



Universidad Autónoma de Sinaloa
Facultad de Ciencias Químico Biológicas
Programa Regional de Posgrado en Biotecnología
Maestría en Ciencias con Orientación en Biotecnología

**Estudio Filogenético y Caracterización
Genómica de la Virulencia de *Salmonella*
enterica Aisladas de Agua de Ríos y
Animales de Granja en Culiacán, Sinaloa**

T E S I S

que presenta

IBT. José Roberto Aguirre Sánchez

como requisito para

obtener el grado de

**Maestro en Ciencias con Orientación
en Biotecnología de la Ambiental**

Director de Tesis

Dra Johana Marcela Soto Beltrán

Dr Juan Ramón Ibarra Rodríguez

Culiacán Rosales, Sinaloa, México

Marzo 2018

PRESENTACIÓN

La presente investigación, titulada “Estudio Filogenético y Caracterización Genómica de la Virulencia de *Salmonella enterica* Aisladas de Agua de Ríos y Animales de Granja en Culiacan, Sinaloa” se llevó a cabo en las instalaciones del Parque de Innovación Tecnológica (PIT) y en los Laboratorios del Programa Regional de Posgrado en Biotecnología , Facultad de Ciencias Químico Biológicas (FQCB), Universidad Autónoma de Sinaloa (UAS). Como Asesores Académicos participaron: Dra. Johana Marcela Soto Beltrán (Directora de tesis), Dr. Juan Ramón Ibarra Rodríguez (Director de tesis), Dra. Bianca Anabel Amézquita López, Dr. Inés Fernando Vega López y el Dr. Jaime Martínez Urtaza. Este trabajo de investigación conforma la base de trabajos que se desarrollarán posteriormente y tendrán como finalidad el estudio “Variaciones Genéticas en *Salmonella enterica* de Muestras Aisladas de Agua de Río y de Animales de Granja de Culiacán, Sinaloa”. La investigación tuvo carácter interinstitucional, participando el Centro de Investigación en Alimentación y Desarrollo, A. C. (CIAD), Universidad de Bath, PIT y la Universidad Autónoma de Sinaloa. El C IBT José Roberto Aguirre Sánchez recibió “Beca Nacional para Estudios de Posgrado” por el Consejo Nacional de Ciencia y Tecnología (CONACyT).

AGRADECIMIENTOS

Agradezco a Dios por los dones y las capacidades que me ha dotado. Por darme vida y fuerza para cumplir mis metas y sueños.

A mis padres, que me han apoyado en cada etapa de mi vida, me han alentado a ser mejor persona y me han dado el mejor ejemplo que han podido. A mi familia, que han estado siempre al cuidado de mí, por la fe que han tenido siempre en mí y por el apoyo incondicional que siempre he recibido de su parte.

A los doctores: Dra. Marcela Soto, Dr. Juan Ramón Ibarra, Dra. Bianca Amézquita, Dr. Inés Vega y Dr. Jaime Martínez, por su paciencia, las facilidades proporcionadas, su apoyo, disponibilidad y por su valiosa guía, en la aventura de darle forma y vida a este proyecto de investigación; específicamente la elaboración de ésta tesis.

A la Universidad Autónoma de Sinaloa, por darme la oportunidad de cumplir una meta más en mi vida y subir un peldaño en mi preparación académica. Por permitirme gozar sus instalaciones y vivir momentos increíbles.

Al Laboratorio Nacional para la Investigación en Inocuidad Alimentaria (LANIIA), por permitirme desarrollar mi proyecto de investigación y por el uso de sus instalaciones.

A mis maestros, por su dedicación en la loable tarea que es la educación; permitiéndome expandir mis horizontes a través de la curiosidad, creatividad y la formulación de nuevas preguntas.

A mis amigos y a las personas que he conocido antes y durante de la realización de este proyecto de investigación, por su apoyo, afecto y compañerismo a través de este camino.

ÍNDICE GENERAL

| | Pág. |
|--|------|
| ÍNDICE DE CUADROS | 8 |
| ÍNDICE DE FIGURAS | 10 |
| I RESUMEN | 13 |
| ABSTRACT | 15 |
| II INTRODUCCIÓN | 17 |
| III REVISIÓN DE LA LITERATURA | 20 |
| A. Características del género <i>Salmonella</i> . | 20 |
| 1 Características biológicas. | 20 |
| a Ambientes hospederos. | 21 |
| b Ambientes no hospederos. | 22 |
| 2 Características genéticas. | 25 |
| 3 Características morfológicas. | 26 |
| B. Taxonomía del género <i>Salmonella</i> . | 26 |
| C. Enfermedades ocasionadas por <i>Salmonella</i> . | 31 |
| D. Patogenicidad de <i>Salmonella</i> . | 34 |
| 1 Virulencia. | 35 |
| 2 Sistema de secreción tipo III (T3SS). | 41 |
| E. Métodos moleculares para la caracterización de patógenos. | 44 |
| 1 Reacción en cadena de la polimerasa (PCR). | 44 |
| 2 Electroforesis en gel de campo pulsado (PFGE) | 45 |
| F. Análisis molecular de bacterias patogénicas mediante biología computacional. | 49 |
| 1 Plataforma de secuenciación Illumina. | 49 |
| 2 Herramientas bioinformáticas para el análisis de secuencias de ácidos nucleicos. | 54 |
| a Ensamblado de fragmentos de ADN. | 54 |
| b Anotación de genomas bacterianos. | 59 |
| c Construcción de árboles filogenéticos. | 59 |
| IV HIPÓTESIS | 61 |
| V JUSTIFICACIÓN | 62 |
| VI OBJETIVOS | 63 |
| A. OBJETIVO GENERAL | 63 |
| B. OBJETIVOS ESPECÍFICOS | 63 |
| VII MATERIALES Y MÉTODOS | 64 |

| | |
|---|-----|
| A. MATERIALES | 64 |
| B. MÉTODOS | 64 |
| 1. Secuenciación de genomas. | 64 |
| 2. Procesamiento de las secuencias (ensamble). | 64 |
| 3. Descripción de la nomenclatura. | 65 |
| 4. Serotipificación <i>in silico</i> . | 67 |
| 5. Análisis del pan-genoma. | 67 |
| 6. Construcción de árboles filogenéticos (fast-tree) | 67 |
| 7. Inferencias filogenéticas. | 69 |
| 8. Identificación de islas de patogenicidad y factores de virulencia. | 69 |
| 9. Identificación de secuencias pertenecientes a profagos en genomas locales. | 70 |
| VIII RESULTADOS Y DISCUSIONES | 71 |
| A. DIVERSIDAD DE LOS SEROTIPOS DE <i>Salmonella</i> AISLADAS EN LA REGIÓN | 71 |
| B. RELACIONES FILOGENÉTICAS DE LOS SEROTIPOS DOMINANTES EN LA REGIÓN CON RESPECTO A LAS REPORTADAS EN DIFERENTES PAÍSES | 80 |
| 1 Filogenia global del serotipo Oranienburg. | 85 |
| 2 Filogenia global del serotipo Saintpaul. | 88 |
| 3 Filogenia global del serotipo Give. | 91 |
| 4 Filogenia global del serotipo Minnesota. | 92 |
| 5 Filogenia global del serotipo Anatum. | 97 |
| C. PREDICCIÓN DE GENES EN LOS AISLADOS DE <i>Salmonella</i> . | 100 |
| 1 Anotación de <i>Salmonella enterica</i> Serotipo Anatum. | 101 |
| 2 Anotación de <i>Salmonella enterica</i> Serotipo Give. | 103 |
| 3 Anotación de <i>Salmonella enterica</i> Serotipo Saintpaul. | 103 |
| 4 Anotación de <i>Salmonella enterica</i> Serotipo Minnesota. | 106 |
| 5 Anotación de <i>Salmonella enterica</i> Serotipo Oranienburg. | 106 |
| D. CARACTERIZACIÓN DE LOS FACTORES DE VIRULENCIA | 111 |
| E. PAN-GENOMA | 118 |
| 1 Pan-genoma de las cepas de <i>Salmonella enterica</i> Serovar Minnesota. | 118 |
| 2 Pan-genoma de las cepas de <i>Salmonella</i> | 123 |

| | | |
|-----------|--|-----|
| | <i>enterica</i> Serovar Saintpaul | |
| | 3 Pan-genoma de las cepas de <i>Salmonella enterica</i> Serovar Anatum. | 127 |
| | 4 Pan-genoma de las cepas de <i>Salmonella enterica</i> Serovar Give. | 131 |
| | 5 Pan-genoma de las cepas de <i>Salmonella enterica</i> Serovar Oranienburg. | 135 |
| | F. APORTE DE LOS FAGOS A LA VIRULENCIA DE <i>Salmonella</i> . | 140 |
| IX | CONCLUSIONES | 147 |
| X | BIBLIOGRAFÍA | 149 |
| XI | ANEXOS | 164 |
| | Abreviaturas | 169 |

ÍNDICE DE CUADROS

| Cuadro | Descripción | Pág. |
|--------|--|------|
| 1 | Sub-especies de <i>Salmonella enterica</i> . | 29 |
| 2 | Principales causas de enfermedades transmisibles en México. | 36 |
| 3 | Entidades federativas con número de casos reportados de salmonelosis e infecciones intestinales por otros organismos y mal definidas en el 2016. | 37 |
| 4 | Descripción de las islas de patogenicidad en <i>Salmonella</i> . | 39 |
| 5 | Abreviación de los serotipos utilizados para renombrar cepas de la región y las encontradas en NCBI. | 68 |
| 6 | Aislados que mostraron incongruencia al ser serotificados por técnicas convencionales y técnicas <i>in silico</i> . | 72 |
| 7 | Distribución de serotipos de los aislados de <i>Salmonella</i> . | 74 |
| 8 | Serotipos predominantes en la región con su ST y perfil alélico. | 79 |
| 9 | Relación de las cepas locales e internacionales de los diferentes serotipos de <i>Salmonella enterica</i> . | 83 |
| 10 | Perfil alélico de los diferentes grupos ST analizados del serotipo Oranienburg. | 87 |
| 11 | Perfil alélico de los diferentes grupos ST analizados del serotipo Saintpaul. | 90 |
| 12 | Perfil alélico de los diferentes grupos ST analizados del serotipo Give. | 94 |
| 13 | Perfil alélico de los diferentes grupos ST analizados del serotipo Minnesota. | 96 |
| 14 | Perfil alélico de los diferentes grupos ST analizados del serotipo Anatum. | 99 |
| 15 | Resumen de los datos obtenidos mediante la anotación de genomas pertenecientes a los serotipos predominantes en la región. | 110 |
| 16 | Resumen de la composición del pan-genoma de cepas locales de <i>Salmonella</i> serotipo Minnesota. | 122 |
| 17 | Resumen de la composición del pan-genoma de cepas locales de <i>Salmonella</i> serotipo Saintpaul. | 126 |
| 18 | Resumen de la composición del pan-genoma de cepas locales de <i>Salmonella</i> para el serotipo Anatum. | 130 |

| | | |
|----|--|-----|
| 19 | Resumen de la composición del pan-genoma de cepas locales de <i>Salmonella</i> para el serotipo Give. | 134 |
| 20 | Resumen de la composición del pan-genoma de cepas locales de <i>Salmonella</i> para el serotipo Oranienburg. | 138 |
| 21 | Porcentaje de los fagos encontrados en los aislados de <i>Salmonella</i> . | 141 |
| 22 | Comparación de la serotipificación convencional vs <i>in silico</i> para las cepas locales. | 164 |

ÍNDICE DE FIGURAS

| Fig | Descripción | Pág. |
|------------|--|-------------|
| 1 | Ruta de dispersión de patógenos. | 23 |
| 2 | Ruta de transmisión para infecciones fecales-orales. | 24 |
| 3 | Pan-genoma representado mediante un diagrama de Venn. | 28 |
| 4 | Localización de antígenos O, H y Vi para la serotipificación de <i>Salmonella</i> . | 30 |
| 5 | Agentes causales de toxiinfección alimentaria en la Unión Europea. | 32 |
| 6 | Estructura del sistema de secreción tipo III en bacterias patogénicas. | 43 |
| 7 | Etapas de la PCR. | 47 |
| 8 | Diagrama de flujo del proceso de electroforesis en gel de campos pulsantes (PFGE). | 48 |
| 9 | Aplicaciones relevantes en genómica comparativa. | 50 |
| 10 | PCR por puente (bridge) en la plataforma de secuenciación Illumina Miseq. | 52 |
| 11 | Diagrama de flujo del proceso general de secuenciación de ácidos nucleicos en plataformas Illumina. | 53 |
| 12 | Panorama general de las aplicaciones bioinformáticas. | 55 |
| 13 | Diagrama para el ensamble de los reads. | 57 |
| 14 | Heatmap del porcentaje de identidad entre los aislados de la región de Culiacán, Sinaloa. | 76 |
| 15 | Árbol filogenético para las cepas locales construido por el algoritmo de máxima verosimilitud. | 78 |
| 16 | Filogenia global de los serotipos Oranienburg, Give, Minnesota, Saintpaul y Anatum de <i>Salmonella enterica</i> subsp. <i>enterica</i> dominantes en la región. | 82 |
| 17 | Árbol filogenético con aislados locales e internacionales de <i>Salmonella enterica</i> serotipo Oranienburg. | 86 |
| 18 | Árbol filogenético con aislados locales e internacionales de <i>Salmonella</i> | 89 |

| | | |
|----|---|-----|
| | <i>enterica</i> serotipo Saintpaul. | |
| 19 | Árbol filogenético con aislados locales e internacionales de <i>Salmonella enterica</i> serotipo Give. | 93 |
| 20 | Árbol filogenético con aislados locales e internacionales de <i>Salmonella enterica</i> serotipo Minnesota. | 95 |
| 21 | Árbol filogenético con aislados locales e internacionales de <i>Salmonella enterica</i> serotipo Anatum. | 98 |
| 22 | Genoma anotado de la cepa <i>Salmonella enterica</i> serotipo Anatum. | 102 |
| 23 | Genoma anotado de la cepa <i>Salmonella enterica</i> serotipo Give. | 104 |
| 24 | Genoma anotado de la cepa <i>Salmonella enterica</i> serotipo Saintpaul. | 105 |
| 25 | Genoma anotado de la cepa <i>Salmonella enterica</i> serotipo Minnesota. | 107 |
| 26 | Genoma anotado de la cepa <i>Salmonella enterica</i> serotipo Oranienburg. | 108 |
| 27 | Frecuencia de las islas de patogenicidad de <i>Salmonella</i> en los distintos serotipos. | 112 |
| 28 | Heatmap de presencia y ausencia de genes de virulencia para los aislados locales. | 117 |
| 29 | Proporción de genes nuevos y únicos para las cepas locales de <i>Salmonella enterica</i> serotipo Minnesota. | 119 |
| 30 | Gráfica del pan-genoma y del genoma core de las cepas locales de <i>Salmonella enterica</i> serotipo Minnesota. | 121 |
| 31 | Proporción de genes nuevos y únicos para las cepas locales de <i>Salmonella enterica</i> serotipo Saintpaul. | 124 |
| 32 | Gráfica del pan-genoma y del genoma core de las cepas locales de <i>Salmonella enterica</i> serotipo Saintpaul. | 125 |
| 33 | Proporción de genes nuevos y únicos para las cepas locales de <i>Salmonella enterica</i> serotipo Anatum. | 128 |
| 34 | Gráfica del pan-genoma y del genoma core de las cepas locales de <i>Salmonella enterica</i> serovar Anatum. | 129 |
| 35 | Proporción de genes nuevos y únicos para las cepas locales de | 130 |

| | | |
|----|---|-----|
| | <i>Salmonella enterica</i> serotipo Give. | |
| 36 | Gráfica del pan-genoma y del genoma core de las cepas locales de <i>Salmonella enterica</i> serotipo Give. | 133 |
| 37 | Proporción de genes nuevos y únicos para las cepas locales de <i>Salmonella enterica</i> serotipo Oranienburg. | 136 |
| 38 | Gráfica del pan-genoma y del genoma core de las cepas locales de <i>Salmonella enterica</i> serotipo Oranienburg. | 137 |
| 39 | Factores de virulencia aportados por los 50 fagos en los serotipos dominantes en la región. | 142 |
| 40 | Gráfica de pastel con los fagos encontrados en los serotipos dominantes en la región. | 145 |

I. RESUMEN

Salmonella es uno de los agentes infecciosos con mayor importancia en el mundo, manteniendo una alta prevalencia en México. Es transmitida por alimentos contaminados o por vía fecal-oral y es responsable de enfermedades como la fiebre tifoidea y la salmonelosis no tifoidea. La identificación oportuna, temprana y específica de estos patógenos promueve su control y erradicación oportuna. Lo que impactará en vigilancia y control epidemiológico, favoreciendo a los sectores salud, económico y social. Este trabajo tiene como objetivo determinar la diversidad genética poblacional de *Salmonella enterica* ambiental e identificar rasgos genéticos de virulencia en aislados de agua de río y animales de granja en Culiacán, Sinaloa, mediante secuenciación masiva y biología computacional. Para ello se construyeron árboles filogenéticos mediante máxima verosimilitud, se serotipificaron los genomas por determinación *in silico* de marcadores antigénicos, se realizaron alineamientos múltiples de genomas y se identificaron marcadores de virulencia. Los resultados mostraron poca diversidad genética entre cepas locales del mismo serotipo; adicionalmente, se encontró alta congruencia en la agrupación de los aislados en clados específicos a su serotipo. Las islas de patogenicidad SPI-1, SPI-2 y SPI-3 se encuentran en el 90% de los aislados, representando un riesgo potencial a la salud pública, mientras que las islas SPI-4 y SPI-5 se encuentran principalmente en aislados del serotipo Saintpaul y Anatum, sugiriendo un mayor potencial patogénico en dichas cepas. Adicionalmente se detectó que *Salmonella* Oranienburg y Give presentan pan-genomas

tentativamente abiertos, lo que indica que son más propensas a sufrir modificaciones genéticas. Se encontró que los fagos tienen un aporte del 9.6% a la virulencia de *Salmonella* confiriendo genes de adherencia e invasión celular. Aquí se muestran las relaciones filogenéticas y la caracterización de la virulencia de genomas de *Salmonella*. Demostrando la eficiencia y versatilidad de herramientas bioinformáticas en biología computacional para la detección, discriminación y caracterización de patógenos.

Palabras clave: *Salmonella*, filogenia, factores de virulencia, fagos.

ABSTRACT

Salmonella is one of the most important infectious agents in the world, maintaining a high prevalence in Mexico. It is transmitted by contaminated food or through the fecal-oral route, and it is responsible for diseases such as typhoid fever and non-typhoidal salmonellosis. The early and specific identification of these pathogens promotes their opportune control and eradication, which will impact on epidemiological surveillance and control, favoring the health, economic and social sectors. The objective of this work was to determine the genetic diversity of *Salmonella enterica* environmental populations and to identify genetic traits of virulence in isolates from river water and farm animals feces samples from Culiacan, Sinaloa, through massive sequencing and computational biology. To this end, phylogenetic trees were constructed by maximum likelihood, genomes were serotyped by *in silico* determination of antigenic markers, multiple genome alignments were made and virulence markers were identified. The results show little genetic diversity among local strains of the same serotype; additionally, high congruence was found in the clustering of genome sequences according to their serotype. The pathogenicity islands SPI-1, SPI-2 and SPI-3 are found in 90% of the isolates, representing a potential risk to public health, while the SPI-4 and SPI-5 islands are mainly found in isolated of serotype Saintpaul and Anatum, suggesting a greater pathogenic potential in such strains. Additionally, it was detected that *Salmonella* Oranienburg and Give present tentatively open pangenomes, indicating that they are more prone to undergo genetic modifications. It was found that phages have a 9.6% contribution to the virulence of *Salmonella*,

conferring adhesion genes and cellular invasion. Here we show the phylogenetic relationships and characterization of the virulence of *Salmonella* genomes. Demonstrating the efficiency and versatility of bioinformatics tools in computational biology for the detection, discrimination and characterization of pathogens.

Keywords: *Salmonella*, phylogeny, virulence factors, fagos

II. INTRODUCCIÓN

Salmonella es una bacteria que pertenece a la familia Enterobacteriaceae y está clasificada como gram-negativa (Jurado, 2015). Se les conoce como bacterias entéricas, debido a su localización habitual como saprófitos en el tubo digestivo; aunque se trata de bacterias que también pueden encontrarse de forma universal en el suelo, el agua y la vegetación (Puerta-García y Mateos-Rodríguez, 2006).

Históricamente, esta bacteria se ha asociado al consumo de alimentos contaminados de origen animal; principalmente en huevos, carne, aves y leche. Sin embargo, también se han detectado brotes de salmonelosis relacionados al consumo de productos vegetales frescos contaminados (Pui y col. 2011). Otro de los mecanismos de infección es la ingesta por vía fecal-oral, la cual ha ocasionado enfermedades como la fiebre tifoidea (*Salmonella Typhi*) y la salmonelosis no tifoidea (*Salmonella Enteritidis* y *Salmonella Typhimurium*); siendo esta última, una de las enfermedades de transmisión alimentaria (ETA) más comunes, encontrándose ampliamente extendida a nivel mundial (OMS, 2013). Dicha enfermedad se caracteriza por fiebres encima de 40°C, dolor abdominal, diarrea, náuseas e inclusive vómitos (Gómez, 2002).

Salmonella es una bacteria que se encuentra distribuida a nivel mundial y con alta prevalencia en México. Esto se debe a que el microorganismo posee condiciones de crecimiento amplias, con una temperatura que va desde 5.2°C hasta los 46.2°C y un pH de 3.8 a 9.5; lo que lo posibilita a resistir en diferentes ambientes. Esta bacteria se encuentra mayoritariamente en los países cercanos al Ecuador, ya que

estos exhiben temperaturas cercanas a 37°C, en la cual *Salmonella* crece adecuadamente (Pascual, 2005).

La evolución de estos genomas bacterianos se caracteriza por una cantidad masiva de inserciones, deleciones y reordenamientos; que permiten la diferenciación y adaptación de los linajes emparentados, promoviendo su supervivencia en una vasta diversidad de ambientes (Romero y Palacios, 1997). La gran cantidad de serotipos que presenta esta bacteria, se debe a que posee genomas altamente dinámicos que adquieren, pierden y reordenan información genética relevante, lo que hace que sean ecológica y fenotípicamente muy diversos (Dutta y Pan, 2002). Dando lugar a cepas con variabilidad genética diversa. Lo que ha servido como pauta para la selección natural, donde las cepas de *Salmonella* más aptas han logrado perpetuar, a través de las mutaciones (cualquier alteración permanente en la secuencia de ADN) y adaptaciones que ha logrado acumular (Health, 2016).

Por muchos años, las técnicas electroforéticas han prevalecido en el estudio de las mutaciones y la identificación de polimorfismos (cambios en una base en la secuencia de ADN). Pero debido al desarrollo, del Proyecto del Genoma Humano (PGH), ha habido una demanda en métodos que mejoren la rapidez, para el descubrimiento y caracterización de los polimorfismos entre los genomas mapeados y los genes de interés. Siendo importantes para explorar la diversidad genética y analizar marcadores moleculares (Walker y Raply, 2009).

La secuenciación masiva de ADN o secuenciación de siguiente generación (NGS), para el análisis genómico de secuencias, ha impulsado dramáticamente la

investigación biológica, facilitando y acelerando el análisis de polimorfismos, el establecimiento de relaciones filogenéticas, la identificación de islas de patogenicidad y la resistencia a antibióticos, entre otras aplicaciones (Shendure y Ji, 2008). Como resultado del rápido desarrollo tecnológico, se ha logrado reducir costos de procesamiento, que a su vez ha estimulado la amplia distribución y uso de estas herramientas entre la comunidad científica (Baker C., 2012).

El uso de las NGS en el análisis genómico, implica el manejo de enormes volúmenes de datos que crecen de manera exponencial en tamaño y complejidad (por la precisión de las NGS). Afortunadamente, el desarrollo tecnológico en el ámbito de la electrónica y en el desarrollo de software, han permitido un avance significativo en las técnicas para el procesamiento y análisis eficiente de los datos (Escobar y col. 2011); es así como surge la Bioinformática, que mezcla tanto herramientas biológicas, como informáticas. Esta nueva área, abarca específicamente el desarrollo de bases de datos o conocimiento para almacenar y recuperar datos biológicos, algoritmos para analizar y determinar sus relaciones con los datos biológicos y las herramientas estadísticas para identificar e interpretar conjunto de datos (Cañedo y Arencibia, 2004).

Esta investigación pretende hacer uso de la bioinformática como una herramienta para el análisis de genomas de aislados ambientales de *Salmonella*. Realizando la secuenciación y tipificación de cepas locales aisladas de agua de río y animales de granja, para posteriormente analizar polimorfismos entre las cepas, establecer relaciones filogenéticas e identificar genes responsables de la virulencia.

III. REVISIÓN DE LA LITERATURA

A. Características del género *Salmonella*

El género *Salmonella* pertenece a la familia Enterobacteriaceae y se encuentra dentro del grupo de las bacterias entéricas; en las que se encuentran también *E. coli*, las especies de *Shigella* (Baker y Dougan, 2007), entre otras. *Salmonella* es de los principales responsables de causar enfermedades por la ingesta de alimentos contaminados (CDC, 2013). Se localiza de forma natural en el intestino de humanos y animales de sangre caliente, caracterizándose por infectar a un amplio rango de hospederos (Winfield y Groisman, 2003). Este género fue nombrado así en honor al veterinario Daniel E. Salmon quien descubrió esta bacteria en 1855, en el intestino de cerdos infectados con cólera, posteriormente dicha bacteria fue nombrada *Salmonella choleraesuis* (Craig y James, 2006). Desde entonces la nomenclatura del género *Salmonella* ha sufrido numerosos cambios hasta llegar a la estructura actual la cual es controversial.

1 Características biológicas

Salmonella es uno de los agentes patogénicos aislados con mayor frecuencia de alimentos como huevos, productos lácteos y carnes crudas (Silva y col. 2011); aunque de forma general, la comida proveniente de animales como el puerco, aves de corral y de res es la primera fuente de infecciones por *Salmonella*. La mayoría de los serotipos encontrados en el género *Salmonella*, tienen la capacidad de adaptarse a una variedad de huéspedes animales, incluido el humano (Nielsen, 2013). La principal ruta de diseminación de este patógeno involucra el consumo de alimentos crudos de origen animal, también se ha

determinado que el proceso de sacrificio en mataderos es una fuente importante para la contaminación de órganos con *Salmonella* (Gillespie y col. 2005).

a) Ambientes hospederos

Salmonella incluye una gran diversidad de serotipos, los cuales se encuentran principalmente distribuidos en hospederos de sangre caliente, peces, aves, anfibios y reptiles (Shu-Kee y col. 2015). Sin embargo, hay ocasiones donde un serotipo presentan predilección por un hospedero en específico, por ejemplo *S. Typhi*, *S. Paratyphi A*, *B* y *C* (solo afectan humanos) (Huang y col. 2005). En el caso de aves de corral se encuentran los serotipos específicos: *S. Gallinarium* y *S. Pullorum*. Estos serotipos pueden causar tifoidea aviar y pullorosis (diarrea blanca). También existen serotipos no tifoideos, los cuales no presentan especificidad por algún hospedero en particular. Destacándose *S. Typhimurium*, *S. Enteritidis* y *S. Heidelberg*, los cuales se han encontrado en el ambiente y animales de granja (Andino y Hanning, 2015).

Para poder persistir o sobrevivir, esta bacteria debe ser capaz de infectar a un hospedero, sobrevivir en su paso por el medio ambiente para posteriormente entrar en un nuevo hospedero susceptible, desarrollándose y multiplicándose en el hospedero recién infectado (Figura 1) (Commission, 2016).

Una de las rutas más importantes para la transmisión de este patógeno, es la vía fecal-oral (Figura 2). Donde el patógeno abandona al huésped a través de las heces y otro hospedero entra en contacto con el material fecal mediante la ingesta; por lo que cualquier ruta que tome materia contaminada con heces y

llegue directamente o indirectamente a la boca de alguien, podría transmitir el patógeno (ELIKA, 2013).

b) Ambientes no hospederos

La supervivencia de *Salmonella* fuera de un hospedero está determinada por una serie de factores, dentro de los cuales se incluyen: temperatura, humedad, tipo de suelo, presencia de plantas y exposición a luz solar (UV), depredación por protozoarios y el número de microorganismos presentes (Jacobsen y Bech, 2012). Se ha reportado que *Salmonella* spp. puede sobrevivir hasta 322 días en suelos donde haya material fecal (Islam y col. 2004).

También, se han reportado diferentes formas en las que este patógeno puede entrar en el suelo, siendo la presencia de estiércol un aspecto importante para el desarrollo del mismo (Garcia y col. 2010). El transporte de la bacteria en suelo es un proceso complejo debido a las interacciones que presenta, tales como: el flujo del agua y la retención de células bacterianas a partículas (Jacobsen y Bech, 2012).

Se han reportado diversos trabajos enfocándose en el estudio de *Salmonella* en ecosistemas acuáticos. Fish y Pettibone en 1995, reportaron que las bacterias se encuentran concentradas en el fondo de los cuerpos acuáticos como resultado de la sedimentación de partículas a las cuales se encuentra unidas dichas bacterias. Interesantemente, encontraron que los sedimentos pudieran proteger a las bacterias entéricas de condiciones de estrés asociadas con ambientes acuáticos (Fish y Pettibone, 1995).



Figura 1. Ruta de dispersión de patógenos. Donde se muestra cada uno de los elementos del ciclo de transmisión de una enfermedad (European Commission, 2016).

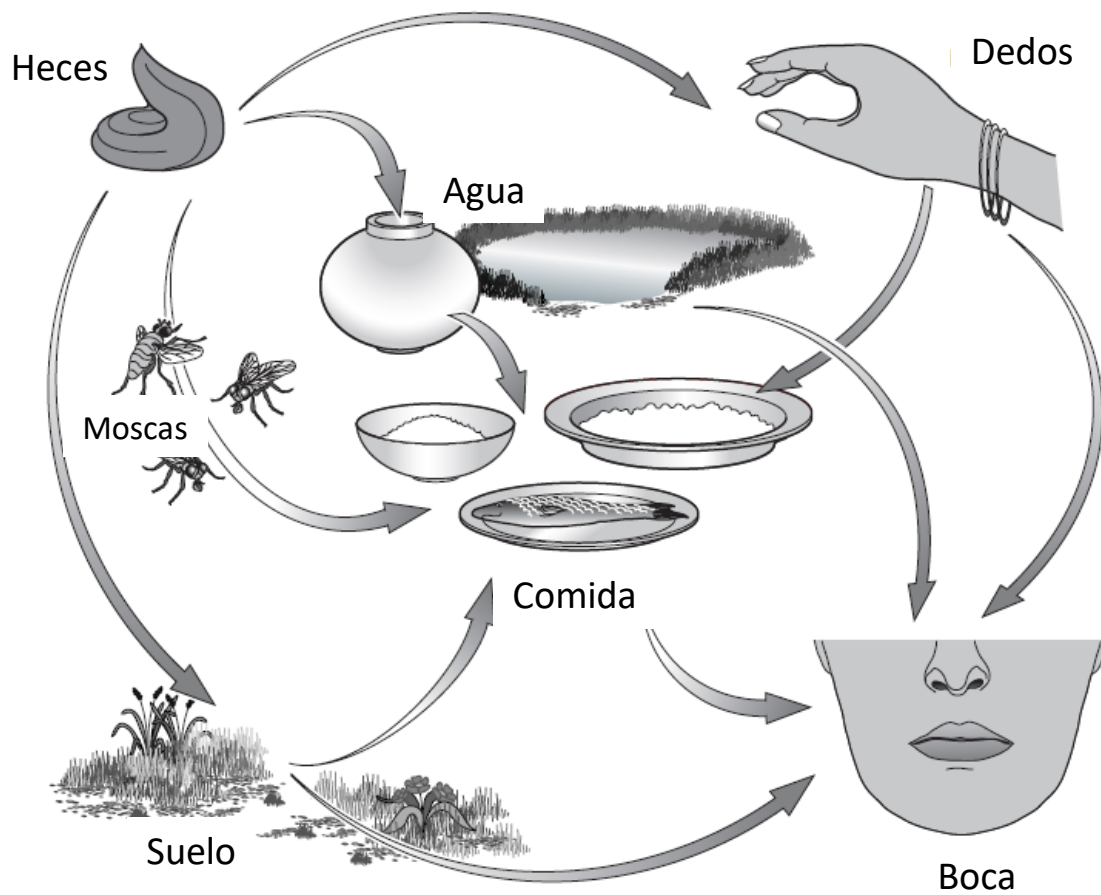


Figura 2. Rutas de transmisión para infecciones fecal-orales (European Commission, 2016).

Salmonella spp. se encuentra generalmente con mayor presencia en los sedimentos, que en agua superpuesta, lo que sugiere que los sedimentos son reservorios importantes para este patógeno. Esto representa un riesgo potencial a la salud de los seres humanos y vivos, mediante su ingesta (Burton y col. 1987). La presencia de materia fecal en cuerpos acuáticos, puede acentuar más la contaminación por este microorganismo (Moore y col. 2003). Su periodo de supervivencia en agua es mayor a las dos semanas, por lo que se puede encontrar en ríos, lagos y océanos que han sido contaminados con materia fecal (Jacobsen y Bech, 2012).

2 Características genéticas

El genoma de *Salmonella* es de aproximadamente 4.3-5.0 Mb, éste varía de acuerdo al serotipo. Algunas cepas pueden tener un ADN extracromosomal en forma de plásmidos, el cual generalmente contiene genes asociados a la virulencia o resistencia a antibióticos (Baker y Dougan, 2007).

Según los proyectos de secuenciación realizados por el Instituto Sanger; indican que el genoma de *Salmonella enterica* serovar Typhi tiene un cromosoma de 4,809,037 pb de longitud, con un porcentaje de 52.09% en el contenido C+G y 4,599 genes que codifican proteínas. En lo que respecta a *Salmonella enterica* serovar Paratyphi A, presenta un genoma de 4,587,797 pb y un porcentaje del 52.2% en el contenido C+G, lo que sustenta el rango (Sanger Institute, 2017).

Realizando comparaciones entre genomas de los diferentes serotipos, se ha identificado un conjunto de genes a los que se les denominó “core” que se refiere a todos los genes compartidos en todas las cepas presentes de una especie. Estos genes son de mantenimiento principalmente y están asociados a funciones vitales en la supervivencia de las cepas (Tettelin y col. 2005). Este genoma core forma parte de un termino denominado “pan-genoma” (Figura 3) que fue utilizado por Tettelin y col. en el 2005, para describir el conjunto de genes que se encuentran en una especie. El pan-genoma está constituido también por el genoma accesorio, el cual incluye genes que no son compartidos en el 100% de las cepas y entre dichos genes se encuentran aquellos que codifican para adaptaciones específicas o factores de virulencia (Ozer y col. 2014).

3 Características morfológicas

La morfología característica de *Salmonella* es de bacilos (forma alargada) o cocos (forma redonda), tienen un tamaño promedio 2-3 x 0.4-0.6 µm (Pui y col. 2011). Son organismos no formadores de esporas y están provistos de flagelos peritricos. Son anaerobios facultativos no encapsulados que utilizan la glucosa como fuente de carbono (Jurado, 2015).

