



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

**Análisis Genómico y Clonal de *Salmonella*
Oranienburg Aisladas de Ríos de Culiacán,
Sinaloa**

T E S I S

que presenta

Cesar Octavio Carreón Gaxiola

como requisito para obtener
el grado académico de

**Maestro en Ciencias con
Orientación en Biotecnología
de Salud**

Director

Dra. Maribel Jiménez Edeza

Dra. Gloria Marisol Castañeda Ruelas

Culiacán de Rosales, SIN, MEX

Mayo de 2017

Presentación

La presente investigación, titulada “Análisis Genómico y Clonal de Cepas de *Salmonella* Oranienburg Aisladas de Ríos de Culiacán, Sinaloa”, se llevó a cabo en el “Laboratorio de Investigación y Diagnóstico Microbiológico”, Programa Regional de Posgrado en Biotecnología, Facultad de Ciencias Químico Biológicas, Universidad Autónoma de Sinaloa. Los C Dr. Héctor Samuel López Moreno, Dra. Claudia del Rosario León Sicaños, Dr. José Geovanni Romero Quintana, Dra. Gloria Marisol Castañeda Ruelas y Dra. Maribel Jiménez Edeza participaron como Miembros del H. Comité Tutorial / Asesores Académicos. Las dos últimas investigadoras fungieron como Director de Tesis. La mencionada investigación forma parte del Proyecto de Investigación “Estudio del Perfil Fenotípico, Genotípico y Potencial Patogénico de Poblaciones Dominantes de *Salmonella enterica* Aisladas de Ríos del Estado de Sinaloa”. Este proyecto fue financiado por el Consejo Nacional de Ciencia y Tecnología (CONACyT), México (Clave CB-2014/235989). La investigación tuvo carácter interinstitucional; en ella participaron “Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria” (Centro Nacional de Referencia en Detección de Organismos Genéticamente Modificados), “Administración para los Alimentos y Medicamentos (FDA, por sus siglas en inglés)” de los Estados Unidos de América, y la Universidad Autónoma de Sinaloa.

El C. César Octavio Carreón Gaxiola recibió Beca Nacional para Estudios de Posgrado del CONACyT durante sus estudios (Maestría en Ciencias con Orientación en Biotecnología de Salud, Programa Regional de Posgrado en Biotecnología, Facultad de Ciencias Químico Biológicas, Universidad Autónoma de Sinaloa).

Agradecimientos

Agradecer primeramente a la Universidad Autónoma de Sinaloa por haberme sido alma mater en mi formación durante el nivel superior y maestría.

A la Facultad de Ciencias Químico Biológicas por brindarme todos los conocimientos, así como a los docentes del nivel superior y maestría durante mi formación como maestro en ciencias.

Al Programa Regional de Posgrado en Biotecnología por ser una plataforma de formación y superación personal.

A la Dra. Maribel Jiménez Edeza por haberme aceptado y formar parte de su equipo de laboratorio, por todo el apoyo durante el desarrollo de esta tesis, así como las facilidades que me dio para completar este ciclo, pero sobre todo doy gracias por esos regaños y presiones que siempre fueron muy acertados, tanto por sus grandes consejos no solo como directora de tesis sino como una amiga que a lo largo de estos dos años me permitieron crecer tanto académicamente como personalmente.

A la Dra. Gloria Marisol Castañeda Ruelas, por su entrega a cada revisión de tesis, por su apoyo incondicional, así como las horas que pasamos trabajando en el documento y presentación y por siempre presionarme y hacerme ver más allá de lo que se lee en las páginas de los artículos, por esos consejos que me ayudaron mucho para resolver situaciones personales y profesionales, porque sé que puedo contar con ella siempre. ¡Gracias!

Al Dr. José Geovanni Romero Quintana Morales por colaborar en el comité de tesis, por su asesoría, apoyo y disposición en todo momento, desde la carrera donde siempre ha estado para sacarme de dudas en metodologías y sobre todo por su gran amistad.

A la Dra. Claudia del Rosario León Sicairos por colaborar en el comité de tesis en este proyecto, y por la asesoría y apoyo brindados en este trabajo.

Al Dr. Héctor Samuel López Moreno por colaborar en el comité de tesis en este proyecto, por la carrilla desde la carrera, así como el apoyo y brindarme un poco de su amistad y camaradería.

A los integrantes del SENASICA del Laboratorio de OGM: Hugo Castelán, Abraham Actazi, Alejandra García, Ángel Romero, Sandra Agüeros, Roberto, Josh, Abigail; a todos ellos muchas gracias por sus experiencias, consejos y por la asesoría en la parte bioinformática y contribución a este trabajo.

Compañeros de Laboratorio LIDiM: Maylet, Kenia, Sergio, Guadalupe, Maritza, Elizabeth, Suluay a todos ellos muchas gracias por su amistad y apoyo, sobre todo por los momentos que pasamos en el Lab.

A mis compañeros y amigos de generación de la Maestría: Evelin, Yuriana, Alex, Leticia, Edgar, Manuel, Nayeli, Mariana, Ruy, German y Raquel. Por esos momentos de estudios fuera y dentro de clases, como también los momentos personales que pasamos juntos; por esto y más los llevare en mi corazón chicos!!.

A mis amigos incondicionales de la vida: Yamilt, Luis, Alejandro, Eveling, Koki, Celina Jessica, Jaime, América. Que sin su amistad y mucho de su apoyo esto no hubiera salido adelante, que siempre podré confiar en ustedes, por sus consejos, abrazos y felicitaciones; sé que, aunque sean pocos, siempre estarán en mi corazón por el resto de mi vida, así como siempre lo han hecho.

A mis padres, hermanas y abuelo, que sin su apoyo y consejos los objetivos de estas tesis jamás se hubieran cumplido, porque estuvieron ahí en cada momento y son mi motor en la vida para seguir adelante, sobre todo porque los amo y

espero me perdonen todas las ausencias que tuve durante estos dos años y sigamos compartiendo todos los momentos en familia. ¡Los amo con todo mi corazón!

A mi novia Paloma, a tí que te conocí en una etapa temprana de la maestría y descubrimos juntos muchos caminos, tanto personales como académicos, por ser una de mis mayores motivaciones; doy gracias por las peleas, berrinches, cariños y consejos que tuvimos durante esta etapa, la cual nos permitió crecer como pareja y permitirte estar a mi lado durante estos dos años y ser partícipe de este proyecto, aun sabiendo que no ha sido nada fácil, sé que siempre estarás ahí, te amo.

ÍNDICE GENERAL

	Pág.
ÍNDICE GENERAL	
ÍNDICE DE CUADROS	
ÍNDICE DE FIGURAS	
I RESUMEN	1
I ABSTRACT	2
II INTRODUCCIÓN	3
III REVISIÓN DE LA LITERATURA	6
A CARACTERÍSTICAS DEL GÉNERO <i>Salmonella</i>	6
1 Historia	6
2 Morfología	7
3 Perfil bioquímico	9
4 Antígenos de superficie	11
5 Taxonomía de <i>Salmonella</i>	12
6 Ciclo de vida de <i>Salmonella</i>	14
a Suelo	17
b Cuerpos de agua	17
c Alimentos	17
d Hospederos	18
B PATOGENIA DE <i>Salmonella</i>	18
1 Ciclo patogénico	19
2 Factores de virulencia e islas de patogenicidad	20
3 Genes de resistencia	26
C EPIDEMIOLOGÍA DE LA SNT	26
1 Tipos de Salmonelosis	27
a Fiebre tifoidea	27
b Salmonelosis no tifoidea (SNT)	28
c Estado del portador	28

	2 Casos y brotes	29
	D PANORAMA DE S. Oranienburg EN SINALOA	31
	1 Fuentes	31
	2 Casos y brotes	31
	3 Caracterización molecular	32
	E EPIDEMIOLOGÍA MOLECULAR DE M.O.	33
	1 Definición y aplicación	33
	2 Métodos de tipificación molecular	33
	3 Electroforesis en Gel de Campo pulsado (PFGE)	34
	4 Tipificación Multilocus de Secuencia (MLST)	35
	5 Secuenciación de nueva generación (NGS)	37
	a Historia de la secuenciación de nueva generación	37
	b Plataforma de secuenciación masiva illumina	40
	c Tipos de bibliotecas o genotecas en la secuenciación masiva	41
	d Secuenciación de nueva generación aplicada para el estudio de microorganismos	43
	e Algoritmos de ensamblado <i>de novo</i>	43
IV	JUSTIFICACIÓN	47
V	HIPÓTESIS	48
VI	OBJETIVOS	49
	A OBJETIVO GENERAL	49
	B OBJETIVOS ESPECÍFICOS	49
VII	MATERIALES Y MÉTODOS	50
	A CEPAS DE S. Oranienburg	50
	B EXTRACCIÓN DE ÁCIDOS NUCLEICOS DE CEPAS DE S. Oranienburg	50
	C GENERACIÓN DE GENOTECAS	52
	1 Tagmentación del ADN de entrada	52

	2 Amplificación de librerías	52
	3 Limpieza de librerías	55
	4 Verificación del control de calidad de las librerías de ADN y cuantificación con Qubit.	55
	D SECUENCIACIÓN	56
	1 Preparación de las genotecas	56
	E ANÁLISIS BIOINFORMÁTICO PRIMARIO	56
	1 Exportación de los datos de corrida de MiSeq	56
	2 Control de calidad de las secuencias genómicas	58
	3 Recorte de adaptadores y recorte por calidad	58
	4 Ensamble de secuencias genómicas <i>de novo</i>	59
	5 Evaluación de los ensamblados <i>de novo</i>	59
	6 Alineamiento de genomas	59
	F ANOTACIÓN DE LOS GENOMAS	60
	1 Obtención del pseudocromosoma	60
	G ARBOL FILOGENÉTICO DE GENOMA COMPLETO (WGS)	60
	H ARBOL FILOGNÉTICO DE GENES DE VIRULENCIA	61
	I VIRULOTIPIACIÓN <i>in-silico</i>	61
	J TIPIFICACIÓN POR MULTILOCUS DE SECUENCIA (MLST) Y RESISTENCIA A ANTIBIÓTICOS	62
	1 Resistencia antibióticos	63
	2 Serotipificación <i>in-silico</i>	63
VII	RESULTADOS Y DISCUSIÓN	64
	A ENSAMBLE Y ANOTACIÓN DE LOS GENOMAS DE <i>S. Oranienburg</i> AISLADAS DE AGUAS DE RÍOS DEL VALLE DE CULIACÁN	64
	1 Obtención los archivos crudos (<i>.fastq</i>) del secuenciador	64
	2 Evaluación de calidad de las cepas	64

	3 Evaluación de las comparaciones por recorte de adaptadores y calidad (Trimmomatic vs PEAT vs CLC)	65
	4 Ensamble del genoma de las cepas de S. Oranienburg.	72
	5 Alineamiento de ensamblajes de S. Oranienburg	77
	6 Anotación de los genomas por RAST.	81
	B DETERMINAR LAS LÍNEAS CLONALES DE LOS AISLAMIENTOS DE S. Oranienburg EN LOS RÍOS DE SINALOA	85
	1 Árbol filogenético de las cepas de S. Oranienburg en función al genoma completo	85
	2 Árbol filogenético de genes de virulencia	88
	C TIPIFICACIÓN POR MULTILOCUS DE SECUENCIA (MLST) DE LAS CEPAS	91
	1 Árbol recubridor mínimo (MST)	91
	D TIPIFICACIÓN DE GENES DE DEFENSA DE LAS CEPAS DE S. Oranienburg MEDIANTE ANÁLISIS <i>in-silico</i> .	96
	1 Virulotipificación	96
	2 Resistencia a antibióticos <i>in silico</i>	100
IX	CONCLUSIONES	102
X	BIBLIOGRAFÍA	103
	Abreviaturas	131

ÍNDICE DE CUADROS

Cuadro	Descripción	Pág.
1	Características bioquímicas diferenciales de especies de <i>Salmonella</i> y subespecies	10
2	Clasificación de <i>Salmonella</i>	13
3	Siete islas de patogenicidad conocidas	22
4	Brotos de enfermedades transmitidas por los alimentos por <i>Salmonella</i>	30
5	Principales características de las plataformas de secuenciación de nueva generación	39
6	Programa de tagmentación	54
7	Códigos obtenidos para el manejo de las 18 muestras	66
8	Parámetros de las lecturas obtenidos con el programa FastQC	67
9	Metadatos de cepas de <i>Salmonella</i> Oranienburg obtenidas de agua de río con RAST	83
10	MLST de las cepas de <i>S. Oranienburg</i> con los 7 genes constitutivos, ST y ST complex.	93
11	Serotipificación in-silico de las muestras con el programa SeqSero.	94

ÍNDICE DE FIGURAS

Figura	Descripción	Pág.
1	Micrografía electrónica de barrido de <i>Salmonella</i> no tifoidea donde se muestra su morfología	8
2	Ciclo de vida de <i>Salmonella</i>	16
3	Ciclo de patogenicidad	25
4	Representación esquemática de los diferentes niveles de resolución para la identificación bacteriana	36
5	Tipos de genotecas	42
6	Algoritmo overlap-layout-consensus.	45
7	Algoritmos basados en grafos de Bruijn	46
8	Flujo de trabajo para generación de genotecas.	53
9	Formato de archivo descomprimido ".fastq"	57
10	Comparación de los programas Trimmomatic, PEAT (softwares libres) y CLC Genomics Workbench (software de licencia) de recorte de adaptadores de las muestras.	69
11	Comparación de los programas Trimmomatic (software libre) y CLC Genomics Workbench (software de licencia) con recortes de calidad.	71
12	Comparación del número de contigs realizados por los programas CLC y A5-MiSeq	74
13	Comparación del N50 realizados por los programas CLC y A5-MiSeq	75
14	Comparación de la métrica de contig más largo obtenidos entre los programas CLC y A5-MiSeq	76
15	Reconstrucción filogenética	77
16	Alineamiento progresivo de las 18 secuencias de <i>Salmonella</i> Oranienburg.	80

17	Representación general de la anotación de las 18 <i>Salmonella</i> Oranienburg.	82
18	Árbol filogenético de genoma completo (WGS)	87
19	Árbol filogenético de genes de virulencia hecho con Mrbayes	90
20	Árbol recubridor mínimo (MST, siglas en inglés) de cepas de <i>Salmonella</i> Oranienburg y otros serotipos no tifoideos, obtenido con el programa PHYLOViZ a partir de MLST	92
21	Dendrograma e índices de virulencia	99
22	Dendrograma de resistencia a antibióticos	102

I RESUMEN

Salmonella Oranienburg representa el serotipo no tifoideo predominante en los ríos y animales domésticos de Sinaloa, lo cual representa un riesgo para la distribución de la bacteria entre el entorno-hospederos. La epidemiología molecular aplicada en microorganismos involucra el uso de métodos moleculares y bioinformáticos como herramienta para determinar su evolución, hospederos y reservorios, así como los factores de virulencia y resistencia que dificultan su control. El objetivo del estudio fue determinar la relación poblacional de cepas de *S. Oranienburg* aisladas de agua de río de Culiacán, Sinaloa. A partir de 18 cepas de *S. Oranienburg* previamente aisladas de río en la región, se prepararon las genotecas (2 ng de ADN genómico) utilizando el kit Nextera-XT, y se secuenciaron con la plataforma Illumina Miseq (2 x 250). Los extremos de lecturas fueron recortados con los programas PEAT-v1-2.1.4 y Trimmomatic-0.36. El análisis de ensamblaje y anotación se realizó con el pipeline A5-miseq y la plataforma Rapid Annotation using Subsystem Technology, respectivamente. Para determinar la relación filogenética de *S. Oranienburg* con base al genoma, virulencia, especie, serotipo y resistencia antimicrobiana se utilizaron los programas Realphy, Mrbayes, Tipificación por Multilocus de Secuencia (MLST), Seqsero y Srst2, respectivamente. Finalmente, se determinó el virulotipo de las cepas mediante la construcción de una matriz de 61 genes de virulencia (5 islas de patogenicidad, fimbrias y plásmidos). El tamaño del genoma de las 18 cepas de *S. Oranienburg* osciló entre 4.57-5.00 Mb con un %G+C de 52.0. El MLST y los árboles filogenéticos permiten inferir que las cepas pertenecen a la misma clona dado la topología obtenida; 3 clonas, 3 virulotipos, 1 serotipo (*S. Oranienburg*), y 1 perfil de resistencia antimicrobiana (aminoglucósidos). La tipificación *in-silico* permite definir un serotipo conservado de *S. Oranienburg* con atributos potencialmente patogénicos y metabólicos que permiten mantener su ciclo de vida.

Palabras clave: *Salmonella* Oranienburg, Río de Culiacán

ABSTRACT

Salmonella Oranienburg represents the non-typhoid serotype prevalent in rivers and domestic animals of Sinaloa, which represents a risk for the distribution of the bacteria among the host-environment. The molecular epidemiology applied in microorganisms involves the use of molecular and bioinformatic methods as a tool to determine their evolution, hosts and reservoirs, as well as virulence factors and resistance that hinder their control. The objective of the study was determined by the population of strains of *S. Oranienburg* isolated from water of river of Culiacán, Sinaloa. From 18 strains of *S. Oranienburg* were isolated from the river in the region, the libraries (2 ng genomic DNA) were prepared using the Nextera-XT kit, and sequenced with the Illumina Miseq (2 x 250) platform. The readout ends were trimmed with the programs PEAT-v1-2.1.4 and Trimmomatic-0.36. The assembly and annotation analysis was performed with the A5-miseq pipeline and the Quick Annotation platform using subsystem technology, respectively. To determine the phylogenetic relationship of *S. Oranienburg* based on the genome, virulence, species, serotype and antimicrobial resistance, the programs Realphy, Mrbayes, Multilocus Sequence Typing (MLST), Seqsero and Srst2, respectively, were used. Finally, the virulotype of the strains was determined by constructing an array of 61 virulence genes (5 pathogenicity islands, fimbriae and plasmids). The genome size of the 18 *S. Oranienburg* strains ranged from 4.57-5.00 Mb with a% G + C of 52.0. The MLST and the phylogenetic trees allow to infer that the strains belong to the same part of the topology obtained; 3 clones, 3 virulotypes, 1 serotype (*S. Oranienburg*), and 1 antimicrobial resistance profile (aminoglycosides). The inelastic typing allows to define a conserved serotype of *S. Oranienburg* with potentially pathogenic and metabolic attributes that allow to maintain its life cycle.

Key words: *Salmonella* Oranienburg, Río de Culiacán

II INTRODUCCIÓN

El género *Salmonella* pertenece a la familia *Enterobacteriaceae* y su hábitat fundamental es el tracto intestinal tanto como de humanos y animales, siendo la salmonelosis una de las causas más frecuentes de enfermedades transmitidas por los alimentos en todo el mundo. Por ello, la transmisión de la salmonelosis a los seres humanos ha sido tradicionalmente ligada al consumo de alimentos contaminados con *Salmonella* (Jiménez y col., 2010; Lyimo y col., 2016; Mała y Popowska, 2016).

Salmonella comprende >2,400 serotipos de *Salmonella* (Popoff y col. 2004), pero, sólo 50 serotipos se han asociado a infecciones en seres humanos o animales. Mientras que, la patogenicidad del resto de los serotipos no ha sido claramente definida (Tavechio y col. 2002). Por otra parte, los serotipos prevalentes en los entornos ambientales rara vez coinciden con los serotipos asociados a infecciones a humanos (Martínez-Urtaza y col. 2004), lo que plantea interrogantes acerca de los orígenes de estos serotipos, la fuente de contaminación y su capacidad para prosperar en el ambiente.

El papel de los entornos no hospederos (medio ambiente) en la transmisión de *Salmonella* a los seres humanos ha adquirido relevancia en los últimos años (Lynch y col., 2009). La capacidad de *Salmonella* para sobrevivir en entornos no hospederos indica que no es un paso crítico para colonizar nuevos anfitriones (Winfield y Groisman, 2003). Estas cuestiones son especialmente importantes en las zonas endémicas y de alta prevalencia de salmonelosis, donde el ambiente puede proporcionar un refugio y las condiciones ideales para nuevas variantes genéticas de *Salmonella* (Jiménez y col. 2014).

En particular, Sinaloa es un área caracterizada por una alta prevalencia de salmonelosis, ya que se han reportado tasas altas de infección por este patógeno. Al respecto, la Dirección General de Epidemiología de México (DGE) reportó a nivel nacional un total de 134,833 infecciones por *Salmonella* (salmonelosis y

fiebres tifoideas) en 2015, lo que representa 62.7 casos por cada 100,000 habitantes. Mientras que, en Sinaloa el número de casos registrados fue de 13,304, lo que representa una incidencia de 76.4 casos por cada 100,000 habitantes (DGE, 2017). Adicionalmente en Sinaloa, Jiménez y col. (2011) reportaron una prevalencia alta (31.7%) de *Salmonella* en animales asintomáticos que son criados para consumo humano (Jiménez y col., 2011).

Diversos trabajos han demostrado que en la región agrícola de la zona centro del estado de Sinaloa, existe la presencia de *Salmonella* en ciertos ambientes acuáticos (ríos, esteros y canales) y suelo, por lo cual se han identificado y caracterizado los serotipos de *Salmonella* circulantes en dichos ambientes; S. Oranienburg se ha identificado como el más prevalente en Sinaloa (López-Cuevas et al., 2009; Estrada, 2011; Jiménez, 2011). Aunado a esto, se ha reportado que este serotipo muestra una heterogeneidad genética alta entre los aislamientos mediante la técnica de Electroforesis en Gel de Campos Pulsados (PFGE). Sin embargo, no se ha podido establecer una relación genética entre los diversos clones de S. Oranienburg, potencial patogénico y su distribución geográfica o temporal.

La epidemiología molecular combina la epidemiología clínica con métodos moleculares de laboratorio para determinar la historia natural de una enfermedad, es decir, su etiología, la frecuencia, vías y patrón distribución, reservorios y factores que incrementen el riesgo de controlar el patógeno y la enfermedad (Vilchíz y Alonso, 2009, Johnson y col., 2005). Dentro de los métodos moleculares se utilizan la secuenciación de ácidos nucleicos, la electroforesis en gel de campos pulsados, la reacción en cadena de la polimerasa y la tipificación de multilocus de secuencia (Atchman y col., 2012; Vilchíz y Alonso, 2009; Vázquez y Berron, 2004; Urwin y Maiden, 2003; Li y col., 2009).

La secuenciación del genoma completo y la tipificación por multilocus de secuencia (MLST, siglas en inglés) basado en la secuenciación de genes de mantenimiento (*house keeping*, terminología en inglés) se han difundido actualmente para el estudio de *Salmonella* (Pérez-Lozada y col., 2012). La

complementación del análisis bioinformático a estas herramientas ha permitido contribuir a la biología de esta bacteria, mostrando su heterogeneidad genética respecto a los atributos como virulencia, resistencia, estrés, metabolismo. Por ellos, se requiere la aplicación y desarrollo de las herramientas de la epidemiología molecular a fin de contribuir a la evolución y tipificación del patógeno, así como la búsqueda de marcadores moleculares (Atchman y col., 2012; Pérez-Lozada y col., 2012; Vilchíz y Alonso, 2009).

Por lo tanto, el objetivo de este trabajo es aplicar el uso de herramientas de epidemiología molecular en aislamientos de *Salmonella* Oranienburg aisladas de agua de ríos de Culiacán, Sinaloa, para establecer la relación poblacional respecto a su clonalidad, variación genética, virulencia, resistencia a antibióticos y serología.

III REVISIÓN DE LITERATURA

A Características del género *Salmonella* spp.

El género *Salmonella* es un grupo de bacterias que pertenece a la familia *Enterobacteriaceae* y su hábitat fundamental es el tracto intestinal de humanos y animales (mamíferos de sangre caliente, aves, anfibios, peces y réptiles). Los miembros de este género destacan por infectar y causar enfermedad a un amplio rango de hospederos (Winfield y Groisman, 2003). Morfológicamente y bioquímicamente este género es un grupo muy homogéneo de bacilos Gram negativos (Farmer y col., 1985). *Salmonella* spp. suele aislarse de fuentes ambientales que permiten la diseminación de la bacteria, y juegan un papel importante en su transmisión entre diferentes hospederos (Liejelbjelke y col., 2005).

Historia

En 1885, Salmon y Smith aislaron una bacteria de cerdos infectados con cólera. La bacteria responsable de la infección fue nombrada *Salmonella choleraesuis* (Le Minor y col., 1984; Humphrey, 2006). Posteriormente, el microbiólogo Daniel Elmer Salmon, definió y nombró el género *Salmonella*, en el cual se incluía todo el grupo de microorganismos relacionados con *Salmonella*. Actualmente, se conoce la clasificación de *Salmonella* en dos especies: *S. enterica* y *S. bongori* (Brenner, 2000). Adicionalmente, *S. enterica* se clasifica en seis subespecies *S. enterica* subs. *enterica* (I), *S. enterica* subs. *salamae* (II), *S. enterica* subs. *arizonae* (IIIa), *S. enterica* subs. *diarizonae* (IIIb), *S. enterica* subs. *houtenae* (IV), *S. enterica* subs. *indica* (VI). Esta especie comprenden >2400 serotipos (Brenner, 2000).

En 1934, se realizó la primera publicación de los serotipos de *Salmonella*, la cual consistía en un listado de 44 serotipos mediante el esquema de Kuffman-White. A finales de 1964, el listado contaba con 958 serotipos. Posteriormente,

Le Minor y MI Popoff incrementaron el esquema de serotipos a >2,400 tipos serológicos diferentes (Popoff, 2001).

Morfología

De acuerdo a las características morfológicas y bioquímicas, el género *Salmonella* es un grupo constituido por bacilos Gram negativos anaerobios facultativos, no formadores de esporas y móviles mediante flagelos peritricos (Fig. 1) (Wray, 2000). *Salmonella* spp. se multiplica en amplios rangos de temperatura (7°C a 48°C), pH de 4-8 actividad de agua por debajo de 0,93. Además, *Salmonella* spp. puede sobrevivir durante meses en diferentes sustratos orgánicos y la inactivación de *Salmonella* spp. se produce a pH >5 y a temperaturas superiores a 60°C (Urías y Suárez, 2006).

Perfil bioquímico

La determinación del perfil bioquímico de *Salmonella* spp. se establece como una técnica de identificación, ya que se trata de una característica distintiva y estable en cada especie bacteriana capaz de diferenciar a la mayoría del grupo de las enterobacterias (Farmer y col., 1985). El Cuadro 1 muestra el perfil bioquímico de las especies y subespecies de *Salmonella*.

Básicamente, *Salmonella enterica* se caracteriza bioquímicamente por utilizar el citrato como única fuente de carbono, es positiva a la reacción de arginina dihidrolasa, lisina, ornitina y catalasa, pero, son oxidasa, indol y urea negativa. La fermentación de la glucosa y otros carbohidratos ocurre con la producción usualmente de gas, y ácido sulfhídrico (con excepción de *S. arizonae* y *S. diarizonae*) (Farmer y col, 1985, Urías y Suárez, 2006). Sin embargo, *S. enterica sub. enterica* cuenta con serotipos que presenta excepción al perfil bioquímico típico, y corresponde a *S. Typhi*, *S. Cholerasuis*, *S. Paratyphi A*, *S. Gallinarum* y *S. Pullorum* (Farmer y col., 1985).

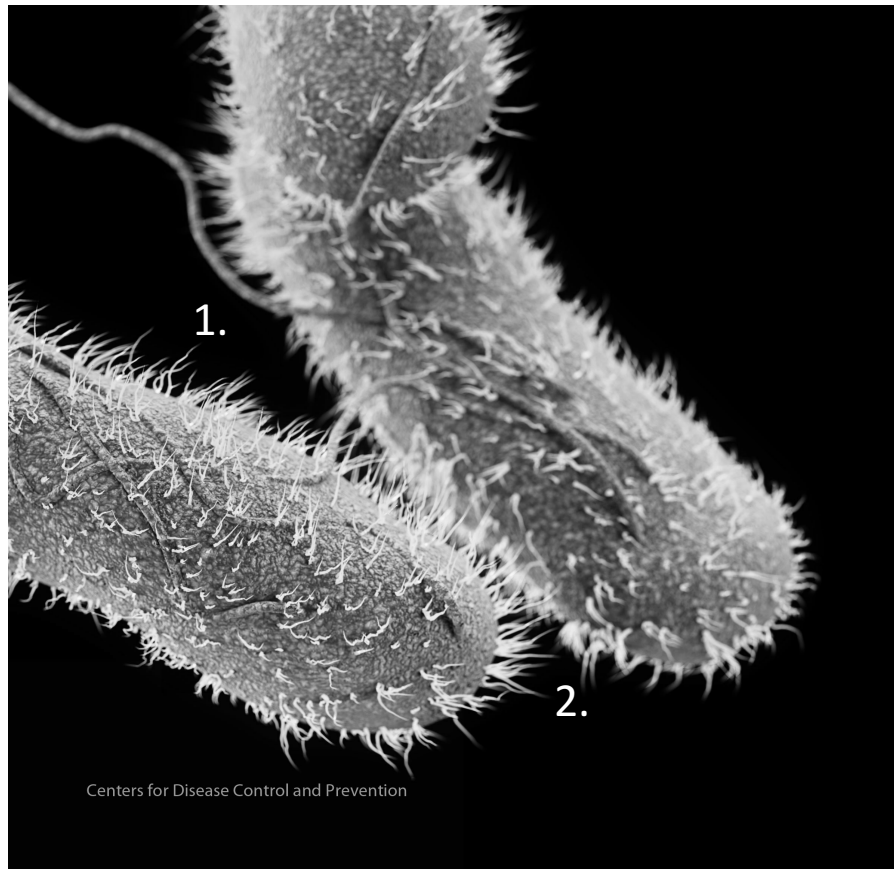


Fig. 1. Micrografía electrónica de barrido de *Salmonella* no tifoidea. donde se muestra su morfología. 1.- Flagelos peritricos, 2.- Pilis.

