

**UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE POSTGRADO**



**LOCUS DE ADHESIÓN Y AUTOAGREGACIÓN (LAA), UNA NUEVA
ISLA DE PATOGENICIDAD PRESENTE EN CEPAS EMERGENTES
DE *Escherichia coli* PRODUCTOR DE SHIGA TOXINA**

DAVID ARTURO MONTERO FORERO

TESIS PARA OPTAR AL GRADO DOCTOR EN CIENCIAS BIOMÉDICAS

Directores de Tesis:

Prof. Dr. ROBERTO MAURICIO VIDAL

Prof. Dr. JOSÉ LUIS PUENTE

2018

**UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE POSTGRADO**

**INFORME DE APROBACIÓN TESIS DE
DOCTORADO EN CIENCIAS BIOMÉDICAS**

Se informa a la Comisión de Grados Académicos de la Facultad de Medicina, que la Tesis de Doctorado en Ciencias Biomédicas presentada por el candidato

DAVID ARTURO MONTERO FORERO

ha sido aprobada por la Comisión Informante de Tesis como requisito para optar al Grado de Doctor en Ciencias Biomédicas en Examen de Defensa de Tesis rendido el día 17 de octubre de 2018.

Prof. Dr. ROBERTO MAURICIO VIDAL
Director de Tesis
Programa de Microbiología y Micología
Instituto de Ciencias Biomédicas
Universidad de Chile

Prof. Dr. JOSÉ LUIS PUENTE
Director de Tesis
Departamento de Microbiología Molecular
Instituto de Biotecnología
Universidad Nacional Autónoma de México

COMISION INFORMANTE DE TESIS

PROF. DRA. MARÍA CARMEN MOLINA

PROF. DR. GUIDO MORA

PROF. DR. CARLOS SANTIVIAGO

PROF. DR. OMAR ORELLANA

Presidente Comisión de Examen

A mi madre Carmenza y
a mi padre Efraín

AGRADECIMIENTOS

Quiero agradecer a todas las personas que durante estos años me han ayudado a realizar mis metas personales y académicas. Su colaboración y compañía es invaluable.

En particular, quiero agradecer a mis padres, quienes han creído siempre en mí, me han dado su amor incondicional y me enseñaron a tomar el camino correcto. Este logro es de ustedes también.

A Juliana Velasco, gracias por tu amor y compañía, por caminar junto a mí y ser mi compañera de aventuras.

Un agradecimiento muy especial a:

El Dr. Roberto Vidal, tutor, maestro, mentor y amigo. Gracias por la paciencia, los consejos, las enseñanzas y ayuda en todo momento. Gracias por enseñarme a vivir y disfrutar el mundo de la ciencia.

Los Doctores José Luis Puente, Juan Carlos Salazar y Felipe del Canto por las enseñanzas, consejos, asesorías y ayuda. Ha sido un placer hacer ciencia con ustedes.

Los compañeros de laboratorio. Su ayuda y colaboración ha sido muy valiosa para la realización de esta Tesis.

Mis amigos por su apoyo en los momentos en los que son necesarias las sonrisas

FINANCIAMIENTO

La presente Tesis fue financiada por el proyecto FONDECYT Regular ID1161161 otorgado al Dr. Roberto Vidal, el proyecto CONACyT 239659 otorgado al Dr. José Luis Puente y la Beca CONICYT de Doctorado Nacional para extranjeros sin permanencia definitiva en Chile - PCHA/2014- 63140238 otorgada a David Montero.

ÍNDICE GENERAL

Contenido	Página
AGRADECIMIENTOS.....	4
FINANCIAMIENTO	5
ÍNDICE DE TABLAS	9
ÍNDICE DE FIGURAS	10
LISTA DE ABREVIACIONES	11
RESUMEN	13
ABSTRACT	15
1. INTRODUCCIÓN	17
1.1 Antecedentes generales.....	17
1.2 <i>Escherichia coli</i> productor de Shiga toxina	17
1.3 Epidemiología de STEC	18
1.4 Patogénesis de la infección por STEC.....	19
1.4.1 Síndrome Hemolítico Urémico.....	21
1.5 Mecanismos de patogenicidad de STEC.....	21
1.5.1 Toxinas.....	21
1.5.2 Isla de patogenicidad “Locus of Enterocyte Effacement”	22
1.5.3 Factores de adhesión y colonización en STEC LEE-negativo	23
1.5.4 Islas de patogenicidad identificadas en STEC LEE-negativo.....	24
1.6 Propósito de este estudio.....	25
2. HIPÓTESIS	26
3. OBJETIVO GENERAL.....	26
4. OBJETIVOS ESPECÍFICOS	27
4.1 Caracterizar funcionalmente la proteína Hes y determinar su contribución a la patogenicidad de cepas STEC.....	27
4.2 Caracterizar la organización genética de LAA en cepas STEC y determinar si existe una región común (core) de este locus entre ellas.	27
4.3 Analizar mediante genómica comparativa los genomas de cepas STEC LEE-negativo e identificar su repertorio de genes de virulencia.....	27
4.4 Determinar la contribución de LAA en la adherencia y colonización de STEC, mediante ensayos <i>in vitro</i> de adhesión y de un modelo murino de colonización, respectivamente.	27
5. MATERIALES Y MÉTODOS.....	28

5.1	Cepas bacterianas, condiciones de cultivo, plásmidos y partidores	28
5.2	Secuenciación del genoma de cepas STEC LEE-negativo.....	28
5.3	Secuencias genómicas disponibles públicamente.....	30
5.4	Detección del gen <i>hes</i> mediante PCR	30
5.5	Detección de LAA mediante PCR múltiple.....	30
5.6	Clonación y expresión de los genes <i>hes</i> y <i>hra1</i>	31
5.7	Construcción de mutantes isogénicos	32
5.8	Ensayo de hemaglutinación.....	33
5.9	Ensayo de autoagregación bacteriana.....	33
5.10	Ensayo de formación de biopelícula.....	33
5.11	Ensayos de adhesión e invasión bacteriana a células epiteliales humanas.....	34
5.12	Extracción de proteínas de membrana externa.....	35
5.13	Electroforesis bidimensional en geles de poliacrilamida (2D-PAGE).....	36
5.14	Ensayo de Western blot	37
5.15	Experimentos en animales.....	37
5.16	Análisis bioinformáticos	38
5.16.1	Determinación de serotipo y detección de genes de virulencia	38
5.16.2	Tipificación molecular	39
5.16.3	Identificación y determinación de la distribución de islas genómicas	39
5.16.4	Análisis filogenéticos.....	40
5.17	Análisis estadísticos.....	41
6.	RESULTADOS.....	42
6.1	Caracterizar funcionalmente la proteína Hes y determinar su contribución a la patogenicidad de cepas STEC.....	42
6.1.1	Análisis de la secuencia de aminoácidos y predicción de la estructura terciaria de la proteína Hes.....	42
6.1.2	Determinación de la distribución del gen <i>hes</i> en STEC.....	42
6.1.3	Caracterización funcional de Hes	44
6.1.4	Determinación de la participación del gen <i>hes</i> en la adhesión de STEC.....	46
6.2	Caracterizar la organización genética de LAA en cepas STEC y determinar si existe una región común (core) de este locus entre ellas.	47
6.2.1	Análisis de la secuencia de nucleótidos del Locus de Adhesión y Autoagregación.....	47
6.2.2	Análisis de la organización genética de LAA	49
6.2.3	Determinación de la distribución de LAA en STEC.....	52

6.3	Analizar mediante genómica comparativa los genomas de cepas STEC LEE-negativo e identificar su repertorio de genes de virulencia.....	57
6.3.1	Análisis <i>in silico</i> de genomas de cepas STEC LEE-negativo.....	57
6.3.1.1	Identificación del serotipo y contenido de genes de virulencia	58
6.3.1.2	Identificación de islas de patogenicidad	60
6.3.1.3	Identificación de elementos integrativos y conjugativos.....	66
6.3.2	Análisis filogenético y tipificación molecular	67
6.4	Determinar la contribución de LAA en la adherencia y colonización de STEC, mediante ensayos <i>in vitro</i> de adhesión y de un modelo murino de colonización, respectivamente.	73
6.4.1	Análisis funcional de LAA en ensayos <i>in vitro</i>	73
6.4.2	Determinación de la contribución de LAA a la colonización intestinal y virulencia de STEC en un modelo murino de infección	75
7.	DISCUSIÓN.....	78
8.	CONCLUSIONES	86
	APÉNDICE.....	87
	PUBLICACIONES ASOCIADAS A ESTA TESIS	102
	PRESENTACIONES EN CONGRESOS	102
	REFERENCIAS.....	104

ÍNDICE DE TABLAS

Tabla 1. Genes de virulencia identificados en STEC LEE-negativo	23
Tabla 2. Plásmidos y partidores usados en este estudio	29
Tabla 3. Distribución del gen <i>hes</i> entre cepas <i>E. coli</i>	44
Tabla 4. Distribución de genes de virulencia entre cepas STEC LEE-negativo	60
Tabla 5. Distribución de islas de patogenicidad entre cepas STEC LEE-negativo....	64
Tabla 6. Potencial de virulencia propuesto para cepas STEC LEE-negativo	85

ÍNDICE DE FIGURAS

Figura 1. Vías de transmisión de STEC al humano.....	19
Figura 2. Patogénesis de la infección por STEC.....	20
Figura 3. Análisis de la secuencia de aminoácidos y predicción de la estructura de la proteína Hes.....	43
Figura 4. Caracterización funcional de la proteína Hes	45
Figura 5. Adhesión de la cepa O113:H21 E045-00 y su mutante isogénica Δhes a células HT-29.....	46
Figura 6. Estructura genética del Locus de Adhesión y Autoagregación.....	48
Figura 7. Identificación de LAA en cepas STEC LEE-negativo	50
Figura 8. Comparación de la estructura genética de LAA presente en seis cepas STEC LEE-negativo.....	51
Figura 9. Comparación de la estructura genética de LAAB2F1 y PAIs relacionadas	52
Figura 10. Detección de los módulos I, II y III de LAA mediante PCR múltiple.....	53
Figura 11. Relaciones filogenéticas y distribución de módulos de LAA entre cepas STEC LEE-negativo y cepas <i>E. coli</i> relacionadas	56
Figura 12. Relaciones genéticas entre 42 secuencias de LAA.....	57
Figura 13. Distribución geográfica y temporal de las cepas STEC LEE-negativo analizadas.....	58
Figura 14. Estructura genética del Locus de Invasión e Inhibición del Crecimiento Dependiente del Contacto	62
Figura 15. Estructura genética del Locus de Adhesión y Colonización.....	63
Figura 16. Distribución de seis islas de patogenicidad entre cepas STEC LEE-negativo y perfiles de toxinas asociados a cada una de ellas	65
Figura 17. Estructura genética de los Elementos Integrativos y Conjugativos (ICE) identificados en cepas STEC LEE-negativo	67
Figura 18. Relaciones filogenéticas entre 367 cepas STEC LEE-negativo	71
Figura 20. Perfiles de proteínas de membrana externa de la cepa E045-00 y su mutante isogénica ΔLAA	74
Figura 21. Adhesión de la cepa O113:H21 E045-00 y su mutante isogénica ΔLAA a células HT-29.....	75
Figura 22. Contribución de LAA a la colonización intestinal y virulencia de STEC	76
Figura 23. Modelo de la evolución de la virulencia en STEC LEE-negativo, mediada por la adquisición de fagos e islas genómicas.....	82

LISTA DE ABREVIACIONES

2D-PAGE	Electroforesis bidimensional en geles de poliacrilamida
A/E	Lesión de adherencia y borrado del enterocito
CDI	Sistema de inhibición del crecimiento dependiente del contacto
CDT	Toxina distensora citoletal
CH	Colitis Hemorrágica
DAEC	<i>Escherichia coli</i> de adherencia difusa
DEC	<i>Escherichia coli</i> diarreogénico
DMEM	“Dulbecco’s Modified Eagle’s Medium”
DR	Secuencias directas repetidas
DTT	Ditiotreitol
EAEC	<i>Escherichia coli</i> enteroagregativo
EDA	Enfermedad diarreica aguda
EIEC	<i>Escherichia coli</i> enteroinvasivo
EPEC	<i>Escherichia coli</i> enteropatogénico
ETEC	<i>Escherichia coli</i> enterotoxigénico
ExPEC	<i>Escherichia coli</i> patógeno extraintestinal
GIs	Islas genómicas
Hes	Hemaglutinina de <i>E. coli</i> productor de Shiga toxina
HPI	“High-Pathogenicity Island”
Hra	Hemaglutininas resistentes al calor
ICE	Elemento Integrativo y Conjugativo
IEF	Isoelectroenfoque
IS	Secuencias de inserción
ISP	Instituto de Salud Pública de Chile
LAA	Locus de Adhesión y Autoagregación
LAC	Locus de Adhesión y Colonización
LB	Luria Bertani
LEE	“Locus of Enterocyte Effacement”

LIC	Locus de Invasión e Inhibición del Crecimiento Dependiente del Contacto
LPA	“Locus of Proteolysis Activity”
MALDI-TOF/TOF	Espectrometría de masas con ionización de la desorción del láser asistida por matriz con analizador de tiempo de vuelo
MGE	Elemento genético móvil
MOI	Multiplicidad de infección
OD	Densidad óptica
OMP	Proteína de membrana externa
ORF	Marcos de lectura abierto
PAI	Isla de Patogenicidad
PBS	Solución salina tamponada con fosfato
PCR	Reacción en cadena de la polimerasa
PMSF	Fenil metil sulfonil fluoruro
SE-PAI	“Subtilase-Encoding Pathogenicity Island”
SHU	Síndrome Hemolítico Urémico
STEC	<i>Escherichia coli</i> productor de Shiga toxina
Stx	Shiga toxina
SubAB	Citotoxina subtilasa
ufc	Unidades formadoras de colonias

RESUMEN

Escherichia coli productor de Shiga toxina (STEC; del inglés “Shiga toxin-producing *Escherichia coli*”) incluye a un grupo de patógenos transmitidos por alimentos que causan diarrea y enfermedades severas como Colitis Hemorrágica (CH) y Síndrome Hemolítico Urémico (SHU), principalmente en niños menores de 5 años. Las Shiga toxinas (Stx) son el principal factor de virulencia de estos patógenos; sin embargo, la adhesión y colonización del intestino humano es requerida en la patogénesis de STEC.

La isla de patogenicidad (PAI; del inglés “Pathogenicity Island”) denominada Locus de Borrado del Enterocito (LEE; del inglés “Locus of Enterocyte Effacement”), es el principal determinante de adhesión y colonización intestinal de un grupo de cepas STEC (LEE-positivo). A nivel mundial, las cepas STEC LEE-positivo (serogrupos O26, O45, O103, O111, O121, O145 y O157) han sido tradicionalmente asociadas con brotes y casos de SHU. En contraste, en las cepas STEC que carecen de LEE (LEE-negativo), el repertorio de genes de virulencia no ha sido completamente caracterizado. Es de destacar que la incidencia de casos de gastroenteritis causados por cepas STEC LEE-negativo (serogrupos O91, O113, O128, O146 y O174) ha aumentado recientemente en diferentes países, lo que resalta la importancia de su estudio.

En esta Tesis se reporta y caracteriza genética y funcionalmente una nueva PAI denominada Locus de Adhesión y Autoagregación (LAA), que está exclusivamente presente en cepas STEC LEE-negativo, incluyendo cepas aisladas de casos de CH y SHU. En LAA se localizan diferentes genes de virulencia que participan en fenotipos de adhesión y colonización. Consistente con lo anterior, los resultados obtenidos proveen evidencia experimental que sugiere la participación de esta PAI en la colonización intestinal en un modelo murino de infección. Adicionalmente, mediante genómica comparativa se analizaron más de 350 genomas de cepas STEC LEE-negativo y se determinó que LAA es la PAI más frecuente en este subgrupo de STEC. A través de este análisis se identificaron otras islas genómicas hasta el momento no

caracterizadas, incluyendo dos PAIs y un Elemento Integrativo y Conjugativo (ICE; del inglés "Integrative and Conjugative Element"). Más importante aún, nuestros análisis filogenéticos indican que en este subgrupo de STEC existen clados geográficamente diseminados que agrupan cepas que contienen más de una PAI, incluyendo LAA, lo que podría contribuir a la emergencia de estos patógenos. Este estudio es un paso adelante en el conocimiento de la evolución de los genomas y los mecanismos patogénicos de STEC LEE-negativo, lo que tiene implicaciones importantes para su vigilancia epidemiológica.

ABSTRACT

LOCUS OF ADHESION AND AUTOAGGREGATION (LAA), A NOVEL PATHOGENICITY ISLAND PRESENT IN EMERGING SHIGA TOXIN-PRODUCING *Escherichia coli* STRAINS

Shiga toxin-producing *Escherichia coli* (STEC) includes a group of foodborne pathogens that cause diarrhea and severe diseases, such as Hemorrhagic Colitis (HC) and Hemolytic Uremic Syndrome (HUS), primarily in children under 5 years of age. Shiga toxins are the major STEC virulence factors, but their ability to colonize the human intestine is a prerequisite to disease.

The Locus of Enterocyte Effacement (LEE), a Pathogenicity Island (PAI), is a major determinant of intestinal epithelium attachment of a group of STEC strains (LEE-positive). The LEE-positive STEC strains (serogroups O26, O45, O103, O111, O121, O145 and O157) have been traditionally associated with outbreaks and HUS cases worldwide. On the contrary, in the case of STEC strains lacking LEE (LEE-negative), their virulence repertoire has not been fully characterized. It is noteworthy that the incidence of gastroenteritis caused by LEE-negative STEC strains (serogroups O91, O113, O128, O146 and O174) has increased in several countries, highlighting the relevance of their study.

In this Thesis a novel PAI denominated Locus of Adhesion and Autoaggregation (LAA), which is exclusively present in LEE-negative STEC strains, is reported and characterized. Several virulence genes that participate in adhesion and colonization phenotypes are localized in LAA. In agreement with the above, the obtained results provide experimental evidence supporting the participation of LAA in the intestinal colonization of a mouse model of STEC infection. Additionally, more than 350 genomes of LEE-negative STEC strains were analyzed by comparative genomics and it was determined that LAA is the most frequent PAI in this subgroup of STEC. Through this analysis, other uncharacterized genomic islands were also identified, including two PAIs and one Integrative and Conjugative Element (ICE). Even more importantly, our phylogenetic analyzes indicate that in this subgroup of STEC there are geographically

disseminated clades in which strains containing more than one PAI, including LAA, are clustered, which could contribute to the emergence of these pathogens. This study is a step forward in our knowledge of the evolution of genomes and pathogenic mechanisms of LEE-negative STEC, with important implications for their epidemiological surveillance.

1. INTRODUCCIÓN

1.1 Antecedentes generales

Más de 2.500 millones de casos de enfermedad diarreica aguda (EDA) ocurren cada año en niños menores de 5 años en el mundo, siendo en este grupo etario la segunda causa de muerte (0,8 a 2 millones por año) después de las infecciones respiratorias (1, 2). Más del 80% de estas muertes ocurren en África y el Sudeste Asiático aunque existen poblaciones en Latinoamérica y en países industrializados en los que la EDA es una causa importante de morbilidad (1, 3, 4). Entre los principales agentes infecciosos causantes de EDA en la población pediátrica se encuentran Rotavirus y *Escherichia coli* diarreogénica (DEC; del inglés “Diarrheagenic *E. coli*”) (5). Específicamente, DEC se subdivide en seis categorías principales o patotipos de acuerdo con sus propiedades de virulencia, interacción con la mucosa intestinal, cuadro clínico y epidemiología: *E. coli* enteropatogénico (EPEC; del inglés “enteropathogenic *E. coli*”), STEC, *E. coli* enterotoxigénico (ETEC; del inglés “enterotoxigenic *E. coli*”), *E. coli* enteroagregativo (EAEC; del inglés “enteroaggregative *E. coli*”), *E. coli* enteroinvasivo (EIEC; del inglés “enteroinvasive *E. coli*”) y *E. coli* de adherencia difusa (DAEC; del inglés “Diffusely adherent *E. coli*”). De estos patotipos, STEC es el que causa las diarreas más graves y que pueden evolucionar a cuadros clínicos extraintestinales (6).

1.2 *Escherichia coli* productor de Shiga toxina

STEC incluye un grupo de patógenos transmitidos por alimentos que causan diarrea y enfermedades graves como CH y SHU (7). Estos patógenos causan brotes y casos esporádicos de gastroenteritis que afectan la salud pública y la industria de alimentos (8). A nivel mundial, se reportan anualmente más de dos millones de casos de gastroenteritis aguda asociados a STEC (9).

La Stx es el principal factor de virulencia y la responsable de las patologías extraintestinales durante el SHU (10): sin embargo, la adhesión a la mucosa intestinal es un paso requerido en la patogenicidad de STEC (11). La PAI LEE contiene genes que median la capacidad de adhesión de un grupo de cepas STEC denominado LEE-positivo de unirse de manera estrecha a la superficie de los enterocitos que colonizan,

dando lugar a un fenotipo conocido como adherencia íntima. Este subgrupo de STEC es clínicamente relevante debido a su asociación con brotes y casos de SHU (12, 13). En consecuencia, en el contexto de la infección por STEC, la presencia de LEE es considerada un factor de riesgo para el desarrollo de SHU (14). No obstante, cepas STEC que no contienen LEE (LEE-negativo) también han sido aisladas de casos de SHU, indicando que factores de virulencia adicionales están involucrados en la patogenicidad de estas bacterias.

1.3 Epidemiología de STEC

Dentro de las DEC, las infecciones por STEC son las únicas definidas como zoonóticas, siendo el ganado bovino el principal reservorio animal (15). Los primeros brotes de STEC durante las décadas de los 80' y los 90' fueron asociados al consumo de carne de bovino contaminada y mal cocida; sin embargo, brotes posteriores han sido asociados a una gran variedad de alimentos contaminados con este patógeno, como, por ejemplo, leche sin pasteurizar, jugo de frutas, lechuga y otros vegetales. De igual manera, debido a la baja dosis infectiva de STEC (10 a 100 bacterias), son posibles otras vías de contagio como contacto directo con animales infectados y transmisión persona a persona (16, 17) (**Figura 1**).

Poco más de 380 serotipos de STEC, definidos a partir de diferencias en su repertorio de antígenos somáticos (O) y flagelares (H), representan las cepas que han sido aisladas de humanos. Los serotipos STEC LEE-positivo implicados con mayor frecuencia con brotes y casos de enfermedad grave son O26:H11, O45:H2, O103:H2, O111:H8, O121:H19, O145:H28, y O157:H7 (18, 19). La frecuencia de aislamiento de estos serotipos varía geográfica y temporalmente, siendo el O157:H7 en general el que casusa más brotes y casos esporádicos de SHU (20–22). En Chile, los serotipos O26:H11 y O157:H7 son prevalentes, aunque esta apreciación pudiera estar sesgada ya que son los más buscados en el país debido a la disponibilidad de antisueros para su detección, por lo que no se puede restar importancia a otros serotipos circulantes a nivel local que no son investigados con igual interés (22).

En los últimos años, el número de reportes de cepas clínicas STEC LEE-negativo pertenecientes a los serogrupos O91, O113, O128, O146 y O174 ha aumentado en

Europa, Argentina y Corea del Sur (23, 24, 33, 25–32). Por ejemplo, en Alemania ha habido un aumento en la detección de cepas O91 aisladas de humanos, desde un 5% en 1999 hasta un 15% en 2012 y 2013 (25); en Holanda desde 2007 a 2012, O91 y O113 estuvieron entre los serogrupos más frecuentemente asociados con enfermedad en humanos (24); en Suiza aumentó la detección de cepas clínicas STEC LEE-negativo de 29.5% en 2010-2014 a 55% en 2017 (30); y recientemente en Argentina, O174 ha sido identificado como uno de los serogrupos comúnmente asociados con SHU (26). A pesar de su creciente prevalencia, se desconocen los mecanismos moleculares por los que estas cepas se adhieren al epitelio intestinal humano en ausencia de la PAI LEE (34). Por otra parte, la vigilancia epidemiológica dirigida contra STEC LEE-positivo y muy limitada contra STEC LEE-negativo, ha impedido una evaluación precisa de su diseminación global e impacto en salud pública (33). Actualmente, las razones para la emergencia de STEC LEE-negativo son desconocidas.

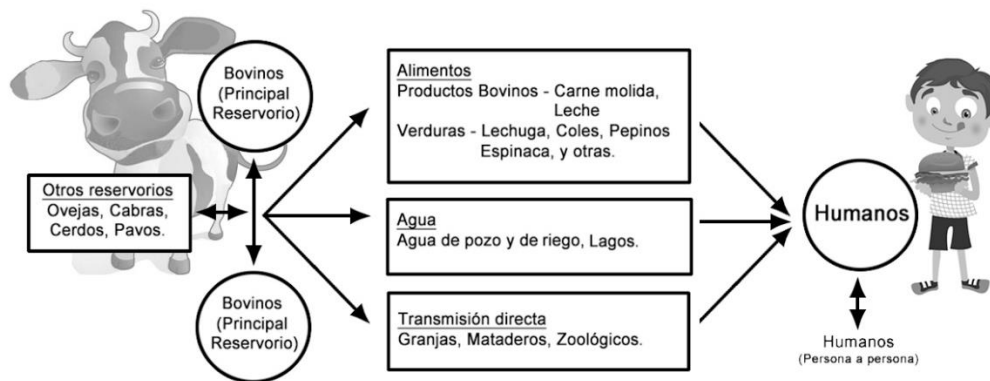


Figura 1. Vías de transmisión de STEC al humano. El ganado bovino constituye el principal reservorio animal de STEC, aunque existen otros reservorios animales. Las bacterias se eliminan en las heces y contaminan el medio ambiente. STEC puede ser transmitida a los humanos a través del consumo de carne, leche, agua y otros alimentos contaminados con heces. Las personas que trabajan o visitan granjas, mataderos y zoológicos pueden infectarse con la bacteria a través del contacto directo con los animales. La transmisión directa de persona a persona también es posible. Modificado y adaptado de (35).

1.4 Patogénesis de la infección por STEC

La **Figura 2** resume la patogénesis y la evolución cronológica de la infección por STEC en el humano. Tras ser ingerida, STEC sobrevive al pH ácido del estómago y posteriormente pasa al intestino delgado y al colon. En este último, se adhiere a los

enterocitos (alterando o no las microvellosidades) y produce la Stx, lo que afecta la función de la barrera epitelial y el transporte de iones ocasionando diarrea o disentería (36, 37). La Stx pertenece a la familia de toxinas AB₅, las que están conformadas por una subunidad B pentamérica y por una subunidad A con actividad enzimática. La subunidad B de la Stx se une al receptor glicolípido Gb3 presente en la superficie epitelial (38), permitiendo la entrada de la toxina por endocitosis al citoplasma celular (39), donde la subunidad A se une al RNA ribosomal y altera la síntesis de proteínas (40). La disminución en la síntesis de proteínas conduce a la muerte por apoptosis de las células blanco de la Stx. Ocasionalmente, en etapas posteriores y si suficiente Stx es producida, ésta puede alcanzar el torrente sanguíneo y diseminarse a tejidos extraintestinales donde existen receptores Gb3: células endoteliales del glomérulo renal, membrana celular de eritrocitos, células del sistema nervioso, entre otras. La evolución de la infección por STEC hacia el desarrollo de SHU está asociada con lesiones en el glomérulo renal que son producto de la acción citotóxica de la Stx, lo que puede resultar en muerte (41, 42).

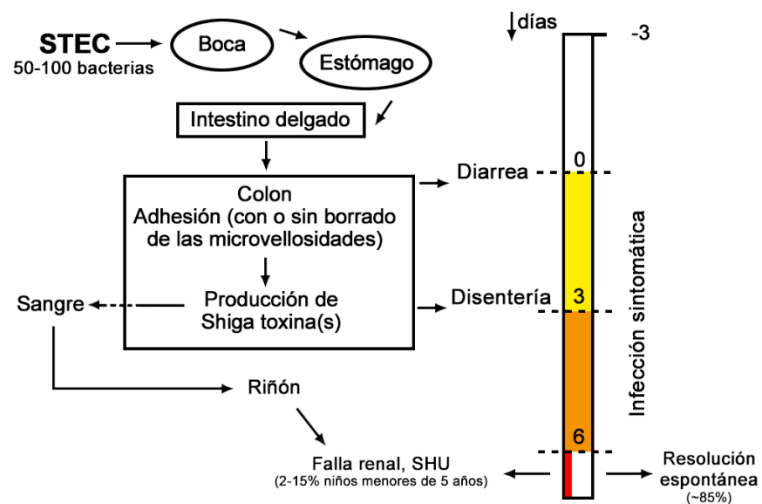


Figura 2. Patogénesis de la infección por STEC. La infección por STEC en el humano puede ser asintomática, producir diarrea acuosa o disentería y/o evolucionar a SHU en aproximadamente 2 a 15% de los niños menores de 5 años que se infectan con esta bacteria. La infección comienza con la entrada de las bacterias a la boca a través de alimentos o agua contaminada (día -3). La resistencia a pH ácido facilita su supervivencia a través del estómago. La bacteria pasa el intestino delgado y llega al colon donde genes de virulencia son activados en respuesta a las señales del entorno. STEC se adhiere a los enterocitos del colon causando diarrea (día 0). En algunos pacientes y si suficiente Stx es producida, el daño local a los vasos sanguíneos en el colon resulta en diarrea con sangre (día 3). Si suficiente Stx llega a la circulación sanguínea, los sitios del endotelio vascular ricos en el receptor de la toxina son dañados y su función es alterada. Los riñones y el sistema nervioso central son sitios frecuentemente afectados, y puede presentarse SHU (día 7). Modificado de (16, 43).

1.4.1 Síndrome Hemolítico Urémico

El SHU es un cuadro clínico caracterizado por anemia hemolítica microangiopática, trombocitopenia e insuficiencia renal aguda que puede progresar a insuficiencia renal crónica (44). Este síndrome presenta una alta incidencia en niños menores de 5 años con una letalidad de 3 a 5%. Se ha observado que cuando el serotipo infectante es O157:H7, el riesgo de desarrollar SHU es de un 10-15%, mientras que para otros serotipos este riesgo disminuye a 2% (43).

1.5 Mecanismos de patogenicidad de STEC

La infección por STEC involucra colonización del intestino y daño debido a toxinas. A continuación, se describen algunos de los principales factores de virulencia que utilizan estos patógenos para causar enfermedad.

1.5.1 Toxinas

1.5.1.1 Shiga toxinas. La producción de Stx es la característica que define a STEC. Existen dos grupos de Stx, denominados Stx1 y Stx2, los que están conformados por subtipos (Stx1a, Stx1c, Stx1d, Stx2a, Stx2b, Stx2c, Stx2d, Stx2e, Stx2f y Stx2g) definidos con base en diferencias de la secuencia de nucleótidos, actividad biológica y reactividad serológica (45). Una cepa de STEC puede producir uno o varios subtipos de Stx. Las cepas STEC que producen Stx1a, Stx2a y Stx2d están asociadas con los casos de diarrea más severa y SHU (46–49). El mecanismo de patogenicidad de la Stx fue descrito en el apartado 1.4.

1.5.1.2 Enterohemolisina. Esta toxina es codificada por el gen *ehxA* que está localizado en un plásmido de virulencia ampliamente distribuido en STEC (50). La enterohemolisina es parte de la familia RTX de citotoxinas con actividad formadora de poros (51). Durante la infección por STEC, el daño a la mucosa puede conllevar a la liberación de sangre al lumen intestinal; en este punto, la enterohemolisina puede lisar eritrocitos liberando el grupo hemo y hemoglobina, lo que podría servir como fuente de hierro y favorecer el crecimiento bacteriano (52).

1.5.1.3 Toxina distensora citoletal. Las toxinas distensoras citoletales (CDTs; del inglés “Cytolethal Distending Toxins”) son una familia de proteínas termolábiles que interfieren con el ciclo celular en la fase G1 o G2, lo que conlleva a la muerte de la

célula blanco (53). Estas toxinas son producidas por diferentes patógenos bacterianos incluyendo a *E. coli*. En general, los miembros de la familia CDT están codificados por un grupo de tres genes (*cdtA*, *cdtB* y *cdtC*) que codifican tres proteínas que conforman la holotoxina CDT (54). De estas proteínas, CdtB es la subunidad con actividad enzimática que es transportada al núcleo, donde genera roturas a la doble cadena de DNA (55). En *E. coli* se han reportado cinco variantes de CDT (CDT-I a CDT-V), que están codificadas en el cromosoma, plásmidos y bacteriófagos (56). Específicamente, las variantes CDT-III y CDT-V han sido identificadas en diferentes serotipos de STEC, incluyendo O91:H21, O128:H8, O157:NM, entre otros (54, 57). La CDT-V producida por STEC causa daño irreversible a las células endoteliales humanas y, por lo tanto, puede contribuir a la patogénesis durante la infección (53).

1.5.1.4 Citotoxina subtilasa. Al igual que la Stx, la citotoxina subtilasa (SubAB) es un miembro de la familia de toxinas AB₅. Las subunidades de esta toxina están codificadas en el operón *subAB* (58). Una vez internalizada a la célula blanco, la subunidad A inactiva la chaperona BiP en el retículo endoplasmático, lo que conlleva la inhibición transitoria de la síntesis de proteínas y a la inducción de apoptosis (59, 60). Esta toxina está distribuida principalmente entre cepas STEC LEE-negativo y se han identificado hasta el momento tres variantes: SubAB1 codificada en un plásmido de virulencia, SubAB2-1 codificada en una PAI y SubAB2-2 codificada en el cromosoma (61, 62). La administración intraperitoneal de SubAB en ratones causa inflamación severa y daño tisular en riñón, hígado y bazo, resultando en la muerte de los animales (60).

1.5.2 Isla de patogenicidad LEE

En esta PAI se localiza el gen *eae* que codifica la proteína Intimina, adhesina no fimbrial considerada el principal factor involucrado en la adherencia de STEC LEE-positivo a células epiteliales intestinales. Adicionalmente, en LEE se codifica un sistema de secreción tipo tres (SST3) que se encarga de translocar distintas proteínas hacia al interior del enterocito. Entre las proteínas secretadas a través del SST3 (algunas de ellas también codificadas en LEE) están el receptor translocado de Intimina (Tir; del inglés “Translocated intimin receptor”), y diversas proteínas

involucradas en mecanismos de transducción de señales que alteran la arquitectura y fisiología de la célula epitelial (6, 7). Como consecuencia de la adhesión mediada por LEE, se produce en los enterocitos la lesión denominada “adherencia y borrado del enterocito” (A/E; del inglés “Attaching and Effacing lesion”), en la que un reordenamiento del citoesqueleto genera una estructura en forma de pedestal en el sitio de adherencia íntima de la bacteria, lo que causa pérdida de las microvellosidades del intestino, inflamación y diarrea (12, 63, 64).

1.5.3 Factores de adhesión y colonización en STEC LEE-negativo

A diferencia de STEC LEE-positivo, en STEC LEE-negativo no se ha establecido si la colonización es mediada por un mecanismo molecular específico o es el resultado de la acción individual de factores de virulencia. Si bien se han identificado PAIs exclusivas de STEC LEE-negativo (ver más adelante), el rol biológico de estos elementos genéticos no se ha investigado a fondo. Por otro lado, entre los factores de virulencia descritos en este grupo de STEC se encuentran genes que codifican adhesinas (fimbriales y no fimbriales) y proteínas autotransportadoras que pueden funcionar como adhesinas, toxinas y/o proteasas (**Tabla 1**). Además, algunos de estos genes (*saa*, *eibG*, *sab* y *epeA*) están distribuidos exclusivamente, aunque no son ubicuos, entre cepas STEC LEE-negativo, por lo que han sido propuestos como marcadores moleculares de estas cepas (11, 65, 66).

Tabla 1. Genes de virulencia identificados en STEC LEE-negativo.