B. Taxonomía del género *Salmonella*

De acuerdo al Centro para Control y Prevención de Enfermedades (CDC), el género *Salmonella* está dividido en dos especies: *enterica* y *bongori*. La especie *Salmonella entérica*, a su vez está dividida en 6 sub-especies que están designadas por nombre o, más comúnmente, por números romanos, como se

observa en la Cuadro 1 (Brenner y col. 2000). Los serotipos que pertenecen a *Salmonella enterica* subespecie *enterica*, generalmente su nombre se establece de acuerdo a la región geográfica donde fue aislada por primera vez (Issenhuth-Jeanjean y col. 2014).

Cada sub-especie esta a su vez dividida en serotipos, de acuerdo a la reacción inmunitaria con los antígenos de superficie: antígeno H, antígeno O y antígeno capsular Vi (Figura 4) según el esquema propuesto por White-Kauffmann (Figuroa y Verdugo, 2005). El primer tipo de antígeno, está formado por la proteína más abundante del flagelo (estructura que permite movimiento) “la flagelina”, la cual puede ser mono o difásicas (cuando se encuentra siempre el mismo antígeno flagelar o cuando el antígeno puede tener una forma alternativa) (CDC, 2011).

El antígeno O (somático) está conformado por una cadena repetida de polisacáridos, que forman parte del lipopolisacárido (LPS), que se genera y sobresale de la membrana externa, actuando como barrera de protección a agentes externos (BIORAD, 2002). El antígeno capsular Vi (virulencia); descubierto por Félix y Pitt en 1934, es un polisacárido capsular en cual solo está presente en *Salmonella typhi*, *S. paratyphi* y *S. dublin*, confiriendo resistencia contra la respuesta inmune celular y humoral del huésped, siendo una determinante en la virulencia de esta bacteria (Kolyva, Waxin, & Popoff, 1992). Este antígeno es un homopolímero lineal el cual está formado por unidades α -1,4-N acetilgalactosaminuronato con una masa molecular superior a los 200 kDa (Zhang y col. 2006).

Existen alrededor de 2,500 serotipos de *Salmonella*, de los cuales aproximadamente el 60% pertenecen a la sub-especie I y representan alrededor del 99% de las infecciones humanas (Lappe, 2011).

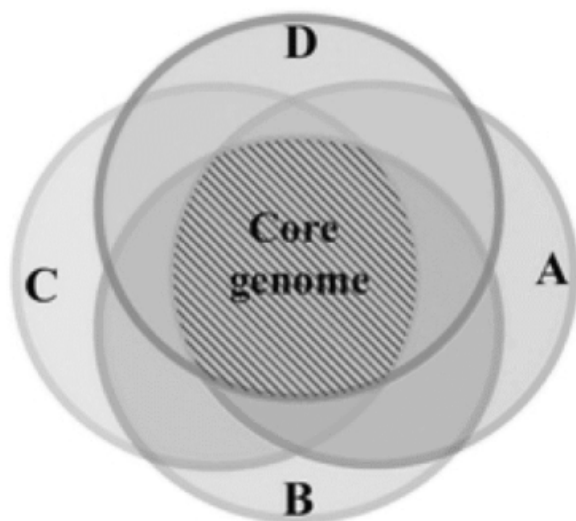


Figura 3. Pan-genoma representado mediante un diagrama de Venn. Este se compone del genoma core, que corresponde a los genes compartidos por las cepas A, B, C y D (sombreado con líneas negras); y el genoma accesorio, que comprenden los genes únicos para pocas cepas (Novogene, 2016).

Cuadro 1. Sub-especies de *Salmonella enterica*

| <i>Salmonella enterica</i> | |
|-----------------------------------|--------------------|
| Número Romano | Sub-especie |
| I | <i>enterica</i> |
| II | <i>salamae</i> |
| III | <i>arizonae</i> |
| IV | <i>diarizonae</i> |
| V | <i>indica</i> |
| VI | <i>houtenae</i> |

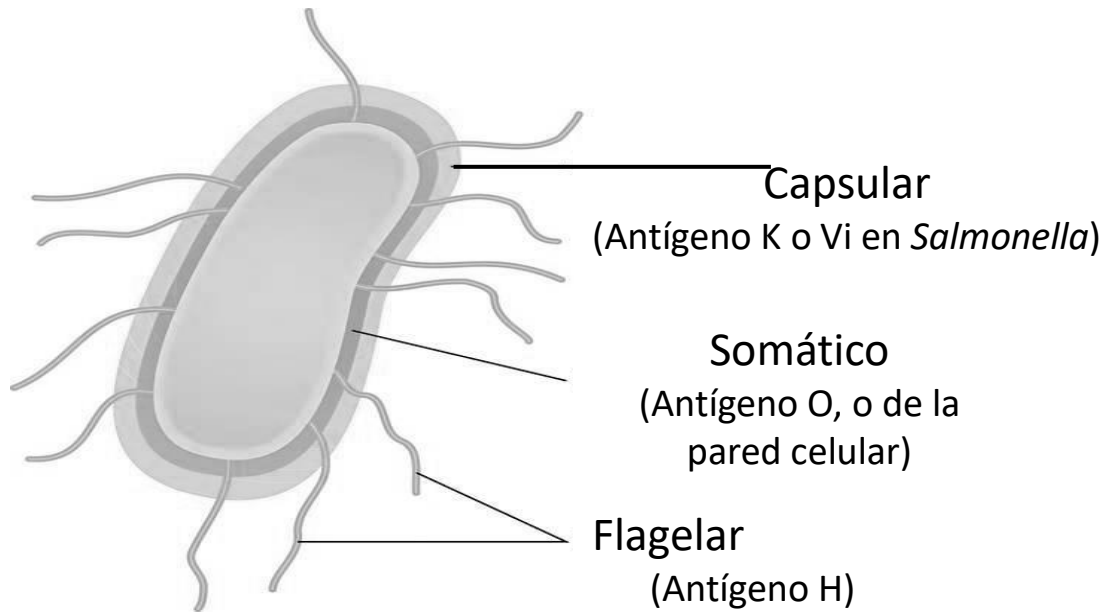


Figura 4. Localización de los antígenos O, H y Vi para serotipificación de *Salmonella* (University Of British Columbia, 2003).

C. Enfermedades ocasionadas por *Salmonella*

Esta bacteria puede causar enfermedades por el consumo de alimentos contaminados de origen animal; principalmente en huevos, carne, aves y leche; o por vía fecal-oral, ocasionando fiebre tifoidea (*Salmonella Typhi*) o salmonelosis no tifoidea (*Salmonella Enteritidis* y *Salmonella Typhimurium*). Siendo esta última, una de las enfermedades de transmisión alimentaria (ETA) más común, encontrándose ampliamente extendida a nivel mundial (OMS, 2013). Dicha enfermedad se caracteriza por fiebres encima de 40°C, dolor abdominal, diarrea, náuseas e inclusive vómitos (Gómez-Saravia, 2002).

De acuerdo al Centro de Control y Prevención de Enfermedades (CDC) de los Estados Unidos, en el periodo de 1998 al 2008, *Salmonella* fue el agente patogénico más común en las ETA's. Siendo responsable de 1,449 casos confirmados de los 7,998 casos presentados. Causante del 44% de las hospitalizaciones y de la mayor cantidad de muertes por este tipo de enfermedades con alrededor de un 30%. Se estima, que aproximadamente 1 millón de casos de salmonelosis ocurre en los Estados Unidos cada dos años, lo que representa un costo de más de 1 billón de dólares (CDC, 2013).

En la Unión Europea, la salmonelosis ocupa el segundo lugar en agentes causantes de infecciones alimentarias (Figura 5); después de campilobacteriosis,

con 150,995 casos confirmados. La principal causa de estos brotes está asociada al consumo de huevos y productos hechos a base de huevos crudos (EFSA, 2010).

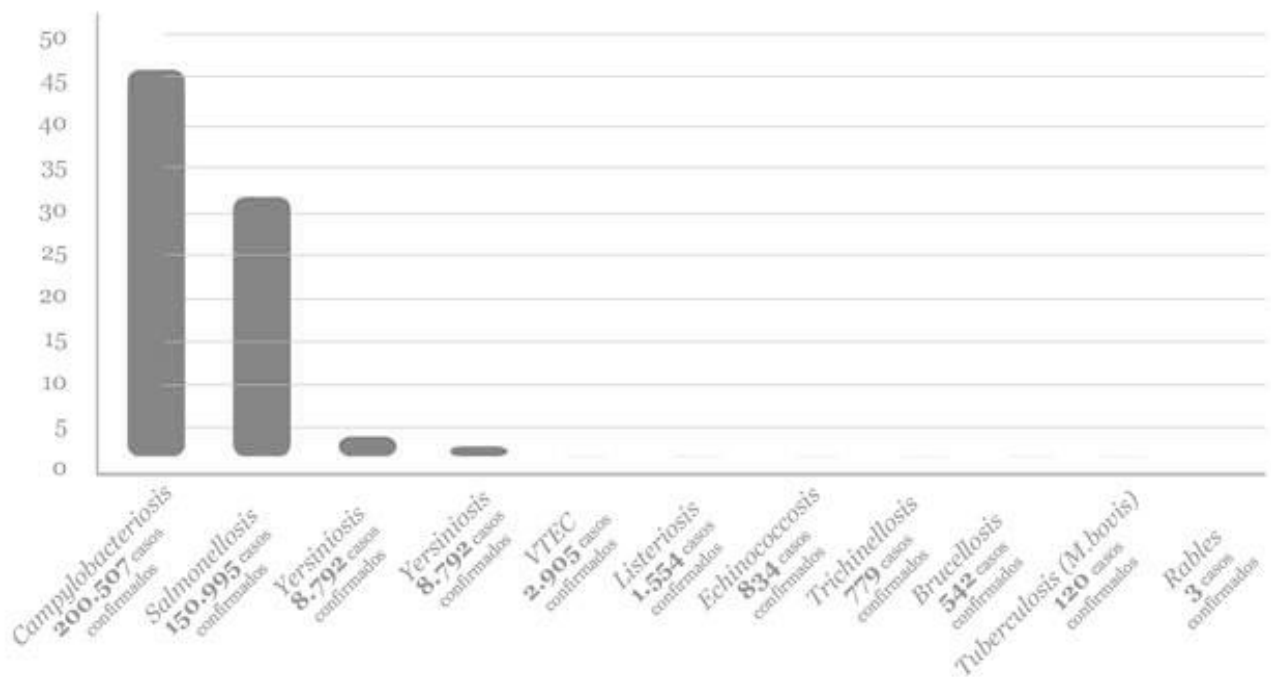


Figura 5. Agentes causantes de toxiinfección alimentaria en la Unión Europea (www.bioseguridad.net, 2007).

En México, según el Sistema Nacional de Vigilancia Epidemiológica (SiNAVE) en el año 2011; reportó que la paratifoidea y otras salmonelosis, ocuparon el décimo primero lugar entre las 20 principales causas de enfermedades transmisibles (Cuadro 2). También ocupó el segundo lugar, en causas provocadas por infecciones intestinales por otros organismos y las mal definidas (SiNAVE, 2011). *Salmonella* se podría encontrar dentro de las enfermedades “mal definidas” debido a las características mencionadas a continuación:

- Casos en los cuales no se puede efectuar un diagnóstico más específico, aún después de haber investigado todos los hechos.
- Hacer diagnósticos provisionales en pacientes que no volvieron a recibir atención o investigación adicional.
- Síntomas que hacen sospechar, con la misma verosimilitud, dos o más enfermedades.

Esta última clasificación, es donde se podrían incluir casos de salmonelosis, ya que a los pacientes se les brinda un diagnóstico provisional y no regresan. Por lo tanto, existe la posibilidad de que salmonelosis se encuentre entre las infecciones mal definidas.

De acuerdo al Boletín Epidemiológico Nacional, durante la semana número 52, la cual comprendió del 24 al 30 de Diciembre del 2017, se registraron 90,477 casos

de salmonelosis en México. El Cuadro 3, describe los casos presentados por estado y sexo. Siendo Sinaloa portador del séptimo lugar con 3,062 casos de salmonelosis confirmados; superando el número de casos reportados en el año 2016. También se incluyen datos de casos por infecciones intestinales por otros microorganismos y las mal definidas, con 5'606,756 casos confirmados en el año 2017. Es por ello, que la caracterización de *Salmonella* juega un papel importante para la disminución del número de casos y la ubicación de focos de infección.

D. Patogenicidad de *Salmonella*

Las infecciones ocasionadas por *Salmonella* comienzan cuando una cepa patogénica es ingerida por medio de agua o alimentos contaminados con material fecal. Los microorganismos ingeridos pasan al estómago y aquellos que logran sobrevivir al pH bajo se mueven al intestino delgado donde empiezan a competir con la flora intestinal por nutrientes y espacio (Giannella, 1996).

El primer paso consiste en la colonización de la superficie de células intestinales recubiertas por microvellosidades. Esto lo logra mediante un mecanismo conocido como disparo, donde la bacteria envía señales a las células epiteliales, induciendo rearrreglos en el citoesqueleto para la formación de ondulamientos en la superficie (Figuroa y Verdugo, 2005).

Después de la invasión al epitelio, la bacteria se multiplica intracelularmente llegando a los ganglios linfáticos, donde se distribuye a todo el cuerpo por medio del sistema circulatorio (Jones, 1997). Dependiendo del serotipo y de la efectividad que este tenga en el hospedero, puede causar daños en el hígado, vesícula, entre

otros órganos (Giannella, 1996). Una vez que ha invadido el intestino, se puede desencadenar una respuesta inflamatoria dando lugar a la formación de úlceras (Figueira y Holden, 2012).

Salmonella tiene la capacidad de invadir las células del hospedero debido a que cuenta con el sistema de secreción tipo III (T3SS) (Lahiri y col. 2010).

En cada paso del proceso infeccioso, este patógeno responde a los cambios en su entorno mediante la expresión de ciertos genes que le permiten sobrevivir a las condiciones desfavorables (Jones, 1997). A estos genes se les conoce como factores de virulencia y en esencia le confieren la habilidad patogénica a *Salmonella* para que pueda llevar a cabo una enfermedad; ya sea por la función en conjunto o individual de los genes. Estos factores están involucrados en interacciones con el tejido del huésped o en mecanismos de defensa. Teniendo como objetivos: la colonización del hospedero, inhibición del sistema inmunológico y la evasión de la respuesta inmune (Wu y col. 2008).

1. Virulencia de *Salmonella*

Se le conoce virulencia a la capacidad de una bacteria para causar una enfermedad; mientras que los factores de virulencia se refieren a las propiedades (productos de genes) que le permiten a un organismo establecerse en un huésped, desencadenando una enfermedad (Chen y col. 2012). Los factores de virulencia incluyen: proteínas de adhesión a una superficie celular, toxinas, enzimas hidrolíticas, carbohidratos y proteínas para la protección de bacterias (Ibarra y Steele-Mortimer, 2009).

Salmonella utiliza un sistema de genes de virulencia denominados islas de patogenicidad para ayudar a la bacteria a invadir, replicarse y sobrevivir en el hospedero (Marcus y col. 2000).

Cuadro 2. Principales causas de enfermedades transmisibles en México.

| | |
|--|--|
| 1. Infecciones respiratorias | 11. Paratifoideas y otras salmonelosis |
| 2. Infecciones intestinales por otros organismos y las mal definidas | 12. Tricomoniasis urogenital |
| 3. Infecciones de las vías urinarias | 13. Faringitis y amigdalitis estreptocócicas |
| 4. Otitis media aguda | 14. Otras enfermedades intestinales debidas a protozoarios |
| 5. Conjuntivitis | 15. Ascariasis |
| 6. Amebiasis intestinal | 16. Escabiosis |
| 7. Varicela | 17. Fiebre tifoidea |
| 8. Otras helmintiasis | 18. Intoxicación alimentaria |
| 9. Cándidas urogenital | 19. Virus del papiloma humano |
| 10. Neumonías y bronconeumonías | 20. Hepatitis aguda tipo A |

Fuente: SINAVE/DGE/Salud 2007

Cuadro 3. Entidades federativas con número de casos reportados de salmonelosis e infecciones intestinales por otros organismos y mal definidas en el 2017.

| ENTIDAD FEDERATIVA | Otras Salmonelosis CIE-10 REV. A02 | | | | Infecciones intestinales por Otros Organismos y las mal definidas CIE-10 REV. A04, A08-A09 excepto A08.0 | | | |
|----------------------------|--|-------|-------|---------------|--|--------|--------|---------------|
| | Sem. | 2017 | | 2016 Acum. | Sem. | 2017 | | 2016 Acum. |
| | | Acum. | | | | Acum. | M | |
| Aguascalientes | 16 | 426 | 709 | 1439 | 996 | 51706 | 53088 | 104322 |
| Baja California | 9 | 374 | 852 | 1725 | 919 | 76348 | 89898 | 114571 |
| Baja California Sur | 5 | 83 | 137 | 176 | 856 | 28729 | 31906 | 42318 |
| Campeche | 17 | 231 | 411 | 1183 | 782 | 23761 | 27508 | 49086 |
| Coahuila | 48 | 1859 | 3700 | 1195 | 1714 | 65013 | 74286 | 136344 |
| Colima | 3 | 57 | 87 | 137 | 478 | 18471 | 20675 | 41495 |
| Chiapas | 272 | 7945 | 14049 | 20678 | 2536 | 98450 | 116957 | 161760 |
| Chihuahua | 30 | 1276 | 2382 | 2740 | 1412 | 85577 | 106953 | 150765 |
| Ciudad de México | 4 | 127 | 271 | 468 | 4056 | 180274 | 210640 | 303377 |
| Durango | - | 5 | 15 | 24 | 839 | 36384 | 43646 | 107238 |
| Guanajuato | 14 | 360 | 684 | 759 | 2380 | 91664 | 105710 | 16150 |
| Guerrero | 23 | 23 | 1102 | 1709 | 1195 | 58369 | 72495 | 111873 |
| Hidalgo | - | 114 | 124 | 241 | 911 | 43892 | 50189 | 80095 |
| Jalisco | 22 | 1069 | 1629 | 2321 | 4302 | 166173 | 188015 | 262961 |
| México | 14 | 672 | 1035 | 1752 | 8287 | 284521 | 327833 | 559785 |
| Michoacán | 18 | 612 | 1362 | 1999 | 1738 | 76808 | 91959 | 142229 |
| Morelos | 7 | 74 | 144 | 213 | 990 | 48105 | 57700 | 93456 |
| Nayarit | 16 | 589 | 1236 | 2242 | 947 | 32766 | 37432 | 67155 |
| Nuevo León | 1 | 109 | 343 | 468 | 3180 | 117679 | 132951 | 214940 |
| Oaxaca | 7 | 452 | 773 | 1409 | 1614 | 75108 | 88333 | 138947 |
| Puebla | 35 | 2042 | 3631 | 5169 | 2533 | 123169 | 143095 | 171065 |
| Querétaro | 2 | 91 | 164 | 313 | 1380 | 47799 | 55159 | 88099 |
| Quintana Roo | 40 | 1195 | 2072 | 2167 | 1576 | 53793 | 56575 | 51274 |
| San Luís Potosí | 15 | 336 | 709 | 658 | 1478 | 65089 | 74141 | 97265 |
| Sinaloa | 47 | 1042 | 2020 | 2167 | 1537 | 73223 | 82021 | 115094 |
| Sonora | - | 219 | 512 | 658 | 592 | 65885 | 78294 | 145411 |
| Tabasco | 70 | 2995 | 5728 | 2139 | 1764 | 59983 | 67582 | 99071 |
| Tamaulipas | 9 | 992 | 2247 | 768 | 4207 | 133619 | 153194 | 140646 |
| Tlaxcala | 5 | 209 | 376 | 8118 | 1222 | 39823 | 44823 | 55363 |
| Veracruz | 163 | 4525 | 9109 | 3096 | 4663 | 136979 | 162223 | 207363 |

| | | | | | | | | |
|------------------|-----|-------|-------|-------|-------|---------|---------|---------|
| Yucatán | 12 | 444 | 673 | 4463 | 2614 | 84432 | 93311 | 207540 |
| Zacatecas | 9 | 381 | 763 | 2614 | 1477 | 58495 | 66080 | 98900 |
| TOTAL | 933 | 31428 | 59049 | 76855 | 65175 | 2602087 | 3004672 | 4417983 |

Fuente: SINAVE/DGE/Salud 2017. Información preliminar, incluye casos probables.

La mayoría de las islas de patogenicidad de *Salmonella* (SPIs) se encuentran adyacentes a genes de ARNt, difieren entre islas en el contenido G+C; lo que sugiere que fueron adquiridas por procesos de transferencia horizontal (Amavisit y col. 2003).

Se han identificado 12 islas de patogenicidad, las cuales se encuentran conservadas entre el género *Salmonella*, mientras que otras son específicas para ciertos serotipos. Del total de islas, solo las primeras 5 islas (Cuadro 4) se encuentran bien definidas (Hensel, 2004).

Cuadro 4. Descripción de las islas de patogenicidad en *Salmonella*.

| Isla de patogenicidad | Tamaño (Kb) | Contenido G+C | Fenotipo | Función | ARNt adyacente |
|------------------------------|--------------------|----------------------|--|---|------------------------------------|
| SPI-1 | 46 | 52% | Codifica las proteínas necesarias para ensamblar el complejo de secreción tipo III | Translocación de proteínas efectoras | No se encuentra asociado a un ARNt |
| SPI-2 | 41 | 43% | Codifica para un segundo sistema de secreción | -Patogénesis sistemática -Respiración anaerobia -Crecimiento intracelular | <i>vaIV</i> ARNt |
| SPI-3 | 17 | 47.5% | Codifica MgtCB para aumentar la afinidad por el Mg ²⁺ | Adaptación a las condiciones limitantes de nutrición | <i>Se/C</i> ARNt |
| SPI-4 | 28 | 50% | Codifica para el tipo secreción I | Secreción de toxinas | <i>ssb</i> ARNt |
| SPI-5 | 9 | 43.6-52% | Codifica proteínas efectoras para SPI-1 y SPI-2 | Asociada a enteropatogénesis | <i>serT</i> ARNt |

Fuente: (Chen y col. 2015), (Hensel, 2004).

SPI-1 codifica al sistema de secreción tipo III que sirve para translocar proteínas efectoras a una célula eucariótica. Estas proteínas median la invasión de las células no fagocíticas por *Salmonella*, proceso que involucra la modificación del citoesqueleto (Galán, 2001).

Se ha identificado la funcionalidad de los genes pertenecientes a esta isla mediante la inserción de mutaciones. Por ejemplo; Lostroh y Lee en el 2001, identificaron el gen *InvA*, siendo responsable de la invasión de este patógeno a células epiteliales (Lostroh y Lee, 2001). Estos genes son expresados en respuesta a las señales percibidas por la bacteria en el intestino del hospedero infectado (Figueira y Holden, 2012).

Diversos autores han encontrado una alta identidad en las secuencias de genes pertenecientes a SPI-1 entre microorganismos como *Salmonella*, *Shigella* y *E. coli*, lo que sugiere que tienen un ancestro en común (Bäumler, 1997; Groisman y Ochman, 1993).

SPI-2, se compone de dos elementos distintos:

- El segmento más grande de 25 kb, el cual es esencial para desencadenar la patogeneidad (se encuentra solamente en *S. enterica*). Este codifica para genes de T3SS, sistema regulatorio, sistema de chaperonas y sistemas efectores (Hensel, 2004).
- El segmento pequeño de 15 kb, responsable de la virulencia sistémica. Codifica para 5 genes *ttr*, los cuales son responsables

de la utilización del tetrionato como aceptor de electrones para la respiración anaerobia de la bacteria (Hensel, 2000).

La distribución de SPI-3 varía entre las cepas de *Salmonella*, pero todas tienen conservado el gen *mgtC* siendo un elemento importante para la supervivencia intramacrófago, ocasionar virulencia en ratón y para crecer en condiciones bajas de Mg^{2+} . El gen *mgtC* es transcripcionalmente controlado por el regulador PhoP-PhoQ, que se activará de acuerdo a las condiciones de baja concentración del ión Mg^{2+} (Blanc-Potard y col. 1999).

El papel de SPI-4 en la virulencia todavía no se ha analizado completamente (Hensel, 2004), sin embargo; Kiss y col. en el 2007, han determinado que la pérdida de esta isla atenua la virulencia de *S. Typhimurium* y *S. Enteritidis* en ratón. SPI-4 se compone de aproximadamente 6 genes, los cuales se han designado como *siiA-F* por ser genes que propician una infección intestinal (Kiss y col. 2007).

SPI-5 fue identificada por primera vez en *Salmonella* Dublin y consiste en 5 genes principalmente: *pipA*, *pipB*, *pipC*, *sopB* y *pipD*. Esta isla juega un papel importante en la virulencia; además codifica efectores de SPI-1 y SPI-2 (Cao y col. 2014). Por ejemplo; el gen *sopB* codifica para una proteína efectora de T3SS en SPI-1 bajo el control de *hilA*. Adicionalmente *pipB* codifica para una proteína efectora para T3SS de SPI-2 (Hensel, 2004).

2. Sistema de secreción tipo III (T3SS)

Este tipo de sistema de secreción se encuentra en bacterias gram-negativas, permitiéndoles la translocación de proteínas efectoras dentro de la célula huésped (Cabañas-Romero y Huerta-Saquero, 2014). T3SS se compone de una estructura especializada parecida a una jeringa (Figura 6) compuesta por múltiples anillos que se extienden a través de la envoltura bacteriana (Galán y col. 2014). Fue identificada originalmente en *Salmonella* Thyphimurium, pero se ha identificado que es codificada por otras bacterias (Blocker y col. 2001).

Este aparato consiste principalmente en una base que cubre la envoltura bacteriana y una aguja que termina en una estructura de punta (Figura 6). La bacteria secreta dos proteínas translocatorias, que se introducen en la membrana de la célula huésped para la formación de un poro; el cual permite la transferencia de otras proteínas necesarias por el patógeno (Lee y col. 2014).

Su función principal es inyectar proteínas bacterianas denominadas efectoras, las cuales pasan al citoplasma del hospedero, afectando la señalización intracelular, la dinámica del citoesqueleto y el tráfico vesicular de la célula a favor del patógeno, lo que constituye un paso importante para el desarrollo de la enfermedad en el huésped (Molina y col. 2017).

Este es un sistema complejo que se componen de más de 20 proteínas estructurales. Se ha reportado que este sistema se encuentra parcialmente conservado en *Salmonella enterica*, por lo que la estructura se encuentra generalizada en toda esta sub-especie (Gonzales-Pedrajo y Dreyfus, 2003).

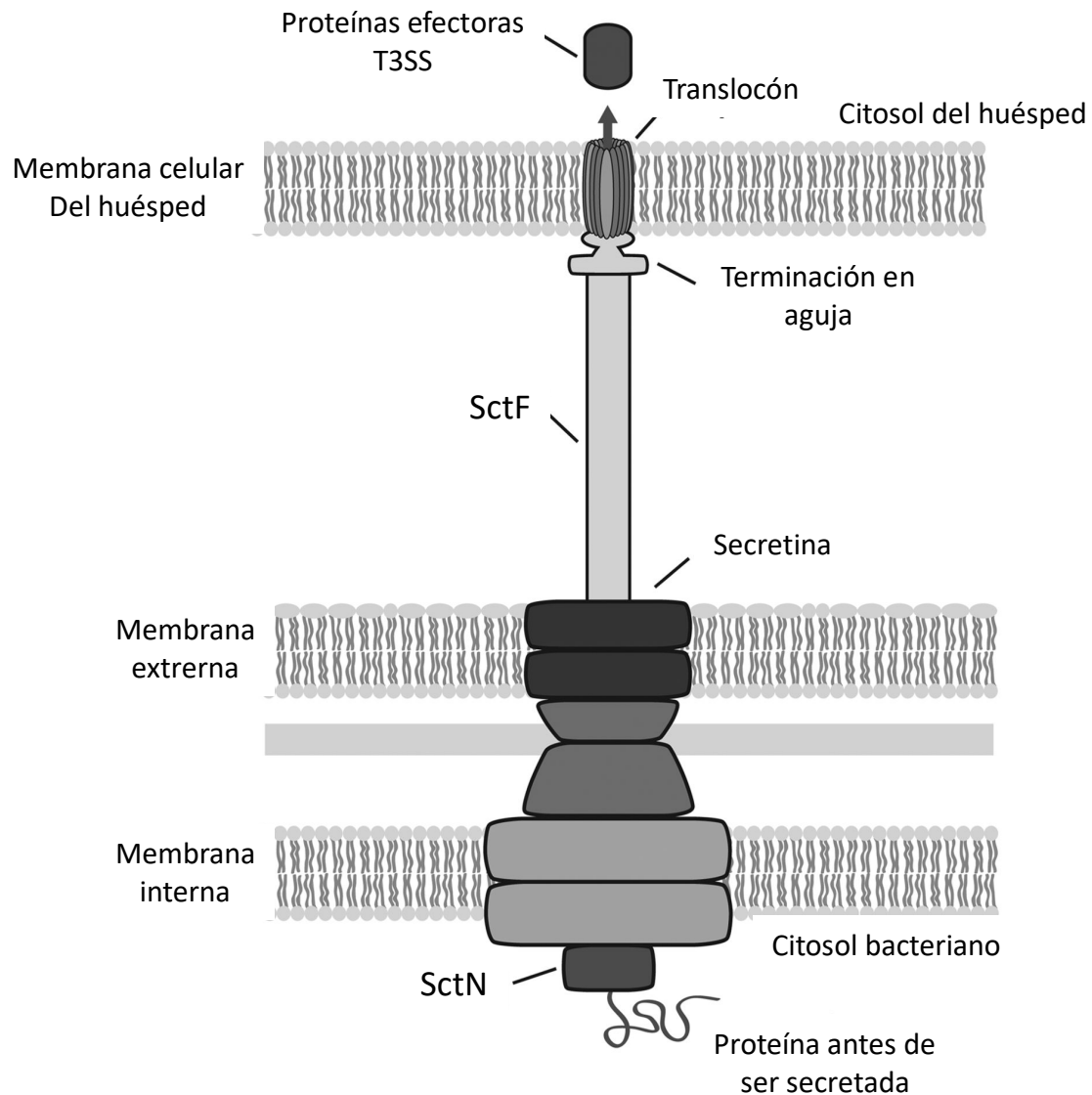


Figura 6. Estructura del sistema de secreción tipo III en bacterias patogénicas. Su estructura consta principalmente de una estructura inmersa en la membrana bacteriana, prolongándose hasta terminar en una especie de punta o aguja (Duncan y col. 2012).

E. Métodos moleculares para la caracterización de patógenos

Debido a la importancia epidemiológica en el estudio de bacterias patogénicas como *Salmonella*, se han desarrollado métodos moleculares que permiten discriminar bacterias de manera precisa y oportuna. Logrando con ello determinar las especies involucradas en un brote, para posteriormente identificar la fuente de infección; con el objetivo de detectar, controlar y/o prevenir los brotes de ETAs. A continuación se describen algunos de los métodos más comunes para la caracterización de bacterias.

1 Reacción en cadena de la polimerasa (PCR)

Es un método enzimático utilizado para sintetizar largas regiones del ADN, generando millones de copias. Fue desarrollado por Kary Mullis en 1983, quien recibió el Premio Nobel en Química por dicho trabajo. Ésta es la técnica de diagnóstico molecular más ampliamente utilizada debido a que su protocolo es fácil y rápido (Olive y Bean, 1999).

La PCR se ha utilizado en la identificación de virus, bacterias, hongos, entre otros. Diferentes autores han desarrollado metodologías específicas para la detección de patógenos, por ejemplo Ayala y col. en el 2004 identificaron la bacteria *Clavibacter michiganensis*, causante de enfermedades en maíz; para lo que desarrollaron cebadores específicos de hibridación a una región conservada en dicha bacteria (Ayala-Labarríos y col. 2004).

Esta metodología se basa en la repetición de ciclos, los cuales incluyen tres etapas (Figura 7):

- Desnaturalización: Consiste en separar la doble cadena del ADN en dos hebras independientes mediante el uso elevado de temperatura (aprox. 94°C).
- Alineación: Aquí se añaden los cebadores (secuencias cortas complementarias a la hebra molde) para que se hibriden con las hebras desnaturalizadas mediante la acción de la enzima *Taq* polimerasa. Para ello, se utilizan temperaturas más bajas.
- Extensión: Se empiezan a extender las cadenas molde del ADN mediante la adición de nucleótidos por acción de la polimerasa.

2. Electroforesis en gel de campo pulsado (PFGE)

Es una técnica de laboratorio usado para generar “huellas” de ADN de aislados bacterianos mediante la utilización de enzimas de restricción. Esta es una variante de la electroforesis convencional donde se aplica un campo eléctrico constante; sin embargo, la PFGE permite analizar grandes fragmentos de ADN (mayores de 40 kb) gracias a que la orientación del campo eléctrico es variable (Cercenado y Cantón, 2005).

PFGE es considerada como “la técnica de oro” para la tipificación molecular (Palomino-Camargo y Gonzalez-Muños, 2014). En la Figura 8 se muestra el diagrama de flujo para llevar a cabo una tipificación por PFGE. El primer paso consiste en tomar una unidad formadora de colonia y posteriormente

homogenizarla con agar líquido. Seguido, se vierte dicha mezcla en moldes donde se añaden agentes bioquímicos que permiten retirar impurezas del agar. El siguiente paso consiste en agregar agentes de lisis para liberar el ADN; posteriormente se fragmenta mediante enzimas de restricción y cargar estas muestras en el gel. Una vez finalizada la corrida electroforética se procede a visualizar el bandeo generado por la migración del ADN en el gel mediante luz ultravioleta (CDC, 2017).

Se han propuesto diferentes sistemas para estandarizar la interpretación de los patrones obtenidos por PFGE con el objetivo de determinar la cercanía entre cepas. En este esquema, todas las bacterias que presenten el mismo perfil de bandas, son consideradas como pertenecientes a la misma cepa (Olive y Bean, 1999). Como se ha mencionado previamente, esta técnica es útil para la mayoría de las especies bacterianas. Sin embargo, en algunos casos carecen del suficiente poder de discriminación. Un ejemplo de ello es el género *Salmonella*, el cual posee una gran cantidad de serovariedades las cuales no pueden ser diferenciadas con suficiente especificidad; adicionalmente presenta una rápida degradación del ADN cromosómico, lo que genera que muchas cepas sean no tipificables (Cercenado y Cantón, 2005).

Adicionalmente, se ha reportado que la ejecución de la técnica es muy tardada y que en ocasiones las bandas que coinciden en el mismo tamaño pueden no provenir de la misma parte del cromosoma (CDC, 2007).

Debido a las limitantes de las técnicas como la PFGE, se han buscado estrategias que tengan la sensibilidad suficiente para medir pequeñas variaciones genéticas.

Lo que permitirá mejorar la caracterización de agentes patogénicos, aportando información de gran utilidad en el área de epidemiología.