(Tomado de la página de la CDC, consultado en octubre del 2017).

<http://www.cdc.gov/media/subtopic/library/diseases.html>

Cuadro 1. Características bioquímicas diferenciales de especies de *Salmonella* y subespecies

Especie	<i>Salmonella enterica</i>						<i>Salmonella bongori</i>
	<i>enterica</i> (I)	<i>salamae</i> (II)	<i>arizonae</i> (IIIa)	<i>diarizonae</i> (IIIb)	<i>houtenae</i> (IV)	<i>indica</i> (VI)	
Dulcitol	+	+	-	-	-	D	+
ONPG (2hrs)	-	-	+	+	-	D	+
Malonato	-	+	+	+	-	-	-
Gelatina	-	+	+	+	+	+	-
Sorbitol	+	+	+	+	-/+	-	+
KCN	-	-	-	-	+	-	+
D-tartrato	+	-	-	-	-	-	-
B-glucoronidasa	D	D	-	+	-	D	-
Mucato	+	+	+	-0.7	-	+	+
Salicina	-	-	-	-	+	-	-
Lactosa	-	-	-0.75	0.75	-	D	-

González-Pedraza y col. (2010)

Antígenos de superficie

Salmonella spp. en su estructura externa tiene antígenos de superficie tales como el antígeno somático (O) y flagelar (H) que corresponden a los lipopolisacáridos (LPS) y moléculas de flagelina, respectivamente (Brenner y col., 2000). Mediante estos antígenos es posible clasificar los serotipos de *Salmonella* utilizando procedimientos de aglutinación (Brenner y col., 2000). Adicionalmente, se emplean antígenos de superficie (Vi) que es un polisacárido capsular que se encuentra en *Salmonella* Typhi (Barat y col., 2012, Santander y col., 2014).

El antígeno O es una parte del componente de LPS de la membrana externa de las bacterias Gram negativas, y es uno de los constituyentes celulares más variables (Reeves, 1993). Los LPS se componen de repeticiones de oligosacáridos (Unidades O), que contienen de dos a ocho residuos de azúcar. La variación es principalmente en los tipos de azúcares presentes, su orden en la estructura, y los vínculos entre ellos (Reeves y Wang, 2002.). El antígeno O está sujeto a una intensa selección por el sistema inmune del hospedero, bacteriófagos, y otros factores ambientales (Reeves y Wang, 2002), lo que puede explicar el mantenimiento de las diversas formas del antígeno O dentro de una especie. La diversidad del antígeno O es una base común en la serotipificación bacteriana, y es importante para las bacterias debido a que permite que cada uno de los diversos clones presenten una superficie que ofrece ventaja selectiva en su nicho específico (Reeves, 1992). Además, la presencia de antígeno O es esencial para la supervivencia de las bacterias en su ambiente natural y desempeña un papel en la virulencia bacteriana (Liu y col., 2012, Ovchinnikova y col., 2013).

El antígeno H se caracteriza por presentar dos fases: fase 1 y fase 2. Estas fases del antígeno son coordinadamente reguladas, de tal manera que sólo un antígeno flagelar se expresa a la vez en una sola célula (Medini y col., 2008). En raras ocasiones, los aislados de *Salmonella* expresan antígenos flagelares adicionales. Algunos de estos antígenos adicionales han sido inestables, los

cuales se han denominado fase 3 y R (Popoff y col., 2001).

Estudios recientes han encontrado que el antígeno Vi es responsable de la virulencia de *Salmonella* Typhi, el cual es diferente a otros serotipos de *Salmonella* (Hornick y col., 1970). El antígeno Vi es un homopolímero lineal formado por unidades de α -1,4-N acetilgalactosaminuronato (Daniels y col., 1989) y tiene una masa molecular >200 kDa (Zhang y col., 2006).

Taxonomía de *Salmonella*

El Centro de Referencia e Investigación de *Salmonella* de la Organización Mundial de la Salud y del Instituto Pasteur de Francia, elaboraron la nomenclatura de *Salmonella* spp. acorde con sus características bioquímicas y genéticas, la cual describe dos especies distintas: *Salmonella bongori* y *Salmonella enterica*. A su vez, *S. enterica* se divide en 6 subespecies (subgéneros), cada una de éstas antecedida por un número romano. Por lo tanto, *S. enterica* comprenden los siguientes subgéneros: I (*S. enterica* subsp *enterica*.), II (*S. enterica* subsp *salamae*.), III (antiguo género de Arizona; subdividido en IIIa, *S. enterica* subsp *arizonae* y IIIb, *S. enterica* subsp *diarizonae*), IV (*S. enterica* subsp *houtenae*), VI (*S. enterica* subsp. *indica*) y *S. bongori* corresponde al subgrupo (Brenner y col., 2000). El Cuadro 2 se muestran las especies, subespecies, subgrupo y el número de serotipos de *Salmonella* que están registrados. *S. enterica* subs. *enterica* comprende una gran variedad de serotipos (>1400 serotipos) y diversidad de nombres que se han utilizado para su nomenclatura.

Ciclo de vida de *Salmonella*

El principal hábitat de *Salmonella enterica* es el tracto gastrointestinal de animales de sangre caliente (Jiménez y col., 2011). La manera de que *Salmonella* sale del hospedero es por medio de la excreción de materia fecal; vía que le permite a *Salmonella* llegar al medio ambiente (agua, suelo, alimentos, entre otros), y donde emplea estrategias para sobrevivir en el nuevo entorno (Winfield

Cuadro 2. Clasificación de *Salmonella*

<i>Salmonella</i> especies y subespecies	Grupo	No. de Serotipos
<i>S. enterica</i> subespecie <i>enterica</i>	I	1,454
<i>S. enterica</i> subespecie <i>salamae</i>	II	489
<i>S. enterica</i> subespecie <i>arizonae</i>	IIIa	94
<i>S. enterica</i> subespecie <i>diarizonae</i>	IIIb	324
<i>S. enterica</i> subespecie <i>houtenae</i>	IV	70
<i>S. enterica</i> subespecie <i>indica</i>	VI	12
<i>S. bongori</i>	V	20
Total		2,463

Brenner y col. (2000)

y Groisman, 2003). *Salmonella enterica* subesp. *enterica* es una de las bacterias patógenas humanas más ubicuas en la naturaleza y con una gran diversidad de hospederos causando un amplio espectro enfermedades. Por ejemplo, *Salmonella* se encuentra como comensal y patógeno del tracto intestinal de humanos, mamíferos domésticos y salvajes, reptiles, aves, insectos y roedores (Perrón y col., 2008). El ciclo de vida completo de *Salmonella* se muestra en la Fig. 2.

Los seres humanos adquieren la infección por *Salmonella* después de la ingestión de alimentos o agua contaminados. Sin embargo, la infección puede transmitirse de persona a persona o por vía fecal oral directa (Foley y col., 2008). Cabe señalar, que los animales de granja (aves de corral, vacas, cerdos) son los principales reservorios de esta bacteria, y son vectores importantes para su transmisión y recirculación en el ambiente (Foley y col., 2008).

Los factores que favorecen a la sobrevivencia en ambientes externos al hospedero son la alta concentración de nutrientes en ecosistemas tropicales, el aire caliente y la temperatura constante del suelo y agua (Lynch y col., 2009). Algunas infecciones por *Salmonella* son producidas por el consumo de vegetales frescos que han sido contaminados por fertilizante con la bacteria. Esto nos indica que *Salmonella* posee una capacidad de adaptación y es capaz de colonizar gran variedad de organismos (Winfield y Groisman, 2003).

Cabe señalar, que la prevalencia de los serotipos de *Salmonella enterica* subesp. *enterica* identificados en el ambiente, raramente coincide con los serotipos aislados de los animales o de casos clínicos (Martinez-Urtaza y col., 2004). Esto plantea las interrogantes de las vías y estrategias que emplea cada uno de los serotipos de *Salmonella* para contaminar y sobrevivir en diferentes entornos fuera de los hospederos (Jiménez y col., 2014).

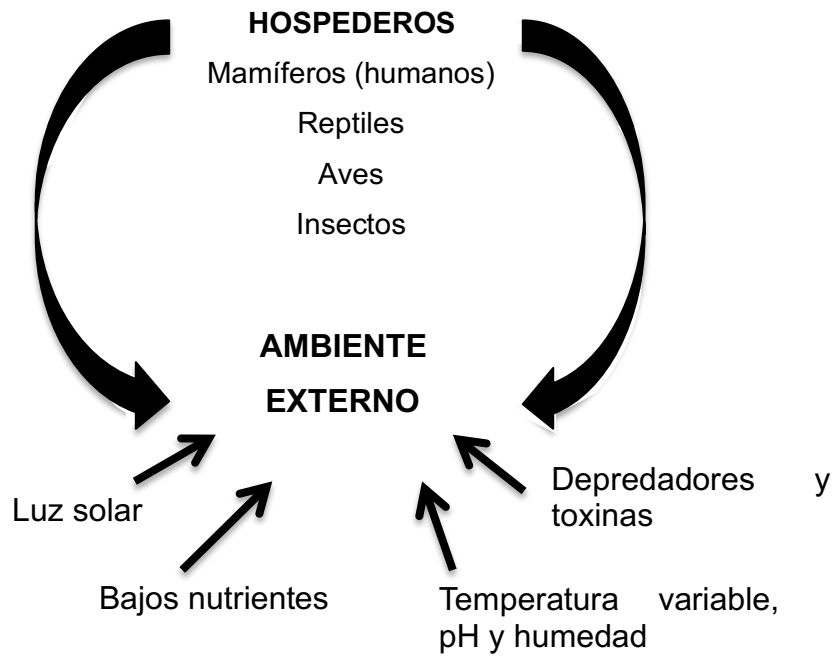


Fig. 2. Ciclo de vida de *Salmonella*

Winfield y Groisman (2003)

a. Suelo

La supervivencia de *Salmonella* spp. en el suelo está determinada por varios factores como: la temperatura, humedad, tipo de suelo, presencia de plantas, la exposición a los rayos del sol (UV), depredación por protozoarios número inicial de organismos presentes, entre otros (Fig. 2). Adicionalmente, se ha reportado que *Salmonella* spp. puede sobrevivir desde unos pocos días hasta 332 días (\approx 11 meses) en suelos con estiércol. (Holley y col., 2006; Islam y col., 2004 y You y col., 2006). El transporte de esta bacteria en el suelo ha sido reconocido como un proceso complejo debido a las interacciones de procesos tales como: flujo de agua, filtración y retención de células bacterianas en el suelo (Haznedaroglu y col., 2012, Jacobsen y col., 2011, Lyimo y col., 2016).

b. Cuerpos de agua

La transmisión de *Salmonella* spp. ocurre primeramente por el consumo de alimentos y agua contaminados, lo cual indica la facultad de la bacteria para permanecer en estas fuentes. En el agua *Salmonella* puede sobrevivir por largos periodos (>2 semanas) por lo que puede ser encontrada en muestras de ríos, lagos, tanques, pozos y grifos que han sido contaminadas principalmente por las actividades humanas (Haznedaroglu y col., 2012, Jacobsen y col., 2011, Lyimo y col., 2016). Estos cuerpos de agua se han identificado como fuente de una gran diversidad de serotipos, principalmente del tipo no tifoideo (Jiménez y col., 2014). Adicionalmente, los ecosistemas acuáticos se han asociado como ambientes que impactan las características morfológicas, bioquímicas y genéticas de esta bacteria (Baudar y col., 2000).

c. Alimentos

La principal ruta de infección de *Salmonella* spp. es el consumo de alimentos contaminados como son huevo, frutas y hortalizas, pollo, entre otros (CDC, 2017). El mayor énfasis está en el consumo de verduras crudas contaminados con *Salmonella*, dado que estos alimentos pueden consumirse sin cocción previa favoreciendo el riesgo de enfermedades (Guo y col., 2001; Cook

y col., 1998). Las principales causas de contaminación de los alimentos por *Salmonella enterica* se debe falta de higiene del personal por parte de los portadores asintomáticos, el inadecuado almacenamiento y manipulación de los alimentos, así como, la limpieza deficiente de los materiales e instalaciones utilizados para la preparación de los alimentos (Greig y col., 2010; Vázquez-Garcidueñas y col., 2014).

d. Hospederos

Salmonella spp. presenta una amplia diversidad de serotipos distribuidos en varios hospederos de sangre caliente, aves, peces, anfibios y reptiles. Sin embargo, los serotipos pueden representar especificidad por un hospedero. Al respecto, hay serotipos que son especie-específicos como: *S. Typhi*, *S. Paratyphi* A, B y C (solo infectan al hombre), *S. Gallinarum*, *S. Pollorum* en aves de corral, *S. Abortusovis* en ovejas. Estos serotipos específicos, pueden ocasionar padecimientos severos en los hospederos (Gonzales, 2000, Andino y col., 2014). Por otra parte, existe un grupo amplio de serotipos no tifoideos (SNT) que no tienen especificidad por algún hospedero o entorno (ecosistemas acuáticos, suelos, vegetación, etc.). Entre ellos destaca *S. Typhimurium*, *S. Enteritidis*, *S. Newport* y *S. Heidelberg*, ya que sus reservorios se han asociado a animales de granja principalmente (Andino y col., 2014).

B Patogenia de *Salmonella*

La estrategia patogénica de *S. enterica* subesp. *enterica* incluye la penetración de la barrera intestinal y la interacción con células del sistema inmune del hospedero, donde actúa como parásito intracelular (Marcus y col., 2000). En consecuencia, los mecanismos de patogenia de *S. enterica* son complejos, y el grado de virulencia puede ser distinto entre las cepas de esta especie (Marcus y col., 2000). De hecho, se estima que el 4% del genoma de *S. Typhimurium* LT2 representa 200 genes de virulencia, los cuales se requieren para la infección letal en modelos murinos (Bowe y col., 1998; McClelland y col.,

2001).

El ciclo de patogenia de *Salmonella enterica* subesp. *enterica* comprende dos grupos de factores de virulencia; estructuras superficiales y genes de virulencia localizados en el cromosoma o plásmidos. En las estructuras superficiales incluyen (i) LPS, con actividad tóxica, debido fundamentalmente al lípido A; (ii) los flagelos, que son estructurales que dirigen a la bacteria hacia el epitelio intestinal, mediante quimiotaxis, permiten atravesar la espesa capa de mucina, y contribuyen al proceso de inflamación; (iii) la cápsula, estructura directamente relacionada con la capacidad invasiva (solo *S. Typhi*); y (iv) fimbrias prolongaciones bacterianas que permiten la adherencia de la bacteria a la célula hospedera (Schmidt y Hensel, 2004). Los genes de virulencia localizados en el cromosoma o en plásmidos codifican factores solubles que modifican la fisiología celular del hospedero o que protegen a la bacteria de sistemas antimicrobianos del mismo. Estos genes pueden estar sueltos, y formando pequeñas agrupaciones (islotos) y/o en agrupaciones mayores (islas de patogenicidad) (Schmidt y Hensel, 2004).

Adicional a la capacidad patogénica de *Salmonella* para causar enfermedad en el hospedero, influye el estado inmunológico del hospedero, la dosis de infección, el serotipo implicado, del inóculo y la exposición del patógeno con este (Heithoff y col., 2008).

Ciclo patogénico

El ciclo patogénico de *Salmonella* comienza cuando la bacteria coloniza la superficie de las células intestinales que se hallan cubiertas por microvellosidades (células M) que incrementan la superficie de absorción de nutrientes y están cercanas a las placas de Peyer (Gebert, 1997; Lahiri y col., 2010). *Salmonella* tiene la capacidad para invadir células hospederas, utilizando el mecanismo sistema secreción Tipo III (Lahiri y col. 2010). para atravesar la membrana plasmática e inyectar proteínas al citoplasma del hospedero. Las

proteínas inyectadas inducen a la membrana plasmática a envolver a la bacteria, y reacomodan el citoesqueleto de la célula infectada. El proceso de invasión de *Salmonella* termina cuando la bacteria es englobada en una vacuola fabricada por la célula, la cual es arrastrada hacia el interior por filamentos de actina (Lahiri y col., 2010).

En circunstancias normales la célula hospedera destruye a la bacteria, es decir, la vacuola se fusiona con un lisosoma que contiene enzimas digestivas que degradan al patógeno introducido. No obstante, la bacteria puede utilizar el sistema de secreción tipo III para inyectar otras proteínas bacterianas en la vacuola que la rodea, y modificarla para evitar la fusión de la vacuola con los lisosomas de la célula hospedera (Lahirir y col., 2010).

Una vez a salvo, *Salmonella* comienza a dividirse por fisión binaria en la vacuola dentro de la célula hospedera y proyectan filamentos de la vacuola que se adhieren a la membrana celular hospedera. La formación de esos filamentos esta motorizada por las bacterias y, a medida que se dividen la vacuola crece favoreciendo la invasión y diseminación de la bacteria (Drecktarh y col., 2008; Ibarra y Steele, 2009).

Factores de virulencia e islas de patogenicidad

El género *Salmonella* tiene factores que contribuyen a la virulencia de este patógeno, unos de los factores más estudiados son las proteínas efectoras secretadas que dependen de los sistemas de secreción de tipo III codificadas por *Salmonella* y las islas de patogenicidad (Ruby y col., 2012). *Salmonella* utiliza un sistema de genes de virulencia denominados islas de patogenicidad para ayudar a la bacteria a invadir, replicarse y difundirse dentro del ambiente del hospedero (Marcus y col., 2000). Actualmente, se conocen 12 islas de patogenicidad y solo se han descrito las funciones de 7 islas de patogenicidad. En el Cuadro 3 se describen las composiciones y los funcionamientos de las islas de patogenicidad descritas para *Salmonella* (Lahiri y col., 2010; Ibarra y Steele, 2009). La virulencia

de las islas de patogenicidad en el desarrollo de la infección en el hospedero sigue siendo incierto (Ruby y col., 2012).

El éxito de la infección por *Salmonella* depende de su capacidad para sobrevivir dentro de las células de mamíferos (macrófagos y células epiteliales no-fagocíticas) (Lahiri y col., 2010). *Salmonella* invade las células epiteliales no-fagocíticas mediante la ondulación de la membrana plasmática, y se internaliza en un endosoma temprano mediado por la vía endocítica de factores EEA1, *Rab5* y receptor de transferrina en un lapso de 15-60 min; la bacteria se encuentra dentro de un fagosoma denominado “*Salmonella* contenida en vacuolas” (SCV) (Vazquez-Torres y col., 2000). Las proteínas del endosoma temprano son reemplazadas por las proteínas LAMPs (proteínas de membrana asociadas a lisosoma) y un tipo de ATPasa vacuolar (Lahiri y col., 2010).

Una vez que las SCV se asocian con el alargamiento de túbulos llamados filamentos inducidos por *Salmonella* (FIS) en las células epiteliales, las SCV logran una posición juxtanclear (Steele-Mortimer, 2008). Por último, al estar las SCV dentro de las células del sistema inmune hospedero puede hacer frente a las especies de oxígeno reactivo (EOR) y las especies de nitrógeno reactivo (ENR) que las células del sistema inmune hospedero produce para controlar al patógeno mediante una serie de factores que utiliza *Salmonella* para que la infección persista dentro del hospedero (Chevaril y Antos, 2001). Las condiciones elevadas de sal y de bajo oxígeno que prevalecen en el medio intestinal, activan la expresión de genes *SPI-1* permitiendo la invasión. Una vez que la bacteria está como parásito intracelular, el entorno intra-fagosomal, incluyendo el pH ácido y las concentraciones limitantes de Ca^{2+} , Mg^{2+} y fosfato inorgánico, inhiben la expresión de genes de las islas *SPI-1* y *SPI-2*, y activa la expresión de otros genes o islas de patogenicidad (Fig. 3) (Lahiri y col., 2010).

Cuadro 3. Siete islas de patogenicidad conocidas

Isla de patogenicidad	Tamaño aprox. (Kb)	Tipo de sistema de secreción	Función
SPI-1	40	Sistema de secreción tipo III (T3SS)	Invasión de células no fagocíticas
SPI-2	40		Sobrevivencia bacteriana en macrófagos
SPI-3	17	Sistema de secreción tipo I (T1SS)	Sobrevivencia en macrófagos
SPI-4	27		Media la adhesión a células epiteliales y sobrevivencia en macrófagos
SPI-5	8	-	Enteropatogénesis
SPI-7	134	-	Invasión de células epiteliales

Hensel (2004)

Finalmente, *Salmonella* cuenta con una gran diversidad de genes que codifican para moléculas necesarias en diferentes fases del proceso patogénico y a su vez forman parte de los llamados “islotos” (Groisman, 1998). Principalmente, entre ellos se encuentran: *sifA* (implicado en la multiplicación en macrófagos), *pagC* y *msgA* (intervienen en la supervivencia en el interior de los fagosomas), *iviVI* (relacionado con la adhesión e invasión de los enterocitos) y algunos genes de fimbrias (Groisman, 1998). Otros genes se encuentran de manera aislada, entre ellos los que codifican toxinas (*stn*, *ast*), sideróforos implicados en la captación de hierro (*iroB*), resistencia al estrés oxidativo (*sodC1*), reguladores transcripcionales de distintos genes de virulencia (*phoP/Q*, *slyA*), etc. (Way y col., 1994, Guerra y col., 2000).

Por otra parte, la colonización sostenida del tracto gastrointestinal es otro aspecto importante de la infección persistente por *Salmonella*, ya que es necesario para que las bacterias vivas que se pasan en las heces se transmitan finalmente a otros potenciales hospederos (Lawley y col., 2008). Al respecto, se han descrito varios factores relacionados con la adherencia de *Salmonella* como los factores de unión a fibronectina *ShdA* (Kingsley y col., 2000, 2003, 2004) y *MisL* (Dorsey y col., 2005), que contribuyen de manera significativa a la persistencia en el tracto gastrointestinal y la excreción fecal de bacterias vivas. Adicionalmente, se han descrito factores en *Salmonella* Typhimurium que le permiten a la bacteria protegerse contra péptidos antimicrobianos derivados del hospedero, y consecuentemente contribuyan a la infección persistente por este patógeno (Dorsey y col., 2005, Kingsley y col., 2000, 2003, 2004).

Los genes de la isla *SPI-2* codifican un segundo sistema de secreción tipo tres (SST3), utilizado para inyectar un segundo conjunto de proteínas efectoras en el citoplasma de la célula hospedera, y modificar el SCV como un lugar propicio para la replicación bacteriana (Lahiri y col., 2010). El gen *sodC1* (superóxido dismutasa) de la *SPI-2* codifica una enzima requerida para la

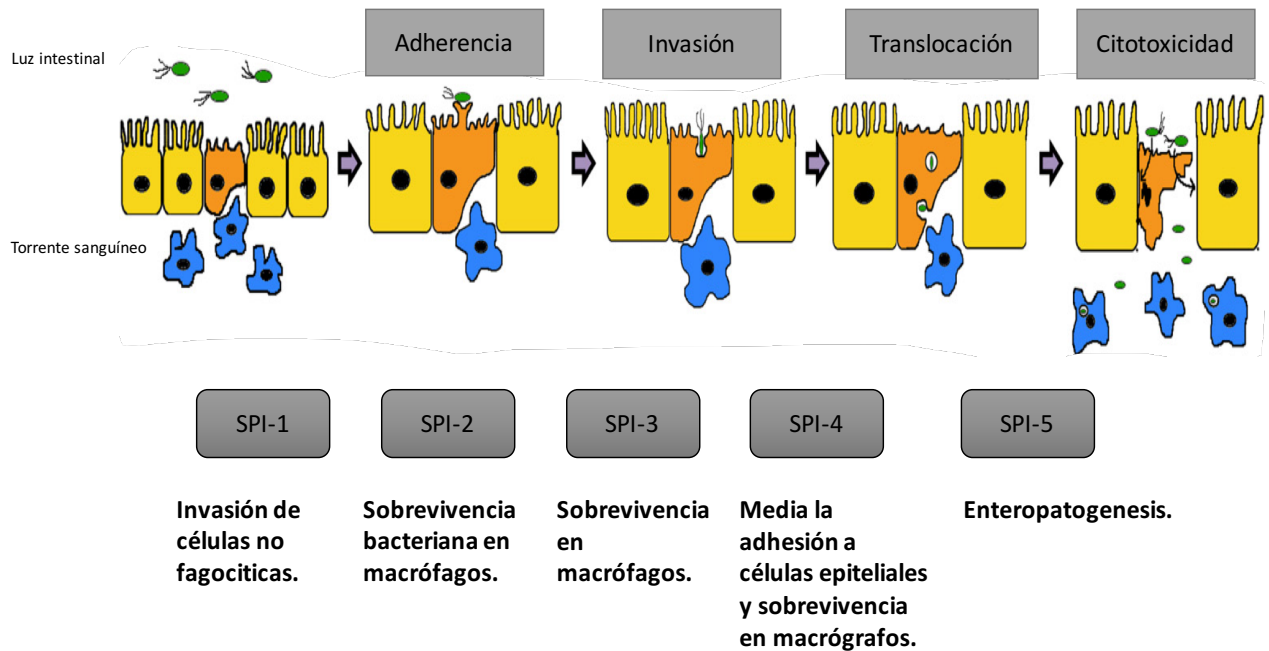


Fig 3. Ciclo de patogenicidad. *Salmonella* tiene preferencia por las células M del intestino (próximas a las placas de Peyer), utiliza 4 fases (adherencia, invasión, translocación y citotoxicidad) para poder invadir al hospedero y así poderse replicar para completar su ciclo

Lahiri y col. (2010)

desintoxicación del anión superóxido y es necesario en la virulencia *in vivo* (Lahiri y col., 2010, Jantsch y col., 2011, Ruby y col., 2012). De acuerdo a lo reportado por Elemfareji y col. (2013), la virulotipificación es un método de cribado para identificar genes, el perfil de virulencia en cepas de *Salmonella*, y mejorar la comprensión de potencial riesgo de infecciones humanas y animales. Entre los genes que se han empleado destaca: *invA*, *perfA*, *spvC*, *sirA*, *gipA*, SEN1417, *trhH* y *prot6E* (Retamal y col., 2015). Todos estos genes se han asociado con virulencia que han sido variablemente detectado en *S. Enteritidis* mediante estudios de reacción en cadena de la polimerasa (PCR, siglas en inglés) (Pan y col. 2009, Huehn y col. 2010).

Genes de resistencia

La resistencia a antibióticos ligada a los patógenos transmitidos por los alimentos está relacionado al uso indiscriminado de antibióticos en animales principalmente. La diseminación de la resistencia a antibióticos es a menudo dada por elementos genéticos móviles tales como plásmidos, transposones y casetes genéticos en integrones. Los integrones más comunes involucrados en la resistencia son los de clase 1 los cuales son abundantes en los genomas de muchas bacterias (Małka y Popowska, 2016). *Salmonella* spp. se ha caracterizado por el incremento de cepas resistentes a antibióticos como ácido nalidixico, tetraciclina, penicilina; antibióticos considerados de primera elección para el tratamiento de padecimientos severos de *Salmonella* (Tsai y col., 2010; Cheng y col., 2013).

C Epidemiología de la SNT

Salmonella es el agente causal de dos padecimientos: salmonelosis no tifoidea (SNT) y fiebre tifoidea, principalmente adquiridas por el consumo de alimentos o la ruta fecal-oral, respectivamente (Adino y col., 2015) La Organización Mundial de la Salud, declara a *Salmonella* como un agente

etiológico de notificación obligatoria, y estima la ocurrencia anual de 1.3 de millones de casos de SNT y 16 millones de casos de fiebre tifoidea, los cuales resultan en >3 millones de muertes a nivel mundial (OMS, 2016).

Las salmonelosis transmitidas por los alimentos pueden ser causada por >2,400 serotipos, y comúnmente asociada al consumo de frutas y vegetales frescos contaminados por los serotipos de esta bacteria. En años recientes la incidencia de salmonelosis asociada a frutas y hortalizas frescas ha incrementado, y se ha convertido en una gran preocupación para los países industrializados (Pui y col., 2011). Mientras los casos de fiebre tifoidea se mantienen estables con números bajos en los países industrializados, el número de casos de SNT ha incrementado alrededor del mundo. Al respecto, en los Estados Unidos de América, el Centro Nacional para la Prevención y Control de Enfermedades (CDC, siglas en inglés) reporta que ocurren de 2 a 4 millones de casos de SNT, generando alrededor de 500 muertes al año (WHO, 2015; CDC, 2017; Scallan 2011). Sin embargo, en países subdesarrollados de Asia, África, Sudamérica y América central, la prevalencia de la SNT y febril no está claramente notificada (Portillo, 2000; Hanes, 2003; Hu y Kopecko, 2003).

