medio líquido donde se generarán millones de burbujas microscópicas -las cuales se expanden y colapsan- que transmiten su energía al ADN, cuyo proceso tiene el nombre de cavitación gaseosa ⁽³⁵⁾.

2.9 Características generales de las bacterias utilizadas en el presente estudio

Las bacterias utilizadas en este proyecto pertenecen al género de *Salmonella*, que es una de las principales bacterias causantes de ETA ⁽⁴⁾. La salmonelosis es una de las más comunes y ampliamente distribuidas en el globo terráqueo. Se estima que afecta anualmente a decenas de millones de personas en todo el mundo y provoca más de cien mil defunciones ⁽³⁶⁾.

Las bacterias *Salmonellas* pertenecen a la familia *Enterobacteriaceae*, son bacilos gram negativos, anaerobios facultativos, que no producen esporas. La clasificación más reciente de *Salmonella*, basada en la secuenciación de ADN, considera solamente dos especies: *S. enterica* y *S. bongori*, a su vez divididas en subespecies y serovariedades. Hasta el presente se han identificado más de 2.500 cepas diferentes (llamadas “serotipos” o “variantes séricas”) de *Salmonella* spp, por lo cual en este proyecto se trabajó con cuatro serotipos: *Salmonella enterica* serovariedad *Muenchen*, *Salmonella enterica* serovariedad *Kentucky* ⁽³⁷⁾.

El serotipo *Salmonella enterica* serovariedad *Muenchen* es uno de los 2500 serotipos de *Salmonella*. En los EUA, *Salmonella Muenchen* es una serovariedad común, se encuentra entre los diez principales serotipos aislados de los seres humanos.

En cuanto a su distribución geográfica podemos decir que la *S. Muenchen* se ha reportado entre los serotipos más comunes en distintos lugares de EUA, Australia y Europa. También hay algunas identificaciones microbiológicas de *S. Muenchen* en algunos países de África (Sudáfrica y Etiopía), Asia (Singapur y Malasia) y América del Sur (Brasil). Las fuentes alimenticias asociadas a la

enfermedad causada por *Salmonella Muenchen* son: melón, cerdo, alfalfa y jugo de naranja ⁽³⁸⁾.

Salmonella enterica serovariedad Kentucky es el serotipo más común identificado en fuentes clínicas no humanas, de acuerdo con el Centro para el Control y Prevención de Enfermedades (CDC) en 2011. Sin embargo, en Europa, la serovariedad Kentucky se encuentra en aproximadamente 1% de los casos humanos. Esta cepa muestra resistencia antibióticos β -lactámicos, carbapenémicos, quinolonas, aminoglucósidos, cotrimoxazol, y a la azitromicina. La caracterización molecular muestra una isla genómica cromosómica, que contiene los genes que confieren resistencia a estos antimicrobianos.

La *S. Kentucky* está ampliamente distribuida en E.U.A., y se encuentra comúnmente en los animales -específicamente en ganado- y aves de corral. En Europa, Asia y África, esta serovariedad también está aislada con regularidad de los seres humanos, con síntomas compatibles con la salmonelosis ⁽³⁹⁾.

Salmonella enterica subsp. enterica serotipo Enteritidis es la serovariedad más común en los E.U.A y Europa, se asocia con el consumo de pollo y blanquillos. La *S. Enteritidis* puede contaminar con éxito los huevos. El principal reservorio animal es el pollo, sin embargo también se encuentra en bajo porcentaje (<1%) en el pavo y el ganado. La distribución geográfica de serovariedad *Enteritidis* es global; según la Organización Mundial de la Salud, es el serotipo más común reportado en casos de salmonelosis en Europa, Norteamérica, Sudamérica, África y Asia ⁽⁴⁰⁾.

Por otro lado *Salmonella entérica* serotipo *Panamá* causa generalmente gastroenteritis en humanos -que es más invasiva que la causada por otros

serotipos- y afecta principalmente a los niños. En cuanto a su distribución geográfica se han reportado brotes en varios países americanos como Chile y México; así como en países europeos. La *S. Panama* se encuentra principalmente en productos de carne de cerdo ^(41, 42, 43). En el Cuadro 2.4 se muestra el tamaño del genoma de los cuatro serotipos de *Salmonella* que se utilizaron para este proyecto.

Cuadro 2.4 Tamaño del Genoma de las bacterias

Nombre de Bacteria	Tamaño de Genoma	Referencia
<i>Salmonella enterica serotipo Panamá</i>	4.5 Mb	Base de Datos Patrick https://www.patricbrc.org/portal/portal/patric/Home
<i>Salmonella enterica serotipo Enteritidis</i>	4.6 Mb	Base de Datos Patrick https://www.patricbrc.org/portal/portal/patric/Home
<i>Salmonella enterica serotipo Kentucky</i>	4.7 Mb	Base de Datos NCBI http://www.ncbi.nlm.nih.gov/genome/genomes/152?http://www.ncbi.nlm.nih.gov/genome/152?genome_assembly_id=23088
<i>Salmonella enterica serotipo Muenchen</i>	5.0 Mb	Base de Datos Patrick https://www.patricbrc.org/portal/portal/patric/Home

2.10. Antecedentes particulares del presente estudio

La utilización de la secuenciación masiva en los laboratorios clínicos ha sido empleada para el diagnóstico de enfermedades en México en el Instituto de Diagnóstico y Referencia Epidemiológicos (InDRE), es utilizada para la identificación de virus de interés epidemiológico, por ejemplo el virus del dengue, virus de la influenza y chikungunya, por lo cual cada uno de los componentes de la técnica de secuenciación -fragmentación del ADN, preparación de la genoteca, métodos de enriquecimiento, secuenciación y análisis de datos bioinformáticas- deben ser optimizados, documentados y validados en el ámbito clínico para obtener resultados confiables y así un diagnóstico preciso ^(44, 45).

La secuenciación es utilizada para el diagnóstico de enfermedades complejas, como cáncer, esclerosis, Síndrome Stickler, aneuploidias, cardiopatías hereditarias, retinitis pigmentaria etc, por tanto, sí esta técnica es una herramienta para el diagnóstico de padecimientos graves, debe estar estandarizada, validada y documentada ^(22, 25, 46).

Por otro lado, en España, Martín Ortiz en colaboración con otros investigadores, en 2012, realizaron la validación clínica de una genoteca de secuenciación masiva para el diagnóstico genético de las cardiopatías, donde consiguieron resultados de especificidad y sensibilidad al 100% de la cobertura de la genoteca, con lo cual obtuvieron conclusiones decisivas para hacer el diagnóstico para este tipo de pacientes ⁽⁴⁷⁾.

Budowle y colaboradores, en el año 2014 realizaron una validación de secuenciación, en el campo de investigación forense, donde determinaron que los

criterios de especificidad y sensibilidad de una validación son necesarios para dar un resultado confiable, así como para el cumplimiento de los fines legales que se requieren en el área forense ⁽²⁹⁾.

En México, hasta el momento, en el área agroalimentaria no se ha realizado una estandarización con procesos de elaboración de genotecas de secuenciación masiva, a pesar de que esta técnica está en aumento, para la tipificación de microorganismos patógenos, y casos complejos de microorganismos contaminantes en los alimentos.

III. PLANTEAMIENTO DEL PROBLEMA

3.1 PLANTEAMIENTO DEL PROBLEMA

La elaboración de genotecas es una etapa crucial para la secuenciación masiva, durante esta etapa se pueden producir errores, por ejemplo una cobertura desigual, cambios en la secuencia, y sesgos de longitud producidos por la fragmentación del ADN. Actualmente no existe una forma de garantizar reproducibilidad en el proceso de generación de genotecas bacterianas con un tamaño específico para la química de 400 pares de bases (pb) de la plataforma de secuenciación masiva Ion Torrent PGM.

Hasta el momento no hay evidencia acerca de la estandarización o validación de la técnica de Secuenciación Masiva aplicada a bacterias patógenas de interés agroalimentario, por lo tanto, en este trabajo se busca realizar una estandarización de las condiciones de fragmentación enzimática y mecánica por sonicación, así como establecer el parámetro óptimo para la selección de tamaño por medio de electroferesis. Todo lo anterior para la elaboración de genotecas en dicha plataforma de secuenciación, y con ello, obtener genotecas con un tamaño específico de 450 a 480 pb, y reducir así el tiempo de los procesos de elaboración.

IV. HIPÓTESIS

Si se controlan todas las condiciones para la elaboración de genotecas bacterianas, entonces se obtendrán genotecas con fragmentos de entre 450 a 480 pb y con una concentración mayor de 5 millones de moléculas que cumplan con los parámetros para una corrida de secuenciación de la plataforma de secuenciación masiva Ion Torrent PGM.

V. OBJETIVOS

5.1 Objetivo General

Estandarizar el proceso de elaboración de genotecas de bacterias de interés agroalimentaria en la plataforma de secuenciación masiva Ion Torrent PGM para la química de 400 pb.

5.2 Objetivos particulares

- Establecer la duración de tratamiento y ciclos de rompimiento de la fragmentación por el método de sonicación para obtener fragmentos en un rango de 350-470 pb.
- Determinar el tiempo óptimo de digestión enzimática para obtener fragmentos en un rango de 350-470 pb.
- Implementar el rango óptimo de pares de bases para la selección de tamaño por medio de electroforesis.
- Obtener genotecas con una concentración mayor de 5 millones de moléculas.

VI. MATERIALES Y MÉTODOS

6.1 Estrategia experimental

La estrategia experimental consistió en determinar los parámetros óptimos para el proceso de formación de las genotecas bacterianas, con el fin de estandarizar su elaboración para la plataforma de secuenciación masiva en paralelo de Ion Torrent PGM de la química de 400 pb.

Se seleccionaron cuatro bacterias provenientes del cepario del Laboratorio de Diagnóstico para la Detección de Organismos Patógenos (LCDDOP) pertenecientes al SENASICA. De las muestras anteriores se extrajo el ADN con el cual se establecieron los parámetros de fragmentación y selección de tamaño de ADN, se elaboraron diez genotecas bacterianas, cinco fragmentadas por digestión enzimática y cinco fragmentadas por sonicación.

Posteriormente se evaluó la calidad de genotecas, por medio de una electroforesis de microfluidos en un chip de Alta sensibilidad de ADN, y se evaluó la concentración de genotecas por medio de una qPCR en el equipo Lyght Cyclor 480 II.