Gen	Función / efecto	Referencia
Marcadores moleculares de STEC LEE-negativo		
<i>saa</i> (STEC autoagglutinating adhesin)	Adhesión	(65)
<i>eibG</i> (Immunoglobulin-binding protein G)	Adhesión	(67)
<i>sab</i> (STEC autotransporter contributing to biofilm formation)	Adhesión, formación de biopelícula	(11)
<i>epeA</i> (EHEC plasmid-encoded autotransporter)	Mucinasas	(66)
Adhesinas		
<i>lpfA</i> (Long polar fimbriae subunit A)	Adhesión	(34)
<i>iha</i> (IrgA homologue adhesin)	Adhesión, receptor de sideróforo	(68)
<i>tia</i> (Toxicogenic invasion loci A)	Adhesión e invasión	(69)
Autotransportadores		
<i>ag43</i> (Antigen 43)	Adhesión, Autoagregación, Biopelícula	(70)
<i>espP</i> (Extracellular serine protease plasmid pO157-encoded)	Adherencia, Citotoxina	(71)
<i>espl</i> (<i>E. coli</i> secreted protease, island encoded)	Proteasa	(72)

1.5.4 Islas de patogenicidad identificadas en STEC LEE-negativo

Dos PAIs han sido identificadas como exclusivas de STEC LEE-negativo. Éstas son el Locus de Actividad Proteolítica (LPA; del inglés “Locus of Proteolysis Activity”) (72) y la SE-PAI (del inglés “Subtilase-Encoding Pathogenicity Island”) (59). Otra PAI, denominada HPI (del inglés “High-Pathogenicity Island”) ha sido reportada tanto en STEC LEE-positivo como en LEE-negativo (73). A continuación, se describen las principales características de estos elementos genéticos.

1.5.4.1 LPA. Es una región de DNA de 33 kb que está adyacente al gen *se/C-tRNA*. LPA contiene 34 genes, entre ellos *iha* y *espl* (72). La contribución directa de LPA a la virulencia de STEC no ha sido determinada aún; sin embargo, se sabe que *Iha* funciona como adhesina y que *Espl* es una serino proteasa que degrada la pepsina A porcina y la apolipoproteína A-I humana (72). LPA está ampliamente distribuida en cepas STEC LEE-negativo, principalmente en las que producen Stx1c y Stx2b [originalmente denominada Stx2d (47)] (74).

1.5.4.2 SE-PAI. Es una región de DNA de 8 kb que está adyacente al gen *pheV-tRNA*. Esta PAI contiene 6 genes, incluyendo el operón *subAB* que codifican la toxina SubAB2-1. Otro factor de virulencia presente en esta PAI es el gen *tia*. Recientemente, se reportó que la proteína Tia está involucrada en la capacidad de invasión de algunas cepas STEC LEE-negativo (75).

1.5.4.3 HPI. Esta PAI fue identificada originalmente en *Yersinia pestis* y otras especies del género *Yersinia*, aunque estudios posteriores han indicado que está ampliamente diseminada entre especies de la familia Enterobacteriaceae, incluida *E. coli* (73). El ortólogo de HPI en *E. coli* tiene 35 kb, 12 genes y está adyacente al gen *asnV-tRNA*. Esta PAI contiene genes involucrados en la biosíntesis y transporte de yersiniabactina (Ybt), un sideróforo acoplable a hierro y además el gen *fyuA*, que codifica una proteína de membrana externa que actúa como receptor de Ybt. Este sistema de captura de hierro favorece el crecimiento de la bacteria en el hospedero mamífero (76). HPI está ampliamente distribuida en EAEC y *E. coli* patógeno extraintestinal (ExPEC; del inglés “Extra-intestinal pathogenic *E. coli*”) (77), aunque también se ha identificado en menor grado en STEC (73).

1.6 Propósito de este estudio

En un estudio previo, nuestro grupo identificó un antígeno reactivo a sueros de pacientes con SHU en el extracto de proteínas de membrana externa de la cepa STEC LEE-negativo O113:H21 E045-00 (78). El análisis de este antígeno mediante espectrometría de masas sugirió que su identidad correspondía a la proteína Hek, que hace parte de la familia de hemaglutininas resistentes al calor (Hra; del inglés “Heat-resistant agglutinin”) (79); sin embargo, no se encontró el gen *hek* en los genomas de cepas STEC que se encuentran disponibles en GenBank. En su lugar, se encontró una variante alélica (GenBank: EGW68377) en la cepa STEC LEE-negativo O91:H21 B2F1, la cual a la fecha no ha sido descrita, por lo que siguiendo la nomenclatura utilizada para otros miembros de la familia Hra, la nombramos Hemaglutinina de *E. coli* productor de Shiga toxina (Hes). Análisis bioinformáticos posteriores, realizados en mi unidad de investigación, indicaron que el gen *hes* está localizado en una región de DNA de 86 kb que exhibe características típicas de las PAIs. Entre los genes de virulencia identificados en esta región están *iha* y *ag43*, que son conocidos por su participación en adhesión y autoagregación, respectivamente. Por lo tanto, esta región fue nombrada como Locus de Adhesión y Autoagregación (LAA). Teniendo en cuenta estas observaciones, el objetivo de esta Tesis fue caracterizar genética y funcionalmente esta región de DNA y determinar su participación en la patogenicidad de STEC. Los resultados de este estudio contribuyen al conocimiento de la evolución, epidemiología y patogenicidad de estas bacterias.

2. HIPÓTESIS

El Locus de Adhesión y Autoagregación está involucrado en la patogenicidad de un grupo emergente de cepas STEC a través de su participación en la adherencia y colonización del epitelio intestinal.

3. OBJETIVO GENERAL

Determinar si el Locus de Adhesión y Autoagregación contribuye a la adherencia y colonización del epitelio intestinal de un grupo de cepas STEC.

4. OBJETIVOS ESPECÍFICOS

4.1 Caracterizar funcionalmente la proteína Hes y determinar su contribución a la patogenicidad de cepas STEC.

Actividades:

- Análisis de la secuencia de aminoácidos y predicción de la estructura terciaria de la proteína Hes
- Determinación de la distribución del gen *hes* en STEC
- Caracterización funcional de la proteína Hes
- Determinación de la participación de *hes* en la adhesión de STEC

4.2 Caracterizar la organización genética de LAA en cepas STEC y determinar si existe una región común (core) de este locus entre ellas.

Actividades:

- Análisis de la secuencia de nucleótidos de LAA
- Análisis de la organización genética de LAA
- Determinación de la distribución de la LAA en STEC

4.3 Analizar mediante genómica comparativa los genomas de cepas STEC LEE-negativo e identificar su repertorio de genes de virulencia

Actividades:

- Análisis *in silico* de genomas de cepas STEC LEE-negativo
- Análisis filogenético y tipificación molecular

4.4 Determinar la contribución de LAA en la adherencia y colonización de STEC, mediante ensayos *in vitro* de adhesión y de un modelo murino de colonización, respectivamente.

Actividades:

- Análisis funcional de LAA en ensayos *in vitro*
- Determinación de la contribución de LAA a la colonización intestinal y virulencia de STEC en un modelo murino de infección.

5. MATERIALES Y MÉTODOS

5.1 Cepas bacterianas, condiciones de cultivo, plásmidos y partidores

Las cepas bacterianas utilizadas se indican en la Tabla S1. Las cepas STEC fueron caracterizadas mediante reacción en cadena de la polimerasa (PCR; del inglés “Polymerase chain reaction”) y serotipificación por el Instituto de Salud Pública de Chile (ISP) y por el Laboratorio de Enteropatógenos, Programa de Microbiología y Micología, Facultad de Medicina, Universidad de Chile. Para el clonamiento de genes se utilizó la cepa de laboratorio *E. coli* DH5 α . En los ensayos de expresión heteróloga se utilizó la cepa de laboratorio *E. coli* HB101. Además, se obtuvieron mutantes espontáneas resistentes a estreptomicina (Str^r) de las cepas STEC O113:H21 E045-00 y O91:H21 V07-4-4. Las cepas se cultivaron en caldo Luria Bertani (LB; Triptona 10 g/l, NaCl 10 g/l y extracto de levadura 5 g/l) con o sin agitación a 37 °C. Para la preparación del medio sólido se agregó agar bacteriológico a una concentración final de 1,5% (p/v). También fue usado agar MacConkey (BD Difco™, USA). Cuando fue necesario, los medios de cultivo se suplementaron con ampicilina (100 μ g/ml), kanamicina (50 μ g/ml), estreptomicina (50 μ g/ml), IPTG (50 μ g/ml), X-Gal (5-bromo-4-cloro-3-indolil- β -D-galactopiranosido) (50 μ g/ml), L-arabinosa (10 Mm) y ácido toluico (2 mM). Los plásmidos y partidores utilizados se indican en la **Tabla 2**. Para el diseño de partidores se utilizó la herramienta Primer-BLAST (80).

5.2 Secuenciación del genoma de cepas STEC LEE-negativo

Se secuenciaron los genomas de treinta y cinco cepas STEC LEE-negativo, aisladas en Argentina (n = 14), Chile (n = 16) y España (n = 5). Las cepas de Argentina y España fueron amablemente donadas por la Dra. Nora Padola (Universidad Nacional del Centro de la Provincia de Buenos Aires, Tandil, Argentina) y el Dr. Jorge Blanco (Universidad de Santiago de Compostela, Lugo, España), respectivamente. Las cepas chilenas hacen parte de la colección de STEC del Dr. Roberto Vidal. La secuenciación y el ensamblaje de genomas fueron realizados por la empresa MicrobesNG (Universidad de Birmingham, Reino Unido; <http://www.microbesng.uk>) utilizando las plataformas Illumina MiSeq o HiSeq 2500. Los genomas parciales fueron proporcionados por la empresa después de eliminar extremos de baja calidad y ensamblar lecturas con SPAdes 3.10 (81). Las estadísticas de ensamblado se

obtuvieron con Quast v4.6.351 (82). Los contigs < 200 nt se eliminaron y las secuencias se depositaron en GenBank con el número de bioproyecto PRJNA448751. Los números de acceso de los genomas se indican en la **Tabla S2**.

Tabla 2. Plásmidos y partidores usados en este estudio.

Plásmidos	Descripción	Fuente
pTZ57R/T	Plásmido para clonación directa de productos de PCR	Fermentas, Lituania
pTZ57R/T_hes	Plásmido derivado con el gen <i>hes</i> clonado	Este estudio
pTZ57R/T_hra1	Plásmido derivado con el gen <i>hra1</i> clonado	Este estudio
pVB1	Plásmido de expresión	Dualsystems Biotech, Suiza
pVB1_hes	Plásmido derivado con el gen <i>hes</i> clonado	Este estudio
pVB1_hra1	Plásmido derivado con el gen <i>hra1</i> clonado	Este estudio
pKD4	Plásmido que contiene el gen de resistencia a kanamicina flanqueado por secuencias FRT (Flp Recombianase Target).	(83)
pKD46	Plásmido que contiene los genes <i>bet</i> , <i>gam</i> y <i>exo</i> que codifican el sistema de recombinación Red del fago lambda	(83)
pCP20	Plásmido que contiene el gen de la recombinasa Flp de levadura.	(84)

Partidores	Secuencia (5' → 3') *	Protocolo	Tamaño (pb)	Fuente
<i>hes_seq1</i>	ATGATTGAAATGAAAAAGTTAT	Amplificación de la secuencia completa del gen <i>hes</i>	750	Este estudio
<i>hes_seq2</i>	TCAGAAGTTATAAGTCATAC			
<i>hes_det1</i>	CAACCAGCGTCTTATCGAT	Detección del gen <i>hes</i>	350	Este estudio
<i>hes_det2</i>	CGGTTGTTTTCTGGTGAAC			
<i>hes_clon1</i>	<u>CATATG</u> ATTGAAATAAAAAAGTTATTGTG	Clonamiento del gen <i>hes</i>	759	Este estudio
<i>hes_clon2</i>	<u>GGATCCT</u> CAGAAGTTATAAGTCATAC			
<i>hra1_clon1</i>	<u>CATATG</u> AGCAACCCCATTTATCGTT	Clonamiento del gen <i>hra1</i>	801	Este estudio
<i>hra1_clon2</i>	<u>GGATCCT</u> CAGAAGTTATAAGTCACAC			
<i>hes_del_for</i>	ATGAAAAAGGTTATTGTGGTTTCAGTAT TTGCAATGGCGGGTATGTTTTAGCCC AGGCCGT GTAGGCTGGAGCTGCTTC	Reemplazo alélico / delección del gen <i>hes</i>	1606	Este estudio
<i>hes_del_rev</i>	AATATACGCAATTATCGTTCAGGAGCAG TGTCGTCAGAAGTTATAAGTCATACCAA GCAT CATATGAATATCCTCCTTAG			
<i>hes_conf_for</i>	ATTCTTCTCAGTGTATCCGC	Confirmación de la delección del gen <i>hes</i>	1329	Este estudio
<i>hes_conf_rev</i>	AGGCTATTCGGCTATGACTG			
<i>LAA1_for</i>	AACCGGGAGCCATATTCAGC	Detección módulo LAA-I	2207	Este estudio
<i>LAA1_rev</i>	ACGCGGTAAGTCGATGTAGAG			
<i>LAA2_for</i>	CTGCAAAAGACATTGCCACAACA	Detección módulo LAA-II	1689	Este estudio
<i>LAA2_rev</i>	TGCACAACAGGGTCAAGTGA			
<i>ms-3_for</i>	GCGGTAGCAGCTCATGCAG	Detección módulo LAA-III	880	(85)
<i>ms-3_rev</i>	GACAACACTGACCGGATAATC			
<i>LAA_del_for</i>	CGGGCACCAAATTCATATCAACGGACCT CCACGGAGGTCCGTTTTTCGT GTTAGG CTGGAGCTGCTTC	Reemplazo alélico / delección de LAA	1599	Este estudio
<i>LAA_del_rev</i>	TTGTCAGATCCAAAAGCAAAAACCCGCC TTGTGGGCGGGTCTTAAGAATTAG CAT ATGAATATCCTCCTTAG			
<i>LAA_conf_for</i>	GGTATTGCGGTATCGGTGAT	Confirmación de la delección de LAA	969	Este estudio (83)
<i>K1</i>	CAGTCATAGCCGAATAGCCT			

* La secuencia subrayada indica los sitios de reconocimiento para las enzimas de restricción NdeI y BamHI. Priming sites 1 y 2 son indicados en negrita.

5.3 Secuencias genómicas disponibles públicamente

Un total de 332 genomas (parciales o completos) de cepas STEC LEE-negativo se descargaron de GenBank (Tabla S2). Los genomas parciales o completos de 13 cepas de *E. coli* relacionadas fueron también descargados. El manejo rutinario de secuencias genómicas y búsquedas BLASTn se realizaron usando el software Geneious (v11.0.5; Biomatters Ltd). La detección de los genes *stx* y *eae* se realizó mediante búsquedas BLASTn.

5.4 Detección del gen *hes* mediante PCR

Los genes codificantes de las proteínas de la familia Hra tienen un alto porcentaje de identidad. Por lo tanto, para detectar y caracterizar el gen *hes* mediante PCR fue necesario diseñar dos pares de partidores distintos. El primer par de partidores, denominados *hes_det1* y *hes_det2*, amplifican una región específica de *hes*. Por el contrario, el segundo par de partidores, denominados *hes_seq1* y *hes_seq2*, amplifican la secuencia completa de *hes* y de algunos de sus homólogos. Inicialmente, los partidores *hes_det1* + *hes_det2* fueron usados para identificar las cepas positivas para *hes*. Posteriormente, las cepas positivas para *hes* fueron analizadas con los partidores *hes_seq1* + *hes_seq2* y los productos de PCR obtenidos fueron enviados a secuenciar a la empresa Macrogen (USA). Las dos reacciones de PCR se realizaron en un volumen final de 25 µl que contenía 1 µl de DNA molde, 0,25 µM de cada partidor, 0,3 µM de cada dNTP (Fermentas, Lituania), 5 µl de buffer “GoTaq DNA-polymerase” 5X y 1,25 U de “GoTaq® DNA Polymerase” (Promega, USA). Las reacciones de amplificación incluyeron desnaturalización inicial a 94 °C durante 5 min, 30 ciclos de desnaturalización a 94 °C durante 30 s, hibridación a 56 °C durante 30 s y extensión a 72 °C durante 1 min, con una extensión final a 72 °C por 10 min. Los productos de PCR se analizaron mediante electroforesis en geles de agarosa al 1% usando buffer Tris-acetato-EDTA y se tiñeron con bromuro de etidio.

5.5 Detección de LAA mediante PCR múltiple

La detección simultánea de los módulos I, II y III de LAA se realizó a través de un ensayo de PCR. Los partidores LAA1_for + LAA1_rev y LAA2_for + LAA2_rev amplifican los módulos I y II, respectivamente. Un tercer par de partidores (*ms3_for* +

ms3_rev), reportado por Girardeau et al., (85), fue usado para la amplificación del módulo III. No fue posible diseñar partidores específicos para el módulo IV debido a que los genes codificados en esta región están ampliamente distribuidos en *E. coli*. Las reacciones de amplificación se realizaron en un volumen final de 25 µl que contenía 1 µl de DNA molde, 0,3 µM de cada partidor, 0,4 µM de cada dNTP (Fermentas, Lituania), 5 µl de buffer “GoTaq DNA-polymerase” 5X y 1,25 U de “GoTaq® DNA Polymerase” (Promega, USA). La reacción de amplificación incluyó desnaturalización inicial a 94 °C durante 5 min, 30 ciclos de desnaturalización a 94 °C durante 30 s, hibridación a 62,5 °C durante 40 s y extensión a 68 °C durante 2 min, con una extensión final a 72 °C por 10 min. Los productos de PCR se analizaron mediante electroforesis en geles de agarosa al 1% usando buffer Tris-acetato-EDTA y se tiñeron con bromuro de etidio.

5.6 Clonación y expresión de los genes *hes* y *hra1*

Los genes *hes* y *hra1* se amplificaron a partir de DNA genómico de las cepas STEC E045-00 y EAEC 042, respectivamente. Las secuencias codificantes de *hes* y *hra1* fueron amplificadas usando los partidores *hes_clon1* + *hes_clon2* y *hra1_clon1* + *hra1_clon2*, respectivamente. Estos partidores permitieron agregar sitios de reconocimiento para las enzimas de restricción NdeI y BamHI en los extremos 5' y 3' de cada gen, respectivamente. Los productos de PCR se ligaron al plásmido pTZ57R/T (Fermentas, Lituania), siguiendo las instrucciones del fabricante de la T4 DNA ligasa (Promega, USA), para obtener los plásmidos pTZ57R/T_*hes* y pTZ57R/T_*hra1*. Estos plásmidos se usaron para transformar la cepa *E. coli* DH5α, y los clones transformantes se seleccionaron de acuerdo con la resistencia a ampicilina y α-complementación. El correcto clonamiento de los genes en el plásmido pTZ57R/T fue confirmado por secuenciación (Macrogen, USA). A continuación, los plásmidos correspondientes fueron purificados y digeridos con NdeI y BamHI. Los productos de la digestión enzimática se analizaron por electroforesis en gel de agarosa, y los insertos (genes *hes* y *hra1*) fueron purificados. Estos fragmentos se ligaron al plásmido pVB1 (Dualsystems Biotech, Suiza) usando la T4 DNA ligasa (Promega, USA) para obtener los plásmidos pVB1_*hes* y pVB1_*hra1*, en los que la expresión de *hes* y *hra1* está regulada por el sistema de expresión Pm/XylS. Estos plásmidos se usaron para

transformar la cepa *E. coli* HB101. Como control, el plásmido vacío también se transformó en la cepa HB101. Para inducir la expresión de *hes* y *hra1* se suplementó el medio con ácido toluico (2 mM) desde el inicio del cultivo.

5.7 Construcción de mutantes isogénicos

El gen *hes* fue inactivado en la cepa E045-00 mediante reemplazo alélico siguiendo la metodología descrita por Datsenko y Wanner (83). Brevemente, la cepa E045-00 fue transformada con el plásmido pKD46, que contiene los genes que codifican el sistema de recombinación Red del fago lambda. Posteriormente, el plásmido pKD4 se utilizó como DNA molde para amplificar por PCR (partidores *hes_del_for* + *hes_del_rev*) un fragmento de DNA que contiene el gen de resistencia a kanamicina flanqueado por 62 y 60 nucleótidos idénticos a los extremos 3' y 5' del gen *hes*, respectivamente. El producto de PCR obtenido fue utilizado para transformar la cepa E045-00 / pKD46 y los clones transformantes se seleccionaron de acuerdo con la resistencia a kanamicina. Además, se verificó el intercambio alélico mediante PCR usando los partidores *hes_conf_for* + *hes_conf_rev*. Posteriormente, el gen de resistencia a kanamicina fue removido transformando la cepa mutante con el plásmido pCP20. La complementación del gen *hes* se realizó en trans utilizando el plásmido pVB1_*hes*.

La inactivación de LAA en las cepas E045-00 y V07-4-4 se realizó como se describió para el gen *hes*, pero utilizando los partidores LAA_*del_for* + LAA_*del_rev* para amplificar un fragmento de DNA que contiene el gen de resistencia a kanamicina flanqueado por 49 y 53 nucleótidos idénticos a los extremos 3' del gen *pheV-tRNA* y de LAA. Los clones resistentes a kanamicina se analizaron por PCR para verificar el intercambio alélico usando los partidores LAA_*conf_for* + K1. Todas las reacciones de PCR se realizaron en un volumen final de 50 µl que contenía 2 µl de DNA molde, 0,25 µM de cada partidor, 0,3 µM de cada dNTP (Fermentas, Lituania), 10 µl de buffer "GoTaq DNA-polymerase" 5X y 2,5 U de "GoTaq® DNA Polymerase" (Promega, USA). Las reacciones de amplificación incluyeron desnaturalización inicial a 94 °C durante 5 min, 30 ciclos de desnaturalización a 94 °C durante 30 s, hibridación a 52 °C durante 45 s y extensión a 72 °C durante 2 min, con una extensión final a 72 °C por 10 min.

Los productos de PCR se analizaron mediante electroforesis en geles de agarosa al 1% usando buffer Tris-acetato-EDTA y se tiñeron con bromuro de etidio.

5.8 Ensayo de hemaglutinación

La aglutinación de eritrocitos de cordero se llevó a cabo siguiendo el protocolo descrito por Gutiérrez et al., (86), con modificaciones menores. Brevemente, la sangre de cordero (comprada al ISP) fue centrifugada tres veces a 500 x g y suspendida en solución salina tamponada con fosfato (PBS; pH 7.2). Las bacterias se cultivaron en caldo LB a 37 °C sin agitación y se ajustaron a una densidad óptica de 600 nm (OD_{600}) de 0,7. Las bacterias fueron mezcladas con las células sanguíneas en una proporción 1:1, 1:2, 1:4 y 1:8, y 100 μ l de cada mezcla fueron dispensados en una placa de poliestireno de fondo redondo de 96 pocillos, seguido de incubación por 1 h a 4 °C. Un resultado positivo se evidenció por la formación de una película roja en la parte inferior, y un resultado negativo por un sedimento rojo en el fondo del pocillo. Este ensayo se realizó dos veces en duplicado.

5.9 Ensayo de autoagregación bacteriana

La autoagregación bacteriana se evaluó como fue descrito por Bhargava et al., (87), con modificaciones. Brevemente, las bacterias se cultivaron en caldo LB durante la noche a 37 °C sin agitación. Los cultivos se centrifugaron a 9000 x g, se suspendieron en PBS y se ajustaron a una OD_{600} de 0,9. Posteriormente, 5 ml de cada suspensión bacteriana fueron depositados por duplicado en tubos estériles de 15 ml. Los tubos se dejaron estáticos a temperatura ambiente. Un tubo permaneció estático y el otro se agitó con vórtex antes de cada medición. Para medir la autoagregación bacteriana a lo largo del tiempo, en puntos de tiempo designados (cada hora durante 8 horas), se tomó 0,5 ml de la superficie de cada suspensión bacteriana y se midió la OD_{600} . Este ensayo se realizó dos veces en duplicado.

5.10 Ensayo de formación de biopelícula

La formación de biopelícula se observó y se cuantificó mediante fijación y tinción con cristal violeta como describió Bhargava et al. (87), con modificaciones. Brevemente, las bacterias fueron cultivadas en caldo LB durante toda la noche a 37 °C sin agitación. Al otro día, 10 μ l del cultivo bacteriano fueron utilizados para inocular

1 ml de caldo LB en una placa de cultivo celular de 24 pocillos. Las placas se incubaron sin agitación a temperatura ambiente durante 24, 48 y 72 h. Después de cada incubación, el medio de cultivo fue retirado, los pocillos lavados tres veces con PBS y las bacterias adheridas a la superficie de poliestireno fijadas durante 7 min con metanol al 70%. Después de dejar secar por completo las placas, las biopelículas se tiñeron con cristal violeta al 0,5% durante 15 min y lavadas dos veces con agua destilada. La cuantificación de la formación de biopelícula se llevó a cabo mediante la dilución del cristal violeta con 0,5 ml de ácido acético glacial al 33% y medición de la solución a OD₅₉₅. Este ensayo se realizó tres veces en triplicado.

5.11 Ensayos de adhesión e invasión bacteriana a células epiteliales humanas

La adhesión bacteriana a células epiteliales Caco-2 (ATCC® HTB-37™), HT-29 (ATCC® HTB-38™) y HEp-2 (ATCC® CCL-23™) se evaluó siguiendo el protocolo descrito por Gutiérrez et al., (86), con modificaciones. Brevemente, las células epiteliales se cultivaron en DMEM (del inglés “Dulbecco’s Modified Eagle’s Medium”) suplementado con 10% de suero fetal bovino y 1% de penicilina-estreptomicina a 37 °C en atmósfera de 5% CO₂. Las células se sembraron en placas de 24 pocillos y se cultivaron hasta 100% de confluencia (aproximadamente 4x10⁵ células/pocillo). Los pre-inóculos bacterianos se cultivaron durante toda la noche en LB con agitación a 37 °C. Una alícuota de cada pre-inóculo se diluyó 50 veces en el mismo medio de cultivo y se incubó a 37 °C durante 4 h con agitación. Las células epiteliales se lavaron tres veces con PBS y se infectaron con una multiplicidad de infección (MOI) de 10 bacterias/célula en DMEM durante 30 min o 1 h a 37 °C. Las bacterias no adherentes (planctónicas) se retiraron mediante cinco lavados con PBS, y las bacterias adherentes se recuperaron por lisis de la monocapa de células epiteliales con 0.1% Tritón X-100. El número de bacterias adheridas se determinó mediante diluciones en serie y recuento de unidades formadoras de colonias (ufc) en agar LB. El resultado final se expresó como el porcentaje de bacterias adheridas en relación con la cantidad de bacterias del inóculo inicial.

Para visualizar el fenotipo de adherencia, las células HT-29 se cultivaron en cubreobjetos de vidrio y se infectaron durante 1 h como se describió anteriormente.

Después de tres lavados con PBS, se añadió DMEM y se incubó durante 1 h. Luego de otros tres lavados con PBS, se añadió metanol al 70% durante 7 min y las células se tiñeron con Giemsa al 5% durante 40 min a temperatura ambiente.

Para determinar la invasión bacteriana, se realizó un ensayo de protección a gentamicina. Una hora después de la infección, las bacterias planctónicas se retiraron mediante tres lavados con PBS y las células epiteliales se incubaron con DMEM suplementado con gentamicina (50 µg/ml) durante 1,5 h. Posteriormente, el medio se removió, se lavó tres veces con PBS y las bacterias invasivas se recuperaron por lisis de la monocapa de células con 0,1% Tritón X-100. Después de diluciones en serie y recuento de ufc en agar LB, se determinó el porcentaje de bacterias invasivas en relación con la cantidad de bacterias del inóculo inicial. Todos los ensayos se realizaron tres veces por duplicado.

5.12 Extracción de proteínas de membrana externa

Las extractos de proteínas de membrana externa (OMPs; del inglés “Outer membrane proteins”) se obtuvieron como describe Rivas *et al.*, (88), con las modificaciones realizadas por Montero *et al.* (78). Brevemente, 2 l de cultivo bacteriano crecido durante la noche en LB, con o sin agitación a 37 °C, se centrifugaron a 9.000 x g por 10 min a 4 °C. El sobrenadante se removió y las bacterias fueron lavadas y suspendidas en Tris-HCl 10 mM pH 8,0 hasta un volumen final de 5 ml suplementado con 1 mM de Fluoruro de fenilmetilsulfonilo (PMSF; del inglés “phenylmethylsulfonyl fluoride”). Cada suspensión se sometió a sonicación en hielo (60 ciclos de 30 s con intervalos de 30 s permitiendo enfriamiento) y se centrifugó a 12.000 x g por 10 min a 4 °C para remover bacterias y restos celulares. El sobrenadante se trató con DNasa/RNasa durante 20 min a 25 °C y se centrifugó a 12.000 x g por 10 min a 4 °C. Para solubilizar la membrana citoplasmática el sobrenadante se incubó con Sarkosyl (Sigma Aldrich Co., USA) al 2% durante 30 min a 25 °C y luego se centrifugó a 20.500 x g por 60 min a 4 °C. El sedimento obtenido se lavó con 1 ml de agua miliQ estéril conteniendo 1 mM PMSF y se centrifugó a 20.500 X g por 40 min a 4 °C. El sedimento final se suspendió en 200 µl de agua miliQ estéril. La concentración de proteínas en la fracción de membrana externa se determinó usando el reactivo “Bradford protein assay

dye reagent” (Bio-Rad) y seroalbúmina de bovino como estándar, siguiendo las instrucciones del fabricante. Los extractos de OMPs se almacenaron a -80 °C hasta su uso.

5.13 Electroforesis bidimensional en geles de poliacrilamida (2D-PAGE)

Los extractos de OMPs fueron analizados mediante 2D-PAGE como describió Montero *et al.*, (78). Brevemente, tiras prefabricadas de 13 cm y de rango lineal de pH 4-7 (GE Healthcare) fueron rehidratadas con 250 µl del reactivo “DeStreak rehydration solution” (GE Healthcare) conteniendo buffer IPG 4-7 (1%) (GE Healthcare), ditiotreitól (DTT; 10 mM) y 200 µg de OMPs, por 16 h a temperatura ambiente de acuerdo con las instrucciones del fabricante. Posteriormente, las tiras rehidratadas fueron sujetas a isoelectroenfoque (IEF) usando el sistema “Ettan IPGphor Isoelectric Focusing System” (GE Healthcare) a 20 °C con un voltaje límite de 50 µA/tira con el siguiente programa: 200 V/1 h, 500 V/1 h, 1.000 V en gradiente/1 h, 8.000 V en gradiente/3,5 h y 8.000 V finales hasta alcanzar un voltaje final de 19 - 20 kVh. Una vez terminado el IEF, las tiras fueron incubadas por 15 min en 3 ml de solución de equilibrio I [Tris-HCl 50 mM pH 8.8, urea 6 M, glicerol 30% (p/v), SDS 2% (p/v) y DTT 10 mg/mL (p/v)] y posteriormente por 15 min en 3 ml de solución de equilibrio II, que se diferencia de la primera en que contiene 2-yodoacetamida (25 mg/ml) en lugar de DTT. Luego, las tiras fueron embebidas en la parte superior de geles de poliacrilamida al 12%. La segunda dimensión se realizó utilizando la cámara de electroforesis vertical SE-600 Ruby Standard Dual Cool (Amersham Biosciences, USA) de acuerdo con las instrucciones del fabricante. Como marcador de peso molecular se usó “Precision Plus Protein™ Kaleidoscope™ Standards” (Bio-Rad). La electroforesis se realizó con una intensidad de corriente de 5 mA/gel durante 20 min, seguido de 15 mA/gel durante 9 h hasta que el frente de azul de bromofenol alcanzó el límite inferior del gel. Los geles se tiñeron con Blue Silver (89) y se digitalizaron en escáner “Epson Perfection V750 Pro”. Las imágenes fueron analizadas usando el software BioNumerics 2D v6.6 (Applied Maths, Bélgica).

Por otro lado, proteínas seleccionadas se cortaron de los geles de poliacrilamida y se identificaron mediante espectrometría de masas con ionización de la desorción

del láser asistida por matriz con analizador de tiempo de vuelo (MALDI-TOF / TOF) en el “Mass Spectrometry Core” de la Universidad de Texas Medical Branch (Galveston, Texas).

5.14 Ensayo de Western blot

Las OMPs analizadas mediante SDS-PAGE fueron transferidas a nitrocelulosa (Millipore) por 60 min a 100 V / 4 °C usando el sistema “Mini Trans-Blot Cell” (Bio-Rad) siguiendo las instrucciones del fabricante. Una vez terminada la transferencia, los sitios de unión libres de la membrana se bloquearon con la solución de bloqueo [Tris buffer salino (TBS) 1X, Tween 20 al 0,03% y albumina de suero bovino (BSA; del inglés “Bovine serum albumin”) al 1%] durante 1 h a temperatura ambiente. Posteriormente, se lavó 3 veces por 5 min con solución TBS-T (TBS 1X, Tween 20 al 0,3%). A continuación, se incubó con agitación durante 1 h en solución de bloqueo que contenía la mezcla de sueros de ratón, diluida en una proporción 1:250. Después de 2 lavados con solución TBS-T durante 10 min, se incubó la membrana en solución de bloqueo conteniendo anti-IgG ratón conjugado con peroxidasa (Cat # A4416, Sigma Aldrich Co., USA) diluido en proporción 1:1.000, durante 1 h a temperatura ambiente y con agitación. El revelado se realizó con el sustrato 3,3',5,5' - Tetrametilbenzidina (Cat # 860336, Sigma Aldrich Co., USA) y la reacción se detuvo con agua destilada.

5.15 Experimentos en animales

Todos los experimentos con animales se realizaron en la Universidad de Concepción, siguiendo protocolos aprobados por el Comité de Bioética de la Facultad de Ciencias Biológicas (Ver anexos). Ratones BALB/c machos (6 a 8 semanas de edad, comprados al ISP) fueron distribuidos aleatoriamente en grupos experimentales. Los ratones se mantuvieron en instalaciones convencionales para animales y recibieron agua y alimentos *ad libitum*. El modelo ratón de infección por STEC se llevó a cabo como describió previamente Lindgren et al. (90), con modificaciones menores. Brevemente, a los ratones se les dio agua *ad libitum* que contenía estreptomicina (5 g/l) por 24 a 48 h antes de la inoculación y durante el experimento. Las cepas bacterianas se cultivaron en caldo LB suplementado con estreptomicina (50 µg/ml), a 37 °C, con agitación y durante 18 h. Los cultivos se centrifugaron, los sedimentos se

lavarón en PBS estéril y se suspendieron en una solución de 20% sacarosa (p/v) y 10% NaHCO₃ (p/v), alcanzando una concentración de ~10¹⁰ ufc/ml. La concentración bacteriana se confirmó mediante diluciones en serie y recuentos de ufc en placas de agar LB.

Los ratones fueron privados de comida y agua durante la noche (12 h). A la mañana siguiente, cada animal se infectó oralmente utilizando una micropipeta con ~10⁹ ufc suspendidas en 100 µl. Después del desafío, los animales se alojaron individualmente y se les proporcionó alimento y agua suplementada con estreptomicina (5 g/l) *ad libitum*. La colonización intestinal de las bacterias desafiantes se determinó mediante el recuento de ufc resistentes a estreptomicina en las heces. Para esto, las heces se colectaron diariamente, se pesaron, se homogeneizaron, se suspendieron en 1 ml de PBS y después de diluciones seriadas en PBS, alícuotas fueron sembradas en placas de agar MacConkey suplementadas con estreptomicina (50 µg/ml). Para los experimentos de supervivencia, los ratones se manipularon e infectaron como se describió para los experimentos de colonización, sin embargo, no se registró el número de ufc en las heces a lo largo del tiempo.

5.16 Análisis bioinformáticos

5.16.1 Determinación de serotipo y detección de genes de virulencia

El serotipo de las cepas se determinó utilizando la herramienta SerotypeFinder (91). Un conjunto de genes de virulencia (*stx*, *subA*, *astA*, *cdtB*, *sta1*, *stb*, *senB*, *lpfA*, *iha*, *espP*, *espl*, *epeA*, *pic*, *sigA*, *cba*, *celb*, *cma*, *mcmA*, *mchB*, *mchC* y *mchF*) fue identificado usando la herramienta VirulenceFinder (92). Otros genes fueron detectados mediante búsquedas BLASTn: *saa* (AF399919.3; posiciones: 6290-7840), *sab* (AY258503.2; posiciones: 118905-123200), *eibG* (GU295813.1), *hes* (CP023541.1; posiciones: 4035512- 4036252), *lesP* (CP023541.1; posiciones: 4063646-4067740), *tia* (JQ994271.1; posiciones: 4726-5472), y *hra1* (U07174.1). La presencia de un gen se confirmó cuando se encontró una cobertura e identidad > 90%. El gen *ag43* tiene varios alelos que se dividen en dos subfamilias (70). La identificación de los alelos *ag43*-I (subfamilia I) y *ag43*-II (subfamilia II) se realizó mediante PCR *in*

silico utilizando partidores dirigidos a regiones conservadas como describió Montero et al., (78).