Tipos de Salmonelosis

La salmonelosis es una de las enfermedades infecciosas más frecuentes a nivel mundial que afecta tanto al hombre como animales, y es causada por *Salmonella* (*S. enterica* y *S. bongori*) (Pignato y col., 1998). Aunque fundamentalmente son bacterias intestinales, *Salmonella* está distribuida en el ambiente, lo cual le permite mantener un ciclo dentro y fuera del hospedero (Winfield y Groisman, 2003). *Salmonella* spp. es considerada una bacteria zoonótica responsable de infecciones intestinales y sistémicas en el hombre, como consecuencia del consumo de alimentos contaminados o la ruta fecal-oral (Andino y Hanning, 2015; Abulreesh, 2012). Los agentes de *Salmonella* se pueden encontrar originando el transporte gastrointestinal asintomático o

enfermedades infecciosas en los animales, particularmente en aves y en cerdos. La salmonelosis comprende: dos cuadros clínicos a) Fiebre tifoidea y b) Gastroenteritis (salmonelosis no tifoidea, SNT) (Andino y Hanning, 2015).

a. Fiebre tifoidea

La fiebre tifoidea es una enfermedad bacteriana causada por *Salmonella* Typhi y en raras ocasiones por *S. Paratyphi* A, B y *S. Paratyphica* C, ya que pueden producir un cuadro clínico similar, pero de menor gravedad. La fiebre tifoidea se transmite por la ingestión de alimentos o bebidas contaminadas por heces u orina de personas o animales infectados con *Salmonella* (OMS, 2015). Los síntomas suelen aparecer entre 1 y 3 semanas después de la exposición con la bacteria y pueden ser leves o graves como son fiebre, malestar general, cefaleas, estreñimiento o diarrea, manchas rosadas en el tórax, hepatomegalia y esplenomegalia. Tras la enfermedad aguda se puede instaurar un estado de portador crónico y es aquí donde se puede transmitir la enfermedad (OMS, 2015). Este tipo de salmonelosis requiere el uso de antibióticos como ácido nalidíxico, tetraciclina, penicilina para el control de la enfermedad (Tsai y col., 2010; Cheng y col., 2013).

b. Salmonelosis no tifoidea (SNT)

La salmonelosis no tifoidea (SNT) o gastroenteritis es causada >2,400 serotipos de *Salmonella* (excepto *S. Typhi*, *S. Paratyphi* A, B y C), los cuales se transmiten a través de alimentos contaminados (Cheng y col., 2013). Entre los serotipos frecuentemente asociados a las SNT están *S. Tiphymurium*, y *S. Enteritidis*. Sin embargo, los síntomas de la SNT son fiebre alta, dolor abdominal, diarrea, náusea, y vómitos. La enfermedad comienza a manifestarse entre 6 y 72 horas después de la ingesta de *Salmonella*, y la enfermedad es autolimitante con un periodo de 2 y 7 días (OMS, 2016). En casos severos, la SNT puede resultar en septicemia e infecciones localizadas, lo cual requiere administración de sueros electrolíticos y antibióticos (Cheng y col., 2013).

c. Estado del portador

Los portadores convalecientes o crónico asintomático refieren a los hospederos que eliminan a *Salmonella* por las heces por periodos persistentes de tiempo (generalmente semanas) (Jones y Falkow, 1996), o períodos superiores a un año, hecho que ocurre entre el 1% y el 3% de los pacientes con fiebre tifoidea, y pueden llegar a excretar de 10^6 - 10^9 UFC/g de heces (Miller y Pegues, 2000). Adicionalmente, se han identificado a *S. Enteritidis* y *S. Typhimurium* como serotipos que se aíslan en portadores crónicos asintomáticos el 1.0% - 5.4% de la población infectada (Jones y Falkow, 1996; Miller y Pegues, 2000).

Casos y brotes

A nivel mundial, La Organización Mundial de la Salud (OMS) estimó la ocurrencia de 1.3 mil millones de casos de SNT y 16 millones de casos de fiebre tifoidea en 2015, las cuales resultan en conjunto en 3 millones de muertes al año. La mayoría de los casos de salmonelosis se asocian al consumo de alimentos contaminados (Scallan y col., 2011). Por lo tanto, *Salmonella* representa un serio problema de salud pública e industria alimentaria. En EUA, la CDC reportó 88,715 casos de SNT por consumo de alimentos en 2014, el 34.5% y % resultaron en hospitalizaciones y defunciones respectivamente (EFSA y ECDC, 2015). La CDC vincula la mayoría de estos casos de SNT como brotes epidemiológicos por el consumo de alimentos contaminados principalmente hortalizas, y en los que destacan como serotipos responsables *S. Poona*, *S. Hiedelberg*, *S. Saintpaul* y *S. Braenderup* (CDC, 2015). El Cuadro 4 describe algunos de los principales brotes de SNT de origen alimentario y el serotipo responsable. Así mismo, el Cuadro 4 señala la participación de México en los brotes de SNT.

Cuadro 4. Brotes de enfermedades transmitidas por los alimentos por *Salmonella*

Año	País de Publicación	Alimento implicado	País de procedencia	Casos/ Muertes	Serotipo
2004	EUA	Tomate	EUA	32/0	S. Braenderup
2007	Holanda	Carnes rojas	Dinamarca	56/0	S. Typhimurium
2008	EUA	Chile y tomate	México	1442/0	S. Saintpaul
2012	EUA	Mango	México	127/0	S. Braenderup
2013	Alemania	Puerco	Alemania	267/1	S. Infantis
2014	EUA	Pollo	EUA	23/0	S. Heidelberg
2015	EUA	Pepino	México	328/4	S. Poona

CDC (2015)

D Panorama de *S. Oranienburg* en Sinaloa

Fuentes

En Sinaloa, *Salmonella* juega un papel importante en la contaminación de ecosistemas acuáticos y colonizando animales domésticos de sangre caliente. Al respecto, *Salmonella* spp. se ha identificado en 31.7% de muestras de heces de vacas, pollos y ovejas (Jiménez y col. en 2011), 80.4% de muestras de agua de río (Jiménez y col., 2014), 39.2% muestras de agua de canal (López-Cuevas y col. 2009), 0% de muestras de sedimento (López-Cuevas y col. 2009) y 8.6% de muestras de cultivos de tomate (Estrada-Acosta y col. 2009); *S. Oranienburg*, *S. Give*, *S. Saintpaul*, *S. Weltevedren*, *S. Minnesota*, *S. Newport* se han identificado mediante serología tradicional como los principales serotipos aislados de heces de animales y agua de río (Jiménez y col., 2011; Jiménez y col., 2014). La coincidencia de serotipo de SNT de origen fecal y acuático sugiere la recirculación de *Salmonella* entre el hospedero y el ambiente. No obstante, esta relación clonal no está esclarecida.

Casos y brotes

En México, la Dirección General de Epidemiología (DGE) declara a la SNT y fiebre tifoidea como enfermedades de notificación obligatoria desde 1986, as cuales tienen una incidencia de un 62.7 y 38.4 por cada 100,000 habitantes a nivel nacional (DGE, 2016). La DGE en su anuario de morbilidad estimó que ocurrieron 75,855 casos de SNT y 46,503 casos fiebre tifoidea en 2015 a nivel nacional (DGE, 2016). Sin embargo, la DGE no declara los serotipos responsables de estos casos, la vía de infección, o la posible asociación entre los casos reportados.

Sinaloa es considerado un estado endémico para *Salmonella*, dada la incidencia de casos de SNT y la prevalencia de *Salmonella* en el ambiente,

alimentos y animales domésticos. Al respecto, la DGE reporta una incidencia anual de SNT por cada 100,000 habitantes en Sinaloa.

Respecto a casos o brotes de origen alimentario asociados a *S. Oranienburg*, Vázquez-Garcidueñas y col. (2014), reportaron un brote de 150 casos por el consumo de alimentos contaminados con este serotipo en una prisión de México. Cabe señalar, que ciertos brotes de SNT reportados en los EUA se han asociado al consumo de frutas y hortalizas que han sido irrigadas con aguas contaminadas procedentes de México (CDC, 2013; CDC, 2008). Al respecto, la CDC declaró un brote de 328 casos de SNT y 4 personas fallecidas por el consumo de pepino exportado de México contaminado con *S. Poona* en los EUA; la fuente de contaminación del alimento fue el agua de irrigación empleada durante el cultivo.

Por otra parte, Gutierrez-Gogco y col. (2002) reportan *S. Typhimurium*, *S. Enteritidis*, *S. Derby*, *S. Agona* y *S. Anatum* como los serotipos responsables de los casos de SNT en México durante el periodo de 1972-1999 exponiendo la diversidad de serotipos de SNT circulantes en México. Estos hechos plantean la posibilidad de diversas vías de infección dado las fuentes de origen reportadas para los serotipos SNT en México (Jiménez y col., 2011; Jiménez y col., 2014).

Caracterización molecular

Los ríos de Sinaloa, se han identificado como reservorios de *Salmonella* Oranienburg, debido a la persistencia de este serotipo en dicho ecosistema (aproximadamente un año). Por ello, Jiménez y col. (2014), determinaron el perfil genético del serotipo mediante el análisis de electroforesis en gel de campos pulsados, identificando una heterogeneidad genética alta (Jiménez y col., 2014). Este panorama sustenta que *S. Oranienburg* ha desarrollado estrategias de sobrevivencia en el ambiente, y el riesgo de representar un problema de salud pública en la región. Por lo tanto, se requiere emplear herramientas de epidemiología molecular que permitan el estudio de *S. Oranienburg*.

E Epidemiología molecular de microorganismos

Definición y aplicación

La epidemiología molecular es una de las ramas de la biología molecular la cual pretende establecer patrones epidemiológicos más precisos utilizando estrategias basadas en la caracterización del ADN proveniente de organismos patógenos, el uso de estas estrategias en la monitorización epidemiológica permite establecer la progresión de especies patógenas hacia la adquisición de resistencia a los agentes terapéuticos disponibles por variaciones genotípicas o mutaciones (Vílchez y Alonso, 2009).

Además, la epidemiología molecular permite determinar con mayor certeza y rapidez las alternativas terapéuticas efectivas en el control de las enfermedades infecciosas y en especial cuando se refiere a infecciones nosocomiales (Lee, 2004). Cabe destacar que la epidemiología molecular no se limita a las enfermedades infecciosas, si no que también es posible usar estas mismas estrategias para el estudio epidemiológico de enfermedades genéticas o que posean un componente genético (Lee, 2004). Por lo que en el presente trabajo, se pretende realizar un análisis de los genes de *Salmonella* Oranienburg para tratar de establecer la relación y las estrategias que emplea el patógeno para colonizar el medio ambiente y sus posibles hospederos.

Métodos de tipificación molecular

Los métodos de tipificación molecular se utilizan para evaluar la relación genómica entre especies del mismo o diferentes géneros, para los cuales hay diferentes técnicas como lo son el método de reacción en cadena de la polimerasa (PCR, siglas en inglés), la electroforesis en gel de campo pulsado (PFGE, siglas en inglés), tipificación por multilocus de secuencia (MLST, siglas en inglés), análisis de secuencia de nucleótidos, secuenciación parcial o completa de genomas entre otros (Harbottle y col., 2006). Una ventaja relevante

de estos métodos es su análisis bioinformático que permite realizar análisis *in-silico* como la selección de marcadores moleculares, diseño de estrategias de tipificación, almacenamiento, intercambio y comparación de datos genotípicos (patrones de bandeo de ADN, secuencias de ADN y perfiles de microarreglos); el desarrollo y mejoramiento de herramientas de análisis, y análisis filogenéticos (Li y col., 2009).

En la Fig. 4 se muestra un esquema comparativo de los diferentes niveles de resolución dentro de una población bacteriana obtenida por varios métodos de tipificación. Donde se nota una tendencia actual a la identificación de bacterias por medio (Medini y col, 2008). Por lo tanto, se señala que la secuenciación de genomas completos, se utiliza cada vez con mayor frecuencia para la identificación y caracterización de bacterias patógenas.

Electroforesis en gel de Campo pulsado (PFGE)

La PFGE es un método establecido para el análisis de grandes fragmentos generados por digestión con endonucleasas de restricción de ADN genómico y se considera actualmente como uno de los procedimientos de tipificación más fiables para bacterias debido al número discreto de fragmentos de la banda bien resueltas permite una comparación visual de los perfiles de restricción (Garaizar y col., 2000). Además, la tipeabilidad adecuada y poder discriminatorio de este método han llevado a la conclusión de que PFGE es una de las herramientas más valiosas para estudios epidemiológicos moleculares de microorganismos (). Estudios realizados por Garaizar y col. (2000), Harbottle y col. (2009), Kubota y col. (2005). Han demostrado ser altamente eficaz para los estudios epidemiológicos de algunos serotipos de *S. enterica*. La limitación de esta técnica es su dificultad para hacer reproducibles los resultados entre laboratorios (Nallapareddy et al., 2002; Johnson et al., 2007).

Tipificación Multilocus de Secuencia (MLST)

El método de MLST permite la caracterización inequívoca de los aislados de agentes infecciosos utilizando secuencias de fragmentos internos de siete genes de mantenimiento o “*housekeeping*” (genes constitutivos necesarios para el mantenimiento de las funciones celulares básicas) (Atchman y col., 2012). El MLST se fundamenta en la caracterización de los alelos en cada uno de los siete *locis*, que constituyen su perfil alélica o tipo de secuencia (ST) (Atchman y col., 2012). El enfoque MLST proporciona una evaluación precisa inter e intra especies (Lewis-Rogers y col. 2009; Crandall y Pérez-Losada, 2008).

Las limitaciones del MLST en la caracterización de bacterias consiste en (i) la presencia de genes constitutivos con poca o nula variación en la(s) secuencia(s), lo cual dificulta proporcionar una discriminación suficiente debido a la diversidad nucleotídica (por ejemplo, *Salmonella Typhi* y *Bacillus anthracis*) (Vazqu ez y col., 2004; Medini y col., 2008), (ii) la baja cobertura sobre el genoma completo, es decir, solo se secuencian peque as fracciones del genoma, lo cual no permite conocer la diversidad nucleot dica del genoma completo (V zquez y col., 2004; Urwin y Maiden, 2003)

Secuenciaci n de nueva generaci n (NGS)

a. Historia de la secuenciaci n de nueva generaci n

Las primeras t cnicas de secuenciaci n fueron descritas en 1977 por Allan Maxam, Walter Gilbert, y Frederick Sanger. El m todo de secuenciaci n de Allan Maxam, Walter Gilbert se basa en la degradaci n qu mica de la cadena de ADN marcada con los extremos 5' y 3' con P³², para su posterior fragmentaci n y separaci n de fragmentos por medio de electroforesis (Maxam y Gilbert, 1977). Mientras que el m todo de Sanger, es un m todo automatizado que se basa en el uso de la ADN-polimerasa para incluir aleatoriamente nucle tidos entre los componentes de la reacci n de nucle tidos carentes de un grupo hidroxilo en su

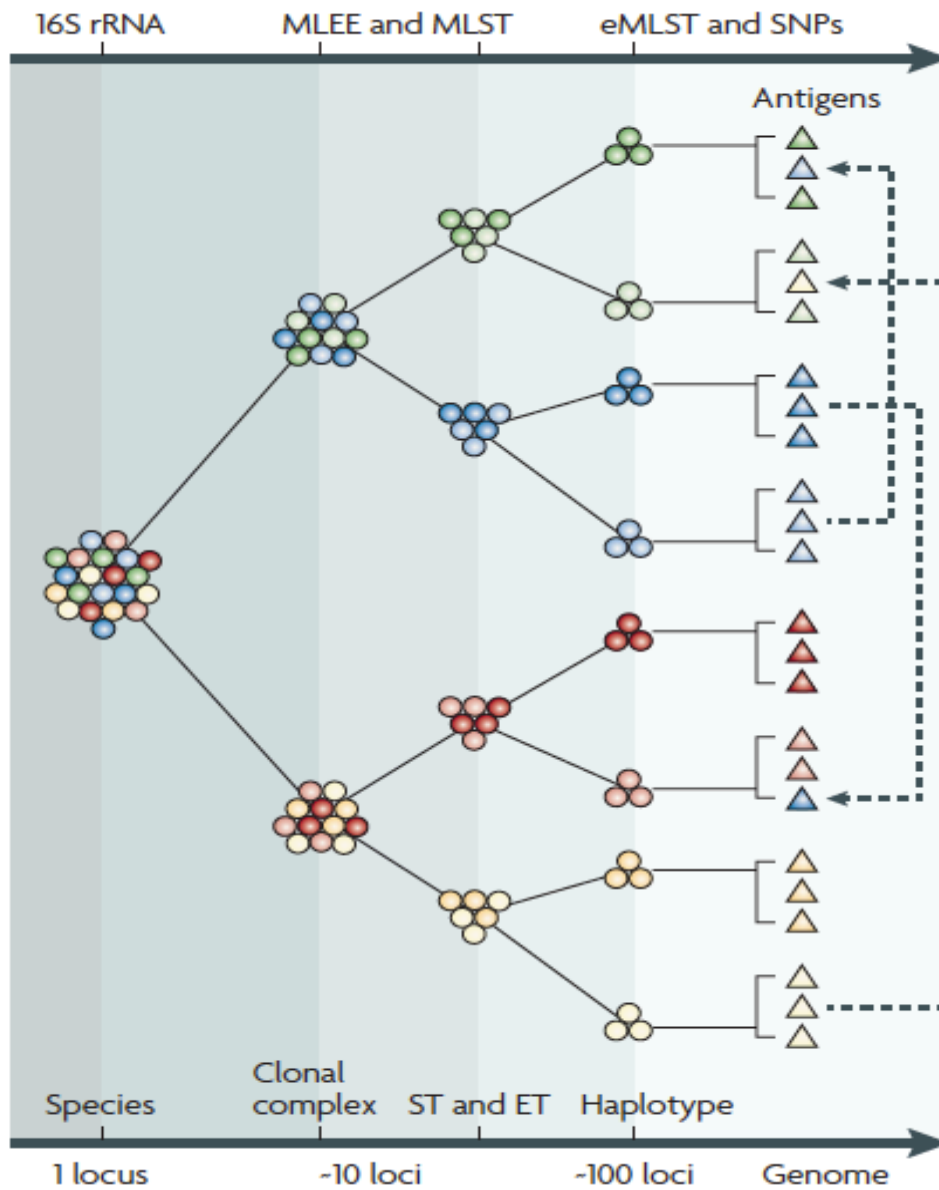


Fig 4. Representación esquemática de los diferentes niveles de resolución para la identificación bacteriana.

Medini y col. (2008)

extremo 3' (ddNTP) y obtener una terminación específica en las cadenas y evita que la cadena de ADN sintetizada continúe extendiéndose. El método de Sanger tuvo muchos logros como la secuenciación del genoma completo del ser humano. Sin embargo, las limitaciones de este método llevaron a investigaciones para buscar nuevas y mejores tecnologías para la secuenciación de un gran número de genomas (Barba y col., 2014).

A inicios del siglo XXI, se realizaron investigaciones para sustituir al método automatizado de Sanger, el cual se consideró como una tecnología de “primera generación” (Barba y col, 2014, Liu y col. 2012). Actualmente, la secuenciación masiva o tecnología de segunda generación o nueva generación o NGS (por sus siglas en inglés Next Generation Sequencing) es aplicada para la secuenciación de genomas (Barba y col, 2014, Rodríguez-Santiago y Armengol, 2012). En 1990, Balasubramanian y Klenerman definieron la NGS basada en el uso de nucleótidos marcados con fluorescencia para observar el movimiento de una polimerasa que sintetizaban ADN inmovilizado en una superficie. Actualmente, la empresa Illumina, Roche y Life Technologies cuentan con las plataformas que han revolucionado la secuenciación de ADN por su capacidad de producir nucleótidos en un tiempo corto (Rodríguez-Santiago y Armengol, 2012). En el Cuadro 5 se muestran algunas de las características de las plataformas de segunda generación NGS (Rodríguez-Santiago y Armengol, 2012).

Actualmente, se están desarrollando tecnologías de tercera generación las cuales pretenden secuenciar una sola molécula de ADN muy larga, sin necesidad de realizar una amplificación clonal, como lo realizan los equipos de segunda generación. Uno de los equipos considerado como un secuenciador de tercera generación es el PacBio de Pacific Bioscience (Rodríguez-Santiago y Armengol, 2012).

Cuadro 5. Principales Características de las plataformas de Secuenciación de Nueva Generación

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

Plataforma de secuenciación	Método de amplificación	Longitud de lectura (pb)	Tiempo de Corrida (h)	Ventajas	Desventajas
454 Roche	Emulsión	700	23	Lecturas largas	Errores homopolímeros. Caro
illumina MiSeq	Puentes de PCR	150	27	Fácil uso	Caro por base
Ion Torrent (Life Technologies)	Emulsión PCR	>100	2	Tiempo corto de corrida y bajo costo de reactivos	Rendimiento medio
PacBio (Pacific Biosciences)	No hay amplificación	>1000	1.5	Lecturas largas, tiempo corto de corrida, secuencia de una sola molécula	Tasa de error alta. Rendimiento bajo

b. Plataforma de secuenciación masiva illumina

La plataforma de secuenciación illumina MiSeq utiliza una tecnología basada en un terminador reversible marcado con fluorescencia que forma la imagen medida que se añade cada dNTP, y luego escindido para permitir la incorporación de la siguiente base. El illumina MiSeq utiliza una “Flow cell” de vidrio la cual utiliza para tener el proceso de clusterización y de esta manera realizar de manera paralela y masiva la secuenciación (illumina, EUA) (Barba y col., 2014; Head y col., 2014).

La celda de flujo contiene miles de cadenas de ADN de cada biblioteca formada previamente, en la celda se lleva a cabo la secuenciación por síntesis (SBS) donde se coloca en un casete de fluidos dentro del secuenciador, cada grupo o clúster se suministra con polimerasa y cuatro nucleótidos con etiquetados fluorescentes que tienen sus 3'-OH inactivado químicamente para asegurar que sólo una única base está constituida por ciclo. Cada ciclo de incorporación de base es seguido por una etapa de formación de imágenes para identificar el nucleótido incorporado en cada grupo y de una etapa de producto químico que elimina el grupo fluorescente y desbloquea el extremo 3' para el ciclo de incorporación base siguiente (Mardis, 2008).

El flujo de trabajo de la secuenciación en illumina MiSeq consta de tres etapas; (i) preparación de librerías o genotecas que se produce mediante la fragmentación del ADN, (ii) ligación de adaptadores a estos fragmentos para iniciar la secuenciación, (iii) conversión de los fragmentos ligados a los adaptadores a genotecas que posteriormente son amplificadas y (iv) secuenciación por síntesis. Es decir, las genotecas se adhieren a la superficie de la celda de flujo estas para su amplificación por puentes donde forman clúster para comenzar la detección por fluorescencia de nucleótidos (Head y col., 2014; Kozarewa y col., 2009; Quail y col., 2008).

c. Tipos de bibliotecas o Genotecas en la Secuenciación Masiva

Existen tres tipos de genotecas de ADN que se pueden generar para la secuenciación de ADN. En la Fig. 5 se muestra los diferentes tipos de genotecas.

(i) Genoteca tipo “single-end”, en donde por cada fragmento de ADN se realiza una sola lectura (unidireccional). Este tipo de genotecas son ideales para el análisis de sustituciones de un solo nucleótido o variaciones, y se requiere una cobertura mayor a 30x (Head y col. 2014).

(ii) Genoteca tipo “paired-end”, que realiza las lecturas en ambas direcciones: sentido y antisentido de un fragmento de ADN o genoteca, en el cual se obtienen dos lecturas por cada segmento, lo que facilita la detección de reordenamientos genómicos, elementos repetitivos, así como la fusión de genes (Head y col., 2014).

(iii) Genoteca tipo “mate-pair”, que es construida por fragmentos genómicos de ADN mayores a 1000 pb, lo cual significa un exceso en la longitud de las lecturas y alto costo en la secuenciación (Head y col., 2014). Este tipo de genoteca provee información acerca de inversiones, sustituciones de un solo nucleótido y variación en la estructura del genoma (Head y col., 2014).

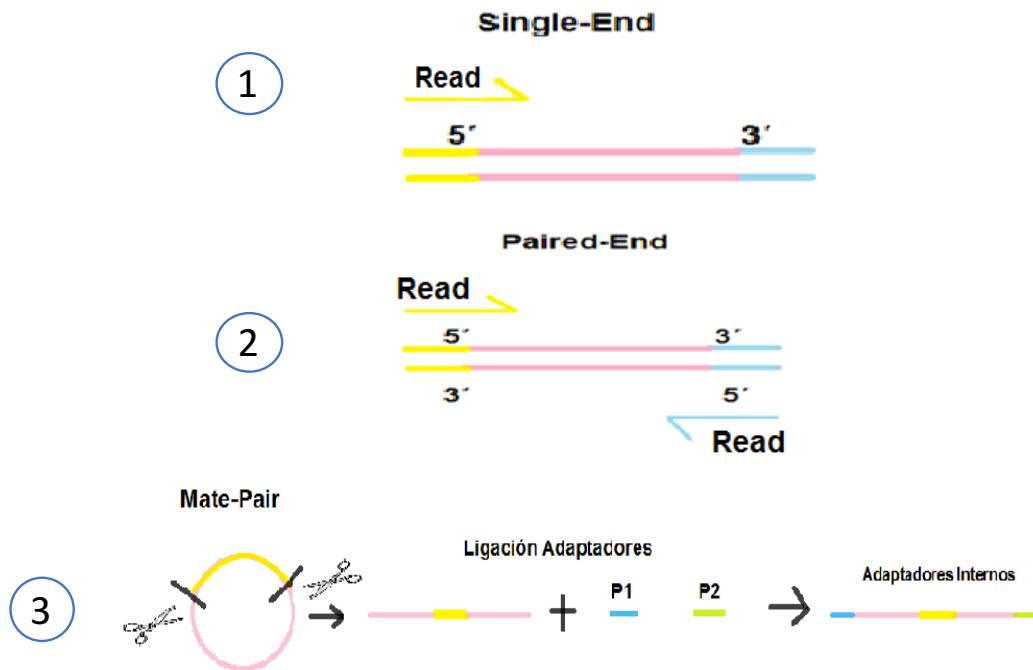


Fig. 5. Tipos de genotecas. 1. Genoteca “single-end”. La secuenciación de la genoteca es del extremo 5’ al extremo 3’ obteniéndose una sola secuencia por genoteca; 2. Genoteca “paired-end”. A partir de una genoteca “paired-end” se obtienen dos secuencias, sentido y antisentido; 3. Genoteca “mate pair”, En la primera etapa de la construcción de estas genotecas el ADN es recircularizado por medio de dos adaptadores.

(Modificado de Head y col., 2014)

e. Secuenciación de nueva generación aplicada para el estudio de microorganismos

El estudio de los microorganismos es esencial para conocer las enfermedades que causan y comprender a cerca de su función y evolución para poder desarrollar medidas y procedimientos de intervención (Lal y Narain, 2014). Una de las herramientas para ayudar en el estudio de estos mismos es la secuenciación de nueva generación la cual ha sido desarrollada para ser más barata y obtener de grandes volúmenes de información por lo que cada tecnología de secuenciación tiene sus fortalezas y debilidades, dependiendo de los factores que tienen que ser considerados para escoger la correcta plataforma como son el análisis de expresión de genes, identificación de SNP, metagenómica, ensamble por alineamiento y ensamble *de novo*.

d. Algoritmos de ensamblado *de novo*

Dado que las tecnologías de secuenciación han evolucionado rápidamente y con ellas también las metodologías para ensamblado *de novo*. Hasta la fecha ninguna tecnología de secuenciación genera suficiente información para secuenciar un organismo completo, por eso es necesario el uso de algoritmos computacionales para armar el “rompecabezas” del genoma a partir de los fragmentos pequeños que genera la secuenciación (Ekblom y col., 2014; Miller y col., 2010).

Los algoritmos más comunes para el ensamblado de lecturas eran basados en el enfoque de “overlap-layout-consensus” (Zhenyu y col., 2011). Este algoritmo calcula las sobre-posiciones entre los bordes de las lecturas. Después, cada lectura es graficada como un nodo y cada sobre-posición como una arista uniendo dos nodos que presenta este sobre-posición. El algoritmo después determina el camino que recorre cada nodo exactamente una vez (camino Hamiltoniano). La información redundante (vértices y aristas no usados) es desechada. El proceso es repetido una infinidad de veces, y las lecturas resultantes son combinadas para sacar un consenso. Algoritmos tales como

ARACHNE, Celera, Atlas, Phusion y PCAP se basan en este método (Havlak y col., 2004, Batzoglou y col., 2002, Myers y col., 2000, Huang y col. 2003, Mullikin y Ning, 2003).

El algoritmo *Brujin* se basa en dividir todas las lecturas en palabras de tamaño k (*k-meros*) y se crean grafos (vértices o nodos) a partir de estos *k-meros* donde cada arista representa una de estas secuencias y los nodos representan sobre-posiciones con otros *k-meros* (Zerbino y col., 2008). De las desventajas que puede tener este tipo de algoritmo es la resolución de secuencias repetidas debido a que la información de conectividad puede ser perdida al separar las lecturas en *k-meros*, especialmente en lecturas grandes.

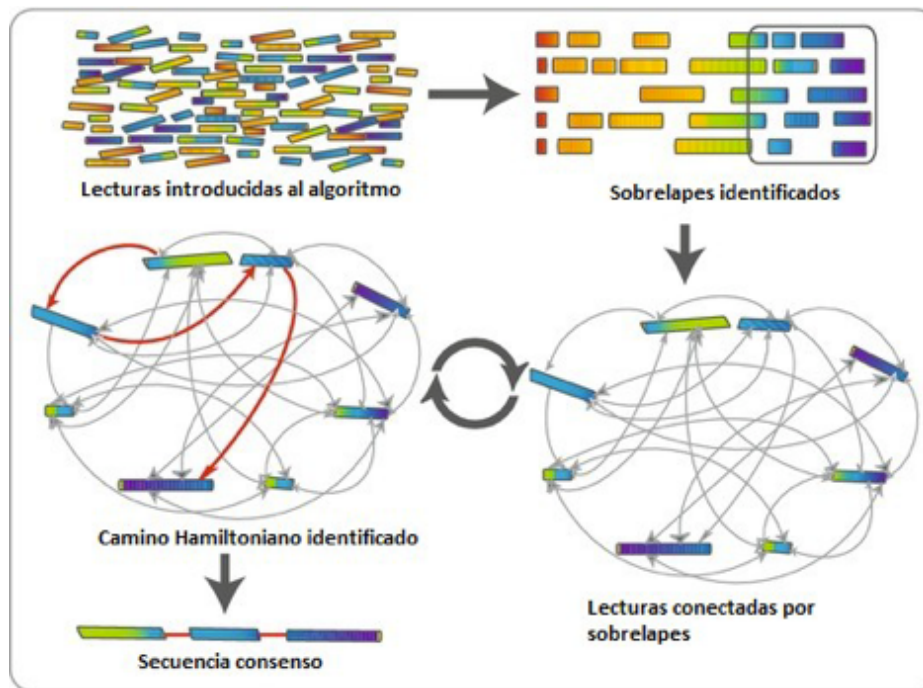


Fig. 6. Algoritmo overlap-layout-consensus. Primero, las lecturas obtenidas del secuenciador son introducidas al algoritmo. Después se procesan sobrelapes entre los bordes de estas lecturas para después hacer un grafo conectando estas lecturas por medio de estos sobrelapes. Se identifican un camino Hamiltoniano nuevamente se repite el proceso de construcción de grafo y resolución de camino Hamiltoniano. Y, por último, se llega a una secuencia consenso basada en los caminos Hamiltonianos encontrados. (Modificado de Comming y col., 2009).