6.2 DIAGRAMA DE FLUJO EXPERIMENTAL

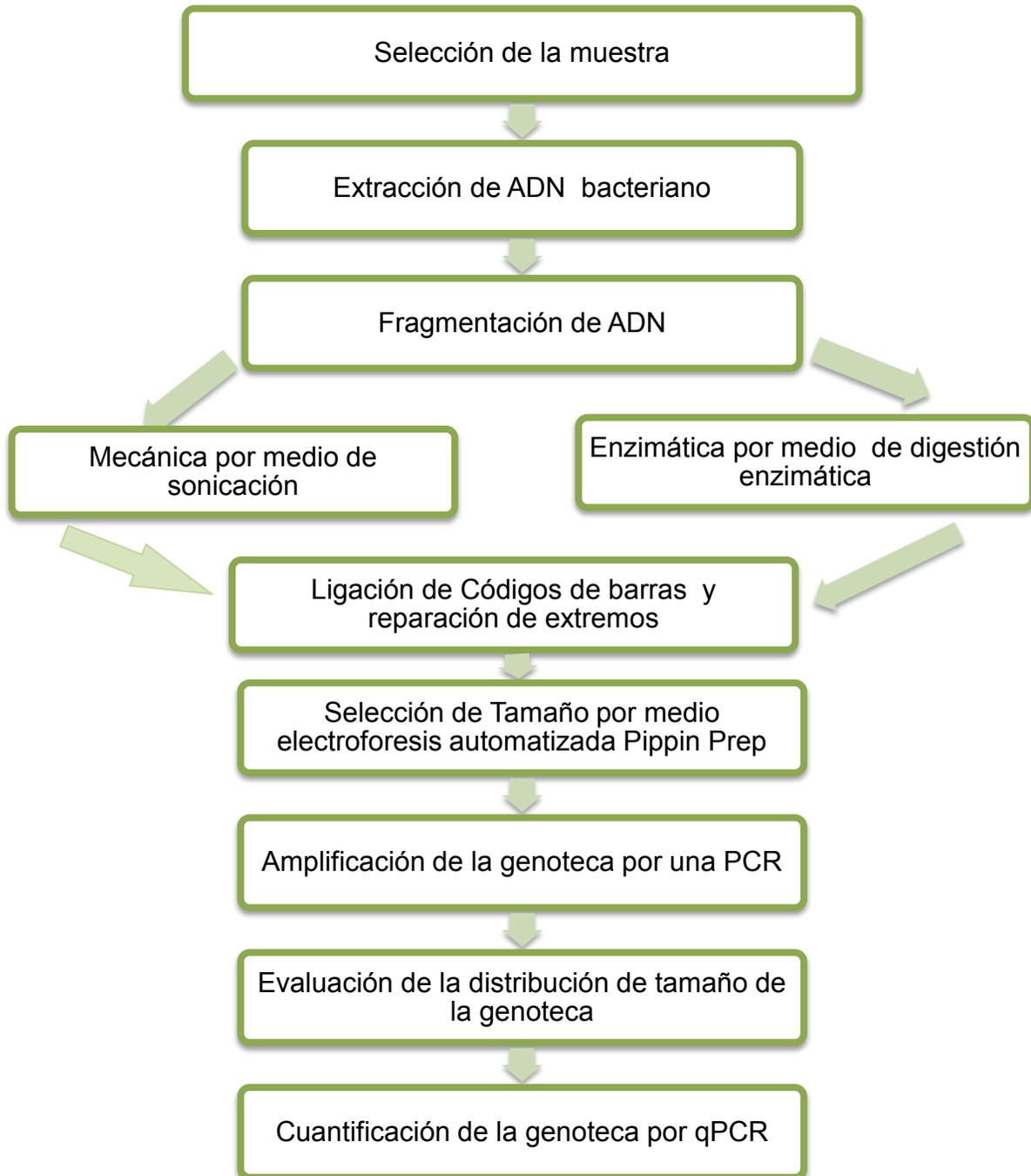


Figura 6.1 Esquema del flujo del trabajo realizado

6.3 Selección de Muestras

Se eligieron cuatro muestras de bacterias patógenas, provenientes del cepario del laboratorio de diagnóstico para la detección de organismos patógenos (LCDDOP) pertenecientes al SENASICA.

De la muestra *Salmonella Panamá* se partió de 100 µL del material genético que había sido extraído previamente, y de las muestras *Salmonella enteritis*, *Salmonella Kentucky* y *Salmonella Muenchen* se recibieron 3 alícuotas de 3 mL de células bacterianas (con una concentración de 1×10^{-9} células/µl) de las cuales se realizó la extracción de ADN. En el Cuadro 6.1 se muestra el nombre de las bacterias y el estatus de las mismas, con las que se trabajó.

Cuadro 6.1 Nombre de las bacterias patógenas

Número de Muestras	Nombre de las bacterias	Estatus de la muestra
1	<i>Salmonella Panamá</i>	ADN
2	<i>Salmonella Enteritidis</i>	Células bacterianas
3	<i>Salmonella Kentucky</i>	Células bacterianas
4	<i>Salmonella Muenchen</i>	Células bacterianas

6.4 Extracción de ADN

La extracción de ADN de las bacterias, se realizó con el kit de ADN Comercial FAST ID. Se utilizó un volumen de 1 mL de la suspensión bacteriana, y se centrifugó a 3000 rpm durante 10 min, para concentrar un pellet de células bacterianas.

Posteriormente se agregaron 500 μ L de Buffer TE 1X, se centrifugó, se resuspendió en 200 μ L de Buffer TE 1X y se adicionó 40 μ L de lisozima (10mg/mL). Se agitó e incubó a 37 °C durante 30 min. Se centrifugó a 10000 g durante 5 min. y se desechó el sobrenadante. Se le agregó 1 mL de Buffer de lisis y 10 μ L de Proteinasa K (10 mg/mL) y se agitó e incubó en el termobloque a 65°C durante 30 min. con agitación de 600 rpm.

Posteriormente se agregó 750 μ L de cloroformo, se agitó con un vórtex y centrifugó a 10,000 g por 5 min para la separación de las fases. De la fase superior se tomaron 800 μ L y se colocan en otro tubo de 2 mL. A esta fase superior se le agregó 800 μ L del Buffer de unión, se agitó con un vórtex y centrifugó a 10000 rpm por 30 s.

Después, se realizaron tres lavados con un volumen 500 μ L de solución de lavado, que se centrifugaron a 10,000 rpm por 30 s. A continuación se llevaron a cabo dos lavados con 700 μ L de etanol al 75 % de. Posteriormente se centrifugaron a 10,000 rpm por 30 s, por último se volvió a centrifugar con el fin de eliminar el exceso de etanol y secar a 10,000 rpm por 30 s.

La columna se colocó en un tubo de 1.7 mL y se le agregaron 50 μ L de Buffer TE 1X para eluir (previamente calentado a 65°C) y se incubó a 65°C durante 10 min. Se le adicionó 50 μ L de Buffer TE 1X y se centrifugó a 10000 rpm por 30 s.

Finalmente se recuperaron, -en un tubo de 1.7 mL- 100 μ L de ADN de las bacterias. La cuantificación de la concentración de ADN se realizó mediante una técnica espectrofotométrica en el equipo Nanodrop y para evaluar la pureza de ADN se tomó en cuenta la relación 260/280 porque el ADN absorbe a la longitud 260 nm y las proteínas a la longitud de 280 nm.

6.5 Fragmentación del ADN

La fragmentación de ADN es la primera etapa del proceso de elaboración de genotecas, se utilizaron dos métodos para la fragmentación: el enzimático y el mecánico por sonicación, para este último se utilizó el equipo covaris S220.

6.5.1 Fragmentación Enzimática

Para la fragmentación enzimática de ADN, se probaron diferentes tiempos de digestión enzimática 6, 7, 7.5, 8 y 9 min, con el fin de establecer el tiempo óptimo de digestión enzimática.

Una vez establecido el tiempo óptimo para la digestión enzimática, se utilizó lo equivalente a 200 ng de ADN bacteriano, se colocó en un tubo de 1.6 mL y se llevó a un volumen de 35 μ L con agua libre de nucleasas y se adicionaron 5 μ L de buffer de reacción Ion Shear Plus 10 x y 10 μ L de enzima Mix II Ion Shear Plus, la cual se mezcló por pipeteo 8 veces .Posteriormente se incubó el ADN a 37 °C

durante 7.5 min, con agitación a 300 rpm, y finalmente, para detener la reacción se adicionó el buffer Ion Shear Stop buffer.

6.5.2 Fragmentación por Sonicación

Para llevar a cabo la fragmentación por sonicación se utilizó el equipo Covaris S220, se realizaron alícuotas de ADN que contenían lo equivalente a 200 ng de la muestra y se llevaron a un volumen de 50 μ L con agua libre de nucleasas y se utilizaron diferentes números de ciclos de rompimiento. En el Cuadro 6.2 se muestran los ciclos de rompimiento utilizados.

Cuadro 6.2 Número de ciclos de rompimiento

Numero de ciclos de rompimiento
100
200
230
250
300

Condiciones del tratamiento de la fragmentación por sonicación PIP=175 w,
Duty Factor (D.F)= 10 %, tiempo 80 s

Se utilizaron diferentes tiempos para el tratamiento de fragmentación por sonicación. En el Cuadro 6.3 se muestran los tiempos utilizados en el tratamiento de sonicación para la fragmentación.

Cuadro 6.3 Número de ciclos de rompimiento

Tiempo (s)
80
160
360

Condiciones del tratamiento de la fragmentación por sonicación
PIP=175 w, D.F. 10 %, Numero de ciclos 100

Una vez fragmentado el ADN, se realizó la reparación de los extremos romos, para ello se realizó la siguiente reacción: se adicionó 29 μL de agua libre de nucleasas, 20 μL buffer 5x end repair y 1 μL la enzima end repair para tener un volumen final de 100 μL y se incubó por 20 min a temperatura ambiente.

6.6 Purificación de fragmentos de ADN

La purificación de los fragmentos se realizó con perlas magnéticas recubiertas con estreptavidina, una vez fragmentado el ADN se le adicionó 99 μL Agencourt Ampure XP, se incubaron por 5 min a temperatura ambiente para que los fragmentos de ADN se unieran a las perlas, después se colocaron en un concentrador de perlas magnéticas (MPC) durante 3 min para separar el material genético unido a las perlas de contaminantes.

Se realizaron dos lavados con 500 μL de etanol al 70 % a las perlas con ADN; para eliminar el exceso de etanol se dejó destapado el tubo durante 3 min dentro del MPC, se retiraron los tubos del MPC y se realizó la elución con 25 μL de buffer Low TE y se incubó a temperatura ambiente durante 1 min, se colocaron nuevamente en el MPC durante 1 min.

6.7 Ligación de adaptadores y reparación de extremos

Para la ligación de adaptadores se preparó, en un tubo de PCR de 0.2 mL, la reacción de ligación de adaptadores y reparación de extremos. En el Cuadro 6.4 se muestran los reactivos y volumen utilizados para preparar la reacción de

ligación y se colocaron en el termociclador con las siguientes condiciones: 25 °C 15 min, 72 °C 5 min, y 4°C.

Cuadro 6.4 Reacción de ligación de adaptadores y reparación.

Reactivos	Volumen (μL)
Ligasa buffer 10 x	10
Ion P1 Adapter	2
Ion Xpress Barcode X	2
dNTP Mix	2
Ligasa ADN	2
Nick Repair Polymerase	8
Agua libre de nucleasas	49

6.8 Purificación de las genotecas

La purificación de las genotecas se realizó con 100 μL de Agencourt Ampure XP, se incubaron durante 5 min. a temperatura ambiente, y posteriormente se pasaron a la MPC, se incubaron durante 3 min. Se realizaron dos lavados con etanol al 70%, se dejó abierto el tubo durante 3 min para eliminar el exceso de etanol. Se retiraron del MPC y se eluyó con 20 μL de Low TE.

6.9 Selección de Tamaño de ADN

La selección de tamaño se realizó por medio de una electroforesis automatizada, en el equipo *Pippin prep*. La muestra se preparó en un tubo de 0.6 mL, se adicionaron 20 μL de la muestra más 10 μL Low TE y 10 μL Loading Solution para tener un volumen final de 40 μL y el marcador se preparó en un tubo de 0.6 mL, 20 μL de marcador peso molecular, más 10 μL Low TE y 10 μL Loading Solution.

Una vez preparadas las muestras y el marcador, se cargaron en el casete

de gel agarosa 2% el cual ya viene preparado. En la Figura 6.2 se muestra el casete de gel de agarosa. En el equipo *Pippin prep* se programó el tiempo de corrida, que fue de 1.5 horas y el objetivo (target) para obtener fragmentos entre el rango 450pb a 480 pb, se probaron tres opciones de target el de 460 pb, 470 pb y 475 pb.