5.16.2 Tipificación molecular

La tipificación molecular de las cepas se llevó a cabo de acuerdo con el esquema de Achtman (93) utilizando la herramienta MLST 1.8 (94). La asignación de filogrupos se realizó con la herramienta ClermonTyping (95).

5.16.3 Identificación y determinación de la distribución de islas genómicas

Los contigs de los genomas parciales se ordenaron y alinearon contra el genoma completo de *E. coli* K-12 MG1655 usando progressiveMauve (96). En otros casos, el ensamblaje de contigs (mapeo) contra secuencias de referencia específicas se realizó mediante el software Geneious v10.0.9 (Biomatters Ltd, USA). La identificación y caracterización de regiones de DNA con características de islas genómicas (GIs) se realizó usando un conjunto de algoritmos y herramientas bioinformáticas. Por ejemplo, las secuencias directas repetidas (DR), los elementos de secuencias de inserción (IS) y los genes de tRNA se identificaron usando REPuter (97), ISfinder (98) y tRNAscan-SE (99), respectivamente. Además, los marcos de lectura abiertos (ORFs) y el contenido de G + C se determinaron usando el software Geneious. Regiones de DNA con características de PAIs fueron buscadas mediante BLASTn en las bases de datos Islander (100), PAIDB (101) y ICEberg (102). Las regiones de DNA correspondientes a nuevas PAIs y ICEs se extrajeron de las secuencias genómicas originales disponibles en GenBank y luego se anotaron utilizando el servidor RAST (103).

En el caso de LAA, secuencias completas (cuatro módulos) fueron comparadas usando progressiveMauve (96) y EasyFig v2.1 (104). Además, se determinó la presencia, ausencia y variaciones en los genes localizados en LAA mediante búsquedas BLASTn realizadas en el software Geneious y usando la secuencia de LAA presente en la cepa STEC O91:H21 B2F1 (LAA_{B2F1}) como referencia. Por defecto, la ausencia de un gen se definió cuando la cobertura y/o identidad fue < 60%. Además, se obtuvo un mapa de calor mostrando la presencia, ausencia y % identidad de estos genes utilizando el paquete gplots (105) en R (106).

Las secuencias de nucleótidos de las islas SE-PAI, LPA y HPI fueron descargadas de GenBank usando los números de acceso Q994271, AJ278144.1 y AL031866 (posiciones: 78113-114560), respectivamente. La distribución de las PAIs entre las cepas fue determinada mediante búsquedas BLASTn usando las secuencias de nucleótidos completas. Por defecto, la usencia de una PAI se definió cuando la cobertura y/o identidad de nucleótidos fue < 70%. La distribución de las PAIs entre las cepas fue posteriormente analizada mediante una red de asociaciones basada en una matriz de presencia / ausencia, usando el paquete igraph (107) en R (106).

5.16.4 Análisis filogenéticos

Tres análisis filogénicos basados en polimorfismos de nucleótido simple (SNPs) fueron realizados. El primero se realizó en el servidor CSI Phylogeny 1.4 (108), que identifica SNPs en genomas parciales o completos, filtra, valida la posición de cada SNP, infiere la filogenia en función de perfiles de SNPs concadenados y construye árboles de máxima verosimilitud. Este análisis incluyó un total de 128 genomas parciales y completos de *E. coli* (115 STEC LEE-negativo y 13 cepas *E. coli* relacionadas) y se realizó utilizando parámetros de entrada por defecto y el genoma de *E. coli* K-12 MG1655 como referencia. Como resultado, se identificaron 167.167 SNPs en 3.008.649 posiciones presentes en todos los genomas analizados. El archivo de salida en formato Newick fue descargado y utilizado para la visualización del árbol filogenético en FigTree v.1.4.2 (<http://tree.bio.ed.ac.uk/>).

En el segundo análisis se utilizó también el servidor CSI Phylogeny 1.4 (108), pero en este caso, el análisis incluyó 42 genomas parciales de cepas que contenían los 4 módulos de LAA y se utilizó LAA_{B2F1} como secuencia de referencia.

El tercer análisis se realizó implementando kSNP3.1 (109) en un servidor 20-core Dell Poweredge R730. En este caso se analizaron 367 genomas parciales o completos de cepas STEC LEE-negativo y *E. coli* K-12 MG1655 como referencia. Como resultado, se identificaron 3,956 SNPs en el genoma central de las cepas analizadas. El árbol filogenético se obtuvo con el método de máxima parsimonia y 100 réplicas Bootstrap. La edición del árbol se realizó con iTOL (110).

5.17 Análisis estadísticos

Los datos de los ensayos de adhesión, invasión y formación de biopelícula fueron analizados mediante la prueba t de Student (dos colas). Las diferencias en la distribución de los genes de virulencia entre las fuentes de aislamiento se analizaron individualmente mediante la prueba exacta de Fisher (dos colas). La magnitud de los patrones de co-ocurrencia o exclusión mutua entre PAIs fue determinada mediante la razón de momios (Odds ratio).

Las diferencias en los niveles de colonización bacteriana para cada día en los estudios de infección con ratones se analizaron usando la prueba U de Mann-Whitney. La diferencia en las tasas de supervivencia entre grupos de ratones se analizó utilizando la prueba de log-rank. Para todas las pruebas estadísticas, un valor $P < 0.05$ fue considerado significativo.

6. RESULTADOS

6.1 Caracterizar funcionalmente la proteína Hes y determinar su contribución a la patogenicidad de cepas STEC.

6.1.1 Análisis de la secuencia de aminoácidos y predicción de la estructura terciaria de la proteína Hes

La familia Hra está conformada por las OMPs Hra1, Hra2, Tia y Hek (79). Hra1 participa en hemaglutinación, autoagregación, formación de biopelícula y adherencia agregativa (87); Hra2 en adhesión (79); Tia en adhesión e invasión (69); y Hek en hemaglutinación, autoagregación, adhesión e invasión (111). Teniendo en cuenta estos antecedentes, se determinaron similitudes y diferencias entre la secuencia de aminoácidos de Hes y los miembros de la familia Hra. El alineamiento de las secuencias de aminoácidos indicó que Hes tiene 90%, 65%, 65% y 86% de identidad con Hra1, Hra2, Tia y Hek, respectivamente (**Figura 3A y 3B**).

Además, la predicción de la estructura de Hes en dos y en tres dimensiones indicó que la mayoría de las substituciones de aminoácidos en relación con los homólogos Hra se localizan en regiones “loops” que se exponen hacia el espacio extracelular (**Figura 3C y 3D**). De esta manera, las diferencias funcionales entre los homólogos Hra podrían ser causadas por la variación de aminoácidos en estos loops.

6.1.2 Determinación de la distribución del gen *hes* en STEC

Se determinó la distribución del gen *hes* mediante PCR en una colección de cepas *E. coli*, incluyendo 12 cepas comensales fecales (n=12), 119 cepas STEC LEE-positivo y 48 cepas STEC LEE-negativo (**Tabla S1**). El gen *hes* fue identificado en 29/48 (60%) de las cepas STEC LEE-negativo, incluyendo aislados de los serotipos O91:H21, O113:H21 y O174:H21 (**Tabla 3**). Además, la cepa STEC O113:H21 E045-00 fue positiva para *hes* indicando que la proteína inicialmente identificada como Hek es de hecho Hes; sin embargo, *hes* no se detectó en otras dos cepas O113:H21 analizadas, indicando que este gen no está presente en todas las cepas de un mismo serotipo. Adicionalmente, la secuenciación del gen *hes* en las 29 cepas indicó que su secuencia

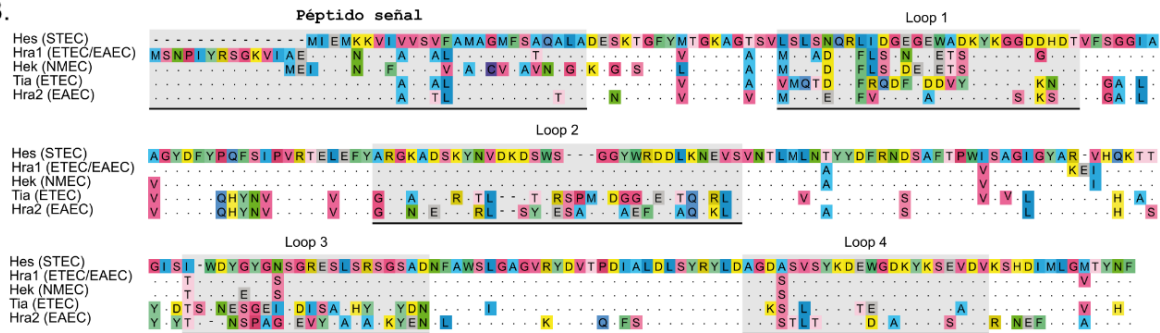
es 99.7% conservada (datos no mostrados). En contraste, *hes* no fue detectado en las cepas comensales y STEC LEE-positivo.

A.

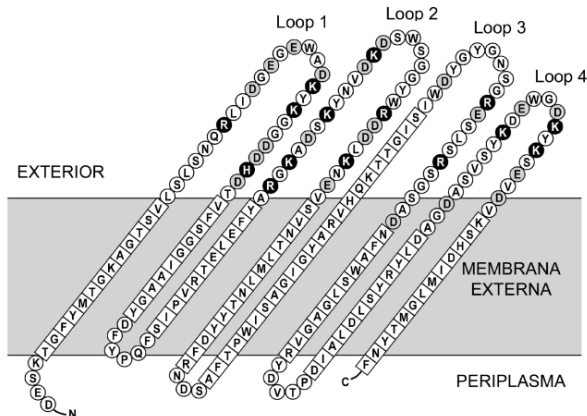
(%) Porcentaje de Identidad

1: Hes	100.00	89.56	86.35	65.06	65.04
2: Hra1	89.56	100.00	90.04	67.20	63.56
3: Hek	86.35	90.04	100.00	61.45	59.35
4: Tia	65.06	67.20	61.45	100.00	69.76
5: Hra2	65.04	63.56	59.35	69.76	100.00

B.



C.



D.

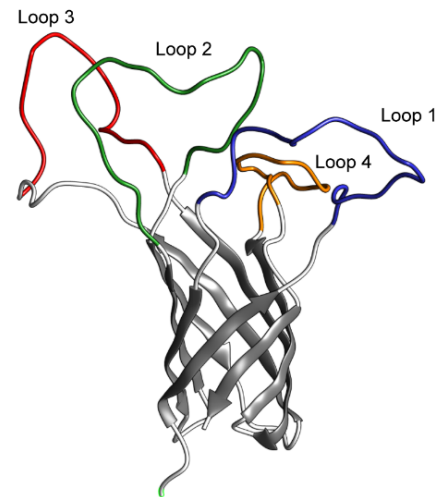


Figura 3. Análisis de la secuencia de aminoácidos y predicción de la estructura de la proteína Hes. **A.** Porcentajes de identidad entre los miembros (*Hes*, *Hra1*, *Hek*, *Tia* y *Hra2*) de la familia *Hra*. **B.** Alineamiento de secuencias de aminoácidos entre miembros de la familia *Hra* usando MUSCLE (112) en Unipro UGENE 1.21 (113). Los residuos de aminoácidos se colorean de acuerdo con el esquema de color de UGENE. Sólo se muestran los residuos que difieren de aquellos en la secuencia de *Hes* para la misma posición. Los puntos indican residuos idénticos. Los números de acceso de las secuencias utilizadas son: Q46678 (*Hra1*), AAK73174 (*Hek*), Q46760 (*Tia*) y AEK70958 (*Hra2*). **C.** Modelo bidimensional de la topología de la proteína *Hes*. Los aminoácidos predichos que forman estructuras β -barril transmembrana están indicados por cuadrados y los que forman estructuras de loops están indicados por círculos. Los aminoácidos con carga positiva y negativa presentes en los loops externos (expuestos al exterior) están indicados por círculos negros y grises, respectivamente. El modelo fue obtenido usando Protter (114) y PredictProtein (115). **D.** Estructura hipotética tridimensional de *Hes*, construida con base a la estructura de la proteína *Opa60* (número de acceso, banco de datos de proteínas: 2MLH) en el servidor Phyre2 (116). La imagen fue editada usando UCSF Chimera 1.10.2 (117).

Tabla 3. Distribución del gen *hes* entre cepas *E. coli*

Cepas	No. de aislados	No. de aislados <i>hes</i> +
Comensales	11	0
STEC LEE-positivo	119	0
STEC LEE-negativo		
<i>Serogrupo / Serotipo</i>		
O2	1	0
O7	1	1
O8:H21	4	0
O20:H19	1	1
O22	1	1
O46	1	1
O76	1	0
O91:H21	2	2
O103:H42	1	1
O104:H7	4	0
O104:H21	2	2
O113:H21	15	13
O130:H11	2	0
O136:H16	1	0
O139	4	1
O156:H-	1	1
O163:H9	1	1
O174	1	0
O171:H2	2	2
O174:H21	1	1
O181:H49	1	1
Total. STEC LEE-negativo	48	29 60%

6.1.3 Caracterización funcional de Hes

Teniendo en cuenta la función biológica de los miembros de la familia Hra, se evaluó si el producto del gen *hes* es capaz de conferir diferentes fenotipos de colonización. Para esto, *hes* fue clonado en el plásmido de expresión pVB1. Como control positivo, el gen *hra1* fue también clonado en este plásmido. Posteriormente, los plásmidos obtenidos fueron utilizados para transformar la cepa de laboratorio *E. coli* HB101. Con las cepas transformantes se realizaron ensayos *in vitro* de hemaglutinación, autoagregación bacteriana, formación de biopelícula, adhesión e invasión a células epiteliales. Los resultados obtenidos indican que cuando *hes* es expresado bajo el control de un promotor inducible, confiere la capacidad de aglutinar eritrocitos de oveja (**Figura 4A**) y de autoagregación (**Figura 4B y 4C**) a la cepa HB101. Además, promueve la formación de biopelícula a 72 h de incubación (**Figura 4D y 4E**). Los ensayos de adhesión e invasión mostraron que Hes incrementa significativamente ($p < 0.001$) la adherencia de la cepa HB101 a células epiteliales humanas (**Figura 4F**), en un patrón agregativo (**Figura 4G**), pero no promueve la invasión de células epiteliales (no mostrado). Estos resultados demuestran que el producto del gen *hes* es funcional y que confiere diferentes fenotipos asociados a colonización.

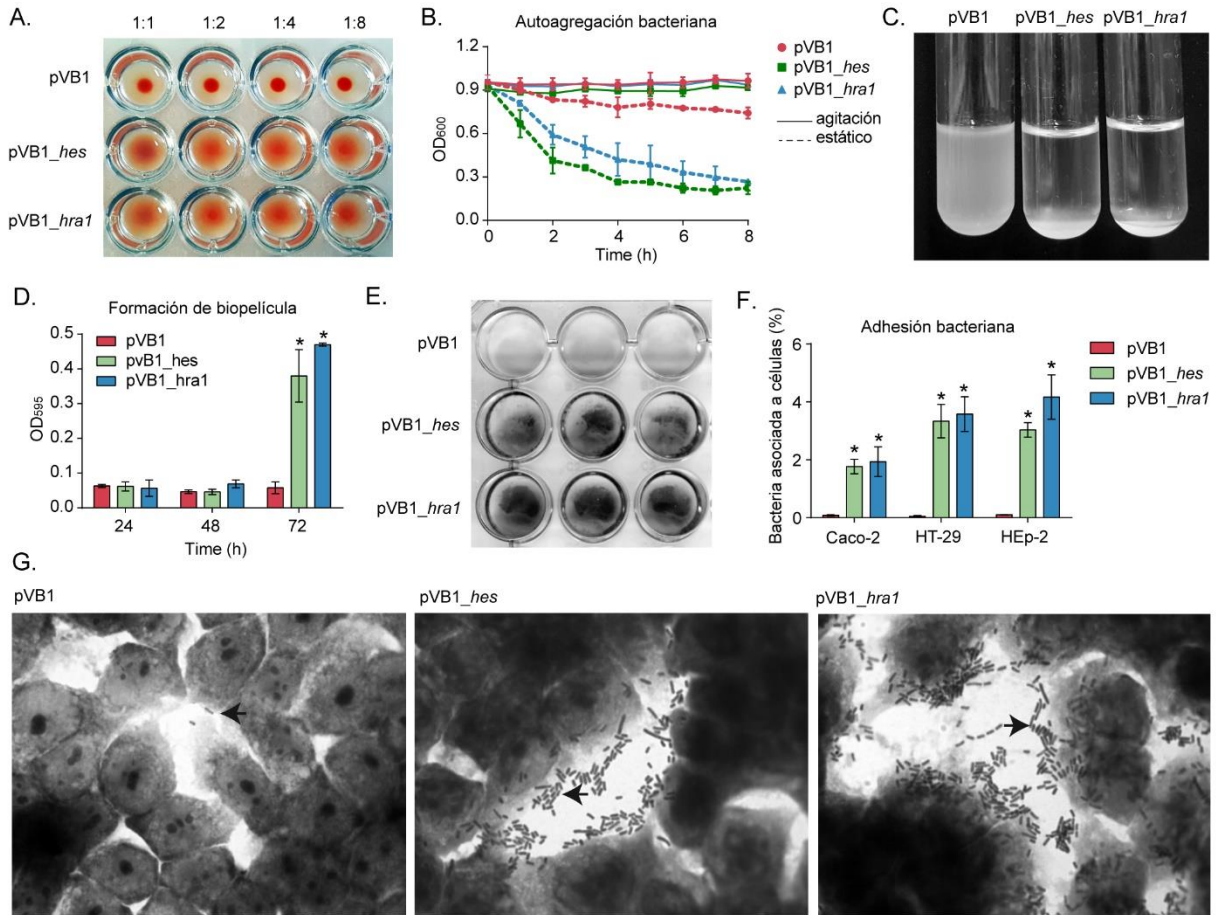


Figura 4. Caracterización funcional de la proteína Hes. La cepa *E. coli* HB101 transformada con los plásmidos pVB1, pVB1_hes o pVB1_hra1 fue evaluada en su capacidad de aglutinación de eritrocitos de oveja (A), autoagregación bacteriana (B, C), formación de biopelícula (D, E) y adhesión a células epiteliales humanas (F, G). **A.** Ensayo de Hemaglutinación. Las bacterias fueron evaluadas en diluciones 1:1, 1:2, 1:4 y 1:8. Un resultado positivo es indicado por la formación de una película roja, y un resultado negativo es indicado por un sedimento rojo en el fondo del pocillo. **B.** Autoagregación bacteriana. Cultivos incubados durante la noche fueron centrifugados y re-suspendidos en PBS hasta una densidad óptica de ~0.9 medida a 600 nm (OD₆₀₀). Las suspensiones de bacteria fueron dejadas en estático durante 8 h. La OD₆₀₀ dentro de un 1 cm de la superficie de cada suspensión fue cuantificada (línea punteada). Para cada cepa, un tubo paralelo fue agitado (línea continua) antes de cada medición. Barras de error representan desviación estándar (s.d) (n=2). **C.** Suspensiones de bacteria luego de 8 h de incubación en estático a temperatura de habitación. **D.** Formación de biopelícula cuantificada a 24, 48 y 72 h mediante tinción con cristal violeta. La biopelícula fue determinada como el OD₅₉₅ de cristal violeta eluido. Barras de error representan s.d. (n=3). **E.** Biopelícula producida sobre superficie de poliestireno luego de 72 h. **F.** Adhesión bacteriana a células epiteliales humanas (Caco-2, HT-29 y HEP-2). Datos son expresados como el porcentaje de inóculo inicial recuperado después de 30 min de infección a una multiplicidad de infección de 10 bacterias por célula. Barras de error representan s.d. (n=3). **G.** Tinción Giemsa de bacterias adheridas a monocapa de células HT-29 visualizada mediante microscopía de luz. Magnificación X1.000. Flechas indican bacterias adheridas. * P < 0.005, calculado mediante prueba t de Student (dos colas) relativo a la cepa HB101/pVB1.

6.1.4 Determinación de la participación del gen *hes* en la adhesión de STEC

Teniendo en cuenta la función biológica de *hes*, se intentó determinar la contribución de este gen a la patogenicidad de STEC. Para esto, *hes* fue removido de la cepa E045-00 mediante reemplazo alélico y se determinó si esta delección afecta la capacidad de la bacteria para adherirse a células epiteliales HT-29. Los resultados obtenidos de estos ensayos indicaron que no hay diferencia en los niveles de adherencia entre la cepa silvestre y su mutante isogénica Δhes (**Figura 5**). La complementación de *hes* en trans usando el vector pVB1_*hes* tampoco afectó la capacidad de adhesión de la mutante isogénica Δhes . En conjunto, estos resultados indican que mientras *hes* confiere diferentes fenotipos asociados a colonización a la cepa HB101, en el contexto genético de la cepa E045-00 podría cumplir un papel redundante.

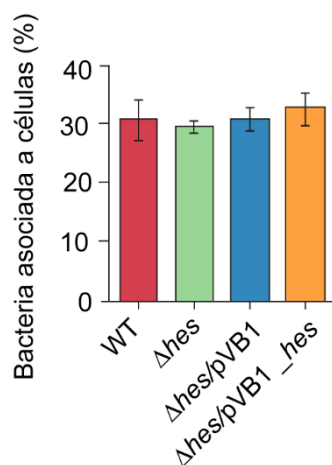


Figura 5. Adhesión de la cepa O113:H21 E045-00 y su mutante isogénica Δhes a células HT-29. Datos son expresados como el porcentaje de inóculo inicial recuperado después de 30 min de infección a una multiplicidad de infección de 10 bacterias por célula. Barras de error representan s.d. (n=3).

6.2 Caracterizar la organización genética de LAA en cepas STEC y determinar si existe una región común (core) de este locus entre ellas.

6.2.1 Análisis de la secuencia de nucleótidos del Locus de Adhesión y Autoagregación

El Locus de Adhesión y Autoagregación fue identificado en las posiciones 385.984 – 472.336 pb en el contig #25 (Número de acceso GenBank: AFDQ01000026) del genoma parcial de la cepa STEC O91:H21 B2F1 (**Figura 6**). En el cromosoma de la cepa B2F1, LAA es una región de DNA de 86 kb adyacente al gen *pheV*-tRNA y flanqueada por secuencias directas repetidas imperfectas (DR1) de 23 pb; los genes *yqgA* y *yghD* están localizados adyacentes a las secuencias DR1, con el primero localizado río arriba y el segundo río abajo. De igual manera, contiene otras secuencias DR (DR2 a DR7) y diversos elementos de secuencias de inserción (IS). Su contenido de G + C es 48%, que es menor al 51% G + C exhibido en promedio por el cromosoma de la cepa B2F1. En total, 80 ORFs fueron identificados en esta región de DNA, incluyendo un gen que codifica una integrasa (ORF1), genes asociados con patogenicidad y diversos genes y pseudogenes que codifican proteínas hipotéticas (**Tabla S3**). Entre los factores de virulencia identificados en esta región de DNA, además de *hes* (ORF5), están genes que codifican adhesinas como *lha* (ORF20) (68) y *Ag43* (ORF67), que además participa en autoagregación (118).

Otros genes con función conocida que fueron identificados en LAA son: el gen *sisA* (ORF2), que codifica una proteína que atenúa la respuesta inflamatoria inducida por cepas *E. coli* uropatogénicas (119); el gen *nmpC* (ORF26), que codifica una proteína asociada con la resistencia a calor (120), que también ha sido caracterizada como antígeno reactivo a sueros de pacientes con SHU (78).

Cuatro ORFs adicionales que codifican potenciales factores de virulencia son: el ORF29, que codifica una nueva variante de la familia de proteínas Serino Proteasas Autotransportadoras de Enterobacteriaceae (SPATE), la cual fue llamada “LAA encoded SPATE” (LesP) (**Tabla S4**); el ORF42, que codifica una proteína con 60.6% de similitud con la proteína PagC de *Salmonella enterica*, la que participa en el fenotipo

de resistencia a suero (121); y los ORF43 y ORF45 (genes *tpsA* y *tpsB*, respectivamente), que codifican un sistema de secreción de dos compañeros que en bacterias Gram negativo participan en diferentes fenotipos de virulencia (122); sin embargo, la función biológica de estas proteínas hipotéticas permanece sin caracterizar.

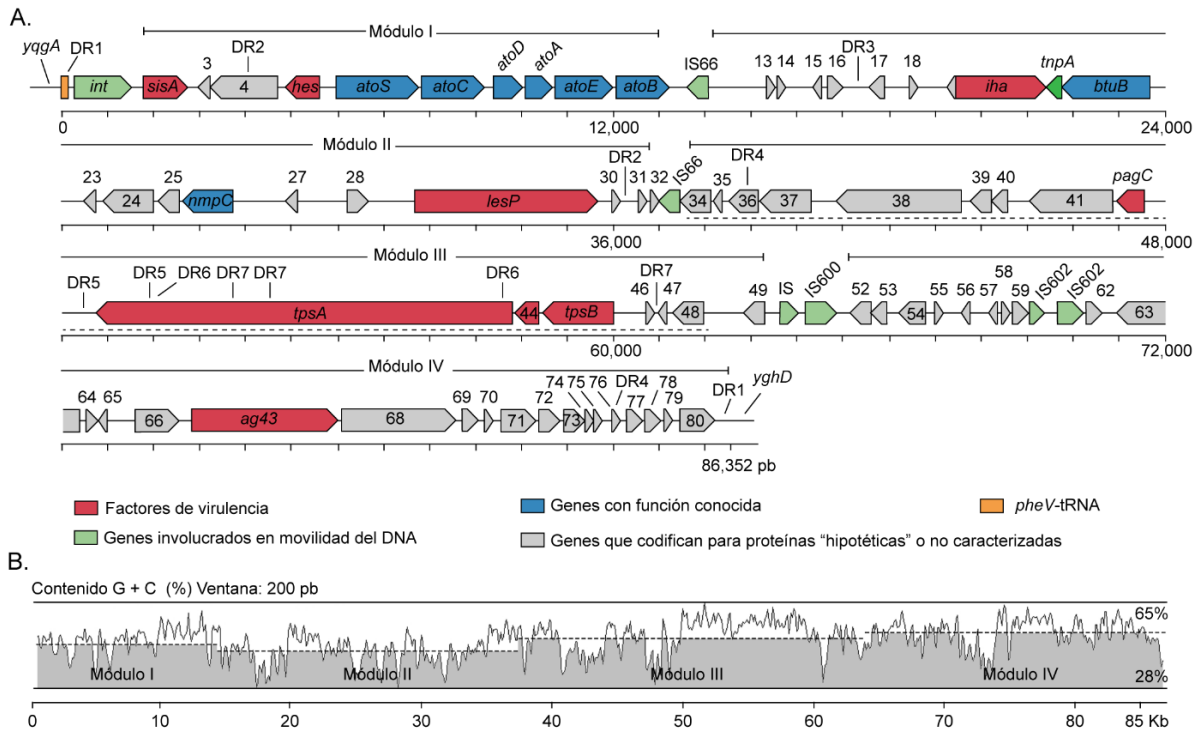


Figura 6. Estructura genética del Locus de Adhesión y Autoagregación. A. Genes predichos y dirección de la transcripción son representados como flechas y son coloreados de acuerdo con la función, como se indica en la leyenda. Se muestran algunos nombres de genes. Esta isla de patogenicidad tiene una estructura de mosaico que está organizada en módulos flanqueados por secuencias de inserción (IS) y secuencias directas repetidas (DR). Línea punteada, región de DNA reportada previamente por Shen et al. (125). **B.** Contenido G + C de cada módulo: I (48.6%), II (43.8%), III (50.2%) y IV (51.7%).

Las islas de patogenicidad generalmente contienen genes que codifican reguladores transcripcionales (123). Estas proteínas pueden controlar la expresión de genes localizados dentro de estos elementos genéticos o en otras partes del genoma. Entre estos reguladores transcripcionales se encuentran miembros de la familia de proteínas tipo AraC y de los sistemas de transducción de señales de dos componentes. En LAA se identificó el ORF28, que codifica una proteína hipotética de la familia AraC que está presente también en la isla LPA (72). Además, los ORF6 y ORF7 codifican el

sistema de dos componentes AtoS-AtoC que regula positivamente la expresión del operón *atoDAEB* (ORF8 – ORF11) (124). No se descarta la posibilidad de que el sistema AtoS-AtoC participe en la regulación de la expresión de otros genes codificados en LAA y también de genes codificados fuera de esta PAI.

6.2.2 Análisis de la organización genética de LAA

El genoma parcial de la cepa STEC O91:H21 B2F1 fue alineado frente a los genomas (parciales o completos) de un grupo de cepas *E. coli* mediante *progressiveMauve* (96). Como se observa en la **Figura 7**, además de la cepa B2F1, LAA está presente en el genoma de otras cuatro cepas STEC LEE-negativo de los serotipos O48:H21, O104:H21, O113:H21 y O174:H21. Adicionalmente, en la cepa STEC LEE-negativo O91:H14 06-3691 fue identificado un segmento de LAA (ORF2 a ORF20) adyacente al gen *selC*-tRNA y un segundo segmento (ORF43-truncado a ORF80) adyacente al gen *pheV*-tRNA. En contraste, LAA no fue identificado en el genoma de cepas comensales y STEC LEE-positivo. Estos resultados sugieren que este locus está asociado con cepas STEC LEE-negativo.

Posteriormente, la secuencia de nucleótidos de LAA identificada en cada genoma fue comparada mediante BLASTn usando la herramienta *EasyFig* (104). A través de este análisis se determinó que, en general, la secuencia de nucleótidos de LAA entre los distintos serotipos STEC LEE-negativo tiene > 90% de identidad (**Figura 8**). Entre las diferencias identificadas se encontraron, particularmente, inserciones de elementos IS y deleciones en el gen *tpsA* (ORF43), probablemente mediadas por procesos de recombinación y escisión de DNA facilitadas por la presencia de varias secuencias DR en este gen (**Figura 6A**).

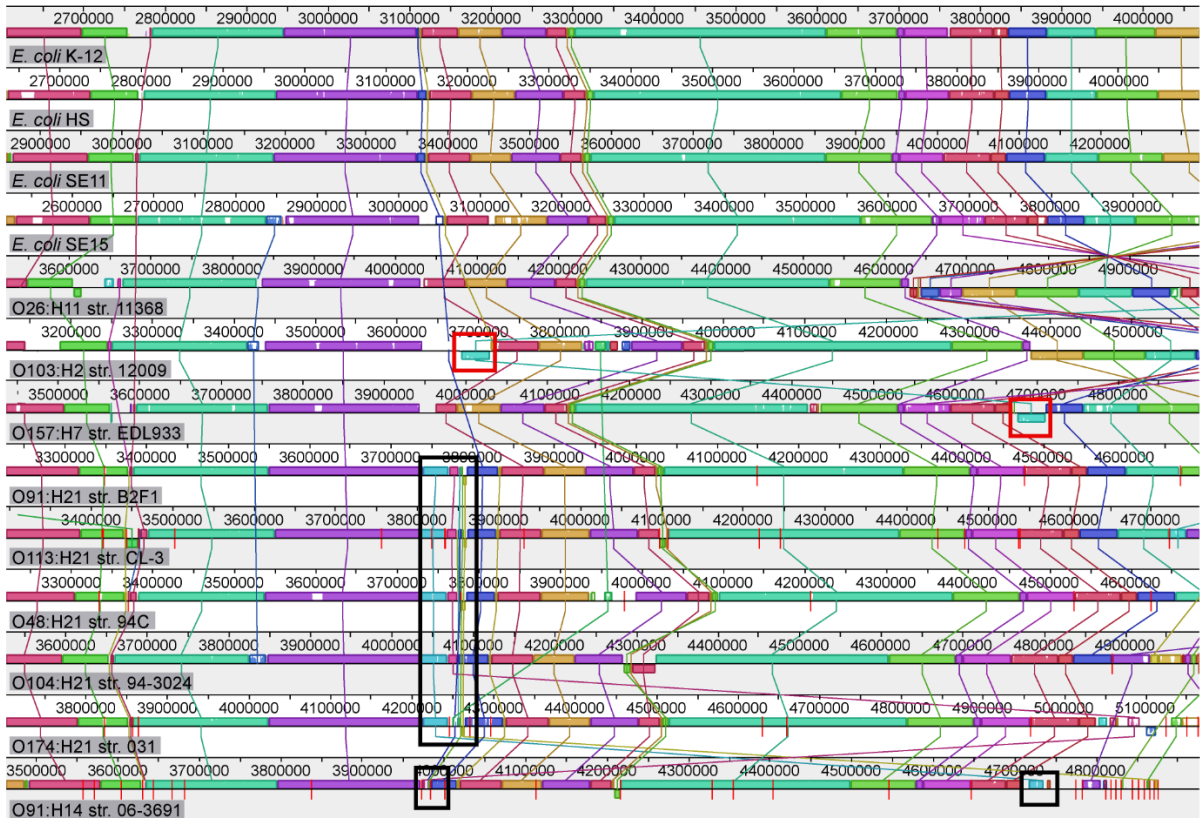


Figura 7. Identificación de LAA en cepas STEC LEE-negativo. Alineamientos entre genomas de cepas STEC (LEE-positivo: O26:H11, O103:H2 y O157:H7; LEE-negativo: O48:H21, O91:H21, O91:H14, O104:H21, O113:H21 y O174:H21), *E. coli* K-12 MG1655 y cepas *E. coli* comensales (HS, SE11 y SE15) mediante progressiveMauve (96). El genoma de la cepa MG1655 se utilizó como referencia. Bloques de colores representan regiones co-lineales y homólogas. Áreas no coloreadas representan secuencias no alineadas que pueden ser específicas para un genoma determinado. Regiones invertidas son identificadas como bloques bajo la línea central. Regiones consecutivas entre líneas rojas indican contigs individuales. La localización de LAA es mostrada como rectángulos negros. La localización de la isla LEE es mostrada como rectángulos rojos.

Por otro lado, la presencia de fragmentos de LAA en lugares diferentes al gen *pheV*-tRNA, junto con la identificación de diversas secuencias DR y elementos IS sugirió que la estructura de este locus podría ser modular. Para probar esto, se analizó la secuencia de nucleótidos de LAA en busca de regiones flanqueadas por elementos IS, secuencias DR y/o teniendo diferentes contenidos G + C, las cuales podrían corresponder a distintos módulos. Cuatro regiones con estas características se identificaron (**Figura 6A**): módulo I (13 kb), desde el ORF1 hasta el ORF11 y flanqueado río abajo por el ORF12, que codifica una transposasa de la familia IS66; módulo II (23 kb), desde el ORF13 hasta el ORF32 y flanqueado río abajo por el ORF33, que codifica una transposasa de la familia IS66; módulo III (26 kb), desde el

ORF34 hasta el ORF49 y flanqueado río abajo por los ORFs 50 y 51, que codifican transposasas de la familia IS600; y módulo IV (21 kb), desde el ORF52 al ORF80 localizado en el extremo 3' terminal de LAA. Además, todas estas regiones de DNA tienen contenidos de G + C diferentes (**Figura 6B**). Posteriormente, se buscaron estos módulos en las bases de datos Islander (100) y PAIDB (101), y se encontró un grupo de otras PAIs que contienen regiones de DNA con identidades > 80%, con excepción del módulo I (**Figura 9**). El módulo II tiene homología con la isla LPA (**Figura 9A**), mientras que el módulo III fue previamente descrito en la cepa STEC LEE-negativo O113:H21 CL3 y llamado PAI I_{CL3} (125) (**Figura 6A**, línea punteada; **Figura 9B**). Un grupo de genes similar al módulo III se encontró en la isla genómica GI_{pheV-CR_{ICC168}} de *Citrobacter rodentium* (85) y como parte de la PAI-IAL862 (con delección de los ORFs 43-45) (126). De igual manera, el módulo IV se asemeja a un grupo de genes presentes en diferentes PAIs. En conjunto, estos resultados confirman la estructura de mosaico de LAA.