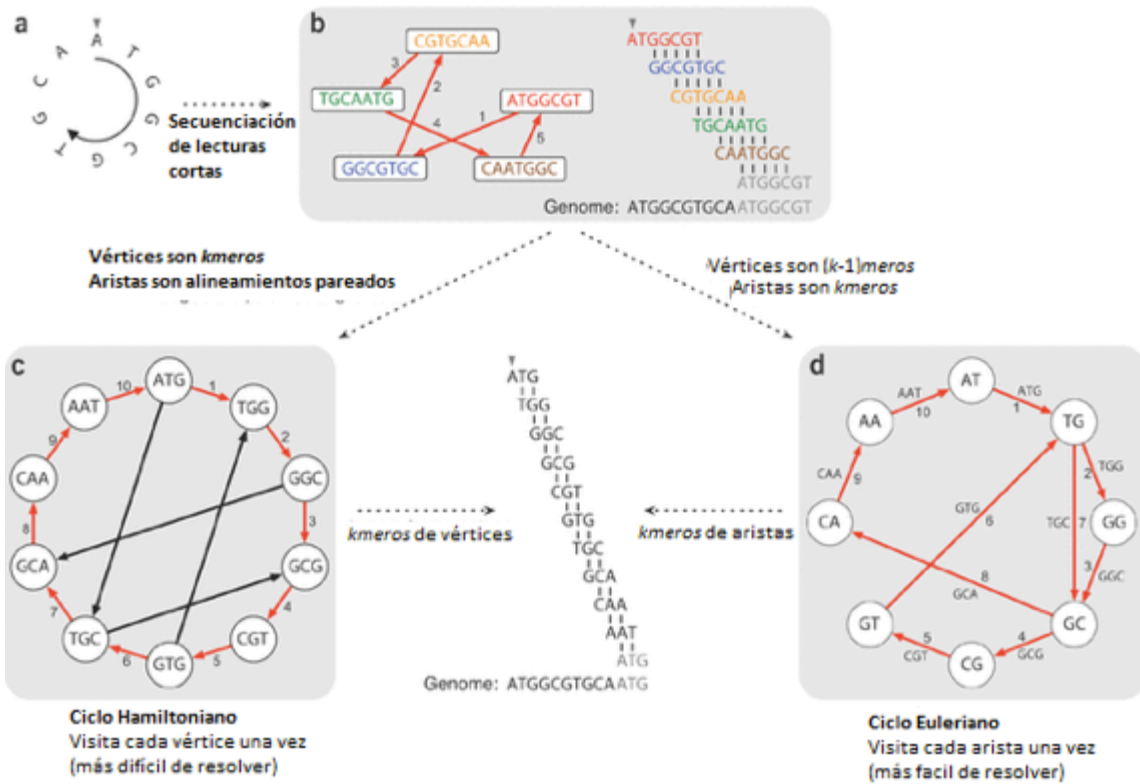


Fig. 7. Algoritmos basados en grafos de Bruijn. a) Muestra un ejemplo de un genoma circular. b) Muestra el acercamiento utilizado anteriormente para lecturas de tecnologías Sanger basado en solapamientos. c) y d) Muestran alternativas donde se parten primero las lecturas en secuencias de tamaño k (k -meros). En c) los k -meros son los vértices mientras que las aristas son alineamientos pareados. La resolución del genoma es por medio de un ciclo Hamiltoniano. En d) Las aristas son k -meros y los vértices son $(k-1)$ -meros. La resolución del genoma es por medio de un ciclo Euleriano. La mayoría de los algoritmos modernos basados en grafos de Bruijn se basan en d) por la mayor facilidad de resolución. (Modificado de Compeau y col., 2011)

IV JUSTIFICACIÓN

Salmonella representa un problema de salud pública a nivel mundial debido a múltiples brotes de transmisión alimentaria y su alta frecuencia de detección en el ambiente y alimentos. Particularmente en Sinaloa, *Salmonella enterica* subesp. *enterica* Oranienburg es reconocida como agente infeccioso de animales domésticos y agente de contaminación de ecosistemas acuáticos, lo cual exhibe el ciclo entorno–hospedero en la región, y sugiere las potenciales estrategias de supervivencia, adaptación y virulencia que emplea la bacteria. Sin embargo, estas estrategias no están definidas en las cepas de *S. Oranienburg* aisladas de agua de ríos de Culiacán, Sinaloa. Los estudios sobre análisis del genoma de *S. Oranienburg* son escasos, por lo cual se requiere del uso de herramientas de epidemiología molecular que permitan determinar la diversidad genética y patrones de distribución de las cepas circulantes de *S. Oranienburg* en la región, definir la biología del serotipo, y su potencial patogénico en animales y humanos.

V HIPOTESIS

La tipificación de patrones de virulencia permite establecer relaciones clonales entre cepas de *Salmonella* Oranienburg aisladas de ríos de Culiacán, con un mayor grado de discriminación comparado con la tipificación de secuencia multilocus, perfiles de resistencia y serología *in silico*.

VI OBJETIVOS

A Objetivo general

Proponer una herramienta de epidemiología molecular que permita establecer las relaciones clonales, patogénicas, resistencia a antibióticos y serológicas de diferentes cepas de *Salmonella* Oranienburg aisladas de agua de río de Culiacán, Sinaloa.

B Objetivos específicos

1. Anotar el genoma de las cepas de *S. Oranienburg* aisladas de aguas de río.
2. Determinar las líneas clonales de los aislamientos de *S. Oranienburg*.
3. Diferenciar a nivel intra-serotipo las cepas de *S. Oranienburg* mediante Tipificación por Multilocus de Secuencia (MLST), resistencia a antibióticos y serotipificación *in-silico*.
4. Tipificar los patrones de virulencia presentes de los aislamientos de *S. Oranienburg* mediante análisis *in-silico*.

VII MATERIALES Y METODOS

A Cepas de *Salmonella* Oranienburg

En este estudio se incluyeron 18 cepas de *Salmonella* Oranienburg aisladas previamente de distintos puntos de los ríos Humaya, Tamazula y Culiacán, localizados en Sinaloa durante el periodo 2008 a 2009. Las cepas de *S.* Oranienburg se caracterizaron previamente mediante serología y electroforesis en gel de campo pulsado (Jiménez y col., en 2014). Las cepas de *S.* Oranienburg se encuentran preservadas a -20 °C en la colección de cepas del Laboratorio de Investigación y Diagnostico Microbiológico de la Facultad de Ciencias Químico Biológicas de la Universidad Autónoma de Sinaloa. Las cepas de *S.* Oranienburg fueron reactivadas en caldo de soya tripticaseína (MCD lab, México) durante 24 h a 37 °C. Posteriormente, el cultivo se sembró sobre agar Hecktoen (Difco, EUA) para el aislamiento bacteriano.

B Extracción de ácidos nucleicos de cepas de *Salmonella*

A partir del agar Hecktoen (Difco, EUA) se inoculó una colonia pura de cada cepa en caldo TSB (Difco, EUA), y se incubó a 37 °C por 24 h. Se realizó una concentración de la pastilla celular a partir de alícuotas de 1.5 ± 0.25 mL del cultivo bacteriano de las suspensiones bacterianas asignadas, por centrifugación a $13,000 \pm 50$ rpm durante aproximadamente 5 min. La pastilla se resuspendió en 180 μ L del buffer ATL, 20 μ L de proteinasa K (Sigma-Aldrich, EUA) y se agitó vigorosamente durante aproximadamente 15 s. Posteriormente, se adicionaron 180 ± 10 μ L de buffer AL y se mezcló durante 15 s. Después se incubó a 56 ± 2 °C durante 20 min. A la mezcla se le adicionó 200 μ L de etanol grado biología molecular, y se homogenizó por 15 s. La suspensión resultante se transfirió a una columna DNesay Mini (Qiagen, Alemania) colocada en un tubo colector de 2 mL para centrifugar a 8,000 rpm por 1 min. Se desechó el tubo colector y se colocó

la columna a un nuevo tubo colector de 2 mL con 500 μ L del buffer AW1 y se centrifugo a 8,000 rpm por 1 min. Nuevamente, se desechó el tubo colector, y se colocó la columna en un tubo colector nuevo de 2 mL con 500 μ L de buffer AW2, se centrifugo a 8,000 rpm por 2 min, y se desechó el tubo colector. La columna se colocó en un tubo estéril de 1.5 mL con 50 μ L de buffer AE, y se incubó a temperatura ambiente (15-25 °C) por 1 min. Finalmente, el tubo se centrifugó a 8,000 rpm por un 1 min (este paso se repitió una vez más). La columna se desechó, y el ADN extraído remanente en el tubo se preservó a -20 °C para su posterior cuantificación [$\text{ng}/\mu\text{L}$] mediante espectrofotometría ($\lambda = 260/280$).

C Generación de genotecas

La Fig 8 muestra el flujo de preparación de genotecas a partir del ADN bacteriano extraído constituido en 6 fases.

Preparación del ADN

Se cuantificó la concentración y el tamaño del ADN empleando el equipo Qubit 2.0 (Thermo Fisher Scientific, EUA), y el kit de Qubit BR dsDNA Assay (Thermo Fisher Scientific, EUA) como los describe el fabricante. Para cada cepa se normalizó una concentración de 2 $\text{ng}/\mu\text{L}$.

Tagmentación del ADN de entrada.

Para la tagmentación, en un tubo de 1.5 ml estéril se adicionó 5 μL de ATM (Amplicon Tagment Mix) (illumina, EUA) y 5 μl de cada muestra, se homogenizó mediante pipeteo, y se centrifugó a 280 rpm por 1 min a 20 °C. Previamente, se programó el termociclador con el programa de tagmentación y se agregaron 5 μL de NT (Neutralize Tagment buffer). Una vez preparada la tagmentación se procedió a evaluar las muestras de ADN bacteriano, utilizando 1 μL de muestra

tagmentada en un chip de ADN de alta sensibilidad del Agilent Technology 2100 Bioanalyzer (Agilent Technology, EUA).

Amplificación de librerías

Este paso amplifica el ADN tagmentado usando un programa PCR de ciclo limitado. La etapa de PCR se suma el índice 1 adaptadores (i7) e índice 2 adaptadores (i5) (proporcionados por el kit Nextera) y secuencias requeridas de las muestras para la formación de agrupaciones. Los ciclos utilizados para la amplificación en el termociclador (Thermo Fisher Scientific, EUA) se muestran en el Cuadro 6. Se pipetió el volumen NPM (Nextera PCR Master Mix, siglas en inglés) requerido para las reacciones y se dividió el volumen equitativamente en alícuotas. Posteriormente, se organizaron los cebadores o adaptadores de índice en el accesorio TruSeq Index Plate Fixture (placa) divididas en filas y columnas de la siguiente manera: 1.- el adaptador índice en las columnas 1-12 (i7) del accesorio TruSeq Index Plate Fixture y 2.- adaptador índice en las filas A-H (i5) del accesorio. Adicionalmente, Las placas fueron cargadas con una pipeta multicanal, agregando 5 μ L de cada adaptador Índice 1 (i7) por debajo de cada columna de la placa, se reemplazó la tapa de cada tubo del adaptador i7 con una nueva tapa naranja para distinguirlas. De igual manera se agregaron 5 μ L de cada adaptador índice 2 (i5) a través de cada fila. se agregaron 15 μ L de NPM para cada tubo que contenía los adaptadores índices. Seguidamente, se centrifugo a 280 rpm a 20 °C por 1 min. Por último, se les aplico una corrida en con el programa de PCR en el termociclador.



Fig. 8. Flujo de trabajo para generación de genotecas.

Cuadro 6. Programa de tagmentación.

Ciclos	Temperatura	Tiempo
-	72 °C	3 min
-	95 °C	30 seg
	95 °C	10 seg
12	55 °C	30 seg
	72 °C	30 seg
-	72 °C	5 min
-	4 °C	∞

Limpieza de librerías

Para limpieza y purificación de las librerías se utilizaron perlas AMPure XP (Beckman coulter, EUA) de acuerdo a las instrucciones del fabricante. Brevemente, se centrifugaron las muestras a 280 rpm por 1 min, y se transfirieron 50 μ L de los productos de PCR a los pocillos de una nueva placa midi (Fisherscientific, EUA). Después, se agregaron 30 μ L de perlas AMPure XP a cada pocillo y se homogenizó a 1800 rpm durante 2 min. La mezcla se incubó a 25 °C por 5 min. Las muestras se colocaron en un soporte magnético y se dejaron en reposo por 2 min. Se descartó el sobrenadante de cada pocillo, y se realizaron 2 lavados a los pocillos con 200 μ L de etanol (Merck Millipore, EUA) frío al 80%, y la placa de pozos se incubó en el soporte magnético por 30 s.

Posteriormente, se descartó el sobrenadante de cada pocillo, y se dejó secar la placa en el soporte magnético a 25 °C por 15 min. La placa de pozos se removió del soporte magnético y se agregaron 52.5 μ L de RSB (Buffer de resuspensión) (illumina, EUA) a cada pocillo, y se centrifugó a 1,800 rpm por 2 min. La mezcla se mantuvo a 25 °C por 2 min en el soporte magnético hasta que el líquido se tornara transparente (~2 min). Por último, se transfirieron 50 μ L del sobrenadante del plato midi a una placa TCY (96 pocillos) (Biorad, EUA).

Verificación del control de calidad de las librerías de ADN y cuantificación con Qubit.

Para determinar la concentración (ng/ μ L) de ADN de cada muestra se utilizó el kit Qubit dsDNA HS Assay kit de acuerdo a las instrucciones del fabricante.

D Secuenciación

1. Preparación de las genotecas

La concentración de las librerías se ajustó a 2 nM, y se realizó la mezcla con todas las librerías. En un tubo estéril se ajustó una concentración final de 10 pM de ADN con un volumen final de 1000 μ L; se mezcló 5 μ L de las librerías, 5 μ L de NaOH 0.2 M y 990 μ L de buffer HT1. De esta mezcla, se colocaron 600 μ L en el pozo de la placa de secuenciación, y se agregaron insumos necesarios en el equipo illumina MiSeq (illumina, EUA) y se inició la corrida bajo los parámetros establecidos en el protocolo de Pulsenet.

E Análisis bioinformático primario

Exportación de los datos de corrida de MiSeq

Los datos fueron obtenidos directamente del equipo Illumina MiSeq (illumina, EUA) siguiendo la ruta de directorio, D drive -> Illumina -> MiSeqOutput -> Run folder -> data -> Intensites -> BaseCalls. En el directorio BaseCalls se encontraron los datos crudos en formato comprimido “*fastq.gz*” (Figura 8) y se obtuvieron 2 archivos por muestra correspondientes a las lecturas tipo paired-end.


```

@SRR014849.1 EIXKN4201CFU84 length=93
GGGGGGGGGGGGGGGGGGCTTTTTTTGTTTGGAAACCGAAAGG
GTTTTGAATTTCAAACCCTTTTCGGTTTCCAACCTTCCAA
AGCAATGCCAATA
+SRR014849.1 EIXKN4201CFU84 length=93
3+&$#"7F@71,'";C?,B;?6B;:EA1EA
1EA5'9B:?:#9EA0D@2EA5':>5?:%A;A8A;?9B;D@
/=<?7=9<2A8==

```

Fig 9. Formato de archivo descomprimido ".fastq". Se caracteriza por estar compuestos por cuatro líneas: (i) Primera línea de título '@' que a menudo tiene sólo un identificador de registro (formato libre sin límite de longitud, (ii) Segunda línea es la secuencia de nucleótidos, (iii) Tercer línea, para señalar el final de las líneas de secuencia y el inicio de la cadena de calidad, viene el símbolo '+', (iv) Cuarto línea para indicar las calidades que usa un subconjunto de los caracteres imprimibles ASCII (como máximo ASCII 33-126), donde se almacena las puntuaciones o valores del índice PHRED, lo cual proporciona una codificación simple y razonablemente eficiente en el espacio (Cock y col., 2010).

Control de calidad de las secuencias genómicas

Una vez obtenidos los datos crudos (*fastq.gz*) del proceso de secuenciación se realizó un control de calidad con el programa libre FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) que permite visualizar la calidad de las lecturas que permiten un ensamblaje adecuado. Los parámetros que se incluyeron fueron:

- Calidad por base secuenciada (índice PHRED).
- Contenido de Adenina, Citosina, Timina y Guanina por secuencia (%).
- Contenido de GC (%).
- Presencia de secuencias sobrerrepresentadas (grafico).
- Contenido de adaptadores (%).
- Contenido de k-meros (%).

Recorte de adaptadores y recorte por calidad

Para la remoción de adaptadores se utilizó el programa libre PEAT (Paired-end Adapter Trimmer, siglas en inglés) (<https://github.com/jhhung/PEAT>) que utiliza el algoritmo de poda de adaptador rápido y de alta precisión, diseñado específicamente para las secuencias tipo paired-end. Adicionalmente, se recortaron las secuencias por calidad utilizando el programa Trimmomatic (Bolger y col. 2014) que se aplica paired-end y single-end. Los parámetros utilizados para el recorte fueron:

- ILLUMINACLIP: Adaptador de corte y otras secuencias-Illumina específica de la lectura.
- SLIDINGWINDOW: Excluye la lectura por debajo de la ventana deslizante de recorte (umbral) establecida.
- MINLEN: Excluye la lectura por debajo de una longitud especificada.
- TOPHRED33: Convertir las puntuaciones de calidad de Phred-33.
- TOPHRED64: Convertir las puntuaciones de calidad de Phred-64.

Ensamble de secuencias genómicas *de novo*

Las secuencias se ensamblaron utilizando el pipeline A5-MiSeq (Tritt y col., 2012). Brevemente, las secuencias limpias y recortadas se utilizaron para construir “contigs” utilizando el ensamblador IDBA (Peng y col., 2012). Posteriormente, se formaron “scaffolds” o andamios (unión de contigs) utilizando el programa SSPACE (Boetzer y col., 2010). Finalmente, el pipeline realizó un control de calidad sobre la construcción correcta de los scaffolds.

Evaluación de los ensamblados *de novo*

Una vez obtenidos los ensamblados *de novo* se realizó una evaluación de los genomas ensamblados con el programa A5-MiSeq QUAST (Quality Assessment Tool, siglas en inglés) (<http://quast.sourceforge.net/quast>). El programa QUAST produjo cuadros sinópticos y diagramas para la representación del tamaño de los contigs, los ensamblados perdidos y estructuras variables, la representación del genoma y sus elementos funcionales y las variaciones del N50 basados en el alineamiento de los bloques.

Alineamiento de genomas

La evaluación de los ensamblados se realizó ordenando los contigs con respecto a un genoma de referencia mediante el programa Mauve (Darling y col., 2010), el cual alinea regiones de ortólogos y xenólogos entre dos secuencias de un genoma, para observar cambios locales, reordenamientos y la sintenia de los genes (Darling y col., 2010). Como cepa de referencia se utilizó una *Salmonella* Bredeney (cepa filogenéticamente más emparentada a las muestras).

F Anotación de los genomas

Después de obtener los ensamblados de los 26 genomas de *S. Oranienburg*. Se

utilizó la herramienta automatizada libre RAST (*Rapid Annotation using Subsystem Technology*, por sus siglas en inglés) para anotar genomas de bacterias. El servicio identifica proteínas que codifican, genes, secuencias de ARNr y ARNt, asigna funciones a los genes, predice cuales subsistemas están representados en el genoma, y utiliza esta información para reconstruir la red metabólica (Azis y col. 2008). Los datos de la anotación (tamaño total del genoma, fuente del aislamiento y el nombre de los genes de las regiones codificantes) se obtienen en archivo “.xls” y archivo “.gbk”.

1. Obtención del pseudocromosoma

Para la construcción del pseudocromosoma se utilizó el programa libre CONTIGuator (<http://contiguator.sourceforge.net/>). Los archivos en formato “FASTA” se sometieron al programa, el cual realiza el mapeo de los contigs contra un genoma de referencia (S. Bredeney) usando un algoritmo de BLASTN que permite la visualización de un mapa de contigs, la pérdida y/o ganancia de elementos genéticos y otras características descriptivas como solapamientos putativos entre contigs mapeadas y regiones sintéticas y no sintéticas dentro y fuera de los contigs permitiendo terminar genomas.

G Árbol filogenético de genoma completo (WGS)

Para la reconstrucción filogenética se utilizó el pipeline REALPHY. Primeramente, se creó un archivo de configuración “*config.txt*” con la ruta necesaria para correr los programas (bowtie2, bowtie2.build y PhyML-3.1) y se colocó en la carpeta donde se depositaron los archivos de salida. El árbol de genoma se construyó con los polimorfismos de un solo nucleótido (SNP's, siglas en inglés) provenientes de los ensamblajes con recorte de adaptadores y calidad. El pipeline utiliza el programa bowtie2 y bowtie-build para hacer un mapeo y un algoritmo de máxima verosimilitud que reconstruye múltiples alineamientos de secuencias a partir del cual los árboles filogenéticos son inferidos con el

programa PhyML (<http://www.atgc-montpellier.fr/phyml/>).

H Árbol filogenético de genes de virulencia

Se tomaron los archivos de los ensamblajes (contigs) con recorte de adaptadores y calidad. Posteriormente, a cada archivo se les convirtió las cabeceras (*headers*) correspondientes al formato “.FASTA”, requerido para que el programa de un solo archivo con un concatenado de todas las secuencias consenso de las cepas. Seguidamente, se cambiaron las extensiones de los archivos “.FASTA” a “.fas” para que el programa Mugsy (Angiuoli y Salzberg, 2011) reconociera el archivo y así poder alinear las secuencias concatenadas. Adicionalmente, se utilizó el programa SeaView (Gouy y col., 2010) para visualizar los correctos alineamientos. Por otra parte, el programa Mugsy otorgo archivos con extensión “.MAF” por lo cual se reconvirtieron a la extensión “.FASTA” (formato requerido para realizar la filogenia) mediante el uso de plataforma abierta Galaxy (<https://usegalaxy.org/>), una vez obtenido el archivo con los consensos de las secuencias concatenadas y el formato correcto. Finalmente, Se utilizó el programa Mrbayes (Ronquist y col., 2012) para crear el árbol filogenético con genes de virulencia.

I Virulotipificación *in silico*

Para construir una base de datos de los genes de virulencia característicos de *Salmonella*, se realizó una búsqueda intencionada de las secuencias de nucleótidos de los genes de virulencia correspondientes a las islas de patogenicidad (ISP-1, ISP-2, ISP-3, ISP-4, ISP-5 y ISP-7), Quorum sensing, fimbrias y plásmidos en la base de datos *GenBank* del *National Center for Biotechnology information* (NCBI) (Benson y col. 2012), Virulence Factors Database (VFD) (Chen y col. 2004) y Protein Data Bank (PDB). Se obtuvieron un total de 65 genes representantes de virulencia en formato “.gbk”, los cuales se depositaron en un cuadro “.xls”. Posteriormente, se procedió a realizar una base

de datos dentro del programa CLC Genomics Workbench para hacer un mapeo en batch de la base de datos de los genes virulencia (.xls) contra las secuencias (previamente sometidas al control de calidad) de cada cepa. Después, se realizó un BLAST directo y un BLAST inverso con los contigs anotados y la base de datos en CLC Genomics Workbench para constatar la presencia o ausencia del gen de interés. Un valor de >99% de similitud se consideró como un gen presente, y un <99% de similitud para realizar pruebas detalladas (mapeo y BLAST) para constatar la ausencia total de gen. Una vez obtenida la matriz en una hoja de cálculo (.csv) se exportó al programa para estadística computacional y gráficas "R" (<https://www.r-project.org/>). Se utilizó la paquetería *heatmap.2* para hacer un mapa de calor, el cual fue adaptado para dar un dendrograma de virulotipificación.

J Tipificación por multilocus de secuencia (MLST) y resistencia a antibióticos.

Para la tipificación por multilocus de secuencias se utilizó el programa *SRST2* (Inouye y col., 2014). El programa cuenta con la base de datos de los siete genes de mantenimiento: *aroC* (corismato sintasa), *dnaN* (ADN polimerera III subunidad beta), *hemD* (uroporfirinogeno III corisintasa), *hisD* (histidinol deshidrogenasa), *purE* (fosforibosilaminoimidazol carboxilasa), *sucA* (alfa cetoglutarato deshidrogenasa), *thrA* (Aspartatocinasa + homoserina deshidrogenasa), descritos por la Universidad de Warwick (Reino Unido). Las secuencias en formato ".fastq" se sometieron al programa *SRST2* para obtener el MLST de las 18 *S. Oranienburg*. Una vez obtenido el MLST se procedió a realizar el árbol recubridor mínimo (MST, siglas en inglés), el cual se hizo a partir de una tabla (formato ".txt",) colocando en la primera columna la secuencia tipo (ST) y en las columnas posteriores el eburst-group de los genes constitutivos mediante el programa PHILOViZ (Ribeiro-Gonçalves y col., 2016) mediante la opción de MLST contenida en el mismo programa.

K Resistencia antibióticos

Para la resistencia a antibióticos se utilizaron las secuencias en formato “.fastq” y se sometieron al programa Srst2 (Inouye y col., 2014) en terminal. El programa proporciona una tabla con la resistencia a antibióticos.

L Serotipificación in-silico.

La serotipificación se llevó a cabo mediante el programa en línea Seqsero (Zhang y col., 2015), utilizando los archivos en formato “.fastq.gz” (previamente sometidos al control de calidad).

VIII RESULTADOS Y DISCUSIÓN

A Ensamble y anotación de los genomas de *Salmonella* Oranienburg aisladas de agua de río del valle de Culiacán.

1 Obtención los archivos crudos (.fastq) del secuenciador

En este trabajo las 18 cepas de *S. Oranienburg* fueron secuenciadas en el *Center for Food Safety and Applied Nutrition* (CFSAN, por sus siglas en inglés) mediante el equipo *illumina MiSeq* (*illumina*, San Diego, CA) utilizando el kit *Nextera sample preparation 2 x 251* (*illumina*, San Diego, CA) con secuencias tipo *Pired-end*. A cada uno de los archivos “.fastq.gz” generados para cada muestra en la secuenciación se les asignó un código proporcionado por el Laboratorio de Investigación y Diagnóstico Microbiológico (LIDiM) de la Universidad Autónoma de Sinaloa, tal como se muestra en el Cuadro 7.

2 Evaluación de calidad de las cepas.

Los archivos “.fastq.gz” de las secuencias de las cepas de *S. Oranienburg* (18x2) se sometieron al análisis de calidades por el programa *FastQC*, el cual consistió en observar las secuencias de baja calidad (Bolger y col., 2014 y Li y col., 2015). Con finalidad de hacer el recorte de calidad y adaptadores para obtener secuencias de buena calidad y poder realizar el ensamble de los genomas por lo cual solo se consideró el índice *PHRED* (>20) como un parámetro significativo aquellas lecturas que cumplieran para el recorte de secuencias de mala calidad.

3 Evaluación de las comparaciones por recorte de adaptadores y calidad (*Trimmomatic* vs *PEAT* vs *CLC*)

Se realizó un análisis comparativo con los programas *PEAT* (Li y col., 2015), *Trimmomatic* (Bolger y col., 2014) y *CLC Genomics Workbench* ([/www.qiagenbioinformatics.com](http://www.qiagenbioinformatics.com)). Cada uno de los programas realizó el recorte

de adaptadores y secuencias de baja calidad, basado en el índice PHRED de las secuencias (Fig. 10).

En la Fig. 10 se muestra una comparación del recorte de las secuencias analizadas por los tres programas. Los programas PEAT y CLC Genomics Workbench mostraron un comportamiento similar, conservando un mayor número de secuencias útiles para el proceso de ensamble (recorte de adaptadores del 1%). Mientras que, el programa Trimmomatic recorta una mayor cantidad de adaptadores en un 48.3% de secuencias.

Los resultados obtenidos se deben a la diferente implementación de la búsqueda de ad adaptadores en las secuencias. Li y col. (2015) demuestran que PAET presenta una mayor precisión y sensibilidad al de eliminar adaptadores, comparado con Trimmomatic. El programa Trimmomatic realiza una búsqueda de los adaptadores basado en tablas hash y en puntuación simple (algoritmos de alineamiento), lo que descuida las inserciones y las supresiones para secuencia de adaptadores (Jiang y col., 2014). El algoritmo del programa PEAT está diseñado para detectar la intersección entre la secuencia y los adaptadores en dos secuencias pareadas (sentido y antisentido), convirtiendo a PEAT un programa más sensible para detección de adaptadores (Li y col., 2015).

Cuadro 7. Códigos obtenidos para el manejo de las 18 muestras.