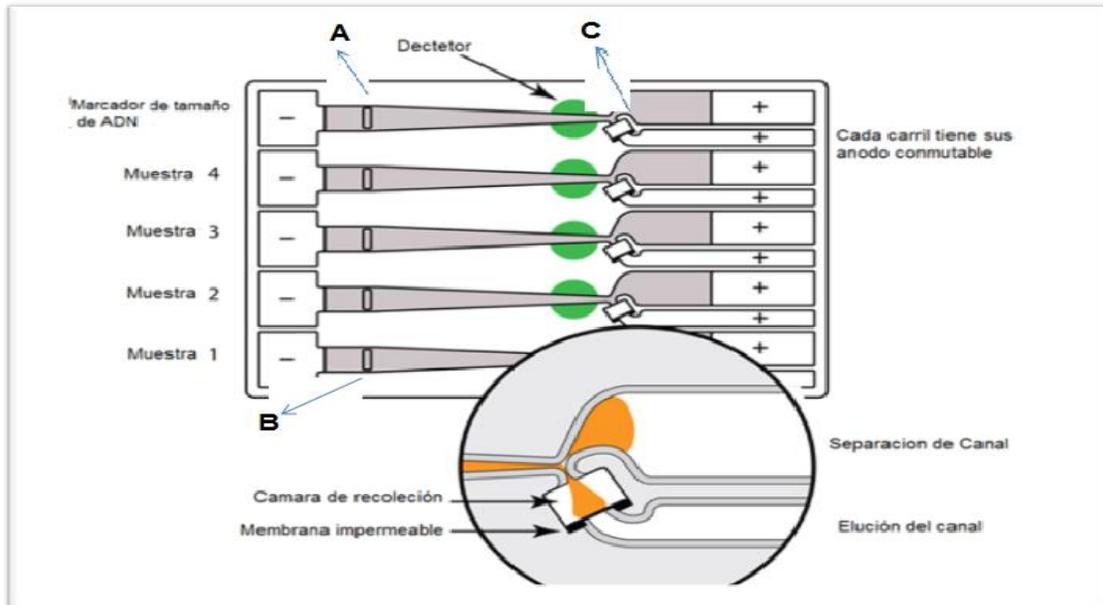


Figura 6.2 Casete de gel de agarosa del equipo Pippin prep. A) Pozo donde se carga el marcador. B) Pozo donde se carga la muestra. C) Pozo donde se recolecta la muestra.

6.10. Purificación de las genotecas después de la selección de tamaño

Las muestras que se recolectaron después de la selección de tamaño, se purificaron con 99 μL de Agencourt Ampure XP; se incubaron durante 5 min. a temperatura ambiente, y se colocaron en el MPC durante 3 min. Se realizaron dos lavados con etanol al 70 %, se dejó abierto el tubo durante 3 min para evaporar el exceso de etanol. Finalmente los tubos se retiraron del MPC y se eluyó con 25 μL de Low TE.

6.11 Amplificación de genoteca bacteriana

Después de la selección de tamaño se realizó la amplificación de la genoteca por medio de una PCR punto final, para ello se preparó la reacción de amplificación con 100 μL Platinum PCR Super Mix más 5 μL de library Amplification Primer Mix y los 25 μL de genoteca bacteriana para tener un volumen final de 130 μL . Se colocó la reacción de amplificación en el termociclador con el protocolo que se muestra en la Figura 6.3 para realizar la amplificación.

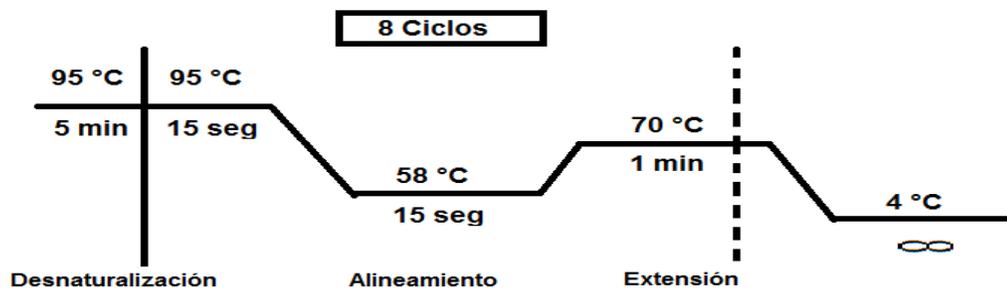


Figura 6.3 Condiciones para la amplificación de las genotecas. Se utilizaron 8 ciclos con las condiciones de 95°C de desnaturalización durante 5 min, 95°C por 15 s, 58°C durante 15 s de alineación, 70°C por 1 min de extensión.

6.12 Purificación de la genoteca amplificada

Una vez amplificada la genoteca, se realizó la última purificación con las perlas magnéticas con 130 μL de Agencourt Ampure XP, se incubaron durante 5 min a temperatura ambiente, y se colocaron en el MPC durante 3 min. Se realizaron dos lavados con etanol al 70 %, para eliminar el exceso de etanol. Se expuso a temperatura ambiente por 3 min. Se retiró del MPC y se eluyó con 20 μL de Low TE. La genoteca ya purificada se recuperó en un tubo de 0.6 mL.

6.13 Evaluación del tamaño de la genoteca

La evaluación de la distribución de tamaño de la genoteca se llevó a cabo por medio de una electroforesis de microfluidos en el equipo Bionalizador Agilent



2100, se utilizó un chip de alta sensibilidad, el cual se colocó en la estación. En la Figura 6.4 se muestra la estación en la cual se coloca el chip de alta sensibilidad,

Figura 6.4 Estación para realizar la presurización.

Para realizar el cargado de chip, los reactivos se colocaron a temperatura ambiente por 30 min; después se realizó el cargado del chip el cual se muestra en la Figura 6.5.

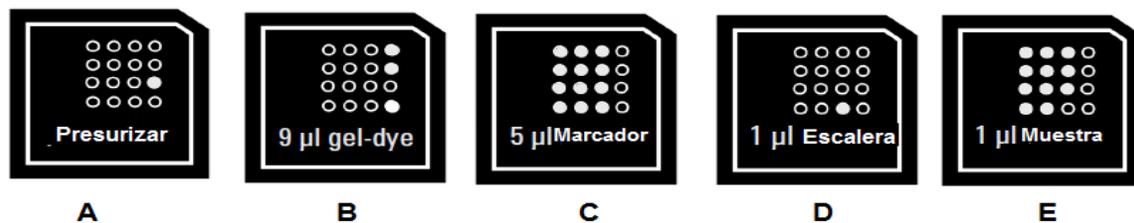


Figura 6.5 Cargado del chip de alta sensibilidad. A y B. Cargado del Chip con matriz gel dye C. Marcador de referencia D. Cargado

En el gráfico A de la Figura 6.5 se muestra en que pozo se adicionaron 9 μ L de gel dye mix y se presionó el émbolo de la jeringa, se colocó el seguro y se dejó

por 1 min para realizar la presurización de la matriz, después se liberó hasta que llegara a su posición inicial en un 1 mL. En el gráfico B se muestra el pozo donde se agregaron 9 μL de gel dye mix. En el gráfico C se muestra en que pozo se adicionaron 5 μL de marcador de ADN de alta sensibilidad. En el Gráfico D se muestra el pozo donde se colocó 1 μL de marcador de talla molecular, finalmente en el gráfico E se muestra en que pozo se adicionó 1 μL de muestra. Una vez cargado el chip, este se colocó en el vórtex durante 1 min. a 2400 rpm, se realizó una revisión visual para que no existieran burbujas; finalmente se colocó en el equipo bionalizador.

6.14 Cuantificación de genoteca por qPCR.

La cuantificación de las genotecas se realizó por medio de la técnica PCR en tiempo real (qPCR) en el equipo *LightCycler*® 480 de Roche, se utilizó un kit KAPA SYBR® Fast qPCR. Para lo cual se prepararon diluciones seriales: 1:500, 1:1000, 1:2000, 1:4000, en la Figura 6.6 se muestran las diluciones. Una vez preparadas las diluciones se realizó la mezcla de reacción, se añadieron 6 μL Master Mix y 4 μL de cada una de las diluciones de la genotecas se agitó con vórtex, se realizó por triplicado cada una de las reacciones y se colocaron 10 μL de la reacción en una placa de 96 pozos.

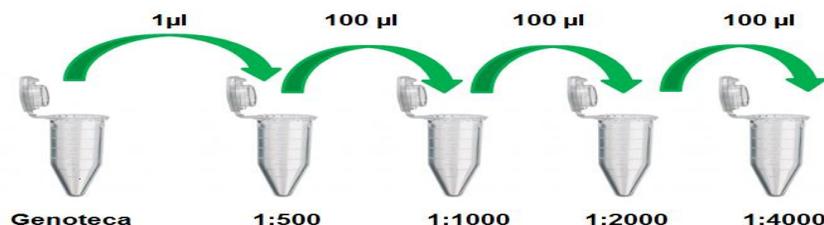


Figura 6.6 Diluciones seriales de genotecas.

VII. RESULTADOS

Selección de las muestras y cuantificación de ADN.

Las bacterias utilizadas en este trabajo son provenientes del cepario del Laboratorio de Diagnóstico para la Detección de Organismos Patógenos (LDDOP) que pertenece al SENASICA, a las cuales se les realizó la extracción de ADN, la concentración se cuantificó por espectrofotometría en el equipo Nanodrop. En el Cuadro 7.1 se muestran la concentración y relación de pureza de ADN de las muestras.

Cuadro 7.1 Concentración y pureza de ADN de las bacterias.

Números	Nombre de Bacteria	Concentración (ng / μ L)	Pureza 260/280	Pureza 260/230
1	<i>Salmonella entérica</i> <i>serotipo Enteritidis</i>	65.1	1.98	2.18
2	<i>Salmonella entérica</i> <i>serotipo Kentucky</i>	54.2	1.97	2.13
3	<i>Salmonella entérica</i> <i>serotipo Muenchen</i>	53.1	2.00	2.19
4	<i>Salmonell entérica</i> <i>serotipo a Panamá</i>	69.7	1.98	2.16

La concentración de las muestras obtenida fue mayor de 50 ng/ μ L -que es la cantidad mínima necesaria para elaborar genotecas- y la relación de pureza de 260/280 de las muestras está en el rango de 1.8-2.0 y para la relación 260/230 las muestras están entre 1.8-2.2 por lo tanto es considerado como un ADN de pureza adecuada, por este motivo se decidió utilizar estas bacterias para la elaboración de las genotecas.

Fragmentación del material genético ADN.

Una vez extraído el material genético, el siguiente paso fue la fragmentación de ADN, se estableció el tiempo de 80 s y 300 ciclos para la ruptura del material genético por medio de sonicación en el equipo Covaris S220 con la muestra *Salmonella Muenchen*. El Cuadro 7.2 muestra las condiciones establecidas para la fragmentación por sonicación.

Cuadro 7.2 Condiciones establecidas para la fragmentación por sonicación.

PIP (w)	DF (%)	Ciclos de rompimiento	Tiempo (s)	Volumen de la Muestra
175	10	300	80	50 μ L

PIP=Potencia Máxima Incidente, D.F.=factor de trabajo

Una vez fraccionado el ADN, se realizó la evaluación de los fragmentos por medio de una electroforesis de microfluidos en el Bionalizador Agilent 2100 en un chip de alta sensibilidad de ADN. Los resultados obtenidos se muestran en la Figura 7.1.

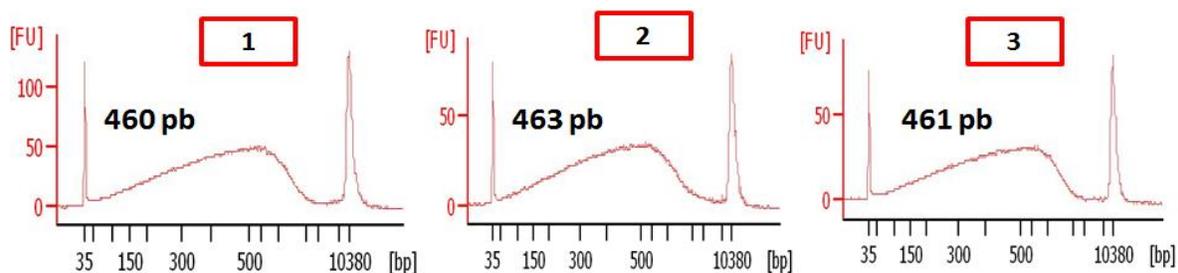


Figura 7.1 Gráfico de fluorescencia contra pares de bases (pb) de ADN de las repeticiones realizadas con las condiciones de la Cuadro 10. Gráfico 1) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella Muenchen*. Gráfico 2) Tamaño promedio de pb de

los fragmentos de ADN de la muestra *Salmonella Muenchen*. Gráfico 3) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella Muenchen*.