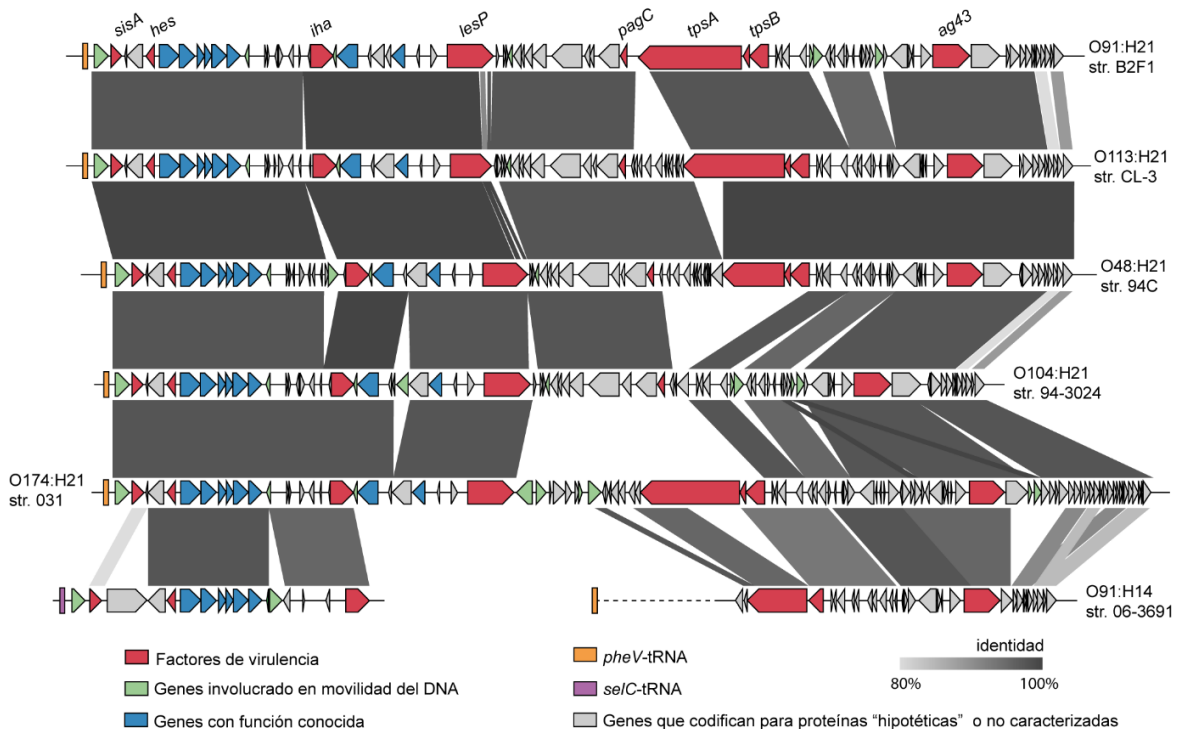


Figura 8. Comparación de la estructura genética de LAA presente en seis cepas STEC LEE-negativo. Genes predichos y la dirección de la transcripción se muestran como flechas. ORFs están coloreados de acuerdo a la función del gen, como se indica en la leyenda. Se muestran los nombres de los genes de virulencia. Regiones conservadas son mostradas en gris y la intensidad del color indica los niveles de identidad de nucleótidos, desde 80 a 100%. La figura se preparó usando EasyFig (104).

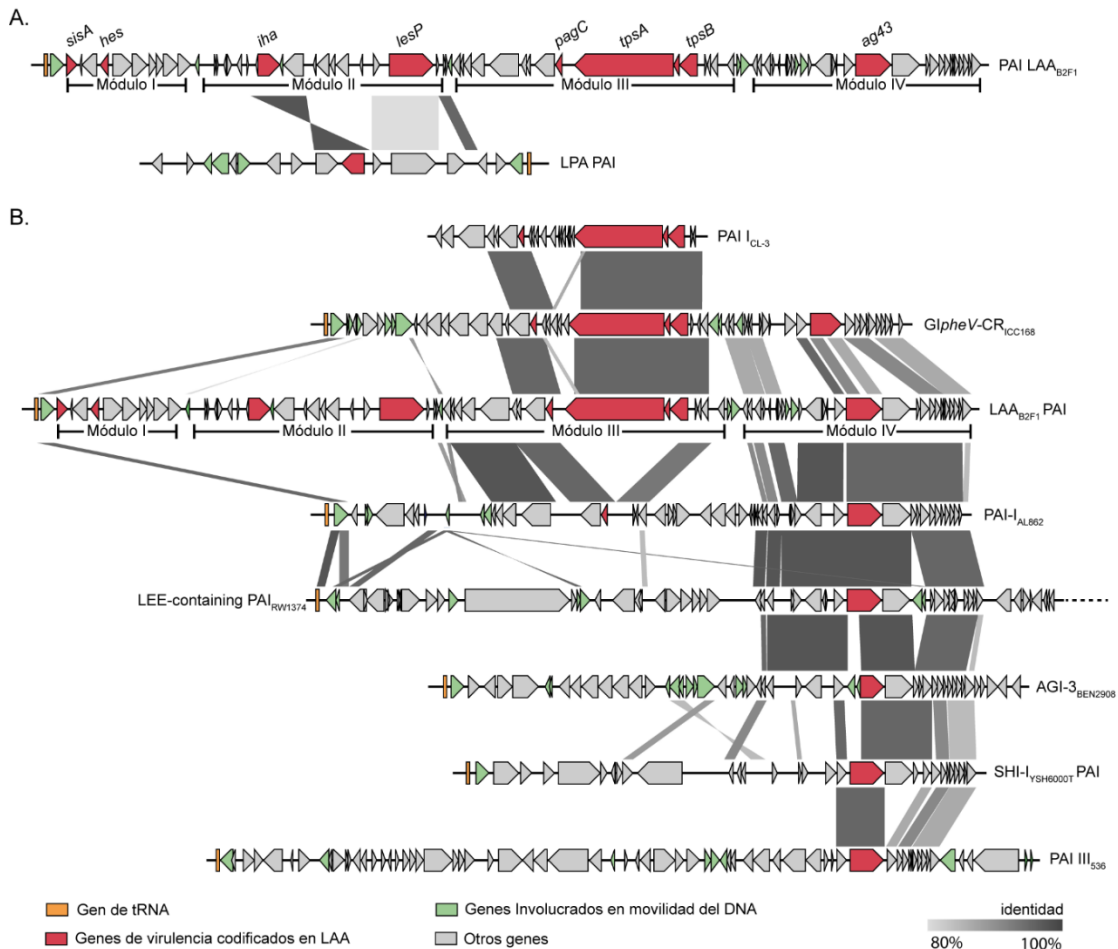


Figura 9. Comparación de la estructura genética de LAA_{B2F1} y PAIs relacionadas. Genes predichos y dirección de la transcripción son representados como flechas y están coloreados de acuerdo a la leyenda. Los nombres de genes de virulencia codificados en LAA son mostrados en el panel superior. Regiones conservadas son mostradas en gris y la intensidad del color indica los niveles de identidad de nucleótidos, desde 80 a 100%. **A.** Comparación entre LAA_{B2F1} y LPA (Genbank: AJ278144). **B.** Comparación entre LAA_{B2F1}, PAI-*ICL3* (Genbank: AY275838), *GlpheV-CR*_{ICC168}, PAI-*I_{AL862}* (Genbank: GQ497943), LEE-containing PAI_{RW1374} (Genbank: AJ303141; obsérvese que esta isla es mostrada parcialmente), AGI-3_{BEN2908} (Genbank: AY857617), SHI-*I_{YSH6000T}* PAI (Genbank: AF200692) y PAI-*III₅₃₆* (Genbank: X16664). La figura fue preparada usando EasyFig (104).

6.2.3 Determinación de la distribución de LAA en STEC

La distribución de LAA fue determinada en una colección de cepas STEC (LEE-positivo: 119 cepas; LEE-negativo: 48 cepas) mediante un ensayo de PCR múltiple que fue diseñado para la identificación de los módulos I, II y III (**Figura 10**). No fue posible obtener partidores específicos para el módulo IV debido a que los genes codificados en esta región están ampliamente distribuidos en *E. coli* (**Figura 9B**). Similar a lo observado con la distribución del gen *hes*, los módulos de LAA fueron identificados solamente en cepas STEC LEE-negativo. Entre ellas, 24/48 (50%) cepas

fueron positivas por PCR para los tres módulos, mientras que 6/48 (12.5%) cepas fueron positivas para dos (**Tabla S1**). Estos resultados confirman que LAA está asociado con cepas STEC LEE-negativo.

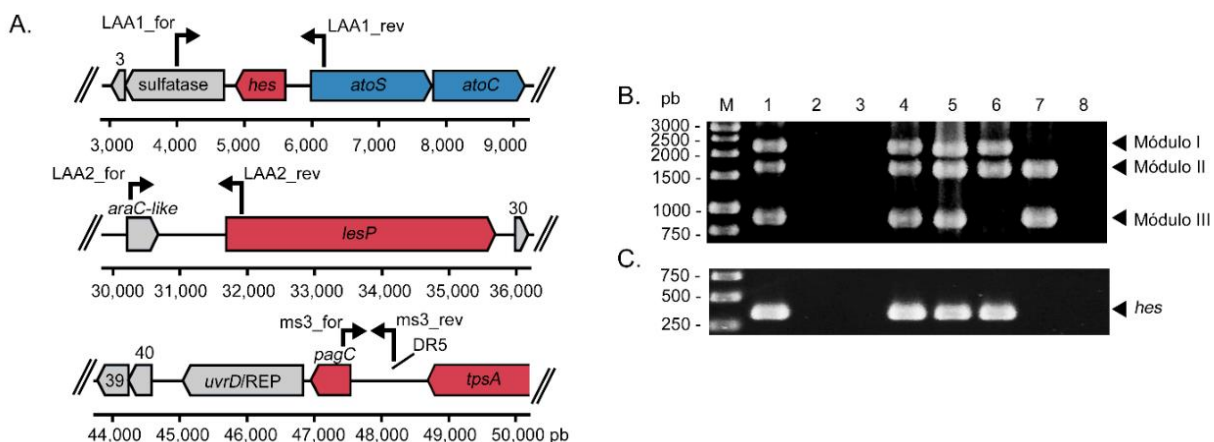


Figura 10. Detección de los módulos I, II y III de LAA mediante PCR múltiple. **A.** Diagrama esquemático mostrando las regiones de LAA que son amplificadas en el ensayo de PCR múltiple. Flechas negras indican los partidores usados para amplificar cada región de DNA. **B y C.** Electroforesis en gel de agarosa (1%) de productos de PCR de cepas representativas. La detección de *hes* se realizó con los partidores *hes_det1* + *hes_det2*. Obsérvese en el carril 7 la ausencia del módulo I y del gen *hes*. Carriles: 1, O113:H21 E045-00; 2, O113:H21 E045-00ΔLAA; 3, O113:H21 5_1; 4, O104:H21 B45; 5, O91:H21 A13; 6, O181:H49 72_1; 7, O113:H21 IH39595/07C; 8, control negativo. M, marcador de tamaño molecular.

De igual manera, en un trabajo colaborativo con la Dra. Nora Padola (Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina), realizado durante una estancia corta de investigación, se determinó la distribución de LAA en cepas STEC LEE-negativo aisladas de bovinos (n = 311), alimentos (n = 37) y el medio ambiente (n=6) durante los años 2000 y 2015 en Argentina (127). Es importante mencionar que, en el estudio mencionado, la presencia de LAA fue confirmada cuando los módulos I, II y III y *ag43* (como marcador del módulo IV) fueron identificados mediante PCR usando los partidores descritos en la **Figura 10**. En dicho trabajo, LAA fue identificado en 161/354 (46%) cepas de los serogrupos O15, O20, O39, O46, O88, O91, O113, O117, O130, O136, O139, O141, O153, O160, O163, O168, O171, O174 y O178. De acuerdo con el origen, LAA fue identificado en 19/37 (56%), 121/241 (50%) y 21/69 (30%) cepas aisladas desde alimentos, bovinos adultos y terneros, respectivamente. Adicionalmente, similar a lo observado en nuestra colección de cepas STEC, un porcentaje de cepas de Argentina contenía menos de 3 módulos. En conjunto, estos

resultados demuestran la amplia distribución de LAA entre cepas STEC LEE-negativo y sugieren que sus módulos pueden movilizarse (ser adquiridos o perdidos) independientemente de la estructura completa de esta PAI.

Posteriormente, se intentó relacionar la distribución de los módulos de LAA con la historia evolutiva de las cepas STEC que los contienen. Para esto, se realizó un análisis genómico comparativo de los genomas (parciales o completos) de 115 cepas STEC LEE-negativo, 7 cepas STEC LEE-positivo, 2 cepas de otros patotipos de *E. coli*, *E. coli* K-12 MG1655 y 3 cepas de *E. coli* comensales, que están disponibles en GenBank (**Tabla S2**). Primero, todas las secuencias genómicas parciales fueron alineadas y ordenadas usando progressiveMauve (96). Posteriormente, se determinó la relación filogenética entre las cepas con base en el análisis de SNPs del genoma completo usando el servidor CSI Phylogeny 1.4 (108). Adicionalmente, se determinó el serotipo y el filogrupo usando las herramientas SerotypeFinder (91) y ClermonTyping (95), respectivamente. Finalmente, mediante búsquedas BLASTn se determinó la presencia de LAA, la distribución de sus módulos y los loci de tRNA cercanos a estas secuencias. Los resultados obtenidos indican que las cepas STEC LEE-negativo son filogenéticamente diversas. En la **Figura 11** se muestra el árbol filogenético de máxima verosimilitud basado en el análisis de SNPs. En general, la designación de serotipo y filogrupo fueron consistentes con la topología del árbol, con la excepción del serotipo O8:H19, en el que sus cepas fueron separadas en tres clados diferentes. En contraste, los serogrupos fueron polifiléticos. Por ejemplo, varios serotipos de los serogrupos O91, O113, O104 y O174 fueron encontrados en clados lejanos, lo que es consistente con estudios previos que demuestran esta diversidad genética (128, 129). Es de destacar que la estructura completa (los cuatro módulos) de LAA fue identificada adyacente al gen *pheV*-tRNA en 40/115 (34.8%) cepas STEC LEE-negativo de diferentes serotipos (-:H25, O8:H19, O22:H8, O38:H21, O48:H21, O74:H42, O79:H7, O88:H25, O91:H21, O96:H19, O104:H21, O113:H21, O116:H21, O130:H38, O134:H38, O163:H19, O168:H8, O171:H2, O174:H2, O174:H21 y O185:H7), incluyendo 3 y 6 cepas aisladas de casos de CH y SHU, respectivamente. En dos cepas adicionales STEC LEE-negativo, una aislada de un caso de CH (O166:H28 FHI92) y la otra de heces humanas (O91:H21 FHI59, diagnóstico no

disponible), se identificaron también cada uno de los cuatro módulos, pero adyacentes a diferentes genes de tRNA (*pheV*, *selC* y *thrW*). Todos los aislados anteriores pertenecen al filogruppo B1, excepto la cepa FHI92, que pertenece al filogruppo E. No se identificó LAA en dos aislados O91:H14, indicando que este locus no está presente en todas las cepas de un mismo serotipo. Por otro lado, se identificaron módulos individuales (en algunos casos haciendo parte de otras estructuras de mosaico) en varias cepas STEC LEE-negativo, incluyendo aislados de filogruppos distintos de B1. En estos casos, los módulos están adyacentes a diferentes genes de tRNA (*pheV*, *selC*, *ileX* y *serX*). De esta manera, cualquiera que sea el mecanismo (adquisición o delección), este análisis demuestra la amplia distribución y movilización de los módulos de LAA entre STEC LEE-negativo. Más importante aún, este análisis confirma que LAA es una PAI modular que está asociada a STEC LEE-negativo.

A continuación, se investigó la historia evolutiva de esta PAI. Para esto, se determinaron las relaciones genéticas entre 42 secuencias de LAA (con los 4 módulos) con base en el análisis de SNPs. Además, se determinó la variabilidad de los genes codificados en cada una estas secuencias. En ambos análisis, LAA_{B2F1} fue usado como referencia. Como resultado, el árbol filogenético de máxima verosimilitud indica que dos linajes principales de LAA (LAA-1 y LAA-2) se encuentran en diferentes cepas de un mismo serotipo, incluyendo O91:H21, O96:H19 y O174:H21 (**Figura 12A**). Por lo tanto, se infiere que en estos serotipos LAA ha sido adquirido más de una vez en eventos independientes. La mayor variabilidad genética entre estos dos linajes fue encontrada en los módulos III y IV, los que a su vez también presentan más variación dentro de cada linaje (**Figura 12B**), indicando que estas regiones de DNA pueden estar sometidas a diferentes presiones evolutivas.

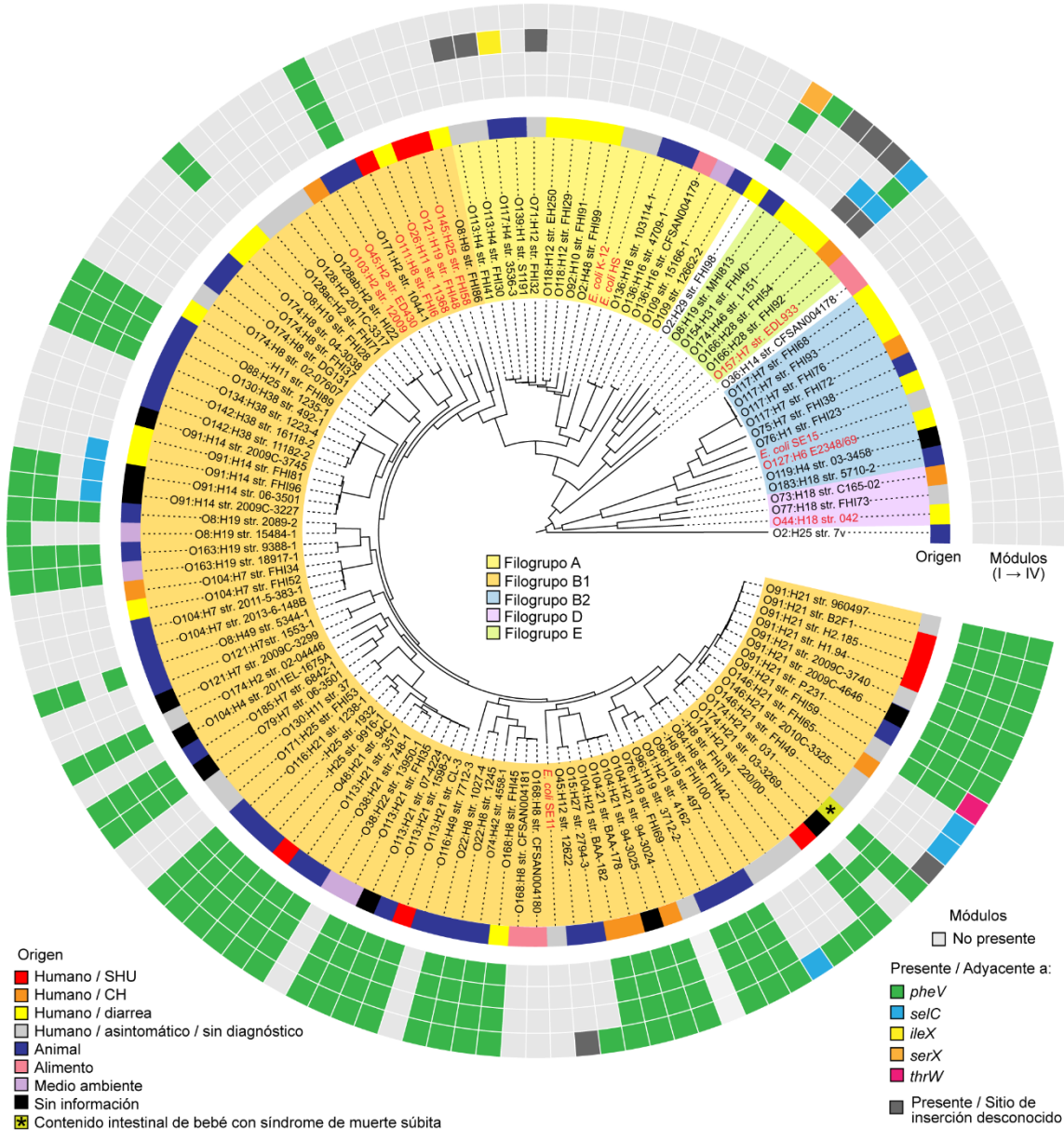


Figura 11. Relaciones filogenéticas y distribución de módulos de LAA entre cepas STEC LEE-negativo y cepas *E. coli* relacionadas. 128 genomas parciales y completos de *E. coli* (115 cepas STEC LEE-negativo y 13 cepas *E. coli* relacionadas) fueron usados para construir el árbol filogenético de máxima verosimilitud basado en el análisis de SNPs (167,167 SNPs dentro de 3,008,649 posiciones encontradas en todos los genomas). Los filogrupos son indicados por colores en el anillo interno de acuerdo con la leyenda. El origen de cada aislado es indicado en el anillo intermedio por cuadros con color de acuerdo con la leyenda. La presencia de LAA, de sus módulos y sitios de inserción son indicados en los anillos externos por cuadros con color de acuerdo con la leyenda. Las cepas STEC LEE-negativo y cepas *E. coli* relacionadas son resaltadas en negro y rojo, respectivamente.

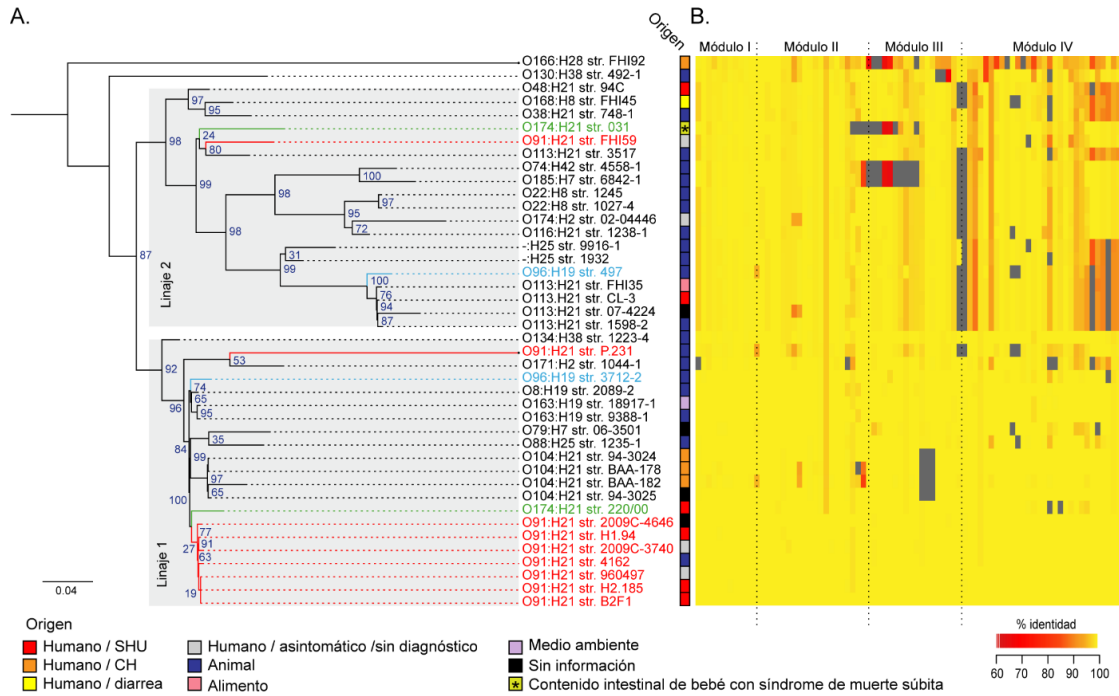


Figura 12. Relaciones genéticas entre 42 secuencias de LAA. **A.** Árbol filogenético de máxima verosimilitud basado en el análisis de SNPs (677 SNPs dentro de 49,427 posiciones encontradas en todas las secuencias) de secuencias de LAA. La secuencia LAA_{B2F1} fue usada como referencia. Los linajes son indicados en las ramas. El origen de cada aislado es indicado por cajas de color de acuerdo con la leyenda. Los valores de Bootstrap se indican en el árbol. Observar que ambos linajes están presentes en cepas que pertenecen a los serotipos O91:H21 (resaltado en rojo), O96:H19 (resaltado en azul) y O174:H21 (resaltado en verde). **B.** Mapa de calor mostrando la presencia, ausencia y % de identidad de los 80 genes codificados en LAA. La presencia y variación (identidad desde 60% a 100%) para cada gen es indicado por la intensidad de color (rojo a amarillo), como se muestra en la leyenda. La ausencia fue definida como una identidad o cobertura de secuencia < 60% y es indicada en color gris. La secuencia LAA_{B2F1} fue usada como referencia. El mapa de calor está dividido en módulos para mostrar la variabilidad genética dentro de éstos.

6.3 Analizar mediante genómica comparativa los genomas de cepas STEC LEE-negativo e identificar su repertorio de genes de virulencia

6.3.1 Análisis *in silico* de genomas de cepas STEC LEE-negativo

Como fue mencionado, el número de casos clínicos asociados con cepas STEC LEE-negativo ha aumentado recientemente en varios países. En ese sentido, la adquisición de elementos genéticos móviles (MGEs; del inglés “Mobile genetic elements”), entre ellos la isla LAA, podría aumentar el potencial de virulencia de estas cepas y favorecer su emergencia. Para evaluar esta hipótesis, se realizó un análisis genómico a gran escala de 367 genomas (parciales o completos) de cepas STEC LEE-negativo, incluyendo 35 genomas secuenciados en esta tesis y 332 disponibles en

GenBank. Las cepas analizadas fueron aisladas desde humanos ($n = 150$), animales ($n = 123$), alimentos ($n = 45$), el medio ambiente ($n = 4$) y un grupo final carecía de información sobre su fuente ($n = 45$) (**Tabla S2, Figura 13**). En la mayoría de los registros genómicos, con excepción de 55, la fecha de aislamiento fue informada (de 1947 a 2016). Un aumento en el número de cepas aisladas, observado entre 2013-2014, correspondió a un estudio prospectivo multicéntrico, STEC-ID-net, llevado a cabo en las regiones holandesas de Groningen y Rotterdam (**Figura 13B**) (23). Cabe señalar que la mayoría de las cepas se aislaron después de 2008, en particular de casos clínicos, por lo que representan un panorama epidemiológico actualizado de este subgrupo de STEC.

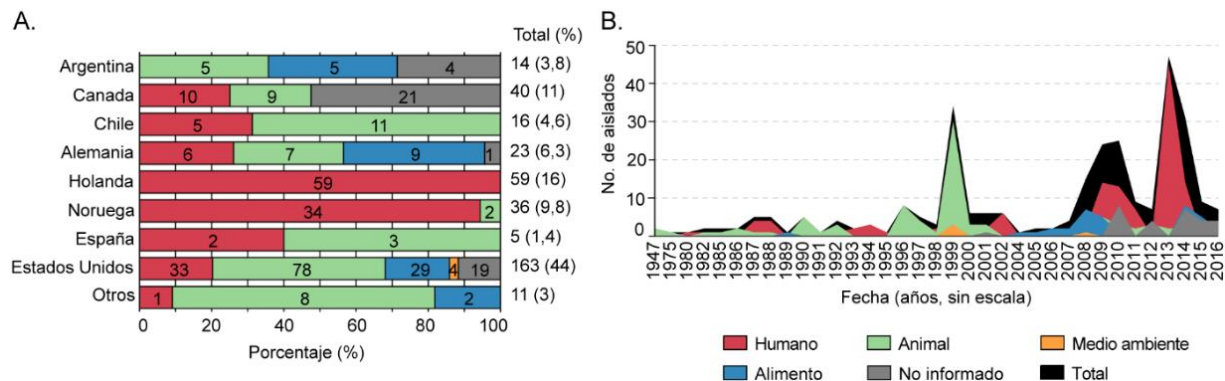


Figura 13. Distribución geográfica y temporal de las cepas STEC LEE-negativo analizadas. Se indica el número de cepas por fuente de aislamiento (ver leyenda) y país. **A.** Distribución geográfica. Los aislamientos enumerados en categorías de países como "Otros" incluyen: 1 aislado clínico de Bélgica, 2 aislados de alimentos de Uruguay y 8 aislados de animales de Australia (1), China (1), India (1), Italia (1), Japón (2), Puerto Rico (1) y el Reino Unido (1). **B.** Distribución temporal. La fecha de aislamiento estaba disponible para 312 de las 367 cepas analizadas (de 1947 a 2016). Se observa un aumento en la secuenciación de cepas clínicas a partir del año 2008.

El análisis *in silico* realizado se basó en primer lugar en la determinación del serotipo y del contenido de genes de virulencia y de PAIs. Posteriormente, se realizó un análisis filogenético para determinar relaciones genéticas entre las cepas. Adicionalmente, todo lo anterior fue relacionado con los datos epidemiológicos de las cepas, incluyendo la fuente de aislamiento y el origen geográfico.

6.3.1.1 Identificación del serotipo y contenido de genes de virulencia

En total, se identificaron 101 serotipos (Tabla S2); sin embargo, en 15 cepas se identificó sólo el antígeno "H". Los serotipos más frecuentes fueron O91:H14 ($n = 29$),

O113:H21 (n = 29), O91:H21 (n = 22), O174:H21 (n = 18), O146:H21 (n = 17), O174:H8 (n = 12), O22:H8 (n = 10) y O8:H19 (n = 9). Estudios previos también han demostrado esta gran diversidad de serotipos entre cepas STEC LEE-negativo (31, 130, 131).

Posteriormente, se determinó la presencia / ausencia de genes *stx* y de otros genes de virulencia presentes estas cepas. La frecuencia de detección para cada gen se presenta en la **Tabla 4**. Entre las cepas (n = 367), 266 (72,5%) contienen sólo un gen *stx* y 110 (27,5%) tienen más de uno (**Tabla S5**). Los genes *stx2b*, *senB*, *eibG*, *tia*, *ag43-l* y *sigA* fueron significativamente más frecuentes en cepas clínicas en comparación con cepas aisladas de animales. En contraste, los genes *stx1d*, *stx2a*, *stx2e*, *sta1*, *saa*, *hra1*, *lpfA*, *lesP* y *espP* se encontraron en una proporción significativamente mayor de cepas aisladas de animales en comparación con cepas clínicas.

Un estudio previo reportó una alta frecuencia de detección de genes que codifican bacteriocinas entre cepas STEC LEE-negativo aisladas de humanos (23). Las bacteriocinas son péptidos antimicrobianos producidos por enterobacterias que son activos contra cepas estrechamente relacionadas (132). *E. coli* produce dos tipos de bacteriocinas, clasificadas con base en su peso molecular como colicinas (25-80 kDa) y microcinas (<10 kDa). Notablemente, los genes *mchB*, *mchC* y *mchF* (clúster *mch*), involucrados en la síntesis de microcinas, fueron identificados en una proporción significativamente mayor de cepas clínicas en comparación con cepas aisladas de otras fuentes. Por lo tanto, es posible que la producción de microcinas represente una ventaja adaptativa para estas bacterias durante la infección en el humano. En conjunto, estos datos indican que las cepas STEC LEE-negativo contienen varios genes que proporcionan fenotipos asociados con toxicidad, adhesión y colonización. Si bien algunos de estos genes contribuyen a la virulencia, otros pueden ser necesarios para la colonización o persistencia de la bacteria en su reservorio animal.

Tabla 4. Distribución de genes de virulencia entre cepas STEC LEE-negativo

Gen	Frecuencia total (n=367), n (%)	Fuente de aislamiento, n (%) ¹				Significancia (valor P) ²	
		Humano (n=150)	Animal (n=123)	Alimento (n=45)	Sin info. (n=45)	Humano vs. Animal	Humano vs. Alimento
Shiga toxina³							
<i>stx1a</i>	101 (27.5)	50 (33.3)	35 (28.5)	8 (17.8)	6 (13.3)	0.431	0.060
<i>stx1c</i>	71 (19.3)	36 (24.0)	20 (16.3)	8 (17.8)	7 (15.6)	0.133	0.424
<i>stx1d</i>	10 (2.7)	1 (0.7)	6 (4.9)	0 (0.0)	3 (6.7)	0.048	-
<i>stx2a</i>	105 (28.6)	27 (18.0)	44 (35.8)	19 (42.2)	12 (26.7)	0.001	0.001
<i>stx2b</i>	62 (16.9)	35 (23.3)	15 (12.2)	3 (6.7)	9 (20.0)	0.019	0.017
<i>stx2c</i>	33 (9.0)	17 (11.3)	8 (6.5)	4 (8.9)	4 (8.9)	0.208	0.788
<i>stx2d</i>	56 (15.3)	25 (16.7)	17 (13.8)	5 (11.1)	9 (20.0)	0.614	0.482
<i>stx2e</i>	16 (4.4)	2 (1.3)	10 (8.1)	1 (2.2)	3 (6.7)	0.007	0.081
<i>stx2g</i>	7 (1.9)	0 (0.0)	3 (2.4)	8 (2.2)	3 (6.7)	-	-
Otras toxinas							
<i>ehxA</i>	179 (48.8)	75 (50.0)	60 (48.8)	25 (55.6)	15 (33.3)	0.903	0.611
<i>subA</i>	130 (35.4)	60 (40.0)	40 (32.5)	14 (31.1)	14 (31.1)	0.209	0.299
<i>cdtB</i>	53 (14.4)	15 (10.0)	22 (17.9)	9 (20.0)	7 (15.6)	0.075	0.117
<i>astA</i>	62 (16.9)	20 (13.3)	21 (17.1)	9 (20.0)	12 (26.7)	0.400	0.338
<i>sta1</i>	23 (6.3)	3 (2.0)	11 (8.9)	7 (15.6)	1 (2.2)	0.012	0.001
<i>stb</i>	4 (1.1)	1 (0.7)	2 (1.6)	0 (0.0)	1 (2.2)	0.590	-
<i>senB</i>	35 (9.5)	26 (17.3)	5 (4.1)	1 (2.2)	2 (4.4)	0.001	0.012
Marcadores moleculares de STEC LEE-negativo							
<i>saa</i>	108 (29.4)	33 (22.0)	50 (40.7)	13 (28.9)	9 (20.0)	0.001	0.423
<i>hes</i>	150 (40.9)	61 (40.7)	57 (46.3)	16 (35.6)	15 (33.3)	0.390	0.604
<i>eibG</i>	33 (9.0)	27 (18.0)	1 (0.8)	1 (2.2)	4 (8.9)	<0.001	0.013
<i>sab</i>	17 (4.6)	9 (6.0)	6 (4.9)	2 (4.4)	0 (0.0)	0.793	1.000
Genes de adhesinas							
<i>tia</i>	121 (33)	70 (46.7)	27 (22.0)	11 (24.4)	12 (26.7)	<0.001	0.009
<i>hra1</i>	43 (11.7)	8 (5.3)	20 (16.3)	6 (13.3)	9 (20.0)	0.004	0.095
<i>lpfA</i>	309 (84.2)	120 (80.0)	111 (90.2)	37 (82.2)	37 (82.2)	0.027	0.832
<i>lha</i>	273 (74.4)	126 (84.0)	91 (74.0)	29 (64.4)	24 (53.3)	0.0502	0.006
<i>ag43-I</i>	277 (75.5)	128 (85.3)	88 (71.5)	24 (53.3)	33 (73.3)	0.007	<0.001
<i>ag43-II</i>	96 (26.2)	36 (24.0)	29 (23.6)	16 (35.6)	14 (31.1)	1.000	0.177
SPATEs							
<i>lesP</i>	140 (38.1)	48 (32.0)	57 (46.3)	17 (37.8)	17 (37.8)	0.018	0.589
<i>espP</i>	94 (25.6)	28 (18.7)	40 (32.5)	15 (33.3)	8 (17.8)	0.011	0.043
<i>espl</i>	61 (16.6)	31 (20.7)	19 (15.4)	5 (11.1)	6 (13.3)	0.340	0.219
<i>epeA</i>	16 (4.4)	10 (6.7)	4 (3.3)	1 (2.2)	1 (2.2)	0.319	0.311
<i>pic</i>	14 (3.8)	8 (5.3)	2 (1.6)	1 (2.2)	2 (4.4)	0.120	0.469
<i>sigA⁴</i>	17 (4.6)	16 (10.7)	0 (0.0)	0 (0.0)	1 (2.2)	-	-
Colicinas							
<i>cma</i>	52 (14.2)	22 (14.7)	12 (9.8)	9 (20.0)	9 (20.0)	0.270	0.485
<i>Cba</i>	70 (19.1)	29 (19.3)	22 (17.9)	9 (20.0)	10 (22.2)	0.876	1.000
<i>celb</i>	93 (25.3)	31 (20.7)	36 (29.3)	10 (22.2)	13 (28.9)	0.120	0.836
Microcinas							
<i>mcmA</i>	16 (4.4)	9 (6.0)	5 (4.1)	0 (0.0)	2 (4.4)	0.585	-
<i>mchB</i>	76 (20.7)	57 (38.0)	10 (8.1)	2 (4.4)	7 (15.6)	<0.001	<0.001
<i>mchC</i>	75 (20.4)	56 (37.3)	10 (8.1)	2 (4.4)	7 (15.6)	<0.001	<0.001
<i>mchF</i>	90 (24.5)	60 (40.0)	17 (13.8)	5 (11.1)	8 (17.8)	<0.001	<0.001

¹Las cepas aisladas del medio ambiente no se muestran debido al bajo número de representantes con el que se cuenta (n = 4). ²Prueba exacta de Fisher (dos colas). Un valor de P < 0.05 se consideró significativo. Los valores significativos se muestran en negrita. ³Ninguna de las cepas fue positiva para el gen *stx2f*. ⁴El gen *sigA* estuvo presente casi exclusivamente en cepas clínicas, aunque con muy baja frecuencia.