Código CFSAN	Código LIDiM	Pulsotipo	Origen	Archivo
CFSAN039514	S-04	X13	Acuático	CFSAN039514.fastq
CFSAN039515	S-08	X07	Acuático	CFSAN039515.fastq
CFSAN039516	S-12	X07	Acuático	CFSAN039516.fastq
CFSAN039518	S-13	X07	Acuático	CFSAN039518.fastq
CFSAN039520	S-17	X12	Acuático	CFSAN039520.fastq
CFSAN039521	S-20	X11	Acuático	CFSAN039521.fastq
CFSAN039522	S-22	X11	Acuático	CFSAN039522.fastq
CFSAN039523	S-29	X06	Acuático	CFSAN039523.fastq
CFSAN039525	S-31	X01	Acuático	CFSAN039525.fastq
CFSAN039527	S-69	X05	Acuático	CFSAN039527.fastq
CFSAN039528	S-71	X10	Acuático	CFSAN039528.fastq
CFSAN039529	S-72	X07	Acuático	CFSAN039529.fastq
CFSAN039530	S-76	X07	Acuático	CFSAN039530.fastq
CFSAN039531	S-82	X07	Acuático	CFSAN039531.fastq
CFSAN039532	S-100	X08	Acuático	CFSAN039532.fastq
CFSAN039536	S-106	X02	Acuático	CFSAN039536.fastq
CFSAN039538	S-110	X04	Acuático	CFSAN039538.fastq
CFSAN039540	S-60	X07	Acuático	CFSAN039540.fastq

Cuadro 8. Parámetros de las lecturas obtenidos con el programa FastQC.

Cepa	Secuencias			% GC	Índice PHRED*	contenido de N	K-mero frecuente
	Total, de secuencias	marcadas como de mala calidad	Longitud de secuencia				
S-04	579856	0	35-251	52	Cumple	0	TATATGG
	579856	0	35-252	52	Cumple	0	ATACTGC
S-08	1326283	0	35-253	52	Cumple	0	TATACCG
	1326283	0	35-254	52	Cumple	0	GTATTAA
S-12	751449	0	35-255	52	Cumple	0	TGAGGAA
	751449	0	35-256	52	Cumple	0	TATACCG
S-13	572401	0	35-257	52	Cumple	0	TTGAAAT
	572401	0	35-258	52	Cumple	0	CGGGGGG
S-17	914060	0	35-259	52	Cumple	0	ACCGTGC
	914060	0	35-260	52	Cumple	0	TATACCG
S-20	553704	0	35-261	52	Cumple	0	ATACTGT
	553704	0	35-262	52	Cumple	0	CCTTAAC
S-22	596204	0	35-263	52	Cumple	0	ACTGTGC
	596204	0	35-264	52	Cumple	0	GTGTAGC
S-29	879805	0	35-265	52	Cumple	0	GTGTATG
	879805	0	35-266	52	Cumple	0	TATACCG

S-31	904735	0	35-267	52	Cumple	0	AGGACAG
	904735	0	35-268	52	Cumple	0	TATACCG
S-69	978736	0	35-269	52	Cumple	0	TATACCA
	978736	0	35-270	52	Cumple	0	ATAAGAC
S-71	564107	0	35-271	52	Cumple	0	TACACCG
	564107	0	35-272	52	Cumple	0	TAATACA
S-72	728110	0	35-273	52	Cumple	0	GTCATAG
	728110	0	35-274	52	Cumple	0	TAACACT
S-76	979829	0	35-275	52	Cumple	0	CTACACC
	979829	0	35-276	52	Cumple	0	GTATTAA
S-82	724143	0	35-277	52	Cumple	0	GTATTAA
	724143	0	35-278	52	Cumple	0	TGGGGGG
S-100	714391	0	35-279	52	Cumple	0	TATACAG
	714391	0	35-280	52	Cumple	0	CCAGGGG
S-106	835393	0	35-281	52	Cumple	0	TATACAG
	835393	0	35-282	52	Cumple	0	GTGTTAA
S-110	841063	0	35-283	52	Cumple	0	GAGGAAA
	841063	0	35-284	52	Cumple	0	GTATTAC
S-60	1306470	0	35-285	52	Cumple	0	GTATTAA
	1306470	0	35-286	52	Cumple	0	ACCGGGG

*solo el parámetro de índice PHRED es significativo para determinar las calidades de las secuencias.

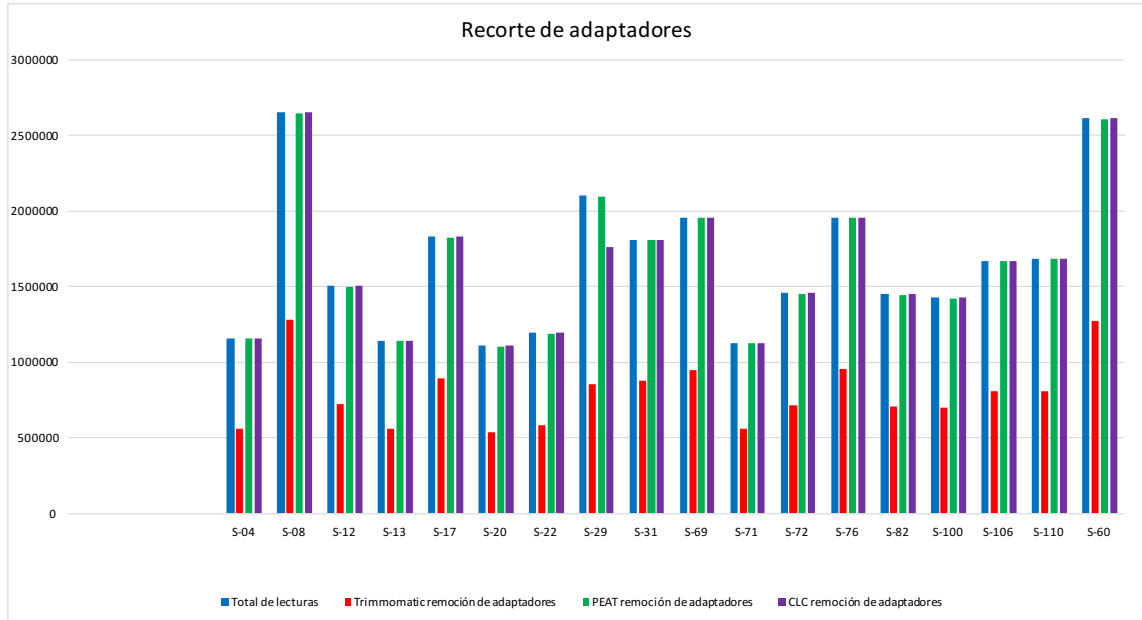


Fig. 10. Comparación de los programas Trimmomatic, PEAT (softwares libres) y CLC Genomics Workbench (software de licencia) de recorte de adaptadores de las muestras. En el eje de las X se muestra el código asignado por el LIDiM de las muestras y en el eje de las Y se muestra el número de lecturas.

En segundo lugar, el recorte de secuencias por calidad, solamente se compararon los programas Trimmomatic (Bolger y col., 2014) y CLC Genomics Workbench ([/www.qiagenbioinformatics.com](http://www.qiagenbioinformatics.com)), esto porque PEAT (Li y col., 2015) no realiza recorte por de las secuencias de baja calidad. Para realizar el recorte de calidad se utilizó un índice PHRED (>20), es decir, las lecturas por debajo de este número se consideran de mala calidad y fueron eliminadas, debido que en los análisis posteriores pueden causar interferencias al momento de realizar los demás análisis. La Fig. 11 muestra la comparación de los programas CLC Genomics Workbench y Trimmomatic. Brevemente, el programa Trimmomatic eliminó una mayor cantidad de secuencias, un 48.3% de las secuencias con respecto al total de las lecturas, debido a que posee un algoritmo que realiza una selección más estricta de las secuencias de adaptadores, mediante algoritmos de alineamientos locales, como se mencionó anteriormente, mientras que CLC emplea el algoritmo modificado de Mott este funciona de la siguiente manera: Para cada base, se resta la probabilidad de error, con base de un valor de corte de probabilidad de error (valor por defecto de 0,05) para formar la puntuación. La probabilidad de error de base se calcula a partir de la puntuación de calidad (Q), tal que la probabilidad de error es igual $P(\text{error}) = 10^{-(Q/10)}$. Esto significa que las bases de baja calidad tienen probabilidades de error altas y por lo tanto pueden tener una puntuación de base negativa y son recortadas.

Dado a los resultados obtenidos en la comparación de los programas para el análisis y selección de secuencias, se seleccionó y se empleó complementariamente los programas PEAT y Trimmomatic para el recorte de adaptadores y secuencias de baja de calidad, respectivamente.

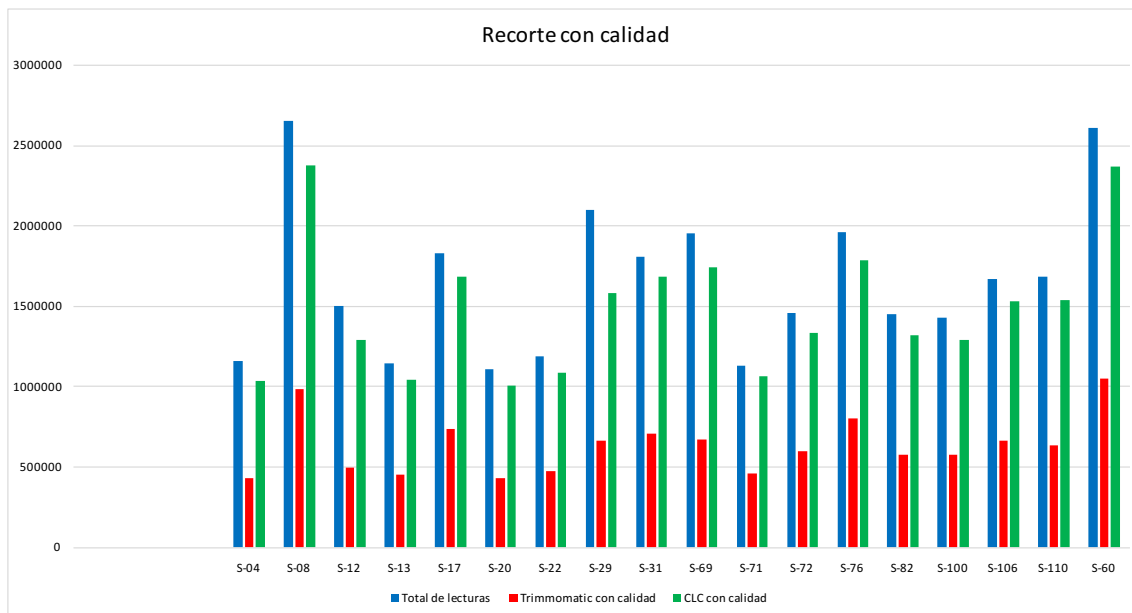


Fig. 11. Comparación de los programas Trimmomatic (software libre) y CLC Genomics Workbench (software de licencia) con recortes de calidad. En el eje de las X se muestra el código asignado por el LIDiM de las muestras y en el eje de las Y se muestra el número de lecturas.

Ensamble del genoma de las cepas de *Salmonella* Oranienburg.

Una vez que las secuencias pasaron por el análisis bioinformático primario se realizó el ensamble de las secuencias de las cepas de *S. Oranienburg* mediante la comparación de dos programas CLC Genomics Workbench (software de licencia) ([/www.qiagenbioinformatics.com](http://www.qiagenbioinformatics.com)) y A5-MiSeq (software libre) (Tritt y col., 2012), ambos programas utilizan el algoritmo de grafos de Bruijn (es un grafo dirigido que representa los solapamientos entre las secuencias). Sin embargo, ambos programas son diferentes, debido a que CLC realiza únicamente un ensamble *de novo* y A5-MiSeq es un pipeline compuesto por diferentes programas que además de realizar el ensamble de las secuencias, también el recorte por calidad de las secuencias, limpieza y formación de scaffolding (Coli y col., 2015).

La calidad de los ensamblados se determina mediante la evaluación de métricas básicas como: el número de contigs, N50 y contig más largo. La Fig. 12 muestra el número de contigs de cada una de las cepas de *S. Oranienburg* analizadas por los dos programas de ensambladores. Todas las cepas mostraron un número de contigs similar, lo cual significa que el ensamblado del genoma no presentó fragmentación. Es decir, a mayor sea el número de contigs, el genoma resulta más fragmentado, mientras que, a menor sea este, menor fragmentación del genoma (Gurevich y col., 2013). EL CLC Genomics Workbench y A5-MiSeq mostraron resultados similares para las cepas en estudio.

El N50 es otra de las métricas utilizadas para medir la calidad del ensamble, el cual se define como el valor de la longitud del último contig añadido para dar el 50% del genoma. La Fig. 13 muestra que el valor del N50 osciló entre 100 mil a 500 mil pares de bases. En general, se obtuvieron contigs grandes lo que indica que el ensamble se llevó a cabo de manera correcta. A pesar de que los dos ensambladores mostraron un comportamiento similar en sus métricas. El programa A5-MiSeq mostró valores de N50 mayores que CLC, indicando un mejor rendimiento durante el ensamblado.

La Fig. 14 muestra el contig más largo que oscilaron entre los 400 mil hasta un millón de pares de bases (pb) por contig, lo cual indica un 1/5 del tamaño del genoma de *Salmonella*, considerando que el tamaño promedio del genoma completo es ≈ 5 millones pb, Ambos programas resolvieron contigs largos para cada una de las cepas de *S. Oranienburg*, pero, con el programa A5-MiSeq se obtuvieron contigs más largos. Esto se debe a la forma particular del pipeline A5-MiSeq que utiliza diferentes programas para mejorar los ensamblajes y construir scaffolds o andamios. El pipeline une contigs para la formar los scaffolds utilizando el programa SSPACE y en esta etapa realiza un control de calidad de los scaffolds verificando que sea una construcción correcta, si no es así los remueve, esto independiente del microorganismo ensamblado.

Finalmente, ambos ensambladores mostraron resultados satisfactorios, sin embargo, al final de la comparación de métricas, gracias a los programas de mejoramiento de ensamblaje de A5-MiSeq, se pudo recuperar contigs más largos, lo que evitó, tener problemas de secuencias repetidas, dentro de los genomas que generalmente causan ruido en los ensamblajes, ya que, este tipo de secuencias, generalmente son sobrelapadas por los ensambladores, como si se tratará de unas solas secuencias y no repetidos contiguos. Por lo cual se optó por utilizar durante todos los análisis posteriores necesarios, los ensamblajes que se obtenidos con A5-MiSeq.

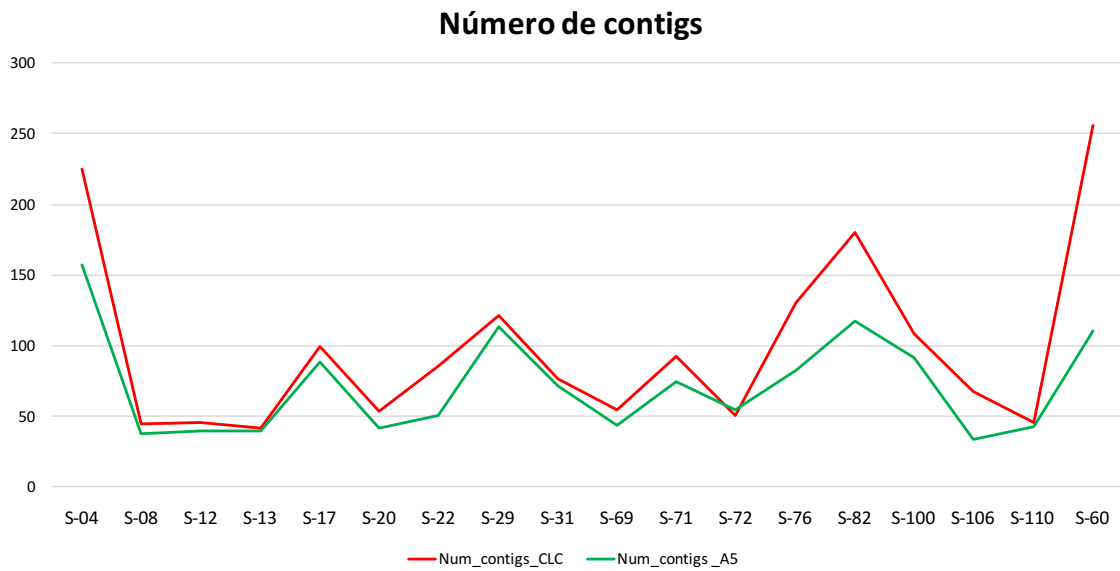


Fig. 12. Comparación del número de contigs realizados por los programas CLC y A5-MiSeq. En el eje de las X se muestra el código asignado por el LIDiM para las muestras y en el eje de las Y se muestra el número de contigs.

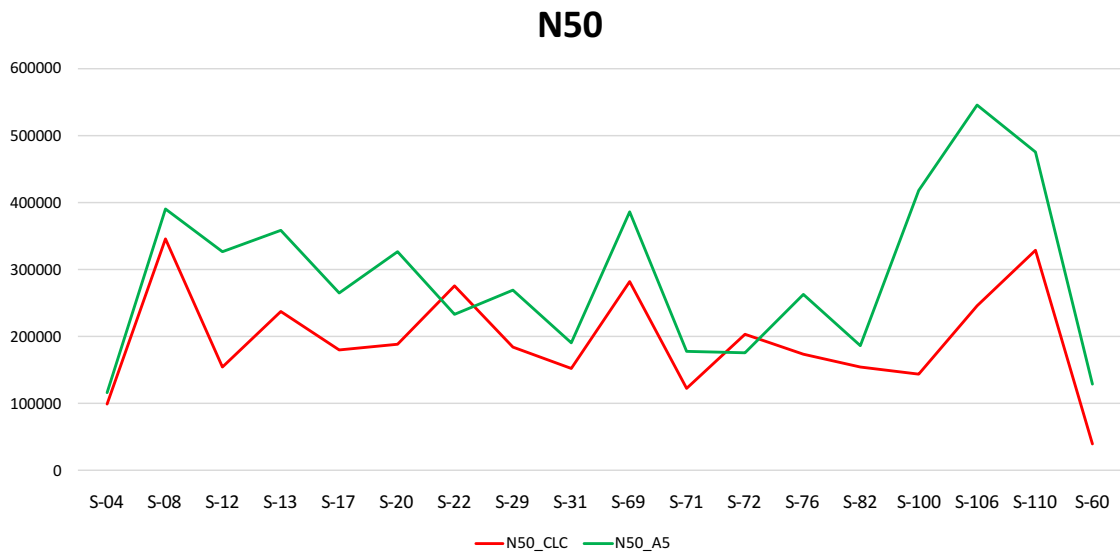


Fig. 13. Comparación del N50 realizados por los programas CLC y A5-MiSeq. En el eje de las X se muestra el código asignado por el LIDiM a las muestras y en el eje de las Y el número de secuencias para formar la mitad del genoma.

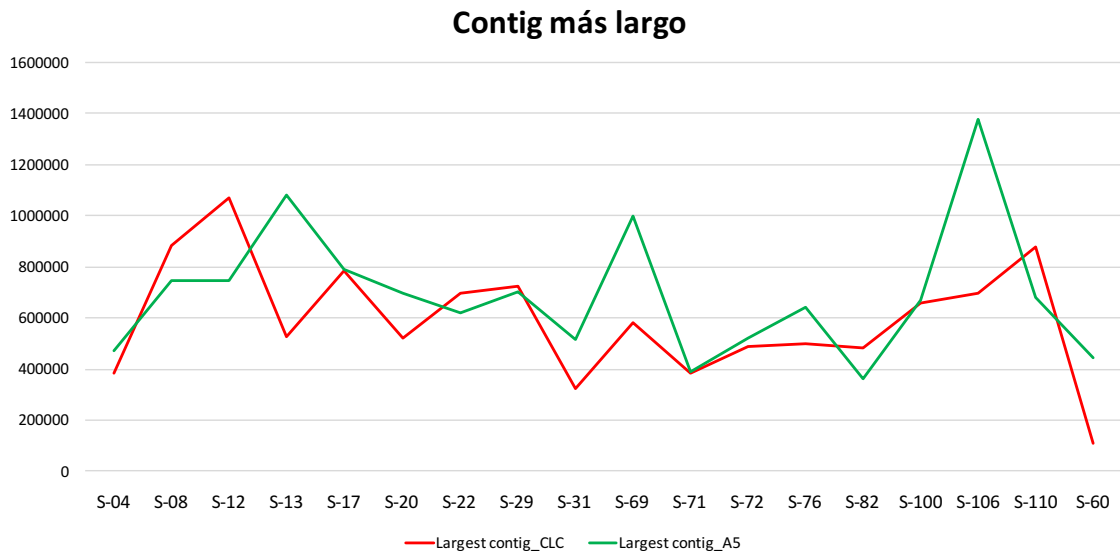


Fig. 14. Comparación de la métrica de contig más largo obtenidos entre los programas CLC y A5-MiSeq. En el eje de las X se muestra el código asignado a las muestras por el LIDiM y en el eje de las Y se muestra el número de secuencias.

Alineamiento de ensamblajes de *Salmonella* Oranienburg

Una vez alineados los contigs, se ordenaron con una secuencia de referencia se utilizó la cepa *S. Bredeney*, dado que no existe una secuencia de referencia de *S. Oranienburg*, el criterio de selección de la *S. Bredeney* como referencia fue en base a lo realizado en el programa Ragout (Kolmogorov y col., 2014) cuyo algoritmo descompone los genomas de entrada en un conjunto de cadenas de bloques de sintenia, a modo que cuando el conjunto de cadenas está concatenado. Mientras que cada cromosoma de referencia contiguo se transforma en una sola secuencia de bloques de sintenia. Ragout infiere una relación filogenética entre los genomas y construye un gráfico de punto de ruptura a partir de los conjuntos de cadenas de bloques de sintenia y luego utiliza un enfoque de reordenación para inferir cuál de las secuencias de referencia utilizadas está más cercana evolutivamente a la secuencia problema o blanco (Kolmogorov y col., 2016) como se muestra en la Fig 15.

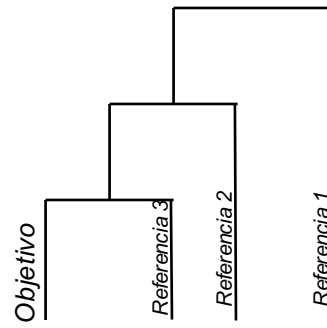


Fig 15. Reconstrucción filogenética

Kolmogorov y col. (2016)

Los resultados de los alineamientos se muestran en la Fig. 16, donde se observan los bloques lineales de colores que representan los genes que está compuesto el genoma, se observó que estos conservan el mismo patrón de sintenia, es decir, la organización espacial de los genes es idéntica en los genomas analizados, con esto se infiere relaciones de homología entre los genomas, por lo tanto, están filogenéticamente emparentados, además con esto se puede inferir que las muestras que se secuenciaron corresponden a un mismo serotipo bacteriano.

Una vez que se determinó la secuencia de referencia, a utilizar se determinaron el número de “gaps” o huecos presentes en las muestras secuencias y con ello poner determinar qué tan cubierto se tuvo el genoma, para ello se realizó el ordenamiento utilizando el programa ABACAS (Assefa y col., 2009), que ordena y orienta los contigs tomando como base una referencia y cerrar los gap entre los contigs ordenados, el utiliza el programa MUMmer (Kurtz et al., 2004), que utiliza el algoritmo de árboles de sufijos para realizar los alineamientos. Un árbol de sufijos es una estructura de datos para representar todas las subcadenas de una cadena, ya sea que esa cadena sea una secuencia de ADN, una secuencia de proteínas o texto plano. Los árboles de sufijo tienen las siguientes características agradables que los convierten en una estructura de datos importante para el análisis de genomas a gran escala: un árbol de sufijo para una cadena “S” de longitud “n” puede representarse en el espacio proporcional de “n”; Dado el árbol de sufijo de “S” y una cadena de consulta “Q” de longitud m, hay algoritmos para calcular todas las coincidencias máximas únicas entre “S” y “Q” de cualquier longitud mínima especificada en tiempo proporcional a “m”, lo que lo hace un algoritmo bastante rápido de alineamiento de genomas (Kurtz y col., 2004).

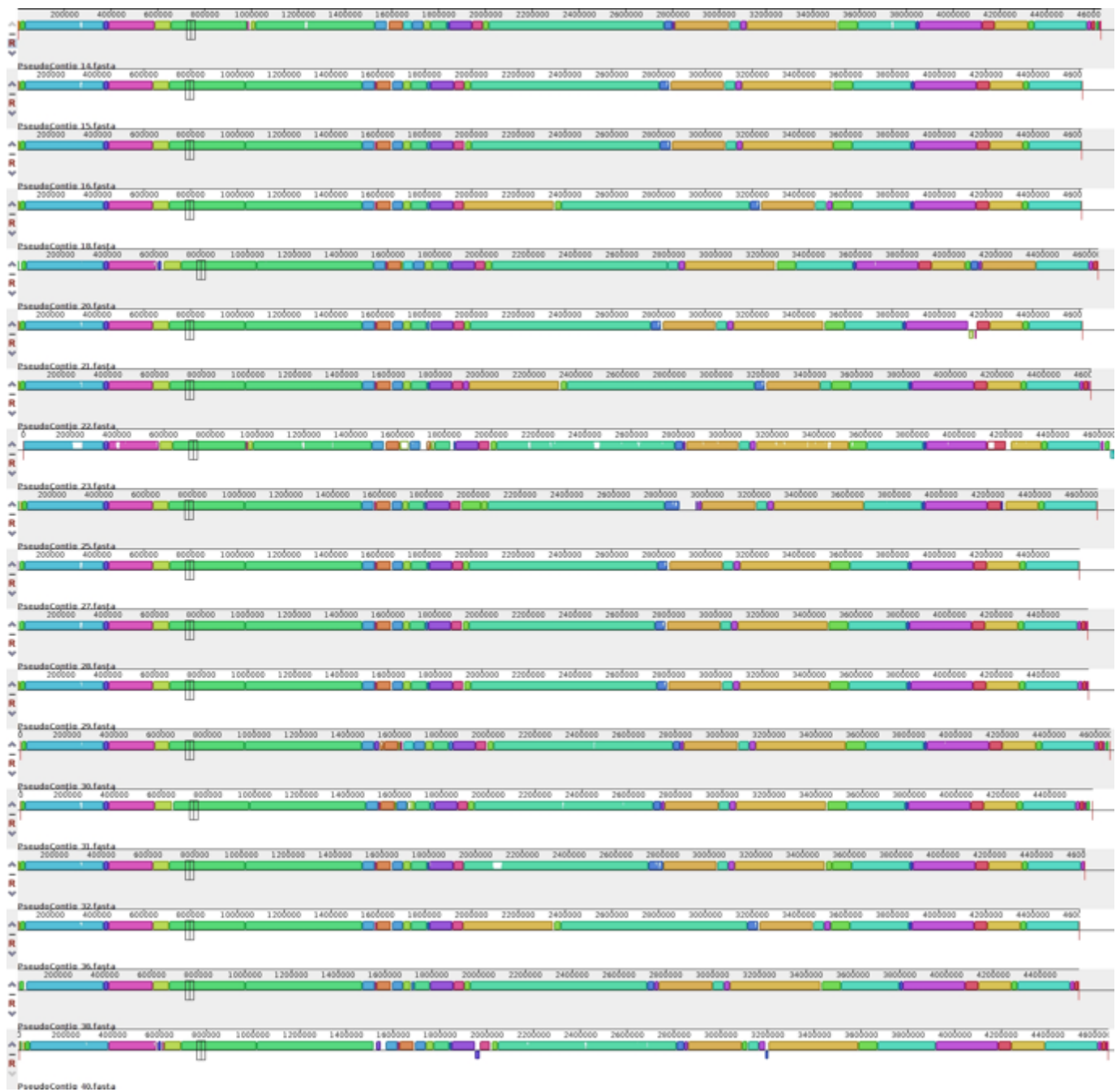


Fig. 16. Alineamiento progresivo de las 18 secuencias de *Salmonella* Oranienburg.

Los resultados del número de “gaps” encontrados se muestran en la Cuadro 9 donde se puede observar que las cepas S-12, S-69, y S-106 fueron las que tuvieron menor número de gaps con tan solo 16, mientras que las cepas que tuvieron un mayor número de gaps fueron la S-31 y S-04 con 30 y 37 respectivamente.

Anotación de los genomas por RAST.

Una vez ensamblados los genomas, estos se anotaron con el objetivo de darle un sentido a las secuencias mediante la herramienta RAST (Aziz y col., 2008). Este estudio reporta los primeros borradores de genomas de *S. Oranienburg* aislados de los ríos Culiacán, Humaya y Tamazula en Sinaloa (Jiménez y col., 2014) en las bases de datos de NCBI. Esta información generada contribuye a al conocimiento de la biología y evolución de *S. Oranienburg*.

La Fig. 17 muestra la representación de uno de los genomas de *S. Oranienburg* (S-04), describiendo la clasificación en subsistemas del contenido genético. Adicionalmente, el Cuadro 9 describe los parámetros de No. de contigs, Tamaño de genoma (pb), No. de gaps, Contenido de G+C (%), Secuencias codificantes de ADN (pb), No. de proteínas, No. de genes virulencia/estrés y Cobertura (fold) de las 18 cepas del serotipo *S. Oranienburg*. Generalmente, el tamaño promedio del cromosoma bacteriano fue de 4759515.5, con 52% GC. Interesantemente, las cepas coinciden en el número de genes de estrés (174), pero, varían en el número de genes de virulencia. Este hecho, permitir inferir que las bacterias emplean un mecanismo de sobrevivencia similar en el entorno de origen, pero, pudiesen emplear virulencia diferente.