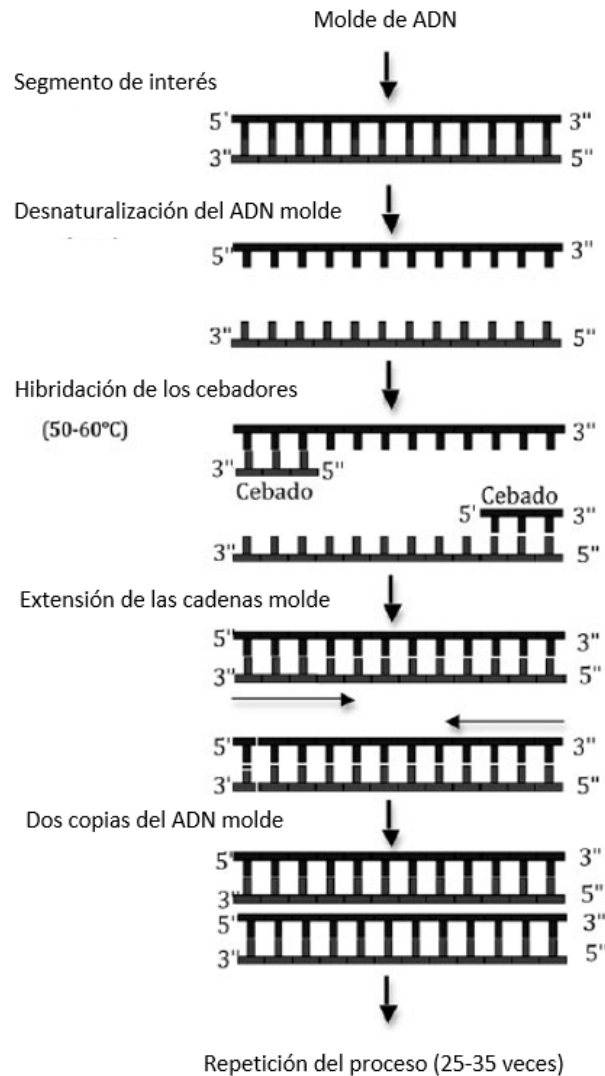


Figura 7. Etapas de la PCR. Esta técnica se basa en el principio de complementariedad de las bases del ADN y en la participación de la enzima polimerasa que permite la extensión del fragmento a amplificar, agregando nucleótidos en la secuencia complementaria al ADN molde por medio de un

proceso cíclico de tres pasos: desnaturalización, alineación y extensión (Palomino-Camargo y Gonzalez-Muños, 2014).

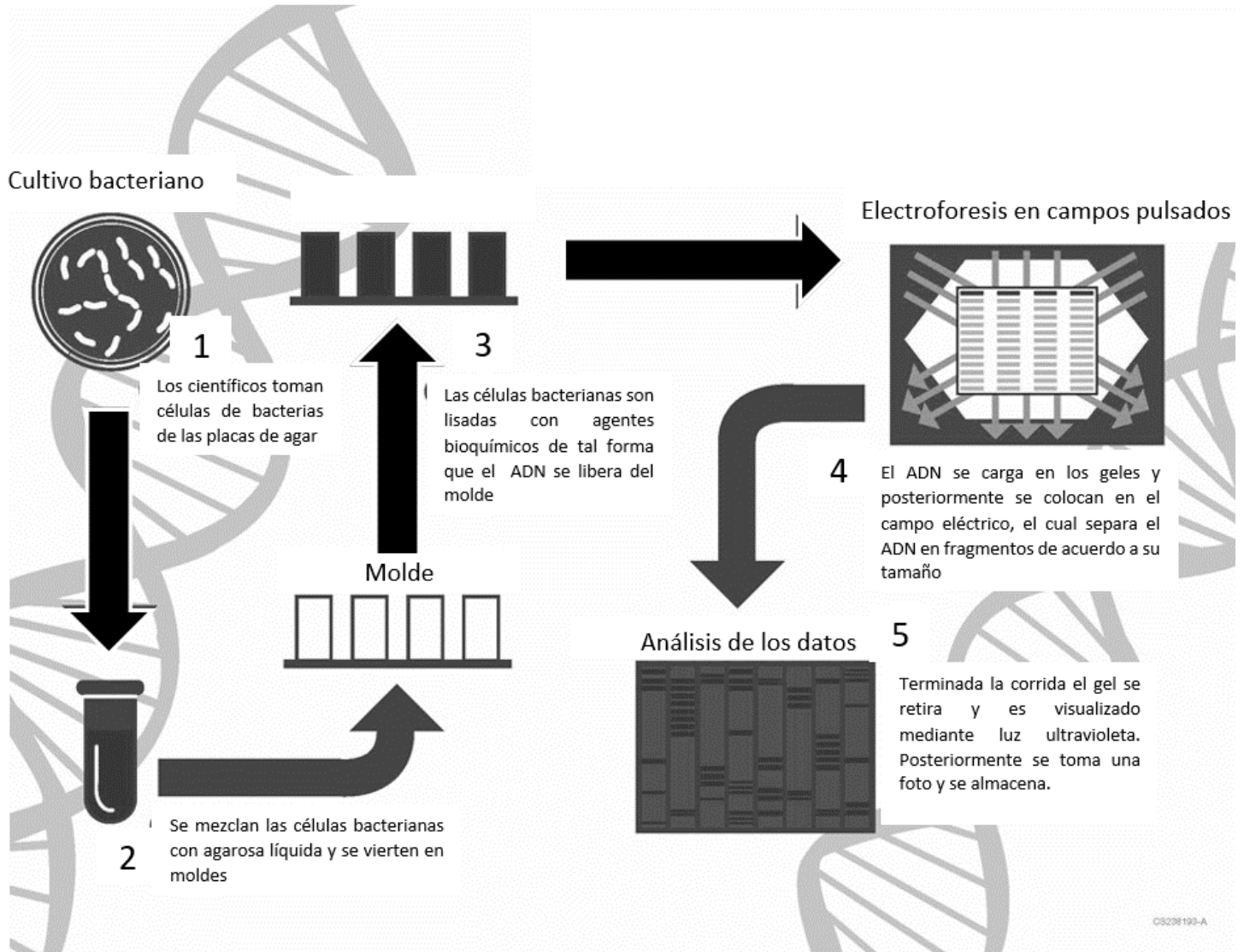


Figura 8. Diagrama de flujo del proceso para realizar una electroforesis PFGE. Se destacan con número cada uno de los pasos (CDC, 2007).

Por lo que el uso de procesos de secuenciación y herramientas bioinformáticas han demostrado ser suficientemente sensibles para diferenciar inclusive hasta serotipos (Wyres y col. 2014).

F. Análisis molecular de bacterias patogénicas mediante herramientas bioinformáticas.

El proceso de secuenciación se ha visto influenciado por el desarrollo tecnológico, dando lugar a la aparición de la secuenciación de siguiente generación (NGS). Término que es empleado para designar a aquellas plataformas de secuenciación que tienen como característica, el procesamiento masivo y en paralelo de muestras; logrando reducir costos, tiempo y mejorando la profundidad del análisis (Behjati y Tarpey, 2013).

Esto ha mejorado la investigación científica; específicamente el área de la genómica comparativa. Permitiendo la tipificación multilocus de secuencias (MLST), la identificación del pan-genoma (core y accesorio), el análisis de polimorfismos de nucleótido simple (SNPs) para la construcción de árboles filogenéticos, la identificación de factores de virulencia y de resistencia a antibióticos (Figura 9), principalmente (Shendure y Ji, 2008).

1 Plataforma de secuenciación Illumina

El sistema de Illumina funciona mediante secuenciación por síntesis; es decir, utiliza bases nitrogenadas marcadas con fluorescencia, las cuales se denominan dNTPs (deoxinucleótidos trifosfatos) que se añaden a hebras crecientes.

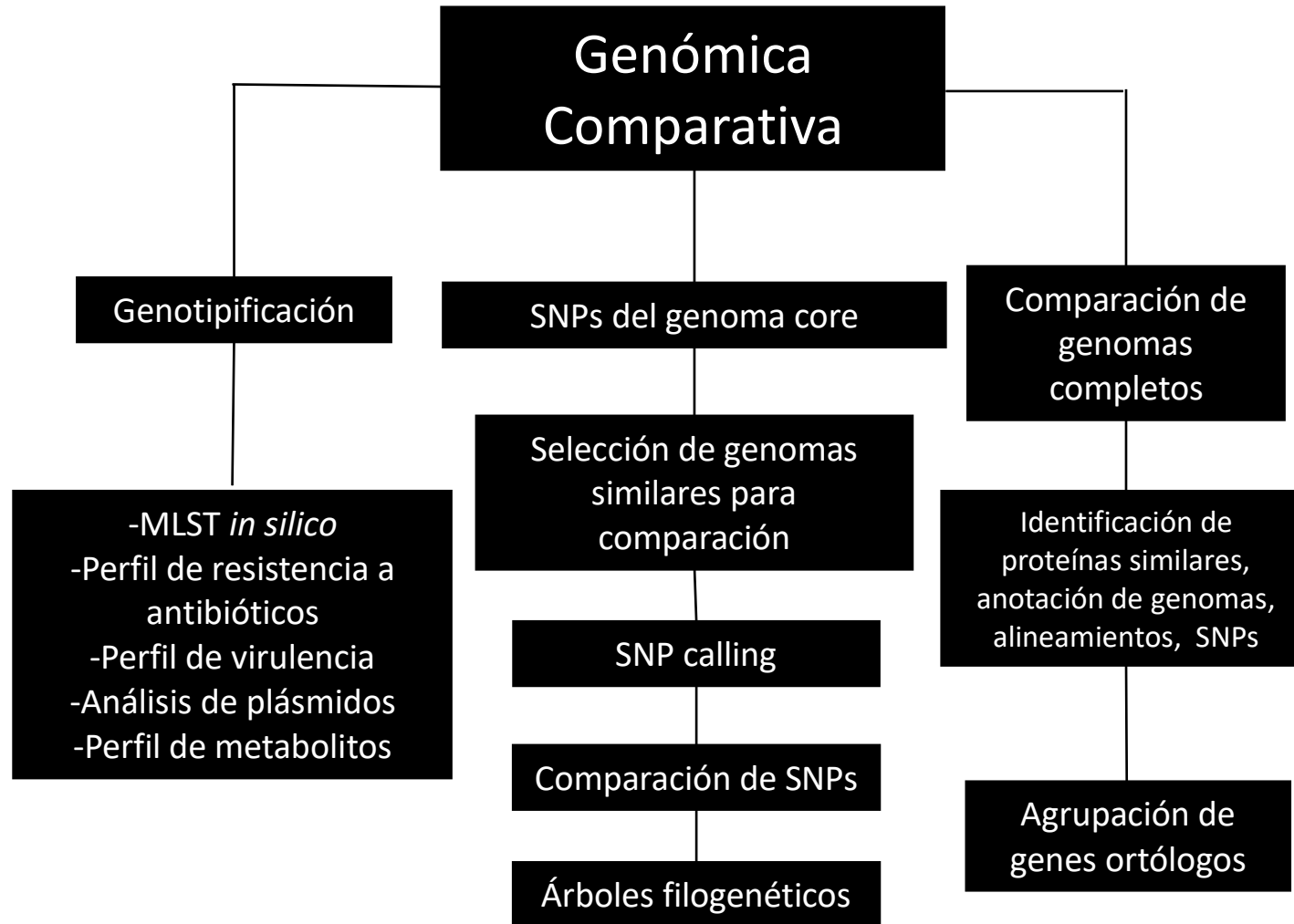


Figura 9. Diagrama con algunas de las aplicaciones más importantes que se pueden realizar en genómica comparativa, gracias al desarrollo tecnológico en el área de secuenciación y bioinformática (www.era7bioinformatics.com, 2016).

El proceso de secuenciación inicia con la generación de clústers; para ello se lleva a cabo la inmovilización de los primers forward y reverso en la superficie de celdas de flujo (flow cell), haciendo asequible la participación de enzimas polimerasas (Shendure y Ji, 2008).

Posteriormente, se da una amplificación llamada “PCR bridge” en fase sólida, donde se crean más de 1,000 copias idénticas a las muestras inmovilizadas (Figura 10). Para la identificación de las bases en las secuencias, se realizan una serie de ciclos donde se van agregando una mezcla de primers, enzimas y dNTPs marcados; los cuales son excitados con láser para la emisión de la fluorescencia (Illumina, 2010). Por último, se recolectan los datos y generan las secuencias cortas denominadas reads. En la Figura 11 se muestra de manera general los pasos englobados en 4 categorías principales para la secuenciación de ácidos nucleicos en la plataforma Illumina.

El empleo de esta nueva tecnología, ha traído consigo la generación de cantidades masivas de datos, dificultando su análisis. Es por ello, que resulta de suma utilidad e importancia, el hacer uso de herramientas que permitan analizar de manera rápida y eficiente las secuencias genómicas (Marchant y col. 2014). La bioinformática es un área que permite el procesamiento y análisis de estos datos para asignarles un significado biológico mediante el empleo de matemáticas, ciencias computacionales y estadística (Luscombe y col. 2001). Lo que resulta indispensable su uso para la organización y entendimiento de la información genómica.

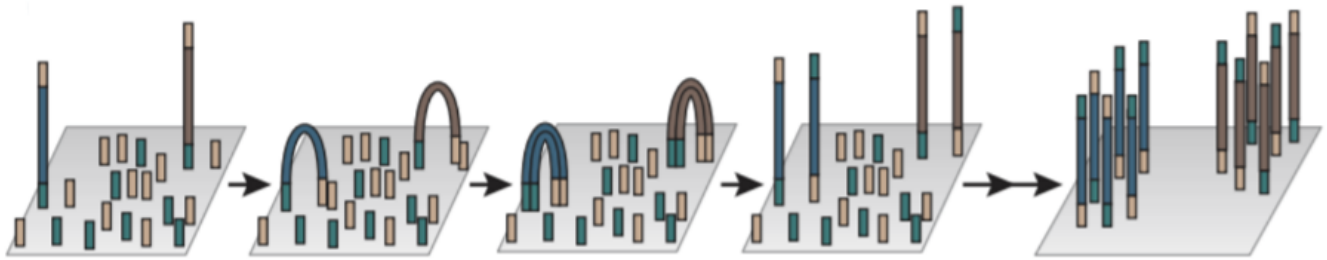


Figura 10. PCR por puente (bridge) en la plataforma de secuenciación Illumina MiSeq. El proceso inicia con la unión de los adaptadores a las secuencia, posteriormente las secuencias se adhieren a una superficie sólida. La secuencia se empieza a amplificar, generándose diferentes grupos de clústers (Shendure y Ji, 2008).



Figura 11. Diagrama de flujo del proceso general de secuenciación de ácidos nucleicos en plataformas Illumina.

El objetivo final de la bioinformática por lo tanto, es mejorar el entendimiento de los organismos vivos y sus funciones a nivel molecular. Para ello se basa en tres aspectos fundamentales (Xiong, 2006) (Figura 12). De los cuales, nos centraremos en el análisis de secuencias genómicas para el desarrollo de esta investigación.

2 Herramientas bioinformáticas para el análisis de secuencias genómicas

La secuenciación de genomas ha llevado a la necesidad de obtener conclusiones de las lecturas de millones de pares de bases; para saber que codifican, como se relacionan y regulan la expresión de distintos productos genéticos; además de generar modelos que permitan estudiar mutaciones puntuales (Walteros y col. 2012).

Para estos fines mencionados, se han desarrollado programas computacionales inteligentes, que utilizan algoritmos diseñados específicamente para el estudio de aspectos biológicos, posibilitando el manejo de una gran cantidad de datos (Franco y col. 2008).

a) Ensamblado de fragmentos de ADN

El uso de plataformas de secuenciación, generan como productos finales fragmentos cortos de ADN denominados *reads*. A partir de los cuales se pretende descifrar la secuencia genómica para obtener segmentos más largos denominados *contigs* y reconstruir así el cromosoma de interés (Gómez, 2012).

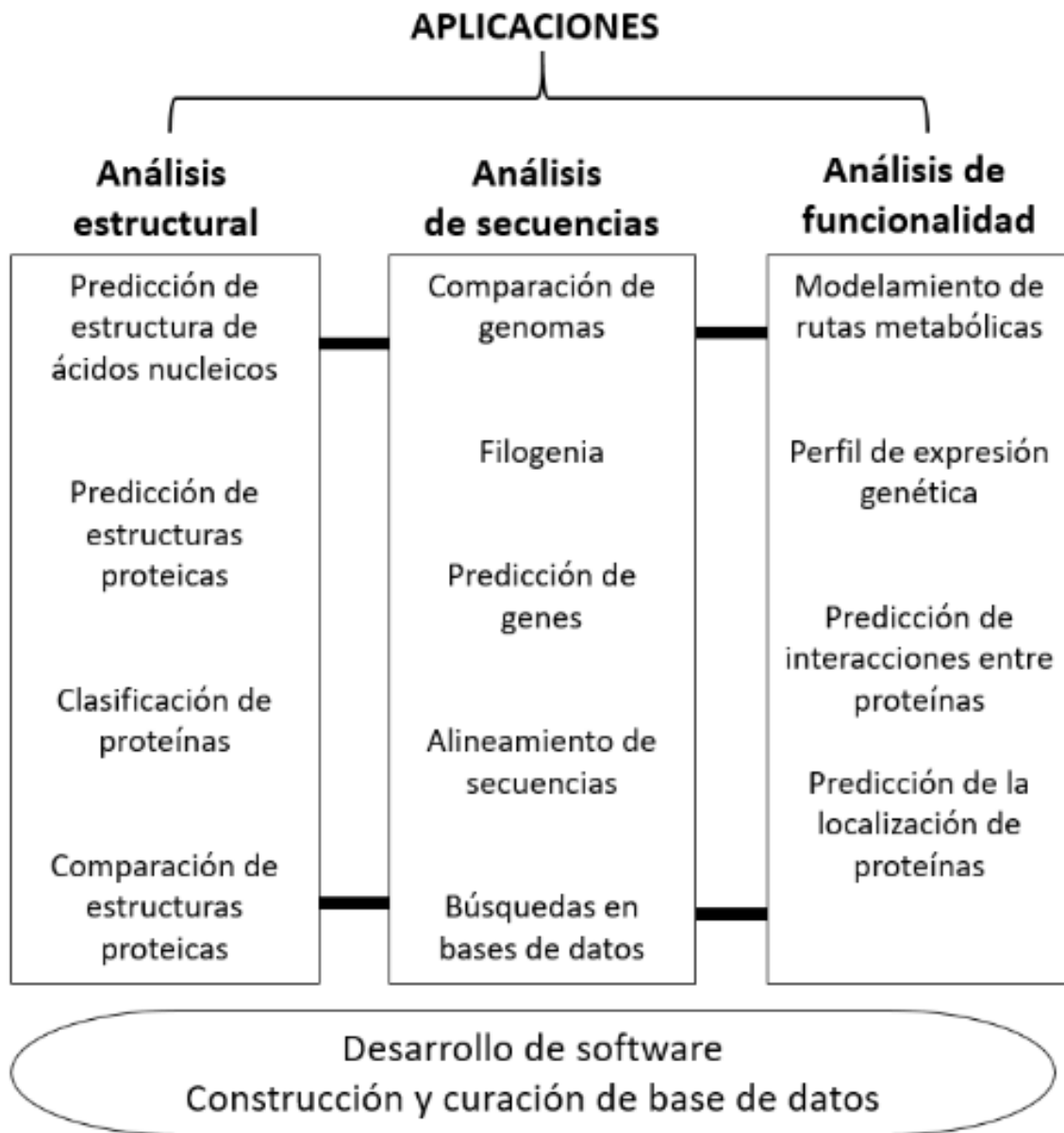


Figura 12. Panorama general de las aplicaciones de la bioinformática. Las cuales se pueden dividir en tres campos: el análisis estructura, de secuencias y de funcionalidad. Para que esto se pueda llevar a cabo es importante el desarrollo de software y la construcción de bases de datos. También se muestra las relaciones intrínsecas entre las diferentes áreas (Xiong, 2006).

Existen dos estrategias para el ensamblaje de secuencias:

- Ensamblaje *de novo*: el cual solo utiliza la información obtenida de la secuenciación para la construcción del genoma, sin contar con conocimiento previo de la secuencia a estructurar (Aguilar-Bultet y Falquet, 2015)
- Ensamblaje por comparación: utiliza un genoma de referencia

Siendo el ensamblaje *de novo*, mayormente usado en proyectos donde se trabaja con genomas microbianos debido a su pequeña talla (Chistoserdova, 2010).

Existen diferentes programas para el ensamble de *reads*, la elección del programa va a estar sujeta a la plataforma de secuenciación utilizada, las características del genoma y la aproximación que utilice el programa bioinformático (Baker, 2012).

El principio para el ensamblado de genomas, parte de la suposición de que los fragmentos de ADN altamente similares se originan de la misma posición dentro del genoma (Ekblom y Wolf, 2014). De esta manera, la similitud entre secuencias de ADN se usa para conectar fragmentos individuales en secuencias más largas (Aguilar-Bultet y Falquet, 2015) (Figura 13).

Es importante que haya suficiente solapación de los *reads* al momento de ensamblar; por lo que se requiere que haya una suficiente profundidad (aproximadamente de 45x para librerías *paired-end*) y que los fragmentos sean lo más largos posibles (Nagarajan y Pop, 2013).

Las estrategias utilizadas por los programas bioinformáticos emplean generalmente tres tipos de algoritmos basados en la construcción de grafos:

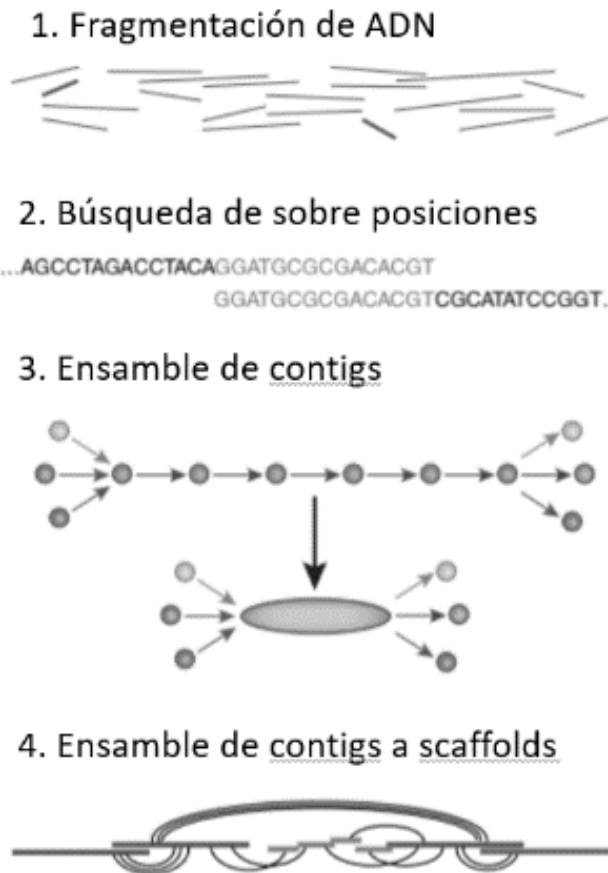


Figura 13. Diagrama para el ensamble de los *reads*. El primer paso requiere la utilización de los productos de secuenciación, se buscan solapaciones entre los pequeños fragmentos para la formación de *contigs*. Posteriormente estos *contigs* se unen en fragmentos más largos denominados *scaffolds* y finalmente se obtiene la secuencia del genoma (Baker, 2012).

- Overlap Layout Consensus: Identifica todos los pares de *reads* que se solapan de manera adecuada, organizando dicha información en un gráfico que se construye de nodos y conectores. Lo que definen caminos que corresponden a los segmentos que están siendo ensamblados (Aguilar-Bultet y Falquet, 2015).
- Gráficos de Bruijn: Es el algoritmo más utilizado para *reads* cortos procedentes de plataformas como Illumina y SOLiD. Se basa en la utilización de k-meros, donde la secuencia original es fragmentada en piezas de un tamaño definido (k-mer). Posteriormente, se construye una gráfica conectando mediante sobrelapaciones los primeros k-meros (nucleótidos k-1) (Miller y col. 2011), para encontrar el orden de la secuencia genética que se está reconstruyendo.
- Greedy: Es uno de los algoritmos más sencillos e intuitivos que realiza la solapación de secuencias cortas de manera iterativa. Sin embargo, no es tan empleado ya que solo utiliza información local de los fragmentos, sin considerar la información global; por lo que no resuelve de manera eficiente el ensamble de regiones largas (Dasgupta y col. 2006).

Una vez que el genoma ha sido construido satisfactoriamente, se evalúa la calidad del mismo (Ekblom y Wolf, 2014). Existen una gran variedad de métricas que se pueden evaluar; sin embargo entre las que se consideran de

mayor relevancia son: el tamaño del genoma generado, el valor de N50 y el fragmento de k-mer utilizado (Bradnam y col. 2013).

b) Anotación de genomas bacterianos

Después de obtener un genoma ensamblado de buena calidad, el siguiente paso consiste en asignarle un significado biológico, a este proceso se le llama anotación (ECOEVOL, 2017). El cual se basa en identificar y encontrar las características más importantes del genoma; lo que le agrega significado a la secuencia, mejorando la facilidad en la visualización y análisis en el contenido (National Library of Medicine, 2017).

Este proceso es un poco difícil, en parte porque existe un gran número de genes dentro de las secuencias genómicas. Es por ello que se hace uso de bases de datos para comparar de manera automática cada uno de los genes en el genoma patogénico contra secuencias parecidas (Gibas y Jambeck, 2001).

c) Construcción de árboles filogenéticos

Se le conoce como filogenia a la relación de parentesco entre especies o taxones en general. Ésta, se encarga de estudiar las relaciones evolutivas entre los organismos mediante la construcción de árboles filogenéticos, los cuales presentan topologías específicas (Peña, 2011). El estudio de la filogenética resulta importante ya que enriquece nuestro conocimiento de cómo secuencias, genes y especies han evolucionado.

Los métodos de reconstrucción filogenética se pueden clasificar de acuerdo al tipo de datos que emplean, los cuales pueden ser: por matrices de distancia o caracteres. Los métodos de matrices de distancia permiten construir árboles filogenéticos basados en la distancia genética entre secuencias de ADN, por lo que requieren realizar alineamientos múltiples (Yang y Rannala, 2012). Mientras que los métodos de caracteres comparan simultáneamente todas las secuencias en el alineamiento considerando un carácter (ubicado en el sitio de alineamiento) para calcular el puntaje para cada árbol (Valcárcel, 2008). Entre estos métodos se encuentra el de máxima verosimilitud (Maximum Likelihood). El cual busca determinar la organización de las ramas de un árbol filogenético y su longitud para evaluar las hipótesis evolutivas de los taxones analizados (Yang y Rannala, 2012).

IV. HIPÓTESIS

La caracterización genética de organismos patógenos mediante el uso de herramientas bioinformáticas como parte de la biología computacional permite analizar genomas de *Salmonella* aislados de agua de río y animales de granja con una sensibilidad suficiente para establecer relaciones filogenéticas, identificar factores de virulencia y agrupar a los serotipos de acuerdo a su perfil alélico.

Las cepas procedentes de la región de Culiacán, Sinaloa presentan poca diversidad genética entre los diferentes serotipos por tratarse de muestras recolectadas en puntos cercanos de la localidad y por corresponder a un periodo de muestreo corto. Adicionalmente, se espera que compartan perfiles de virulencia similares entre los distintos serotipos.

V. JUSTIFICACIÓN

El avance científico y tecnológico en los últimos años, ha puesto al alcance del investigador la posibilidad de conocer a detalle las secuencias de los diferentes microorganismos de manera rápida, eficiente y económica. Abriendo la puerta a procesos como la caracterización genética de los microorganismos en general.

El presente trabajo de investigación, está enfocado en la caracterización de los aislados de *Salmonella enterica* de agua de río y animales de granja de la región de Culiacán, Sinaloa. Mediante el empleo de técnicas de secuenciación y biología computacional basada en herramientas bioinformáticas actuales, con la finalidad de conocer la diversidad genética ambiental, origen global y marcadores genéticos de virulencia para dichas cepas.

La información generada en esta investigación y el desarrollo de un flujo de análisis con base en estas herramientas permitirá profundizar en conocimiento acerca de la dispersión, variabilidad, adaptabilidad, demografía y evolución de este patógeno, lo cual podría aplicarse al monitoreo, trazabilidad e identificación oportuna de brotes potenciales de *Salmonella*. Por lo que el desarrollo de este trabajo resulta de interés para instituciones de salud pública, sector agropecuario y dependencias económicas del país. Esto se debe a que *Salmonella* se ubica entre

los principales agentes causantes de enfermedades infecciosas, siendo responsable de numerosos casos de hospitalizaciones; repercutiendo en la economía del país.

VI. OBJETIVOS

A. OBJETIVO GENERAL

Determinar la diversidad genética poblacional de *Salmonella enterica* ambiental e identificar rasgos genéticos de virulencia en aislados de agua de río y animales de granja en Culiacán, Sinaloa.

B. OBJETIVOS ESPECÍFICOS

- Generar una base de datos que contenga los genomas y metadatos de *Salmonella* descargadas de NCBI.
- Analizar las relaciones filogenéticas entre los genomas de aislados locales e internacionales de *Salmonella enterica* mediante la construcción de árboles filogenéticos.
- Realizar una tipificación *in silico* de los serotipos locales de *Salmonella enterica*.
- Identificar las islas de patogenicidad así como los factores de virulencia que encuentran incluidos en el genoma accesorio de las bacterias.

VII. MATERIALES Y MÉTODOS

A. MATERIALES

Se utilizaron 154 cepas de *Salmonella enterica* pertenecientes al cepario del Laboratorio Nacional para la Investigación en Inocuidad Alimentaria (LANIIA). Estas cepas fueron aisladas entre Julio de 2008 y Junio de 2009 a partir de muestras de agua de río y animales domésticos asintomáticos (vacas, cabras y pollos), recolectadas en Culiacán, Sinaloa en los sitios de Jotagua, Aguacaliente, Puente Negro, La Cofradía, Iraguato y El Castillo.

La extracción de ADN de todas las cepas se realizó con el kit DNeasy Blood & Tissue Kit. Para la secuenciación de los genomas se utilizó el equipo Illumina miseq y la preparación de las librerías con el Kit Nextera XT DNA simple preparation kit (Illumina).

Para llevar a cabo los experimentos *in silico* se utilizó una Workstation HP Z240 con memoria RAM de 32 GB y un procesador Intel® Xeon (R) CPU E5-1603.

B. MÉTODOS

1. Secuenciación de genomas

El genoma de las 153 cepas de *Salmonella enterica* aisladas de agua de río y de animales de granja fue secuenciado por sistema MiSeq desarrollado por Illumina, con una cobertura mínima de 40-120X, en la Food and Drug Administration (FDA)

2. Procesamiento de las secuencias (ensamble)

Una vez obtenidos los productos de secuenciación, se procedió a realizar un proceso de filtración de reads para la remoción de secuencias por debajo del valor 15 en la escala de calidad Phred, mediante la herramienta Trimmomatic v 0.36. Adicionalmente, se eliminaron secuencias de adaptadores que pudieron haber quedado como residuos del proceso de secuenciación. Cada genoma pre-tratado fue ensamblado por el método *de novo* mediante el uso la herramienta A5-miseq v 20160825 (Andrew Tritt and Aaron Darling). Las secuencias generadas fueron anotadas utilizando la herramienta Prokka v 1.11.

3. Descripción de la nomenclatura

Cada archivo fue renombrado con la intención de facilitar su gestión. Para ello se siguieron dos estilos diferentes. El primero de ellos se aplicó para las cepas internacionales descargadas de NCBI y consistió en la siguiente estructura:

SSSPPAANNNN

Donde las primeras tres letras (SSS) corresponden al serotipo (Cuadro 5), las siguientes dos (PP) a la abreviación del país donde se obtuvo la cepa (ver anexos), las siguientes dos (AA) corresponden al año de aislamiento y por último los cuatro caracteres (NNNN) indican un número consecutivo para mantener el orden y saber el número de cepas con las que se cuentan.

A continuación se muestra un ejemplo

ORAUS080015

Donde:

- ORA: Abreviación del serotipo.
- US: Abreviación del país donde se obtuvo la cepa.
- 08: Año de aislamiento representado en dos dígitos.
- 0015: Número consecutivo representado en cuatro dígitos.

Para las cepas locales se utilizó la siguiente estructura:

CY-SSSAANNN

Donde el primer carácter (C) corresponde a una etiqueta para indicar que pertenece a un aislado local. El segundo carácter denota la fuente de donde se obtuvo la cepa, pudiendo ser animal (denotado con una A) o de río (no se le adiciona carácter). Seguido, se agregan tres caracteres correspondientes al serotipo (Cuadro 5). Posteriormente, dos dígitos que corresponden al año de aislamiento y por último se agregan 4 caracteres correspondientes al número secuencial.

A continuación se muestra un ejemplo:

CA-SAI08001

Donde:

C: Indica que es una cepa local

A: Fue aislada de un animal

SAI: Pertenece al serotipo Saintpaul

08: Corresponde al año de aislamiento

001: Indica el número secuencial

4. Serotipificación *in silico*

Para realizar la serotipificación de los aislados de *Salmonella* de la región, se utilizó el programa *Salmonella In Silico Typing Resource* (SISTR). El cual integra la tipificación por métodos de análisis de tipificación multilocus de secuencias (MLST), MLST ribosomal (rMLST) y genoma core MLST (cMLST) (Yoshida y col. 2016).

5. Análisis del pan-genoma

Se identificaron las características de calidad de las secuencias procesadas y su respectivo tamaño de genoma, excluyendo aquellas que tuvieran un número de scaffolds mayor a 100 y un tamaño de genoma superior a 5 Mb. Se determinó el pan genoma de las cepas locales mediante la herramienta Roary. Adicionalmente, se generaron gráficas con información del conjunto de genes core y accesorios utilizando el script `roary_plots.py`

6. Construcción de un árbol filogenético (fast-tree)

Para generar árboles filogenéticos provisionales mediante la alineación del genoma core de las cepas internacionales y locales, se utilizó la herramienta Harvest V 1.1.2, construyéndose los siguientes árboles:

- Árbol filogenético global: el cual incluyó todas las cepas internacionales (447) y todas las locales (153) de los 5 serotipos utilizando un genoma al azar como referencia.

Cuadro 5. Abreviación de los serotipos utilizados para renombrar cepas de la región y las encontradas en NCBI.

| <i>Serotipo</i> | <i>Abreviación</i> |
|------------------------|---------------------------|
| Oranienburg | ORA |
| Saintpaul | SAI |
| Give | GIV |
| Minnesota | MIN |
| Anatum | ANA |

- Árbol filogenético local: para ello se utilizaron solamente las cepas locales de los 5 serotipos.
- Árbol filogenético para cada serotipo: utilizando cepas locales e internacionales para cada uno de los 5 serotipos respectivamente, generando 1 árbol filogenético para cada serotipo

Se utilizó FigTree v 1.4.2 para visualizar el árbol construido. Para la eliminación de los sitios con huecos (denominados gaps) del alineamiento múltiple de genomas generado, se utilizó la herramienta trimAl v 3.

7. Inferencias filogenéticas

Se realizaron mediante el método de máxima verosimilitud considerando los genes core y accesorios, usando el programa RAxML v 8.2.9 y el modelo GTRGAMMA (100 réplicas bootstrap). Las búsquedas subsecuentes para la elección del mejor árbol fue conducido por el modelo de aproximación GTRCAT.