6.3.1.2 Identificación de islas de patogenicidad

Se determinó la distribución de las islas SE-PAI, LPA, HPI y LAA entre las cepas. Es importante mencionar que varios de los genes de virulencia detectados previamente están contenidos en PAIs. Por ejemplo: *tia* y *subA* se localizan en SE-PAI; *iha* y *espl* en LPA; y *hes*, *iha*, *lesP* y *ag43-I* en LAA; sin embargo, no se puede descartar que estos genes también estén localizados en otras regiones del cromosoma o en elementos extracromosómicos. Por lo tanto, las secuencias de nucleótidos

completas de las PAIs se usaron para realizar búsquedas BLASTn en cada genoma. Además, cuando una cepa positiva para cualquiera de estos genes carecía de la PAI correspondiente, se analizaba entonces el contexto genético del gen. Como resultado, se identificaron dos regiones de DNA, hasta el momento no caracterizadas, que contienen el gen *tia* y que exhiben características comunes a las PAIs.

Una de estas regiones de DNA tiene 48 kb y está adyacente al gen *se/C-tRNA* en el contig 13 del genoma parcial de la cepa STEC O174:H21 MOD1-EC1633 (Número de acceso GenBank: NJVC01000013) (**Figura 14**). Los ORFs identificados en esta región se describen en la **Tabla S6**. Esta región tiene las siguientes características: i) contiene un gen que codifica una integrasa (ORF1), localizado río abajo del locus *se/C-tRNA*; ii) está flanqueada por secuencias DR de 23 pb, iii) posee un contenido G + C de 48% que es menor al 51% (valor promedio) del cromosoma de la cepa MOD1-EC1633; iv) contiene varios elementos IS; y v) contiene genes de virulencia, entre ellos el gen *tia* (ORF4), implicado en la invasión de algunas cepas STEC LEE-negativo (75), y un grupo de genes (ORFs 14-21) que codifica un miembro del sistema de inhibición del crecimiento dependiente del contacto (CDI). Las bacterias CDI⁺ exportan proteínas filamentosas (CdiA) que se unen a bacterias vecinas y translocan toxinas inhibitoras del crecimiento (133). Específicamente, el ORF20 codifica una proteína que tiene 89,3% de similitud con la proteína CdiA prototípica (Número de acceso GenBank: AAZ57198.1) de la cepa *E. coli* EC93, mientras que el ORF21 codifica un homólogo de CdiB que es la proteína β -barril que exporta a CdiA a través de la membrana externa. Debido a que esta región de DNA contiene el gen *tia* relacionado con la invasión bacteriana y codifica un sistema CDI, se denominó como Locus de Invasión e Inhibición del Crecimiento Dependiente del Contacto (LIC).

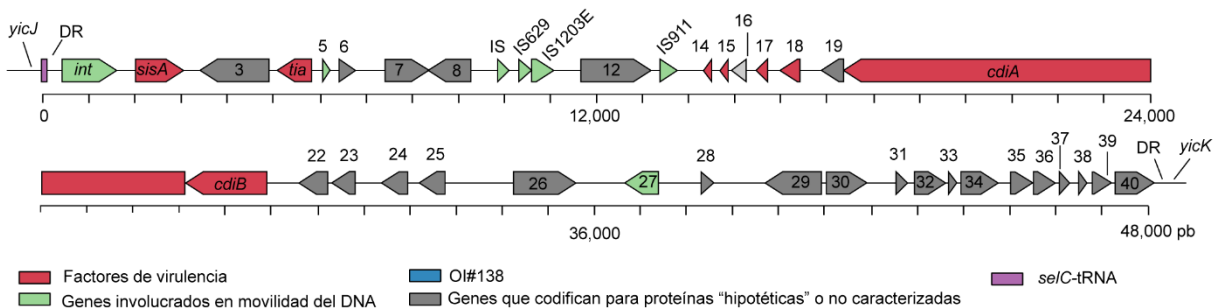


Figura 14. Estructura genética del Locus de Invasión e Inhibición del Crecimiento Dependiente del Contacto. Los genes predichos y la dirección de la transcripción se representan como flechas y son coloreados de acuerdo con la función, como se indica en la leyenda. Se muestran los nombres de algunos genes. Esta región de DNA está localizada adyacente al gen *se/C-tRNA* en el contig 13 (Número de acceso GenBank: NJVC01000013) del genoma parcial de la cepa STEC O174:H21 MOD1-EC1633.

La segunda región de DNA que contiene el gen *tia* se identificó en el contig 57 del genoma parcial de la cepa STEC O22:H8 MOD1-EC3763 (Número de acceso GenBank: NJSE01000057.1), el cual tiene sólo 13 kb y no contiene toda la secuencia de la PAI. Esta región se buscó en los otros genomas de cepas STEC LEE-negativo, pero no se encontró ninguno que tuviera la secuencia completa de la PAI dentro de un contig. Una búsqueda adicional en GenBank permitió encontrar una secuencia homóloga (99% de identidad) en el contig 3 (214 kb) del genoma parcial de la cepa *E. coli* FCH1 (Número de acceso GenBank: AYKD01000003.1). Análisis adicionales revelaron que el fragmento de 13 kb hace parte de una PAI de 71 kb (**Figura 15**) adyacente al gen *pheV-tRNA* en el cromosoma de la cepa FCH1. Los ORFs identificados en esta región se enumeran en la **Tabla S7**. Posteriormente, los contigs de la cepa MOD1-EC3763 se mapearon frente al genoma parcial de la cepa FCH1 (para detalles, ver métodos). A través de este análisis se logró identificar el 99,6% de la PAI (70.923 pb, 99,8% de identidad) dentro de 7 contigs concadenados en el genoma parcial de la cepa MOD1-EC3763 (**Figura S1**). Posteriormente, se mapearon otros genomas parciales de cepas STEC LEE-negativo y se logró identificar en algunos de ellos la estructura completa de esta PAI. Es importante mencionar que la cepa FCH1 es un aislado humano que carece de genes *stx* (134), lo que indica que esta PAI no está restringida a STEC. Entre las características identificadas en esta región de DNA están: i) contiene un gen que codifica para una integrasa (ORF1), localizado río abajo del locus *pheV-tRNA*; ii) la región está flanqueada por secuencias DR

imperfectas de 23 pb; iii) contiene elementos IS; y iv) contiene varios genes de virulencia. Su contenido de G + C (50.5%) es similar al valor promedio exhibido por el cromosoma de la cepa FCH1. Entre los genes de virulencia contenidos en esta PAI están: *tia* (ORF3); el ORF14, que codifica una proteína que presenta 88.1% y 86.5% de similitud con la proteína prototípica CdiA y el homólogo CdiA codificado en la isla LIC, respectivamente; los operones *vpe* (ORFs 34-37) y *deoK* (ORFs 39-42), localizados en varias islas genómicas y que promueven la colonización intestinal de cepas *E. coli* patógenas en modelos murinos de infección (135, 136); y el ORF54, que codifica una proteína que tiene 87.6% y 79.8% de similitud con Cah (Número de acceso GenBank: AAG55356.1) y Ag43 (Número de acceso GenBank: AUG16753.1), respectivamente, que participan en adhesión, autoagregación y formación de biopelícula (70). Además, se identificó un grupo de genes (ORFs 17-33) con 91% de identidad con la isla OI#138 (Número de acceso GenBank: AE005571) de STEC O157:H7. La isla OI#138 no se ha caracterizado previamente, pero su función predicha es la biosíntesis de ácidos grasos y el ensamblaje de policétidos (137). Debido a que esta región contiene genes relacionados con adhesión y colonización, fue nombrada como Locus de Adhesión y Colonización (LAC).

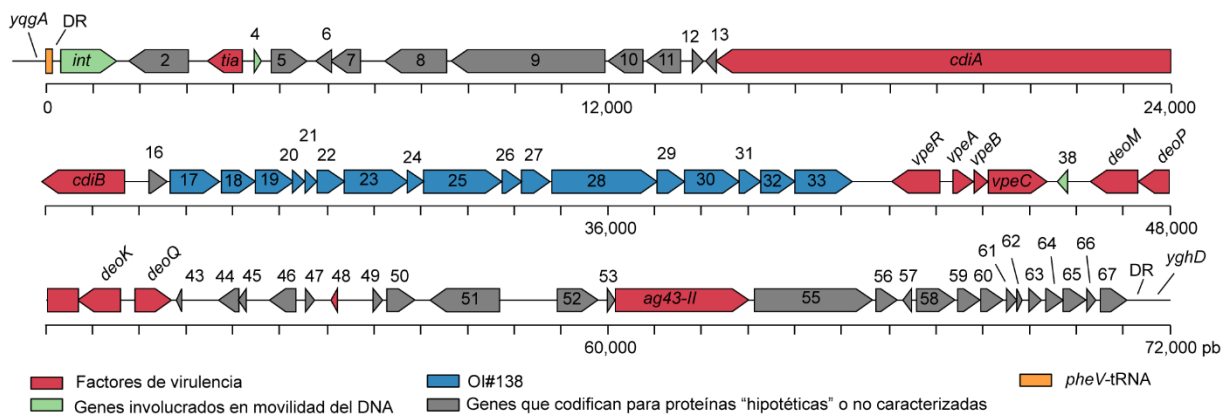


Figura 15. Estructura genética del Locus de Adhesión y Colonización. Los genes predichos y la dirección de la transcripción se representan como flechas y son coloreados de acuerdo con la función, como se indica en la leyenda. Se muestran los nombres de algunos genes. Esta región de DNA está localizada adyacente al gen *pheV-tRNA* en el contig 3 (Número de acceso GenBank: AYKD0100003.1) del genoma parcial de la cepa *E. coli* FCH1.

Por otro lado, se determinó que la PAI más frecuente entre las cepas analizadas es LAA (151/367; 41%), seguida de SE-PAI (65/367; 18%), LPA (59/367; 16%) y HPI (43/367; 12%). Las PAIs menos frecuentes son LIC (25/367; 7%) y LAC (11/367; 3%) (**Figura 16A**), mientras que en 106/367 (29%) cepas no se identificó ninguna de estas PAIs (**Tabla 5**). Debido a que LAA es modular, para este análisis particular se consideró una cepa positiva para este locus cuando al menos tres de sus módulos fueron identificados. Notablemente, todas estas PAIs se identificaron en cepas pertenecientes a diferentes serotipos, lo que demuestra su amplia distribución en STEC LEE-negativo. Por otra parte, sólo SE-PAI y LIC fueron significativamente más frecuentes en cepas clínicas en comparación con cepas aisladas de animales (**Figura 16B**).

Tabla 5. Distribución de islas de patogenicidad entre cepas STEC LEE-negativo

PAI	Serotipos (No. de cepas)	Total (%)
LAA	O2:H29 (1), O8:H19 (1), O8:H49 (1), O22:H8 (6), O38:H21 (1), O48:H21 (2), O74:H42 (2), O79:H7(2), O88:H25 (6), O91:H14 (17), O91:H21 (22), O104:H21 (6), O112ac:H19 (1), O113:H21 (27), O116:H21 (4), O116:H49 (1), O130:H38 (1), O134:H38 (1), O146:H10 (1), O146:H21 (3), O153/O178:H19 (4), O163:H19 (5), O166:H28 (2), O168:H8 (2), O171:H2 (4), O174:H8 (1), O174:H21 (14), O179:H8 (1), O181:H49 (1), O185:H7 (4), H10 (1), H21 (2), H25 (4)	151 (41)
SE-PAI	O5:H19 (1), O22:H8 (1), O27:H30 (1), O71:H12 (1), O76:H19 (6), O81:H21 (1), O84:H8 (1), O91:H14 (7), O113:H4 (4), O128ab:H2 (6), O128ac:H2 (3), O140:H21 (1), O146:H21 (15), O151:H12 (2), O174:H8 (11), O174:H21 (1), Ont:H8 (1), H8 (2)	65 (18)
LPA	O2:H29 (1), O5:H19 (1), O6:H34 (2), O15:H21 (1), O15:H27 (4), O21:H21 (1), O22:H8 (2), O38:H26 (1), O75:H8 (1), O76:H19 (1), O81:H21 (1), O91:H14 (8), O112ab:H21 (1), O113:H4 (4), O128ab:H2 (5), O128ac:H2 (5), O140:H21 (1), O146:H21 (7), O151:H12 (2), O174:H8 (6), O174:H21 (4)	59 (16)
HPI	O6:H10 (4), O8:H16 (1), O15:H27 (4), O45:H12 (1), O50/O2:H48 (1), O75:H7 (2), O75:H8 (1), O76:H19 (1), O92:H10 (1), O104:H4 (5), O109 (2), O109:H16 (1), O112ac:H19 (1), O117:H7 (5), O119:H4 (1), O128ab:H2 (5), O128ac:H2 (1), O136:H16 (3), O151:H12 (3)	43 (12)
LIC	O2:H9 (1), O21:H21 (1), O76:H19 (4), O91:H21 (1), O113:H4 (2), O113:H21 (8), O174:H21 (8)	25 (7)
LAC	O22:H8 (1), O113:H21 (5), O174:H21 (3), O168:H8 (1), H10 (1)	11 (3)
Sin identificar	O2:H25 (4), O8:H9 (2), O8:H10 (1), O8:H19 (8), O8:H28 (2), O9:H7 (1), O17/O77:H18 (5), O22:H8 (2), O28ac/O42:H25 (1), O36:H14 (2), O38:H21 (1), O41:H26 (2), O46:H38 (1), O55:H12 (2), O75:H31 (1), O88:H19 (1), O89:H9 (2), O91:H14 (4), O100:H25 (1), O100:H30 (2), O104:H7 (6), O10:H9 (1), O112ab:H2 (1), O113:H8 (1), O113:H21 (2), O116:H49 (1), O117:H4 (1), O121:H7 (3), O121:H10 (1), O130:H11 (4), O136:H12 (1), O138:H14 (2), O139:H1 (4), O141ac:H4 (1), O142:H38 (2), O146:H21 (2), O149:H12 (1), O152:H19 (1), O153:H31 (1), O154:H31 (1), O163:H19 (2), O166:H15 (1), O168:H8 (5), O169/O183:H46 (1), O171:H25 (1), O174:H25 (1), O174:H25 (1), O183:H18 (1), O187:H52 (4), H11 (1), H19 (1), Ont:H20 (3), OXY24:H16 (2)	106 (27)

Además, se encontró la existencia de patrones de co-ocurrencia y exclusión mutua entre las PAIs (**Figura 16C**). Específicamente, se encontró que la co-ocurrencia de LAA, LIC y LAC, SE-PAI y LPA, así como LPA y HPI, es estadísticamente probable ($p < 0.005$, $p < 0.0001$ and $p < 0.05$, respectivamente). Por el contrario, tanto LAA como LAC son mutuamente excluyentes con SE-PAI, LPA y HPI ($p < 0,0001$). No está claro si estas asociaciones proporcionan ventajas selectivas a cepas particulares o están

relacionadas con el fondo genético de cada cepa. Finalmente, las PAIs se asociaron con genes de toxinas específicas (**Figura 16E**). Por ejemplo, LAA se asoció con *stx1a*, *stx2a*, *stx2d*, *ehxA* y *cdtB* que codifican las toxinas asociadas a los casos más graves de infección por STEC (46–49).

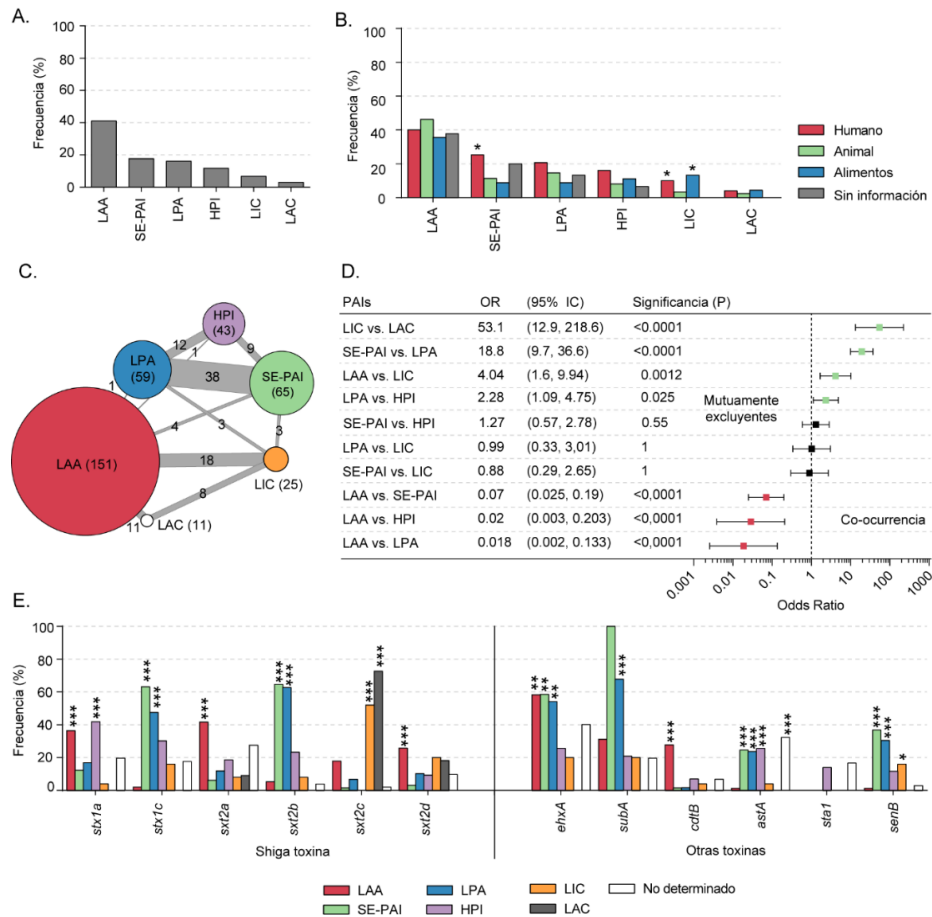


Figura 16. Distribución de seis islas de patogenicidad entre cepas STEC LEE-negativo y perfiles de toxinas asociados a cada una de ellas. **A.** Frecuencia de detección de las islas LAA, SE-PAI, LPA, HPI, LIC y LAC. **B.** Distribución de las PAIs entre cepas aisladas de diferentes fuentes. Las barras de colores representan la frecuencia de detección de cada PAIs en cepas aisladas de una fuente específica (ver leyenda). Diferencias estadísticas determinadas individualmente con la prueba exacta de Fisher (dos colas). * P <0.05. **C.** Gráfico de módulos que muestran las interacciones entre las PAIs. Los enlaces (líneas grises) se ponderan tanto por el número de cepas vinculadas entre los módulos como por la cantidad de cepas dentro de los módulos. El número de cepas positivas para cada PAI se muestra entre paréntesis. El número de cepas en las que se identificaron dos PAIs específicas (co-ocurrencia) se muestra al lado o dentro de los enlaces. La figura se preparó utilizando el paquete igraph (107) en R (106). **D.** Odds Ratios (OR) con intervalos de confianza del 95%, destacando la asociación entre PAIs específicas. Los OR entre LAA vs LAC, LPA vs LAC, SE-PAI vs LAC, HPI vs LIC y HPI vs LAC no se calcularon porque un cuadrante de la tabla de contingencia 2x2 era cero. Diferencias estadísticas determinadas con la prueba chi-cuadrado de Pearson o la prueba exacta de Fisher (dos colas), en función del tamaño de la muestra. **E.** Frecuencia de detección de genes de toxinas entre las cepas portadoras de las islas LAA, SE-PAI, LPA, HPI, LIC y LAC. También se muestran las cepas en las que no se identificó ninguna de estas PAIs (no determinado). Las diferencias significativas entre grupos se determinaron individualmente con la prueba exacta de Fisher (dos colas). * P <0.05, ** P <0.005, *** P <0.0005.

6.3.1.3 Identificación de elementos integrativos y conjugativos

El gen *sigA* fue detectado casi exclusivamente en un grupo de cepas clínicas STEC LEE-negativo (**Tabla 4**). Este gen fue originalmente identificado en la PAI *she* de *Shigella flexneri* 2a (138), y más tarde en dos PAIs albergadas por la cepa STEC/EAEC O104:H4, causante del brote de SHU en Europa durante el 2011 (139). Teniendo en cuenta estas observaciones, se analizó el contexto genético de este gen en STEC LEE-negativo. En el genoma parcial de la cepa O117:H7 FHI72 (Número de acceso de ensamblaje GenBank: GCA_000939255.1), el gen *sigA* fue identificado 20,168 y 20,442 pb río abajo del locus *pheV*-tRNA y de un gen que codifica una integrasa, respectivamente (**Figura 16A**). Además, una secuencia DR de 52 pb, que corresponde al extremo 3' terminal del locus *pheV*-tRNA, se identificó a 66,066 pb río abajo del codón de inicio de *sigA*. Estas secuencias DR flanquean una región de DNA de 86.482 pb que tiene un contenido de G + C (51%) similar al valor promedio exhibido por el cromosoma de la cepa FHI72. En esta región se identificaron 103 ORFs, entre los que se encuentran genes implicados en conjugación y biogénesis de un pilus tipo IV (**Tabla S8**), características que son exhibidas por los ICEs (140). Hasta el momento, se han identificado seis ICEs en *E. coli*, que están registrados en la base de datos ICEberg (102). En consecuencia, de acuerdo con la nomenclatura propuesta por Burrus et al. (141), se denominó esta región de DNA como ICE*Ec*7. Este elemento genético fue identificado en cepas pertenecientes a los serotipos O91:H14 (7/29; 24%) y O117:H7 (5/5; 100%). Además de *sigA* (ORF23), *iha* (ORF52) es otro gen de virulencia contenido en el ICE*Ec*7 presente en las cepas O117:H7, pero no en el de las cepas O91:H14, en el que además existe una inversión de aproximadamente 40 kb (**Figuras 16B y 16C**). La movilidad (transferencia horizontal) de este elemento genético debe investigarse en el futuro.

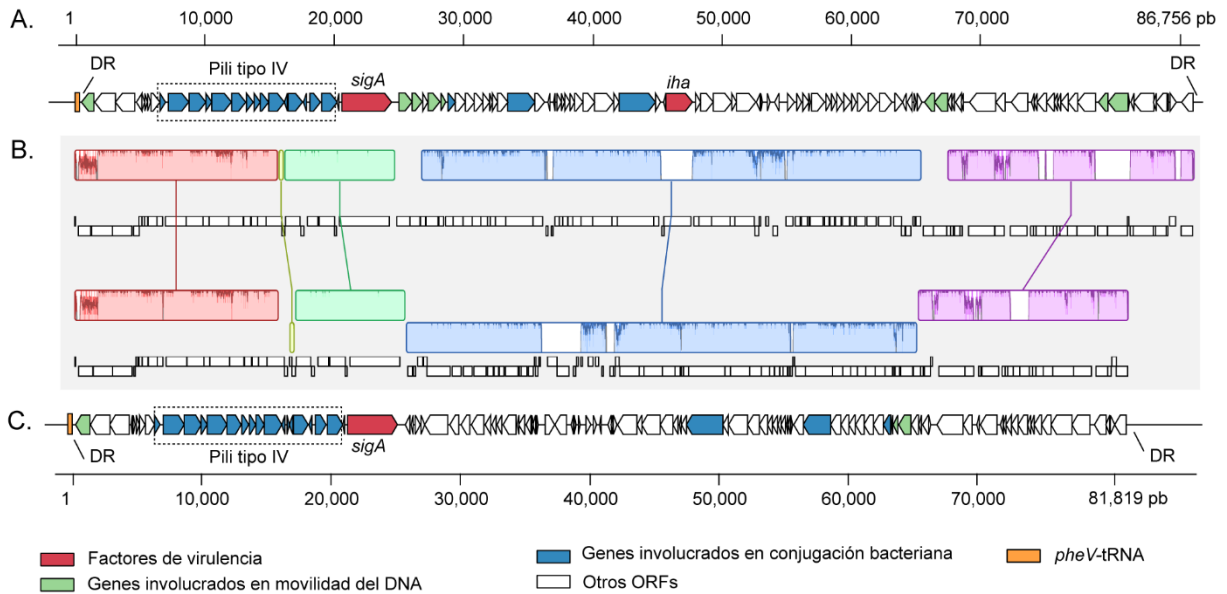


Figura 17. Estructura genética de los Elementos Integrativos y Conjugativos (ICE) identificados en cepas STEC LEE-negativo. **A.** ICE adyacente al gen de *pheV*-tRNA en el genoma parcial de la cepa STEC O117:H7 FHI72 (Número de acceso ensamblaje GenBank: GCA_000939255.1). **B.** Alineación entre los ICEs mostrados en (A) y (C). La alineación se realizó usando progressiveMauve (96). Los segmentos coloreados representan regiones homólogas. Las áreas no coloreadas representan secuencias no alineadas que pueden ser específicas del genoma. Las inversiones se identifican por cuadros debajo de la línea central. **C.** ICE adyacente al gen de *pheV*-tRNA en el contig 26 (Número de acceso GenBank: LOGV01000031.1) del genoma parcial de la cepa STEC O91:H14 2174. Los genes predichos y la dirección de transcripción se representan como flechas y están codificados por colores según la función, como lo indica la leyenda. Los nombres de algunos genes se muestran.

6.3.2 Análisis filogenético y tipificación molecular

Entre las cepas STEC LEE-positivo existen clados con diferente potencial de virulencia (142, 143). Además, se ha sugerido que las diferencias por país en la incidencia de SHU pueden estar asociadas a la circulación de clados hipervirulentos de STEC (144). Por lo tanto, se intentó determinar si hay clados con mayor potencial de virulencia entre STEC LEE-negativo y, de ser así, conocer su distribución geográfica. Para esto, se investigó la historia evolutiva de las cepas por medio de un análisis filogenético basado en SNPs. El árbol filogenético de máxima parsimonia obtenido se muestra en la **Figura 18**. Para una mejor visualización y comprensión, se mapeó en el árbol filogenético la distribución de determinantes de virulencia y los datos epidemiológicos de las cepas pertenecientes a los serotipos clínicamente relevantes (O91:H14, O91:H21, O113:H21, O128:H8, O146:H21 y O174:H21) (**Figura 19**). Notablemente, se encontró una estrecha relación entre la filogenia y los repertorios de genes de virulencia, y en algunos casos, con el origen geográfico. Por ejemplo, el

serotipo O91:H21 se separó en dos clados principales (**Figura 19A**); el clado 1 agrupa cepas de Argentina y Canadá, mientras que las cepas del clado 2 son en su mayoría de USA. Todas las cepas de este serotipo contienen LAA. Una cepa aislada en Alemania, ubicada fuera de estos dos clados, también contiene LIC.

El serotipo O113:H21 se separó en múltiples clados, siendo los clados 1 y 2 los que contienen la mayoría de las cepas (**Figura 19B**); el clado 1 agrupa cepas de Norte América mientras que el clado 2 agrupa cepas de Norte y Sur América y Europa. El clado 2 incluye dos cepas O38:H21 y una cepa O21:H21, lo que indica que estos serotipos son filogenéticamente cercanos a O113:H21. La mayoría de las cepas O113:H21 contienen LAA y un número importante de cepas del clado 2, incluyendo la cepa E045-00, contienen además LIC y LAC. La adquisición acumulativa de PAIs en STEC LEE-positivo se ha asociado con una mayor capacidad para causar enfermedad en humanos, lo que sugiere un efecto aditivo o sinérgico (145). Por lo tanto, con base en la presencia de múltiples PAIs y la distribución geográfica, las cepas O113:H21 del clado 2 parecen tener un mayor potencial de virulencia y estar diseminadas globalmente.

Se obtuvieron resultados similares para el serotipo O174:H21 (**Figura 19C**), donde tres clados fueron identificados. Los clados 1 y 2 agruparon cepas aisladas principalmente en Norte América, mientras que el clado 3 agrupó cepas con un origen geográfico más diverso. Las cepas del clado 1 contienen LAA, las del clado 2 LPA y las del clado 3 LAA, LIC y LAC. Como se observó para el serotipo O113:H21, las cepas O174:H21 que contienen más de una PAI parecen estar diseminadas en áreas geográficas más grandes.

En 2013 y 2014, O91:H14 fue el serotipo de STEC más prevalente en un estudio prospectivo multicéntrico realizado en Holanda (23). La mayoría de las cepas O91:H14 analizadas en esta Tesis se aislaron en ese estudio. Este serotipo se separó en tres clados (**Figura 19D**); los clados 1 y 2 agruparon cepas de Norte América y Europa, mientras que el clado 3 sólo de Holanda. Las cepas del clado 1 contienen SE-PAI y LPA; las del clado 2 LAA; mientras que las del clado 3 LAA y el ICE*Ec7*. Estos resultados indican que hay varias subpoblaciones de este serotipo circulando en

Europa, particularmente en Holanda. Es importante mencionar que los genes implicados en la producción de microcinas (clúster *mch*) son prevalentes en este serotipo, tal vez contribuyendo a su emergencia en Europa.

El serotipo O146:H21 se separó en tres clados (**Figura 19E**). Los clados 1 y 2 agruparon cepas de Norte América y el clado 3 de Norte América y Europa. Las cepas del clado 1 contienen SE-PAI; las del clado 2 LAA y SE-PAI; mientras que las del clado 3 SE-PAI y LPA. Además, el clúster *mch* fue prevalente en los clados 1 y 3. Como en los serotipos anteriores, las cepas que contienen más de una PAI (clado 3) tienen un origen geográfico más amplio.

El serotipo O128:H8 se agrupó en dos clados (**Figura 19F**), en los que se encuentran cepas tanto de Norte América como de Europa. Mientras que las cepas del clado 1 contienen SE-PAI y LPA, las cepas del clado 2 contienen además HPI. El clúster *mch* es también prevalente en este serotipo. En conjunto, estos resultados sugieren una relación entre la adquisición acumulativa de PAIs y una distribución geográfica más amplia.

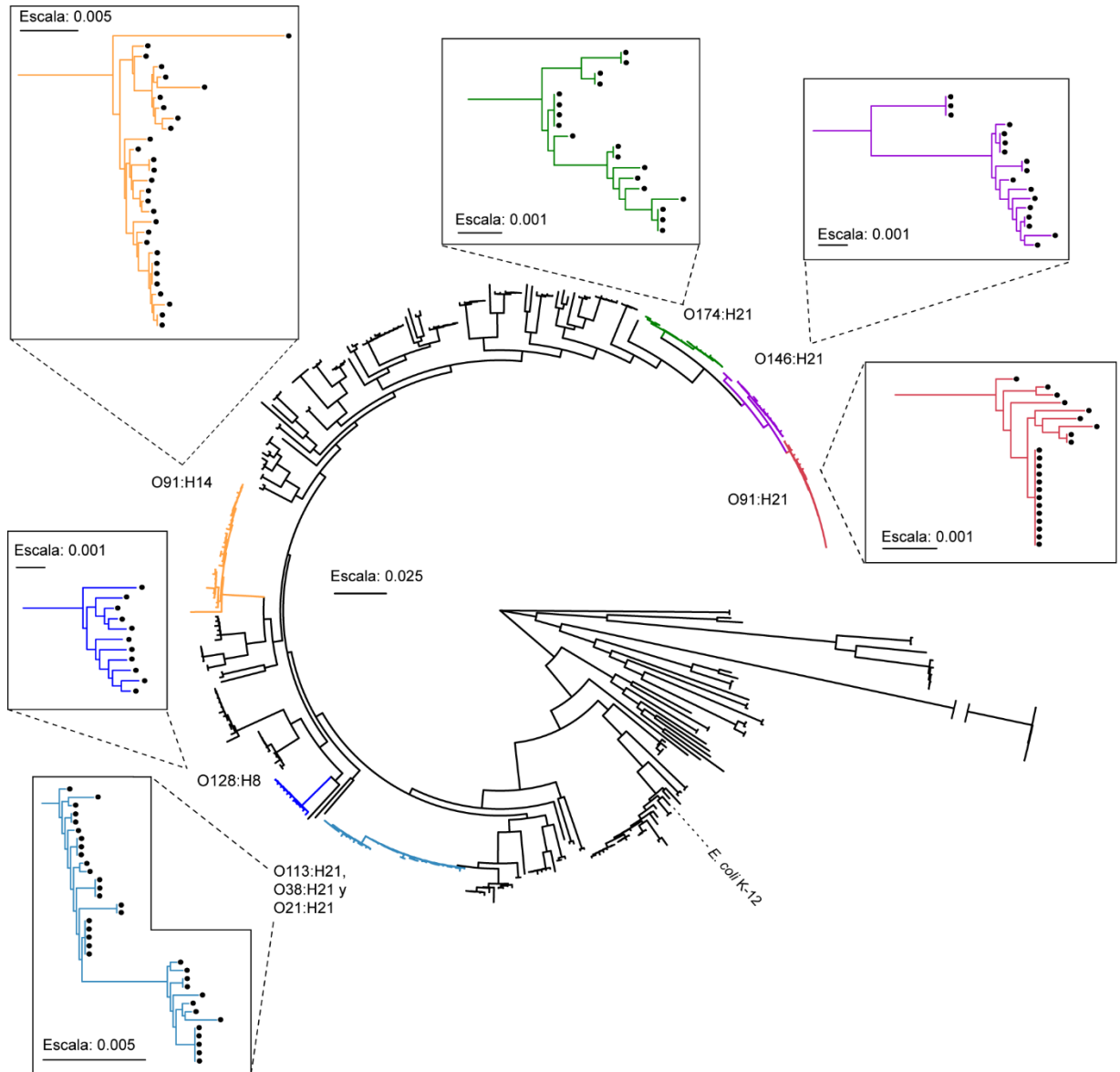


Figura 18. Relaciones filogenéticas entre 367 cepas STEC LEE-negativo. El árbol filogenético de máxima parsimonia se basa en 3.956 SNPs del genoma “core” identificados en las cepas STEC LEE-negativo y *E. coli* K-12 MG1655. El árbol se construyó con kSNP3.1 (109) y se editó con iTOL (110). Las ramas para cepas más distantes se acortaron con fines gráficos. Se muestran clústeres que contienen cepas de serotipos clínicamente relevantes.