Anotación de la muestra S-04 (CFSAN039514)

Subsystem Information

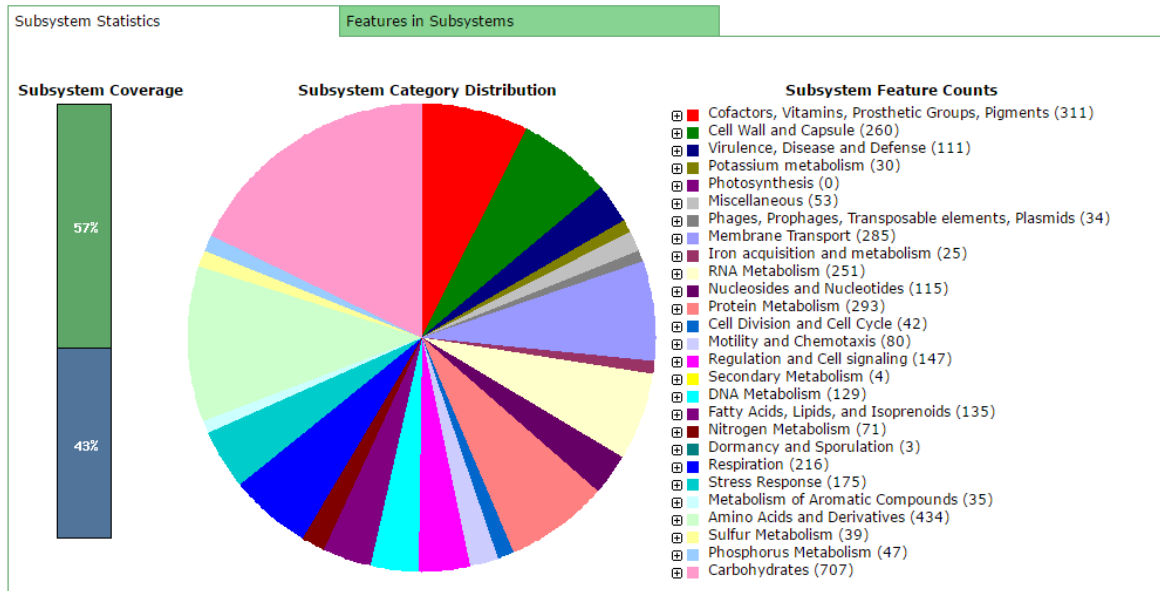


Fig. 17. Representación general de la anotación de las 18 *Salmonella* Oranienburg

Cuadro 9. Metadatos de cepas de *Salmonella* Oranienburg obtenidas de agua de río con RAST

Código LIDiM	No. de contigs	Tamaño de genoma (bp)	No. de gaps	Contenido de G+C (%)	Secuencias codificantes de ADN (bp)	No. de proteínas	No. de genes virulencia/estrés	Cobertura (fold)
S-04	157	4994902	37	51.9	4956	5076	111/175	34
S-08	38	4623009	17	52.1	4495	4610	108/174	87
S-12	40	4622245	16	52.1	4497	4614	109/174	45
S-13	40	4621080	20	52.1	4493	4610	107/174	38
S-17	89	4959125	22	51.9	4932	5054	109/175	57
S-20	42	4621690	21	52.1	4503	4620	108/174	36
S-22	51	4620771	24	52.1	4493	4612	108/174	38
S-29	113	5018503	17	52.1	5090	5201	102/175	59
S-31	72	4706253	30	52	4566	4685	108/175	68
S-69	44	4623133	16	52.1	4493	4607	108/174	75
S-71	75	4624848	23	52.1	4499	4620	108/174	49
S-72	55	4622567	21	52.1	4504	4623	108/174	47
S-76	83	4984035	23	52	4978	5095	109/175	56
S-82	117	5004735	22	52	4991	5102	108/175	43
S-100	92	4852085	17	52	4726	4838	108/175	42
S-106	34	4619639	16	52.1	4496	4607	108/174	55
S-110	43	4583664	22	52.1	4442	4557	108/174	58
S-60	110	4968996	29	51.9	4946	5052	109/175	31
Promedio	72	4759516	22	52	4672	4788	108/174	51

Previamente, Medrano-Félix y col. (2013) reportaron un borrador de genoma (draft genome) de *S. Oranienburg* aislado de origen acuático en Sinaloa; la secuenciación del genoma fue mediante la plataforma Ion Torrent reportando una cobertura promedio de 35.8, y un tamaño de cromosoma de 4,609,551 pb, incluyendo 5,020 secuencias codificantes y un contenido de G+C de 52.2%. Estos datos son similares a las cepas de *S. Oranienburg* de este estudio (Cuadro 9). Cabe señalar, que el promedio de cobertura obtenido (38.5 vs 51) en este estudio fue mayor debido a la plataforma de secuenciación empleada. La plataforma illumina MiSeq permite obtener mayor información (lecturas) ya que esta plataforma secuencia bajo la generación de sentido y anti-sentido (pired-end).

Adicionalmente, la base de datos de genoma NCBI incluye genomas (borradores o completos) sobre diversos serotipos de *Salmonella* como *S. Typhimurium* (Nguyen y col., 2016), *S. Typhi* (Ong y col., 2012), *S. Enteritidis* (Moreau y col., 2015), *S. Java* (Fazal y col., 2016) y *S. Heidelberg* (Lebbé y col., 2016), los cuales han sido aislados principalmente de casos clínicos o agentes etiológicos de brotes epidemiológicos. El enriquecimiento de las bases de datos de genomas de microorganismos puede contribuir al conocimiento de la biología de los microorganismos (genes de virulencia, toxinas y genes de resistencia) e incluir nuevas medidas que permitan el control (bio-protección) de estos microorganismos (Chen y col., 2004; Ecker y col., 2005; Liu y col., 2009; Zhou y col., 2006).

B Determinación de las líneas clónales de los aislamientos de S. Oranienburg en los ríos de Sinaloa.

1 Árbol filogenético de las cepas de S. Oranienburg en función al genoma completo

El análisis de filogenia del genoma completo de las 18 cepas de S. Oranienburg mediante el programa Realphy, permitió determinar un mismo origen evolutivo entre las cepas de este serotipo, es decir, son genomas con una alta homología (Bertels y col., 2014). El programa Realphy, es un programa que permite realizar inferencias filogenéticas basado en un método de modelos evolutivos de máxima verosimilitud, donde la verosimilitud es la probabilidad de los datos moldeados al azar para generar la relación filogenética entre los objetos de estudio (Bertels y col., 2014).

El programa Realphy identificó tres clados (ramificaciones) principales, *i*) clado-II compuesto por la mayoría de las cepas de S. Oranienburg (n=17), *ii*) clado-I correspondiente a la cepa S29_A10_X06, y el *iii*) clado-III correspondiente a la muestra de S. Montevideo (cepa para enraizar el árbol de filogenia). Los clados identificados constituyen un grupo de taxones filogenéticamente emparentado con el grupo de estudio, que corresponde a un ancestro común, por lo tanto, se observa mínima heterogeneidad genética entre las cepas de S. Oranienburg.

Adicionalmente, se realizó una comparación del árbol filogenético de los genomas completos obtenidos por medio de la secuenciación masiva y el programa Realphy con los pulsotipos de PFGE previamente reportados (Jiménez y col., 2014) para las cepas de S. Oranienburg de este estudio. La comparación arrojó que el método de PFGE muestra una mayor heterogeneidad genética (15 pulsotipos), mientras que el árbol de filogenia muestra dos clados (I y II) para la clasificación de los 15 pulsotipos de PFGE.

La PFGE es un método considerado el estándar de oro para *Salmonella*, dado

que permite determinar huella o perfiles genéticos. No obstante, la relación genética debe ser utilizado como una guía, y no como una verdadera medida filogenética, dado que las bandas del mismo tamaño pueden no provenir de la misma parte del cromosoma, o bien, el cambio en un sitio de restricción puede resultar en más de un cambio de banda (Salipante y col., 2015). Mientras que la secuenciación masiva del genoma completo, y la reconstrucción filogenética permite el alineamiento y ordenamiento de los genes, conservando la sintenia en los genes (Salipante y col., 2015). En este sentido, las cepas tipificadas por secuenciación permiten inferir que las cepas aisladas en diferentes puntos de los ríos circundantes del valle de Culiacán son de la misma clona, como consecuencia de su serotipo y/o origen de aislamiento. Por lo tanto, se sugiere la incorporación cepas de *S. Oranienburg* de diferentes fuentes de origen, a fin de esclarecer esta hipótesis.

La no concordancia de los diferentes pulsotipos en el PFGE y el árbol filogenético de las cepas de *S. Oranienburg* (Fig. 16), podrían explicarse por polimorfismo presentes en las secuencias donde realiza el corte la enzima *XbaI* (Davis y col., 2003; Salipante y col., 2015). Cabe señalar, que las tasas de mutación en bacterias son generalmente bajas, y se estima aproximadamente en 1×10^{-8} base errónea, lo que significa un error por cada 100 millones de bases que se replican (Galán y col., 2005). Esto, podría indicar que el serotipo *S. Oranienburg* mantiene una estructura genética conservada, y cuyos cambios (polimorfismos) estarán influenciados al entorno que coloniza.

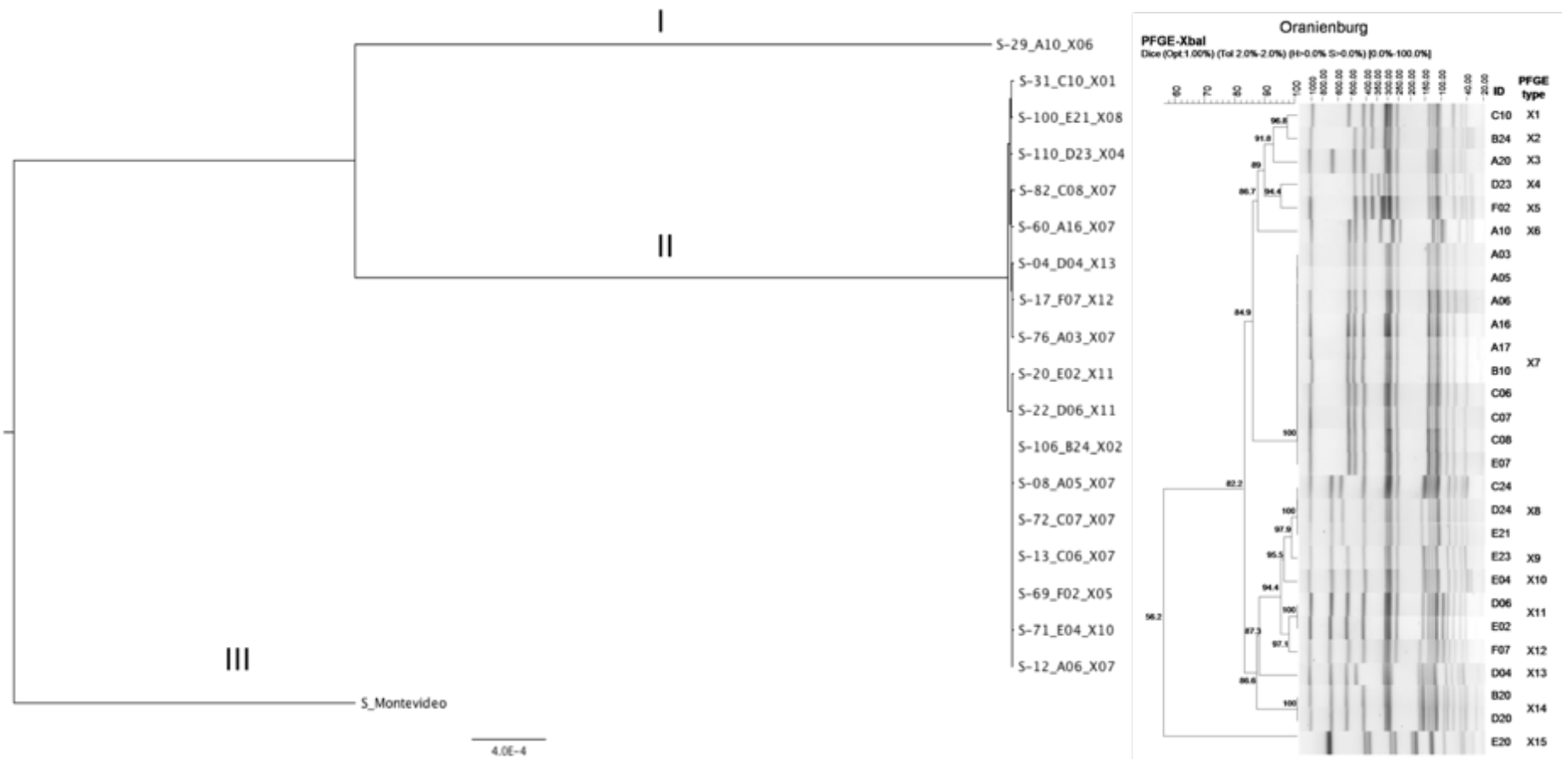


Fig. 18. Árbol filogenético de genoma completo (WGS). El código de muestra asignado por LIDiM y perfil de PFGE.

Árbol filogenético de genes de virulencia

Existen múltiples herramientas fenotípicas y genotípicas para evaluar la clonalidad entre organismos. Sin embargo, las herramientas bioinformáticas, permiten evaluar la presencia de genes (virulencia, resistencia a antibióticos, metabolismo, replicación, entre otros) *in silico* e inferir la relación de filogenia entre ellos de forma precisa y rápida (Salipante y col., 2015; Vílchez y col., 2009). En este contexto, se realizó un base de datos con 65 genes de virulencia de las cepas de *Salmonella*, y posteriormente se realizó un mapeo de las secuencias de las cepas con esta base de datos, para inferir las filogenias de las cepas de acuerdo al atributo de virulencia con el programa Mrbayes. El programa Mrbayes se basa en la estadística bayesiana, que estima la probabilidad de que determinadas topologías hayan sido producidas por un alineamiento, obteniendo una población de árboles, a partir de los cuales se construye un árbol consenso (Peña, 2011).

La Fig. 19, muestra la relación filogenética de las 18 cepas de *S. Oranienburg*, distinguiéndose dos clados diferentes (I y II), pero ambos emparentados filogenéticamente. Cabe señalar que, las cepas de los clados de virulencia corresponden idénticamente a los clados de genoma completo. Adicionalmente, el clado-I presenta subclados con patrones similares de virulencia (Figura 21). Sin embargo, las cepas S-17_F07_X12 (I.b1), S-31_C10_X01 (I.b2.1), S-110_D23_X04 (I.b2.1), y S-82_C08_X07 (I.b2.2), presentan mayores diferencias en los genes de virulencia con respecto a las demás cepas del clado-I. Las diferencias de genes de virulencia es posiblemente debido a factores que han provocado la pérdida o ganancia de genes como consecuencia de la presión ambiental donde se han desarrollado, o bien, al tiempo de aislamiento (Retamal y col., 2015). El perfil de genes virulencia de *Salmonella* puede determinar dos aspectos relevantes: *i*) la filogenia basada en virulencia como un atributo de tipificación, y *ii*) las cepas de *S. Oranienburg* aisladas de río son capaces de

establecer un cuadro patogénico en el hospedero. En este sentido, Asakura y col. (2002) y Miller y col. (2016) han definido la capacidad patogénica en *S. Oranienburg* en modelos murinos y cultivos celulares, respectivamente. Cabe señalar, que estos autores han mostrado la diversidad de la virulencia de este serotipo, lo cual se relaciona con la presencia o ausencia de genes.

C Tipificación por multilocus de secuencia (MLST) de las cepas.

1 Árbol recubridor mínimo (MST)

La técnica de Tipificación por Multilocus de Secuencias a partir de la secuenciación masiva se realizó mediante, el programa SRST2 (Inouye y col., 2014), con lo cual se obtuvo un solo ST para las 18 cepas de *S. Oranienburg* (Cuadro 10). A partir del ST obtenido se elaboró un árbol recubridor mínimo con el programa PHYLOViZ (Ribeiro-Gonçalves y col., 2016). Adicionalmente, las secuencias fueron analizadas por el programa SeqSero para determinar el serotipo de las cepas, lo cual permitió clasificarlas como *S. Oranienburg* (Cuadro 11). La Fig. 20 señala una agrupación clonal compartida entre las cepas de *S. Oranienburg* correspondiente al ST-23. Esto indica que el estudio de MLST no mostró diferencias intraserotipo de las cepas de *S. Oranienburg*, es decir, no se presentaron diferencias en los alelos ni en la secuencia tipo que se obtuvo a partir del análisis de éstos. Atchman y col. (2012) señalan que la conservación de la clonalidad puede ser atribuido a que los genes constitutivos tienen pocas variaciones en su composición a lo largo del tiempo, por lo que la divergencia clonal se debe a la constante presión selectiva a la cual pueda estar expuesta la bacteria. Por lo tanto, la clonalidad en las cepas de *S. Oranienburg* de este estudio es un indicador del origen de las cepas, y la adaptabilidad que han establecido en este ecosistema. Se sugiere la comparación de MLST entre cepas de *S. Oranienburg* de orígenes diversos (hospederos, alimentos, suelo, entre otros).

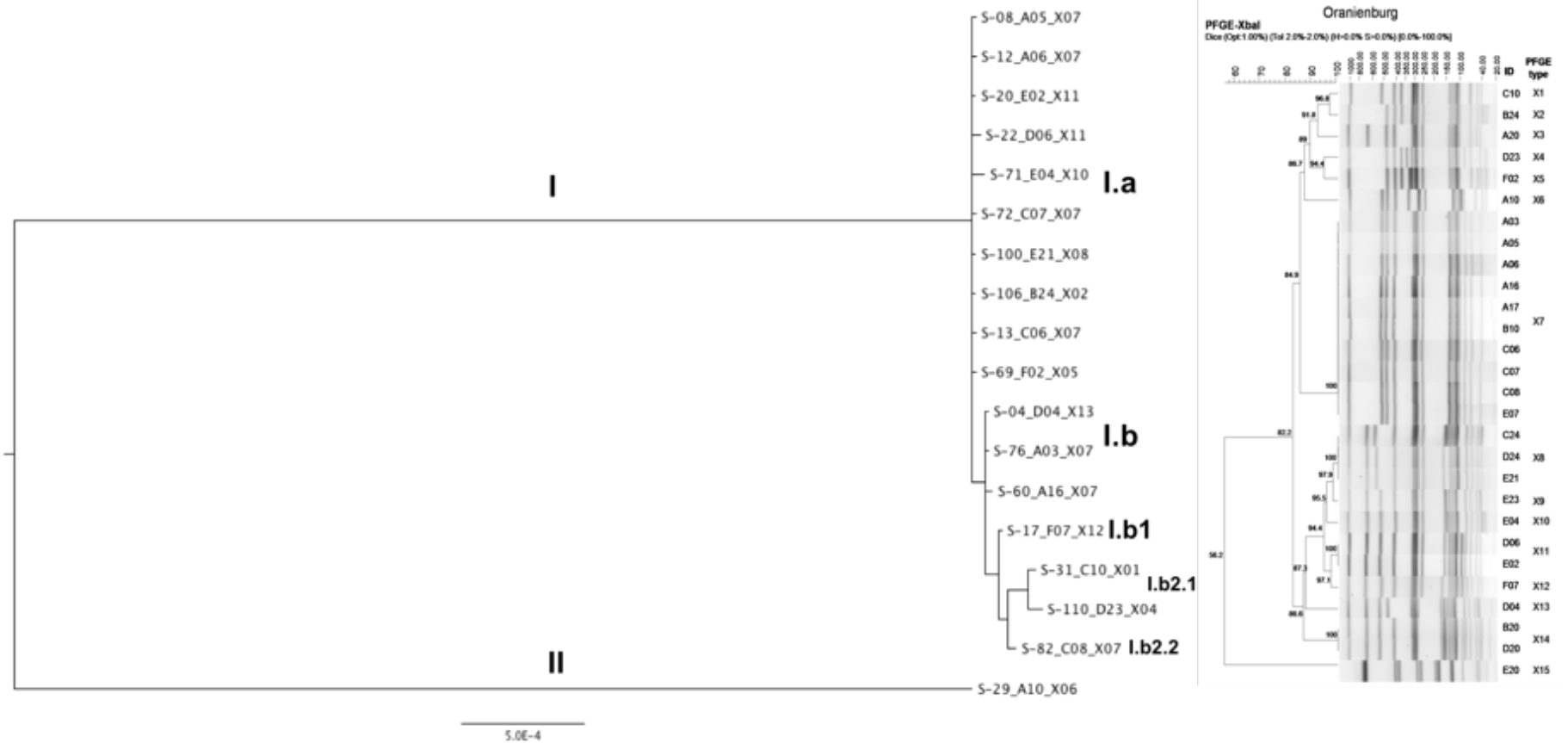


Fig. 19. Árbol filogenético de genes de virulencia hecho con Mrbayes. Donde se muestra la comparación con la electroforesis en gel de campos pulsados.

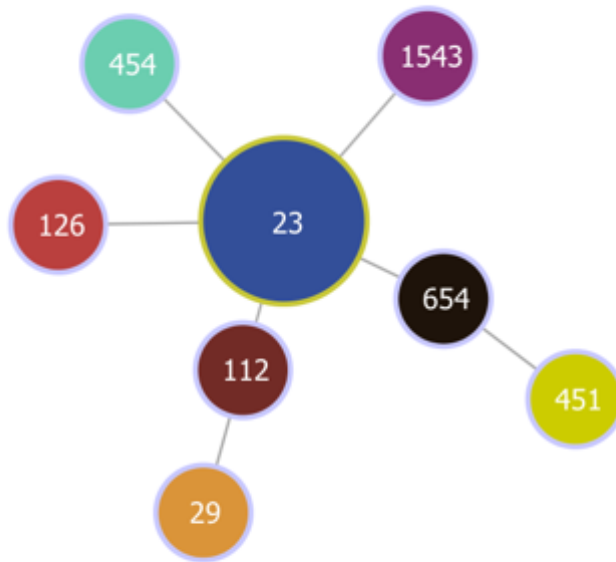


Fig. 20. Árbol recubridor mínimo (MST, siglas en inglés) de cepas de *Salmonella* Oranienburg y otros serotipos no tifoideos, obtenido con el programa PHYLOViZ a partir de MLST. Serotipos no tifoideos de acuerdo a la secuencia tipo: ST-23 (18 cepas de *S. Oranienburg*), ST-29 (1 cepa de *S. Stanley*), ST-112 (1 cepa de *S. Muenchen*), ST-126 (cepa de *S. Sandiego*), ST-1543 (1 cepa de *S. Kedougou*), ST-454 (1 cepa de *S. houtenae*), ST-654 (1 cepa de *S. Give*), ST-451 (1 cepa de *S. Pomona*).

Cuadro 10. MLST de las cepas de *S. Oranienburg* con los 7 genes constitutivos, ST y ST complex.

Muestra	<i>aroC</i>	<i>dnaN</i>	<i>hemD</i>	<i>hisD</i>	<i>purE</i>	<i>sucA</i>	<i>thrA</i>	ST	ST complex
S-04	13	11	16	15	12	15	4	23	41
S-08	13	11	16	15	12	15	4	23	41
S-12	13	11	16	15	12	15	4	23	41
S-13	13	11	16	15	12	15	4	23	41
S-17	13	11	16	15	12	15	4	23	41
S-20	13	11	16	15	12	15	4	23	41
S-22	13	11	16	15	12	15	4	23	41
S-29	13	11	16	15	12	15	4	23	41
S-31	13	11	16	15	12	15	4	23	41
S-69	13	11	16	15	12	15	4	23	41
S-71	13	11	16	15	12	15	4	23	41
S-72	13	11	16	15	12	15	4	23	41
S-76	13	11	16	15	12	15	4	23	41
S-82	13	11	16	15	12	15	4	23	41
S-100	13	11	16	15	12	15	4	23	41
S-106	13	11	16	15	12	15	4	23	41
S-110	13	11	16	15	12	15	4	23	41
S-60	13	11	16	15	12	15	4	23	41

Cuadro 11. Serotipificación *in-silico* de las muestras con el programa SeqSero.

Muestra	Antígeno	fljC	fljB	Perfil antigénico	sdf	Serotipo
S-04	O-7	m,t	-	7:m,t:-		Oranienburg
S-08	O-7	m,t	-	7:m,t:-		Oranienburg
S-12	O-7	m,t	-	7:m,t:-		Oranienburg
S-13	O-7	m,t	-	7:m,t:-		Oranienburg
S-17	O-7	m,t	-	7:m,t:-		Oranienburg
S-20	O-7	m,t	-	7:m,t:-		Oranienburg
S-22	O-7	m,t	-	7:m,t:-		Oranienburg
S-29	O-7	m,t	-	7:m,t:-		Oranienburg
S-31	O-7	m,t	-	7:m,t:-		Oranienburg
S-69	O-7	m,t	-	7:m,t:-		Oranienburg
S-71	O-7	m,t	-	7:m,t:-		Oranienburg
S-72	O-7	m,t	-	7:m,t:-		Oranienburg
S-76	O-7	m,t	-	7:m,t:-		Oranienburg
S-82	O-7	m,t	-	7:m,t:-		Oranienburg
S-100	O-7	m,t	-	7:m,t:-		Oranienburg
S-106	O-7	m,t	-	7:m,t:-		Oranienburg
S-110	O-7	m,t	-	7:m,t:-		Oranienburg
S-60	O-7	m,t	-	7:m,t:-		Oranienburg

Un estudio similar realizado por Retamal y col. (2015) caracterizó fenotípicamente y genotípicamente cepas de *S. Enteritidis* aisladas de humano y animal en Chile, demostrando que las cepas de origen animal comparten el mismo ST, correspondiente al ST-11, con perfiles fenotípicos y genotípicos diversos. Sin embargo, Martínez-Gamboa y col. (2015) utilizando MLST y el marcador molecular IS200 mediante PCR, encontraron dos ST diferentes dentro de la población de *S. Typhi* aislada de casos clínicos de diferentes países; identificando la variante *hemD340* que corresponde a un polimorfismo de un solo nucleótido (SNP).

Adicionalmente, Atchman y col. (2012) propone la metodología de MLST como una alternativa a la serotipificación tradicional dado su poder discriminativo entre serotipos. Con la finalidad de evaluar este atributo se serotipificaron *in-vitro* con MLST. Por lo que, en este estudio al árbol recubridor de *S. Oranienburg* se le adicionaron cepas de *Salmonella* de diferentes serotipos, observándose el mantenimiento de las cepas de *S. Oranienburg* en el ST-23, y su diferenciación respecto a los serotipos seleccionados (Fig. 20). Brevemente, en el árbol recubridor mínimo se observa la clonalidad de este serotipo representado en el círculo de mayor tamaño (ST-23) que refleja la frecuencia de la ST (numero de muestras). Adicionalmente, el árbol recubridor mínimo muestra que los serotipos SNT están estrechamente emparentados dada la topología del árbol. Sin embargo, el algoritmo de BURST no encontró ninguna diferencia entre los 7 genes de las muestras de *S. Oranienburg* (intraserotipo), pero si entre serotipos de SNT.

Debido a los resultados obtenidos con el MLST en las 18 cepas de *S. Oranienburg*, se pueden considerar que las cepas pertenecen a la misma clona (basado en los genes constitutivos). No obstante, las cepas de *S. Oranienburg* no están exentas de variaciones en funciones de su genoma (virulencia, resistencia antimicrobiana, respuesta ante estrés, etc.), debido a la dinámica del ciclo de vida de *Salmonella* (Winfield y Groisman, 2003).

D Tipificación de genes de defensa de las cepas de S. Oranienburg mediante análisis *in-silico*.

Virulotipificación

La virulotipificación es una prueba rápida para identificar los perfiles de genes de virulencia en *Salmonella*, y con ello mejorar la comprensión del potencial riesgo de infección en el hombre y animales (Khoo y col., 2009). Diversos autores como Elemfareji y col. (2013), Khoo y col. (2009), Huehn y col. (2010), y Fresno y col. (2013), han diseñado y aplicado la virulotipificación en cepas de *Salmonella*, mediante la amplificación por medio de cada uno de los genes de virulencia mediante PCR, para verificar la presencia o ausencia y con ello realizar un perfil de virulencia. Lo cual exhibe un procedimiento laborioso y costoso a nivel de técnica y análisis.

En nuestro estudio, se realizó una virulotipificación *in-silico*, a partir de la construcción de una matriz de genes de virulencia que conforman las islas de patogenicidad de *Salmonella*, y su posterior alineamiento con las cepas sometidas al proceso de secuenciación para la obtención de un dendrograma. La Fig. 21 se muestra las 18 cepas de S. Oranienburg, las cuales presentaron tres perfiles diferentes de virulencia (A, B y C). La Fig. 21 muestra el porcentaje de cepas por isla de patogenicidad.

Los genes de virulencia identificados se relacionaron principalmente como procesos de adherencia, invasión y sobrevivencia de *Salmonella* en las células fagocíticas y epiteliales, lo cual implica la capacidad de S. Oranienburg para causar infección en el hospedero. No obstante, la enfermedad dependerá adicionalmente de factores como el sistema inmunológico del hospedero, y/o factores de predisposición (Lokken y col., 2016).

El perfil-A, perfil-B y perfil-C mostro un 65%, 63% y 60% de genes de virulencia, respectivamente. Primordialmente, la mayoría de los aislados de S. Oranienburg presentaron los genes, de SPI-1 (*hilA*, *orgA*, *sopD*, *ttrB*, *spaS*, *sitA*,

sipA, *invA*, *invI* y *prgH*), la *SPI-2* (*sifA*, *ssrB*, *ttrB*, *ssaQ*, *sseC*, *sseF*, *sseJ* y *ttrC*, *ssek1*) y la *SPI-5* (*pipA*, *pipB*, *pipD*, *sopB*, y *sopD*) cuya función en general son proteínas de invasión (Collier-Hyams y col., 2002; Ben-Barak y col., 2005), sobrevivencia en macrófagos y la enteropatogénesis, respectivamente. Solamente, la cepa S-29 presentó un gen de virulencia adicional *avrA* (cistein proteasa) (perteneciente a la *SPI-1*) y una delección del gen *ssek1* (perteneciente a la *SPI-2*), los cuales se relacionan con la muerte celular programada, así como la inhibición de la producción de IL-18 y respuesta inflamatoria de los hospederos, por lo cual se garantiza la infección de *Salmonella* y la sobrevivencia en macrófagos (Collier-Hyams y col., 2002; Ben-Barak y col., 2005).