Después de que se fragmentó el ADN, con las condiciones que se establecieron para la sonicación en el equipo Covaris S220, se realizó estadística descriptiva, para comprobar la repetibilidad. En el Cuadro 7.3 se muestra el tamaño promedio del ADN fragmentado de la bacteria *Salmonella Muenchen*.

Cuadro 7.3 Tamaño del ADN fragmentado por Sonicación en el equipo Covaris S220.

Número de Repetición	Tamaño promedio de los fragmentos (pb)
1	460
2	463
3	461
Promedio	461.33
Desviación estándar	1.52
C.V(%)	0.331

Para confirmar que las condiciones establecidas para la fragmentación de ADN por sonicación fueran óptimas se probaron tres muestras de bacterias diferentes: *Salmonella Panama*, *Salmonella Kentucky* y *Salmonella Enteritidis*, las cuales se evaluaron por medio de electroforesis de microfluidos. En la Figura 7.2 se observa el tamaño promedio del ADN fragmentado por sonicación.

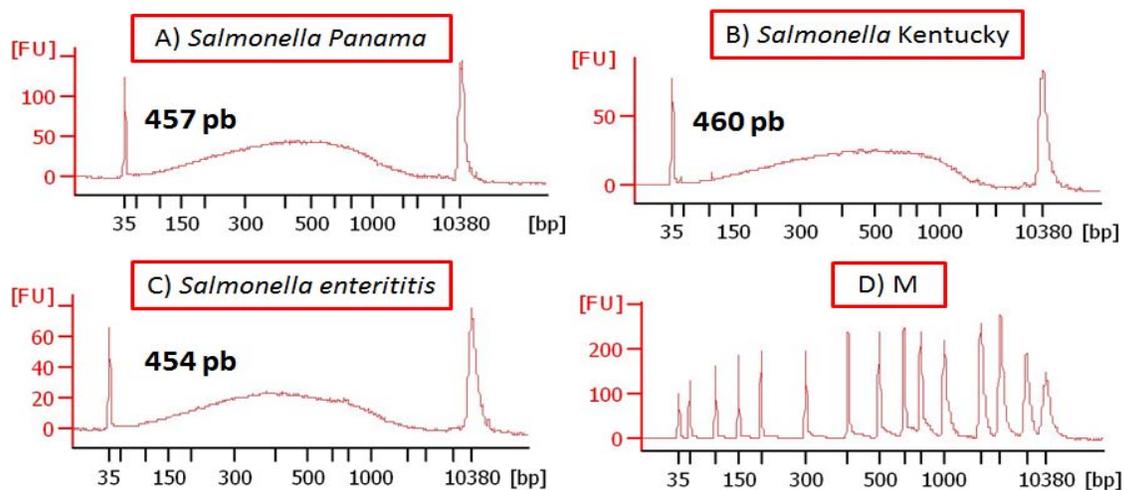


Figura 7.2 Gráfico de fluorescencia contra pares de bases (pb) de ADN fragmentado por sonicación con tres diferentes cepas de *Salmonella* en el bioanalizador Agilent en un chip Alta sensibilidad ADN. Gráfico A) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella Panama*. Gráfico B) Tamaño promedios de pb de los fragmentos de ADN de la muestra *Salmonella Kentucky*. Gráfico C) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella enteritidis*. Gráfico D) marcador de talla molecular.

Una vez establecidas las condiciones de fragmentación mecánica por sonicación, se realizaron pruebas para establecer el parámetro óptimo para la fragmentación enzimática, el tiempo que se estableció de digestión enzimática fue de 7.5 min.

Para confirmar que en el tiempo 7.5 min se obtuvieron fragmentos de tamaño óptimo para la elaboración de genotecas, se realizó la fragmentación enzimática de la muestra tres veces. Después se analizaron los fragmentos, en un chip de alta sensibilidad de ADN del Bionalizador Agilent 2100. En la Figura 7.3 se puede observar el gráfico de fluorescencia contra pares de bases del ADN fragmentado por enzimas y el tamaño promedio de cada una de las muestras.

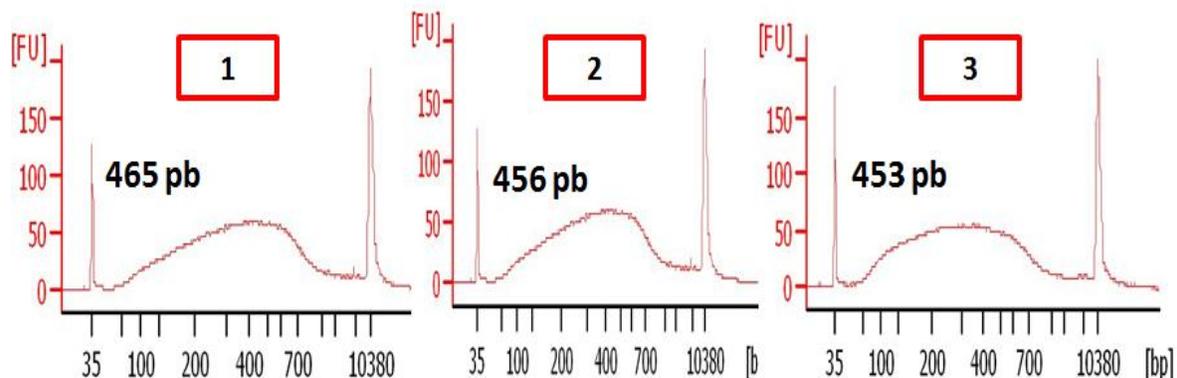


Figura 7.3 Gráfico de fluorescencia contra pares de bases (pb) de ADN fragmentado por enzimas de las repeticiones realizadas con el tiempo de 7.5 min de fragmentación enzimática en el bioanalizador Agilent en un chip de alta sensibilidad de ADN. Gráfico 1) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella Muenchen*. Gráfico 2) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella Muenchen*. Gráfico 3) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella Muenchen*.

Con la finalidad de comprobar la repetibilidad del método de fragmentación enzimática, se le aplicó estadística descriptiva al tamaño promedio de los fragmentos de ADN, para ello se utilizaron medidas de tendencia central (la media); y de dispersión (desviación estándar). En el Cuadro 7.4 se muestran los resultados de la estadística realizada.

Cuadro 7.4 Tamaño de ADN fragmentado por digestión enzimática.

Numero de Repetición	Tamaño promedio de los fragmentos (pb)
1	465
2	456
3	453
Promedio	468
Desviación estándar	6.24
C.V(%)	1.36

El coeficiente que se obtuvo fue de 1.36 y el promedio del tamaño de los fragmentos de ADN fue de 468 pb, por lo cual se concluyó que la fragmentación enzimática con el tiempo de digestión de 7.5 min es reproducible. Para confirmar que este parámetro es óptimo, se realizaron otras tres fragmentaciones con tres bacterias diferentes y se evaluó el tamaño en un chip de alta sensibilidad de ADN. En la Figura 7.4 se muestran los gráficos de ADN fragmentado de las tres muestras.

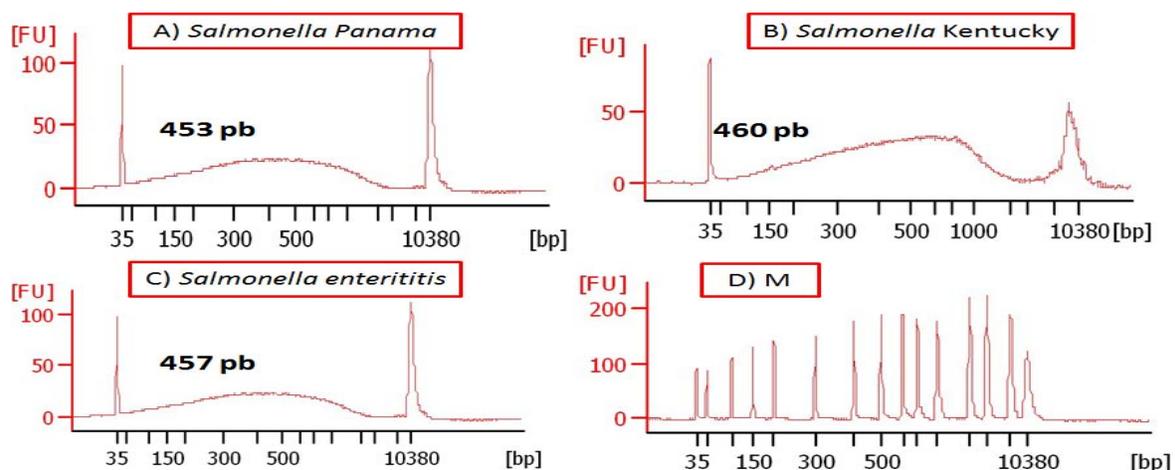


Figura 7.4. Gráfico de fluorescencia contra pares de bases (pb) de ADN fragmentado por el método enzimático con tres diferentes cepas de *Salmonella* en el bioanalizador Agilent en un chip alta sensibilidad de ADN. Gráfico A) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella Panama*. Gráfico B) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella Kentucky*. Gráfico C) Tamaño promedio de pb de los fragmentos de ADN de la muestra *Salmonella enteritidis*. Gráfico D) Marcador de talla molecular.

Selección de tamaño.

Una vez establecidos los parámetros óptimos de fragmentación de ADN por el método enzimático y mecánico por sonicación, se realizaron pruebas con el fin de establecer el parámetro para la selección de tamaño de ADN por medio de una electroforesis automatizada.

El parámetro “target” que se estableció para la selección de tamaño de ADN en el equipo Pippin Prep® fue de 470 pb, para ello se utilizaron las muestras fragmentadas por el método enzimático, posteriormente se realizó la ligación de adaptadores, que son secuencias de nucleótidos conocidos que se adhieren a cada uno de los extremos del material genético, con esto se convierten en genotecas y posteriormente se realizó la selección de tamaño.

Para confirmar si el “target” de 470 pb era el óptimo para la selección de tamaño, se elaboraron tres genotecas con este parámetro y se evaluaron en chip de alta sensibilidad en bionalizador Agilent 2100. En la Figura 7.5 se muestran las genotecas después de la selección de tamaño.

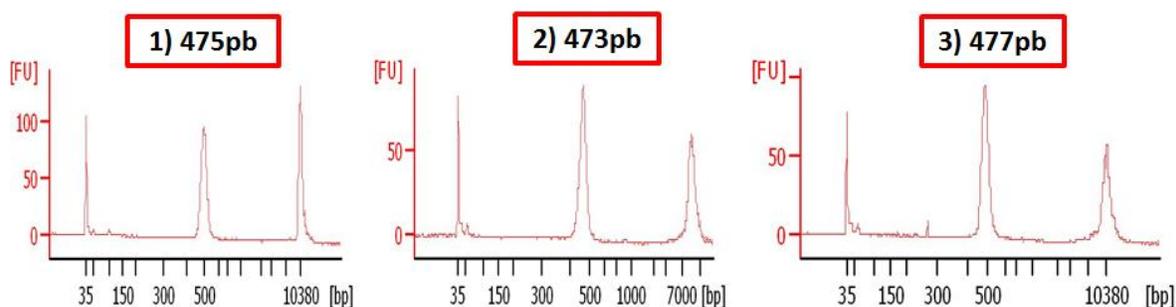


Figura 7.5 Gráfico de fluorescencia contra pares de bases (pb) de genotecas en un ensayo de tres repeticiones utilizando el “target” de 470 para la selección tamaño. Gráfico 1) Tamaño promedio de pb de genoteca de *Salmonella Muenchen*. Gráfico 2) Tamaño promedio de pb de genoteca de *Salmonella Muenchen*. Gráfico 3) Tamaño promedio de pb de genoteca de *Salmonella Muenchen*,

Posteriormente se realizó estadística descriptiva a los resultados obtenidos mediante el “target” de 470 pb para la selección tamaño, los resultados se muestran en el Cuadro 7.5.