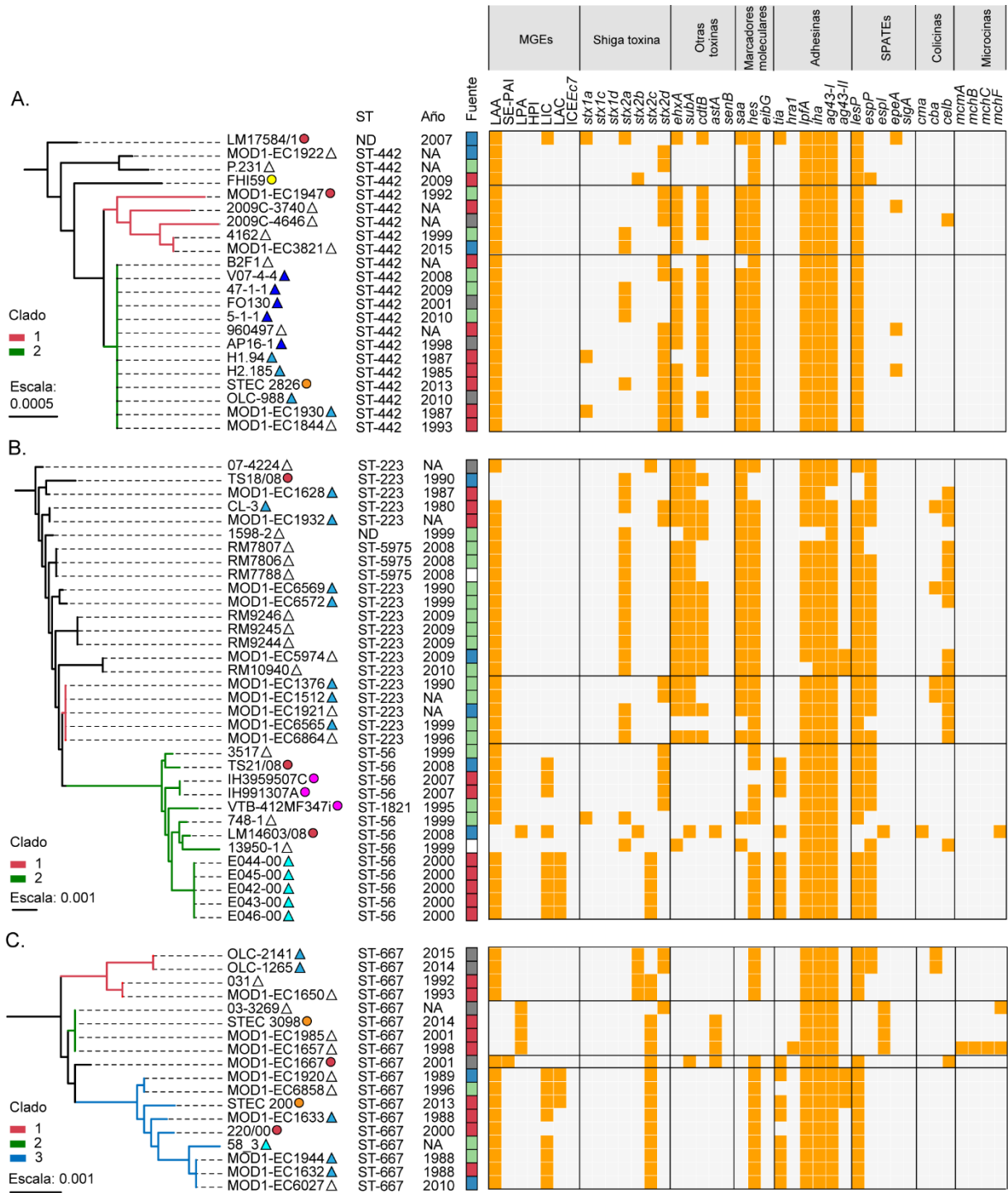


Figura 19. Relaciones filogenéticas entre cepas STEC LEE-negativo pertenecientes a serotipos clínicamente relevantes. Clústeres correspondientes a los serotipos O91:H21 (A), O113:H21 (B), O174:H21 (C), O91:H14 (D), O146:H21 (E) y O128:H8 (F) fueron extraídos del árbol de máxima parsimonia mostrado en la Figura 17. La distribución de los determinantes de virulencia se muestra en el panel derecho. Se indica la presencia (naranja) o ausencia (gris) de cada MGE o gen. Se muestran los datos epidemiológicos de cada cepa, incluidos el país, año y fuente de aislamiento (ver leyenda). Los clados identificados se indican por color de acuerdo con la leyenda a la izquierda.

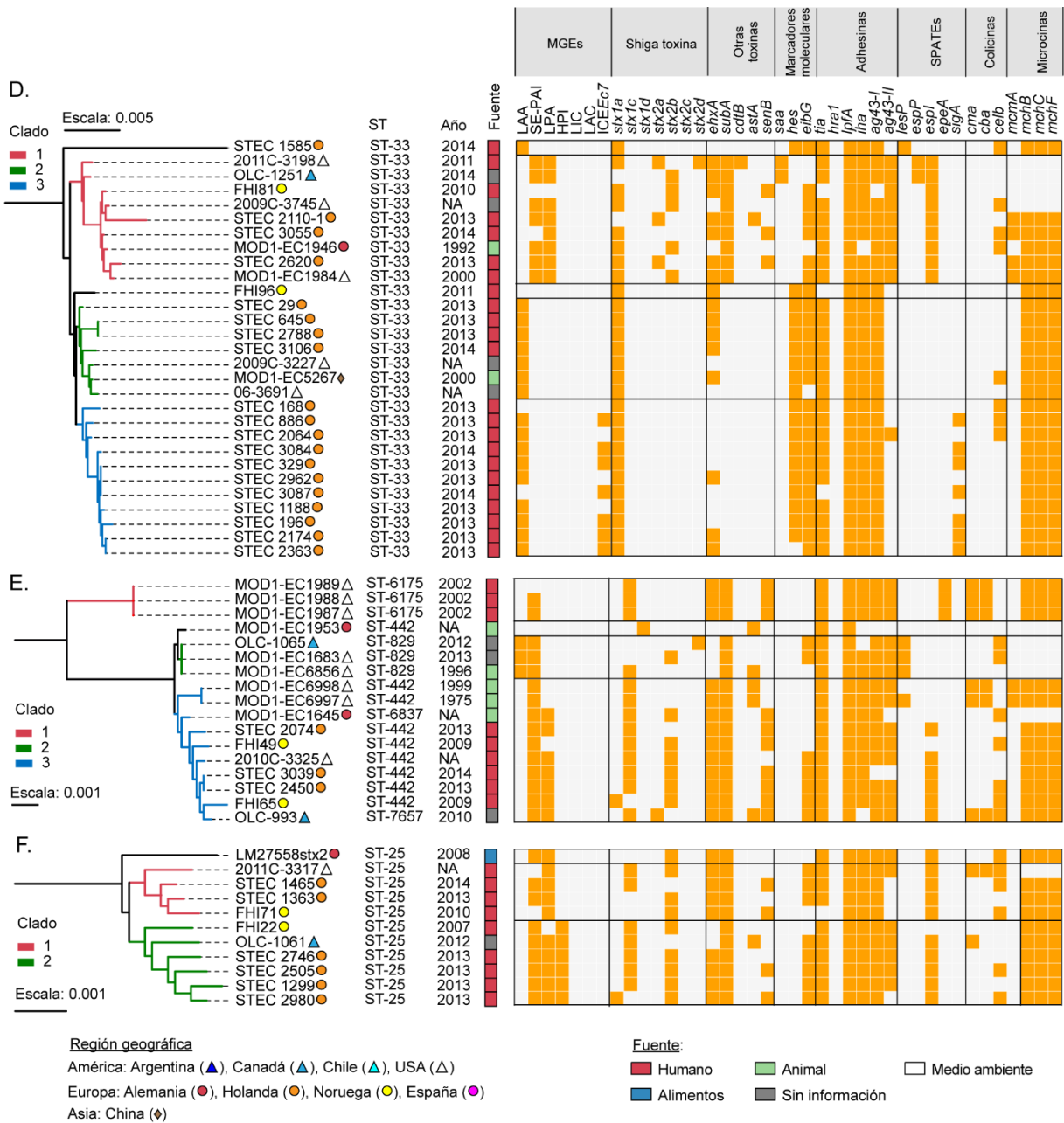


Figura 19. (Continuación)

6.4 Determinar la contribución de LAA en la adherencia y colonización de STEC, mediante ensayos *in vitro* de adhesión y de un modelo murino de colonización, respectivamente.

6.4.1 Análisis funcional de LAA en ensayos *in vitro*

La alta prevalencia de LAA entre cepas STEC LEE-negativo sugiere que juega un papel biológico importante. En consecuencia, se investigó su contribución a los mecanismos patogénicos de STEC. Para esto, LAA fue removida desde el ORF1 hasta el ORF 80 (aprox. 86 kb) en la cepa STEC O113:H21 E045-00 mediante reemplazo alélico. Posteriormente, la cepa E045-00 y su mutante isogénica Δ LAA fueron cultivadas en LB con o sin agitación a 37°C y sus perfiles de OMPs se analizaron mediante 2D-PAGE (**Figura 20**). Las OMPs sintetizadas diferencialmente se identificaron mediante espectrometría de masa. El análisis de los perfiles de OMPs de la cepa E45-00 indica que la proteína Ag43 codificada en LAA se sintetizó durante el cultivo con agitación, pero no en el cultivo estático. Por el contrario, la proteína flagelar (FliC-H21) fue identificada en el perfil del cultivo estático, pero no fue detectable en el perfil de cultivo con agitación. Este efecto regulador opuesto en la expresión de *fliC* y *ag43* también fue observado en un estudio previo (146). Notablemente, en ausencia de LAA, FliC-H21 fue detectada en cantidades similares en extractos obtenidos de cultivos agitados y estáticos, lo que sugiere la existencia de una red de regulación entre LAA y los genes flagelares que merece un análisis futuro. Como se esperaba, en los perfiles de OMPs de la mutante Δ LAA no se identificaron los spots correspondientes a las proteínas NmpC y Ag43 (ambas codificadas en LAA). Un solo spot correspondió a las proteínas Tia y Hes (codificada en LAA). Ambas proteínas pertenecen a la familia Hra y tienen 65% de identidad en su secuencia de aminoácidos (**Figura 3**), con puntos isoeléctricos y pesos moleculares casi idénticos. La cepa E045-00 contiene dos copias del gen *tia*, localizadas en LIC y LAC (**Tabla S1**). En la cepa E045-000 y su mutante Δ LAA, Tia fue más abundante en los perfiles de OMPs obtenidos de los cultivos sin agitación. Queda por determinar si ambas copias del gen están reguladas positivamente en esta condición de crecimiento. Otra proteína que aumentó la síntesis en cultivo agitado fue OmpW, la que se ha asociado con la

protección contra la fagocitosis y mecanismos de expulsión de sustancias tóxicas (147, 148). Teniendo en cuenta estos resultados, los experimentos funcionales posteriores se realizaron cultivando las bacterias en LB con agitación y a 37 °C.

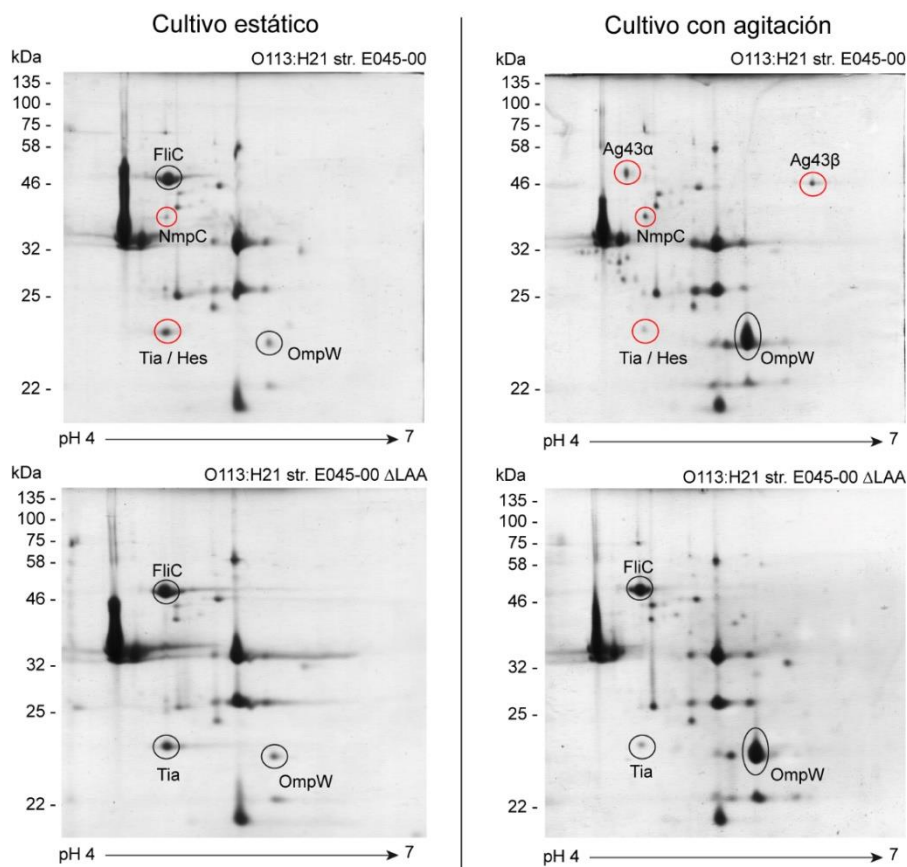


Figura 20. Perfiles de proteínas de membrana externa de la cepa E045-00 y su mutante isogénica Δ LAA. 2D-PAGE de extractos de OMPs de las cepas E045-00 y su mutante Δ LAA cultivadas en LB a 37 °C con agitación o en estático. Los gels de poliacrilamida al 12% (13 cm, intervalo de pH: 4 - 7) fueron teñidos con Coomassie G-250 y las proteínas seleccionadas fueron identificadas mediante espectrometría de masas MALDI-TOF/TOF. Se indican en círculos rojos las proteínas que están codificadas en LAA. Las barras de escala a la izquierda indican pesos moleculares en kDa.

A continuación, se determinó si la delección de LAA afecta la adhesión *in vitro* de la cepa E045-00 a células HT-29. Como se muestra en la **Figura 21**, no se observaron diferencias en los niveles de adhesión a 30 min entre la cepa E045-00 y su mutante Δ LAA. Este resultado no descarta la participación de las proteínas codificadas en LAA en la adhesión *in vitro*, pero sí indica que esta cepa contiene genes adicionales implicados en este fenotipo.

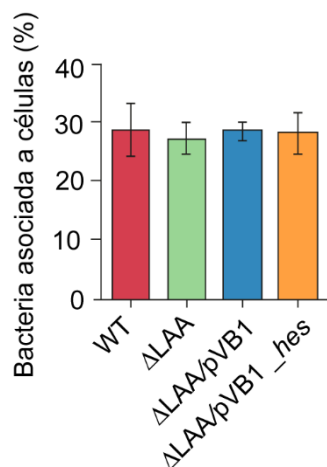


Figura 21. Adhesión de la cepa O113:H21 E045-00 y su mutante isogénica Δ LAA a células HT-29. Datos son expresados como el porcentaje de inóculo inicial recuperado después de 30 min de infección a una multiplicidad de infección de 10 bacterias por célula. Barras de error representan s.d. (n=3).

6.4.2 Determinación de la contribución de LAA a la colonización intestinal y virulencia de STEC en un modelo murino de infección

El papel de LAA en la colonización intestinal fue evaluado en ratones tratados con estreptomicina. Para esto, grupos de cinco ratones BALB/c fueron inoculados oralmente con $\sim 10^9$ cfu de la cepa E045-00 o su mutante Δ LAA. La colonización intestinal se evaluó durante 12 días consecutivos mediante el recuento de bacterias viables en heces fecales. Los niveles de colonización de la cepa E045-00 fueron mayores en comparación con la mutante Δ LAA desde el día 2 hasta el día 10, con una diferencia significativa en los días 2, 3, 6, 8 y 9 (**Figura 22A**). En los días 11 y 12, los niveles de colonización de ambas cepas fueron similares. Es importante mencionar que la cepa E045-00 contiene la PAI LAC, donde se localizan los operones *vpe* y *deoK*, que promueven la colonización del intestino murino (135, 136). Por lo tanto, no se puede descartar la posibilidad que los genes contenidos en LAC u otras PAIs puedan compensar la falta de LAA. Adicionalmente, en el día 13 post-infección se tomaron muestras de sangre de los ratones infectados con la cepa E045-00 y se obtuvo una mezcla de sueros. Un ensayo de Western blot usando esta mezcla de sueros indicó que dos proteínas presentes en el extracto de OMPs de la cepa E045-00 pero no en el de la mutante Δ LAA, son reactivas a IgG (**Figura 22B**; asteriscos). Teniendo en

cuenta el peso molecular de las proteínas inmunodetectadas, éstas proteínas podrían corresponder a α Ag43 (48 kDa) y NmpC (38 kDa).

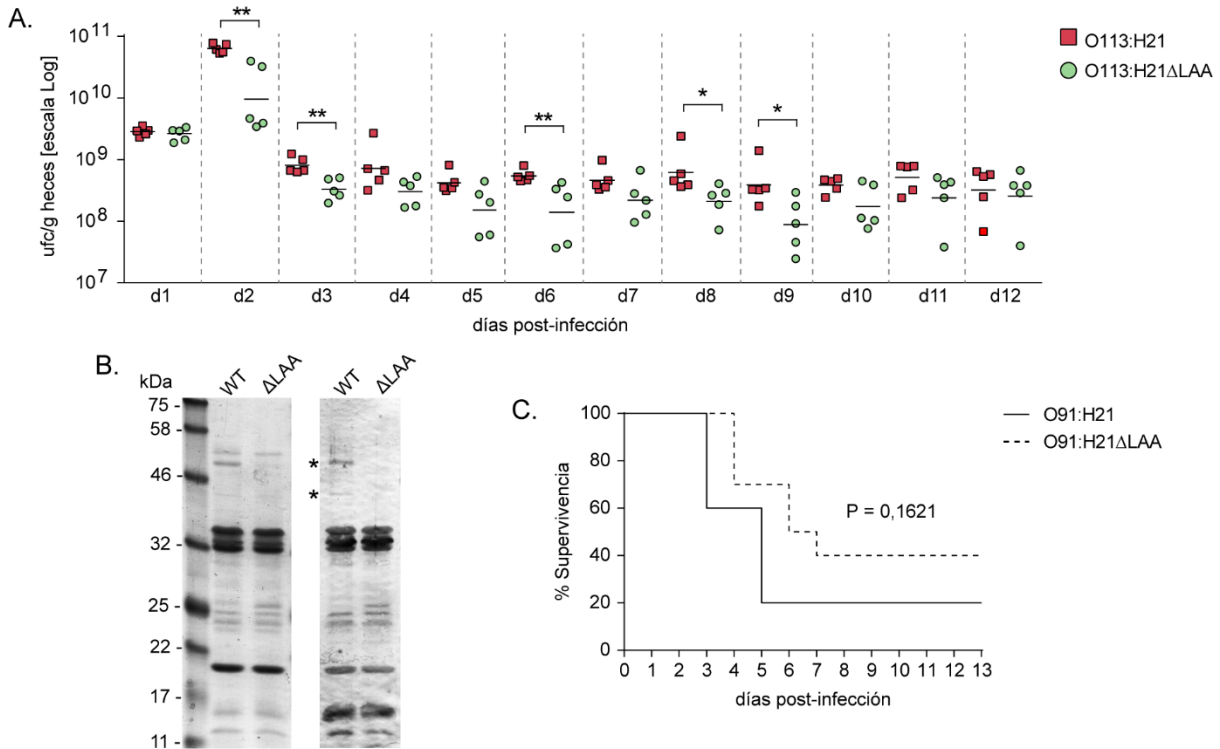


Figura 22. Contribución de LAA a la colonización intestinal y virulencia de STEC. A. Ensayos de colonización intestinal. Ratones tratados con estreptomicina fueron inoculados oralmente con $\sim 10^9$ ufc de la cepa E045-00^{SR} o su mutante Δ LAA^{SR}. Los puntos representan el promedio de ufc/g de muestra fecal de cada ratón en el día indicado. Las líneas representan las medias. Las diferencias en los niveles de colonización para cada día se analizaron con la prueba U de Mann-Whitney (* P < 0.05 ** P < 0.005). **B.** SDS-PAGE de OMPs de la cepa E045-00 (izquierda) y Western blot (derecha) usando la mezcla de sueros (obtenidos al día 13) de los ratones inoculados con la cepa E045-00. Dilución de mezcla de sueros 1:250. Anticuerpo secundario Anti-IgG ratón conjugado con peroxidasa (dilución 1:1000). Los asteriscos indican proteínas reactivas a IgG presentes en el perfil de OMPs de la cepa silvestre (WT) y ausentes en la mutante Δ LAA. **C.** Tasa de supervivencia de ratones tratados con estreptomicina que fueron inoculados oralmente con 10^9 ufc de la cepa V07-4-4^{SR} o su mutante Δ LAA^{SR}. La diferencia en la tasa de supervivencia entre ambos grupos (n=10) se analizó utilizando la prueba de log-rank. La significancia estadística fue definida como P < 0.05.

Por otro lado, se evaluó la participación de LAA en la virulencia de STEC utilizando la cepa O91:H21 V07-4-4, que produce Stx2d (el subtipo de Stx con la mayor toxicidad en el modelo murino (149)) y que sólo contiene LAA de las seis PAIs analizadas en este estudio. Para esto, LAA fue removida desde el ORF1 hasta el ORF 80 (aprox. 86 kb) en la cepa V07-4-4 mediante reemplazo alélico y se determinó si esta delección afecta la virulencia de la bacteria en el modelo de ratón tratado con estreptomicina. Se

inocularon grupos de diez ratones BALB/c con $\sim 10^9$ cfu de la cepa V07-4-4 o su mutante Δ LAA. Los ratones de ambos grupos mostraron signos de enfermedad (letargo y consumo reducido de alimentos y agua) desde el segundo al octavo día del experimento. En los ratones que murieron, se observó parálisis de las extremidades posteriores, ataxia y dificultades para respirar. Como se muestra en la **Figura 22C**, dos de diez (20%) y cuatro de diez (40%) ratones desafiados con la cepa V07-4-4 o la mutante Δ LAA sobrevivieron hasta el final del experimento (día 13), respectivamente; sin embargo, no se observó una diferencia significativa ($P = 0.1621$) en las curvas de supervivencia entre los grupos experimentales. Estudios adicionales utilizando un inóculo más bajo, determinando el índice de competencia entre la cepa silvestre y la mutante Δ LAA o determinando la DL_{50} podrían complementar estos resultados y proporcionar más información sobre el papel de esta PAI en la patogenicidad de STEC.

Tomados en conjunto, estos resultados proporcionan las bases para futuros estudios que aborden el papel de LAA en adhesión y virulencia, así como el potencial papel redundante de otras PAIs que puedan compensar la falta de LAA, usando modelos *in vitro* e *in vivo* alternativos. Además, la regulación observada entre LAA y *fliC*, sugiere que factores transcripcionales codificados en LAA reprimen la expresión de genes de virulencia que pueden estar activados en la cepa Δ LAA, compensando así la pérdida de LAA. Aunado a estas posibilidades, debemos también tener en cuenta que el modelo murino es un sustituto ante la poca o nula posibilidad de realizar estudios en hospederos naturales de STEC, en los que las funciones codificadas en LAA podrían ser más evidentes

7. DISCUSIÓN

A pesar de los persistentes esfuerzos por aumentar la inocuidad de los alimentos, la incidencia de enfermedades causadas por patógenos transmitidos por alimentos continúa siendo una preocupación tanto en países desarrollados como en vías de desarrollo. La evolución de estos patógenos y el impacto de este proceso biológico en su epidemiología constituye un desafío creciente para las autoridades de salud pública en el mundo entero (150). La emergencia de la cepa epidémica EAEC/STEC O104:H4 causante del brote del 2011 en Europa, mostró que la plasticidad genética y transferencia horizontal de genes favorecen la adaptación de *E. coli* a una gran variedad de ambientes (151). Actualmente, el número de casos de gastroenteritis asociados a cepas STEC LEE-negativo va en aumento en Europa, Argentina y Corea del Sur (23, 24, 33, 25–32), lo que sustenta la necesidad de estudiar los mecanismos moleculares que favorecen la emergencia de estos patógenos.

En esta Tesis se reporta y caracteriza la PAI denominada Locus de Adhesión y Autoagregación (**Figura 6**). Esta PAI contiene genes de virulencia previamente caracterizados como *sisA* (119), *iha* (68), *nmpC* (120) y *ag43* (118); potenciales genes de virulencia como *lesP* (ORF29) y *tpsA* (ORF43); y un nuevo miembro de la familia Hra denominado Hes. La caracterización funcional de Hes usando la cepa HB101 indicó que es una proteína multifuncional que participa en hemaglutinación, adhesión, autoagregación y formación de biopelícula (**Figura 4**); sin embargo, la delección del gen *hes* no afectó la capacidad de adhesión *in vitro* de la cepa STEC O113:H21 E045-00, indicando que este gen podría desempeñar un papel redundante en esta cepa particular. Es importante mencionar que las proteínas Hes, NmpC y Ag43 son reactivas a sueros de pacientes con SHU (78), indicando que son sintetizadas durante la infección en el humano.

La determinación de la distribución de LAA mediante PCR y BLASTn indicó que está exclusivamente presente en cepas STEC LEE-negativo. Esta PAI consta de cuatro módulos y, por lo tanto, para esta Tesis, una cepa STEC fue considerada LAA-positivo cuando al menos 3 de sus módulos fueron identificados. Es importante mencionar que un resultado positivo en la PCR múltiple que se desarrolló solamente

sugiere y no indica necesariamente la presencia de un módulo. Lo anterior debido a que a través de esta técnica se detecta sólo un fragmento específico y no el módulo completo. En consecuencia, solamente mediante la secuenciación del genoma de la bacteria es posible determinar con certeza la presencia de LAA. Interesantemente, la estructura completa de LAA (cuatro módulos) fue identificada en cepas aisladas de casos de CH y SHU (**Figura 11**). En particular, el módulo III, descrito previamente por Shen et al. (125), fue propuesto como un marcador molecular de cepas STEC LEE-negativo asociadas con enfermedad (85). No obstante, nuestros resultados indican que la estructura completa de LAA es posiblemente un mejor marcador molecular de riesgo. Cabe señalar que la estructura completa de esta PAI fue identificada casi exclusivamente en cepas STEC que pertenecen al filogrupo B1 (**Figura 11**), lo que sugiere que un contexto genético específico es requerido para su adquisición y/o mantención. Adicionalmente, se identificaron dos linajes principales de LAA que pueden estar presentes en diferentes cepas de un mismo serotipo (**Figura 12**), indicando que este locus ha sido adquirido más de una vez a través de eventos independientes. Estudios adicionales dirigidos a comprender la variabilidad genética entre linajes de LAA podrían proporcionar más información sobre la evolución de esta PAI y su posible papel en la adaptación de STEC a nuevos nichos, como el intestino humano o bovino.

A la fecha, no ha sido posible identificar un marcador molecular o un grupo de éstos, incluyendo la determinación del serotipo, que permita predecir con certeza el potencial de una cepa STEC para causar enfermedad en el humano (152). La “European Food Safety Authority Panel on Biological Hazards” ha sugerido que la presencia de genes como *stx2* y *eae* o *aaiC* más *aggR*, junto con otros marcadores moleculares aún no caracterizados, puede ser asociada con un mayor riesgo de enfermedad severa (152). De hecho, en un reporte del año 2018, la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) incluyó a LAA entre los factores de adherencia relevantes que pueden ser utilizados en la vigilancia epidemiológica de STEC (46).

En esta Tesis se aprovechó el aumento de los datos proporcionados por la secuenciación masiva de genomas y se realizó el análisis genómico comparativo más grande reportado hasta la fecha en cepas STEC LEE-negativo. Mediante este análisis no solamente se examinó la distribución de LAA, sino además de otros determinantes de virulencia, e incluso se identificaron nuevas GIs. Los genes *stx1a*, *stx2a* y *stx2d* codifican los subtipos de Stx asociados con los casos de diarrea más severos (46); sin embargo, ninguno de ellos fue significativamente más frecuente en cepas clínicas en comparación con cepas aisladas de animales (**Tabla 4**). Esto confirma que la identificación del subtipo de Stx, en ausencia de otros marcadores moleculares, no es suficiente para predecir el potencial de virulencia de STEC. Otros genes, como *eibG*, *tia*, *ag43-1* y *sigA* se identificaron con una frecuencia significativamente mayor en cepas clínicas en comparación con las cepas de otras fuentes. El clúster *mch* también se asoció con cepas clínicas, aunque la participación de las microcinas como factores de virulencia es un punto que merece ser investigado. Notablemente, el gen *hes* fue detectado con mayor frecuencia en comparación con el gen *saa*, que hasta el momento ha sido el marcador molecular de STEC LEE-negativo más utilizado (65). Por lo tanto, la identificación de estos genes puede complementar esquemas existentes de evaluación molecular de riesgo (153).

Entre los nuevos MGEs identificados están las islas LAC, LIC y ICE*Ec7*, que están ampliamente distribuidas en STEC LEE-negativo (**Tabla 5**). Cabe señalar que LIC y LAC contienen el gen *tia* (**Figuras 14 y 15**, respectivamente), un factor de virulencia de ETEC (69), que se había asociado a la presencia de SE-PAI en STEC (59). Por lo tanto, el reporte en esta Tesis, de dos nuevas PAIs que también contienen *tia* es indicativo de que este gen se ha propagado en STEC, lo que representa otro ejemplo de transferencia horizontal de genes y plasticidad genómica entre patotipos de *E. coli* (154). Otros genes ubicados en LAC y LIC, como los que codifican sistemas CDI, podrían proporcionar ventajas competitivas para STEC frente a otras bacterias. Del mismo modo, la identificación de ICE*Ec7* que contiene *sigA* es un hallazgo importante (**Figura 16**), ya que este gen participa en la colonización intestinal de la cepa epidémica STEC/EAEC O104:H4 (139). Un estudio reciente reportó la identificación del gen *sigA* en 36 cepas STEC LEE-negativo O117:H7 aisladas en el Reino Unido

(155), lo que sugiere que ICE $Ec7$ podría estar presente en ellas. En consecuencia, es importante que estudios futuros aborden la caracterización funcional de estos MGEs.

Este es el primer estudio que muestra la acumulación y amplia distribución de GIs en STEC LEE-negativo, lo que enfatiza el papel de estos elementos genéticos en la evolución de estos patógenos. Incluidos los resultados de esta Tesis, existen seis PAIs y un ICEs distribuidos entre STEC LEE-negativo y es probable que haya un número importante aún por identificar. En la **Figura 23** se muestra como estas cepas pueden tener más de una GIs en su genoma. Sin embargo, algunos de estos MGEs pueden ser co-ocurrentes o mutuamente excluyentes. Por ejemplo, LIC y LAA son identificadas frecuentemente en cepas que además contienen LAA. Por el contrario, LAA es mutuamente excluyente con SE-PAI, LPA y HPI. En consecuencia, existen subpoblaciones entre las cepas STEC LEE-negativo que contienen diferentes repertorios de genes de virulencia.

Vale la pena resaltar que LAA fue la PAI más prevalente entre las cepas analizadas, incluyendo cepas aisladas de casos clínicos y de animales (**Figura 16**). Lo anterior sustenta la idea que esta PAI podría representar una ventaja adaptativa tanto en la infección del humano como en la persistencia en reservorios animales, cumpliendo un doble rol biológico que también cumple la isla LEE (156). De hecho, en ensayos preliminares, realizados en conjunto con el grupo de laboratorio de la Dra. Nora Padola (Argentina), se observó una reducción en la capacidad de adhesión a explantes de colon bovino después de eliminar LAA en una cepa STEC O91:H21 (127). Los ensayos de infección realizados en ratones tratados con estreptomycin sugieren la participación de LAA en la colonización intestinal (**Figura 21A**). Aunque los resultados obtenidos en este modelo de infección fueron insuficientes para establecer la participación directa de LAA en virulencia (**Figura 21C**), si se pudo observar una ligera atenuación de la cepa V07-4-4 Δ LAA. En cualquier caso, la participación de LAA en virulencia merece ser investigada usando dosis infectivas más bajas en el modelo murino, determinando la dosis letal 50, realizando ensayos de competencia entre la cepa silvestre y la mutante o probando modelos *in vivo* alternativos. Otro resultado a destacar es la asociación de LAA con genes de toxinas que están asociados con

enfermedad grave, como *stx1a*, *stx2a*, *stx2d*, *ehxA* y *cdtB* (Figura 16E), lo que debe considerarse en la vigilancia de STEC.

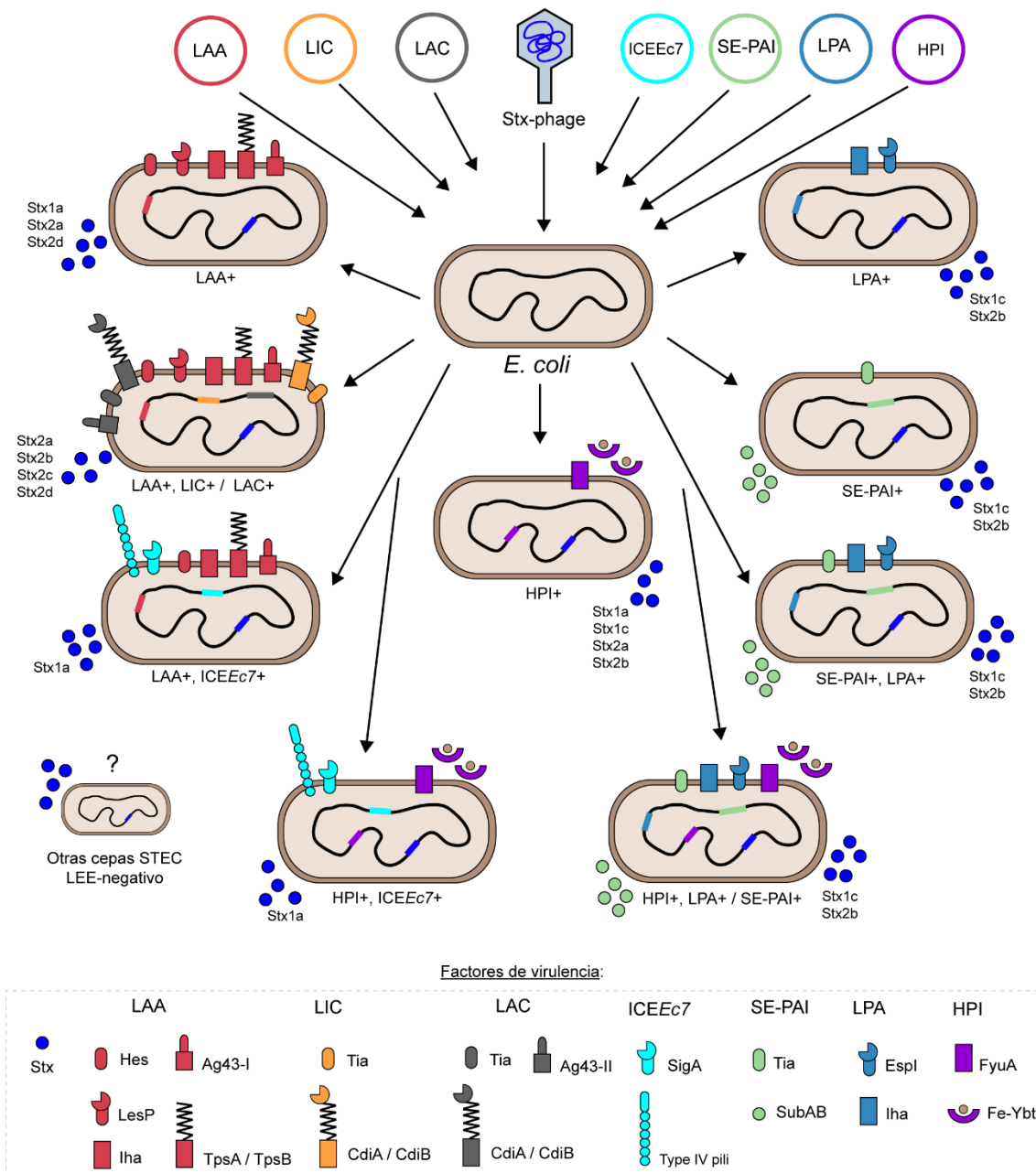


Figura 23. Modelo de la evolución de la virulencia en STEC LEE-negativo, mediada por la adquisición de fagos e islas genómicas. La adquisición de MGEs proporciona nuevos mecanismos moleculares a la cepa STEC LEE-negativo receptora, lo que aumentan su potencial de virulencia y puede contribuir a su emergencia. En la figura se indican los MGEs caracterizados hasta el momento en este subgrupo de STEC. Además, se muestra como estas cepas pueden tener más de una PAI en su genoma. Los factores de virulencia codificados en cada MGE se indican de acuerdo con la leyenda al final de la figura. También se muestran los perfiles de Stx más comunes. LAA, Locus de Adhesión y Autoagregación; LIC, Locus de invasión e inhibición del crecimiento dependiente de contacto; LAC, Locus de Adhesión y Colonización; ICEE7, Elemento Integrativo y Conjugativo No. 7 de *E. coli*; SE-PAI, "Subtilase-Encoding Pathogenicity Island"; LPA, "Locus of Proteolysis Activity"; HPI, "High-Pathogenicity Island".