En cuanto al resto de las islas de patogenicidad *SPI-3* (función de sobrevivencia a macrófagos) e *SPI-4* (función de adherencia), las cepas de *S. Oranienburg* mostraron diversidad. Brevemente, las cepas presentaron solo el 67% (4/6) y 50% (1/2) de los genes implicados en las *SPI-3* y *SPI-4*, respectivamente. Por otra parte, la *SPI-7* se encontró ausente en todas las muestras, esta isla de patogenicidad está asociada a enfermedades sistémicas. En un estudio realizado por Bueno y col. (2004) encontraron la falta de *SPI-7* en *S. Typhi* y *S. Paratyphi*, lo cual la falta de esta isla hace que pierda la habilidad de producir el antígeno Vi (capsular), sugiriendo que la falta de la *SPI-7* está relacionado con los pasos de estas cepas en el laboratorio. Por otro lado, otros serotipos de *Salmonella* demuestran tener un ordenamiento de genes más conservado lo que indica una estabilidad genómica únicamente en *S. Typhi* (Liu y col., 2006). Por lo cual, se ha propuesto que la inestabilidad en el genoma tiene consecuencias como la inserción de islas de patogenicidad tales como *SPI-7*. En poblaciones de *S. Typhi* se ha observado que sufren reordenamientos espontáneos del genoma, y que las están más cerca de restaurar la polaridad orienter (ya que esta polaridad puede afectar adversamente la replicación) tienen los genomas más estables y los tiempos de generación más rápidos (Liu y col, 2006).

Interesantemente, todas las cepas de *S. Oranienburg* presentaron el gen *sdiA*, relacionado con la percepción de *quórum sensing*, el cual es un mecanismo de regulación de la expresión genética, es decir, es una forma de comunicación dentro de la población de células bacterianas (Almeida y col., 2016). Por otro lado, los genes relacionados con fimbrias, se encontró presente en todas las cepas los siguientes genes: *bcfC*, *csgA*, *fimA*, *stdD*, *stdB*, *stiC*, STM4595. Mientras, que los genes *safC*, *stjB*, *tcfA* se encontró presente en el 99 % de las cepas, excepto la cepa S-29. En un estudio similar realizado por Huehn y col. (2010) encontraron el mismo perfil en común de genes fimbriales (*bcf*, *agf*, *csg*, *fim*, *lpf*, *saf*, *stb*, *stf*, y STM4595), los cuales contribuyen a la colonización de los hospederos.

En general, se observó que la mayoría de las cepas de *S. Oranienburg* presentaron el mismo perfil de virulencia. Sin embargo, la cepa S-29, fue la única que presentó deleciones o adiciones en diferentes genes, lo cual puede explicarse por la presión selectiva del entorno que coloniza (Albalat y Cañestro, 2016; Whitaker, 2006). No obstante, este resultado es congruente con la topología del árbol de genoma completo y el árbol las secuencias consenso de genes de virulencia, confirmando que las cepas de *S. Oranienburg* son clonales, es decir, *S. Oranienburg* es un serotipo conservado.

Resistencia a antibióticos *in silico*

La resistencia a antibióticos es una prueba que permite evaluar la clonalidad en las muestras a partir de un antibiograma que permita determinar los perfiles de resistencia a antibióticos en las cepas denominadas “antibiotipo”, dado que se espera que las cepas clonales presenten el mismo perfil de resistencia (Labarca, 2002).

La resistencia a antibióticos de las 18 cepas de *S. Oranienburg* se analizó mediante el programa *SRST2* (Inouye y col., 2014), el cual realiza mapeos específicos contra la base de datos de genes de resistencia a antibióticos de diversas familias dependiendo del microorganismo, y genera un antibiotipo (Inouye y col., 2014). La Fig. 22 muestra el dendrograma de las cepas de *S. Oranienburg*, en el cual se encontró que el 100 % de las muestras fueron resistentes un antibiótico correspondiente a la familia de aminoglucósidos (*Aac6.laa_AGly*). Basados en el antibiotipo *in silico*, se pudo argumentar que las cepas mantienen de manera conservada la resistencia a aminoglucósidos, dado que al material genético analizado solo incluye ADN cromosomal.

Estos resultados están en concordancia con lo que el Centro de Prevención y Control de Enfermedades de Estados Unidos, es decir, se identifica a la familia de aminoglucósidos como antibióticos a los cuales *Salmonella* exhibe resistencia actualmente. Algunos estudios previos sobre resistencia antimicrobiana en cepas de *S. Oranienburg* han reportado fenotipos de resistencia a ampicilina, cloranfenicol, kanamicina, estreptomina, sulfametoxazol, tetraciclina y ceftriaxona (Kaneene y col., 2010; Yang y col., 2014), lo cual evidencia la dinámica para la adquisición de estrategias de defensa de la bacteria en el hospedero. Particularmente, en la región de Sinaloa, Jiménez y col. (2011) reportaron cepas resistentes de *S. Oranienburg* a ampicilina, neomicina, cloranfenicol y estreptomina aisladas de heces de animales domésticos. Estos

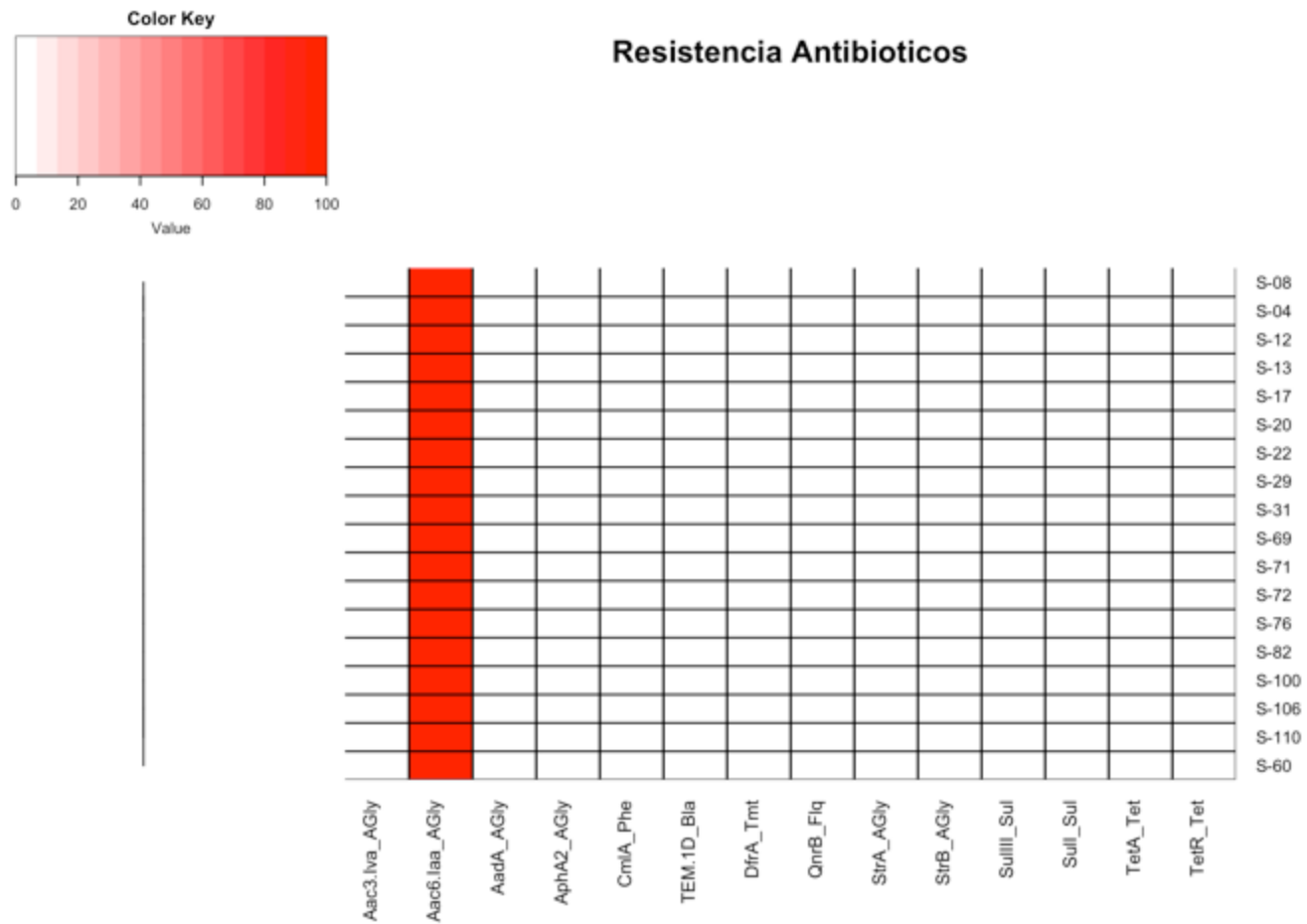


Fig. 22.- Dendrograma de resistencia a antibióticos. Las 18 cepas de *Salmonella* Oranienburg mostraron resistencia a solo una familia de antibióticos mediante el programa SRST2.

hechos plantean la dinámica del serotipo para la adquisición de genes de resistencia antimicrobiana, lo cual pueda dificultar su control.

Finalmente, el tratamiento actual para la salmonelosis no tifoidea se basa en la administración de ampicilina, cloranfenicol y gentamicina, por lo tanto, la resistencia a antibióticos observada en las cepas de estudio no representa un problema de salud pública. No obstante, los genotipos bacterianos son dinámicos ya que estos van cambiando con el tiempo y ecosistema donde se desarrollan (Martín y col., 2003). En este contexto, la resistencia a los antibióticos es un proceso dinámico, ya que el ambiente puede estar contribuyendo a la propagación de cepas resistentes a antibióticos derivada de las actividades antropogénicas (granjas de animales y aguas de desecho) (Tello y col., 2012; Anderson y Hughes, 2014).

IX CONCLUSIONES

1. Los ríos de Culiacán, Sinaloa son fuente de clones de *S. Oranienburg* con homología genética alta a nivel de genoma, serotipo y virulencia.
2. Los clones de *S. Oranienburg* aisladas de aguas de ríos de Culiacán sugieren el riesgo de esta fuente para la transmisión de cepas resistentes y con potencial patogénico.
3. La tipificación por multilocus de secuencia (MLST) es una herramienta útil para conocer la clonalidad de diferentes poblaciones, más no es lo suficientemente detallada como para discernir entre clones de una misma población.
4. La virulotipificación es una herramienta capaz de inferir el potencial patogénico de esta bacteria, e identificar la heterogeneidad de esta facultad.

X BIBLIOGRAFÍA

- Achtman, M.; Wain, J.; Weill, F. X.; Nair, S.; Zhou, Z.; Sangal, V.; Krauland, M. G.; Hale, J. L.; Harbottle, H.; Uesbeck, A.; Dougan, G.; Harrison, L. H.; Brisse, S.; Group, S. E. M. S., Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. *PLoS Pathog* **2012**, *8*, e1002776.
- Albalat, R.; Canestro, C., Evolution by gene loss. *Nat Rev Genet* **2016**, *17*, 379-91.
- Almeida, F. A.; Pinto, U. M.; Vanetti, M. C., Novel insights from molecular docking of SdiA from *Salmonella Enteritidis* and *Escherichia coli* with quorum sensing and quorum quenching molecules. *Microb Pathog* **2016**, *99*, 178-190.
- Álvarez-Ordóñez, A.; Prieto, M.; Bernardo, A.; Hill, C.; López, M., The Acid Tolerance Response of *Salmonella* spp.: An adaptive strategy to survive in stressful environments prevailing in foods and the host. *Food Research International* **2012**, *45*, 482-492.
- Amezquita-Lopez, B. A.; Quinones, B.; Cooley, M. B.; Leon-Felix, J.; Castro-del Campo, N.; Mandrell, R. E.; Jimenez, M.; Chaidez, C., Genotypic analyses of shiga toxin-producing *Escherichia coli* O157 and non-O157 recovered from feces of domestic animals on rural farms in Mexico. *PLoS One* **2012**, *7*, e51565.
- Andersson, D. I.; Hughes, D., Microbiological effects of sublethal levels of antibiotics. *Nat Rev Microbiol* **2014**, *12*, 465-78.
- Andino, A.; Hanning, I., *Salmonella enterica*: survival, colonization, and virulence differences among serovars. *ScientificWorldJournal* **2015**, *2015*, 520179.
- Angiuoli, S. V.; Salzberg, S. L., Mugsy: fast multiple alignment of closely related whole genomes. *Bioinformatics* **2011**, *27*, 334-42.

- Asakura H., W. M., Shirahata T. and Makino S., Viable but Nonculturable Salmonella Species Recovery and Systemic Infection in Morphine-Treated Mice. *the journal of Infectious Diseases* **2002**, *186*, 1526-9.†
- Assefa, S.; Keane, T. M.; Otto, T. D.; Newbold, C.; Berriman, M., ABACAS: algorithm-based automatic contiguation of assembled sequences. *Bioinformatics* **2009**, *25*, 1968-9.
- Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma, K.; Gerdes, S.; Glass, E. M.; Kubal, M.; Meyer, F.; Olsen, G. J.; Olson, R.; Osterman, A. L.; Overbeek, R. A.; McNeil, L. K.; Paarmann, D.; Paczian, T.; Parrello, B.; Pusch, G. D.; Reich, C.; Stevens, R.; Vassieva, O.; Vonstein, V.; Wilke, A.; Zagnitko, O., The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **2008**, *9*, 75.
- Barat, S.; Willer, Y.; Rizos, K.; Claudi, B.; Maze, A.; Schemmer, A. K.; Kirchhoff, D.; Schmidt, A.; Burton, N.; Bumann, D., Immunity to intracellular Salmonella depends on surface-associated antigens. *PLoS Pathog* **2012**, *8*, e1002966.
- Barba, M.; Czosnek, H.; Hadidi, A., Historical perspective, development and applications of next-generation sequencing in plant virology. *Viruses* **2014**, *6*, 106-36.
- Beatus, L.; Joram, B.; Woutrina, S.; Murugan, S.; Douglas, R. C., Surface waters in northern Tanzania harbor fecal coliform and antibiotic resistant Salmonella spp. capable of horizontal gene transfer. *African Journal of Microbiology Research* **2016**, *10*, 348-356.
- Ben-Barak, Z.; Streckel, W.; Yaron, S.; Cohen, S.; Prager, R.; Tschape, H., The expression of the virulence-associated effector protein gene *avrA* is dependent on a Salmonella enterica-specific regulatory function. *Int J Med Microbiol* **2006**, *296*, 25-38.
- Bergey, D. H.; Krieg, N. R.; Holt, J. G., *Bergey's manual of systematic bacteriology*. Williams & Wilkins: Baltimore, MD, 1984.

- Berglund EC., K. A. a. S. A., Next generation sequencing technologies and applications for human genetic history and forensics. *Investigative Genetics* **2011**, 2:23.
- Bertels, F.; Silander, O. K.; Pachkov, M.; Rainey, P. B.; van Nimwegen, E., Automated reconstruction of whole-genome phylogenies from short-sequence reads. *Mol Biol Evol* **2014**, 31, 1077-88.
- Boetzer, M.; Henkel, C. V.; Jansen, H. J.; Butler, D.; Pirovano, W., Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **2011**, 27, 578-9.
- Bolger, A. M.; Lohse, M.; Usadel, B., Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, 30, 2114-20.
- Brenner, F.; Villar, R.; Angulo, F.; Tauxe, R.; Swaminathan, B., Salmonella nomenclature. *Journal of clinical microbiology* **2000**, 38, 2465-2467.
- Bueno, S. M.; Santiviago, C. A.; Murillo, A. A.; Fuentes, J. A.; Trombert, A. N.; Rodas, P. I.; Youderian, P.; Mora, G. C., Precise Excision of the Large Pathogenicity Island, SPI7, in Salmonella enterica Serovar Typhi. *Journal of Bacteriology* **2004**, 186, 3202-3213.
- Burton, P. R.; Tobin, M. D.; Hopper, J. L., Key concepts in genetic epidemiology. *The Lancet* **2005**, 366, 941-951.
- Cárdenas, L.; Sastoque, A.; Peña, M.; Association, I. W. In *Métodos moleculares en muestras ambientales: una alternativa para establecer la relación entre la estructura de la comunidad microbiana y la operación del sistema*, Agua 2003: Usos Múltiples del Agua, para la Vida y el Desarrollo Sostenible, 2003; IWA: 2003; pp 1-8.
- Cecilia, H. C.; Guadalupe, A. A. M.; Graciela, C. E., Situación de las enfermedades gastrointestinales en México. *Enfermedades infecciosas y microbiología* **2011**, 31, 137.

- Chacón, L.; Barrantes, K.; García, C.; Achí, R., Artículo original Estandarización de una PCR para la detección del gen invA de Salmonella spp. en lechuga. *Revista de la Sociedad Venezolana de Microbiología* **2010**, *30*, 18-23.
- Chen, H. M.; Wang, Y.; Su, L. H.; Chiu, C. H., Nontyphoid salmonella infection: microbiology, clinical features, and antimicrobial therapy. *Pediatr Neonatol* **2013**, *54*, 147-52.
- Chen, L.; Yang, J.; Yu, J.; Yao, Z.; Sun, L.; Shen, Y.; Jin, Q., VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res* **2005**, *33*, D325-8.
- Cherayil, B. J.; Antos, D., Inducible nitric oxide synthase and Salmonella infection. *Microbes Infect* **2001**, *3*, 771-6.
- Cock, P. J.; Fields, C. J.; Goto, N.; Heuer, M. L.; Rice, P. M., The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res* **2010**, *38*, 1767-71.
- Coil, D.; Jospin, G.; Darling, A. E., A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* **2015**, *31*, 587-9.
- Collier-Hyams, L. S.; Zeng, H.; Sun, J.; Tomlinson, A. D.; Bao, Z. Q.; Chen, H.; Madara, J. L.; Orth, K.; Neish, A. S., Cutting Edge: Salmonella AvrA Effector Inhibits the Key Proinflammatory, Anti-Apoptotic NF- B Pathway. *The Journal of Immunology* **2002**, *169*, 2846-2850.
- Commins, J.; Toft, C.; Fares, M. A., Computational biology methods and their application to the comparative genomics of endocellular symbiotic bacteria of insects. *Biol Proced Online* **2009**, *11*, 52-78.
- Compeau, P. E.; Pevzner, P. A.; Tesler, G., How to apply de Bruijn graphs to genome assembly. *Nat Biotechnol* **2011**, *29*, 987-91.
- Control, C. f. D.; Prevention, Standard operating procedure for PulseNet PFGE of Escherichia coli O157: H7, Escherichia coli non-O157 (STEC), Salmonella

- serotypes, *Shigella sonnei* and *Shigella flexneri*. *Centers for Disease Control and Prevention, Atlanta* **2013**.
- Cook, K. A.; Dobbs, T. E.; Hlady, W. G.; Wells, J. G.; Barrett, T. J.; Puh, N. D.; Lancette, G. A.; Bodager, D. W.; Toth, B. L.; Genese, C. A.; Highsmith, A. K.; Pilot, K. E.; Finelli, L.; Swerdlow, D. L., Outbreak of *Salmonella* serotype Hartford infections associated with unpasteurized orange juice. *JAMA* **1998**, *280*, 1504-9.
- Cooper, J. E.; Feil, E. J., Multilocus sequence typing--what is resolved? *Trends Microbiol* **2004**, *12*, 373-7.
- Darling, A. E.; Mau, B.; Perna, N. T., progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* **2010**, *5*, e11147.
- Davis, M. A.; Hancock, D. D.; Besser, T. E.; Call, D. R., Evaluation of Pulsed-Field Gel Electrophoresis as a Tool for Determining the Degree of Genetic Relatedness between Strains of *Escherichia coli* O157:H7. *Journal of Clinical Microbiology* **2003**, *41*, 1843-1849.
- Day-Williams, A. G.; Zeggini, E., The effect of next-generation sequencing technology on complex trait research. *Eur J Clin Invest* **2011**, *41*, 561-7.
- den Bakker, H. C.; Switt, A. I. M.; Govoni, G.; Cummings, C. A.; Ranieri, M. L.; Degoricija, L.; Hoelzer, K.; Rodriguez-Rivera, L. D.; Brown, S.; Bolchacova, E., Genome sequencing reveals diversification of virulence factor content and possible host adaptation in distinct subpopulations of *Salmonella enterica*. *BMC genomics* **2011**, *12*, 425.
- Dobrindt, U.; Hochhut, B.; Hentschel, U.; Hacker, J., Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* **2004**, *2*, 414-24.

- Doerks, T.; Copley, R. R.; Schultz, J.; Ponting, C. P.; Bork, P., Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res* **2002**, *12*, 47-56.
- Drecktrah, D.; Knodler, L. A.; Ireland, R.; Steele-Mortimer, O., The mechanism of Salmonella entry determines the vacuolar environment and intracellular gene expression. *Traffic* **2006**, *7*, 39-51.
- Drecktrah, D.; Levine-Wilkinson, S.; Dam, T.; Winfree, S.; Knodler, L. A.; Schroer, T. A.; Steele-Mortimer, O., Dynamic behavior of Salmonella-induced membrane tubules in epithelial cells. *Traffic* **2008**, *9*, 2117-29.
- Ecker, D. J.; Sampath, R.; Willett, P.; Wyatt, J. R.; Samant, V.; Massire, C.; Hall, T. A.; Hari, K.; McNeil, J. A.; Buchen-Osmond, C.; Budowle, B., The Microbial Rosetta Stone Database: a compilation of global and emerging infectious microorganisms and bioterrorist threat agents. *BMC Microbiol* **2005**, *5*, 19.
- Ekblom, R.; Wolf, J. B., A field guide to whole-genome sequencing, assembly and annotation. *Evol Appl* **2014**, *7*, 1026-42.
- Elemfareji, O. I.; Thong, K. L., Comparative Virulotyping of Salmonella typhi and Salmonella enteritidis. *Indian J Microbiol* **2013**, *53*, 410-7.
- Estrada-Acosta, M.; Jimenez, M.; Chaidez, C.; Leon-Felix, J.; Castro-Del Campo, N., Irrigation water quality and the benefits of implementing good agricultural practices during tomato (*Lycopersicum esculentum*) production. *Environ Monit Assess* **2014**, *186*, 4323-30.
- Fakhr, M. K.; Nolan, L. K.; Logue, C. M., Multilocus sequence typing lacks the discriminatory ability of pulsed-field gel electrophoresis for typing Salmonella enterica serovar Typhimurium. *J Clin Microbiol* **2005**, *43*, 2215-9.
- Farmer, J.; Davis, B. R.; Hickman-Brenner, F.; McWhorter, A.; Huntley-Carter, G.; Asbury, M.; Riddle, C.; Wathen-Grady, H.; Elias, C.; Fanning, G.,

- Biochemical identification of new species and biogroups of Enterobacteriaceae isolated from clinical specimens. *Journal of clinical microbiology* **1985**, *21*, 46-76.
- Fazal, M. A.; Alexander, S.; Burnett, E.; Deheer-Graham, A.; Oliver, K.; Holroyd, N.; Parkhill, J.; Russell, J. E., Complete Whole-Genome Sequence of *Salmonella enterica* subsp. *enterica* Serovar Java NCTC5706. *Genome Announc* **2016**, *4*.
- Feil, E. J.; Li, B. C.; Aanensen, D. M.; Hanage, W. P.; Spratt, B. G., eBURST: Inferring Patterns of Evolutionary Descent among Clusters of Related Bacterial Genotypes from Multilocus Sequence Typing Data. *Journal of Bacteriology* **2004**, *186*, 1518-1530.
- Fresno, M.; Barrera, V.; Gornall, V.; Lillo, P.; Paredes, N.; Abalos, P.; Fernandez, A.; Retamal, P., Identification of diverse *Salmonella* serotypes, virulotypes, and antimicrobial resistance phenotypes in waterfowl from Chile. *Vector Borne Zoonotic Dis* **2013**, *13*, 884-7.
- Galán, J. C.; Baquero, M. R.; Morosini, M. I., Bacterias con alta tasa de mutación: los riesgos de una vida acelerada. *Infectio* **2011**, *10*.
- Galardini, M.; Biondi, E. G.; Bazzicalupo, M.; Mengoni, A., CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med* **2011**, *6*, 11.
- Garaizar, J.; López-Molina, N.; Laconcha, I.; Baggesen, D. L.; Rementeria, A.; Vivanco, A.; Audicana, A.; Perales, I., Suitability of PCR fingerprinting, infrequent-restriction-site PCR, and pulsed-field gel electrophoresis, combined with computerized gel analysis, in library typing of *Salmonella enterica* serovar Enteritidis. *Applied and environmental microbiology* **2000**, *66*, 5273-5281.
- García, J.; Paniagua, J.; Pelayo, R.; Isibasi, A.; Kumate, J., Factores de virulencia de *Salmonella typhi* en relación al desarrollo de nuevas vacunas. *Salud Pública de México* **1992**, *34*, 262-267.

- Gebert, A., The role of M cells in the protection of mucosal membranes. *Histochem Cell Biol* **1997**, *108*, 455-70.
- Gnerre, S.; Maccallum, I.; Przybylski, D.; Ribeiro, F. J.; Burton, J. N.; Walker, B. J.; Sharpe, T.; Hall, G.; Shea, T. P.; Sykes, S.; Berlin, A. M.; Aird, D.; Costello, M.; Daza, R.; Williams, L.; Nicol, R.; Gnirke, A.; Nusbaum, C.; Lander, E. S.; Jaffe, D. B., High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A* **2011**, *108*, 1513-8.
- Gong, H.; Vu, G. P.; Bai, Y.; Chan, E.; Wu, R.; Yang, E.; Liu, F.; Lu, S., A Salmonella small non-coding RNA facilitates bacterial invasion and intracellular replication by modulating the expression of virulence factors. *PLoS Pathog* **2011**, *7*, e1002120.
- Gonzales, L., Salmonella abortus ovis infection. *WB Martin and ID Aitken. Malden, MA: Blackwell Science* **2000**, 102-7.
- Gouy, M.; Guindon, S.; Gascuel, O., SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **2010**, *27*, 221-4.
- Graziani, C.; Busani, L.; Dionisi, A. M.; Caprioli, A.; Ivarsson, S.; Hedenstrom, I.; Luzzi, I., Virulotyping of Salmonella enterica serovar Napoli strains isolated in Italy from human and nonhuman sources. *Foodborne Pathog Dis* **2011**, *8*, 997-1003.
- Greig, J. D.; Lee, M. B.; Harris, J. E., Review of enteric outbreaks in prisons: effective infection control interventions. *Public Health* **2011**, *125*, 222-8.
- Guerra, B.; Laconcha, I.; Soto, S. M.; González-Hevia, M. Á.; Mendoza, M. C., Molecular characterisation of emergent multiresistant Salmonella enterica serotype [4, 5, 12: i:-] organisms causing human salmonellosis. *FEMS microbiology letters* **2000**, *190*, 341-347.

- Gunn, J. S.; Marshall, J. M.; Baker, S.; Dongol, S.; Charles, R. C.; Ryan, E. T., Salmonella chronic carriage: epidemiology, diagnosis, and gallbladder persistence. *Trends Microbiol* **2014**, *22*, 648-55.
- Guo, X.; Chen, J.; Brackett, R. E.; Beuchat, L. R., Survival of Salmonellae on and in tomato plants from the time of inoculation at flowering and early stages of fruit development through fruit ripening. *Applied and Environmental Microbiology* **2001**, *67*, 4760-4764.
- Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G., QCAST: quality assessment tool for genome assemblies. *Bioinformatics* **2013**, *29*, 1072-5.
- Gutiérrez-Cogco, L.; Montiel-Vázquez, E.; Aguilera-Pérez, P.; del Carmen González-Andrade, M., Serotipos de Salmonella identificados en los servicios de salud de México. *Salud pública de México* **2000**, *42*, 490-495.
- Haase, J. K.; Didelot, X.; Lecuit, M.; Korkeala, H.; Group, L. m. M. S.; Achtman, M., The ubiquitous nature of *Listeria monocytogenes* clones: a large-scale Multilocus Sequence Typing study. *Environ Microbiol* **2014**, *16*, 405-16.
- Haley, B. J.; Cole, D. J.; Lipp, E. K., Distribution, diversity, and seasonality of waterborne salmonellae in a rural watershed. *Appl Environ Microbiol* **2009**, *75*, 1248-55.
- Hall, N., Advanced sequencing technologies and their wider impact in microbiology. *J Exp Biol* **2007**, *210*, 1518-25.
- Harbottle, H.; White, D. G.; McDermott, P. F.; Walker, R. D.; Zhao, S., Comparison of multilocus sequence typing, pulsed-field gel electrophoresis, and antimicrobial susceptibility typing for characterization of *Salmonella enterica* serotype Newport isolates. *J Clin Microbiol* **2006**, *44*, 2449-57.
- Haznedaroglu, B. Z.; Yates, M. V.; Maduro, M. F.; Walker, S. L., Effects of residual antibiotics in groundwater on *Salmonella typhimurium*: changes in antibiotic resistance, in vivo and in vitro pathogenicity. *J Environ Monit* **2012**, *14*, 41-7.