Cuadro 7.5 Tamaño de genotecas después de la selección de tamaño de ADN.

Número de Repetición	Tamaño promedio de las genotecas
1	475
2	473
3	477
Promedio	475
Desviación estándar	2
C.V(%)	0.4215

Elaboración de Genotecas.

Después de haber establecido los parámetros de fragmentación y selección de tamaño de ADN, se realizó la elaboración de genotecas bacterianas, para corroborar que todo el proceso es repetible y reproducible; para esto se utilizaron 200 ng de ADN de la muestra y se elaboraron cinco genotecas fragmentadas de forma enzimática y cinco genotecas fragmentadas por sonicación.

Primero se realizaron cinco genotecas fragmentadas de forma mecánica por sonicación, en la Figura 7.6 se muestra el ADN fragmentado de cada una de estas muestras.

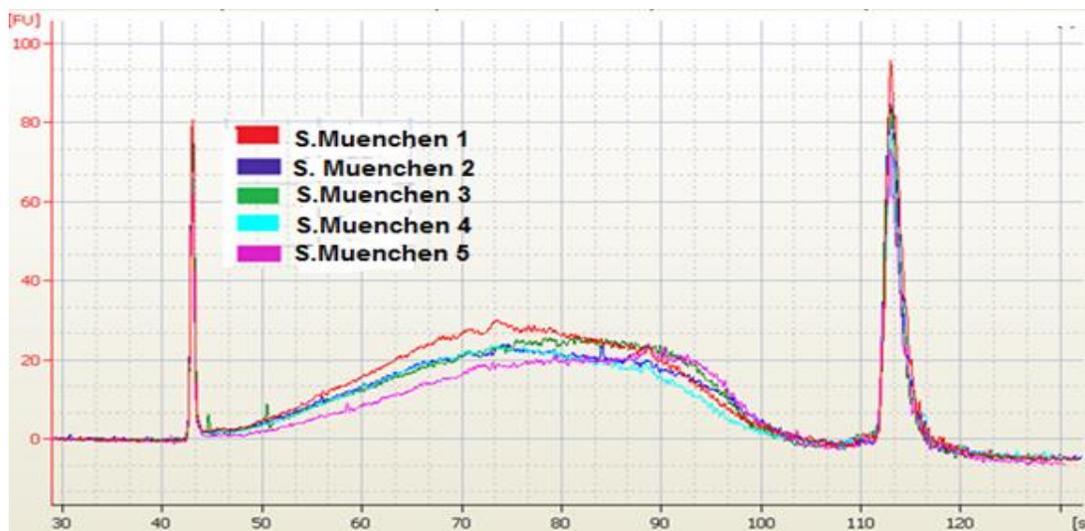


Figura 7.6 Gráfico de fluorescencia contra segundos(s) de ADN fragmentado por sonicación de las cinco genotecas proporcionado por el equipo bionalizador. *Salmonella Muenchen 1*: tamaño promedio de 459. *Salmonella Muenchen 2*: tamaño promedio de 454 pb. *Salmonella Muenchen 3*: tamaño promedio de 453 pb. *Salmonella Muenchen 4*: tamaño promedio de 453 pb. *Salmonella Muenchen 5*: tamaño promedio de 453 pb.

Con los parámetros de sonicación estandarizados previamente, se observó que los resultados de los fragmentos de ADN tuvieron un promedio de 454 pb, para los cinco materiales genéticos de la bacteria *Salmonella Muenchen*.

Después de la fragmentación, se ligaron los adaptadores, se realizó la selección de tamaño, y posteriormente se amplificó la genoteca con el objetivo de tener una mayor cantidad de biblioteca genómica para secuenciar.

Posteriormente, como en todas las reacciones de amplificación, se purificó el producto, y finalmente se realizó la evaluación de calidad de las genotecas fragmentadas por sonicación por medio de una electroforesis de microcanales en el bionalizador Agilent 2100. En la Figura 7.7 se observan las cinco genotecas, así como el tamaño promedio de cada una de ellas, mientras que en la Figura 7.8 aparece la imagen de gel de las genotecas proporcionado por el bionalizador.

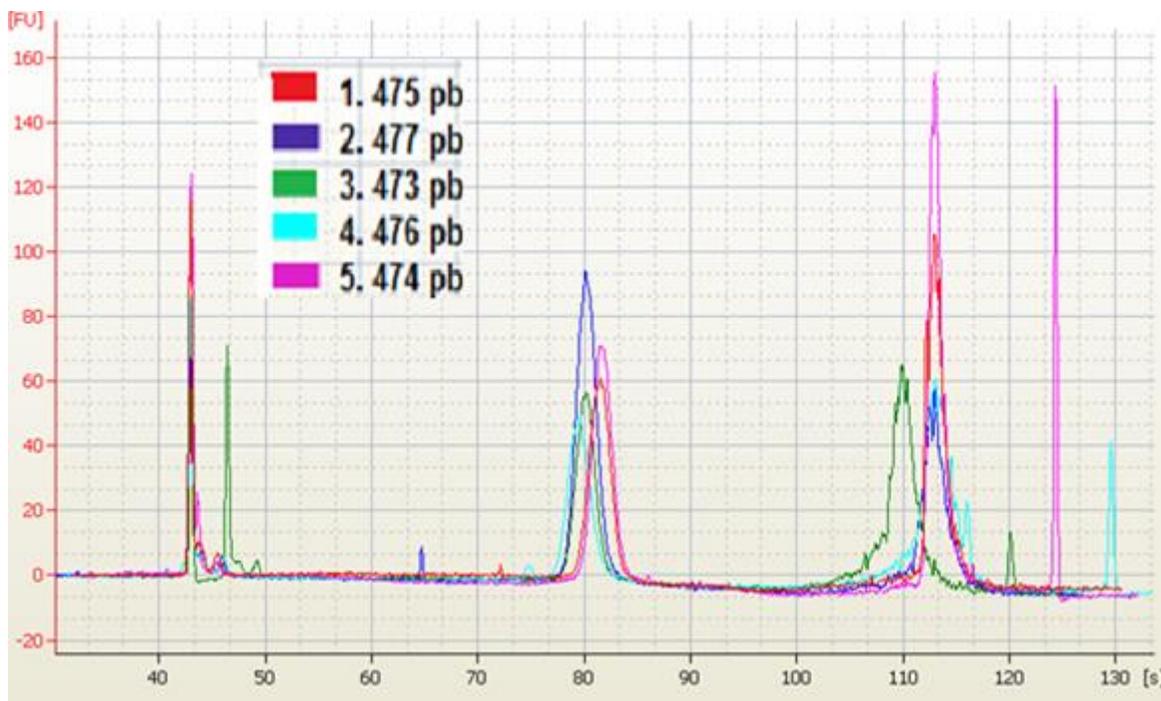


Figura 7.7 Gráfico de fluorescencia contra segundos(s) de las cinco genotecas de la bacteria *S. Muenchen* fragmentadas por sonicación.

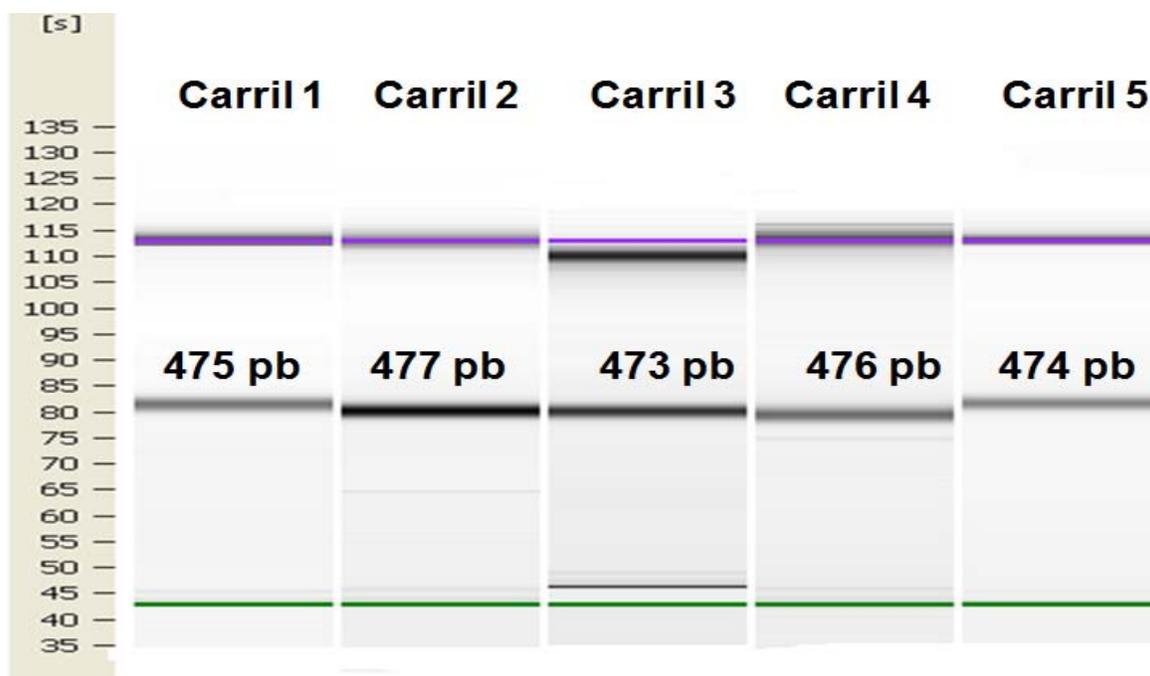


Figura 7.8 Imagen de gel bionalizador Agilent en chip Alta sensibilidad ADN de las genotecas fragmentadas por sonicación. Carril 1) *Salmonella Muenchen*, Carril 2) *Salmonella Muenchen*, Carril 3) *Salmonella Muenchen*, Carril 4) *Salmonella Muenchen*, Carril 5) *Salmonella Muenchen*. Tanto en la Figura 7.7 como en la Figura 7.8, se observa que las cinco genotecas

que se fragmentaron por sonicación, presentan intensidad de fluorescencia y tamaño en pb similares, en cada una de las cinco repeticiones realizadas, lo que indica que el proceso es reproducible, para comprobar lo anterior se realizó la respectiva estadística descriptiva, para verificar la repetibilidad del proceso y en la Cuadro 7.6 se muestran los resultados.

Cuadro 7.6. Estadística descriptiva para las genotecas bacterianas generadas por medio de la fragmentación mecánica.

Numero de repetición	Tamaño de Genoteca (pb)
1	475
2	477
3	473
4	476
5	474
Media del tamaño de la genoteca (pb)	475
Desviación estándar	1.581
Coefficiente de Variación (%)	0.332
Error Reproducibilidad	0.665
Error Precisión	1.063
Tamaño Genoteca ideal (pb)	470

Con esto se comprobó la repetibilidad del proceso de la genotecas bacterianas bajo las mismas condiciones, ya que se obtuvo un coeficiente de variación de 0.33 %, es decir, menor a 15 %, por lo tanto se dice que es repetible el proceso para la elaboración de genotecas ⁽⁴⁸⁾.

Una vez elaboradas las genotecas, se realizó la cuantificación por medio de una PCR en tiempo real. En la Figura 7.9 se muestra la gráfica de la curva patrón y la curva de una de las muestras, donde la pendiente obtenida fue de $m= 3.24$. En el Cuadro 7.7 se muestra la concentración de las genotecas fragmentadas por

sonicación.