Quizás aún más importante es la evidencia epidemiológica y genómica que respalda la idea que hay subpoblaciones de cepas STEC LEE-negativo que se distribuyen a nivel mundial. Específicamente, los análisis filogenéticos revelaron que los serotipos O91:H14, O91:H21, O113:H21, O128:H8, O146:H21 y O174:H21 tienen clados con mayor potencial de virulencia, principalmente como consecuencia del subtipo Stx y del número de PAIs que contienen (**Figura 19**). Lo anterior puede ser observado en el hecho que clados que agruparon cepas con más de una PAI tienen un origen geográfico más diverso. Esto sugiere que la adquisición acumulativa de PAIs, y la potencial interacción reguladora o funcional entre ellas, confiere ventajas selectivas a las bacterias, favoreciendo la diseminación de estas cepas a nivel regional, o incluso en algunos casos a nivel mundial. Posteriormente, la adaptación a un nuevo hospedero o reservorio conduce el proceso evolutivo, y dadas las condiciones adecuadas (selección natural), una cepa adaptada emergerá en una nueva región geográfica (radiación adaptativa). Lo anterior está de acuerdo con Reid *et al.* (157), quienes establecieron que la emergencia de clones *E. coli* hipervirulentos es un estado derivado de la adquisición de genes en lugar de una condición ancestral de la especie.

Finalmente, teniendo en cuenta los subtipos de Stx, el tipo y número de PAIs y los datos epidemiológicos de las cepas aisladas de humanos que se analizaron, se propone un sistema de clasificación de tres niveles (alto, medio y bajo) para el potencial de virulencia de STEC LEE-negativo (**Tabla 6**). En esta clasificación una cepa STEC LEE-negativo tiene un potencial de virulencia bajo si produce las toxinas Stx1c y/o Stx2b y contiene solamente una PAI como SE-PAI, LPA y HPI; un potencial de virulencia medio si produce las toxinas Stx1a, Stx1c y/o Stx2b, contiene LAA o más de una PAI; y un potencial alto si produce Stx2a, Stx2c, Stx2d, contiene el gen *aggR*, LAA o más de una PAI. Un grupo de cepas (n = 27) no pudo ser clasificado bajo este esquema debido a que en sus genomas no se identificó ninguna de las PAIs analizadas en esta Tesis. El enfoque que se propone puede ayudar a definir las cepas STEC LEE-negativo que representan un mayor riesgo para la salud y, por lo tanto, contribuye a mejorar la vigilancia de estos patógenos. Los estudios prospectivos multicéntricos podrán confirmar y refinar esta clasificación.

En resumen, los resultados obtenidos en esta Tesis contribuyen al conocimiento sobre la patogenicidad de STEC LEE-negativo y establecen una línea base para su vigilancia epidemiológica.

Tabla 6. Potencial de virulencia propuesto para cepas clínicas STEC LEE-negativo ¹

Serotipos ² (n)	PAIs y/o ICEs	Perfiles de <i>stx</i> ³	Otras toxinas ³	Otros genes de virulencia ³	Cuadro clínico (n)	Potencial de virulencia
O104:H4 (4)	Eco587.50F + HPI	<i>stx2a</i>	-	<i>aggR, pic, sigA, iha, ag43-I, lpfA</i>	SHU (1) Diarrea (2) Sin diagnóstico (1)	Alto
O22:H8 (1), O48:H21 (1), O79:H7 (1), O88:H25 (3), O91:H21 (9), O104:H21 (3), O112ac:H19 (1), O113:H21 (2), O116:H21 (2), O146:H10 (1), O166:H28 (2), O168:H8 (1), O174:H2 (2), O174:H21 (3), O181:H49 (1), O185:H7 (1), H25 (1)	LAA	<i>stx2a, stx2d, stx1a + stx2a, stx1a + stx2d</i>	<i>ehxA, subAB, cdtB</i>	<i>hes, iha, lesP, ag43-I, saa, lpfA, espP</i>	SHU (6) CH (5) Diarrea (5) Dolor abdominal (1) Asintomático (1) Contenido intestinal de bebé con SIDS (1) Sin diagnóstico (15)	Alto
O2:H29 (1), O113:H21 (7), O174:H21 (3)	LAA + LIC + LAC ⁴	<i>stx2c, stx2d, stx2a + stx2b</i>	-	<i>hes, tia, iha, lesP, ag43-I, ag43-II, lpfA, espP</i>	CH (5) Diarrea (4) Sin diagnóstico (2)	Alto
O91:H14 (8) O146:H10 (1)	LAA (Sin módulo II)	<i>stx1a</i>	-	<i>hes, iha, ag43-I, lpfA, eibG, tia, mch</i> clúster genes	CH (1) Diarrea (3) Dolor abdominal (3) Sin diagnóstico (2)	Medio
O91:H14 (7)	LAA (sin módulo II) + ICE <i>Ec7</i>	<i>stx1a</i>	-	<i>sigA, hes, iha, ag43-I, eibG, tia, mch</i> clúster genes	Diarrea (4) Dolor abdominal (2) Sin diagnóstico (1)	Medio
O15:H27 (1), O76:H19 (1), O128:H2 (5), O151:H12 (2)	SE-PAI + LPA + HPI	<i>stx2b, stx1c + stx2b</i>	<i>SubAB, ehxA</i>	<i>tia, iha, espl, eibG, ag43-I, lpfA, ag43-II, mch</i> clúster	Diarrea (5) Asintomático (1) Sin diagnóstico (3)	Medio
O5:H19 (1), O91:H14 (4), O113:H4 (1), O128:H8 (2), O146:H21 (5), O174:H8 (2)	SE-PAI + LPA	<i>stx1c, stx1a + stx2b, stx1c + stx2b</i>	<i>subAB, ehxA, senB</i>	<i>tia, iha, espl, eibG, ag43-I, lpfA, ag43-II, mch</i> clúster	CH (2) Diarrea (3) Dolor abdominal (5) Sin diagnóstico (5)	Medio
O117:H7 (5)	HPI + ICE <i>Ec7</i>	<i>stx1a</i>	-	<i>sigA, iha, ag43-I</i>	CH (1) Diarrea (3) Dolor abdominal (1)	Medio
O27:H30 (1), O71:H12 (1), O76:H19 (3), O84:H8 (1), O113:H4 (1), O146:H21 (3), O174:H8 (2), H8 (3)	SE-PAI	<i>stx1c, stx2b, stx1c + stx2b</i>	<i>subAB, ehxA, senB</i>	<i>tia, iha, ag43-I, lpfA, mch</i> clúster genes	Diarrea (2) Asintomático (2) Sin diagnóstico (11)	Bajo
O128:H2 (2), O38:H26 (1), O91:H14 (1), O174:H21 (3)	LPA	<i>stx1c, stx1c + stx2b</i>	<i>subAB, ehxA</i>	<i>iha, espl, ag43-I, lpfA, mch</i> clúster	CH (1) Diarrea (1) Sin diagnóstico (5)	Bajo
O6:H10(1), O50/O2:H48 (1), O75:H7 (1), O92:H10 (1), O151:H12 (1)	HPI	<i>stx1a, stx1c, stx2a, stx2b</i>	-	-	Diarrea (4) Sin diagnóstico (1)	Bajo
O8:H9 (1), O8:H19 (1), 17/O77:H18 (3), O28ac/O42:H25 (1), O55:H12 (1), O91:H14 (3), O100:H30 (1), O104:H7 (2), O110:H9 (1), O113:H21 (1), O130:H11 (1), O146:H21 (1), O154:H31 (1), O169/O183:H46 (1), O171:H25 (1), O174:H25 (1), O187:H52 (2), Ont:H20 (3), H11 (1)	No identificados	<i>stx1a, stx1c, stx2a</i>	-	<i>iha, ag43-I, lpfA</i>	CH (2), Diarrea (14), Sin diagnóstico (11)	Desconocido

¹Necesita estudios epidemiológicos para su confirmación. Este esquema de clasificación se basa sólo en cepas aisladas de humanos que se analizaron en esta Tesis. ² Determinación in silico con SerotypeFinder (91). ³ Perfiles de genes virulencia más comunes. ⁴ LAC fue identificado en 6 de 11 cepas. SIDS; Síndrome de muerte súbita infantil.

8. CONCLUSIONES

La proteína Hes confiere a la cepa *E. coli* HB101 los fenotipos de hemaglutinación, autoagregación, formación de biopelícula y adhesión *in vitro* a células epiteliales humanas en cultivo.

La estructura genética completa de LAA se compone de 4 módulos. Sin embargo, una cepa STEC se puede considerar como LAA-positivo cuando al menos 3 módulos están presentes.

El Locus de Adhesión y Autoagregación es una isla de patogenicidad exclusivamente presente en cepas STEC LEE-negativo, principalmente en aquellas del filogrupo B1. Además, es la PAI más frecuente en este subgrupo de STEC.

LAA se asocia con los genes *stx1a*, *stx2a* y *stx2d*, *ehxA* y *cdtB*, que codifican las toxinas asociadas a los casos más graves de infección por STEC. Por lo tanto, la presencia de LAA junto con la de estos genes puede ser utilizada como un marcador molecular de riesgo en la vigilancia epidemiológica de STEC.

La inactivación de LAA en la cepa STEC O113:H21 E045-00 no afectó su capacidad de adhesión *in vitro* a células epiteliales humanas en cultivo. Lo anterior no descarta la participación de LAA en adhesión, y sugiere que esta cepa contiene genes adicionales que participan en este fenotipo.

LAA participa, aunque no es indispensable, en la colonización intestinal de STEC en el modelo de ratón tratado con estreptomycin.

En los serotipos STEC LEE-negativo que son clínicamente relevantes se observa una relación directa entre la presencia de más de una PAI y una distribución geográfica más amplia. Lo anterior sugiere que la adquisición acumulativa de MGEs contribuye al potencial de virulencia y emergencia de este subgrupo de STEC.

APÉNDICE

Tabla S1. Cepas *E. coli* usadas en este estudio.

Cepa	Serogrupo/serotipo	Origen	Diagnóstico *	stx1	stx2	eae	hes	LAA I	LAA II	LAA III
Cepas STEC										
E059-01	O8:H21	Humano	BD	+	+	-	-	-	-	-
E060-01	O8:H21	Humano	AD	+	-	-	-	-	-	-
E088-01	O8:H21	Humano	BD	+	-	-	-	-	-	-
E265-03	O8:H21	Humano	BD	+	-	-	-	-	-	-
A18	O22	Humano	BD	+	+	-	+	+	+	+
A13	O91:H21	Humano	BD	-	+	-	+	+	+	+
A21	O91:H21	Humano	BD	-	+	-	+	+	+	+
E040-00	O113:H21	Humano	BD	-	+	-	+	+	+	+
E042-00	O113:H21	Humano	BD	-	+	-	+	+	+	+
E043-00	O113:H21	Humano	BD	-	+	-	+	+	+	+
E044-00	O113:H21	Humano	BD	-	+	-	+	+	+	+
E045-00	O113:H21	Humano	BD	-	+	-	+	+	+	+
E46-00	O113:H21	Humano	BD	-	+	-	+	+	+	+
IH21228B	O113:H21	Humano	AD	-	+	-	+	+	+	+
IH39595/07C	O113:H21	Humano	AD	-	+	-	+	+	+	+
IH9913/07A	O113:H21	Humano	AD	-	+	-	+	+	+	+
A16	O113:H21	Humano	BD	-	+	-	+	+	+	+
A11	O139:H21	Humano	BD	-	+	-	-	-	-	-
B2P3 4	O174	Humano	AD	+	+	-	-	-	-	-
26_1	O2	Bovino	NA	-	+	-	-	-	-	-
365_1	O7	Bovino	NA	-	+	-	+	+	+	+
116_1	O20:H19	Bovino	NA	+	+	-	+	+	+	-
348_3	O46	Bovino	NA	+	+	-	+	+	+	-
310_2	O76	Bovino	NA	+	+	-	-	-	-	-
211_1	O103:H42	Bovino	NA	+	+	-	+	+	+	-
VTO-173	O104:H7	Oveja	NA	+	-	-	-	-	-	-
VTO-183	O104:H7	Oveja	NA	+	-	-	-	-	-	-
VTO-16	O104:H7	Oveja	NA	+	-	-	-	-	-	-
VTO-392-145	O104:H7	Oveja	NA	+	-	-	-	-	-	-
B45	O104:H21	Bovino	AD	+	+	-	+	+	+	+
VTB-288	O104:H21	Bovino	NA	+	+	-	+	+	+	+
VTB-104/BC12-4A	O113:H21	Bovino	NA	-	+	-	+	+	+	+
VTB-117/BC13-6A	O113:H21	Bovino	NA	-	+	-	+	+	+	+
VTB-412/MF347i	O113:H21	Bovino	NA	-	+	-	+	+	+	+
J4	O113:H21	Bovino	NA	-	+	-	+	+	+	+
5_1	O113:H21	Bovino	NA	-	+	-	-	-	-	-
6_6	O130:H11	Bovino	NA	-	+	-	-	-	-	-
13_1	O130:H11	Bovino	NA	-	+	-	-	-	-	-
27_9	O136:H16	Bovino	NA	+	-	-	-	-	-	-
J7	O139	Bovino	NA	-	+	-	+	+	+	+
64_10	O139:H19	Bovino	NA	-	+	-	-	-	-	-
208_3	O139:H19	Bovino	NA	-	+	-	-	-	-	-
175_1	O156:H-	Bovino	NA	-	+	-	+	+	+	+
218_8	O163:H9	Bovino	NA	-	+	-	+	+	+	+
J5	O171:H2	Bovino	NA	-	+	-	+	+	+	+
115_4	O171:H2	Bovino	NA	-	+	-	+	+	+	+
58_3	O174:H21	Bovino	NA	-	+	-	+	+	+	+
72_1	O181:H49	Bovino	NA	+	+	-	+	+	+	-
E107-04	O157:H7	Humano	AD	-	+	+	-	-	-	-
E111-03	O157:H7	Humano	AD	-	+	+	-	-	-	-
E112-01	O157:H7	Humano	AD	-	+	+	-	-	-	-
E113-01	O157:H7	Humano	AD	-	+	+	-	-	-	-
E114-01	O157:H7	Humano	AD	-	+	+	-	-	-	-
E115-01	O157:H7	Humano	AD	-	+	+	-	-	-	-
E116-01	O157:H7	Humano	AD	-	+	+	-	-	-	-
E117-01	O157:H7	Humano	AD	-	+	+	-	-	-	-
E118-01	O157:H7	Humano	AD	-	+	+	-	-	-	-
E120-05	O157:H7	Humano	AD	-	+	+	-	-	-	-
E128-04	O157:H7	Humano	AD	-	+	+	-	-	-	-
E158-04	O157:H7	Humano	AD	-	+	+	-	-	-	-
E168-04	O157:H7	Humano	AD	-	+	+	-	-	-	-
E211-03	O157:H7	Humano	AD	-	+	+	-	-	-	-
E310-03	O157:H7	Humano	AD	-	+	+	-	-	-	-
E62-05	O157:H7	Humano	AD	-	+	+	-	-	-	-
E014-03	O157:H7	Humano	BD	-	+	+	-	-	-	-
E016-03	O157:H7	Humano	BD	-	+	+	-	-	-	-
E102-04	O157:H7	Humano	BD	-	+	+	-	-	-	-
E104-05	O157:H7	Humano	BD	-	+	+	-	-	-	-
E121-03	O157:H7	Humano	BD	-	+	+	-	-	-	-
E123-01	O157:H7	Humano	BD	-	+	+	-	-	-	-
E186-04	O157:H7	Humano	BD	-	+	+	-	-	-	-
E205-03	O157:H7	Humano	BD	-	+	+	-	-	-	-
E32-04	O157:H7	Humano	BD	-	+	+	-	-	-	-
E99-04	O157:H7	Humano	BD	-	+	+	-	-	-	-
E107-02	O157:H7	Humano	BD	-	+	+	-	-	-	-
E111-02	O157:H7	Humano	BD	-	+	+	-	-	-	-
E161-02	O157:H7	Humano	BD	-	+	+	-	-	-	-
E42-02	O157:H7	Humano	BD	-	+	+	-	-	-	-
E015-03	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E030-00	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E089-01	O157:H7	Humano	HUS	-	+	+	-	-	-	-

E09-02	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E092-03	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E104-01	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E108-01	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E109-01	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E110-01	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E111-01	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E309-02	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E313-03	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E321-97	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E33-98	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E3-99	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E44-02	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E87-02	O157:H7	Humano	HUS	-	+	+	-	-	-	-
E92-02	O157:H7	Humano	HUS	-	+	+	-	-	-	-
IH28275/06A	O5:HNM	Humano	AD	+	-	+	-	-	-	-
IH391114/07F	O5:HNM	Humano	AD	+	+	+	-	-	-	-
E017-03	O26:H-	Humano	BD	+	-	+	-	-	-	-
E308-02	O26:H-	Humano	AD	+	-	+	-	-	-	-
3 O26/175	O26:H11	Humano	AD	+	-	+	-	-	-	-
6 O26/159	O26:H11	Humano	AD	+	-	+	-	-	-	-
E020-03	O26:H11	Humano	AD	+	-	+	-	-	-	-
E024-03	O26:H11	Humano	AD	+	-	+	-	-	-	-
E037-03	O26:H11	Humano	AD	+	-	+	-	-	-	-
E1-04	O26:H11	Humano	AD	-	+	+	-	-	-	-
E112-03	O26:H11	Humano	AD	+	-	+	-	-	-	-
E124-04	O26:H11	Humano	AD	+	-	+	-	-	-	-
E178-04	O26:H11	Humano	AD	+	-	+	-	-	-	-
E180-02	O26:H11	Humano	AD	+	-	+	-	-	-	-
E196-02	O26:H11	Humano	AD	+	-	+	-	-	-	-
E199-02	O26:H11	Humano	AD	+	-	+	-	-	-	-
E200-02	O26:H11	Humano	AD	+	-	+	-	-	-	-
E298-03	O26:H11	Humano	AD	+	-	+	-	-	-	-
E312-03	O26:H11	Humano	AD	+	-	+	-	-	-	-
E59-05	O26:H11	Humano	AD	+	-	+	-	-	-	-
E92-04	O26:H11	Humano	AD	+	-	+	-	-	-	-
E238	O26:H11	Humano	BD	+	-	+	-	-	-	-
E116-02	O26:H11	Humano	BD	+	-	+	-	-	-	-
E186-02	O26:H11	Humano	BD	-	+	+	-	-	-	-
E187-02	O26:H11	Humano	BD	-	+	+	-	-	-	-
E188-02	O26:H11	Humano	BD	+	-	+	-	-	-	-
E266-03	O26:H11	Humano	BD	+	-	+	-	-	-	-
E34-02	O26:H11	Humano	BD	+	-	+	-	-	-	-
E43-02	O26:H11	Humano	BD	+	-	+	-	-	-	-
E45-02	O26:H11	Humano	BD	+	-	+	-	-	-	-
E46-02	O26:H11	Humano	BD	+	-	+	-	-	-	-
E88-02	O26:H11	Humano	BD	+	-	+	-	-	-	-
IH10496/04A	O103:H2	Humano	AD	+	-	+	-	-	-	-
IH49694/05A	O103:H2	Humano	AD	+	-	+	-	-	-	-
VTH-74/IH2526A	O103:H2	Humano	AD	+	-	+	-	-	-	-
IH31696B	O111:[H8]	Humano	AD	+	+	+	-	-	-	-
IH24497/06A	O111:[H8]	Humano	AD	+	+	+	-	-	-	-
IH4215/07A	O111:[H8]	Humano	AD	+	-	+	-	-	-	-
IH39241/07A	O111:[H8]	Humano	AD	+	+	+	-	-	-	-
IH31928/04A	O118:H16	Humano	AD	+	-	+	-	-	-	-
VTH-28/IH781-1	O118:H16	Humano	AD	+	-	+	-	-	-	-
IH2536/07A	O118:H16	Humano	AD	+	-	+	-	-	-	-
IH56637/04a	O121:H19	Humano	AD	-	+	+	-	-	-	-
IH40478C	O145:[H28]	Humano	AD	-	+	+	-	-	-	-
IH36956/05A	O145:[H28]	Humano	AD	+	-	+	-	-	-	-
VTH-106/IH3538A	O145:[H28]	Humano	AD	+	-	+	-	-	-	-
IH38354/05A	O145:[H28]	Humano	AD	+	-	+	-	-	-	-
IH37847/06A	O145:[H28]	Humano	AD	-	+	+	-	-	-	-
VTH-105A/IH3536A	O145:[H28]	Humano	AD	-	+	+	-	-	-	-
IH32862A	O156:H25	Humano	AD	+	-	+	-	-	-	-
IH38296A	O156:H25	Humano	AD	+	-	+	-	-	-	-
VTH-56A2/IH2044B	O156:[H25]	Humano	AD	+	-	+	-	-	-	-
O26/16	O26:H11	Bovino	NA	+	-	+	-	-	-	-
O26/86	O26:H11	Oveja	NA	+	+	+	-	-	-	-
VTB-376/MF268A	O103:[H2]	Bovino	NA	+	-	+	-	-	-	-
B40/BD199B	O103:H2	Bovino	AD	+	-	+	-	-	-	-
VTB-103/BC12-3D	O103:H2	Bovino	NA	+	-	+	-	-	-	-
VTB-247/BC53-3Ta	O118:HNM	Bovino	NA	+	-	+	-	-	-	-
VTB-248/BC53ATa	O118:HNM	Bovino	NA	+	-	+	-	-	-	-
VTB-113/BC12-9Ta	O118:H16	Bovino	NA	+	-	+	-	-	-	-
Cepas <i>E. coli</i> fecales aisladas de personas sanas										
C1	-	Humano	-	-	-	-	-	-	-	-
C2	-	Humano	-	-	-	-	-	-	-	-
C3	-	Humano	-	-	-	-	-	-	-	-
C4	-	Humano	-	-	-	-	-	-	-	-
C5	-	Humano	-	-	-	-	-	-	-	-
C6	-	Humano	-	-	-	-	-	-	-	-
C7	-	Humano	-	-	-	-	-	-	-	-
C8	-	Humano	-	-	-	-	-	-	-	-
C9	-	Humano	-	-	-	-	-	-	-	-
C10	-	Humano	-	-	-	-	-	-	-	-
DM	-	Humano	-	-	-	-	-	-	-	-
HS	-	Humano	-	-	-	-	-	-	-	-

* NA, No disponible; AD, Diarrea Aguda; BD, Diarrea con sangre; HUS, Síndrome Hemolítico Urémico

Tabla S2. Genomas de cepas STEC analizados en este estudio y presencia / ausencia de genes de virulencia (Por su tamaño, la tabla se presenta en versión electrónica en formato Excel).

Tabla S3. Marcos de lectura abiertos identificados en el Locus de Adhesión y Autoagregación de la cepa STEC O91:H21 B2F1.

ORF	Posición (pb) *	Denominación en LAA	Proteína más cercana (homólogo)	No. de residuos idénticos / No. de residuos total (% Identidad)	Número de acceso de homólogo	Función descrita o predicha del homólogo
	1-76	pheV-tRNA				
	54 - 76 86428 - 86450	DR1	TTCGATTCCGAGTCCGGGCACCA TTCGATTCCGAGTCCGG-CACCA			
1	274 - 1539	Integrasa	Integrasa del bacteriófago P4	421/421 (100%)	EIH32040.1	Movilidad del DNA
2	1764 - 2742	SisA	ShiA-like	313/315 (99%)	AAN82005.1	Atenuación de la respuesta inflamatoria del hospedero
3	c3013 - 3174	Proteína hipotética	YhbX	50/53 (94%)	EGB65631.1	Desconocida
4	c3197 - 4699	Sulfatasa	AriIsulfatasa	483/484 (99%)	EIH33024.1	Desconocida
	4131 - 4146 36258 - 36273	DR2	AACGGTACACCCGGCA AACGGAACACCCGGAA			
5	c4881 - 5630	Hes	Hra1	223/250 (89%)	AAC13752.1	Autoagregación, formación de biopelícula y adherencia
6	6009 - 7835	AtoS	AtoS	595/608 (98%)	Q06067.1	Activa AtoC mediante fosforilación
7	7832 - 9217	AtoC	AtoC	460/461 (99%)	Q06065.2	Regulador transcripcional
8	9413 - 10075	AtoD	AtoD	219/220 (99%)	P76458.1	
9	10075 - 10725	AtoA	AtoA	214/216 (99%)	P76459.1	Metabolismo de ácidos grasos de cadena corta
10	10722 - 12044	AtoE	AtoE	440/440 (100%)	P76460.1	
11	12075 - 13259	AtoB	AtoB	388/394 (98%)	P76461.1	
12	c13669 - 14025	IS66	Transposasa IS66	118/118 (100%)	ERF94613.1	Movilidad del DNA
13	15414 - 15593	Proteína hipotética	Proteína hipotética	57/57 (100%)	EGW68387.1	Desconocida
14	15645 - 15818	Proteína hipotética	Proteína hipotética	57/57 (100%)	EIH33102.1	Desconocida
15	c16333 - 16548	Proteína hipotética	Proteína hipotética	61/64 (95%)	AlF92868.1	Desconocida
16	16618 - 17016	ShiD	ColV	130/132 (98%)	AAN45157.1	Colicina
	17407 - 17517	DR3	TTTCGTAG			
17	c17534 - 17950	ShiA-like	ShiA-like	101/108 (94%)	EZJ33232.1	Desconocida
18	18447 - 18599	Proteína hipotética	Proteína hipotética	49/50 (98%)	EFW62877.1	Desconocida
19	c19283 - 19465	Proteína hipotética	Proteína hipotética	60/60(100%)	EGW68391.1	Desconocida
20	19404 - 21491	lha	lha	664/696 (95%)	NP_309387.1	Receptor específico de la proteína enterobactina / adhesina
21	c21482 - 21808	TnpA	Transposasa TnpA	108/108 (100%)	ERF87109.1	Movilidad del DNA
22	c21837 - 23681	BtuB	BtuB	614/614 (100%)	EGW81509.1	Transportador de membrana externa de vitamina B12 / cobalamina
23	c24582 - 24839	Proteína hipotética	Proteína hipotética	85/85 (100%)	EZA22235.1	Desconocida
24	c24908 - 26026	Sulfatasa	Sulfatasa	370/372(99%)	EJK95236.1	Desconocida
25	c26142 - 26594	AriIsulfatasa	AriIsulfatasa	150/150(100%)	EIH33112.1	Desconocida
26	c26689 - 27846	NmpC	NmpC	375/378 (99%)	ELE48963.1	Resistencia al calor
27	c28888 - 29133	Proteína hipotética	Proteína hipotética	71/81 (88%)	EGW92529.1	Desconocida
28	30142 - 30699	AraC-like	Regulador transcripcional de la familia AraC	245/245 (100%)	EYZ57499.1	Regulación transcripcional (predicho)
29	31598 - 35692	LesP	Autotransportador	1363/1364 (99%)	WP_032317126.1	Serino proteasa (predicho)
30	35954 - 36187	Proteína hipotética	Proteína hipotética	63/63 (100%)	WP_000789660.1	Desconocida
31	36573 - 36764	Proteína hipotética	Proteína hipotética	63/63 (100%)	WP_000264910.1	Desconocida
32	36798 - 37007	Proteína hipotética	FlgM	69/69 (100%)	WP_000323309.1	Desconocida
33	c37034 - 37381	IS66	Transposasa, trunca	115/115 (100%)	WP_000855098.1	Movilidad del DNA
34	c37505 - 38149	Proteína hipotética	Proteína hipotética	214/214 (100%)	WP_001376509.1	Desconocida
35	c38170 - 38439	PerC-like	PerC-like	89/89 (100%)	KEO26191.1	Regulación transcripcional (predicho)
36	c38518 - 39156	lbrB	lbrB	212/212 (100%)	EIH32727.1	Regulación transcripcional (predicho)
	39023 - 39038 84116 - 84131	DR4				
37	c39141 - 40373	lbrA	lbrA	410/410 (100%)	EIH33651.1	Regulación transcripcional (predicho)
38	c40873 - 43554	UvrD Helicasa	UvrD Helicasa	886/893 (99%)	EKI41417.1	Superfamilia I Helicasas de DNA y RNA
39	c43805 - 44263	ImpA	ImpA	152/152 (100%)	WP_001367976.1	Proteína de secreción de tipo VI. Activación de hemolisina (predicho)
40	c44244 - 44621	Proteína hipotética	Proteína hipotética	115/115 (100%)	KDV76601.1	Desconocida
41	c45046 - 46872	UvrD Helicasa	UvrD Helicasa	604/604 (100%)	EIH32041.1	Superfamilia I Helicasas de DNA y RNA
42	c46969 - 47544	PagC-like	PagC-like	182/182 (100%)	ABG02919.1	Desconocida
	48181 - 48304 49774 - 49887	DR5	TTC TTACCC TGTGCCAGAGACAGGGCTGCATTCACTACTGCATGAGCTGCTACACGCCCTGCTCCGTCGGG GTGGTCTGTTTGATAATTCAGCAATATACGGTGC TGACCAGC			
43	c48729 - 57857	TpsA	Hemaglutinina filamentosa	3042/3042 (100%)	WP_001081255.1	Hemaglutinación, adherencia (predicho)
	50133 - 50148 57786 - 57801	DR6	GGCAATATCTGCCACC			
	51541 - 51556 52639 - 52654 60937 - 60952	DR7	TCCCTGCCGGCAATCA			
44	c57873 - 58400	HlyC-like	Toxina RTX	175/175 (100%)	WP_001243916.1	Familia de toxina RTX aciltransferasa (predicho)
45	c58410 - 60062	TpsB	ShB / FhaC / HecB	550/550 (100%)	WP_001376511.1	Activación / secreción de hemolisina (predicho)
46	60738 - 60929	Proteína hipotética	Z1637-like	63/63 (100%)	AAQ19123.1	Desconocida

47	c60982 - 61215	Proteína hipotética	Z1636-like	77/77 (100%)	AAQ19122.1	Desconocida
48	c61311 - 61934	Proteína hipotética	Proteína hipotética	207/207 (100%)	WP_001367955.1	Desconocida
49	c62870 - 63436	Proteína hipotética	Proteína hipotética	188/188 (100%)	WP_000287892.1	Desconocida
50	63739 - 64041	IS600	Transposasa	100/100 (100%)	WP_000088311.1	Movilidad del DNA
51	64182 - 64895	IS600	Transposasa	272/272 (100%)	WP_001367783.1	Movilidad del DNA
52	c65169 - 65594	Aec59	Aec59	141/141 (100%)	WP_001013311.1	Desconocida
53	c65591 - 65974	Aec60	Aec60	127/127 (100%)	WP_000271013.1	Desconocida
54	c66235 - 66801	Aec61	Aec61	188/188 (100%)	WP_000221486.1	Desconocida
55	67005 - 67202	Proteína hipotética	Proteína hipotética	65/65 (100%)	WP_000236763.1	Desconocida
56	c67548 - 67688	Proteína hipotética	Proteína hipotética	46/46 (100%)	WP_000623359.1	Desconocida
57	c68195 - 68308	Proteína hipotética	Proteína hipotética	37/37 (100%)	EFZ40174.1	Desconocida
58	68449 - 68655	Proteína hipotética	Proteína hipotética	92/92 (100%)	EHV18472.1	Regulador transcripcional predicho AlpA (predicho)
59	68736 - 69056	Aec62, trunco	Aec62	104/106 (98%)	WP_033810310.1	Desconocida
60	69099 - 69410	IS602	IS602	103/103 (100%)	WP_001036724.1	Movilidad del DNA
61	69650 - 70276	IS602	IS602	208/208 (100%)	BAI34548.1	Movilidad del DNA
62	70297 - 70605	Aec62, trunco	Aec62	102/102 (100%)	WP_032235550.1	Desconocida
63	c70927 - 72411	Aec63	Aec63	494/494 (100%)	WP_001297234.1	Desconocida
64	72561 - 72803	proteína hipotética	proteína hipotética	79/80 (99%)	WP_024231225.1	Desconocida
65	c72843 - 72965	proteína hipotética	proteína hipotética	40/40 (100%)	EDV62536.1	Desconocida
66	73641 - 74513	YeeP	GTPasa	290/290 (100%)	ENC40329.1	GTPasa (predicho)
67	74841 - 77960	Ag43	Ag43	1039/1039 (100%)	EIH33235.1	Adhesión, autoagregación, formación de biopelícula
68	78081 - 80597	Aec68	Aec68	838/838 (100%)	WP_001367924.1	Desconocida
69	80673 - 81128	Aec69	Aec69	148/151 (98%)	WP_000581504.1	Desconocida
70	81207 - 81440	Aec70	Aec70	77/77 (100%)	WP_001119729.1	Desconocida
71	81540 - 82358	Aec71	Aec71	271/272 (99%)	WP_001234620.1	Desconocida
72	82413 - 82898	KlcA-like	KlcA-like	161/161 (100%)	WP_000849590.1	Desconocida
73	82899 - 83390	YeeS	YeeS	163/163 (100%)	CAE85201.1	Desconocida
74	83453 - 83674	YeeT	YeeT	73/73 (100%)	WP_000692312.1	Desconocida
75	83674 - 83787	proteína hipotética	proteína hipotética	37/37 (100%)	WP_000488311.1	Desconocida
76	83939 - 84211	YeeU	YeeU	89/90 (99%)	WP_001280952.1	Antitoxina
77	84258 - 84635	YeeV	YeeV	125/125 (100%)	WP_000854902.1	Toxina
78	84632 - 85120	proteína hipotética	proteína hipotética	162/162(100%)	WP_000779175.1	Desconocida
79	85132 - 85329	Aec78	Aec78	76/80 (95%)	CAX18608.1	Desconocida
80	85414 - 86256	Aec79	Aec79	280/280(100%)	WP_001280537.1	Desconocida

* c: indica ORFs transcritos en la hebra complementaria.

Tabla S4. Homólogos de LesP entre miembros de la familia de proteínas Serino Proteasas Autotransportadoras de Enterobacteriaceae.

Homólogo	Origen	No. de acceso	% Identidad (% Similitud) ¹			Función	Fuente
			Proteína completa	Dominio pasajero	Dominio Autotransportador		
EspI	STEC	CAC39286.1	76.1 (84.9)	69.6 (81.2)	96.6 (97.7)	Proteasa	(72)
EpeA	pO113	AAL18821.1	58.8 (72.7)	47.6 (65.1)	99.6 (99.6)	Mucinasas	(66)
SepA	<i>S. flexneri</i>	CAA88252.1	52.8 (67.7)	46.9 (62.2)	78.4 (91.7)	Citotoxina	(158)
EspP	pO157	CAA66144.1	52.1 (64.0)	40.9 (55.1)	100 (100)	Citotoxina	(71)
Pic	<i>S. flexneri</i> y <i>E. coli</i> 042	AAD23953.1	50.5 (66.2)	43.5 (60.2)	79.6 (90.6)	Mucinasas, Resistencia a suero, hemaglutinación	(159)
EatA	ETEC	AAO17297.1	50.6 (66.9)	45.9 (62.3)	73.1 (86.7)	Proteasa	(160)
Tsh	APEC	Q47692.1	44.4 (61.6)	40.7 (57.9)	60.2 (77.1)	Hemaglutinación	(161)
TieA	ETEC	KF494347.1	43.2 (60.1)	39.0 (55.8)	60.2 (77.1)	Mucinasas, Adhesión	(86)

¹ Determinado mediante EMBOSS Matcher version 2.0u4 (http://www.ebi.ac.uk/Tools/psa/emboss_matcher/). Porcentaje de identidad y similitud con respecto a la secuencia de aminoácidos de LesP (Número de Acceso: EGW68403.1) presente en la cepa B2F1.