Para la edición de los árboles filogenéticos generados, se utilizó el programa Interactive Tree Of Life v4 (iTOL)

8. Identificación de islas de patogenicidad y factores de virulencia

Para llevar a cabo la identificación de las islas de patogenicidad se utilizó la herramienta SPIFinder desarrollada por el Centro de Epidemiología Genómica (CGE) con un porcentaje de identidad mayor igual al 95% y con una longitud mínima del 60%, utilizando genomas ensamblados como datos de entrada. Los genes de virulencia fueron identificados mediante la utilización de ncbi-blast v 2.6.0 Para el análisis se utilizó un porcentaje de identidad del 95% y un e-value de $1*10^{-5}$

9. Identificación de secuencias pertenecientes a profagos en genomas locales

Se utilizó el programa PHASTER (Phage Search Tool Enhanced Release) desarrollado por la Universidad de Alberta, para la búsqueda e identificación de secuencias pertenecientes a profagos dentro de los genomas de aislados locales. Se utilizaron los parámetros predeterminados por la herramienta.

VIII. RESULTADOS Y DISCUSIÓN

A. Diversidad de los serotipos de *Salmonella* aisladas de la región

Con el propósito de caracterizar el serotipo de las distintas cepas de *Salmonella*, se procedió a realizar la serotipificación *in silico* mediante el uso de la plataforma SISTR, la cual integra una tipificación con base en MLST, rMLST y cgMLST.

Para llevar a cabo estos análisis, se tomaron 153 cepas aisladas previamente por el grupo de Investigación del Laboratorio de Microbiología Ambiental y de Alimentos del CIAD, A.C. coordinación regional Culiacán (Jiménez y col. 2008; Jiménez y col. 2011) de muestras de agua de río y heces frescas de animales de granja de la región de Culiacán, Sinaloa; y las cuales fueron previamente caracterizadas por pruebas serológicas

Los genomas de las cepas locales fueron secuenciados y ensamblados *de novo*, y posteriormente sometidos a una serie de análisis *in silico* para corroborar los serotipos determinados por los métodos serológicos. La tabla de los resultados con el serotipo y ST predicho para cada genoma se muestran en el anexo (Cuadro 22). Los resultados muestran que existe una incongruencia en la asignación del

serotipo en 11 cepas de 153 (Cuadro 6) con respecto a los resultados obtenidos por serotipificación convencional.

Cuadro 6. Aislados que mostraron incongruencia al ser serotipificados por técnicas convencionales y técnicas *in silico*.

| Nomenclatura | Serotipo predicho por serología | Serotipo predicho <i>in silico</i> | Sequence type (ST) |
|---------------------|--|---|---------------------------|
| C-THO08001 | Infantis | Thompson | 26 |
| C-ORA08005 | Anatum | Oranienburg | 23 |
| C-THO08002 | Infantis | Thompson | 26 |
| C-GIV08001 | Oranienburg | Give | 654 |
| C-SAN09001 | Oranienburg | Sandiego | ? |
| C-POM08001 | Oranienburg | Pomona | 451 |
| C-ORA08014 | Give | Oranienburg | 23 |
| C-POO09004 | Oranienburg | Poona | 447 |
| CA-THO08001 | C1 monofasica | Thompson | 26 |
| CA-THO08002 | C1 monofasica | Thompson | 26 |
| CA-THO08003 | C1 monofasica | Thompson | 26 |

Se identificaron 23 serotipos *in silico*, de los cuales se determinaron los predominantes, entre los que se incluyen: Oranienburg (42 cepas, 27.5%), Saintpaul (19, 12.4%), Give (16, 10.5%), Minnesota (15, 9.8%) y Anatum (8, 5.2%). La prevalencia de los otros serotipos estuvo en el rango del 0.7 al 4.6% (Cuadro 7). Durante este proyecto se llevó a cabo el análisis específico por biología computacional de los 5 serotipos predominantes.

Uno de los primeros pasos en la determinación de la diversidad genética de un grupo de cepas consiste en conocer el grado de similitud que presentan los aislados entre sí, de esta forma se puede evaluar de forma preliminar que el género y especie (en algunos casos hasta la variedad) asignados sean correctos.

Para comparar las similitudes genéticas entre los aislados de *Salmonella enterica* de la región, se realizó el cálculo del *average nucleotide identity* (ANI) mediante la utilización de la herramienta bioinformática ANItools. Este proceso permite agrupar las cepas de acuerdo a la similitud nucleotídica de las secuencias; incluso se ha reportado que permite distinguir bacterias muy relacionadas de acuerdo a la divergencia del ADN (Han y col. 2016). Esta herramienta taxonómica involucra la fragmentación de las secuencias, seguido de una búsqueda de las secuencias

nucleotídicas, un alineamiento y posteriormente el cálculo del % de identidad (Ritcher y Rosselló-Móra, 2018).

A continuación se presenta la gráfica del porcentaje de identidad para los distintos serotipos, donde las cepas pertenecientes al mismo serotipo se encuentran delimitadas por rectángulos de los distintos colores (Figura 14).

Cuadro 7. Distribución de serotipos de los aislados de *Salmonella*.

| Serotipos de <i>Salmonella</i> | No. de aislamientos | | Aislados totales | Porcentaje (%) |
|--------------------------------|---------------------|-----------|------------------|----------------|
| | Río | Animal | | |
| Oranienburg | 24 | 18 | 42 | 27.5 |
| Saintpaul | 10 | 9 | 19 | 12.4 |
| Give | 5 | 11 | 16 | 10.5 |
| Minnesota | 7 | 8 | 15 | 9.8 |
| Anatum | 4 | 4 | 8 | 5.2 |
| Weltevreden | 4 | 3 | 7 | 4.6 |
| Poona | 6 | 0 | 6 | 3.9 |
| Pomona | 6 | 0 | 6 | 3.9 |
| Thompson | 2 | 3 | 5 | 3.3 |
| Montevideo | 4 | 1 | 5 | 3.3 |
| Muenster | 5 | 0 | 5 | 3.3 |
| Luciana | 0 | 3 | 3 | 2.0 |
| Infantis | 3 | 0 | 3 | 2.0 |
| Newport | 0 | 2 | 2 | 1.3 |
| Agona | 0 | 2 | 2 | 1.3 |
| Sahanina | 0 | 2 | 2 | 1.3 |
| Sandiego | 1 | 0 | 1 | 0.7 |
| Albany | 0 | 1 | 1 | 0.7 |
| Javiana | 0 | 1 | 1 | 0.7 |
| Muenchen | 0 | 1 | 1 | 0.7 |
| Cayar | 0 | 1 | 1 | 0.7 |
| Typhimirium | 0 | 1 | 1 | 0.7 |
| Gaminara | 0 | 1 | 1 | 0.7 |
| TOTAL | 81 | 72 | 153 | |

Como se puede observar en la Figura 14, la diagonal principal se encuentra marcada con el color rojo intenso (se utilizó una línea negra punteada para facilitar la identificación) que corresponde a la escala 1 y denota que las secuencias son idénticas. Esto se debe a que compara todas las secuencias contra todas, por lo que dentro del análisis se llega a comparar cualquier secuencia contra sí misma y dicha similitud (100%) es representada con el rojo más intenso, el cual disminuye gradualmente su intensidad de acuerdo al porcentaje de similitud.

Se pueden observar zonas de tonalidad homogénea que coinciden con la agrupación de los serotipos en el árbol. Este fenómeno habla de la poca diversidad que existe entre las cepas del mismo serotipo en todo el universo de muestras.

Diferentes investigadores han determinado un umbral típico del 95% como límite de especie (ya que este valor es equivalente al 70% utilizado en hibridación ADN-ADN para la delimitación de especies) (Richter y Rossello-Mora, 2009). Los resultados indican que los porcentajes de nuestras cepas se encuentran en el rango del 95 al 99%; por lo que estos resultados refuerzan la taxonomía descrita previamente, la cual indica que las cepas analizadas pertenecen al género *Salmonella* y a la especie *enterica*.

No se ha reportado que la técnica del ANI sirva para explorar el nivel taxonómico “serotipo” entre los diferentes aislados (Ozen y col. 2013). Sin embargo, nuestros resultados sugieren que esta técnica posee suficiente sensibilidad para discriminar entre los genomas analizados siguiendo un patrón coincidente con la agrupación de los serotipos predichos.

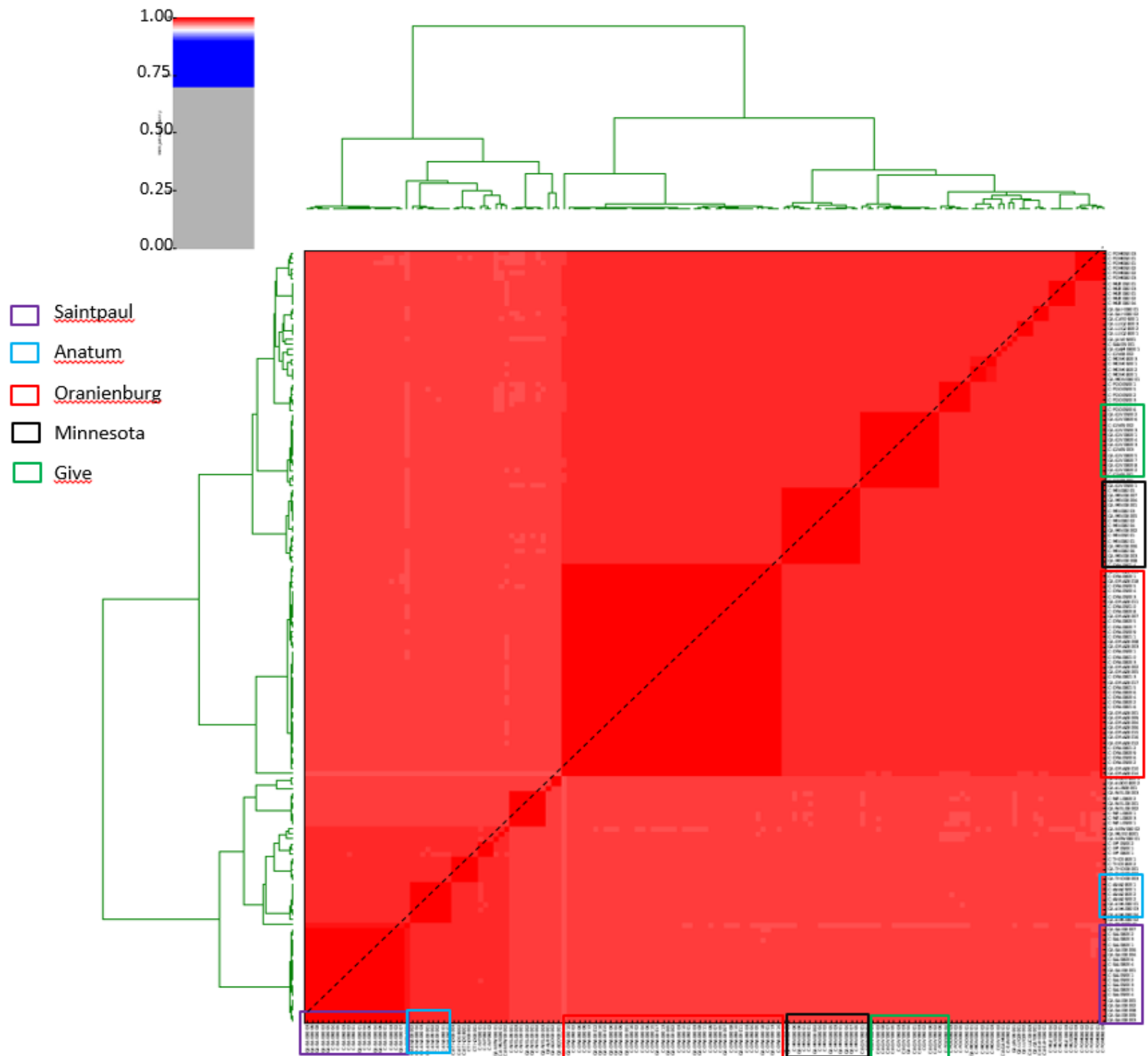


Figura 14. Heatmap del porcentaje de identidad entre los aislados de la región de Culiacán, Sinaloa

Para explorar con mayor detalle la variabilidad genética y poder establecer relaciones filogenéticas entre las cepas, se procedió a construir un árbol filogenético mediante el algoritmo de máxima verosimilitud para los aislados locales (Figura 15). Como se puede apreciar, las distintas cepas se agrupan en clados; siendo estos específicos para el serotipo correspondiente. Por ejemplo, el clado I (denotado en color morado) agrupa a todas las cepas pertenecientes al serotipo Saintpaul; mientras que el clado II (azul) agrupa a las cepas que integran al serotipo Anatum. Lo que demuestra una alta congruencia en la agrupación de los taxones según el serotipo perteneciente.

En la misma figura, la línea exterior denota el ST obtenido para cada cepa. Para la mayoría de los serotipos se obtuvo un único ST, siendo el serotipo Give la única excepción en la comparación de aislados locales; por ejemplo, para el serotipo Oranienburg se obtuvo el ST23 como prevalente. De acuerdo a lo establecido por Urwin y Maiden en el 2003, los aislados que presentan un mismo ST se encuentran estrechamente relacionadas y pertenecen al mismo grupo clonal. Sin embargo, las cepas pertenecientes al serotipo Give presentan dos grupos ST, siendo el ST654 el genotipo central, del cual pueden derivar grupos minoritarios cercanos; como es el caso del ST516 (Urwin, 2003). En el Cuadro 8 se presentan los diferentes serotipos con sus respectivos grupos ST y perfiles alélicos.

Cuadro 8. Serotipos predominantes en la región con su ST y perfil alélico.

| Serotipo | Sequence type (ST) | Perfil alélico | | | | | | |
|--------------------|--------------------|----------------|------|------|------|------|------|------|
| | | aroC | dnaN | hemD | hisD | purE | sucA | thrA |
| Oranienburg | 23 | 13 | 11 | 15 | 15 | 12 | 15 | 4 |
| Saintpaul | 50 | 5 | 21 | 18 | 9 | 6 | 12 | 17 |
| Minnesota | 548 | 13 | 11 | 25 | 197 | 12 | 71 | 4 |
| Give | 654 | 111 | 47 | 49 | 42 | 12 | 58 | 3 |
| | 516 | 84 | 11 | 16 | 42 | 40 | 71 | 4 |
| Anatum | 64 | 10 | 14 | 15 | 31 | 25 | 20 | 33 |

Los resultados sugieren que en la región de Culiacán existe un alto grado de conservación genética entre los aislados. Esto se comprueba ya que se obtuvo un solo ST representativo para la mayoría de los serotipos; con excepción del serotipo Give, en el cual se identificaron dos ST. Asimismo se observó que la diversidad de STs se encuentra asociado con el serotipo (no se encuentra un mismo ST representado a más de un serotipo, aunque si se encuentran distintos ST caracterizando a un mismo serotipo), lo que indica una relación cercana entre estas características, información que concuerda por Liu y col. en el 2011.

Estos resultados sugieren que el uso de los grupos clonales (que integran perfiles de ST relacionados) puede complementar la serotipificación en el proceso de discriminación de cepas patógenas, incrementando o potenciando la resolución intra-específica. Asimismo, el poder de resolución otorgado por la utilización de genomas completos para la serotipificación *in silico* de *Salmonella* puede superar las limitaciones de la serotipificación tradicional, dando lugar a una identificación confiable de los serotipos para este género taxonómico (Achtman y col. 2012).

B. Relaciones filogenéticas de los serotipos dominantes en la región con respecto a las reportadas en diferentes países.

Con el objetivo de colocar a los aislados locales en el contexto de la diversidad genética global dentro de los mismos serotipos de *Salmonella*, y profundizar respecto al posible origen de los grupos clonales encontrados, se construyó un árbol filogenético empleando el algoritmo de máxima verosimilitud (ML), donde se compararon los genomas de los aislados locales con secuencias internacionales de los 5 serotipos predominantes de *Salmonella* (Oranienburg, Anatum, Give,

Minnesota y Saintpaul), las cuales fueron previamente descargadas de NCBI (Cuadro 9).

A continuación se presenta el árbol generado para la monofilia de *S. enterica* subsp. *enterica* que englobó los 550 genomas contenidos en la base de datos generada (Figura 16).

Los resultados muestran que cada cepa se agrupa dentro del clado perteneciente a su serotipo específico. Sin embargo, existe una región (denotada con un rectángulo rojo) en la que se aprecian diferentes cepas que no se agruparon con el clado correspondiente a su serotipo. Esto se debe a que las cepas fueron rotuladas de acuerdo al serotipo reportado originalmente en la base de datos por los diversos autores mediante el uso de técnicas moleculares; pero al realizar la serotipificación *in silico* se encontró que pertenecían a otro serotipo distinto al previamente asignado. Las razones del error se puede atribuir a que los protocolos de serotipificación empleados en dichos casos, están sujetos a problemas metodológicos como el empleo de gran cantidad de sueros específicos, su realización requiere personal capacitado y la falta de sensibilidad intra-específica (González y col. 2010); además son técnicas laboriosas que necesitan ser realizadas por personas capacitadas (Jenkins, 2015). Por lo tanto, dichas cepas se agruparon en otro clado.

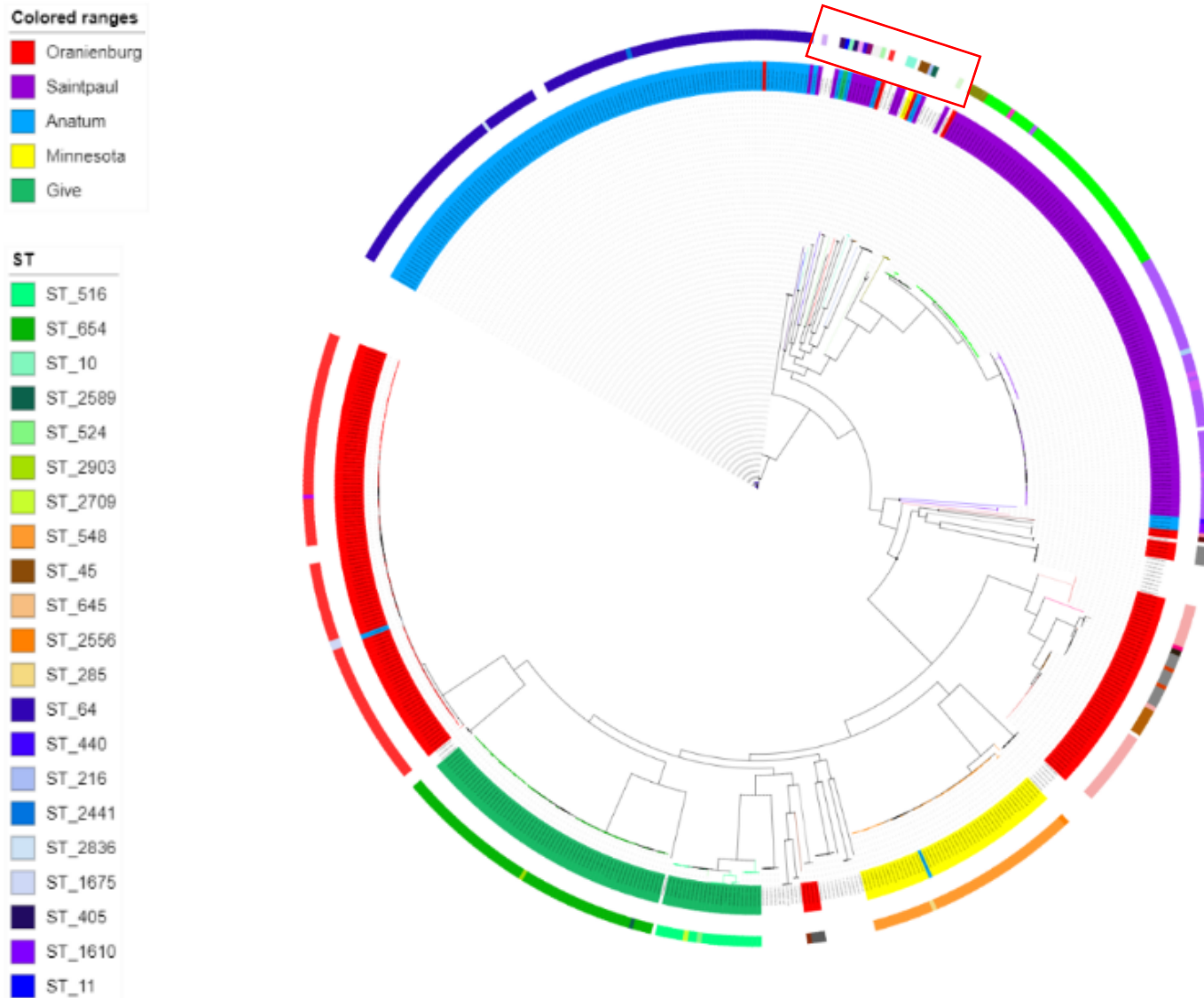


Figura 16. Filogenética global de los serotipos Oranienburg, Give, Minnesota, Saintpaul y Anatum de *Salmonella enterica* subsp. *enterica* dominantes en la región. Se presenta un árbol filogenético construido por medio del método de máxima verosimilitud, utilizando cepas locales e internacionales. Los colores sobre la leyenda indican el serotipo, mientras que los colores del borde exterior indican el grupo ST

Cuadro 9. Relación de las cepas locales e internacionales de los diferentes serotipos de *Salmonella enterica*.

| SEROTIPO | LOCALES | INTERNACIONALES |
|-----------------|----------------|------------------------|
| Oranienburg | 44 | 119 |
| Saintpaul | 19 | 106 |
| Give | 16 | 69 |
| Minnesota | 16 | 34 |
| Anatum | 8 | 119 |
| TOTAL | 103 | 447 |

Prueba de ello es la cepa rotulada como SAIUS090042, la cual fue reportada como perteneciente al serotipo Saintpaul, pero al analizarla *in silico* se determinó que corresponde al serotipo Litchfield y no al reportado. Es por ello que se puede apreciar una diversidad de grupos ST que pertenecen a diferentes serotipos y no a los predichos por técnicas moleculares.

Estos resultados, al igual que los mencionados previamente para las cepas locales, sugieren que la realización de análisis filogenéticos y el empleo de herramientas bioinformáticas pudieran tener potencial en la determinación de serotipos de especies bacterianas.

Adicionalmente, se encontró que los serotipos Oranienburg y Give se encuentran organizados en grupos parafiléticos, lo que coincide con reportes donde se ha observado una alta divergencia para estos serotipos (Didelot y col. 2011).

Con la finalidad de explorar con mayor detalle la filogenia observada en el árbol global y la evolución de los distintos serotipos de relevancia local, se llevó a cabo la construcción de árboles filogenéticos por serotipos. A continuación se detallan los resultados para cada serotipo.

1. Filogenia global del serotipo Oranienburg.

De acuerdo a la Figura 17, se puede observar que los aislados de *S. Oranienburg* de la región se agrupan dentro del clado 1. Siendo el grupo ST23 el representativo, aunque también incluye el ST1675 siendo un *single locus variant* (SLV) del ST23 por diferir en un alelo; mientras que el ST47 es un SLV con respecto al ST23 por diferir en un alelo (Cuadro 10). Además estos tres grupos de ST, pertenecen al mismo eBURST group (eBG) 44, indicando una relación cercana en el perfil alélico de dichos ST. El clado 1 también incluye a los ST1576 y ST864, los cuales no comparten ningún alelo con el ST23. No se ha reportado eBG para el grupo ST864, pero nuestros resultados sugieren que se encuentra muy relacionado filogenéticamente con el grupo eBG 203 (perteneciente al ST1576).

Adicionalmente, los resultados sugieren que las cepas locales pueden tener dos posibles orígenes. El primero de ellos se debe a comparte un ancestro común más probable con cepas procedentes de Estados Unidos, Ecuador y Reino Unido, lo que sugiere que el grupo de aislados locales se pudo derivar de dicho grupo. El segundo posible origen, comparte ancestro con cepas de Estados Unidos, Reino Unido y Corea del Sur, por lo que este linaje se pudo derivar de este grupo.

En la Figura 15, también se puede apreciar la formación de otros dos clados. El segundo clado se encuentra conformado en su mayoría por cepas del Reino Unido. El ST predominante en dichas cepas es el 174, siendo SLV del ST1515 y ST169, mientras que es DLV del ST320. Así mismo, se determinó que estos ST pertenecen al eBG 44. En el caso del ST2575, se encontró que difería en 3 alelos

con respecto al ST174 (Cuadro 10), aunque no se ha determinado el eBG del ST2575, nuestros

- Locales
- México
- Estados Unidos
- Ecuador
- Reino Unido
- Korea del Sur

- ST**
- ST_23
 - ST_40
 - ST_174
 - ST_1675
 - ST_64
 - ST_1517
 - ST_864
 - ST_169
 - ST_1515
 - ST_2575
 - ST_1510
 - ST_33
 - ST_320
 - ST_45
 - ST_1576
 - ST_47

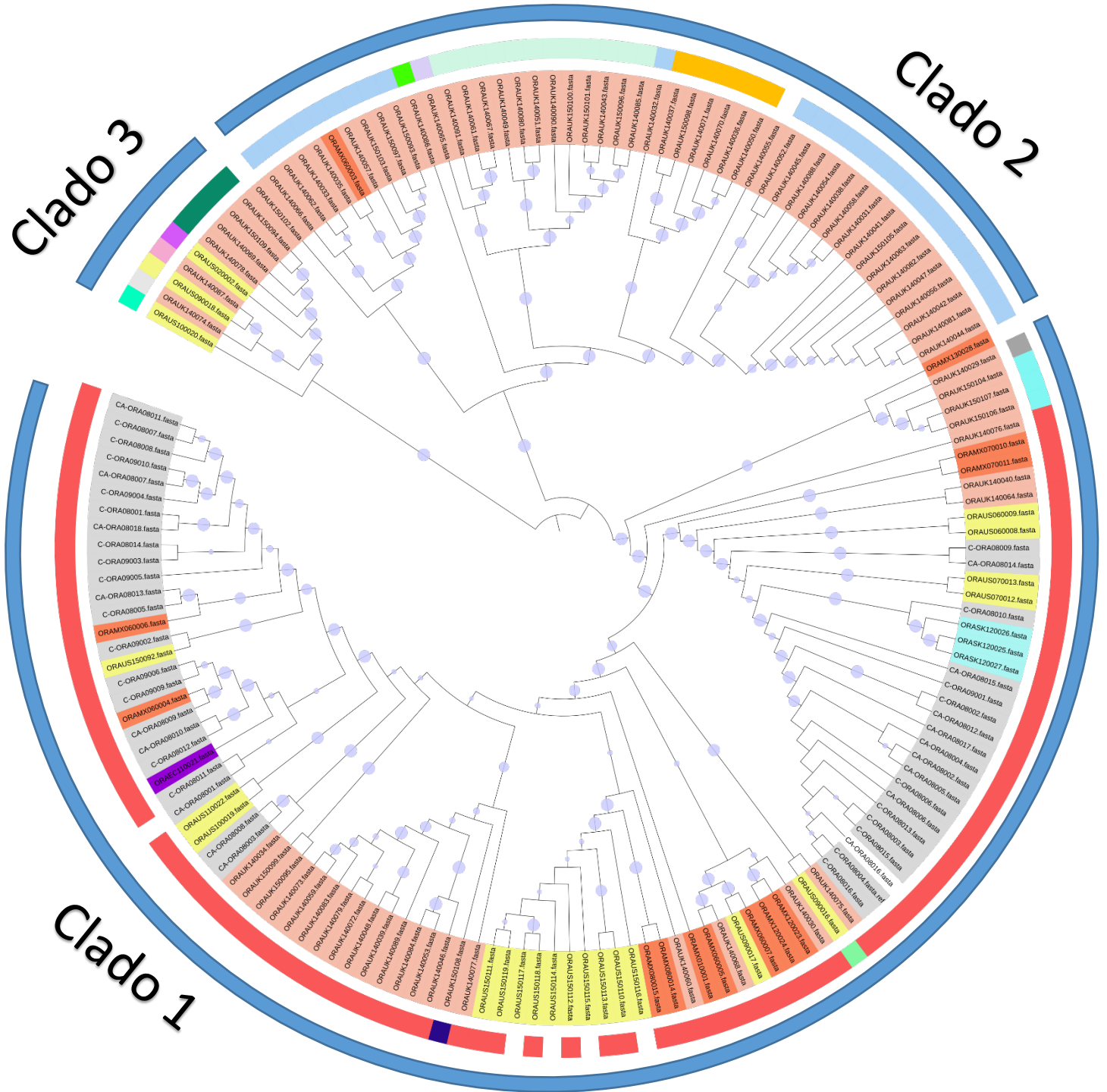


Figura 17. Árbol filogenético con aislados locales e internacionales de *Salmonella enterica* serotipo Oranienburg.

Cuadro 10. Perfil alélico de los diferentes grupos ST analizados del serotipo Oranienburg.

| ST | EBG | aroC | dnaN | hemD | hisD | purE | sucA | thrA |
|-------------|-----|------|------|------|------|------|------|------|
| 23 | 41 | 13 | 11 | 16 | 15 | 12 | 15 | 4 |
| 1675 | 41 | 13 | 11 | 315 | 15 | 12 | 15 | 4 |
| 47 | 41 | 21 | 11 | 16 | 15 | 12 | 15 | 4 |
| 1576 | 203 | 43 | 41 | 16 | 493 | 40 | 58 | 3 |
| 864 | - | 301 | 272 | 48 | 16 | 40 | 13 | 3 |
| 174 | 44 | 50 | 63 | 25 | 72 | 59 | 71 | 63 |
| 1515 | 44 | 50 | 63 | 25 | 72 | 59 | 292 | 63 |
| 169 | 44 | 50 | 63 | 25 | 72 | 59 | 71 | 62 |
| 320 | 44 | 13 | 63 | 25 | 72 | 59 | 118 | 63 |
| 2575 | - | 50 | 63 | 25 | 72 | 40 | 292 | 63 |
| 1510 | - | 41 | 86 | 215 | 358 | 89 | 109 | 48 |
| 45 | 3 | 10 | 7 | 21 | 14 | 15 | 12 | 12 |
| 33 | 22 | 2 | 5 | 6 | 7 | 5 | 7 | 12 |
| 64 | 65 | 10 | 14 | 15 | 31 | 25 | 20 | 33 |
| 40 | 57 | 19 | 20 | 3 | 20 | 5 | 22 | 22 |
| 1517 | - | 10 | 391 | 8 | 494 | 216 | 132 | 172 |

resultados sugieren que este grupo indeterminado presenta cercanía filogenética con el eBG 44.

En el clado 3 se agrupan cepas que no corresponden al serotipo Oranienburg. Esto se debe a que el proceso de renombre para cepas internacionales, se realizó con base a los serotipos reportados por los diferentes investigadores quienes utilizaron técnicas serológicas de laboratorio; sin embargo al realizar la serotipificación *in silico*, se determinó que pertenecían a otro serotipo y no al previamente reportado. Por lo que estos serotipos presentan diferencia en todos los alelos y no comparten eBG.

2. Filogenia global del serotipo Saintpaul

Las cepas de *Salmonella* aisladas de la región se agrupan dentro del clado 2 y tienen como ST representativo al 50 (Figura 18). Adicionalmente, este clado incluyen cepas pertenecientes al ST27, ST49 y ST2438; siendo todos estos SLV del ST50 y pertenecientes al eBG 14, con excepción el ST2438 para el cual no se ha determinado eBG pero nuestros resultados sugieren un eBG filogenéticamente cercano al 14 (Cuadro 11).

Interesantemente, los aislados locales comparten un ancestro común más reciente con cepas procedentes de Estados Unidos y de otras localidades de México, por lo que se sugiere que el origen de los aislados pudo derivar de estos países.

En el clado 1 se agrupan cepas pertenecientes a países como Estados Unidos, India y Francia. Teniendo al ST27 como representativo; sin embargo se incluyen grupos

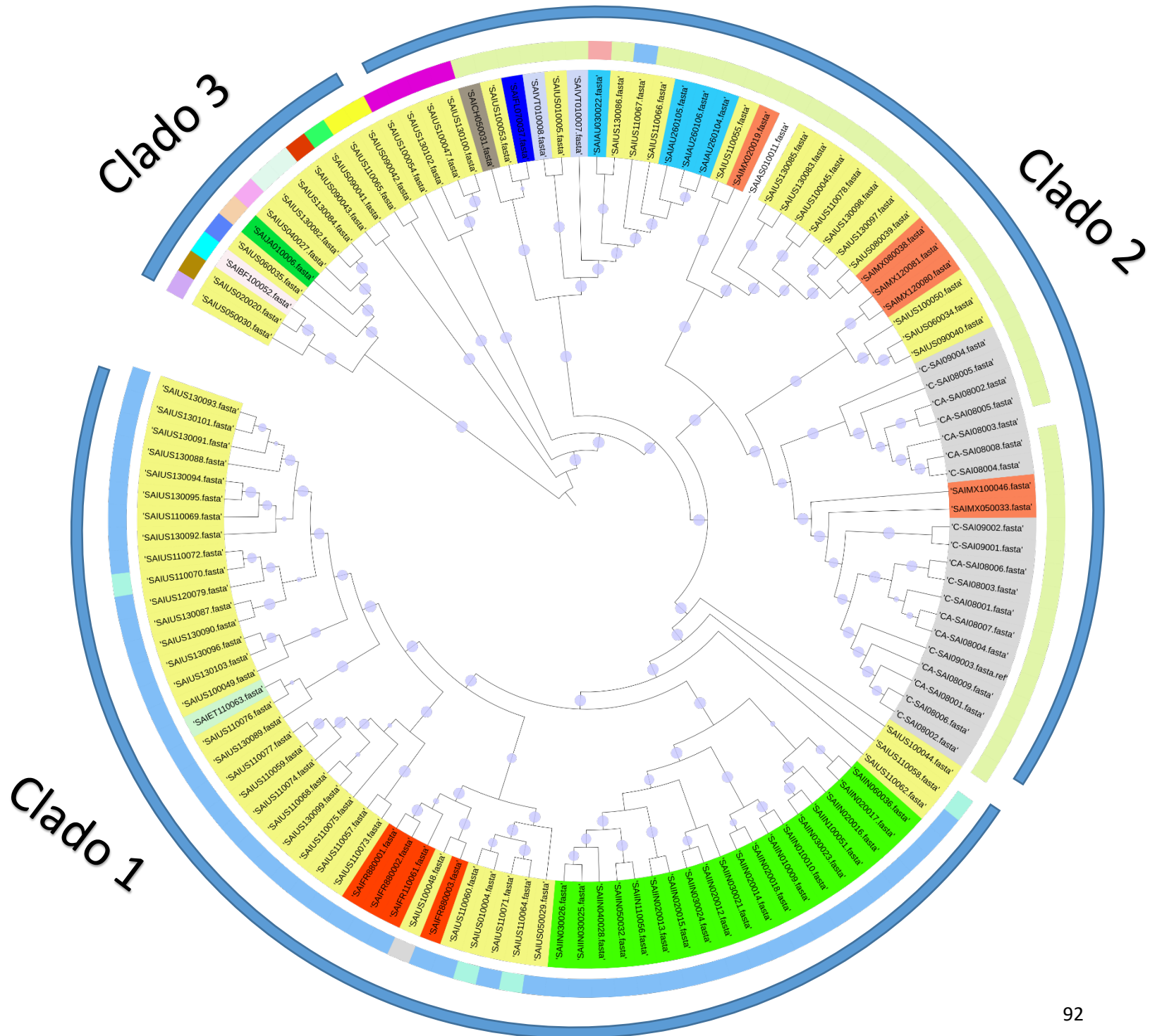
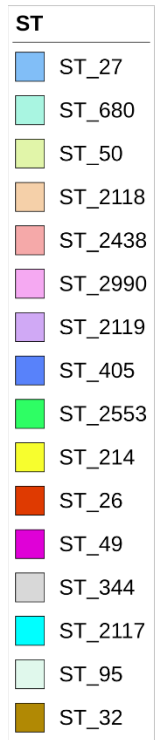


Figura 18. Árbol filogenético con aislados locales e internacionales de *Salmonella enterica* serotipo Saintpaul.