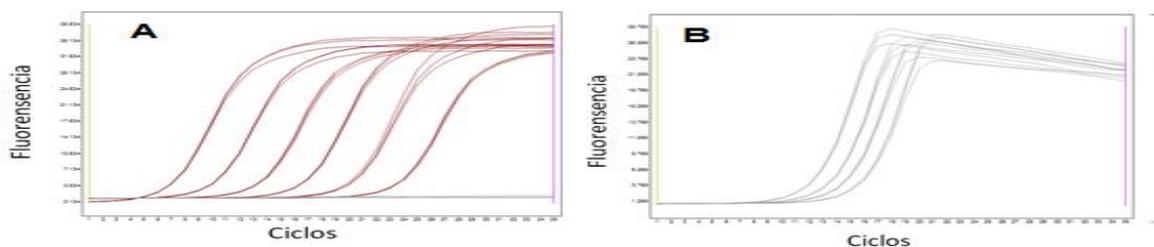


Figura 7.9 Gráfico de fluorescencia contra número de ciclos proporcionado por equipo light cycler 480 II Roche. A) Curva patrón. B) Curva de una de las genotecas.

Cuadro 7.7 Tamaño promedio y concentración de las genotecas bacterianas fragmentadas por sonicación.

Número repetición	T.G. (p b)	Concentración (moléculas/ μ L)
1	475	9.11E 07
2	477	7.72 E 08
3	473	6.41 E 08
4	476	2.09E 08
5	474	4.40E 08
	Promedio de Concentración (moléculas/ μ L)	4.31E.08

La concentración que se obtuvo de las cinco genotecas -que se muestra en el Cuadro 7.7- está arriba de 5 millones de moléculas/ μ L, que es la concentración necesaria para realizar la emulsión en PCR, que es el siguiente paso del proceso de secuenciación masiva. Por lo tanto, las condiciones de fragmentación por sonicación que se establecieron para el proceso de elaboración de genotecas, así como el parámetro de selección de tamaño “target”, brindan buenos resultados para obtener genotecas con los parámetros óptimos para la secuenciación.

Por otro lado también se elaboraron cinco genotecas fragmentadas enzimáticamente, una vez fragmentado el ADN se tomó una alícuota de 1 μ L para

evaluar el tamaño de los fragmentos, que se muestra en la Figura 7.10.

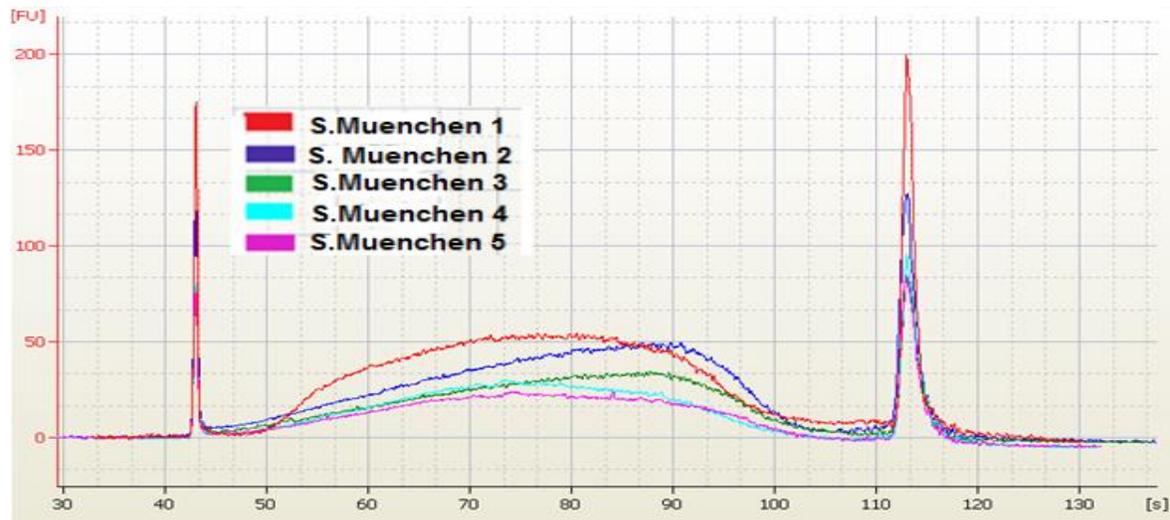


Figura 7.10. Gráfico de fluorescencia contra segundos (s) de los fragmentos de ADN de las cinco genotecas con el método enzimático. *Salmonella Muenchen 1*: tamaño promedio de 458 pb. *Salmonella Muenchen 2*: tamaño promedio de 456 pb. *Salmonella Muenchen 3*: tamaño promedio de 457 pb. *Salmonella Muenchen 4*: tamaño promedio de 452 pb. *Salmonella Muenchen 5*: tamaño promedio de 449 pb.

En la Figura 7.10 se muestra el ADN fragmentado por el método enzimático, el tamaño que se obtuvo fue óptimo -ya que estaba en el rango de 350 pb a 470 pb para la elaboración de genotecas-, por lo cual se decidió realizar las genotecas.

Una vez elaboradas, se evaluó la calidad de las genotecas por medio de una electroforesis de microfluidos en un chip de alta sensibilidad de ADN, los resultados se observan en la Figura 7.11, se muestran las genotecas sobrepuestas y el tamaño promedio de cada una, mientras que en la Figura 7.12 se observa el corrimiento electroforético proporcionado por el bionalizador Agilent 2100.

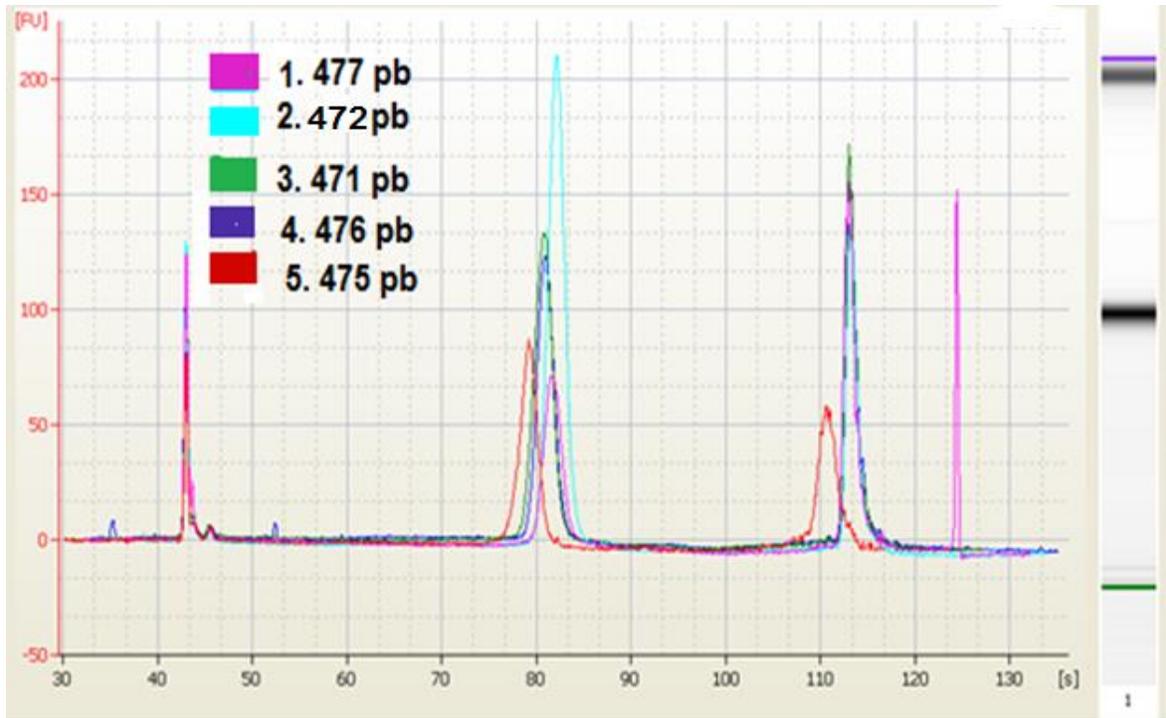


Figura 7.11 Gráfico de fluorescencia contra segundos (s) de las cinco genotecas bacterianas fragmentadas por digestión enzimática y el marcador de talla molecular.

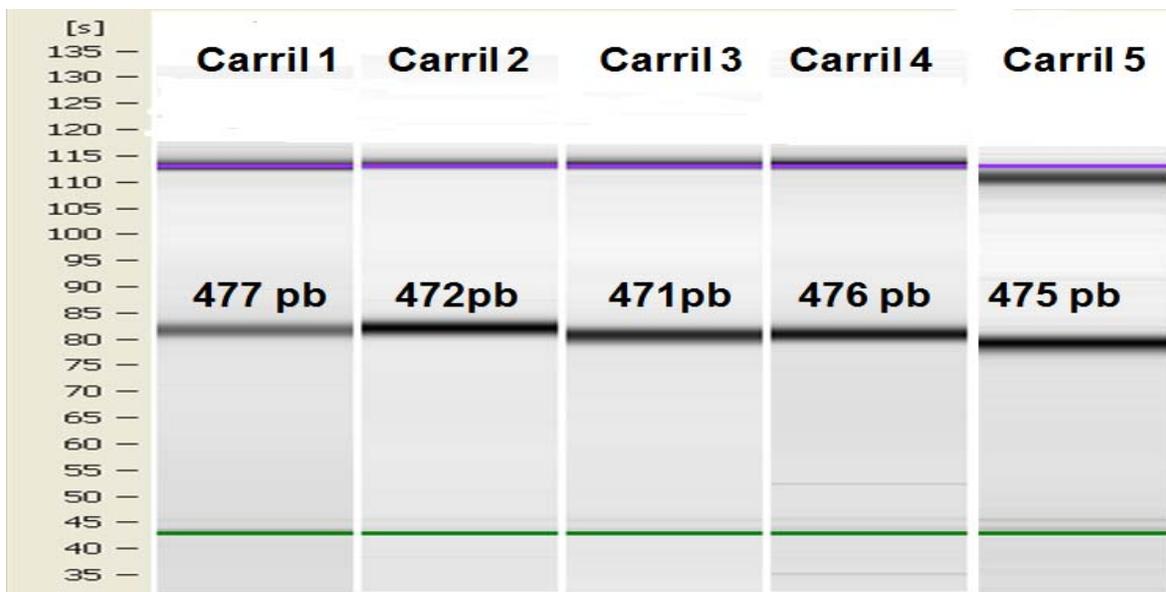


Figura 7.12 Corrimiento electroforético de las genotecas realizadas por digestión enzimática. Carril 1) *Salmonella Muenchen* Carril 2) *Salmonella Muenchen* Carril 3) *Salmonella Muenchen* Carril 4) *Salmonella Muenchen* Carril 5) *Salmonella Muenchen*.

En la Figura 7.11 se observa que el tamaño de las genotecas está entre el 471 a 477 pb, y que son muy similares en las cinco repeticiones y en la Figura 7.12 del corrimiento electroforético se observa la banda de las genotecas, que aparecen en tiempos cercanos, lo que nos proporciona indicios acerca de que el proceso es reproducible. Para comprobar lo anterior se realizó la respectiva estadística descriptiva, para verificar la repetibilidad del proceso y en el Cuadro 7.8 se muestran los resultados.

Cuadro 7.8. Media y desviación estándar de las genoteca bacteriana Fragmentados por digestión enzimática.