- Head, S. R.; Komori, H. K.; LaMere, S. A.; Whisenant, T.; Van Nieuwerburgh, F.; Salomon, D. R.; Ordoukhanian, P., Library construction for next-generation sequencing: overviews and challenges. *Biotechniques* **2014**, *56*, 61-4, 66, 68, passim.
- Heithoff, D. M.; Shimp, W. R.; Lau, P. W.; Badie, G.; Enioutina, E. Y.; Daynes, R. A.; Byrne, B. A.; House, J. K.; Mahan, M. J., Human Salmonella clinical isolates distinct from those of animal origin. *Appl Environ Microbiol* **2008**, *74*, 1757-66.
- Hensel, M., Evolution of pathogenicity islands of Salmonella enterica. *Int J Med Microbiol* **2004**, *294*, 95-102.
- Hiriart, Y.; Serradell, M.; Martínez, A.; Sampaolesi, S.; Maciel, D. G.; Chabalgoity, J. A.; Yim, L.; Algorta, G.; Rumbo, M., Generation and selection of anti-flagellin monoclonal antibodies useful for serotyping Salmonella enterica. *SpringerPlus* **2013**, *2*, 640.
- Holley, R. A.; Arrus, K. M.; Ominski, K. H.; Tenuta, M.; Blank, G., Survival in Manure-Treated Soils during Simulated Seasonal Temperature Exposure. *Journal of Environmental Quality* **2006**, *35*, 1170-1180.
- Hurley, D.; McCusker, M. P.; Fanning, S.; Martins, M., Salmonella-host interactions - modulation of the host innate immune system. *Front Immunol* **2014**, *5*, 481.
- Hyytiä-Trees, E.; Smole, S. C.; Fields, P. A.; Swaminathan, B.; Ribot, E. M., Second generation subtyping: a proposed PulseNet protocol for multiple-locus variable-number tandem repeat analysis of Shiga toxin-producing Escherichia coli O157 (STEC O157). *Foodborne Pathogens & Disease* **2006**, *3*, 118-131.
- Ibarra, J. A.; Steele-Mortimer, O., Salmonella--the ultimate insider. Salmonella virulence factors that modulate intracellular survival. *Cell Microbiol* **2009**, *11*, 1579-86.

- Inouye, M.; Dashnow, H.; Raven, L.-A.; Schultz, M. B.; Pope, B. J.; Tomita, T.; Zobel, J.; Holt, K. E., SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome medicine* **2014**, *6*, 90.
- Isea, R., The Present-Day Meaning of the Word Bioinformatics. *Global Journal of Advanced Research* **2015**, *2*, 70-73.
- Islam, M.; Morgan, J.; Doyle, M. P.; Phatak, S. C.; Millner, P.; Jiang, X., Fate of *Salmonella enterica* Serovar Typhimurium on Carrots and Radishes Grown in Fields Treated with Contaminated Manure Composts or Irrigation Water. *Applied and Environmental Microbiology* **2004**, *70*, 2497-2502.
- Jacobsen, C. S.; Bech, T. B., Soil survival of *Salmonella* and transfer to freshwater and fresh produce. *Food Research International* **2012**, *45*, 557-566.
- Jantsch, J.; Chikkaballi, D.; Hensel, M., Cellular aspects of immunity to intracellular *Salmonella enterica*. *Immunological reviews* **2011**, *240*, 185-195.
- Jimenez, M.; Martinez-Urtaza, J.; Chaidez, C., Geographical and temporal dissemination of salmonellae isolated from domestic animal hosts in the Culiacan Valley, Mexico. *Microb Ecol* **2011**, *61*, 811-20.
- Jimenez, M.; Martinez-Urtaza, J.; Rodriguez-Alvarez, M. X.; Leon-Felix, J.; Chaidez, C., Prevalence and genetic diversity of *Salmonella* spp. in a river in a tropical environment in Mexico. *J Water Health* **2014**, *12*, 874-84.
- Johnson, J. K.; Arduino, S. M.; Stine, O. C.; Johnson, J. A.; Harris, A. D., Multilocus sequence typing compared to pulsed-field gel electrophoresis for molecular typing of *Pseudomonas aeruginosa*. *J Clin Microbiol* **2007**, *45*, 3707-12.
- Jokinen, C.; Edge, T. A.; Ho, S.; Koning, W.; Laing, C.; Mauro, W.; Medeiros, D.; Miller, J.; Robertson, W.; Taboada, E.; Thomas, J. E.; Topp, E.; Ziebell, K.; Gannon, V. P., Molecular subtypes of *Campylobacter* spp., *Salmonella enterica*, and *Escherichia coli* O157:H7 isolated from faecal and surface

- water samples in the Oldman River watershed, Alberta, Canada. *Water Res* **2011**, *45*, 1247-57.
- Jones, B. D.; Falkow, S., SALMONELLOSIS: Host Immune Responses and Bacterial Virulence Determinants¹. *Annual review of immunology* **1996**, *14*, 533-561.
- Jurado Jiménez, R.; Arenas Muñoz, C.; Doblaz Delgado, A.; Rivero, A.; Torre Cisneros, J., Fiebre tifoidea y otras infecciones por salmonellas. *Medicine* **2010**, *10*, 3497-501.
- Kaneene, J. B.; Miller, R.; May, K.; Hattey, J. A., An outbreak of multidrug-resistant *Salmonella enterica* serotype Oranienburg in Michigan dairy calves. *Foodborne pathogens and disease* **2010**, *7*, 1193-1201.
- Khoo, C. H.; Cheah, Y. K.; Lee, L. H.; Sim, J. H.; Salleh, N. A.; Sidik, S. M.; Radu, S.; Sukardi, S., Virulotyping of *Salmonella enterica* subsp. *enterica* isolated from indigenous vegetables and poultry meat in Malaysia using multiplex-PCR. *Antonie Van Leeuwenhoek* **2009**, *96*, 441-57.
- Kidgell, C.; Reichard, U.; Wain, J.; Linz, B.; Torpdahl, M.; Dougan, G.; Achtman, M., *Salmonella typhi*, the causative agent of typhoid fever, is approximately 50,000 years old. *Infect Genet Evol* **2002**, *2*, 39-45.
- Kirk, M. D.; Pires, S. M.; Black, R. E.; Caipo, M.; Crump, J. A.; Devleeschauwer, B.; Dopfer, D.; Fazil, A.; Fischer-Walker, C. L.; Hald, T.; Hall, A. J.; Keddy, K. H.; Lake, R. J.; Lanata, C. F.; Torgerson, P. R.; Havelaar, A. H.; Angulo, F. J., World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLoS Med* **2015**, *12*, e1001921.
- Kiss, T.; Morgan, E.; Nagy, G., Contribution of SPI-4 genes to the virulence of *Salmonella enterica*. *FEMS Microbiol Lett* **2007**, *275*, 153-9.
- Kolmogorov, M.; Raney, B.; Paten, B.; Pham, S., Ragout-a reference-assisted assembly tool for bacterial genomes. *Bioinformatics* **2014**, *30*, i302-9.

- Kondrashov, F. A., Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc Biol Sci* **2012**, 279, 5048-57.
- Koskiniemi, S.; Sun, S.; Berg, O. G.; Andersson, D. I., Selection-driven gene loss in bacteria. *PLoS Genet* **2012**, 8, e1002787.
- Kotetishvili, M.; Stine, O. C.; Kreger, A.; Morris, J. G.; Sulakvelidze, A., Multilocus Sequence Typing for Characterization of Clinical and Environmental Salmonella Strains. *Journal of Clinical Microbiology* **2002**, 40, 1626-1635.
- Kozarewa, I.; Ning, Z.; Quail, M. A.; Sanders, M. J.; Berriman, M.; Turner, D. J., Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat Methods* **2009**, 6, 291-5.
- Kubota, K.; Barrett, T. J.; Ackers, M. L.; Brachman, P. S.; Mintz, E. D., Analysis of Salmonella enterica serotype Typhi pulsed-field gel electrophoresis patterns associated with international travel. *J Clin Microbiol* **2005**, 43, 1205-9.
- LABARCA L., J., Utilización del antibiotipo como marcador epidemiológico en infecciones intrahospitalarias: Comparación con la epidemiología molecular. *Revista chilena de infectología* **2002**, 19, 157-160.
- Labbe, G.; Ziebell, K.; Bekal, S.; Macdonald, K. A.; Parmley, E. J.; Agunos, A.; Desruisseau, A.; Daignault, D.; Slavic, D.; Hoang, L.; Ramsay, D.; Pollari, F.; Robertson, J.; Nash, J. H.; Johnson, R. P., Complete Genome Sequences of 17 Canadian Isolates of Salmonella enterica subsp. enterica Serovar Heidelberg from Human, Animal, and Food Sources. *Genome Announc* **2016**, 4.
- Lahiri, A.; Lahiri, A.; Iyer, N.; Das, P.; Chakravorty, D., Visiting the cell biology of Salmonella infection. *Microbes Infect* **2010**, 12, 809-18.
- Lan, R.; Reeves, P. R.; Octavia, S., Population structure, origins and evolution of major Salmonella enterica clones. *Infect Genet Evol* **2009**, 9, 996-1005.

- Li, W.; Raoult, D.; Fournier, P. E., Bacterial strain typing in the genomic era. *FEMS Microbiol Rev* **2009**, *33*, 892-916.
- Li, Y.-L.; Weng, J.-C.; Hsiao, C.-C.; Chou, M.-T.; Tseng, C.-W.; Hung, J.-H., PEAT: an intelligent and efficient paired-end sequencing adapter trimming algorithm. *BMC bioinformatics* **2015**, *16*, S2.
- Lim, S.; Yun, J.; Yoon, H.; Park, C.; Kim, B.; Jeon, B.; Kim, D.; Ryu, S., Mlc regulation of Salmonella pathogenicity island I gene expression via hilE repression. *Nucleic Acids Res* **2007**, *35*, 1822-32.
- Liu, B.; Knirel, Y. A.; Feng, L.; Perepelov, A. V.; Senchenkova, S. N.; Reeves, P. R.; Wang, L., Structural diversity in Salmonella O antigens and its genetic basis. *FEMS Microbiol Rev* **2014**, *38*, 56-89.
- Liu, B.; Pop, M., ARDB--Antibiotic Resistance Genes Database. *Nucleic Acids Res* **2009**, *37*, D443-7.
- Liu, L.; Li, Y.; Li, S.; Hu, N.; He, Y.; Pong, R.; Lin, D.; Lu, L.; Law, M., Comparison of next-generation sequencing systems. *J Biomed Biotechnol* **2012**, *2012*, 251364.
- Liu, L.; Tan, S.; Jun, W.; Smith, A.; Meng, J.; Bhagwat, A. A., Osmoregulated periplasmic glucans are needed for competitive growth and biofilm formation by Salmonella enterica serovar Typhimurium in leafy-green vegetable wash waters and colonization in mice. *FEMS Microbiol Lett* **2009**, *292*, 13-20.
- López Cuevas, O.; León Félix, J.; Jiménez Edeza, M.; Chaidez Quiroz, C., Detección y resistencia a antibióticos de Escherichia coli y Salmonella en agua y suelo agrícola. *Revista fitotecnia mexicana* **2009**, *32*, 119-126.
- Lynch, M. F.; Tauxe, R. V.; Hedberg, C. W., The growing burden of foodborne outbreaks due to contaminated fresh produce: risks and opportunities. *Epidemiol Infect* **2009**, *137*, 307-15.

- Maiden, M. C., Multilocus sequence typing of bacteria. *Annu Rev Microbiol* **2006**, *60*, 561-88.
- Maiden, M. C.; Jansen van Rensburg, M. J.; Bray, J. E.; Earle, S. G.; Ford, S. A.; Jolley, K. A.; McCarthy, N. D., MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat Rev Microbiol* **2013**, *11*, 728-36.
- Mąka, Ł.; Popowska, M., Antimicrobial resistance of Salmonella spp. isolated from food. *ROCZNIKI PAŃSTWOWEGO ZAKŁADU HIGIENY* **2016**, *67*, 343-358.
- Malorny, B.; Hoorfar, J.; Bunge, C.; Helmuth, R., Multicenter Validation of the Analytical Accuracy of Salmonella PCR: towards an International Standard. *Applied and Environmental Microbiology* **2003**, *69*, 290-296.
- Marcus, S. L.; Brumell, J. H.; Pfeifer, C. G.; Finlay, B. B., Salmonella pathogenicity islands: big virulence in small packages. *Microbes and Infection* **2000**, *2*, 145-156.
- Mardis, E. R., The impact of next-generation sequencing technology on genetics. *Trends Genet* **2008**, *24*, 133-41.
- Martinez-Gamboa, A.; Silva, C.; Fernandez-Mora, M.; Wiesner, M.; Ponce de Leon, A.; Calva, E., IS200 and multilocus sequence typing for the identification of Salmonella enterica serovar Typhi strains from Indonesia. *Int Microbiol* **2015**, *18*, 99-104.
- McQuiston, J. R.; Parrenas, R.; Ortiz-Rivera, M.; Gheesling, L.; Brenner, F.; Fields, P. I., Sequencing and Comparative Analysis of Flagellin Genes *fliC*, *fliB*, and *flpA* from Salmonella. *Journal of Clinical Microbiology* **2004**, *42*, 1923-1932.
- McWhorter, A. R.; Chousalkar, K. K., Comparative phenotypic and genotypic virulence of Salmonella strains isolated from Australian layer farms. *Front Microbiol* **2015**, *6*, 12.

- Medini, D.; Serruto, D.; Parkhill, J.; Relman, D. A.; Donati, C.; Moxon, R.; Falkow, S.; Rappuoli, R., Microbiology in the post-genomic era. *Nat Rev Microbiol* **2008**, *6*, 419-30.
- Medrano-Felix, A.; Estrada-Acosta, M.; Jimenez, M.; Gomez-Gil, B.; Leon-Felix, J.; Amarillas, L.; Chaidez, C., Draft Genome Sequence of *Salmonella enterica* subsp. *enterica* Serotype Oranienburg Strain S-76, Isolated from an Aquatic Environment. *Genome Announc* **2013**, *1*.
- Miliotis, M. D.; Bier, J. W., *International handbook of foodborne pathogens*. CRC Press: 2003; Vol. 125.
- Miller, J. R.; Koren, S.; Sutton, G., Assembly algorithms for next-generation sequencing data. *Genomics* **2010**, *95*, 315-27.
- Moore, B. C.; Martinez, E.; Gay, J. M.; Rice, D. H., Survival of *Salmonella enterica* in Freshwater and Sediments and Transmission by the Aquatic Midge *Chironomus tentans* (Chironomidae: Diptera). *Applied and Environmental Microbiology* **2003**, *69*, 4556-4560.
- Moreau, M. R.; Wijetunge, D. S.; Kurundu Hewage, E. M.; Jayarao, B. M.; Kariyawasam, S., Genome Sequences of Two Strains of *Salmonella enterica* Serovar Enteritidis Isolated from Shell Eggs. *Genome Announc* **2015**, *3*.
- Myers, E. W.; Sutton, G. G.; Delcher, A. L.; Dew, I. M.; Fasulo, D. P.; Flanigan, M. J.; Kravitz, S. A.; Mobarry, C. M.; Reinert, K. H.; Remington, K. A., A whole-genome assembly of *Drosophila*. *Science* **2000**, *287*, 2196-2204.
- Nagarajan, N.; Pop, M., Sequence assembly demystified. *Nat Rev Genet* **2013**, *14*, 157-67.
- Nallapareddy, S. R.; Duh, R. W.; Singh, K. V.; Murray, B. E., Molecular Typing of Selected *Enterococcus faecalis* Isolates: Pilot Study Using Multilocus Sequence Typing and Pulsed-Field Gel Electrophoresis. *Journal of Clinical Microbiology* **2002**, *40*, 868-876.

- Nguyen, S. V.; Harhay, D. M.; Bono, J. L.; Smith, T. P.; Fields, P. I.; Dinsmore, B. A.; Santovenia, M.; Kelley, C. M.; Wang, R.; Bosilevac, J. M.; Harhay, G. P., Complete, Closed Genome Sequences of 10 *Salmonella enterica* subsp. *enterica* Serovar Typhimurium Strains Isolated from Human and Bovine Sources. *Genome Announc* **2016**, *4*.
- Ochman, H.; Lawrence, J. G.; Groisman, E. A., Lateral gene transfer and the nature of bacterial innovation. *Nature* **2000**, *405*, 299-304.
- Ochoa, I. M. F.; Rodríguez, A. V., Mecanismos moleculares de patogenicidad de *Salmonella* sp. *Revista latinoamericana de microbiología* **2005**, *47*, 25-42
- Olmos, A. F.; de la Fuente, C. G.; Nieto, J. A. S.; Ramos, S. V., Metodos de identificación bacteriana en el laboratorio de microbiología.
- Ong, S. Y.; Pratap, C. B.; Wan, X.; Hou, S.; Abdul Rahman, A. Y.; Saito, J. A.; Nath, G.; Alam, M., Complete genome sequence of *Salmonella enterica* subsp. *enterica* serovar Typhi P-stx-12. *J Bacteriol* **2012**, *194*, 2115-6.
- Ovchinnikova, O. G.; Rozalski, A.; Liu, B.; Knirel, Y. A., O-antigens of bacteria of the genus *Providencia*: structure, serology, genetics, and biosynthesis. *Biochemistry (Mosc)* **2013**, *78*, 798-817.
- Pan, Z.; Carter, B.; Nunez-Garcia, J.; Abuoun, M.; Fookes, M.; Ivens, A.; Woodward, M. J.; Anjum, M. F., Identification of genetic and phenotypic differences associated with prevalent and non-prevalent *Salmonella* Enteritidis phage types: analysis of variation in amino acid transport. *Microbiology* **2009**, *155*, 3200-13.
- Parra, M.; Durango, J.; Máttar, S., Microbiología, patogénesis, epidemiología, clínica y diagnóstico de las infecciones producidas por *Salmonella*. *Revista MVZ Córdoba* **2002**, *7*, 187-200.
- Pedraza, J. B. G.; Varela, Z. S.; Hernández, E.; Camacho, J. V., Aislamiento de *Salmonella* spp y herramientas moleculares para su detección. *Salud Uninorte* **2014**, *30*.

- Peng, Y.; Leung, H. C.; Yiu, S. M.; Chin, F. Y., IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **2012**, *28*, 1420-8.
- Perez-Losada, M.; Cabezas, P.; Castro-Nallar, E.; Crandall, K. A., Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. *Infect Genet Evol* **2013**, *16*, 38-53.
- Perron, G. G.; Quessy, S.; Bell, G., A reservoir of drug-resistant pathogenic bacteria in asymptomatic hosts. *PLoS One* **2008**, *3*, e3749.
- Pevzner, P. A.; Tang, H.; Waterman, M. S., An Eulerian path approach to DNA fragment assembly. *Proceedings of the National Academy of Sciences* **2001**, *98*, 9748-9753.
- Pfeifer, S. P., From next-generation resequencing reads to a high-quality variant data set. *Heredity (Edinb)* **2017**, *118*, 111-124.
- Pignato, S.; Giammanco, G.; Santangelo, C.; Giammanco, G., Endemic presence of *Salmonella bongori* 48: z35:-causing enteritis in children in Sicily. *Research in microbiology* **1998**, *149*, 429-431.
- Pires, S. M.; Evers, E. G.; van Pelt, W.; Ayers, T.; Scallan, E.; Angulo, F. J.; Havelaar, A.; Hald, T., Attributing the human disease burden of foodborne infections to specific sources. *Foodborne Pathogens and Disease* **2009**, *6*, 417-424.
- Popoff, M.; Le Minor, L., Antigenic formulas of the *Salmonella* serovars, WHO Collaborating Centre for Reference and Research on *Salmonella*. *World Health Organ* **2001**.
- Popoff, M. Y.; Bockemühl, J.; Hickman-Brenner, F. W., Supplément 1996 (no. 40) to the Kauffmann-White scheme. *Research in Microbiology* **1997**, *148*, 811-814.

- Pui, C. F.; Wong, W. C.; Chai, L. C.; Robin, T.; Ponniah, J.; Sahroni, M.; Hidayah, N.; Anyi, U.; Mohamad Ghazali, F.; Cheah, Y. K., Salmonella: A foodborne pathogen. *International Food Research Journal* **2011**, *18*, 465-473.
- Quail, M. A.; Kozarewa, I.; Smith, F.; Scally, A.; Stephens, P. J.; Durbin, R.; Swerdlow, H.; Turner, D. J., A large genome center's improvements to the Illumina sequencing system. *Nat Methods* **2008**, *5*, 1005-10.
- Quick, J.; Ashton, P.; Calus, S.; Chatt, C.; Gossain, S.; Hawker, J.; Nair, S.; Neal, K.; Nye, K.; Peters, T.; De Pinna, E.; Robinson, E.; Struthers, K.; Webber, M.; Catto, A.; Dallman, T. J.; Hawkey, P.; Loman, N. J., Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. *Genome Biol* **2015**, *16*, 114.
- Raza, A.; Muhammad, G.; Rahman, S.; Rashid, I.; Hanif, K.; Atta, A.; Sharif, S., Evaluation of humoral immunity and protective efficacy of biofilm producing Staphylococcus aureus bacterin-toxoid prepared from a bovine mastitis isolate in rabbit. *Iranian Journal of Veterinary Research* **2015**, *16*, 69.
- Reeves, P., Evolution of Salmonella O antigen variation by interspecific gene transfer on a large scale. *Trends in Genetics* **1993**, *9*, 17-22.
- Reeves, P.; Wang, L., Genomic organization of LPS-specific loci. In *Pathogenicity islands and the evolution of pathogenic microbes*, Springer: 2002; pp 109-135.
- Retamal, P.; Fresno, M.; Dougnac, C.; Gutierrez, S.; Gornall, V.; Vidal, R.; Vernal, R.; Pujol, M.; Barreto, M.; Gonzalez-Acuna, D.; Abalos, P., Genetic and phenotypic evidence of the Salmonella enterica serotype Enteritidis human-animal interface in Chile. *Front Microbiol* **2015**, *6*, 464.
- Ribeiro-Goncalves, B.; Francisco, A. P.; Vaz, C.; Ramirez, M.; Carrico, J. A., PHYLOViZ Online: web-based tool for visualization, phylogenetic inference, analysis and sharing of minimum spanning trees. *Nucleic Acids Res* **2016**, *44*, W246-51.

- Rodríguez-Santiago, B.; Armengol, L., Tecnologías de secuenciación de nueva generación en diagnóstico genético pre- y postnatal. *Diagnóstico Prenatal* **2012**, *23*, 56-66.
- Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D. L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M. A.; Huelsenbeck, J. P., MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **2012**, *61*, 539-42.
- Ruby, T.; McLaughlin, L.; Gopinath, S.; Monack, D., Salmonella's long-term relationship with its host. *FEMS Microbiol Rev* **2012**, *36*, 600-15.
- Salipante, S. J.; SenGupta, D. J.; Cummings, L. A.; Land, T. A.; Hoogestraat, D. R.; Cookson, B. T., Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology. *J Clin Microbiol* **2015**, *53*, 1072-9.
- Sangal, V.; Harbottle, H.; Mazzoni, C. J.; Helmuth, R.; Guerra, B.; Didelot, X.; Paglietti, B.; Rabsch, W.; Brisse, S.; Weill, F. X.; Roumagnac, P.; Achtman, M., Evolution and population structure of Salmonella enterica serovar Newport. *J Bacteriol* **2010**, *192*, 6465-76.
- Santander, J. M.; Roland, K. L.; Roy Curtiss, I., Regulation of Vi capsular polysaccharide synthesis in Salmonella enterica serotype Typhi. *Journal of infection in developing countries* **2008**, *2*, 412.
- Schmidt, H.; Hensel, M., Pathogenicity Islands in Bacterial Pathogenesis. *Clinical Microbiology Reviews* **2004**, *17*, 14-56.
- Seth-Smith, H. M., SPI-7: Salmonella's Vi-encoding pathogenicity island. *The Journal of Infection in Developing Countries* **2008**, *2*, 267-271.
- Simpson, J. T.; Durbin, R., Efficient de novo assembly of large genomes using compressed data structures. *Genome Res* **2012**, *22*, 549-56.
- Simpson, J. T.; Wong, K.; Jackman, S. D.; Schein, J. E.; Jones, S. J.; Birol, I., ABySS: a parallel assembler for short read sequence data. *Genome Res* **2009**, *19*, 1117-23.

- Sims, D.; Sudbery, I.; Illott, N. E.; Heger, A.; Ponting, C. P., Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet* **2014**, *15*, 121-32.
- Sommer, M. O.; Dantas, G.; Church, G. M., Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* **2009**, *325*, 1128-31.
- Spratt, B. G.; Hanage, W. P.; Li, B.; Aanensen, D. M.; Feil, E. J., Displaying the relatedness among isolates of bacterial species -- the eBURST approach. *FEMS Microbiol Lett* **2004**, *241*, 129-34.
- Toro, E. S.; O Miró, A.; Ugarte, C. I.; Larrea, F. J., Electroforesis de campo eléctrico pulsado en la tipificación molecular de cepas aisladas de un brote de Salmonelosis ocasionado por consumo de queso contaminado con Salmonella Javiana. *Revista del Instituto Nacional de Higiene Rafael Rangel* **2010**, *41*, 22-26.
- Tritt, A.; Eisen, J. A.; Facciotti, M. T.; Darling, A. E., An integrated pipeline for de novo assembly of microbial genomes. *PLoS One* **2012**, *7*, e42304.
- Tsai, M. H.; Huang, Y. C.; Lin, T. Y.; Huang, Y. L.; Kuo, C. C.; Chiu, C. H., Reappraisal of parenteral antimicrobial therapy for nontyphoidal Salmonella enteric infection in children. *Clin Microbiol Infect* **2011**, *17*, 300-5.
- Turner, F. S., Assessment of insert sizes and adapter content in fastq data from NexteraXT libraries. *Front Genet* **2014**, *5*, 5.
- Uribe, C.; Suárez, M. C., Salmonelosis no tifoidea y su transmisión a través de alimentos de origen aviar. **2013**.
- Urwin, R.; Maiden, M. C. J., Multi-locus sequence typing: a tool for global epidemiology. *Trends in Microbiology* **2003**, *11*, 479-487.
- Valdelamar, L. M. M.; Díaz, A. S.; Flores, R. C., Secuenciación de fragmentos de ADN.

- van Belkum, A.; Tassios, P. T.; Dijkshoorn, L.; Haeggman, S.; Cookson, B.; Fry, N. K.; Fussing, V.; Green, J.; Feil, E.; Gerner-Smidt, P.; Brisse, S.; Struelens, M.; European Society of Clinical, M.; Infectious Diseases Study Group on Epidemiological, M., Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin Microbiol Infect* **2007**, *13 Suppl 3*, 1-46.
- Vázquez, J. A.; Berrón, S., Multilocus sequence typing: el marcador molecular de la era de Internet. *Enfermedades Infecciosas y Microbiología Clínica* **2004**, *22*, 113-120.
- Vazquez-Garciduenas, M. S.; Romero-Perez, N. L.; Figueroa-Aguilar, G. A.; Jaime-Sanchez, J. L.; Vazquez-Marrufo, G., Investigation of a food-borne Salmonella Oranienburg outbreak in a Mexican prison. *J Infect Dev Ctries* **2014**, *8*, 143-53.
- Vazquez-Torres, A.; Jones-Carson, J.; Mastroeni, P.; Ischiropoulos, H.; Fang, F. C., Antimicrobial actions of the NADPH phagocyte oxidase and inducible nitric oxide synthase in experimental salmonellosis. I. Effects on microbial killing by activated peritoneal macrophages in vitro. *Journal of Experimental Medicine* **2000**, *192*, 227-236.
- Vílchez, G.; Alonso, G., Scope and limitations of molecular methods applied to epidemiological studies. *Revista de la Sociedad Venezolana de Microbiología* **2009**, *29*, 6-12.
- Walters, S. P.; González-Escalona, N.; Son, I.; Melka, D. C.; Sassoubre, L. M.; Boehm, A. B., Salmonella enterica diversity in central Californian coastal waterways. *Applied and environmental microbiology* **2013**, *79*, 4199-4209.
- Way, J. S.; Josephson, K. L.; Pillai, S. D.; Abbaszadegan, M.; Gerba, C. P.; Pepper, I. L., Specific detection of Salmonella spp. by multiplex polymerase chain reaction. *Applied and environmental microbiology* **1993**, *59*, 1473-1479.

- Whitaker, R. J., Allopatric origins of microbial species. *Philos Trans R Soc Lond B Biol Sci* **2006**, 361, 1975-84.
- Wiedemann, A.; Virlogeux-Payant, I.; Chausse, A. M.; Schikora, A.; Velge, P., Interactions of Salmonella with animals and plants. *Front Microbiol* **2014**, 5, 791.
- Winfield, M. D.; Groisman, E. A., Role of Nonhost Environments in the Lifestyles of Salmonella and Escherichia coli. *Applied and Environmental Microbiology* **2003**, 69, 3687-3694.
- Wray, C.; Wray, A., *Salmonella in domestic animals*. Cabi: 2000.
- Yoon, H.; Ansong, C.; McDermott, J. E.; Gritsenko, M.; Smith, R. D.; Heffron, F.; Adkins, J. N., Systems analysis of multiple regulator perturbations allows discovery of virulence factors in Salmonella. *BMC Syst Biol* **2011**, 5, 100.

Abreviaturas

ADN	Ácido desoxirribonucleico
CDC	Centro de prevención y control de enfermedades, por sus siglas en inglés
ddNTP	Desoxinucleotidos
DGE	Dirección General de Epidemiología
ENR	Especies reactivas de Nitrógeno
EOR	Especies reactivas de Oxígeno
EUA	Estados Unidos de América
Fig.	Figura
LPS	Lipopolisacaridos
Mb	Mega bases
mg	miligramo
min	minutos
mL	mililitro
MLST	Tipificación por Multilocus de Secuencia, por sus siglas en inglés
MO	Microorganismo
ng	nano gramo
NGS	Secuenciación de Nueva Generación, por sus siglas en inglés
nM	nano molar
°C	Grado(s) Celsius
OMS	Organización Mundial de la Salud
PCR	Reacción en Cadena de la Polimerasa
PFGE	Electroforesis en gel por Campos Pulsados, por sus siglas en inglés
pH	Potencial de hidrogeno
pM	pico molar
rpm	revoluciones por minuto
SNP	Polimorfismo de un solo nucleótido, por sus siglas en inglés
SNT	Salmonelosis no tifoidea

SPI	Isla de patogenicidad, por sus siglas en inglés
ST	Secuencia tipo, por sus siglas en inglés
TSB	caldo de soya tripticaseín
UAS	Universidad Autónoma de Sinaloa
UFC/g	Unidades formadoras de colonia por gramo.
UV	Ultravioleta
WGS	Secuenciación de Genoma Completo, por sus siglas en inglés
μL	micro litro