Cuadro 11. Perfil alélico de los diferentes grupos ST analizados del serotipo Saintpaul.

| ST | eBG | aroC | dnaN | hemD | hisD | purE | sucA | thrA |
|------|-----|------|------|------|------|------|------|------|
| 50 | 14 | 5 | 21 | 18 | 9 | 6 | 12 | 17 |
| 49 | 14 | 5 | 14 | 21 | 9 | 6 | 12 | 17 |
| 27 | 14 | 5 | 14 | 18 | 9 | 6 | 12 | 17 |
| 2438 | - | 5 | 21 | 18 | 648 | 6 | 12 | 17 |
| 680 | 14 | 201 | 14 | 18 | 9 | 6 | 12 | 17 |
| 344 | 14 | 5 | 14 | 18 | 9 | 6 | 12 | 116 |
| 214 | 61 | 14 | 72 | 21 | 12 | 6 | 19 | |
| 2117 | - | 18 | 7 | 35 | 9 | 5 | 9 | 12 |
| 2990 | - | 3 | 36 | 43 | 38 | 16 | 577 | 38 |
| 26 | 28 | 14 | 13 | 18 | 12 | 14 | 18 | 1 |
| 2553 | - | 543 | 69 | 101 | 12 | 370 | 130 | 347 |
| 2119 | - | 2 | 31 | 10 | 62 | 14 | 19 | 34 |
| 32 | 31 | 17 | 18 | 22 | 17 | 5 | 21 | 19 |
| 405 | 169 | 2 | 15 | 21 | 138 | 130 | 139 | 12 |
| 2118 | - | 46 | 71 | 365 | 130 | 6 | 162 | 115 |
| 95 | 209 | 3 | 36 | 43 | 38 | 16 | 42 | 38 |

minoritarios como el ST680 y ST344. Siendo ambos SLV del ST27 y estando incluidos los tres ST en el grupo eBG 14.

De manera general podemos observar que los grupos ST27 y ST50 se encuentran distribuidos de manera internacional. Información que resulta útil para estudios de epidemiología, donde se requiere identificar agrupaciones de cepas con alta similitud genética.

Por su parte, en el clado 3 se puede apreciar una diversidad en cuanto a la cantidad de ST (Figura 18), esto se debe a que pertenecen a un serotipo diferente el cual fue caracterizado mediante técnicas de serotipificación *in silico*.

3. Filogenia global del serotipo Give

Como se puede observar en la Figura 19, las cepas locales se agrupan mayoritariamente dentro del clado 1. Siendo el ST654 dominante; sin embargo, se incluyen grupos minoritarios como el ST2903 (presentando diferencia en dos alelos con respecto al ST654) y ST589 (presentando diferencia en todos los alelos) (Cuadro 12). Para el caso del ST2903 no se ha determinado su eBG, pero nuestros resultados sugieren una cercanía filogenética con el grupo 67.

Adicionalmente, los resultados muestran que las cepas locales comparten el ancestro común más reciente con aislados procedente de países como Estados Unidos, Brasil, Indonesia y con cepas de México. Por lo que el origen de este grupo clonal de la región de Culiacán, Sinaloa pudiera derivar de ahí.

Solo una cepa local se agrupó dentro del clado 2. Donde el ST516 figura como dominante, pero se incluyen el ST2709 (presentando diferencia en un alelo con respecto al ST516) y ST524 (presentando dos diferencias en el perfil) como minoritarios. No se logró identificar el eBG para el ST2709, pero debido a la cercanía filogenética se cree que pueda pertenecer al grupo 67.

Esta única cepa comparte un ancestro común más probable con cepas reportadas por India, Estados Unidos, México y Reino Unido principalmente.

La filogenia global para este serotipo sugiere la existencia de dos grupos clonales principales. Uno de ellos representado por el ST654 y encontrado principalmente en países del continente Americano; mientras que el segundo grupo se caracteriza por el ST516 y su distribución en países Europeos, Asiáticos y Americanos.

4. Filogenia global del serotipo Minnesota

Los aislados locales del serotipo Minnesota se agrupan tanto en el clado 1 como en el 2. El primer clado se compone principalmente del ST548 (Figura 20), pero también incluye al ST285 (presentando diferencia en un alelo con respecto al ST548), ST2556 (diferencia en dos alelos) y ST45 (diferencia en todos los alelos). Adicionalmente se encontró que los ST548 y ST285 pertenecen al mismo eBG; para el ST2556 no ha sido determinado, pero los resultados sugieren un eBG filogenéticamente cercano al 77. Para el caso del ST45, se determinó que pertenecía al serotipo Newport, por lo que muestra diferencia en todos los alelos y en el eBG (Cuadro 13).

| Colored ranges | |
|--|----------------|
| ■ | India |
| ■ | Malasia |
| ■ | Malta |
| ■ | México |
| ■ | China |
| ■ | Locales |
| ■ | Reino Unido |
| ■ | Ghana |
| ■ | Estados Unidos |
| ■ | Brasil |
| ■ | Costa Rica |
| ■ | Honduras |
| ■ | Indonesia |

| ST | |
|---|---------|
| ■ | ST_516 |
| ■ | ST_654 |
| ■ | ST_524 |
| ■ | ST_589 |
| ■ | ST_2903 |
| ■ | ST_2709 |

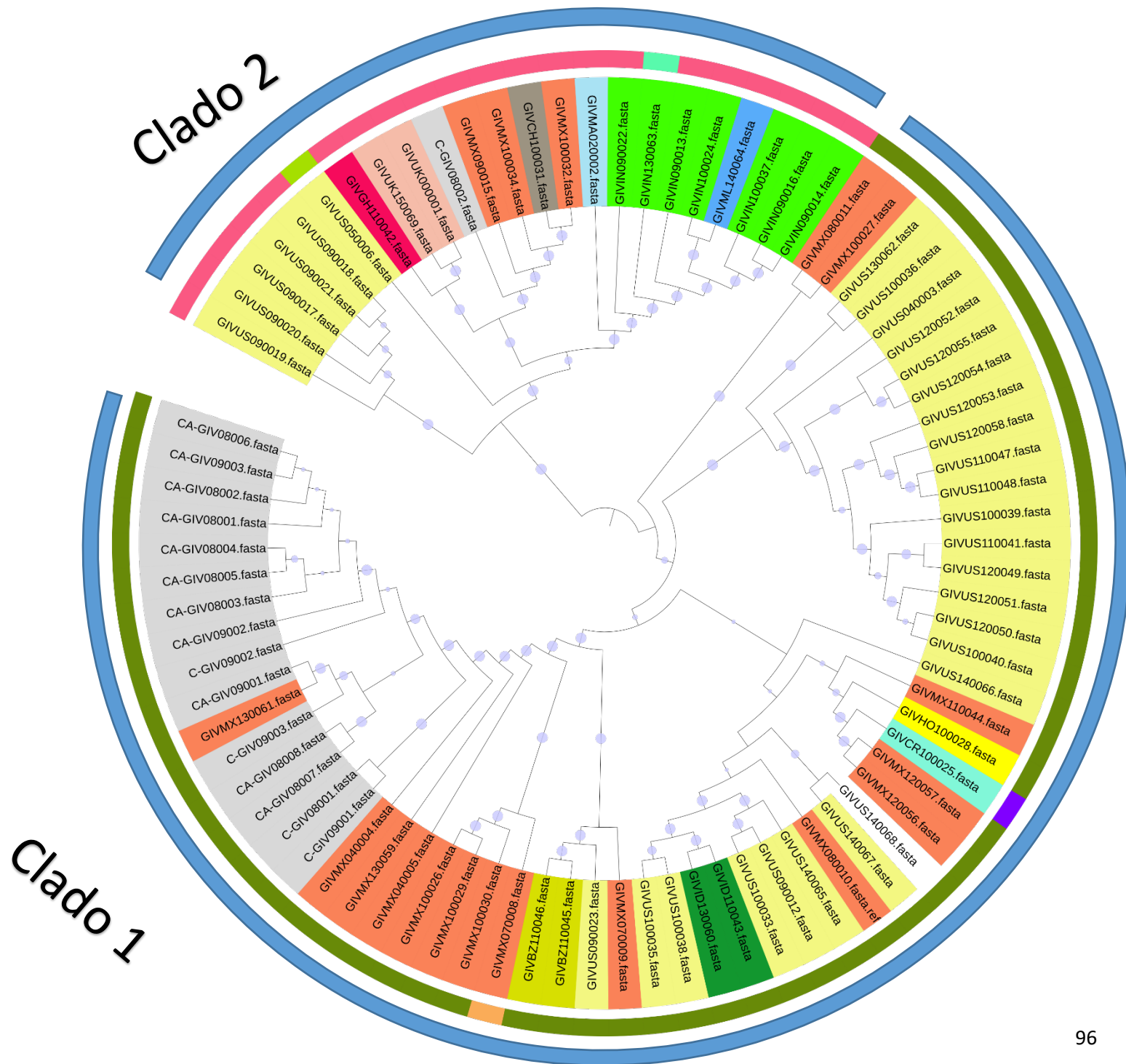


Figura 19. Árbol filogenético con aislados locales e internacionales de *Salmonella enterica* serotipo Give.

Cuadro 12. Perfil alélico de los diferentes grupos ST analizados del serotipo Give.

| ST | eBG | aroC | dnaN | hemD | hisD | purE | sucA | ThrA |
|-----------|------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 654 | 239 | 111 | 47 | 49 | 42 | 12 | 58 | 3 |
| 2903 | - | 111 | 47 | 480 | 738 | 12 | 58 | 3 |
| 589 | 17 | 187 | 12 | 17 | 16 | 13 | 16 | 4 |
| 516 | 67 | 84 | 11 | 16 | 42 | 40 | 71 | 4 |
| 2709 | - | 84 | 11 | 444 | 42 | 40 | 71 | 4 |
| 524 | 67 | 84 | 150 | 16 | 42 | 40 | 166 | 4 |

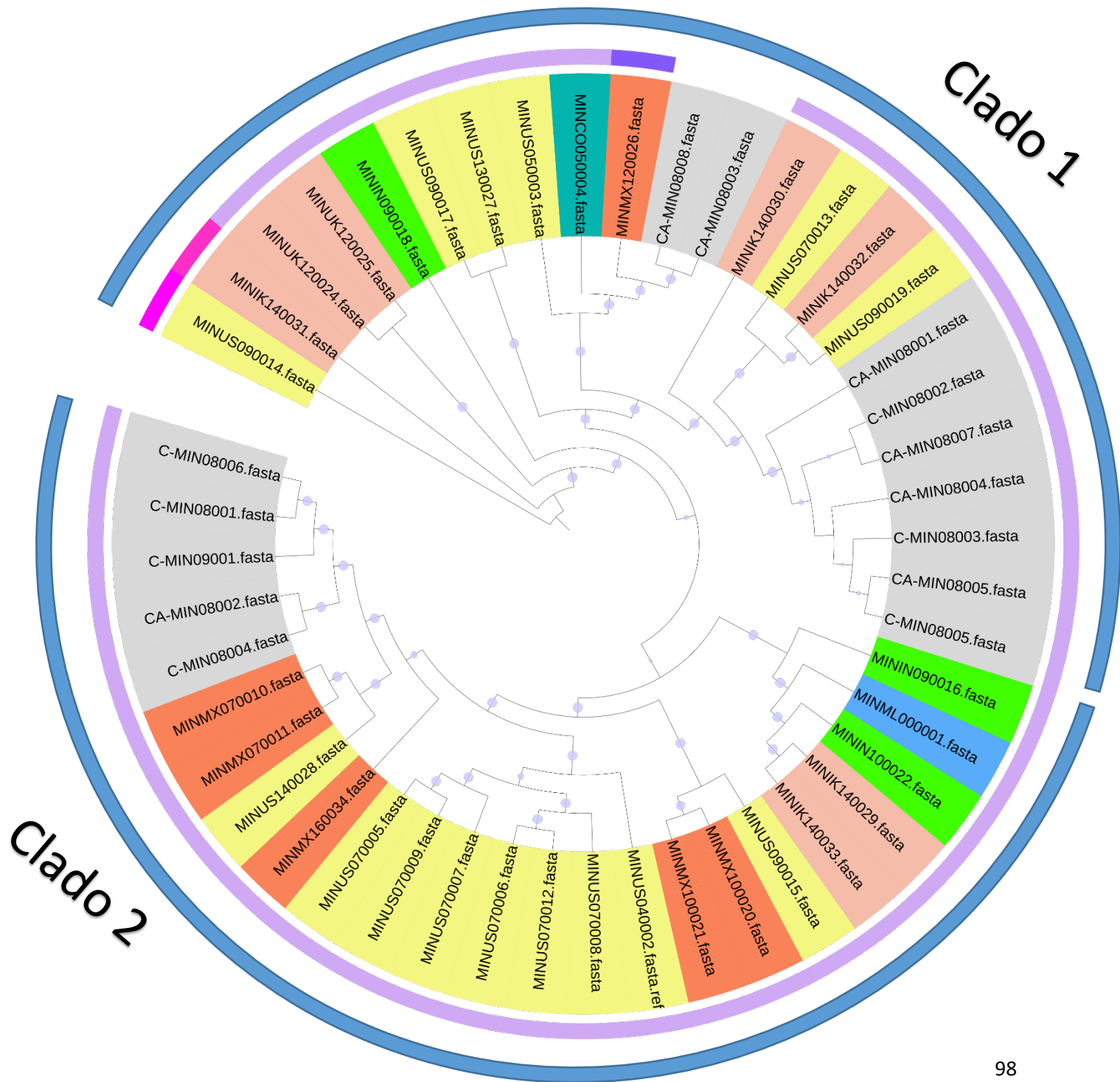
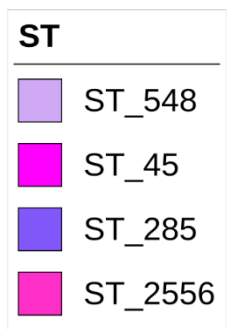


Figura 20. Árbol filogenético con aislados locales e internacionales de *Salmonella enterica* serotipo Minnesota.

Cuadro 13. Perfil alélico de los diferentes grupos ST analizados del serotipo Minnesota.

| ST | eBG | aroC | dnaN | hemD | hisD | purE | sucA | ThrA |
|-------------|------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 548 | 77 | 13 | 11 | 25 | 197 | 12 | 71 | 4 |
| 285 | 77 | 13 | 11 | 25 | 102 | 12 | 71 | 4 |
| 2556 | | 13 | 11 | 25 | 197 | 12 | 71 | 83 |
| 45 | 3 | 10 | 7 | 21 | 14 | 15 | 12 | 12 |

El origen de los aislados locales comparte ancestro común más probable con cepas de Estados Unidos, Colombia y Reino Unido. Por lo que se piensa que este grupo se pudo haber originado de allí.

El clado 2 se compone en su totalidad del ST548, donde el ancestro común más probable esta compartido con Estados Unidos y cepas reportadas en otros puntos de muestreo de México.

Adicionalmente, estos resultados sugieren que existe poca variabilidad a nivel internacional del serotipo Minnesota; ya que el ST548 se encuentra ampliamente distribuido y figura como el dominante (el 88% de los aislados pertenecen al ST548).

5. Filogenia global del serotipo Anatum

De acuerdo a la Figura 21, los aislados locales se agrupan dentro del clado 2. Siendo el ST64 el dominante, con excepción de una cepa perteneciente al ST2441. Siendo esta SLV del ST64; aunque no se ha determinado su grupo, nuestros resultados sugieren que se encuentra filogenéticamente emparentado con el eBG 65 (Cuadro 14).

Nuestro árbol filogenético sugiere que las cepas locales pueden provenir de Estados Unidos, dada la gran diversidad de cepas norteamericanas y a que los ancestros comunes de las cepas locales caen dentro de esta diversidad.

Que las cepas locales pudieron tener un origen común con aislados procedentes de países como Estados Unidos, Colombia, México, China y Vietnam. Esto se debe a que comparten un ancestro común más probable.

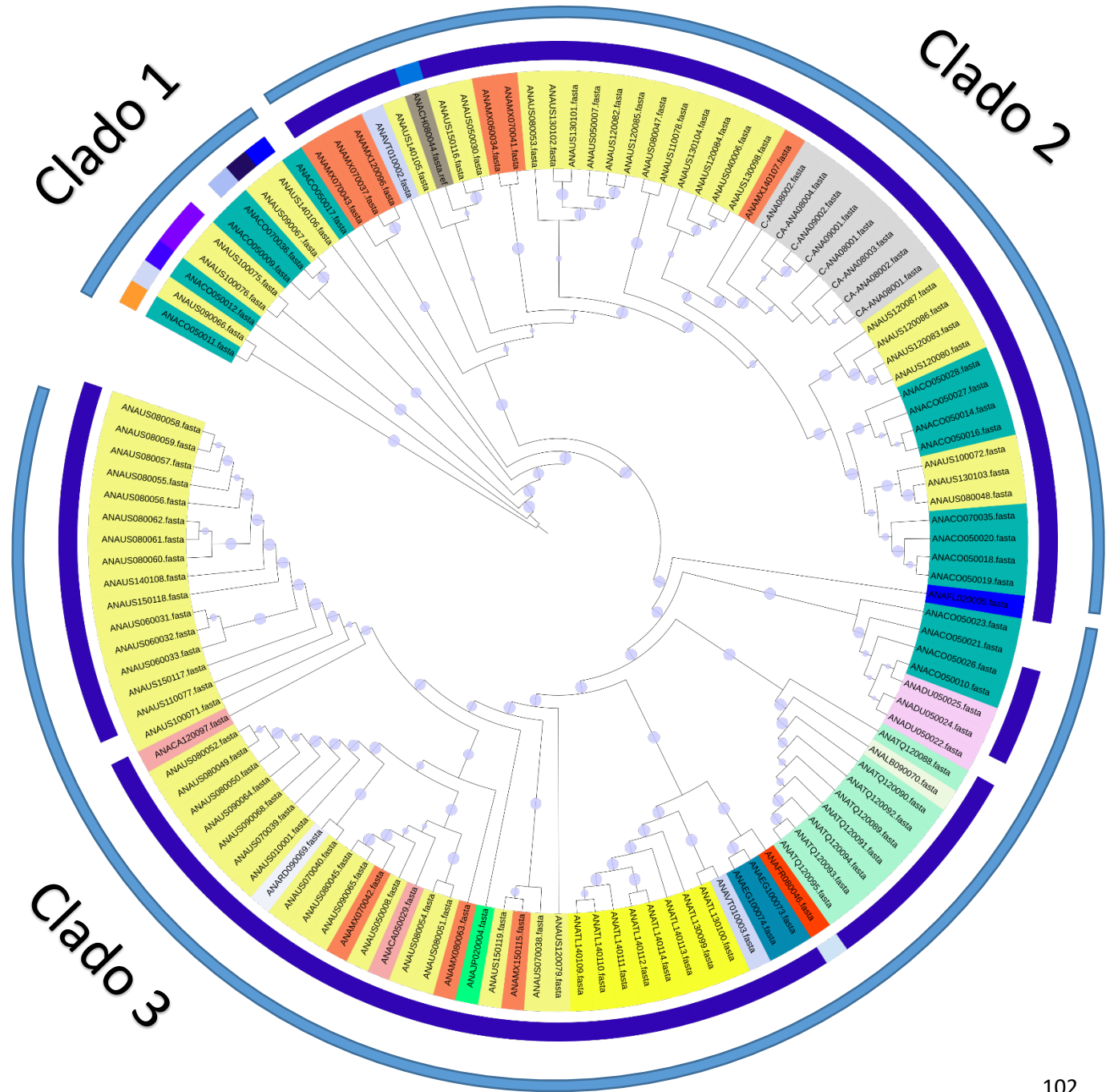
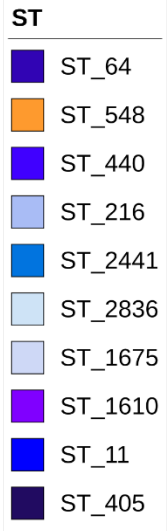


Figura 21. Árbol filogenético con aislados locales e internacionales de *Salmonella enterica* serotipo Anatum.

Cuadro 14. Perfil alélico de los diferentes grupos ST analizados del serotipo Anatum.

| ST | EBG | AroC | dnaN | hemD | hisD | purE | sucA | thrA |
|-------------|-----|------|------|------|------|------|------|------|
| 64 | 65 | 10 | 14 | 15 | 31 | 25 | 20 | 33 |
| 2441 | - | 10 | 500 | 15 | 31 | 25 | 20 | 33 |
| 2836 | - | 10 | 539 | 15 | 31 | 25 | 20 | 33 |
| 548 | 77 | 13 | 11 | 25 | 197 | 12 | 1 | 4 |
| 1675 | 41 | 13 | 11 | 315 | 15 | 12 | 15 | 4 |
| 440 | 82 | 140 | 105 | 85 | 93 | 14 | 10 | 3 |
| 216 | 25 | 2 | 73 | 41 | 79 | 71 | 65 | 72 |
| 1610 | - | 92 | 14 | 18 | 506 | 6 | 386 | 14 |
| 11 | 22 | 6 | 5 | 6 | 7 | 5 | 7 | 12 |
| 405 | 169 | 2 | 15 | 21 | 138 | 130 | 139 | 12 |

Analizado la filogenia global, se puede apreciar que en el clado 3 se agrupan cepas pertenecientes al ST64 y una cepa representativa del ST2836. Siendo un SLV con respecto al ST64, adicionalmente no se ha determinado el grupo eBG pero nuestros resultados muestran una cercanía filogenética con el eBG 65. Este clado está compuesto por aislados de países de América, África y Asia. Como se puede observar el ST64 es el imperante a nivel mundial (Figura 21).

En el clado 1 se puede apreciar una gran variedad de ST, los cuales no corresponden al serotipo Anatum. Es por ello que se agrupan en ramas independientes, compartiendo un ancestro común más probable con la cepa origen del serotipo Anatum.

C. Predicción de genes en aislados de *Salmonella*

Con el objetivo de realizar la caracterización de los genomas locales de *Salmonella*, se realizó el proceso de anotación. El cual consistió en interpretar la información contenida en cada secuencia. Para ello fue importante realizar una predicción de los genes mediante la búsqueda de marcos de lectura abiertos y posteriormente asignarles la información biológica a los genes previamente identificados (conocido como anotación estructural y funcional, respectivamente).

1. Anotación de *Salmonella enterica* Serotipo Anatum

En la Figura 22, se identifican los subsistemas que componen a una cepa local del serotipo Anatum, así como el número de genes pertenecientes a cada uno de estos. Los resultados muestran que el cromosoma de esta cepa se compone de 4'786,616 pb, que incluye 4,556 secuencias codificantes, 108 ARNts y un contenido de G+C del 52.2%. Resultados que concuerdan con los reportados por Marasini y col. en el 2016, quienes reportaron el genoma completo de dos cepas de *Salmonella entérica* serotipo Anatum procedente de aves de corral.

Haciendo una descripción general se encontraron 693 genes relacionados al metabolismo de los carbohidratos, estando específicamente 150 genes involucrados en el metabolismo central de los carbohidratos. Adicionalmente, se encontraron 107 involucrados en mecanismos de virulencia; teniendo participación en la adherencia de la bacteria, la invasión intracelular y la resistencia de antibióticos. También se determinaron 45 genes en la categoría de elementos móviles; perteneciendo todos a la presencia de fagos. Así mismo la cepa presenta 174 genes responsables de la respuesta al estrés; lo que le permite a la bacteria vivir en condiciones adversas, como las medioambientales.

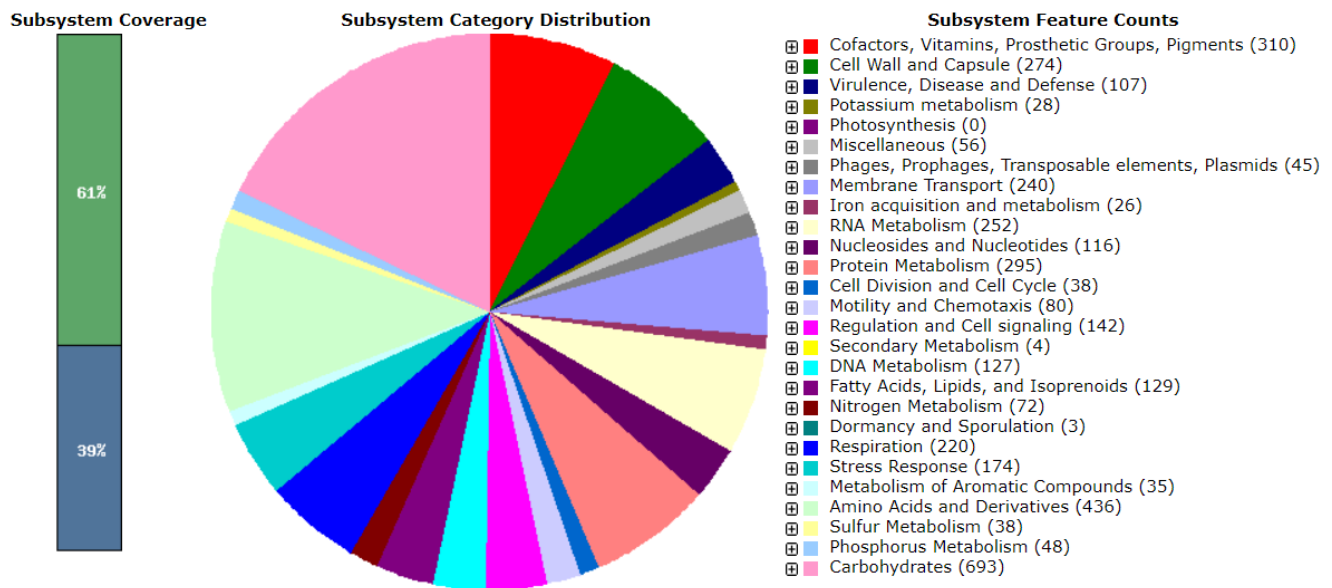


Figura 22. Genoma anotado de la cepa *Salmonella enterica* serotipo Anatum. La columna derecha indica las categorías de los subsistemas, mientras que el número entre paréntesis indica el número de genes pertenecientes a dicha categoría. Gráfica obtenida mediante el servidor RAST.

2. Anotación de *Salmonella enterica* Serotipo Give

Se identificó un tamaño de genoma de 4'903,604 pb con un contenido G+C del 52.1%, 95 ARNt y 5,095 secuencias codificantes. Adicionalmente, en la Figura 23, se muestran los diferentes subsistemas que lo componen, así como el número de genes pertenecientes a dicha categoría. Como se puede observar, se identificaron 717 genes responsables del metabolismo de carbohidratos, 104 genes en la categoría de virulencia y resistencia a antibióticos, 35 genes responsables de mecanismos responsables de transferencia horizontal (siendo el total de genes para fagos) y 171 correspondientes a la respuesta contra el estrés.

3. Anotación de *Salmonella enterica* Serotipo Saintpaul

Se identificó un genoma aproximado de 4'697,656 pb con 4,593 secuencias codificantes, 100 ARNt y un contenido G+C de 52.2%. Datos que coinciden con los ya mencionados en la anotación de los serotipos Oranienburg y Anatum.

En la Figura 24, se muestran los diferentes subsistemas que se identificaron. Encontrándose 722 genes responsables del metabolismo de los carbohidratos. Además, 105 genes responsables de la virulencia y resistencia de antibióticos. También 44 genes en el subsistema de elementos móviles, siendo todos genes de fagos.

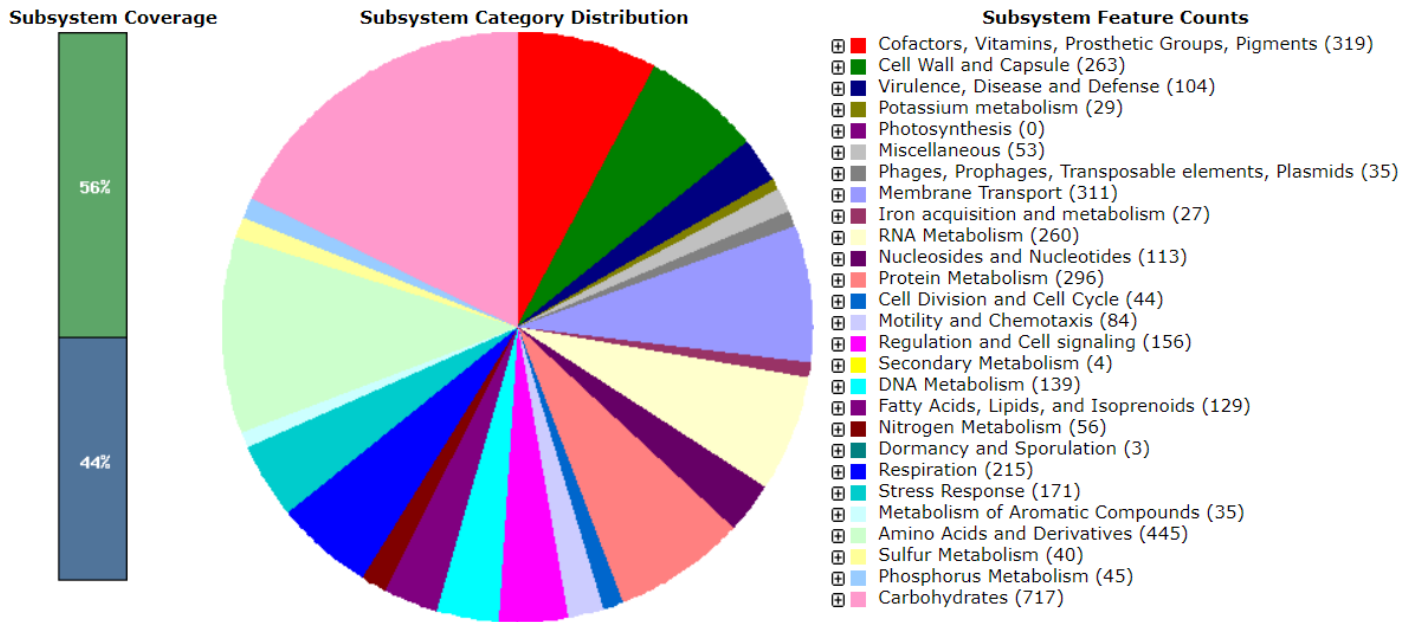


Figura 23. Genoma anotado de la cepa *Salmonella enterica* serotipo Give. La columna derecha indica las categorías de los subsistemas, mientras que el número entre paréntesis indica el número de genes pertenecientes a dicha categoría. Gráfica obtenida mediante el servidor RAST.

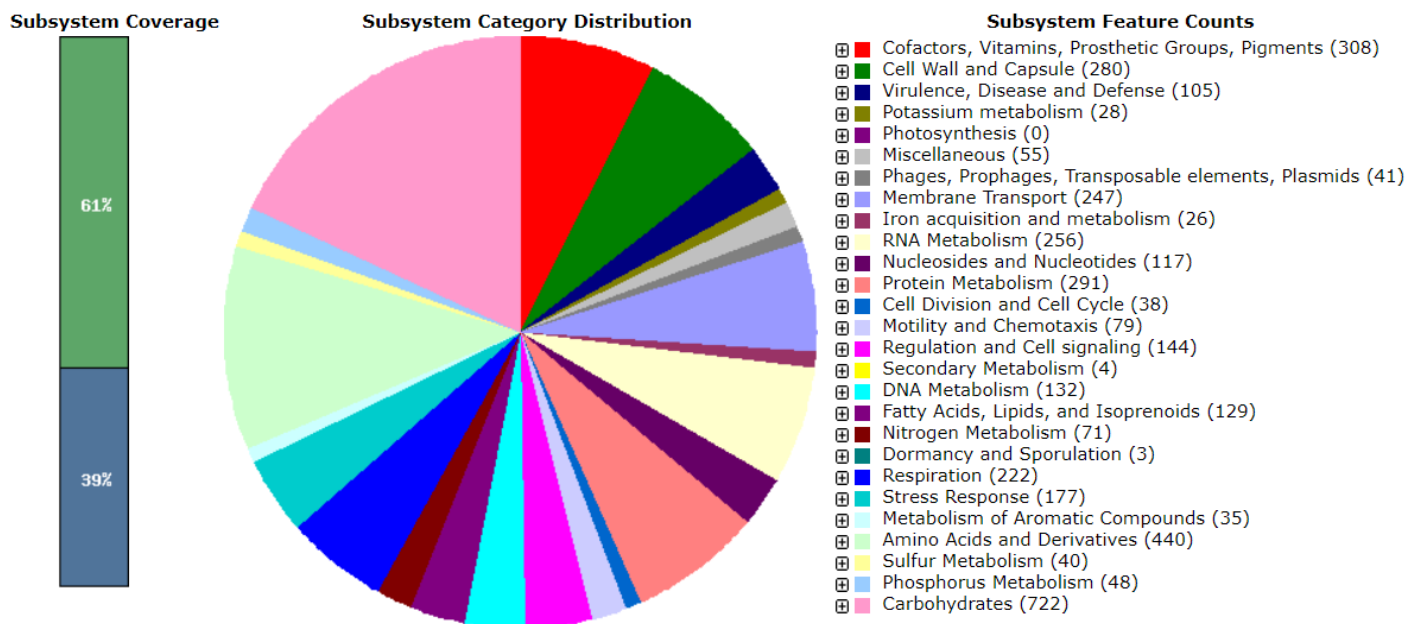


Figura 24. Genoma anotado de la cepa *Salmonella enterica* serotipo Saintpaul. La columna derecha indica las categorías de los subsistemas, mientras que el número entre paréntesis indica el número de genes pertenecientes a dicha categoría. Gráfica obtenida mediante el servidor RA

4. Anotación de *Salmonella enterica* Serotipo Minnesota

Los resultados mostraron un genoma de 4'809,127 pb con 4,934 secuencias codificantes, 115 ARNt y un contenido de G+C del 52.1%. Resultados que coinciden con los encontrados por Wang y col. 2017 quienes analizaron una cepa de *Salmonella* Minnesota aislada de una fruta de mango fresco procedente de México. Encontrando un genoma de 4'716,739 pb con un número de secuencias codificantes de 4,747.

En la Figura 25, se puede apreciar los diferentes subsistemas que componen el genoma de esta cepa. Se identificaron 688 genes responsables del metabolismo de los carbohidratos, 104 genes responsables de la virulencia y resistencia a antibióticos. Adicionalmente, se identificaron 39 genes en el subsistema de elementos móviles; perteneciendo todos a fagos.