Muestra	Tamaño de Genoteca (pb)
1	477
2	472
3	471
4	476
5	475
Media del tamaño de la genoteca (pb)	474.2
Desviación estándar	2.58
Coefficiente de Variación (%)	0.54
Error Reproducibilidad	1.09
Error Precisión	0.89
Tamaño Genoteca ideal (pb)	470

Con base en estos resultados se puede comprobar la repetibilidad del proceso de la genotecas bacterianas bajo las mismas condiciones, ya que se obtuvo un coeficiente de variación de 0.54 %, es decir, menor a 15 %, por lo tanto se dice que es repetible el proceso de elaboración de genotecas fragmentadas enzimáticamente ⁽⁴⁸⁾.

También se puede afirmar que el proceso de elaboración de genotecas es reproducible, ya que estas genotecas se realizaron en diferente tiempo, en relación a las primeras pruebas donde se estandarizaron las condiciones y se obtuvieron resultados similares.

Por último, con las genotecas elaboradas se realizó la cuantificación de las mismas por medio de una PCR en tiempo real, los resultados se muestran en el Cuadro 7.9.

Cuadro 7.9. Tamaño promedio y concentración de las genotecas bacterianas.

Número de repetición	Tamaño de genoteca (p b)	Concentración (moléculas/ μ L)
1	477	5.6 e 08
2	472	1.49 e09
3	471	3.67 e 08
4	476	7.89 e08
5	475	4.92 e 08
	Promedio de Concentración	7.38E+08
	(moléculas/ μ L)	

Para comprobar que el proceso de elaboración de genotecas bacterianas es repetible, reproducible y que las condiciones que se establecieron para la fragmentación y selección tamaño del ADN son óptimos para elaborar genotecas de un tamaño específico de 450 a 480 pb, para ello se realizó estadística descriptiva que se muestra en el Cuadro 7.10.

Cuadro 7.10. Media y desviación estándar de las genotecas bacterianas fragmentados por digestión enzimática.

Muestra	Tamaño de Genoteca (pb)
<i>S.Muenchen 1</i>	475
<i>S.Muenchen 2</i>	477
<i>S.Muenchen 3</i>	473
<i>S.Muenchen 4</i>	476
<i>S.Muenchen 5</i>	474
<i>S.Muenchen 6</i>	477
<i>S.Muenchen 7</i>	472
<i>S.Muenchen 8</i>	471
<i>S.Muenchen 9</i>	476
<i>S.Muenchen 10</i>	475
Media del tamaño de la genoteca (pb)	474.6
Desviación estándar	2.065591118
Coefficiente de Variación (%)	0.435227796
Error Reproducibilidad	0.870455591
Error Precisión	0.978723404
Tamaño Genoteca ideal (pb)	470

El coeficiente de variación que se obtuvo fue de 0.43, de las diez genotecas elaboradas. Se concluyó que con los dos métodos -enzimático y mecánico por sonicación- las genotecas tienen las condiciones óptimas de tamaño y concentración para realizar una corrida de secuenciación masiva para la plataforma Ion Torrent PGM.

VIII. DISCUSIÓN

El uso de la técnica de secuenciación de nueva generación SNG ha incrementado debido a su gran número aplicaciones en diferentes áreas. Una de las controversias que ha surgido es la de poder garantizar que los resultados de la prueba de SNG son confiables, así como comprobar su validez analítica.

Los estudios al respecto, están enfocados al área clínica y forense ^(29,46,47,49), éstas investigaciones hablan sobre la relevancia de asegurar la calidad de SNG, la importancia de controlar cada una de las etapas del proceso en la parte técnica de la elaboración de genotecas, la emulsión en PCR, la corrida de secuenciación, así como el análisis biofarmático.

Por lo tanto, en este trabajo se estandarizó el proceso de elaboración de genotecas para la plataforma de secuenciación de nueva generación Ion Torrent PGM, la cual es una etapa fundamental y crucial para la SNG, debido a los errores que se pueden generar en su construcción como son: coberturas desiguales, cambios en la secuencia e interrupciones en la secuencia de ADN ^(34, 50).

En este estudio se trabajó con las bacterias *Salmonella* que provenían del Laboratorio de Diagnóstico para la Detección de Organismos Patógenos (LCDDOP); la *Salmonella* es uno de los principales agentes contaminantes de alimentos en nuestro país, razón por la cual se estandarizó el proceso de elaboración de genotecas con este microorganismo. Además, hoy en día la identificación y tipificación bacterias, por medio de la SNG tiene una alta especificidad. ^(14, 15, 16, 17, 18, 51).

El primer paso que se realizó fue la extracción del ADN de las muestras, y se cuantificó la concentración de material genético en el equipo Nano Drop®, cuyo principio se basa en la espectrofotometría, el parámetro que se utilizó para la evaluación de pureza de ADN fue la relación de 260/280 que proporciona el equipo Nano Drop®.

Las muestras tuvieron un cociente de la relación absorbancia de 260/280 nm entre 1.8-2.0, por lo tanto se consideran puras.

El ADN tiene un máximo de absorción de luz UV a la longitud de 260 nm, esto es por la presencia de grupos aromáticos de las bases nitrogenadas a lo largo de la cadena de ADN; si las muestras se encontraran fuera del rango 1.8-2.0, indicaría la presencia de contaminantes provenientes del proceso de extracción tales como proteínas, fenoles, entre otros, que absorben a la longitud de onda de 280 nm ^(52,53).

Pero en este caso las muestras utilizadas se localizaron dentro de los valores considerados de alta pureza, por lo cual se decidió utilizar estas muestras para elaboración de genotecas, además que la concentración que se muestra en el Cuadro 7.1 está arriba de 50 ng/μL, que es la cantidad mínima requerida para la elaboración de genotecas de la plataforma de secuenciación masiva Ion Torrent PGM.

Después de realizar la extracción, se realizó la fragmentación de ADN, proceso indispensable, ya que distintos estudios comunican la importancia de la fragmentación de ADN para la elaboración de genotecas porque puede afectar a la secuenciación ión produciendo errores en la longitud de las lecturas ^(34,35, 50).

En este estudio se establecieron los parámetros de fragmentación enzimática y

mecánica por sonicación. Para la ruptura de ADN por sonicación -que se realiza por medio de ondas de energía acústica- se estableció 300 ciclos y 80 s, se decidió establecer este tiempo porque fue con el que se obtuvieron fragmentos de ADN en un rango de 350 a 470 pb como se muestra en la Figura 7.2, lo cual es óptimo para elaborar genotecas de Ion Torrent PGM para la química de 400 pb.

Se evaluaron tiempos más prolongados, el óptimo fue 80 s. además se reportó en una tesis doctoral de Greendy, en 2014, que un tiempo más prolongado no mejora la ruptura de ADN. En la referencia citada, el autor utilizó la cavitación acústica para la ruptura celular de microalgas para obtener biocombustible - aunque la pared celular es más resistente que el ADN-, observó que a los pocos segundos de sonicación a alta intensidad de frecuencia fija se produjo una disrupción celular significativa y, en tiempos de exposición más largos, la eficacia declinó, debido a una atenuación de energía ultrasónica en la nube resultante de las burbujas de cavitación, lo cual lo determinó mediante mediciones acústicas ^(34, 54, 55).

En el caso de la fragmentación enzimática, la ruptura se realizó por medio de un coctel de endonucleasas inespecíficas que rompen los enlaces fosfodiéster entre los nucleótidos dentro de una cadena ADN; el tiempo que se estableció fue de 7.5 min; como se observa en la Figura 7.3 el promedio del tamaño de los fragmentos que se obtuvo fue de 468 pb ⁽³⁴⁾.

Cuando se establecieron las condiciones para fragmentar el ADN, otro de los parámetros que se estableció para la elaboración de las genotecas fue la selección de tamaño en el equipo *Pippin prep*, que se basa en una electroforesis automatizada, que realiza la separación por medio de carga y masa en un campo

eléctrico.

Para la selección de tamaño se utilizó un casete de gel de agarosa desechable al 2% teñido con bromuro de etidio (es un compuesto fluorescente que se intercala en la molécula de ADN) ⁽⁵⁶⁾.

Se evaluó la calidad de las genotecas por medio de una electroforesis de microfluidos en un chip de alta sensibilidad de ADN. Esta técnica permite conocer la calidad, tamaño y concentración del ADN con una elevada resolución. La técnica emplea chips constituidos por dos laminillas y pozos, con microcanales en el interior, donde se efectúa un proceso de separación electroforética por carga y masa, además de realizar la detección de material genético por medio de fluorescencia inducida por un láser que permite detectar fragmentos ADN con una concentración de hasta 0.5 ng ⁽⁵⁷⁾.

En la Figura 7.7 y 7.11 se observan los gráficos de la relación de la fluorescencia emitida (FU) contra la longitud de los fragmentos de ADN, de cada una de las genotecas. La media obtenida fue de 475 pb para las genotecas fragmentadas por sonicación, esto es por la incorporación de adaptadores de amplificación y secuenciación en cada uno de los extremos del fragmento cuya longitud es de 35 pb.

La concentración de las genotecas se evaluó por medio de la ecuación de la recta generada a partir de la curva estándar, por medio de la metodología SYBR Green® el cual es una molécula intercalante que tiene afinidad por ADN de doble cadena ⁽⁵⁸⁾. En el Cuadro 7.7 y 7.9 observamos una concentración en las genotecas fragmentadas enzimáticamente y por sonicación mayor a 5 millones de moléculas en cada una, la cual es la cantidad mínima necesaria para una corrida

de secuenciación en la plataforma de Ion Torrent PGM⁽³³⁾.

Finalmente, para comprobar la repetibilidad del proceso de la elaboración de genoteca se aplicó un análisis estadístico para obtener la media del tamaño promedio de genotecas, el cual fue de 474.6 pb y el coeficiente de variación (que es una medida de dispersión que hace referencia a la relación entre el tamaño de la media y la desviación) fue de 0.43 %. Como el coeficiente de variación obtenido es menor de 15 %, indica que es repetible el proceso de elaboración de genotecas.⁽⁴⁸⁾.

Por lo tanto, las genotecas que se elaboraron en este trabajo cumplieron con las características óptimas de tamaño y concentración para la secuenciación, ya que el tamaño de las genotecas estaba en un rango de 450 a 480, y presentaron una concentración arriba de 5 millones de moléculas, ambas son condiciones óptimas para una corrida de secuenciación para la plataforma de secuenciación masiva Ion Torrent PGM

IX. CONCLUSIONES

Para la fragmentación mecánica por sonicación se estableció el tiempo de 80 segundos con 300 ciclos de rompimiento, mientras que para la fragmentación enzimática el tiempo de digestión fue de 7.5 minutos.

La fragmentación por medio de sonicación ofrece resultados más precisos y reproducibles, que la fragmentación enzimática de acuerdo a la estadística realizada.

El parámetro de selección de tamaño de fragmentos en las genotecas quedó establecido como un “target” 470 pb, con el que se obtuvieron fragmentos con un tamaño mayor de 450 pb y menor de 480 pb.

El proceso de elaboración de genotecas es repetible y reproducible, porque los coeficientes de variación obtenidos fueron menores del 15% para las 10 genotecas elaboradas.