5. Anotación de *Salmonella enterica* Serotipo Oranienburg

Los resultados muestran un cromosoma con un tamaño de 4'823,013 pb, teniendo 4,572 secuencias codificantes, 105 ARNt y un contenido G+C del 52.1% (Figura 26). Existen trabajos donde se ha reportado borradores de genomas para este serotipo; tal es el caso del realizado por Medrano-Félix y col. 2013 quienes trabajaron con una cepa ambiental de la región de Culiacán, encontrado un genoma de 4'609,551 pb con 5,020 secuencias codificantes, 103 ARNt y un contenido G+C del 52.2%.

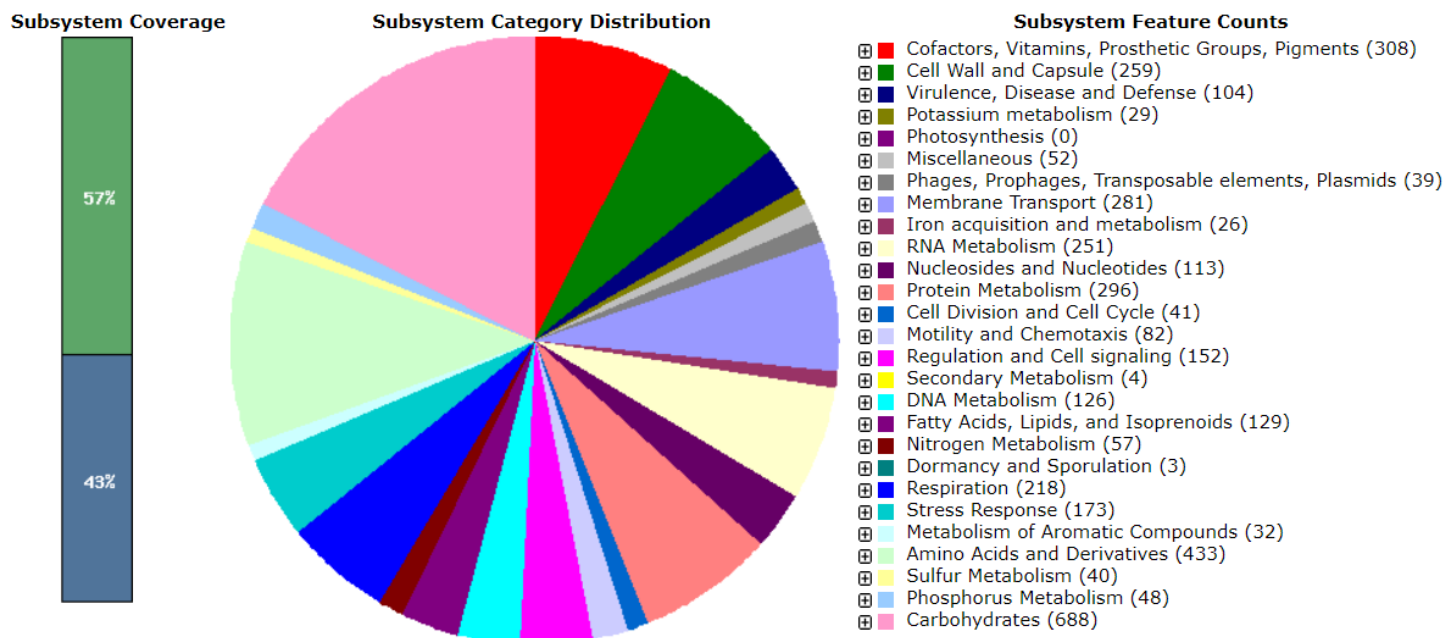
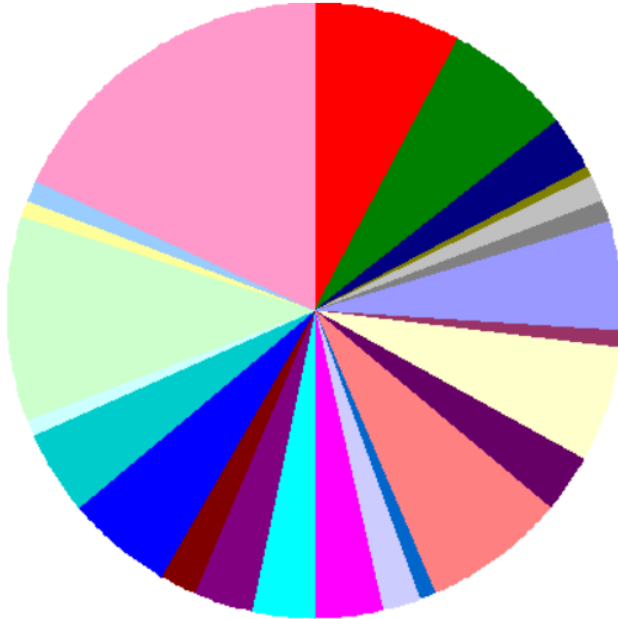


Figura 25. Genoma anotado de la cepa *Salmonella enterica* serotipo Minnesota. La columna derecha indica las categorías de los subsistemas, mientras que el número entre paréntesis indica el número de genes pertenecientes a dicha categoría. Gráfica obtenida mediante el servidor RAST.

Subsystem Coverage



Subsystem Category Distribution



Subsystem Feature Counts

- ☒ Cofactors, Vitamins, Prosthetic Groups, Pigments (310)
- ☒ Cell Wall and Capsule (260)
- ☒ Virulence, Disease and Defense (109)
- ☒ Potassium metabolism (29)
- ☒ Photosynthesis (0)
- ☒ Miscellaneous (53)
- ☒ Phages, Prophages, Transposable elements, Plasmids (38)
- ☒ Membrane Transport (237)
- ☒ Iron acquisition and metabolism (25)
- ☒ RNA Metabolism (249)
- ☒ Nucleosides and Nucleotides (113)
- ☒ Protein Metabolism (294)
- ☒ Cell Division and Cell Cycle (38)
- ☒ Motility and Chemotaxis (78)
- ☒ Regulation and Cell signaling (137)
- ☒ Secondary Metabolism (4)
- ☒ DNA Metabolism (127)
- ☒ Fatty Acids, Lipids, and Isoprenoids (129)
- ☒ Nitrogen Metabolism (71)
- ☒ Dormancy and Sporulation (3)
- ☒ Respiration (217)
- ☒ Stress Response (174)
- ☒ Metabolism of Aromatic Compounds (33)
- ☒ Amino Acids and Derivatives (426)
- ☒ Sulfur Metabolism (39)
- ☒ Phosphorus Metabolism (46)
- ☒ Carbohydrates (692)

Figura 26. Genoma anotado de la cepa *Salmonella enterica* serotipo Oranienburg. La columna derecha indica las categorías de los subsistemas, mientras que el número entre paréntesis indica el número de genes pertenecientes a dicha categoría. Gráfica obtenida mediante el servidor RAST.

Siendo muy similares a los datos encontrados en nuestro estudio. Cabe señalar que la cobertura promedio utilizada por dichos investigadores fue del 35.8; mientras que la utilizada para este estudio fue del 98, debido a que utilizamos una secuenciación en ambos sentidos mediante la plataforma Illumina.

Adicionalmente, se encontraron 692 genes relacionados con el metabolismo de los carbohidratos, estando 142 genes involucrados en el metabolismo central de los carbohidratos. También se identificaron 109 genes referentes a la virulencia y resistencia de antibióticos; así mismo se obtuvieron 38 genes para elementos móviles, siendo todos para fagos.

Como se puede observar en los resultados previamente mencionados, los genomas de *Salmonella enterica* de los serotipos predominantes en la región tienen un tamaño de genoma muy similar a los reportados por diversos autores (Medrano-Félix y col, 2013; Wang y col, 2017). Además de poseer un número de secuencias codificantes similar, al igual que el número de ARNt.

Nuestros resultados permiten observar que la cantidad de genes que se incluyen en cada subsistema para las cepas locales son muy parecidas; por ejemplo los genes referentes al metabolismo de carbohidratos, los de factores de virulencia y resistencia a antibióticos, los responsables de las respuestas al estrés y los de elementos móviles (Cuadro 15). Lo que lleva a suponer que las cepas que se encuentran en la localidad son muy parecidas entre ellas y por lo tanto comparten una gran cantidad de genes. Lo que sustenta los resultados previos obtenidos, donde se determinaron grupos ST específicos para las cepas locales.

Cuadro 15. Resumen de los datos obtenidos mediante la anotación de genomas

| Subsistemas | Anatum | Give | Saintpaul | Minnesota | Oranienburg |
|-------------------------------------|---------------|-------------|------------------|------------------|--------------------|
| Tamaño de genoma (pb) | 4'786,616 | 4'903,640 | 4'697,656 | 4'909,127 | 4'823,013 |
| Secuencias codificantes | 4,556 | 5,095 | 4,593 | 4,934 | 4,572 |
| G+C% | 52.2 | 52.1 | 52.2 | 52.1 | 52.1 |
| Virulencia | 107 | 104 | 105 | 104 | 107 |
| Fagos | 45 | 35 | 41 | 39 | 45 |
| Respuesta al estrés | 174 | 171 | 177 | 173 | 174 |
| Metabolismo de carbohidratos | 693 | 717 | 722 | 688 | 692 |

pertenecientes a los serotipos predominantes en la región.

D. Caracterización de los factores de virulencia

Los resultados obtenidos por el proceso de anotación mostraron que los genomas locales poseen genes de factores de virulencia que le permiten a *Salmonella* invadir al hospedero y desencadenar un proceso infeccioso. Por lo que resulta de relevancia e interés hacer una caracterización de las islas de patogenicidad y posteriormente de los genes involucrados. Este trabajo es el primero que reporta la virulencia de los aislados de Culiacán, Sinaloa pertenecientes a los serotipos Anatum, Give, Saintpaul y Minnesota.

Para la identificar las islas de patogenicidad se utilizó el servidor VirulenceFinder 1.5 del Centro Genómico de Epidemiología de Denmark, como se describe en Materiales y Métodos. Los resultados mostraron que todos los serotipos predominantes en la región presentan al menos las primeras tres islas de patogenicidad denominadas como SPI-1, SPI-2 y SPI-3, respectivamente (Figura 27). Siendo la SPI-1, responsable de codificar las proteínas necesarias para el ensamblado del complejo T3SS; permitiéndole a la bacteria iniciar el proceso de invasión mediante la secreción de proteínas efectoras.

La SPI-2 codifica un segundo T3SS como parte esencial del crecimiento intracelular, además es necesaria para llevar a cabo la parte sistemática de la infección. SPI-3 se encarga de secuestrar Mg^{2+} como parte de la adaptación cuando hay limitaciones nutricionales (Mendoza y col. 2008). Se ha reportado que

las islas SPI-1 y SPI-2 son requeridas por *Salmonella* serotipo Typhimurium para desencadenar enfermedades como gastroenteritis (Lawley y col. 2007).

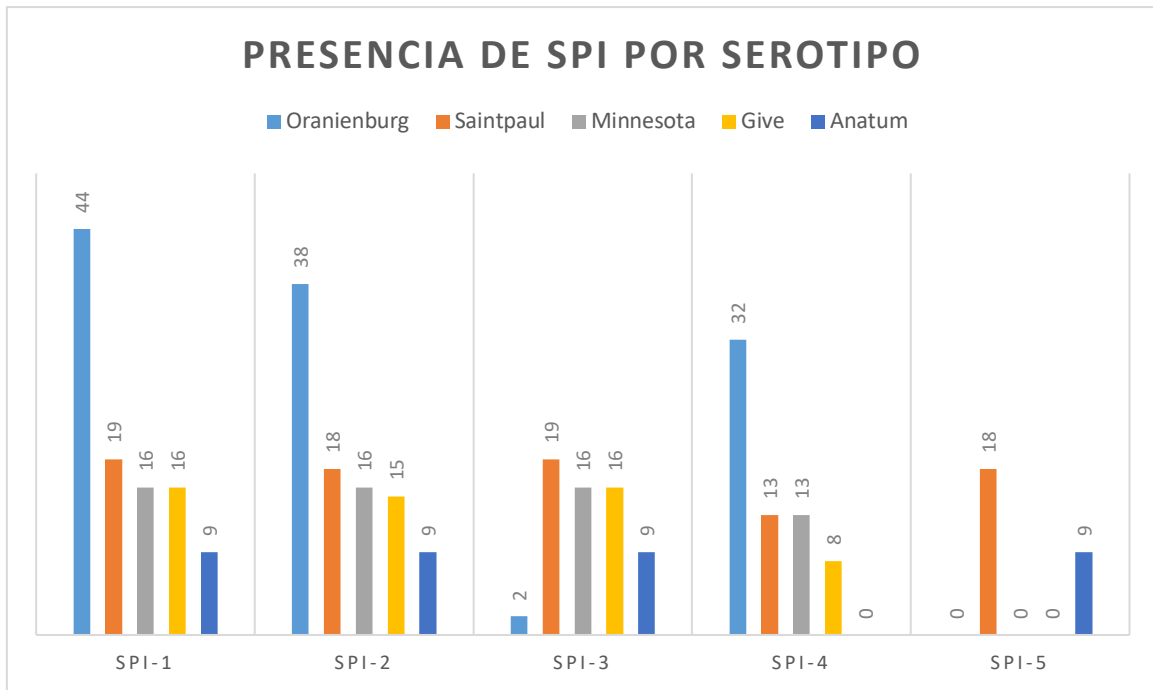


Figura 27. Frecuencia de las islas de patogenicidad de *Salmonella* en los distintos serotipos. Se representa cada serotipo con un color diferente. En el eje de las X se indican las islas de patogenicidad; mientras que en el eje de las Y se especifica el número de serotipos.

SPI-4 y SPI-5 no se han caracterizadas completamente. Sin embargo, estudios han demostrado que la SPI-4 consiste en 6 genes designados *siiA-F*, los cuales intervienen en el tipo de secreción I (Morgan y col. 2004). Siendo importantes para colonizar el intestino de bovinos. Se ha demostrado que la ausencia de esta isla de patogenicidad en *S. Typhimurium* y *S. Enteritidis* atenúa la virulencia en ratones (Kiss y col. 2007); por lo que la presencia de esta isla pudiera tener un mayor impacto en la patogénesis en animales bovinos. Lo que supone un riesgo, para tanto en animales como en humanos (por el consumo de productos infectados), ya que se encontró *Salmonella* Saintpaul (que contiene SPI-4) en vacas de la localidad.

SPI-5 fue identificada por primera vez en un genoma de *Salmonella* Dublin y consiste en 5 genes que juegan un papel importante en la patogenicidad mediante la secreción de proteínas efectoras para SPI-1 y SPI-2 involucradas en la secreción de fluidos y reacción inflamatoria en la mucosa intestinal (Wood y col. 1998). Se ha reportado que esta isla contribuye a la colonización de *Salmonella* en el bazo de gallinas. La mutación de los genes de SPI-5, disminuyen la enteropatogénesis de *Salmonella* (Cao y col. 2014).

De acuerdo con nuestros resultados, no se identificó la presencia de la SPI-4 en el serotipo Anatum. Adicionalmente, se determinó que no se encuentra la SPI-5 en los serotipos Oranienburg, Minnesota y Give. Estos resultados se pueden deber a que se utilizó una identidad mayor al 95%, por lo que si los genes identificados

para estas islas presentaron un porcentaje menor de identidad, fueron retiradas del resultado.

Estos resultados sugieren que las cepas encontradas en agua de río y en animales de granja en la región de Culiacán, cuentan con islas de patogenicidad que les confieren el potencial para llevar a cabo la adherencia, invasión y desencadenar un proceso infeccioso en huéspedes tanto humanos como animales. Lo que puede suponer un riesgo a la población mediante la adquisición directa de enfermedades o por el consumo de productos contaminados.

Adicionalmente, se encontró que las islas de patogenicidad previamente descritas se encuentran flanqueadas por diferentes ARNt. Por ejemplo, la SPI-1 se encuentra flanqueada por el ARNt-pheV, SPI-2 por ARNt-val y la SPI-3 por ARNt-seIC. De acuerdo a Hensel (2004), el que las islas presenten este ordenamiento es un indicio muy fuerte de que fueron adquiridas por mecanismos de transferencia horizontal.

Para identificar los genes contenidos en las islas de patogenicidad, se realizó una alineación de los genomas de las cepas con respecto a los factores de virulencia contenidos en la base de datos del Instituto de Biología de Patógenos de Beijing, China. Donde se identificaron un promedio de 85 genes de virulencia para cada serotipo, en comparación con el proceso de anotación donde se identificaron un promedio de 105. Esta disparidad en número obtenido, se puede deber a los criterios experimentales utilizados; como lo es porcentaje de identidad.

En la Figura 28 se muestran los diferentes genes de virulencia que presentan las cepas locales, englobados en su respectiva categoría. Como se puede observar la mayoría de las cepas tienen genes de adherencia fimbrial tipo 1 y las pertenecientes al grupo Agf (Thin aggregative fimbriae). Siendo éstas importantes para la adherencia de *Salmonella* a las células de tejido; si no se presentarán estos genes, la bacteria no podría desencadenar su patogénesis y por lo tanto sería eliminada mediante secreciones mucosas.

Interesantemente, se determinó que sólo el serotipo Saintpaul presentó el operón *Lpf* (long polar fimbriae), siendo genes que tienen la función de adherirse a las placas de peyer en modelos murinos (Bäumler y col. 1996). Este operón fue reportado previamente solo en *Salmonella* Thyphimurium (Kingsley, 2002) y *S. Enteritidis* (Norris, 1996), pero en el alcance de nuestra revisión documental no hay reportes que sugieran la presencia de este operón en otros serotipos como Saintpaul.

Adicionalmente, se encontró que sólo los serotipos Saintpaul y Anatum presentan genes que codifican para proteínas de adherencia no fimbrial (proteínas que tienen afinidad por los azúcares y su principal función es la adherencia). Estos resultados sugieren que estos serotipos pudieran tener mayor capacidad de adherencia y por lo tanto, mayores probabilidades de desencadenar una enfermedad.

También se encontró que la mayoría de los serotipos presentan el gen *mgtB* y *mgtC*, los cuales son requeridos para la proliferación intracelular en macrófagos

(donde existe una baja concentración de Mg^{2+}) y para la virulencia en ratones (Lee y Lee, 2015).

Con respecto a la SPI-1, se encontró que la mayoría de los genes que la componen se encuentran presentes en los 5 serotipos. Ejemplo de ello, son los genes *inv* que regulan el mecanismo de aguja; siendo característico en el mecanismo de secreción III que presenta esta bacteria. Solamente se identificó el gen *sipD* para los serotipos Anatum y Saintpaul. Dicho gen es requerido para la translocación de proteínas efectoras, siendo esenciales para la internalización de bacterias (Desein y col. 2010). Lo que sugiere que estos serotipos podrían eficientizar el proceso de invasión celular.

Para la SPI-2 se encontró al igual que la SPI-1, que la mayoría de los genes se encuentran presentes en los diferentes serotipos. Sin embargo, solo se identificaron los genes *ssrA* y *ssrB* para solo un aislado de animal perteneciente al serotipo Saintpaul. Estos genes han sido reportados por Choi y col. en el 2010 como dos componentes esenciales que regulan la supervivencia y replicación de *Salmonella* en las células hospederas. Además promueven la transcripción de múltiples genes de la SPI-2.

Se identificaron diversos genes como *SipA*, *SopE*, *SopE2* y *SopB* que codifican para proteínas efectoras que pertenecen a las distintas islas de patogenicidad como la SPI-1, SPI-2 y SPI-5. Teniendo actividades diversas como el rearrreglo del citoesqueleto, secreción fluida e inflamación y apoptosis de macrófagos (Figueroa y Verdugo, 2005).

FACTORES DE VIRULENCIA

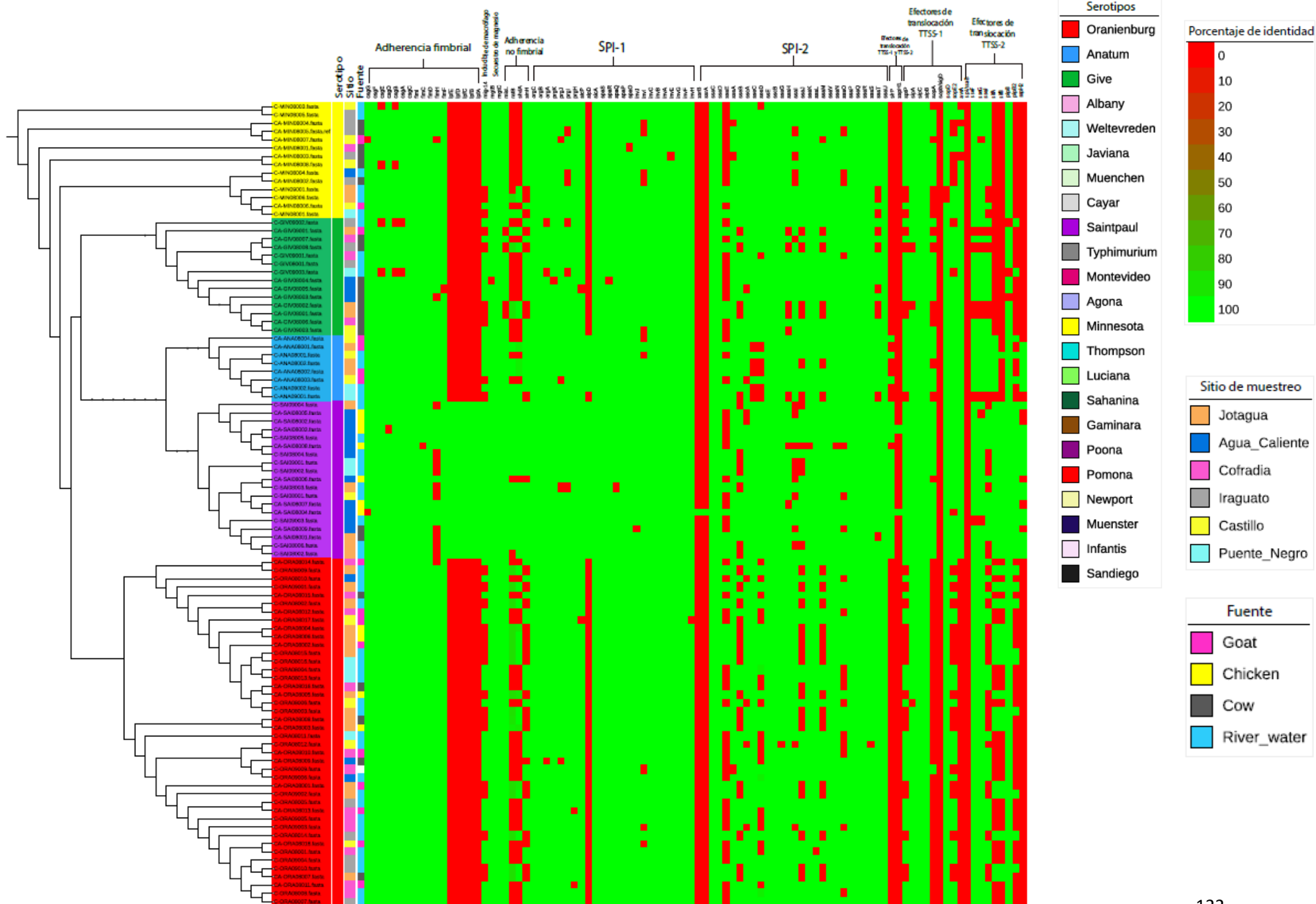


Figura 28. Heatmap de presencia y ausencia de genes de virulencia, denotando con verde la presencia del gen. En la parte superior se muestra las categorías a las cuales pertenecen los genes. En la parte derecha se muestran las acotaciones.

E. Pan-genoma

Para elucidar las diferencias y similitudes para cada serotipo y determinar la cantidad de genes únicos, se realizó la construcción del pan-genoma; el cual indica el número de genes totales compartidos por la especie. Agrupándolos de manera general en genoma accesorio (genes que no se encuentran compartidos en todas las cepas) y genoma core (genes compartidos en el 100% de los genomas).

A continuación se presentan las gráficas de los análisis realizados por Roary para cada uno de los cinco serotipos dominantes en la región. Estas gráficas representan la distribución de los genes conforme se agregan genomas en los análisis, lo que permite diferenciar con gran detalle las características fenotípicas y el aporte en el número de genes de cada cepa al pan-genoma (Jacobsen y col. 2011).

1. Pan-genoma de las cepas de *Salmonella enterica* Serovar Minnesota

De acuerdo a la Figura 29, se puede apreciar que la línea que representa los nuevos genes se hace asintótica a partir del genoma 9. Lo que indica que en ese punto ya se identificó el grupo de genes pertenecientes al core; y por lo tanto la adición de genomas nuevos al análisis incrementará el tamaño del core. Esto se comprueba mediante la adición de los genomas 10, 11 y 12, los cuales no generan un cambio en la identificación de genes nuevos.

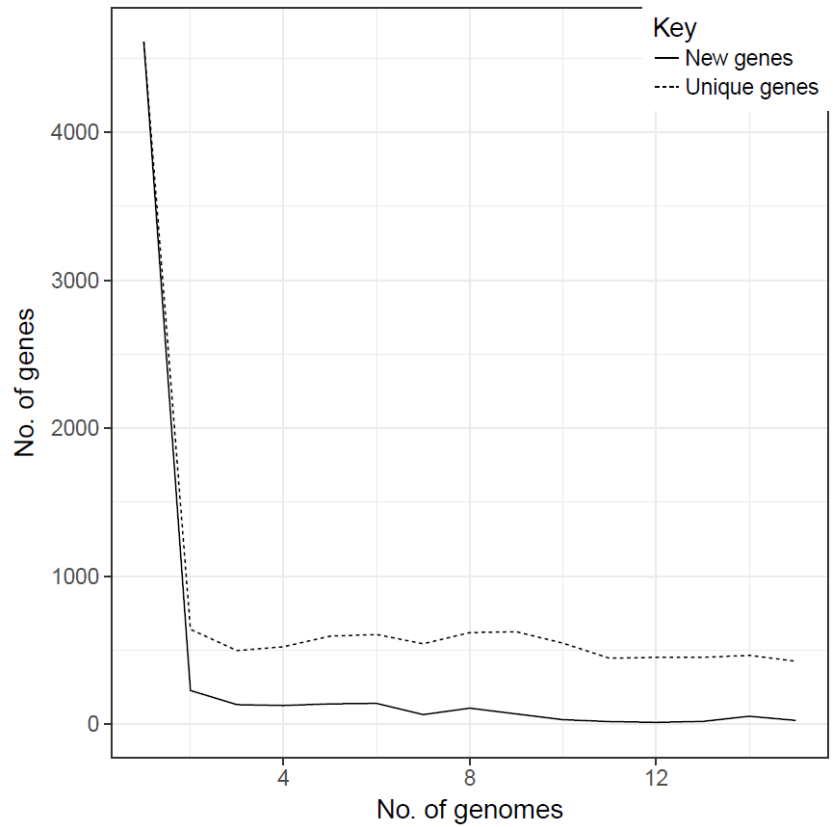


Figura 29. Proporción de genes nuevos y únicos para las cepas locales de *Salmonella enterica* serotipo Minnesota. En el eje de las X aparece el número de genomas analizados, mientras que en el eje Y aparece el número de genes. Se denota con línea punteada la proporción de genes únicos y con línea sólida la proporción de genes nuevos conforme se van agregando nuevos genomas al análisis.

Las cepas locales pertenecientes al serovar Minnesota presentan un tentativo pan-genoma cerrado (Figura 30), es decir conforme nuevas cepas son secuenciadas y agregadas al análisis, estas no aportarán cambios significativos al pan-genoma de la especie. Este serotipo se constituye por un pan-genoma de aproximadamente 5,767 genes; mientras que el genoma core está integrado por 3,989 genes (Cuadro 16).

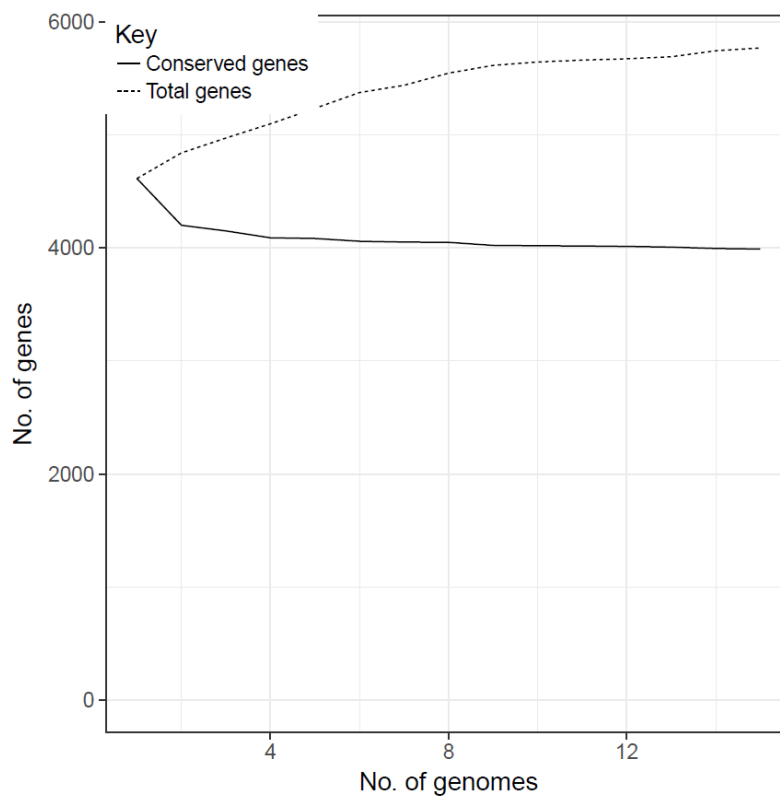


Figura 30. Gráfica del pan-genoma y del genoma core de las cepas locales de *Salmonella enterica* serotipo Minnesota. En el eje X se muestra el número de genomas analizados y en Y el número de genes. La línea punteada indica la cantidad de genes totales, mientras que la línea sólida indica la cantidad de genes conservados.

Cuadro 16. Resumen de la composición del pan-genoma de cepas locales de *Salmonella* para el serotipo Minnesota.

| <i>Tipo de genes</i> | Presencia en cepas | Número de genes |
|-----------------------------|---------------------------|------------------------|
| <i>Genes core</i> | 99-100% | 3,989 |
| <i>Genes soft core</i> | 95-99% | 0 |
| <i>Genes Shell</i> | 15-95% | 854 |
| <i>Genes cloud</i> | 0-15% | 924 |
| <i>Total de genes</i> | 0-100% | 5,767 |

2. Pan-genoma de las cepas de *Salmonella enterica* Serovar Saintpaul

En la Figura 31, se puede apreciar una línea asintótica a partir del genoma 7, lo que sugiere que si se agregan nuevos genomas locales, lo más probable es que no se presente una variación en cuanto al número de genes nuevos.

La Figura 32, sugiere que el serotipo Saintpaul presenta un genoma cerrado. Esto se debe que a partir de la adición el genoma 7 en adelante, no hay un cambio significativo en el conjunto de genes core. Estos resultados sirven como evidencia de que la constitución genética de las cepas de *Salmonella* aisladas de la localidad presentan poca variación genética entre ellas.

Las cepas locales del serotipo Saintpaul presentan un pan-genoma constituido por 4,784 genes. De los cuales 4,213 se encuentran compartidos por todas las cepas y 357 figuran como únicos para dichas cepas (Cuadro 17).

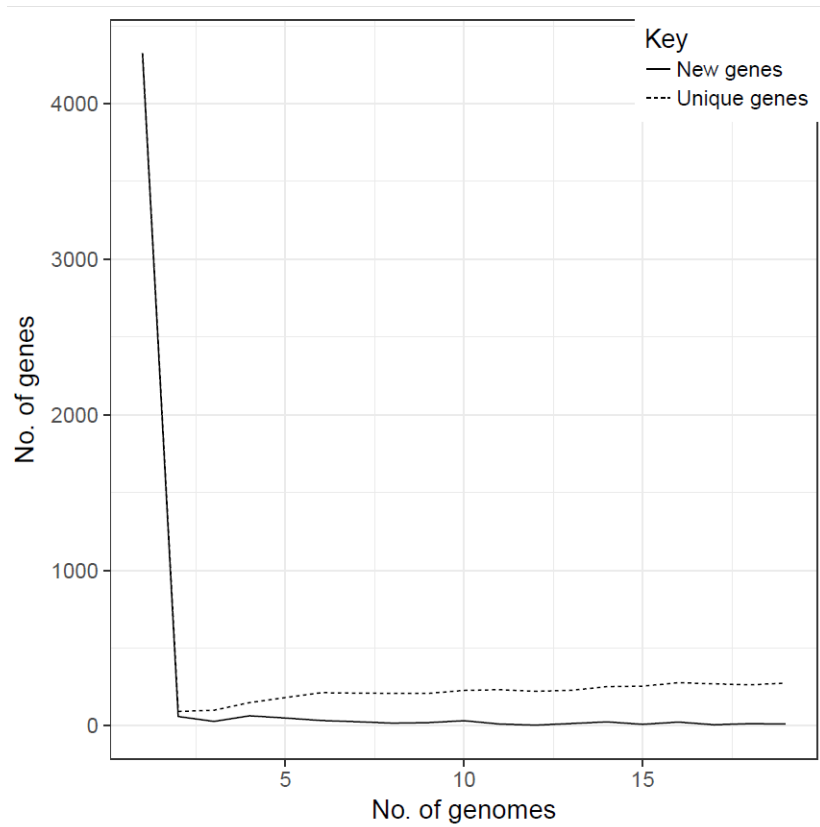


Figura 31. Proporción de genes nuevos y únicos para las cepas locales de *Salmonella enterica* serotipo Saintpaul. En el eje de las X aparece el número de genomas analizados, mientras que en el eje Y aparece el número de genes. Se denota con línea punteada la proporción de genes únicos y con línea sólida la proporción de genes nuevos conforme se van agregando nuevos genomas al análisis.

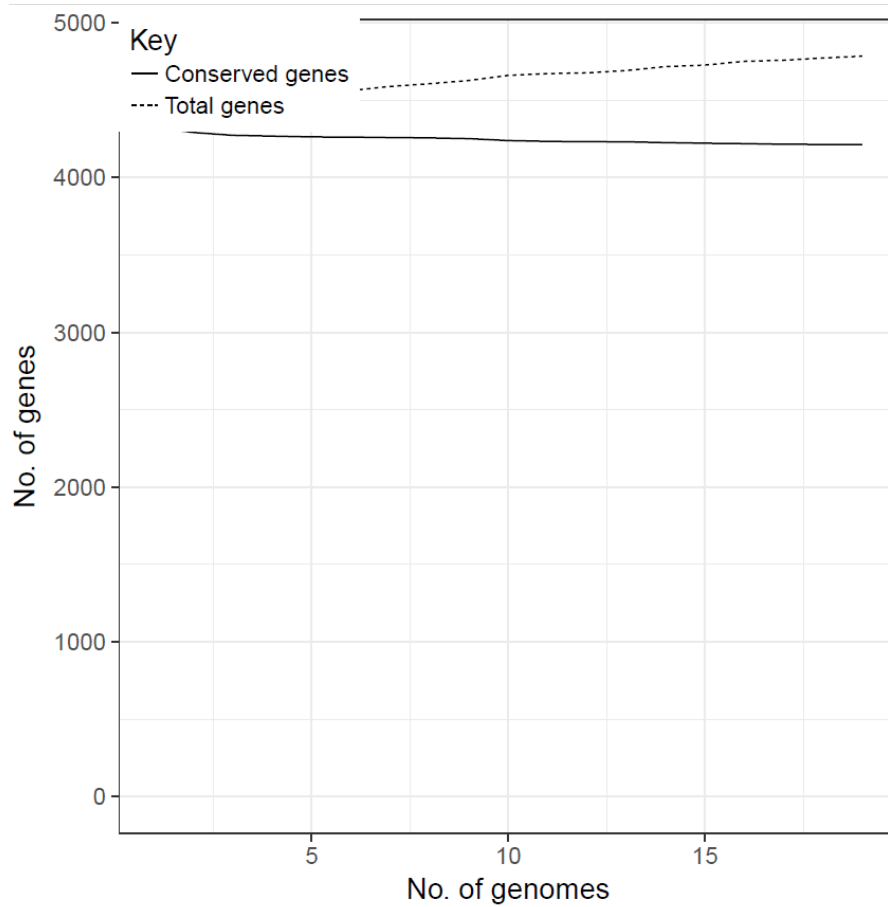


Figura 32. Gráfica del pan-genoma y del genoma core de las cepas locales de *Salmonella enterica* serotipo Saintpaul. En el eje X se muestra el número de genomas analizados y en Y el número de genes. La línea punteada indica la cantidad de genes totales, mientras que la línea sólida indica la cantidad de genes conservados.