REFERENCIAS

1. - World Health Organization (WHO) [internet]. (2015) Las enfermedades de transmisión alimentaria [citado el 3 abril 2015]. disponible en http://www.who.int/topics/foodborne_diseases/es/
- 2.-Flores, T. G., & Herrera, R. A. R. Enfermedades transmitidas por alimentos y PCR: prevención y diagnóstico. *Salud pública Méx.* 2005; 47(5), 388-390.
- 3.- Asociación Mexicana de Gastroenterología (AMG). [Internet]. (2012) Enfermedades transmitidas por alimentos; [citado 19 marzo 2015] disponible en: <https://www.gastro.org.mx/gastro-para-todos/?p=86>
- 4.- Food and Drug Administration (FDA). [Internet].U.S. (2014): FDA; [citado 12 febrero 2015]. Disponible en:<http://www.fda.gov/Food/ResourcesForYou/HealthEducators/ucm091976.htm>
- 5.-United States Department of Agriculture (USDA). [Internet].U.S. (2013): USDA; [citado 14 Marzo 2015]. Disponible en:<http://www.fsis.usda.gov/wps/portal/informational/en-espanol/hojasinformativas/enfermedades-por-alimentos>
- 6.- Sistema Nacional de Vigilancia Epidemiológica (SINAVE). [Internet]. México; 2014: [citado 12 marzo 2015]. Boletín Epidemiológico. Disponible en: <http://www.epidemiologia.salud.gob.mx/doctos/boletin/2014/sem14.pdf>
- 7.-Organización Panamericana de la Salud (OPS) [Internet]. Buenos Aires; 1996: [citado 20 abril 2015] Guía Vigilancia Epidemiológica de Enfermedades Transmitidas por alimentos (VETA).Disponible en: <http://www.bvsops.org.uy/pdf/etas.pdf>
- 8.-Enrique G.B. Intoxicaciones por alimentos. [Internet]. 2004; [citado 19 marzo 2015] disponible en: http://salud.edomexico.gob.mx/html/doctos/ueic/educacion/intox_alim.pdf
- 9.- Maribel J. E. La inocuidad de los alimentos en México [internet]. 2013; [Consultado 07 abril 2015] Disponible en: <http://www.infoaserca.gob.mx/claridades/revistas/240/ca240-28.pdf>

- 10.- Sistema Nacional de Vigilancia Epidemiológica (SINAVE). [Internet]. México; 2015: [citado 12 abril 2015]. Boletín Epidemiológico. Disponible en:<http://www.epidemiologia.salud.gob.mx/doctos/boletin/2015/sem14.pdf>
- 11.- Sistema Nacional de Vigilancia Epidemiológica (SINAVE). [Internet]. México; 2012: [citado 15 abril 2015]. Perfil Epidemiológico de las Enfermedades Infecciosas Intestinales. Disponible en:
http://www.epidemiologia.salud.gob.mx/doctos/infoepid/publicaciones/2012/Monografias4_Enf_Infecciosas_Intestinales_%20junio12.pdf
- 12.-Sistema Único de Información para la Vigilancia Epidemiológica (SUIVE). [Internet]. México; 2011: [citado 15 abril 2015]. Información Epidemiológica de Morbilidad. Disponible en:
http://www.epidemiologia.salud.gob.mx/doctos/infoepid/publicaciones/2012/ver_ejecutiva_2011.pdf
- 13.- Sistema Único de Información para la Vigilancia Epidemiológica (SUIVE). [Internet]. México; 2013: [citado 20 abril 2015]. Incidencia de Intoxicación alimentaria bacteriana México 2010. Disponible en:
<http://www.epidemiologia.salud.gob.mx/anuario/html/anuarios.html>
- 14.-Bou, G., Fernández-Olmos, A., García, C., Sáez-Nieto, J. A., & Valdezate, S. Métodos de identificación bacteriana en el laboratorio de microbiología. *Enferm Infecc Microbiol Clin*. 2011; 29(8), 601-608.
- 15.-Reller, L. B, Weinstein, M. P., & Petti, C. A. Detection and identification of microorganisms by gene amplification and sequencing. *Clin Infect Dis*. 2007; 44(8), 1108-1114.
- 16.-Lan, R., & Reeves, P. R. Escherichia coli in disguise: molecular origins of Shigella. *Microbes Infect*. 2002; 4(11), 1125-1132.
- 17.- Del Rosario Rodicio, M., & del Carmen Mendoza, M. Identificación bacteriana mediante secuenciación del ARNr 16S: fundamento, metodología y aplicaciones en microbiología clínica. *Enferm Infecc Microbiol Clin*. 2004; 22(4), 238-245.
- 18.- Medini, D., Serruto, D., Parkhill, J., Relman, D. A., Donati, C., Moxon, R., & Rappuoli, R. Microbiology in the post-genomic era. *Nat Rev Microbiol*. 2008; 6(6), 419-430.

- 19.-Vázquez, J. A., & Berrón, S. (2004). Multilocus sequence typing: el marcador molecular de la era de Internet. *Enferm Infecc Microbiol Clin*.2004; 22(2), 113-120.
- 20.- Urwin, R., & Maiden, M. C. Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol*.2003; 11(10), 479-487.
- 21.-Tiessen Favier Alex. Fundamentos y Metodologías Innovadoras para el Mejoramiento Genético de Maíz. Fundación Ciencia Activa. Cinvestav Irapuato .México.2009
- 22.-Rodríguez-Santiago, B., & Armengol, L. Tecnologías de secuenciación de nueva generación en diagnóstico genético pre-y postnatal. *Diagn Prenat*. 2012; 23(2), 56-66.
- 23.-Barba, M., Czosnek, H., & Hadidi, A. Historical perspective, development and applications of next-generation sequencing in plant virology. *Viruses*. 2014; 6(1), 106-136.
- 24.- Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., ... & Law, M. Comparison of next-generation sequencing systems. *Biomed Res Int*. 2012; 111
- 25.-Muñoz, R. C. ¿Cómo funciona?: Aplicaciones de las nuevas tecnologías de secuenciación. *Encuentros en la Biología*. 2012; 5(138), 41-42.
26. - Watson, J. D., & Crick, F. H. C. Molecular structure of nucleic acids. *Nature*. 1953; (4356), 737–738.
27. - Handelsman, J., Rondon, M., Brady, S., Clardy, J., & Goodman, R. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol*.1998; 5(10), 245–249.
28. - Wooley, J. C., Godzik, A., & Friedberg, I. A primer on metagenomics. *PLOS Comput Biol*. 2010; 6(2),
- 29.-Berglund, E. C., Kiialainen, A., & Syvänen, A. C. Next-generation sequencing technologies and applications for human genetic history and forensics. *Investig Genet*. 2011; 2, 23.
- 30.-Merriman, B., Torrent, I., Rothberg, J. M., & R&D Team. Progress in ion torrent semiconductor chip based sequencing. *Electrophoresis*.2012; 33(23), 3397-3417.

- 31.- Ion Torrent plataforma. [internet]. [Consultado 28 abril 2015] Disponible en http://www.genomics.cn/en/navigation/show_navigation?nid=2640
- 32.- Eberhard, P. (2003). GENÉTICA texto y atlas. *Editorial Panamericana*.
- 33.- Life Technologies™. *Ion Xpress™ Plus gDNA Fragment Library Preparation. 2014.*
- 34.- Head, S. R., Komori, H. K., LaMere, S. A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D.R., & Ordoukhanian, P. Library construction for next-generation sequencing: overviews and challenges. *Biotechniques*. 2013; 56 (2), 61-4.
35. - Knierim, E., Lucke, B., Schwarz, J. M., Schuelke, M., & Seelow, D. Systematic comparison of three methods for fragmentation of long-range PCR products for next generation sequencing. *PloS One*. 2011; 6(11).
- 36.- World Health Organization (WHO) [internet]. (2015). *Salmonella (no tifoidea)* [citado el 3 abril 2015]. Disponible en <http://www.who.int/mediacentre/factsheets/fs139/es>
- 37.- Echeita, M. A., Aladueña, A. M., Díez, R., Arroyo, M., Cerdán, F., Gutiérrez, R, & Usera, M. Á. Distribución de los serotipos y fagotipos de *Salmonella* de origen humano aislados en España en 1997-2001. *Enferm Infecc Microbiol Clin*. 2005; 23(3), 127-134.
- 38.- Cornell University. [internet]. (2015). *Salmonella Muenchen*. [citado el 19 abril 2015]. disponible en <https://confluence.cornell.edu/display/FOODSAFETY/Salmonella+Muenchen>
- 39.- Cornell University. [internet]. (2013). *Salmonella Kentucky*. [citado el 01 abril 2015]. disponible en <https://confluence.cornell.edu/display/FOODSAFETY/Salmonella+Kentucky>
- 40.- Cornell University. [internet]. (2013). *Salmonella Enteritidis*. [citado el 21 abril 2015]. Disponible en: <https://confluence.cornell.edu/display/FOODSAFETY/Salmonella+Enteritidis>
- 41.- Chen, T. L., Thien, P. F., Liaw, S. C., Fung, C. P., & Siu, L. K. First report of *Salmonella enterica* serotype Panama meningitis associated with consumption of contaminated breast milk by a neonate. *J Clin Microbiol*. 2005; 43(10), 5400-5402.

- 42.- Gutiérrez-Cogco, L., Montiel-Vázquez, E., Aguilera-Pérez, P., & González-Andrade, M. C. Serotipos de Salmonella identificados en los servicios de salud de México. *Salud Pública Méx.*2000; 42(6), 490-495.
- 43.- Centers for Disease Control and Prevention (CDC). (2011). Multistate outbreak of Salmonella Panama infections linked to Cantaloupe. Disponible en: <http://www.cdc.gov/salmonella/panama0311/062311/>
- 44.-Instituto de Diagnóstico y Referencia Epidemiológicos(InDRE) [internet]. (2012). Coordinacion de Referencia Molecular. Disponible en http://www.indre.salud.gob.mx/interior/coordinacion_referencia_molecular.htm
- 45.-Hugo.C.S: Búsqueda de cuasiespecies virales en el gen NA del virus de influenza H3N2 por medio de pirosecuenciación masiva paralela[Tesis]. UNAM. Facultad de Estudios Superiores Cuautitlán. 2011.
- 46.- Gargis, A. S., Kalman, L., Berry, M. W., Bick, D. P., Dimmock, D. P., Hambuch, T., & Lubin, I. M. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat biotechnol.*2012; 30(11), 1033-1036
- 47.-Ortiz.M.F, Mazzanti A. Validación clínica y analítica de una librería de secuenciación masiva para el diagnóstico genético de las cardiopatías familiares.*Rev Esp Cardiol.* 2012.
48. - Wille, S. M., Peters, F. T., Di Fazio, V., & Samyn, N. Practical aspects concerning validation and quality control for forensic and clinical bioanalytical quantitative methods. *Accred Qual Assur.* 2011; 16(6), 279-292
- 49 - Wong, L. J. C. (2013). *Next Generation Sequencing: Translation to Clinical Diagnostics*. Springer Science & Business Media.
- 50.- Robasky, K., Lewis, N. E., & Church, G. M. (2014). The role of replicates for error mitigation in next-generation sequencing. *Nat Rev Genet.*2014;15(1), 56-62.
- 51.-Sherry, N. L., Porter, J. L., Seemann, T., Watkins, A., Stinear, T. P., & Howden, B. P. Outbreak investigation using high-throughput genome sequencing within a diagnostic microbiology laboratory. *J Clin Microbiol.*2013; 51(5), 1396-1401.
- 52.-Scientific, T. F. NanoDrop 1000 Spectrophotometer V3. 7 User's Manual. *Thermo Fisher Scientific.*2008; 1-105.

- 53.-Wilfinger, W. W., Mackey, K., & Chomczynski, P. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *Biotechniques*.1997; 22(3), 474-6.
- 54-2.-Justin Greenly. Processing Of Microalgae: Acoustic Cavitation And Hydrothermal Conversion, [tesis doctoral].Universidad de Cornell Unversity.2014
- 55-3.-Universidad de Barcelona. [internet]. (2013). Activación de reacciones químicas mediante ultrasonidos. Disponible en <http://www.ub.edu/talq/es/node/252>
- 56.- Lee, P. Y., Costumbrado, J., Hsu, C. Y., & Kim, Y. H. Agarose gel electrophoresis for the separation of DNA fragments. *J Vis Exp*.2011; (62), 9400-9405.
- 57.- Castagnino, J. M. Nanotecnología, microchips y microarreglos. *Acta Bioquím Clín Latinoam*.2006; 40(2), 161-163.
- 58.- Tamay de Dios, L., Ibarra, C., & Velasquillo, C. Fundamentos de la reacción en cadena de la polimerasa (PCR) y de la PCR en tiempo real.*Invest Discapacidad*.2013; 2(2), 70-78.