Cuadro 17. Resumen de la composición del pan-genoma de cepas locales de *Salmonella* serotipo Saintpaul.

| <i>Tipo de genes</i> | Presencia en cepas | Número de genes |
|-----------------------------|---------------------------|------------------------|
| <i>Genes core</i> | 99-100% | 4,213 |
| <i>Genes soft core</i> | 95-99% | 0 |
| <i>Genes Shell</i> | 15-95% | 214 |
| <i>Genes clound</i> | 0-15% | 357 |
| <i>Total de genes</i> | 0-100% | 4,784 |

3. Pan-genoma de las cepas de *Salmonella enterica* Serovar Anatum

De acuerdo a la Figura 33, se puede observar que no existen grandes variaciones en la presencia de genes nuevos conforme se contemplan nuevos genomas. Esto se puede deber a que muy pocas cepas de las muestreadas fueron pertenecientes al serotipo Anatum y por lo tanto se contemplaron muy pocas en el análisis. Sería conveniente agregar nuevas cepas para determinar si realmente el serotipo Anatum presenta un genoma cerrado; y por lo tanto no se contabilizaran genes nuevos si continuamos agregando genomas.

Se puede apreciar que el serotipo Anatum presenta un pan-genoma aparente cerrado (Figura 34); pero como se mencionó anteriormente, sería conveniente agregar nuevas cepas locales del mismo serotipo para ver cuál es la dinámica del pan-genoma y determinar si se sigue tratando de uno cerrado o abierto, en su defecto.

El pan-genoma del serotipo Anatum se encuentra constituido por 4486 genes, de los cuales 4,260 pertenecen al genoma core. Se determinaron 226 genes que se encuentran en el rango del 15 a 95% de los genomas analizados (Cuadro 18); sin embargo no se encontraron genes accesorios, lo cual se puede deber al número de genomas analizados.

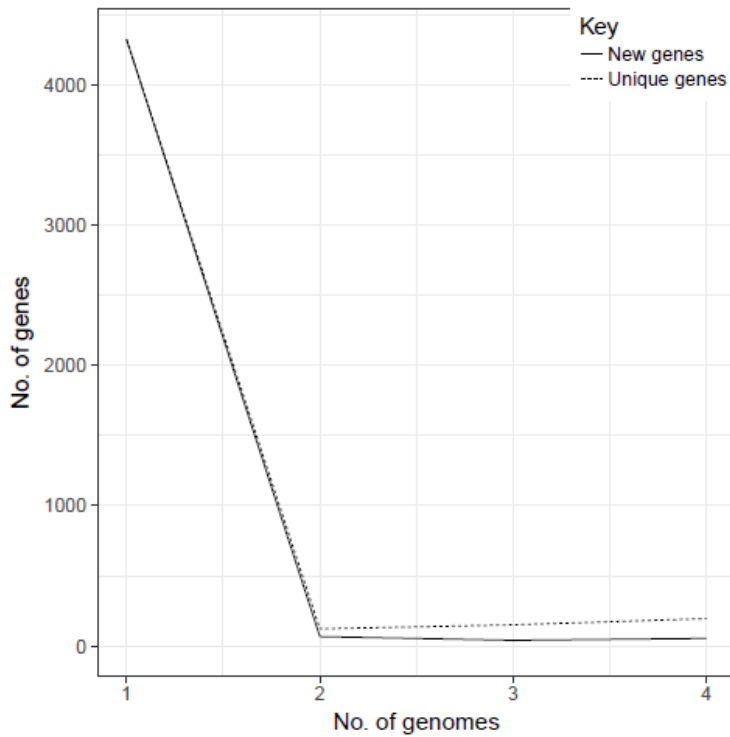


Figura 33. Proporción de genes nuevos y únicos para las cepas locales de *Salmonella enterica* serotipo Anatum. En el eje de las X aparece el número de genomas analizados, mientras que en el eje Y aparece el número de genes. Se denota con línea punteada la proporción de genes únicos y con línea sólida la proporción de genes nuevos conforme se van agregando nuevos genomas al análisis.

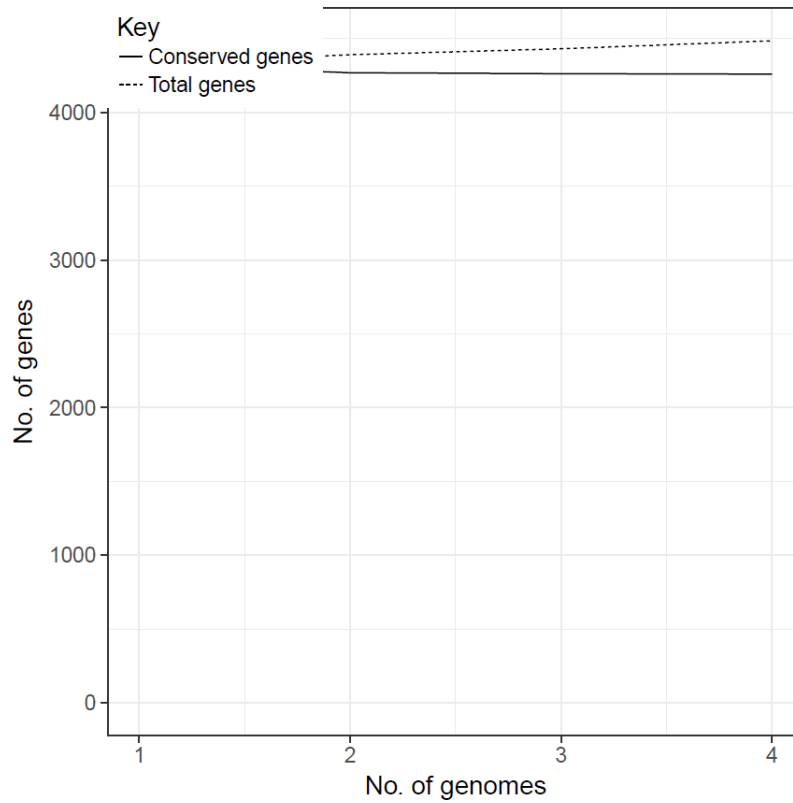


Figura 34. Gráfica del pan-genoma y del genoma core de las cepas locales de *Salmonella enterica* serovar Anatum. En el eje X se muestra el número de genomas analizados y en Y el número de genes. La línea punteada indica la cantidad de genes totales, mientras que la línea sólida indica la cantidad de genes conservados.

Cuadro 18. Resumen de la composición del pan-genoma de cepas locales de *Salmonella* para el serotipo Anatum.

| <i>Tipo de genes</i> | Presencia en cepas | Número de genes |
|-----------------------------|---------------------------|------------------------|
| <i>Genes core</i> | 99-100% | 4,260 |
| <i>Genes soft core</i> | 95-99% | 0 |
| <i>Genes Shell</i> | 15-95% | 226 |
| <i>Genes cloud</i> | 0-15% | 0 |
| <i>Total de genes</i> | 0-100% | 4,486 |

4. Pan-genoma de las cepas de *Salmonella enterica* Serovar Give

En la Figura 35, se puede observar una serie de picos en diferentes momentos, siendo resultado de que al analizar un nuevo genoma se incrementa la cantidad de genes nuevos; lo que indica la diversidad genética entre estas cepas. Esta idea se ve reforzada si observamos la línea punteada, la cual muestra que conforme aumenta la cantidad de genomas, incrementa la cantidad de genes accesorios.

De acuerdo a la Figura 36, se puede apreciar como al aumentar el número de genomas analizados aumenta el número de genes totales y disminuye el número de genes core. Lo que indica que las cepas de *Salmonella enterica* serovar Give presentan genomas dinámicos, permitiéndoles reordenar su material genético mediante la eliminación o la adquisición de nuevos genes. Esto también se puede demostrar mediante el árbol filogenético global (Figura 15) donde se aprecia que el serotipo Give se agrupa en dos grandes clados distintos, poniendo de manifiesto su diversidad genética.

El pan-genoma de este serotipo se compone de 7,489 genes. De los cuales 3,751 forman parte del core y 2,562 genes del accesorio (Cuadro 19). Como se puede

apreciar, la cantidad de genes accesorio es muy alta; lo que refuerza la idea de la diversidad genética para este serotipo.

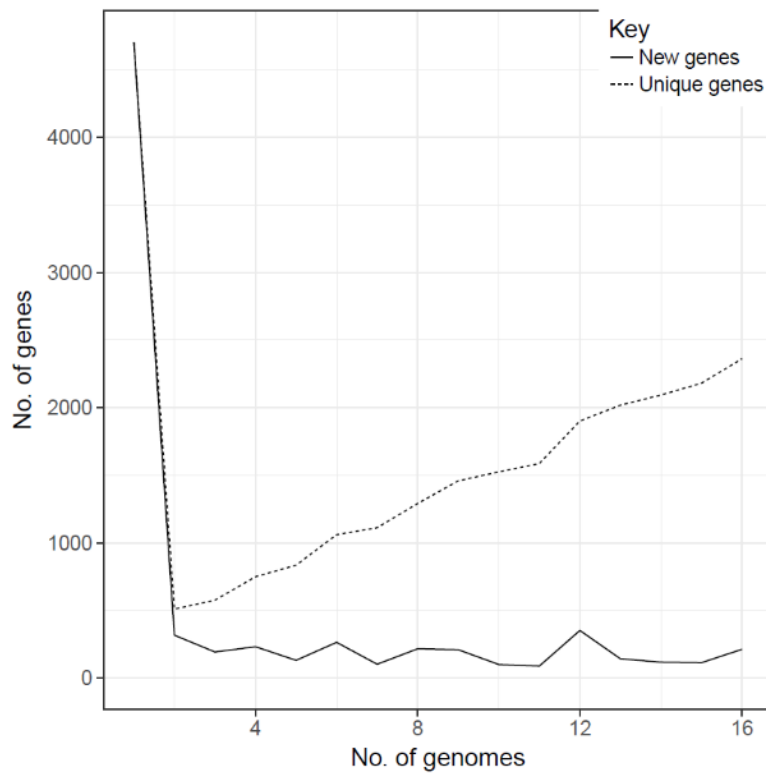


Figura 35. Proporción de genes nuevos y únicos para las cepas locales de *Salmonella enterica* serotipo Give. En el eje de las X aparece el número de genomas analizados, mientras que en el eje Y aparece el número de genes. Se denota con línea punteada la proporción de genes únicos y con línea sólida la

proporción de genes nuevos conforme se van agregando nuevos genomas al análisis.

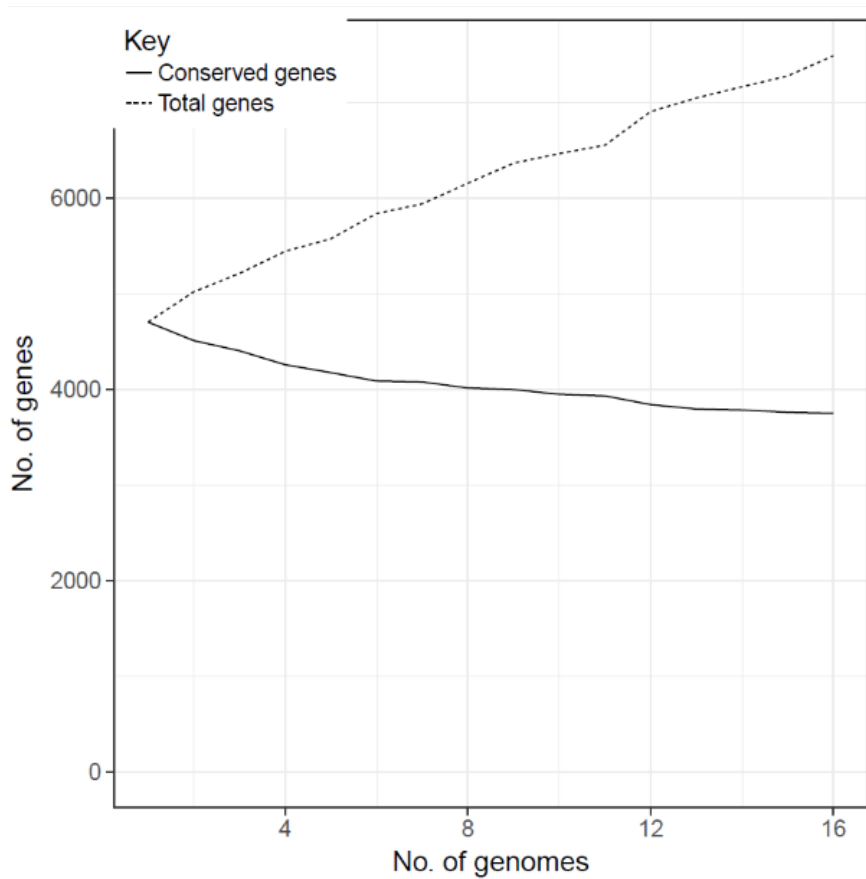


Figura 36. Gráfica del pan-genoma y del genoma core de las cepas locales de *Salmonella enterica* serotipo Give. En el eje X se muestra el número de genomas

analizados y en Y el número de genes. La línea punteada indica la cantidad de genes totales, mientras que la línea sólida indica la cantidad de genes conservados.

Cuadro 19. Resumen de la composición del pan-genoma de cepas locales de *Salmonella* para el serotipo Give.

| <i>Tipo de genes</i> | Presencia en cepas | Número de genes |
|-----------------------------|---------------------------|------------------------|
| <i>Genes core</i> | 99-100% | 3,751 |
| <i>Genes soft core</i> | 95-99% | 0 |
| <i>Genes Shell</i> | 15-95% | 1,176 |
| <i>Genes cloud</i> | 0-15% | 2,562 |
| <i>Total de genes</i> | 0-100% | 7,489 |

5. Pan-genoma de las cepas de *Salmonella enterica* Serovar Oranienburg

Al igual que el serotipo Give, las cepas de Oranienburg presentan diversidad genética (Figura 37). Esto se debe a que con cada genoma analizado se identifican genes nuevos, es por ello que en la parte inferior de la gráfica se muestra una línea sólida no asintótica.

De acuerdo a la Figura 38, se puede observar como aumenta pronunciadamente la cantidad de genes totales mientras aumenta el número de genomas analizados; presentado un pan-genoma abierto. El cual es típico de especies que viven en ambientes donde existen mezclas microbianas, donde se puede propiciar el intercambio de material genético por mecanismos de transferencia horizontal.

El pan-genoma de este serotipo está constituido por la cantidad de 8,375 genes. De los cuales 3,331 genes pertenecen al genoma core y 3,684 al genoma accesorio (Cuadro 20). Como se observa, el número de genes accesorios es mayor a la cantidad de genes core.

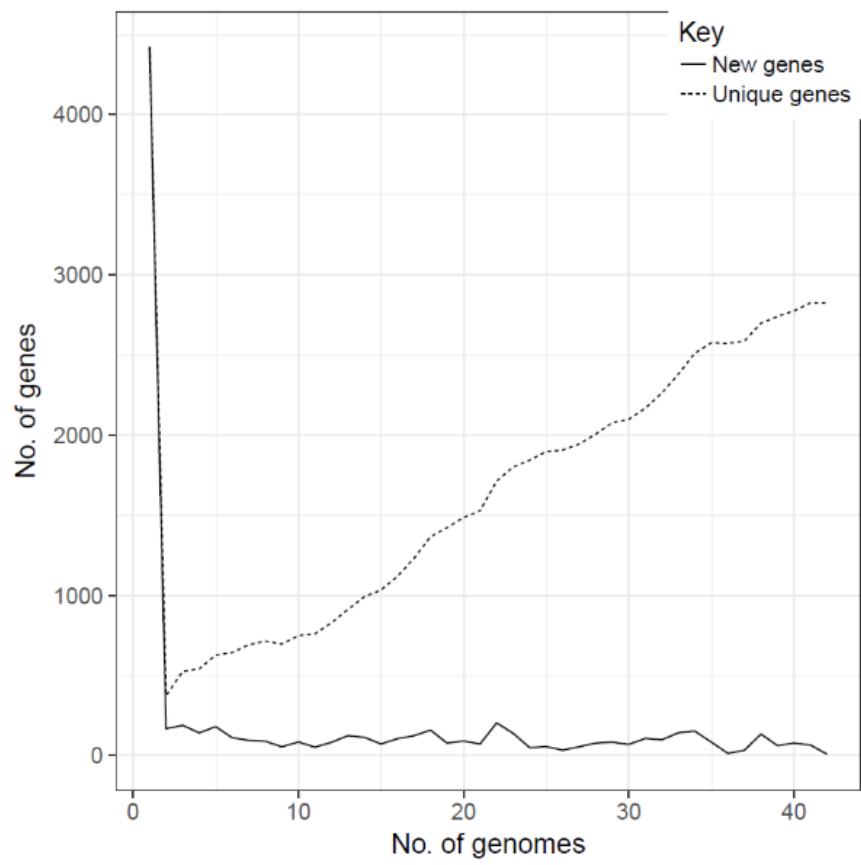


Figura 37. Proporción de genes nuevos y únicos para las cepas locales de *Salmonella enterica* serotipo Oranienburg. En el eje de las X aparece el número de genomas analizados, mientras que en el eje Y aparece el número de genes. Se denota con línea punteada la proporción de genes únicos y con línea sólida la proporción de genes nuevos conforme se van agregando nuevos genomas al análisis.

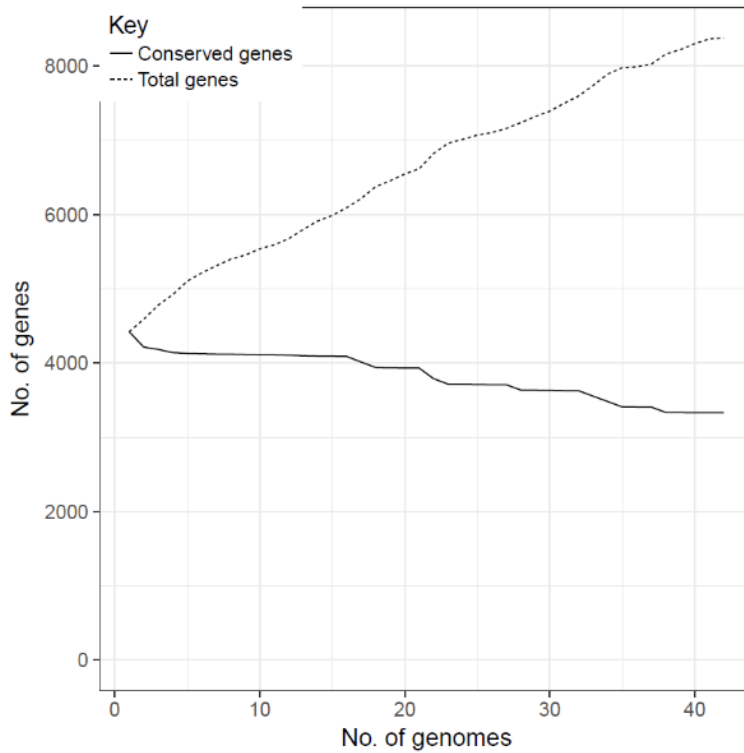


Figura 38. Gráfica del pan-genoma y del genoma core de las cepas locales de *Salmonella enterica* serotipo Oranienburg. En el eje X se muestra el número de

genomas analizados y en Y el número de genes. La línea punteada indica la cantidad de genes totales, mientras que la línea sólida indica la cantidad de genes conservados.

Cuadro 20. Resumen de la composición del pan-genoma de cepas locales de *Salmonella* para el serotipo Oranienburg.

| <i>Tipo de genes</i> | Presencia en cepas | Número de genes |
|-----------------------------|---------------------------|------------------------|
| <i>Genes core</i> | 99-100% | 3,331 |
| <i>Genes soft core</i> | 95-99% | 781 |
| <i>Genes Shell</i> | 15-95% | 579 |
| <i>Genes cloud</i> | 0-15% | 3,684 |
| <i>Total de genes</i> | 0-100% | 8,375 |

Estos resultados sugieren que serotipos como Oranienburg y GIVE presenten alta diversidad genética; ya que poseen pan-genomas abiertos. Permitiéndoles intercambiar material genético con el medio exterior, por lo que se esperaría que en el análisis detallado de esta transferencia se encontraran genes adquiridos mediante fagos, plásmidos o transposones. Además, el alto número de genes accesorios encontrados en estas cepas, sugieren que pueden presentar una cantidad mayor de factores de virulencia y/o resistencia a antibióticos.

Para poder determinar el tipo de pan-genoma que presenta el serotipo Anatum, es necesario incorporar un número mayor de cepas locales; sin embargo los resultados hasta el momento sugieren que se trata de un pan-genoma cerrado.

Estos resultados coinciden en parte con los obtenidos por Jacobsen y col. en el 2011. Quienes determinaron un pan-genoma cerrado para 35 genomas de diferentes serotipos pertenecientes a *Salmonella* subsp. *enterica*. Dicho análisis solo contempló al serotipo Saintpaul, el cual coincide con nuestros resultados, ya que también se determinó que poseía un pan-genoma cerrado.

En el presente trabajo se obtuvieron una cantidad mayor de genes core que los reportados por Jacobsen en el 2012, esto se debe a que en el análisis del pan-genoma en el presente trabajo se hizo por serotipos; por lo que se esperaría que las cepas sean más parecidas entre si y por lo tanto se obtenga un número mayor de genes pertenecientes al core. Sin embargo, también reporta que dos genomas pertenecientes al mismo serotipo pueden ser muy diferentes; tal es el caso de algunas de las cepas para los serotipos Oranienburg y Give.

En general nuestros resultados sugieren que la especie *Salmonella enterica* posee un genoma cerrado con excepción de muy pocos serotipos (Gordienko y col. 2013). Adicionalmente, se han reportado que este tipo de resultados pueden estar en función del estilo de vida y lugar donde se desarrolla el microorganismo (Tettelin y col. 2008).

F. Aporte de los fagos a la virulencia de *Salmonella*

Los resultados obtenidos en el proceso de anotación y en la identificación de islas de patogenicidad, sugieren la presencia de mecanismos de transferencia horizontal para el aporte de genes de virulencia a las cepas locales. Es por ello que se analizó el aporte de los fagos a la virulencia de *Salmonella*. Para lograr este propósito se utilizó la herramienta PHASTER, que permite la identificación y anotación de secuencias profago en genomas bacterianos (Arndt y col. 2016).

Se encontraron 953 fagos en los 153 genomas locales de *Salmonella*; de los cuales solo 92 fagos presentaron al menos un gen de virulencia. Lo que representa un aporte del 9.6% por parte de los fagos a la virulencia de los aislados locales. En el Cuadro 21, se muestra el porcentaje de los fagos presentes en los

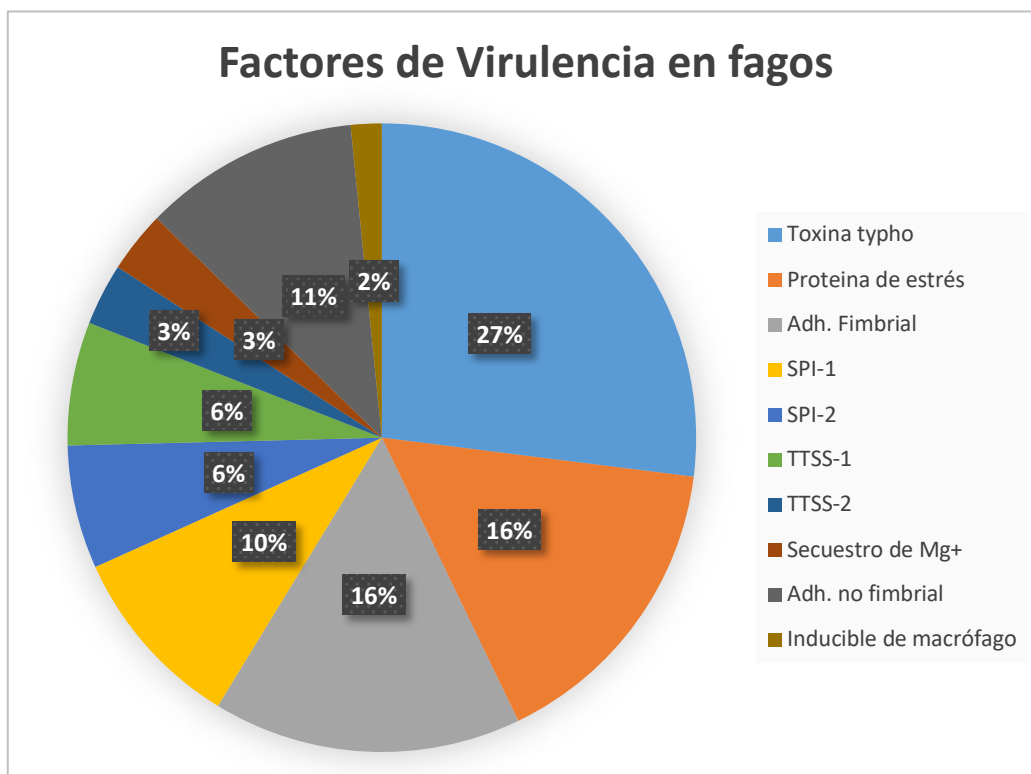
distintos serotipos. Como se puede observar, se encontró una mayor proporción de fagos en los serotipos imperantes en la región.

En lo que respecta a la caracterización de los 5 serotipos dominantes en la región, se encontraron 50 secuencias de fago que contenían al menos un gen de virulencia. El siguiente aspecto consistió en investigar el tipo de genes que aportaban estos fagos. En la Figura 39, se muestra una gráfica de pastel en la que se puede apreciar los factores de virulencia por categoría que poseen estos 50 fagos.

Cuadro 21. Porcentaje de los fagos encontrados en los aislados de *Salmonella*.

| Serotipo | % de fagos presentes |
|--------------------|-----------------------------|
| Oranienburg | 24.03% |
| Give | 18.05% |
| Minnesota | 13.75% |
| Saintpaul | 8.71% |
| Pomona | 8.39% |
| Poona | 7.97% |
| Weltevreden | 6.82% |
| Muenster | 4.62% |
| Anatum | 4.09% |
| Montevideo | 2.10% |
| Infantis | 1.47% |
| TOTAL | 100% |

Figura 39. Factores de virulencia aportados por los 50 fagos en los serotipos dominantes en la región.



Nuestros resultados sugieren que el gen *cdtB* es el principal aporte de los fagos a la virulencia de *Salmonella*. Este gen se ha identificado en *Salmonella enterica* serotipo Typhi y codifica una proteína que sirve como toxina. Información que apoya una vez más a lo establecido previamente, lo que sugiere que estas cepas representan un riesgo a la salud pública.

Interesantemente, se identificó que otro de los principales factores que aportan los fagos a la virulencia es el gen *sodC*. El cual codifica para la superóxido dismutasa, siendo la mayor defensa antioxidante contra radicales O_2^- lo que posibilita a patógenos a vivir en condiciones de estrés y juega un papel importante en la patogénesis (Fukai y Ushio-Fukai, 2018).

Dentro de la categoría de los genes de adhesión fimbrial aportadas por los fagos, se identificaron aquellos que codifican para las proteínas de adhesión no fimbrial (*lpfC*, *lpfD*, *lpfB*, *lpfE* y *lpfA*) que se caracterizaron solamente para el serotipo Saintpaul (Figura 28); lo que sugiere que este serotipo adquirió estos factores de

virulencia por mecanismos de transferencia horizontal mediados por fagos (transducción).

Adicionalmente se encontró que los fagos también aportan genes pertenecientes a la SPI-1 (10%) como *invJ* y *invA*, siendo proteínas importantes en la invasión de células del tejido epitelial. También se encontraron genes de la SPI-2 (6%) como *ssaO* y *ssaP*, siendo reguladores importantes para el desarrollo del mecanismo de aguja.

En la Figura 40, se muestra una gráfica de pastel con los fagos encontrados en los serotipos dominantes. Como se puede apreciar el fago Erwinia fue el que se encontró en mayor proporción, aportando un gen que codifica para una toxina. La cual se ha identificado en *Salmonella enterica* Typhi.

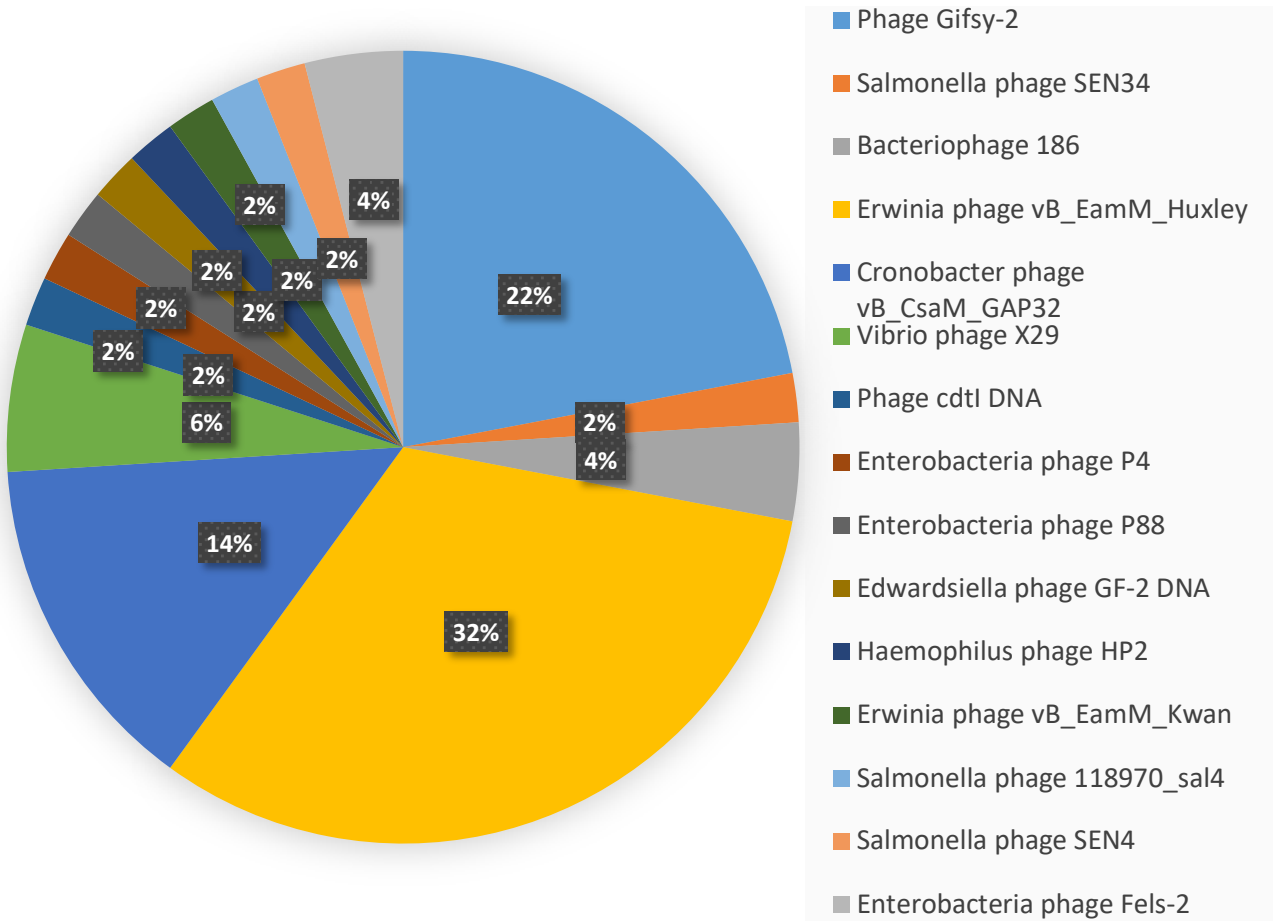
Diversos autores han reportado que el fago Gifsy-2 infecta principalmente a *Salmonella* Typhimurium contribuyendo a su virulencia; además se ha determinado que posee el gen *sodC* responsable de la resistencia al estrés oxidativo (Ho y Slauch, 2001), información que concuerda con los resultados obtenidos. De manera interesante se determinó que solo este fago se encontraba en aislados de río; lo que podría sugerir que las cepas de río presentan condiciones de estrés, por lo que necesitan de este gen.

Para el caso del fago Cronobacter GAP32, ha sido caracterizado en agua procedente de plantas de tratamiento. Este fago pertenece a la familia *Myoviridae* y se ha identificado que posee genes de adherencia fimbrial (Abbasifar y col. 2014).

Este es el primer trabajo de investigación donde se reportan la caracterización genómica a detalle del perfil de virulencia de los serotipos predominantes de la región; así como el aporte de los fagos a esa virulencia. Nuestros resultados sugieren un aporte importante por parte de fagos en cuestión de genes de virulencia necesarios en las cepas aisladas de la región, para sobrevivir a condiciones de estrés y efectuar el mecanismo de patogénesis.

Figura 40. Gráfica de pastel con los fagos encontrados en los serotipos dominantes en la región.

Fagos encontrados en las cepas locales



IX. CONCLUSIONES

- Las relaciones filogenéticas y el porcentaje de identidad ANI sugieren poca variabilidad genética en las cepas de *Salmonella* aisladas de la región, debido a la cercanía geográfica y al espacio temporal del muestreo.
- La construcción de árboles filogenéticos específicos para cada serotipo muestran que el origen global de Oranienburg esta compartido con países como Estados Unidos, Reino Unido, Ecuador y Corea del Sur; mientras que los serotipos Saintpaul y Minnesota lo comparten con Estados Unidos. Por su parte, el serotipo Give lo comparte con Brasil, India y Estados Unidos. Por último el serotipo Anatum tiene su origen compartido con Estados Unidos y China.
- Las cepas de *Salmonella* Oranienburg y Give presentan pan-genomas tentativamente abiertos; lo que indica que son más propensas a sufrir modificaciones genéticas.
- La identificación de SPI-1, SPI-2 y SPI-3 dentro de los genomas analizados sugiere que las cepas ambientales de *Salmonella* representa un riesgo potencial para la salud pública.
- El aporte de los fagos a la virulencia a los aislados de *Salmonella enterica* es del 9.6% confiriendo genes pertenecientes a SPI-1, SPI-2, de adherencia, secuestro de Mg⁺ y antioxidantes.
- El uso de herramientas bioinformáticas para la serotipificación, construcción de árboles filogenéticos, anotación y marcadores genéticos de virulencia, permite la caracterización rápida y eficiente de genomas

bacterianos, constituyendo una estrategia confiable para la discriminación de microorganismos, con aplicación en ecología y epidemiología.

X. BIBLIOGRAFÍA

- Achtman, M., Wain, H. (2012). Multilocus Sequence Typing as a Replacement for Serotyping in *Salmonella enterica*. *PLOS PATHOENS*.
- Aguilar-Bultet, L., & Falquet, L. (2015). Secuenciación y ensamble de novo de genomas bacterianos: una alternativa para el estudio de nuevos patógenos. *Revista de Salud Animal*, 125-132.
- Amavisit, P., Lightfoot, D., Browning, G. F., & Markham, P. F. (2003). Variation between Pathogenic Serovars within *Salmonella* Pathogenicity Islands. *Journal of Bacteriology*, 3624-3635.
- Andino, A., & Hanning, I. (2015). *Salmonella enteric*: Survival, Colonization, and Virulence Differences among Serovars. *The Scientific World Journal*.
- Arahal, D. R. (2014). Whole-genome analysis: average nucleotide identity. *Methods in Microbiology* , 103-122.
- Ayala-Labarríos, L. A., Rodríguez-Herrera, R., Aguilar-González, C. N., Lara-Victoriano, F., & Quero-Carrillo, A. R. (2004). Detección de *Clavibacter michiganensis* subs *nebraskensis* usando la reacción en cadena de la polimerasa. *Revista de fitopatología mexicana*, 239-245.
- Baker, C. (2012). Next-Generation Sequencing vs. Microarrays. *Genetic Engineering & Biotechnology News*.

- Baker, M. (2012). *De novo* genome assembly: what every biologist should know. *Nature Methods*, 333-337.
- Baker, S., & Dougan, G. (2007). The Genome of *Salmonella enterica* Serovar Typhi. *Clinical Infectious Diseases*, 29-33.
- Bäumler, A. J. (1997). The record of horizontal gene transfer in *Salmonella*. *Trends in Microbiology*, 318-322.
- Bäumler, A. T. (1996). The lpf fimbrial operon mediates adhesion of *Salmonella* Typhimurium to murine Peyer's patches. *Proceedings of the National Academy of Sciences of the United States of America*, 279-283.
- Behjati, S., & Tarpey, P. S. (2013). What is next generation sequencing? *Archives of Disease in Childhood Education and Practice Edition*, 236-238.
- BIORAD. (2002). Bacterial Serotyping Guide for *Salmonella*. *Food Science*, 2-8.
- Blanc-Potard, A.-B., Solomon, F., Kayser, J., & Groisman, E. A. (1999). The SPI-3 Pathogenicity Island of *Salmonella enterica*. *Journal of Bacteriology*, 998-1004.
- Blocker, A., Jouihri, N., Larquet, E., Gounon, P., & Ebel, F. (2001). Structure and composition of the *Shigella flexneri* "needle complex", a part of its type III secretion. *Molecular Microbiology*, 652-663.
- Bradnam, K., Fass, J., Alexandrov, A., Baranay, P., Bechner, M., Birol, I., & Boisvert, S. (2013). Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. *GigaScience*, 2-10.

- Brenner, F., Villar, R., Angúlo, F., Tauxe, R., & Swamithan, B. (2000). *Salmonella* Nomenclature. *Journal of Clinical Microbiology*.
- Burton, G. A., Gunnison, D., & Lanza, G. R. (1987). Survival of pathogenic bacteria in various freshwater sediments. *Applied and Environmental Microbiology*, 633-638.
- Cabañas-Romero, P. V., & Huerta-Saquero, A. (2014). Nanomáquinas biológicas: los sistemas de secreción bacterianos. *Mundo Nano*, 28-36.
- Cañedo, R., & Arencibia, J. (2004). Bioinformática: en búsqueda de los secretos moleculares de la vida. *Revista de Salud Animal*.
- Cao, G., Allard, M., Strain, E., Stones, R., Zhao, S., Brown, E., & Meng, J. (2014). Genetic diversity of *Salmonella* pathogenicity islands SPI-5 and SPI-6 in *Salmonella* Newport. *Foodborne pathogens and disease*, 798-807.
- CDC. (2011). Surveillance System Overview: National *Salmonella* Surveillance . *Enteric Diseases Epidemiology Branch*, 1-12.
- Cercenado, E., & Cantón, R. (2005). Métodos moleculares de la tipificación epidemiológica en bacteriología. *Procedimientos en microbiología clínica*.
- Chen, L. H., Zheng, D. D., Liu, B., Yang, J., & Jin, Q. (2015). VFDB 2016: hierarchical and refined dataset for big data analysis-10-years on. *Nucleic Acids Research*, 694-697.

- Chen, L., Zhaohui, X., Sun, L., Ynag, J., & Jin, Q. (2012). VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial factors. *Nucleic Acids Research*, 641-645.
- Chistoserdova, L. (2010). Recent progress and new challenges in matagenomics for biotechnology. *Biotechnology Letters*, 1351-1359.
- Choi, J. S. (2010). *Salmonella* pathogenicity island 2 expression negatively controlled by EllANtr-SsrB interaction is required for *Salmonella* virulence. *Proceedings of the National Academy of Sciences of the United States of America*, 20506-20511.
- Commission, E. (28 de Enero de 2016). *European Civil Protection And Humanitarian Aid Operations*. Obtenido de <http://ec.europa.eu/echo/files/evaluation/watsan2005>
- Craig, E., & James, S. (2006). The Genus *Salmonella*. *The Prokaryotes*, 123-158.
- Dasgupta, S., Papadimitriou, C. H., & Vazirani, U. V. (2006). *Algorithms*.
- Desein, T. M. (2010). Protection of epithelial cells from *Salmonella enterica* serovar Enteritidis invasion by antibodies against the SPI-1 type III secretion system. *Canadian Journal of Microbiology*., 522-526.
- Duncan, M., Linington, R., & Auerbuch, V. (2012). Chemical Inhibitors of the Type Three Secretion System: Disarming Bacterial Pathogens. *Antimicrobial Agents and Chemotherapy*, 5433-5441.

- Dutta, C., & Pan, A. (2002). Horizontal gene transfer and bacterial diversity. *Journal of biosciences*, 27-33.
- ECOEVOL. (11 de 09 de 2017). *Ecology and Evolution Unit Page*. Obtenido de www.ecoevo.unit.oist.jp/lab/
- EFSA. (2010). *Annual Report*. European Food Safety Authority.
- Ekblom, R., & Wolf, J. B. (2014). A field guide to whole-genome sequencing, assembly and annotation. *Evolutionary Applications*, 1026-1042.
- ELIKA. (28 de Febrero de 2013). *Fundación Vasca para la Seguridad Agroalimentaria*. Obtenido de http://www.elika.eus/datos/pdfs_agrupados/Documento82/1.Salmonella.pdf
- Figueira, R., & Holden, D. W. (2012). Functions of the *Salmonella* pathogenicity island 2 (SPI-2) type III secretion system effectors. *Microbiology*, 1147-1161.
- Figuerola, I. M., & Verdugo, A. (2005). Mecanismos moleculares de patogenicidad de *Salmonella* spp. *Revista Latinoamericana de Microbiología*, 25-42.
- Fish, J. T., & Pettibone, G. W. (1995). Influence of freshwater sediment on the survival of *Escherichia coli* and *Salmonella* spp. as measured by three methods of enumeration. *Applied and Environmental Microbiology*, 277-281.
- Franco, L., Cediñ, F., & Payán, C. (2008). Breve Historia de la Bioinformática. *Colombia Médica*, 117-120.

- Fukai, T., & Ushio-Fukai, M. (2018). Superoxide dismutase: role in redox signaling, vascular function and diseases. *Antioxidants & Redox Signaling*, 1583-1606.
- Galán, J. E. (2001). *Salmonella* interactions with Host Cells: Type III secretion at Work. *Annual Review of Cell and Developmental Biology*, 53-86.
- Galán, J. E., Lara-Tejero, M., Marlovits, T. C., & Wagner, S. (2014). Bacterial Type III Secretion Systems: Specialized Nanomachines for Protein Delivery into Target Cells. *Annual Review of Microbiology*, 415-438.
- Garcia, R., Baelum, J., Fredslund, L., Santorum, P., & Jacobsen, C. S. (2010). Influence of temperature and predation on survival of *Salmonella enterica* serovar Typhimurium. *Applied and Environmental Microbiology*, 5025-5031.
- Giannella, R. A. (1996). *Salmonella*. Galveston, Texas: University of Texas Medical Branch at Galveston.
- Gibas, C., & Jambeck, P. (2001). *Developing Bioinformatics Computer Skills*. California: O'Reilly.
- Gillespie, I. A., O'Brien, S. J., Adak, G. K., Ward, L. R., & Smith, H. R. (2005). Foodborne general outbreaks of *Salmonella* Enteritidis phage type 4 infection, England and Wales, 1992-2002: where are the risks? *Epidemiological Infections*, 795-801.
- Gómez Bombarely, A. (2012). Genomas y rompecabezas: una visión sobre el ensamblaje de genomas. *Biosciences*, 151-156.

- Gómez Saravia, J. (2002). Salmonellosis. *Universidad Nacional de Colombia*, 1167-1175.
- Gonzales-Pedrajo, B., & Dreyfus, G. (2003). Sistemas de secreción de proteínas en las bacterias gram negativas: biogénesis flagelar y translocación de factores de virulencia. *Mensaje Bioquímico*, 45-63.
- González, J., Sanandres, N., Varela, Z., Aguirre, E. (2010). Aislamiento microbiológico de *Salmonella* spp. y herramientas moleculares para su detección. *Salud Uninorte*, 73-94.
- Groisman, E. A., & Ochman, H. (1993). Cognate gene clusters govern invasion of host epithelial cells by *Salmonella* Typhimurium and *Shigella flexneri*. *EMBO Journal*, 3779-3787.
- Han, N., Quiang, Y., & Zhang, W. (2016). ANItools web: a web tool for fast genome comparison within multiple bacterial strains. *Database: The Journal of Biological Databases and Curation*, 1-5.
- Hayford, A. B. (2015). Genetic and resistance phenotypic subtyping of *Salmonella* Saintpaul isolates from various food sources and humans: Phylogenetic concordance in combinatory analyses. *Infection, Genetics and Evolutions*, 92-107.
- Health, N. I. (23 de Agosto de 2016). Help Me Understand Genetics. *Genetics Home Reference*, 1-12.
- Hensel, M. (2000). *Salmonella* Pathogenicity Island 2. *Molecular Microbiology*, 1015-1023.

- Hensel, M. (2004). Evolution of pathogenicity islands of *Salmonella enterica*. *International Journal of Medical Microbiology* , 95-102.
- Huang, D. B., White, N. J., & DuPont, H. L. (2005). *Salmonella typhi* (Typhoid Fever) and *S. paratyphi* (Paratyphoid Fever). *Infectious Diseases*, 623-628.
- Ibarra, J. A., & Steele-Mortimer, O. (2009). *Salmonella* - the ultimate insider. *Salmonella* virulence factors that modulate intracellular survival. *Cellular Microbiology*, 1579-1586.
- Illumina. (2010). Technology Spotlight: Illumina Sequencing . *Illumina Inc*, 1-5.
- Sanger Institute. (17 de Julio de 2017). *Wellcome trust sanger institute*. Obtenido de <http://www.sanger.ac.uk/resources/downloads/bacteria/salmonella.html>
- Islam, M., Morgan, J., Doyle, M. P., Phatak, S. C., Millner, P., & Jiang, X. P. (2004). Fate of *Salmonella enterica* serovar Typhimurium on carrots and radishes grown in fields treated with contaminated manure compost or irrigation water. *Applied and Environmental Microbiology* , 2497-2502.
- Issenhuth-Jeanjean, S., Roggentin, P., Mikoleit, M., Guibourdenche, M., Pinna, d. E., Nair, S., . . . Francois-Xavier, W. (2014). Supplement 2008-2010 (no. 48) to the White-Kauffmann-Le Minor. *Research in Microbiology*, 1-5.
- Jacobsen, A. H. (2011). The *Salmonella enterica* pan-genome. *Microbial Ecology*.
- Jacobsen, C. S., & Bech, T. B. (2012). Soil survival of *Salmonella* and transfer to freshwater and fresh produce. *Food Research International* , 557-566.

- Jenkins, C. (2015). Whole-Genome Sequencing Data for Serotyping *Escherichia coli*—It's Time for a Change! *Journal of Clinical Microbiology*, 2402-2403.
- Jones, B. D. (1997). Host responses to pathogenic *Salmonella* infection. *Genes and Development*, 679-687.
- Kingsley, R. W. (2002). Population heterogeneity of *Salmonella enterica* serotype Typhimurium resulting from phase variation of the *lpf* operon in vitro and in vivo. *Journal of Bacteriology*, 2352-2359.
- Kiss, T., Morgan, E., & Nagy, G. (2007). Contribution of SPI-4 genes to the virulence of *Salmonella enterica*. *FEMS Microbiology Letters*, 153-159.
- Kiss, T., Morgan, E., Nagy, & Gábor. (2007). Contribution of SPI-4 genes to the virulence of *Salmonella enterica*. *Microbiol Letters*, 153-159.
- Kolyva, S., Waxin, H., & Popoff, M. (1992). The Vi antigen of *Salmonella typhi*: molecular analysis of the *viaB* locus. *Journal of General Microbiology*, 297-304.
- Lahiri, A., Lahiri, A., Iyer, N., Das, P., & Chakravorty, D. (2010). Visting the cell biology of *Salmonella* infection. *Microbes and Infection*, 809-818.
- Lappe, N. (2011). *Salmonella* Taxonomy. *Dept of Medical Microbiology*, 1.
- Lawley, T. B. (2007). Host transmission of *Salmonella enterica* serovar Typhimurium is controlled by virulence factors and indigenous intestinal microbiota. *Infection and Immunity*, 403-416.

- Lee, P. C., Zmina, S. E., Stopford, C. M., Toska, J., & Rietsch, A. (2014). Control type III secretion activity and substrate specificity by the cytoplasmic regulator PcrG. *Proceedings of the National Academy of Sciences*, 2027-2036.
- Lee, J.-W., Lee, E.-J., (2015). Regulation and function of the *Salmonella* MgtC virulence protein. *Journal of Microbiology*, 667-672.
- Liu, W. B., Liu, B., Zhu, XN., Yu, SJ., Shi, XM., (2011). Diversity of *Salmonella* isolates using serotyping and multilocus sequence typing. *Food Microbiology*, 1182-1189.
- Loströh, C. P., & Lee, C. A. (2001). The *Salmonella* pathogenicity island-1 type III secretion system. *Microbes and Infections*, 1281-1291.
- Luscombe, N. M., Greenbaum, D., & Gerstein, M. (2001). What is bioinformatics? An introduction and overview. *Yearbook of Medical Informatics*, 83-100.
- Marasini, D. A.-S. (2015). Complete Genome Sequences of *Salmonella enterica* Serovars Anatum and Anatum var. 15+, Isolated from Retail Ground Turkey. *American Society for Microbiology*, 1-2.
- Marchant, G. E., Bartlett-Esquilant, G., Dawes, M., & Issa, A. M. (2014). Personalized Medicine and Whole Genome Sequencing in the Era of Big Data: Challenges and Opportunities. *ISPOR CONNECTIONS*, 4-6.
- Marcus, S. L., Brumell, J. H., Pfeiter, C. G., & Finlay, B. (2000). *Salmonella* pathogenicity islands: big virulence in small packages. *Microbes and Infection*, 145-156.

Medicine, N. L. (24 de 07 de 2017). *National Center for Biotechnology Information*.

Obtenido de www.ncbi.nlm.nih.gov

Medrano-Félix, A. E.-A.-G.-F. (2013). Draft Genome Sequence of *Salmonella enterica* subsp. *enterica* Serotype Oranienburg Strain S-76, Isolated from an Aquatic Environment. *Genome Announcements*, 1-2.

Mendoza, M. (2008). Ingeniería evolutiva en *Salmonella*: la emergencia de plásmidos híbridos de virulencia-resistencia a antimicrobianos en serotipos no tifoideos. *Enfermedades infecciosas y microbiología clínica*, 37-43.

Miller, J. R., Koren, S., & Sutton, G. (2011). Assembly Algorithms for Next-Generation Sequencing Data. *Genomics*, 315-327.

Molina, M., Martín, H., Cid, V. J., Rotger, R., & Rodríguez, I. (13 de Agosto de 2017). *Universidad Complutense Madrid*. Obtenido de Microbiología: <http://www.ucm.es/microbiologia2/virulencia-bacteriana>

Moore, B. C., Martinez, E., Gay, J., & Rice, D. H. (2003). Survival of *Salmonella enterica* in Freshwater and Sediments and Transmission by the Aquatic Midge *Chironomus tentans*. *Applied and Environmental Microbiology*, 4556-4560.

Morgan, E., Campbell, J., Rowe, S., Bispham, J. S., Bowen, A., Barrow, P. (2004). Identification of host-specific colonization factors of *Salmonella enterica* serovar Typhimurium. *Molecular Microbiology*, 994-1010.

Nagarajan, N., & Pop, M. (2013). Sequence assembly demystified. *Nature Reviews Genetics*, 157-167.

- Nielsen, R. L. (2013). *Salmonella* Dublin in cattle. Denmark: University of Copenhagen.
- Norris, T. B. (1996). Phase variations of the *l*pf operon is a mechanism to evade cross-immunity between *Salmonella* serotypes. *Proceedings of the National Academy of Sciences of the United States of America*, 13393-13398.
- Olive, D. M., & Bean, P. (1999). Principles and Applications of Methods for DNA-Bases Typing of Microbial Organisms. *Journal of Clinical Microbiology*, 1661-1669.
- OMS. (4 de Agosto de 2013). Organización Mundial de la Salud. Recuperado el 27 de Agosto de 2016, de <http://www.who.int/mediacentre/factsheets/fs139/es/>
- Ozen, A., Vesth, T., & Ussery, D. (2013). Comparative Genomics. *Prokaryotic Biology and Symbiotic Associations* , 209-227.
- Palomino-Camargo, C., & Gonzalez-Muños, Y. (2014). Técnicas moleculares para la detección e identificación de patógenos en alimentos: ventajas y limitaciones. *Revista Peruana de Medicina Experimental y Salud Publica*, 535-546.
- Pascual, M. (2005). Enfermedades de origen Alimentario, su prevención. Madrid: Díaz de Santos S.A.
- Peña, C. (2011). Metodos de inferencia filogenética . *Revista Peruana de Biología*, 265-267.

- Puerta-García, A., & Mateos-Rodríguez, F. (2006). Enterobacterias. *Unidad de Enfermedades Infecciosas*, 1-6.
- Pui, C. F., Chai, L. C., Tunung, R., Jeyaletchumi, P., Ubong, A., Farinazlenn, M. G., Son, R. (2011). Review Article *Salmonella*: A foodborne pathogen. *International Food Research Journal*, 465-475.
- Pui, C. W. (2011). Simultaneous detection of *Salmonella* spp., *Salmonella* Typhimurium in sliced fruits using multiplex PCR. *FOOD CONTROL*, 337-342.
- Ritcher, M., & Rosselló-Móra, R. (2018). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences of the United States of America*, 19126-19131.
- Romero, D., & Palacios, R. (1997). Gene Amplification and Genomic Plasticity in Prokaryotes. *Annual Review of Genetics*, 91-111.
- Shendure, J., & Ji, H. (2008). Next generation DNA sequencing. *Nature Biotechnology*, 1135-1145.
- Shu-Kee, E., Priyia, P., Nurul-Syakima, A. M., Hooi-Leng, S., Kok-Gan, C., & Learn-Han, L. (2015). *Salmonella*: A review on pathogenesis, epidemiology and antibiotic resistance. *Frontiers in Life Science*, 1-11.
- Silva, J., Leite, D., Fernandes, M., C, M., Gibbs, P., & Teixeira, P. (2011). *Campylobacter* spp. as a foodborne pathogen: a review. *Frontiers in Microbiology*, 200.

- Urwin, R. M. (2003). Multi-locus sequence typing: a tool for global epidemiology. *TRENDS in Microbiology*, 479-487.
- Valcárcel, V. (2008). Breve introducción a las técnicas y métodos de reconstrucción filogenética . Madrid, España: Universidad Autónoma de Madrid.
- Walker, J., & Raply, R. (2009). *Molecular Biology and Biotechnology*. Cambridge: RSC Publishing.
- Walteros, J., Urregon, F., Montoya, P., & Suarez, J. (2012). Recopilación Bioinformática. *Universidad Distrital Francisco Jose de Caldas*, 15-20.
- Wang, H. H. (2017). Complete Genome Sequence of *Salmonella enterica* subsp. *enterica* serovar Minnesota. *Genome Biology and Evolution*, 2727-2731.
- Winfield, M. D., & Groisman, E. A. (2003). Role of Nonhost Enviroments in the Lifestyles of *Salmonella* and *Escherichia coli*. *Applied and Enviromental Microbiology*, 3687-3694.
- Wood, M., Jones, M., Watson, P., Hedges, S., Wallis, T., & Galyov, E. (1998). Identification of a pathogenicity island required for *Salmonella* enteropathogenicity. *Molecular Microbiology*, 883-891.
- Wu, H.-J., Wang, A., & Jennings, M. P. (2008). Discovery of virulence factors of pathogenic bacteria. *Current Opinion in Chemical Biology* , 1-9.
- Wyres, K., Conway, T., Garg, S., Queiroz, C., Reumann, M., Kathryn, H., & Rusu, L. (2014). WGS Analysis and interpretation in clinical and public health

microbiology laboratories: what are the requirements and how do existing tools compare? *Pathogens*, 437-458.

Xiong, J. (2006). *Essential Bioinformatics*. United States: Cambridge University Press.

Yang, Z. R. (2012). Molecular phylogenetics: principles and practice. *Nature Reviews Genetics*, 303-314.

Zhang, H., Zhou, Y., Bao, H., & Liu, H. W. (2006). Vi antigen biosynthesis in *Salmonella typhi*: characterization of UDP-N-acetylglucosamine C-6 dehydrogenase (TviB) and UDP-N-acetylglucosaminuronic acid C-4 epimerase (TviC). *Biochemistry*, 8163-8173.

XI. ANEXOS

Cuadro 22. Comparación de la serotipificación convencional vs *in silico* para las cepas locales. Se muestra color amarillo las cepas aisladas de río; mientras que las aisladas de animal en gris. Se encuentran sombreadas de amarillo aquellas cepas que si presentaron diferencia en el serotipo asignado.

| Nomenclatura río | Serotipo predicho por serología | Serotipo predicho <i>in silico</i> | ST |
|-------------------|---------------------------------|------------------------------------|-----------|
| C-WEL08001 | Weltevreden | Weltevreden | 365 |
| C-ORA08001 | Oranienburg | Oranienburg | 23 |
| C-SAI08001 | Saintpaul | Saintpaul | 50 |
| C-THO08001 | Infantis | Thompson | 26 |
| C-ORA08002 | Oranienburg | Oranienburg | 23 |
| C-MIN08001 | Minnesota | Minnesota | 548 |
| C-ANA08001 | Anatum | Anatum | 64 |
| C-MIN08002 | Minnesota | Minnesota | 548 |
| C-ORA08003 | Oranienburg | Oranienburg | 23 |
| C-ORA08004 | Oranienburg | Oranienburg | 23 |
| C-ORA08005 | Anatum | Oranienburg | 23 |
| C-MON08001 | Montevideo | Montevideo | 316 |
| C-SAI08002 | Saintpaul | Saintpaul | 50 |
| C-ORA08006 | Oranienburg | Oranienburg | 23 |
| C-SAI08003 | Saintpaul | Saintpaul | 50 |
| C-ORA08007 | Oranienburg | Oranienburg | 23 |
| C-SAI08004 | Saintpaul | Saintpaul | 50 |
| C-ORA08008 | Oranienburg | Oranienburg | 23 |
| C-MIN08003 | Minnesota | Minnesota | 548 |
| C-ANA08002 | Anatum | Anatum | 64 |
| C-MIN08004 | Minnesota | Minnesota | 548 |
| C-MON08002 | Montevideo | Montevideo | 316 |
| C-THO08002 | Infantis | Thompson | 26 |
| C-MIN08005 | Minnesota | Minnesota | 548 |
| C-ORA08009 | Oranienburg | Oranienburg | 23 |
| C-ORA08010 | Oranienburg | Oranienburg | 23 |

| | | | |
|------------|-------------|-------------|-----|
| C-ORA08011 | Oranienburg | Oranienburg | 23 |
| C-WEL08002 | Weltevreden | Weltevreden | 365 |
| C-INF08001 | Infantis | Infantis | 32 |
| C-MIN08006 | Minnesota | Minnesota | 548 |
| C-GIV08001 | Oranienburg | Give | 654 |
| C-WEL08003 | Weltevreden | Weltevreden | 365 |
| C-GIV08002 | Give | Give | 516 |
| C-MUE08001 | Muenster | Muenster | 321 |
| C-MUE09001 | Muenster | Muenster | 321 |
| C-MUE08002 | Muenster | Muenster | 321 |
| C-MUE08003 | Muenster | Muenster | 321 |
| C-MON08003 | Montevideo | Montevideo | 81 |
| C-WEL09001 | Weltevreden | Weltevreden | 365 |
| C-MUE08004 | Muenster | Muenster | 321 |
| C-MON09001 | Montevideo | Montevideo | 81 |
| C-ANA09001 | Anatum | Anatum | 64 |
| C-ANA09002 | Anatum | Anatum | 64 |
| C-POO09001 | Poona | Poona | 447 |
| C-GIV09001 | Give | Give | 654 |
| C-ORA09001 | Oranienburg | Oranienburg | 23 |
| C-SAN09001 | Oranienburg | Sandiego | ? |
| C-SAI09001 | Saintpaul | Saintpaul | 50 |
| C-SAI09002 | Saintpaul | Saintpaul | 50 |
| C-GIV09002 | Give | Give | 654 |
| C-ORA09002 | Oranienburg | Oranienburg | 23 |
| C-SAI08005 | Saintpaul | Saintpaul | 50 |
| C-ORA08012 | Oranienburg | Oranienburg | 23 |
| C-SAI08006 | Saintpaul | Saintpaul | 50 |
| C-POM08001 | Oranienburg | Pomona | 451 |
| C-ORA08013 | Oranienburg | Oranienburg | 23 |
| C-POM08002 | Pomona | Pomona | 451 |
| C-ORA08014 | Give | Oranienburg | 23 |
| C-POM08003 | Pomona | Pomona | 451 |
| C-ORA08015 | Oranienburg | Oranienburg | 23 |
| C-POO09002 | Poona | Poona | 447 |
| C-POO09003 | Poona | Poona | 447 |
| C-INF09001 | Infantis | Infantis | 32 |
| C-ORA08016 | Oranienburg | Oranienburg | 23 |
| C-SAI09003 | Saintpaul | Saintpaul | 50 |
| C-POM09001 | Pomona | Pomona | 451 |
| C-POO09004 | Oranienburg | Poona | 447 |
| C-ORA09003 | Oranienburg | Oranienburg | 23 |
| C-INF09002 | Infantis | Infantis | 32 |
| C-MIN09001 | Minnesota | Minnesota | 548 |

| | | | |
|-------------------|-------------|-------------|-----|
| C-POM09002 | Pomona | Pomona | 451 |
| C-GIV09003 | Give | Give | 654 |
| C-ORA09004 | Oranienburg | Oranienburg | 23 |
| C-POO09005 | Poona | Poona | 447 |
| C-SAI09004 | Saintpaul | Saintpaul | 50 |
| C-POO09006 | Poona | Poona | 447 |
| C-POM09003 | Pomona | Pomona | 451 |
| C-ORA09005 | Oranienburg | Oranienburg | 23 |
| C-ORA09006 | Oranienburg | Oranienburg | 23 |
| C-ORA09009 | Oranienburg | Oranienburg | 23 |
| C-ORA09010 | Oranienburg | Oranienburg | 23 |

| Nomenclatura animal | Serotipo predicho por serología | Serotipo predicho <i>in silico</i> | ST |
|----------------------------|--|---|-----------|
| CA-ORA08001 | Oranienburg | Oranienburg | ? |
| CA-ANA08001 | Anatum | Anatum | 64 |
| CA-ORA08002 | Oranienburg | Oranienburg | 23 |
| CA-ANA08002 | Anatum | Anatum | 64 |
| CA-GIV09001 | Give | Give | 654 |
| CA-ORA08003 | Oranienburg | Oranienburg | 23 |
| CA-ORA08004 | Oranienburg | Oranienburg | 23 |
| CA-ALB08001 | Albany | Albany | 292 |
| CA-ORA08005 | Oranienburg | Oranienburg | 23 |
| CA-WEL08001 | Weltevreden | Weltevreden | 365 |
| CA-ORA08006 | Oranienburg | Oranienburg | 23 |
| CA-WEL08002 | Weltevreden | Weltevreden | 365 |
| CA-WEL08003 | Weltevreden | Weltevreden | 365 |
| CA-JAV09001 | Javiana | Javiana | 24 |
| CA-ORA08007 | Oranienburg | Oranienburg | 23 |
| CA-NEW08001 | Newport | Newport | 118 |
| CA-SAI08001 | Saintpaul | Saintpaul | 50 |
| CA-ORA08008 | Oranienburg | Oranienburg | 23 |
| CA-GIV08001 | Give | Give | 654 |
| CA-GIV08002 | Give | Give | 654 |
| CA-MUE08001 | Muenchen | Muenchen | 112 |
| CA-CAY09001 | Cayar | Cayar | |
| CA-SAI08002 | Saintpaul | Saintpaul | 50 |
| CA-SAI08003 | Saintpaul | Saintpaul | |
| CA-SAI08004 | Saintpaul | Saintpaul | 50 |
| CA-SAI08005 | Saintpaul | Saintpaul | 50 |
| CA-SAI08006 | Saintpaul | Saintpaul | 50 |

| | | | |
|-------------|---------------|-------------|------|
| CA-SAI08007 | Saintpaul | Saintpaul | 50 |
| CA-SAI08008 | Saintpaul | Saintpaul | 50 |
| CA-TYP08001 | Typhimirium | Typhimirium | 19 |
| CA-GIV08003 | Give | Give | 654 |
| CA-SAI08009 | Saintpaul | Saintpaul | 50 |
| CA-ORA08009 | Oranienburg | Oranienburg | 23 |
| CA-GIV08004 | Give | Give | 654 |
| CA-GIV08005 | Give | Give | 654 |
| CA-ORA08010 | Oranienburg | Oranienburg | 23 |
| CA-ORA08011 | Oranienburg | Oranienburg | 23 |
| CA-ORA08012 | Oranienburg | Oranienburg | 23 |
| CA-ORA08013 | Oranienburg | Oranienburg | 23 |
| CA-MON08001 | Montevideo | Montevideo | 316 |
| CA-ORA08014 | Oranienburg | Oranienburg | 23 |
| CA-AGO08001 | Agona | Agona | 13 |
| CA-ORA08015 | Oranienburg | Oranienburg | 23 |
| CA-GIV08006 | Give | Give | 654 |
| CA-GIV08007 | Give | Give | 654 |
| CA-AGO08002 | Agona | Agona | 13 |
| CA-MIN08001 | Minnesota | Minnesota | 548 |
| CA-ORA08016 | Oranienburg | Oranienburg | 23 |
| CA-GIV09002 | Give | Give | 654 |
| CA-GIV08008 | Give | Give | 654 |
| CA-MIN08002 | Minnesota | Minnesota | 548 |
| CA-MIN08003 | Minnesota | Minnesota | |
| CA-MIN08004 | Minnesota | Minnesota | 548 |
| CA-THO08001 | C1 monofasica | Thompson | 26 |
| CA-MIN08005 | Minnesota | Minnesota | 548 |
| CA-THO08002 | C1 monofasica | Thompson | 26 |
| CA-NEW08002 | Newport | Newport | 45 |
| CA-ANA08003 | Anatum | Anatum | 64 |
| CA-ANA08004 | Anatum | Anatum | 64 |
| CA-MIN08006 | Minnesota | Minnesota | |
| CA-ORA08017 | Oranienburg | Oranienburg | 23 |
| CA-MIN08007 | Minnesota | Minnesota | 548 |
| CA-ORA08018 | Oranienburg | Oranienburg | 23 |
| CA-MIN08008 | Minnesota | Minnesota | |
| CA-LUC08001 | Luciana | Luciana | 3243 |
| CA-LUC08002 | Luciana | Luciana | 3243 |
| CA-LUC08003 | Luciana | Luciana | 3243 |
| CA-THO08003 | C1 monofasica | Thompson | 26 |
| CA-SAH08001 | Sahanina | Sahanina | 970 |

| | | | |
|--------------------|----------|----------|-----|
| CA-GAM08001 | Gaminara | Gaminara | 239 |
| CA-SAH08002 | Sahanina | Sahanina | 970 |
| CA-GIV09003 | Give | Give | 654 |

Abreviaturas

| Abreviatura | Descripción |
|-------------|--|
| ADN | Ácido desoxirribonucleico |
| ANI | Identidad de nucleótido promedio |
| ARN | Ácido ribonucleico |
| ARNt | Ácido ribonucleico transferencia |
| BLAST | Basic Local Alignment Tool |
| CDC | Centro para el Control y Prevención de Enfermedades |
| cgMLST | Tipificación multilocus de secuencias con base en el genoma core |
| DLV | Variante en dos locus |
| dNTPs | Deoxinucleótidos trifosfatos |
| eBG | Grupo eburst |
| EFSA | Autoridad Europea de Seguridad Alimentaria |
| ETA | Enfermedades de transmisión alimentaría |
| iTOL | Interactive Tree of Life |
| Kb | Kilo bases |
| LPS | Lipopolisacárido |
| MLST | Tipificación multilocus de secuencias |
| NCBI | Centro Nacional para la Información Biotecnológica |
| NGS | Secuenciación de siguiente generación |
| OMS | Organización Mundial de la Salud |

| | |
|---------|--|
| PCR | Reacción en cadena de la polimerasa |
| PFGE | Electroforesis en gel de campo pulsado |
| PHASTER | Phage Search Tool Enhanced Release |
| PHG | Proyecto del Genoma Humano |
| rMLST | Tipificación multilocus de secuencias con base en ribosoma |
| SiNAVE | Sistema Nacional de Vigilancia Epidemiológica |
| SISTR | <i>Salmonella</i> in silico Typing Resource |
| SLV | Variante en un solo locus |
| SNPs | Polimorfismos de nucleótido simple |
| SPI | Isla de patogenicidad de <i>Salmonella</i> |
| SPI-1 | Isla de patogenicidad 1 de <i>Salmonella</i> |
| SPI-2 | Isla de patogenicidad 2 de <i>Salmonella</i> |
| SPI-3 | Isla de patogenicidad 3 de <i>Salmonella</i> |
| SPI-4 | Isla de patogenicidad 4 de <i>Salmonella</i> |
| SPI-5 | Isla de patogenicidad 5 de <i>Salmonella</i> |
| ST | Tipo de secuencia |
| T3SS | Sistema de secreción tipo III |