Tabla S5. Perfiles de Shiga toxina (> 1 subtipo) identificados entre las cepas analizadas en este estudio.

Subtipos de Stx	No de cepas (%)
Stx1a, Stx2a	32 (8,7)
Stx1c, stx2b	30 (8,2)
Stx1a, stx2d	13 (3,6)
Stx1a, stx2b	8 (2,7)
Stx2a, Stx2a	6 (1,6)
Stx2a, Stx2d	2 (0,55)
Stx2b, Stx2c	2 (0,55)
Stx2b, Stx2d	2 (0,55)
Stx1c, Stx2a	1(0,27)
Stx1c, Stx2c	1(0,27)
Stx1c, Stx2d	1 (0,27)
Stx2a, Stx2b	1 (0,27)
Stx1a, Stx2a, Stx2d	1(0,27)
Stx1c, Stx2b, Stx2b	1 (0,27)
Total	101 / 367 (27,5)

Tabla S6. Marcos de lectura abiertos identificados en el Locus de Invasión e Inhibición del Crecimiento Dependiente del Contacto de la cepa STEC O174:H21 MOD1-EC1633.

ORF	Posición (pb) *	Denominación en LIC	Proteína más cercana (homólogo)	No. de residuos idénticos / No. de residuos total (% Identidad)	Número de acceso de homólogo	Función descrita o predicha del homólogo
	1-91	se/C-tRNA				
	72 - 95 48318- 48341	DR	TTCGACTCCTGTGATCTCCGCCA TTCGACTCCTGTGATCT-CCGCCA			
1	395 - 1579	Integrasa	Integrasa	394/394 (100%)	WP_001218900.1	Movilidad del DNA
2	11991 – 3034	sisA	ShiA	343/347 (99%)	AHA68728.1	Atenuación de la respuesta inflamatoria del hospedero
3	c3385 - 4887	Proteína hipotética	Hidrolasa	500/500 (100%)	YP_002410067	Desconocida
4	C5062 - 5808	Tia	Tia	248/248 (100%)	YP_002410068	Adhesión and invasión
5	6065 - 6216	Integrasa (trunco)	Integrase (trunco)	52/53 (98%)	YP_002410069.1	Desconocido
6	6410 - 6763	Proteína hipotética	Proteína hipotética	116/117 (99%)	EFJ81903.1	Desconocida
7	7405 - 8361	Proteína hipotética	Proteína hipotética	318/318 (100%)	EHN85659.1	Desconocida
8	c8342-9271	Proteína hipotética	DNA citosina metiltransferasa	309/309 (100%)	WP_001328257.1	Desconocida
9	9845-10099	ISRaql-like	ISRaql	74/82 (90%)	AY528232	Movilidad del DNA
10	10312-10587	IS629	IS629	87/90 (97%)	X51586	Movilidad del DNA
11	10587 - 11066	IS1203E	IS1203E	132/137 (96%)	X97542	Movilidad del DNA
12	11658 - 13181	Proteína hipotética	Proteína hipotética	502/502 (100%)	YP_002406484.1	Desconocida
13	13376 - 13750	IS911	IS911	122/125 (98%)	X17613	Movilidad del DNA
14	14217 - 14678	Proteína hipotética	Dominio de la proteína membrana externa de la familia hemaglutinina filamentosa	153/153 (100%)	EKI34454.1	Inhibición del crecimiento dependiente del contacto (predicho)
15	14716- 14892	Pre-toxina, dominio VENN	Pre-toxina, dominio VENN	58/58 (100%)	EKI34455.1	
16	c14927 - 15250	Proteína hipotética	Proteína con dominio DUF2569	106/107(99%)	WP_077779681.1	Desconocida
17	c15463 – 15714	Proteína hipotética	Proteína con Dominio hemolisina	83/83 (100%)	KDU01438.1	Inhibición del crecimiento dependiente del contacto (predicho)
18	c15978 - 16412	Proteína hipotética	Hemaglutinina filamentosa (dominio)	144/144 (100%)	KDU01311.1	
19	c16876 - 17361	Proteína hipotética	Proteína hipotética	161/161 (100%)	WP_032260910.1	Desconocida
20	c17358 - 27137	CdiA	CdiA	3182/3259 (98%)	EKI34460.1	Inhibición del crecimiento dependiente del contacto (predicho)
21	c27150 - 28916	CdiB	CdiB	588/588 (100%)	WP_077166514.1	
22	c29610 - 30233	Proteína hipotética	Proteína hipotética	207/207 (100%)	KHG70693.1	Desconocida
23	c30322 - 30834	Proteína hipotética	Proteína hipotética	169/170 (99%)	WP_001680714.1	Desconocida
24	c31403 - 31975	Proteína hipotética	Proteína hipotética	190/190 (100%)	WP_032260913.1	Desconocida
25	c32221 – 32787	Proteína hipotética	Proteína hipotética	187/188 (99%)	EDV68605.1	Desconocida
26	34274 - 35626	Proteína hipotética	Proteína hipotética	450/450 (100%)	WP_032260914.1	Desconocida
27	c36691 - 37422	Transposasa	Transposasa	243/243 (100%)	KHG90946.1	Movilidad del DNA
28	38347 – 38622	Proteína hipotética	Proteína hipotética	91/91 (100%)	WP_001551885.1	Desconocida
29	39738 - 40961	Proteína hipotética	Proteína hipotética	407/407 (100%)	WP_001107214.1	Desconocida
30	41062 - 41946	Proteína hipotética	Proteína de la familia GTPasa de unión a ribosoma 50S	294/294 (100%)	WP_032260917.1	Desconocida
31	42578 - 42829	Proteína hipotética	Proteína hipotética	83/83 (100%)	PKZ09602.1	Desconocida
32	42977 – 43654	Proteína hipotética	Proteína hipotética	225/225 (100%)	WP_033801393.1	Desconocida
33	43717 - 43893	Proteína hipotética	Proteína con dominio DUF905	58/58 (100%)	WP_033801392.1	Desconocida
34	43983 – 44801	Proteína hipotética	Proteína con dominio DUF945	272/272 (100%)	WP_033802047.1	Desconocida
35	45067 - 45546	KclA	Proteína anti-restricción	159/159 (100%)	WP_000706975.1	Desconocida
36	45561 - 46037	YeeS	RadC	158/158 (100%)	WP_001318130.1	Desconocida
37	46124 – 46345	YeeT	Proteína con dominio DUF987	72/73 (99%)	WP_000692350.1	Desconocida
38	46545 – 46724	YeeU (trunco)	Antitoxina (trunco)	44/45 (98%)	WP_071852595.1	Antitoxina
39	46834 – 47256	Proteína hipotética	Proteína con dominio DUF957	140/140 (100%)	WP_001521668.1	Desconocida
40	47341 - 48183	Proteína hipotética	Proteína con dominio DUF4942	280/280 (100%)	WP_001521669.1	Desconocida

* c: indica ORFs transcrito en la hebra complementaria

Tabla S7. Marcos de lectura abiertos identificados en el Locus de Adhesión y Colonización de la cepa *E. coli* FCH1.

ORF	Posición (pb) *	Denominación en LAC	Proteína más cercana (homólogo)	No. de residuos idénticos / No. de residuos total (% Identidad)	Número de acceso de homólogo	Función descrita o predicha del homólogo
	1 - 76	<i>pheV</i> -tRNA				
	54 - 76	DR1	TTCGATTCCGAGTCCGGGCACCA			
	71335 - 71356		TTCGATTCCGAGTCCGGC-ACCA			
1	349 - 1536	Integrasa	Integrasa	395/395(100%)	EFK51858.1	Movilidad del DNA
2	c1801-3072	Proteína hipotética	Hidrolasa	411/423 (97%)	WP_001560767.1	Desconocida
3	c3478 - 4224	Tia	Tia	237/248(96%)	OAC05898.1	Adhesión and invasión
4	4471 - 4632	Integrase (trunca)	Integrase (trunca)	50/53 (94%)	YP_002410069.1	Desconocida
5	4838 - 5599	Proteína hipotética	Endonucleasa de restricción	253/253 (100%)	WP_074435527.1	Endonucleasa de restricción (predicha)
6	c5791 - 6132	Proteína hipotética	Proteína hipotética	113/113 (100%)	WP_028132771.1	Desconocida
7	c6102 - 6746	Proteína hipotética	Proteína hipotética	179/180 (99%)	WP_028132771.1	Desconocida
8	c7265 - 8587	Proteína hipotética	Proteína hipotética	440/440 (100%)	WP_029488460.1	Desconocida
9	c8683 - 11964	Proteína hipotética	Proteína hipotética	1093/1093 (100%)	WP_029488459.1	Desconocida
10	c12032 - 12775	Proteína hipotética	Proteína hipotética	241/245 (98%)	EKI46689.1	Desconocida
11	c12839 - 13579	Proteína hipotética	Proteína hipotética	198/201 (99%)	ONG20068.1	Desconocida
12	13828 - 14067	SymE-like toxina	SymE-like toxina	79/79 (100%)	WP_074435529.1	Sistema toxina-antitoxina tipo I (predicho)
13	c14112 - 14339	Proteína hipotética	Proteína hipotética	75/75 (100%)	PAU22824.1	Desconocida
14	c14336 - 24049	CdiA	CdiA	2747/3094 (89%)	EKI34460.1	Inhibición del crecimiento dependiente del contacto
15	c24062 - 25834	CdiB	CdiB	590/590 (100%)	WP_029488455.1	
16	26350 - 26739	Proteína hipotética	Proteína hipotética	129/129 (100%)	WP_000833174.1	
17	26804 - 27862	Metiltransferasa	Metiltransferasa	352/352 (100%)	WP_029488454.1	
18	27903 - 28625	Proteína hipotética	Proteína hipotética	240/240 (100%)	WP_000263010.1	
19	28622 - 29443	Aciltransferasa	Aciltransferasa	273/273 (100%)	WP_077696946.1	
20	29418 - 29675	Proteína transportadora de grupos acilos	Proteína transportadora de grupos acilos	85/85 (100%)	WP_001148685.1	
21	29687 - 29938	Proteína transportadora de grupos acilos	Proteína transportadora de grupos acilos	83/83 (100%)	WP_000132059.1	
22	29943 - 30524	Proteína de membrana putativa	Proteína de membrana putativa	193/193 (100%)	WP_001442994.1	
23	30521 - 31879	Proteína hipotética	Proteína hipotética	452/452 (100%)	WP_001077064.1	
24	31866 - 32219	Acil-[proteína-transportadora-de-acilo] desaturasa	Acil-[proteína-transportadora-de-acilo] desaturasa	117/117 (100%)	OSK21478.1	O-Island#138. Biosíntesis de ácidos grasos (Predicho)
25	32210 - 33886	Aciltransferasa	Aciltransferasa	558/558 (100%)	WP_029488452.1	
26	33890 - 34312	Acil tioesterasa	Acil tioesterasa	119/119 (100%)	WP_029488451.1	
27	34309 - 34914	LolA	LolA	201/201(100%)	WP_000670563.1	
28	34952 - 37201	Proteína de membrana putativa	Proteína de membrana putativa	745/749 (99%)	WP_000180163.1	
29	37198 - 37782	Proteína hipotética	Proteína con dominio DUF3261	194/194 (100%)	WP_000597701.1	
30	37784 - 38953	Beta-cetoacil sintasa	Beta-cetoacil sintasa	389/389 (100%)	WP_001350102.1	
31	38950 - 39414	Acil-ACP deshidratasa	Acil-ACP deshidratasa	153/154 (99%)	WP_000020242.1	
32	39414 - 40145	FabG	FabG	243/243 (100%)	WP_000091658.1	
33	40142 - 41371	Beta-cetoacil sintasa	Beta-cetoacil sintasa	409/409 (100%)	WP_000198472.1	
34	c42211 - 43242	Regulador VpeR	Regulador VpeR	343/343 (100%)	WP_000416157.1	
35	43513 - 43956	VpeA	VpeA	147/147(100%)	WP_000916811.1	
36	44972 - 44259	VpeB	VpeB	95/95 (100%)	WP_000705928.1	Operón <i>vpe</i> , Virulencia y colonización intestinal
37	44272 - 45528	VpeC	VpeC	418/418 (100%)	WP_000345347.1	
38	c45744 - 45965	Transposasa	Transposasa	73/73 (100%)	EYZ96274.1	Movilidad del DNA
39	c46450 - 47463	DeoM	DeoM	337/337(100%)	WP_000107474.1	
40	c47475 - 48791	DeoP	DeoP	438/438 (100%)	WP_000998349.1	Ciúster <i>deoK</i> . Promueve el crecimiento <i>de E. coli</i> durante la colonización del huésped
41	c48819 - 49739	DeoK	DeoK	306/306 (100%)	WP_000350265.1	
42	50042 - 50824	DeoR	Regulador DeoR	260/260 (100%)	WP_001315616.1	
43	c50925 - 51041	Proteína hipotética	Proteína hipotética	37/38 (97%)	EEJ46496.1	Desconocida
44	c51827 - 52261	Proteína hipotética	Proteína hipotética	144/144 (100%)	WP_096271942.1	Desconocida
45	c52249 - 52656	Proteína hipotética	Proteína hipotética	135/135 (100%)	WP_001443055.1	Desconocida
46	c52910 - 53479	Proteína hipotética	Proteína hipotética	189/189 (100%)	WP_000221502.1	Desconocida
47	53683 - 53880	Proteína hipotética	Proteína hipotética	65/65 (100%)	WP_000236763.1	Desconocida
48	c54225 - 54365	Proteína hipotética	Proteína hipotética	46/46 (100%)	WP_071527034.1	Proteína moduladora de la expresión de hemolisina (predicho)
49	55126 - 55332	AlpA	AlpA	68/68 (100%)	WP_001398320.1	Regulador transcripcional (predicho)
50	55419 - 56021	Proteína hipotética	Proteína hipotética	200/200 (100%)	WP_077248524.1	Desconocida
51	c56343 - 57827	Proteína hipotética	Proteína con dominio DUF3987	494/494 (100%)	WP_077792018.1	Desconocida
52	59058 - 59930	YeeP	YeeP	290/290 (100%)	WP_078207435.1	Desconocida
53	60134 - 60292	Proteína hipotética	Proteína hipotética	52/52 (100%)	WP_001332431.1	Desconocida
54	60303 - 63149	Ag43-II	Ag43	948/948 (100%)	WP_029488448.1	Adhesión, autoagregación, formación de biopelícula
55	63270 - 65786	Proteína hipotética	Proteína hipotética	838/838 (100%)	WP_029488447.1	Desconocida

56	65862 – 66317	<i>aec69</i>	Aec69	151/151 (100%)	WP_000581504.1	Desconocida
57	c66438 – 66626	Proteína hipotética	Proteína con dominio DUF905	63/63 (100%)	WP_042036194.1	Desconocida
58	66729 – 67547	Proteína hipotética	Proteína con dominio DUF945	272/272 (100%)	WP_029488446.1	Desconocida
59	67602 – 68087	KclA	KclA	159/161 (99%)	WP_000849582.1	Desconocida
60	68103 – 68579	YeeS	YeeS	158/158 (100%)	ABE06628.1	Desconocida
61	66648 – 68869	YeeT	YeeT	73/73 (100%)	WP_000692329.1	Desconocida
62	68869 – 68982	Proteína hipotética	Proteína hipotética	37/37 (100%)	WP_000488318.1	Desconocida
63	69134 – 69400	YeeV	YeeV	88/88 (100%)	KYR89313.1	Antitoxina
64	69490 – 69864	YeeU	YeeU	124/124 (100%)	WP_000854753.1	Toxina
65	69861 – 70349	Proteína hipotética	Proteína hipotética	162/162 (100%)	WP_000777541.1	Desconocida
66	70361 – 70558	Proteína hipotética	Proteína con dominio DUF957	65/65 (100%)	WP_001317562.1	Desconocida
67	70655 - 71224	Proteína hipotética	Proteína con dominio DUF4942	188/189 (99%)	WP_001290240.1	Desconocida

* c: indica ORFs transcrito en la hebra complementaria

Tabla S8. Marcos de lectura abiertos identificados en el Elemento Integrativo y Conjugativo No. 7 de la cepa STEC O117:H7 FHI72.

ORF	Posición (pb) *	Denominación en LAC	Proteína más cercana (homólogo)	No. de residuos idénticos / No. de residuos total (% Identidad)	Número de acceso de homólogo	Función descrita o predicha del homólogo
	1 - 76	<i>pheV</i> -tRNA				
	26 - 77 86457 - 86508	DR	AGGGGATTGAAATCCCCGTGTCCTTGGTTCGATTCCGAGTCCGGGCACCAC			
1	c275 - 1231	Integrasa	Recombinasa XerD	314/318 (98%)	WP_000777206.1	Movilidad del DNA
2	c1301 - 2911	Helicasa / relaxasa	Helicasa / relaxasa	528/532 (99%)	WP_001625914.1	Movilidad del DNA (predicho)
3	c2904-4412	Helicasa	Helicasa	502/502 (100%)	WP_001120272.1	Replicación, recombinación y reparación del DNA (predicho)
4	C4540 - 4971	Proteína hipotética	Proteína hipotética	117/118 (99%)	EZJ31791.1	Desconocida
5	4952 - 5182	Proteína hipotética	Proteína hipotética	76/76 (100%)	WP_001138606.1	Desconocida
6	5191 - 5385	Proteína hipotética	Proteína hipotética	63/64 (98%)	WP_042104503.1	Desconocida
7	5395 - 5631	Proteína hipotética	Proteína hipotética	72/78 (92%)	WP_074014985.1	Desconocida
8	5664 - 6353	Proteína hipotética	Proteína hipotética	73/78 (94%)	WP_021538137.1	Desconocida
9	6354 - 6788	PilM	PilM	144/144 (100%)	WP_033807106.1	Biosíntesis de Pili tipo IV
10	6997 - 8598	PilN	PilN	533/533 (100%)	WP_000713048.1	
11	8610 - 9917	PilO	PilO	431/431 (100%)	WP_001169832.1	
12	9907 - 10368	PilP	PilP	152/153 (99%)	WP_000543768.1	
13	10378 - 11898	PilQ	PilQ	506/506 (100%)	WP_001115303.1	
14	11900 - 13000	PilR	PilR	366/366 (100%)	WP_000876118.1	
15	13052 - 13585	PilS	PilS	177/177 (100%)	WP_001011155.1	
16	13644 - 14120	PilT	PilT	158/158 (100%)	WP_000871802.1	
17	14127 - 14783	Proteína hipotética	Proteína hipotética	214/216 (99%)	WP_001292581.1	
18	14780 - 15997	PilV	PilV	404/405 (99%)	WP_016237204.1	
19	c15994 - 16227	Proteína hipotética	Proteína hipotética	74/74 (100%)	WP_106482758.1	
20	16282 - 17406	Proteína hipotética	Proteína hipotética	374/374 (100%)	WP_001672437.1	
21	17967 - 18788	TraE	TraE	273/273 (100%)	WP_033808608.1	
22	18878 - 20080	TraF	TraF	400/400 (100%)	WP_000979995.1	
23	20442 - 24299	SigA	SigA	1267/1285 (98%)	AAF67320.1	Proteasa citopática exportada implicada en la acumulación de fluido intestinal
24	24872 - 25894	IS100	IS100	340/340 (100%)	WP_000255944.1	Movilidad del DNA
25	25894 - 26673	Transposasa	Transposasa	259/259 (100%)	WP_001323403.1	Movilidad del DNA
26	26712 - 27089	Proteína hipotética	Proteína hipotética	114/115 (99%)	WP_001682055.1	Desconocida
27	27136 - 28059	Transposasa	Proteína hipotética	307/307 (100%)	WP_033808929.1	Movilidad del DNA
28	28133 - 28501	Transposasa	Transposasa	122/122 (100%)	WP_096955009.1	Movilidad del DNA
29	28655 - 29233	PilL	PilL	192/192 (100%)	WP_032308781.1	Conjugación (predicho)
30	29230 - 2997	Proteína hipotética	Proteína hipotética	255/255 (100%)	WP_032308782.1	Desconocida
31	30002 - 30724	Proteína hipotética	Proteína hipotética	239/240 (99%)	WP_000782640.1	Desconocida
32	30703 - 31365	Proteína hipotética	Proteína hipotética	219/220 (99%)	WP_032308783.1	Desconocida
33	31371 - 31892	Proteína hipotética	Proteína hipotética	168/168 (100%)	WP_033807331.1	Desconocida
34	31892 - 32104	Proteína hipotética	Proteína hipotética	30/30 (100%)	WP_032308785.1	Desconocida
35	32101 - 32490	Proteína hipotética	Proteína hipotética	128/129 (99%)	WP_032308785.1	Desconocida
36	32504 - 33292	Proteína hipotética	Proteína hipotética	262/262 (100%)	WP_032308786.1	Desconocida
37	33285 - 35402	TraD	TraD	701/705 (99%)	WP_062864066.1	Conjugación (predicho)
38	35383 - 36144	Proteína hipotética	Proteína hipotética	249/253 (98%)	WP_000796666.1	Desconocida
39	c36401 - 36559	Proteína hipotética	Proteína hipotética	51/52 (98%)	WP_032308788.1	Desconocida
40	c36826 - 36948	Proteína hipotética	Proteína hipotética	38/39 (97%)	WP_000997863.1	Desconocida
41	37094 - 37435	Proteína hipotética	Proteína hipotética	113/113 (100%)	WP_000492099.1	Desconocida
42	37435 - 37677	Proteína hipotética	Proteína hipotética	80/80 (100%)	WP_032308789.1	Desconocida
43	37708 - 38085	Proteína hipotética	Proteína hipotética	125/125 (100%)	WP_001681767.1	Desconocida
44	38105 - 38470	Proteína hipotética	Proteína hipotética	120/121 (99%)	WP_001681768.1	Desconocida
45	38467 - 39114	Proteína hipotética	Proteína hipotética	215/215 (100%)	WP_000086084.1	Desconocida
46	39111 - 40019	Proteína hipotética	Proteína hipotética	300/302 (99%)	WP_001681769.1	Desconocida
47	40009 - 41487	Proteína hipotética	Proteína hipotética	491/492 (99%)	WP_001681770.1	Desconocida
48	41505 - 41918	Proteína hipotética	Proteína hipotética	135/137 (99%)	WP_032308793.1	Desconocida
49	41918 - 44734	Proteína hipotética	Proteína hipotética	931/938 (99%)	WP_001681772.1	Desconocida
50	44731 - 45117	Proteína hipotética	Proteína hipotética	127/128 (99%)	WP_032308795.1	Desconocida
51	c45363 - 45536	Proteína hipotética	Proteína hipotética	57/57 (100%)	WP_033807339.1	Desconocida
52	45508 - 47598	Iha	Iha	686/696 (99%)	WP_001223350.1	Receptor sideróforo / adhesión
53	47825 - 48223	Proteína hipotética	Proteína hipotética	132/132 (100%)	WP_033807341.1	Desconocida
54	48220 - 49194	Proteína hipotética	Proteína hipotética	324/324 (100%)	WP_033807342.1	Desconocida
55	49203 - 50636	Proteína hipotética	Proteína hipotética	475/476 (99%)	WP_033807343.1	Desconocida
56	50629 - 50994	Proteína hipotética	Proteína hipotética	121/121 (100%)	WP_001681777.1	Desconocida
57	50998 - 52500	TraG	TraG	500/500 (100%)	WP_033807344.1	Conjugación (predicho)
58	c52547 - 52846	Proteína hipotética	Proteína hipotética	97/99 (98%)	WP_077953133.1	Desconocida
59	53383 - 53604	Proteína hipotética	Proteína hipotética	73/73 (100%)	WP_033807345.1	Desconocida
60	c53948 - 54292	Proteína hipotética	Proteína hipotética	113/114 (99%)	WP_033807353.1	Desconocida
61	54952 - 55494	Proteína hipotética	Proteína hipotética	131/133 (98%)	WP_001672448.1	Desconocida
62	55665 - 56045	Proteína hipotética	Proteína hipotética	126/126 (100%)	WP_001681645.1	Desconocida
63	56146 - 56625	Proteína hipotética	Proteína hipotética	157/159 (99%)	WP_021544261.1	Desconocida
64	56603 - 56740	Proteína hipotética	Proteína hipotética	43/45 (96%)	KGM62803.1	Desconocida
65	56818 - 57816	Proteína hipotética	Proteína hipotética	325/327 (99%)	WP_001672446.1	Desconocida
66	57927 - 58286	Proteína hipotética	Proteína hipotética	119/119 (100%)	WP_000287297.1	Desconocida
67	58354 - 58692	Proteína hipotética	Proteína hipotética	112/112 (100%)	WP_001247170.1	Desconocida
68	58766 - 59203	Proteína hipotética	Proteína hipotética	144/144 (100%)	WP_001101430.1	Desconocida
69	59311 - 59853	Proteína hipotética	Proteína hipotética	175/175 (100%)	WP_021531919.1	Desconocida
70	59919 - 60257	IS21	IS21	109/112 (97%)	ABV19510.1	Movilidad del DNA
71	60380 - 61066	Proteína hipotética	Proteína hipotética	228/228 (100%)	WP_000206397.1	Desconocida

72	61198 - 61929	Proteína hipotética	Proteína hipotética	243/243 (100%)	WP_001121871.1	Desconocida
73	62084 - 63034	Proteína hipotética	Proteína hipotética	316/316 (100%)	WP_001672445.1	Desconocida
74	63266 - 63877	Proteína hipotética	Proteína hipotética	203/203(100%)	WP_021531923.1	Desconocida
75	c63867 - 64148	Proteína hipotética	Proteína hipotética	93/93 (100%)	WP_000856156.1	Desconocida
76	c64208 - 64600	Proteína hipotética	Proteína hipotética	130/130 (100%)	WP_000423941.1	Desconocida
77	64997 - 65356	Proteína hipotética	Proteína hipotética	119/119 (100%)	WP_001672443.1	Desconocida
78	c65556 - 66335	Transposasa	Transposasa	259/259 (100%)	WP_001323403.1	Movilidad del DNA
79	c66335 - 67357	IS100	IS100	340/340 (100%)	WP_000255944.1	Movilidad del DNA
80	c67384 - 67656	Proteína hipotética	Proteína hipotética	73/89 (82%)	WP_047665679.1	Desconocida
81	c67666 - 67872	Proteína hipotética	Proteína hipotética	68/68 (100%)	WP_032308778.1	Desconocida
82	c67966 - 68430	Proteína hipotética	Proteína hipotética	154/154 (100%)	WP_001682053.1	Desconocida
83	c69055 - 71067	DNA topoisomerasa III	DNA topoisomerasa III	668/670 (99%)	WP_032308776.1	Proteína de unión a DNA (predicho)
84	c71090 - 71809	Proteína hipotética	Proteína hipotética	239/239 (100%)	WP_032308774.1	Desconocida
85	c72288 - 73574	Proteína hipotética	Proteína hipotética	428/428 (100%)	WP_001682049.1	Desconocida
86	c73838 - 74065	Proteína hipotética	Proteína hipotética	75/75 (100%)	WP_021564883.1	Desconocida
87	c74062 - 74277	Proteína hipotética	Proteína hipotética	71/71(100%)	WP_001682047.1	Desconocida
88	c74274 - 75011	Proteína hipotética	Proteína hipotética	245/245 (100%)	WP_044707195.1	Desconocida
89	c75008 - 75616	Proteína hipotética	Proteína hipotética	202/202 (100%)	WP_021577937.1	Desconocida
90	c75603 - 75860	Proteína hipotética	Proteína hipotética	84/85 (99%)	WP_000063336.1	Desconocida
91	c75870 - 76118	Proteína hipotética	Proteína hipotética	82/82 (100%)	WP_001682043.1	Desconocida
92	c76115 - 76693	Proteína hipotética	Proteína hipotética	192/192 (100%)	WP_000069531.1	Desconocida
93	c76690 - 77421	Proteína hipotética	Proteína hipotética	243/243 (100%)	WP_032308772.1	Desconocida
94	c77411 - 78820	ParB-like	ParB-like	461/469 (98%)	WP_032308770.1	Desconocida
95	c79022 - 79777	IS408	IS408	250/251 (99%)	KGM76225.1	Movilidad del DNA
96	c79794 - 81329	IS21	IS21	511/511 (100%)	WP_021577119.1	Movilidad del DNA
97	81331 - 81447	Proteína hipotética	Proteína hipotética	38/38 (100%)	CDK82228.1	Desconocida
98	c81545 - 81778	Proteína hipotética	Proteína hipotética	69/73 (95%)	PNB86652.1	Desconocida
99	c81768 - 83321	DnaB-like	DnaB-like	458/462 (99%)	WP_001682040.1	Helicasa (predicho)
100	c83281 - 83499	Proteína hipotética	Proteína hipotética	66/72(92%)	WP_033806496.1	Desconocida
101	c83504 - 84352	Proteína hipotética	Proteína hipotética	275/300 (92%)	WP_001682039.1	Desconocida
102	84579 - 85052	Proteína hipotética	Proteína hipotética	149/158 (94%)	WP_001354423.1	Desconocida
103	c85483 - 86361	ParA	ParA	291/292 (99%)	WP_001682036.1	Desconocida

* c: indica ORFs transcrito en la hebra complementaria

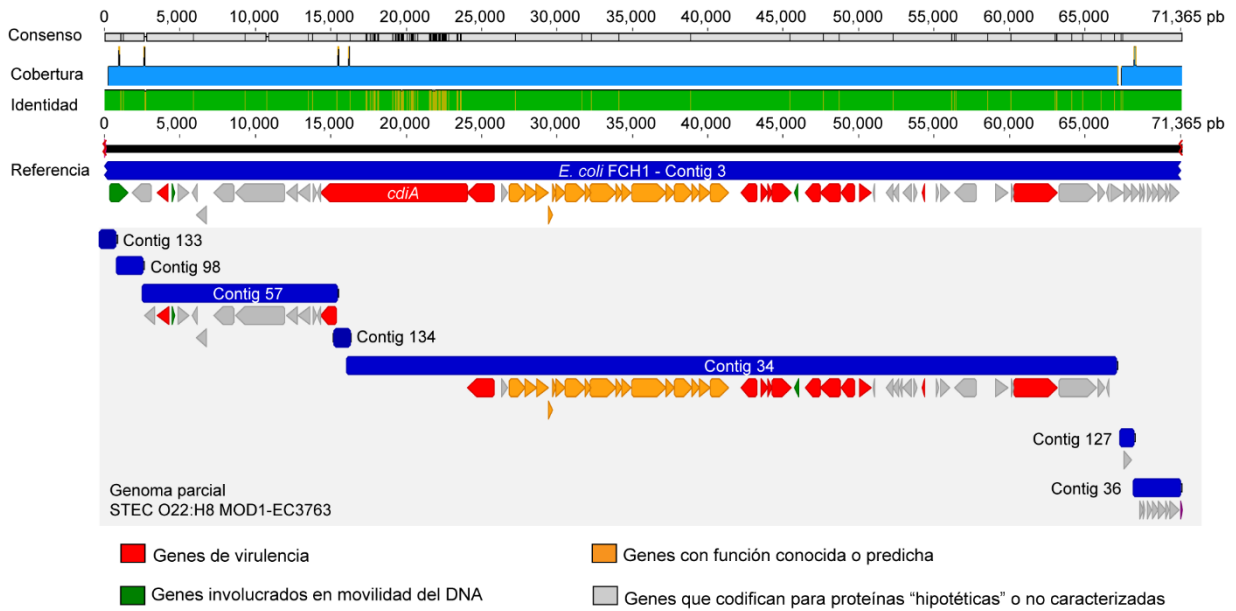


Figura S1. Locus de adhesión y colonización identificado en el genoma parcial de la cepa *E. coli* FCH1 (Número de acceso GenBank: AYKD01000003.1) y STEC O22:H8 MOD1-EC3763 FCH1 (Número de acceso GenBank: NJSE01000000). Los contigs del genoma parcial de la cepa MOD1-EC3763 se mapearon frente al genoma parcial de la cepa FCH1 usando el software Geneious. Como se muestra en la figura, la isla LAC se encontró en 7 contigs concadenados (Contigs 133, 98, 57, 134, 34, 127 y 36) del genoma parcial de la cepa MOD1-EC3763. Los genes predichos y la dirección de la transcripción se representan como flechas de bloque y están codificados por colores según la función del gen, como lo indica la leyenda en la parte inferior.

Concepción, 23 de junio de 2017

INFORME COMITÉ DE BIOÉTICA Y BIOSEGURIDAD

<p>Comité de Ética y Bioseguridad Secretaría de Investigación Facultad de Ciencias Biológicas Universidad de Concepción</p> <ul style="list-style-type: none"> ❖ Dra. Roxana Pincheira B. ❖ Dra. Verónica Madrid V. ❖ Dra. Mafalda Maldonado V. ❖ Sr. Joaquín Rojas F. ❖ Dra. M^a de los Angeles García 	PROYECTO	CONICYT Doctoral
	TÍTULO	"LOCUS DE ADHESIÓN Y AUTOAGREGACIÓN (LAA), UNA ISLA DE PATOGENICIDAD PRESENTE EN CEPAS EMERGENTES DE ESCHERICHIA COLI PRODUCTOR DE SHIGA TOXINA"
	INVESTIGADOR PRINCIPAL	- Sr. David Arturo Montero Forero

A solicitud del Sr. David Arturo Montero Forero, se recibe documento (adjunto) con precisiones a la metodología incluida en la postulación al proyecto CONICYT antes mencionado, solicitando que este Comité emita una carta de aprobación.

Estas precisiones, así como la metodología del Proyecto, han sido analizadas de acuerdo a las recomendaciones de GUIDE for the Care and Use of Laboratory Animals (Institute of Laboratory Animal Resources Commission on Life Science National Research Council), los principios elaborados por el Consejo de Organizaciones Internacionales de las Ciencias Médicas, Organización Mundial de la Salud (1985), el texto "Aspectos Bioéticos de la Experimentación Animal", editado por CONICYT en 2009, y el "Manual de Bioseguridad" editado por CONICYT en 2008.

De acuerdo a lo señalado en el documento, las precisiones a la metodología son consecuentes con el principio de las "3 Rs" empleado para la investigación con animales. En particular:

1. Dado que no existe un modelo alternativo para realizar el estudio, el Investigador necesita emplear animales Raton (*Mus Musculus*) Cepa Balb/c, correspondiendo a especies apropiadas para realizar la investigación, lo que representa una buena aproximación al requisito de REEMPLAZAR.
2. En términos de REDUCIR, se utilizará un número de animales dentro de los rangos admisibles.
3. El uso de analgesia y eutanasia son apropiados para la especie en términos de REFINAR. Además, el Proyecto incluye el uso de un protocolo de supervisión de animales de experimentación.
4. En consecuencia, los protocolos existentes son adecuados.
5. El estudio propuesto permitirá realizar avances en los conocimientos en beneficio humano y animal.
6. El personal involucrado es idóneo en el manejo y uso de las muestras animales.



7. Los estudios serán desarrollados en los LABORATORIO DE INMUNOLOGÍA MOLECULAR, donde se encuentran disponibles tanto las facilidades como el equipamiento necesarios para realizar los estudios con animales incluidos en el Proyecto.

8. Cabe mencionar que nuestra Facultad cuenta con los servicios de la Unidad de Seguridad MATPEL (www2.udec.cl/matpel), quienes colectan y procesan en forma rutinaria los residuos biológicos y químicos generados por nuestros investigadores, de acuerdo a las indicaciones del "Manual de Bioseguridad" de CONICYT.

En consecuencia, el comité de Bioética y Bioseguridad de la Facultad de Ciencias Biológicas de la Universidad de Concepción valida los estudios señalados.

Sin otro particular, les saluda muy atentamente,

Dra. Roxana Pincheira B.
Departamento de Bioquímica Biología Molecular
Facultad de Ciencias Biológicas
Universidad de Concepción

Dra. María de los Angeles García R.
Departamento de Biología Celular
Facultad de Ciencias Biológicas
Universidad de Concepción

Dra. Verónica Madrid V.
Departamento de Microbiología
Facultad de Ciencias Biológicas
Universidad de Concepción

Dra. Mafalda Maldonado Villagrán
Departamento de Fisiopatología
Facultad de Ciencias Biológicas
Universidad de Concepción

Sr. Joaquín Rojas R.
Médico Veterinario
Facultad de Ciencias Biológicas
Universidad de Concepción



PUBLICACIONES ASOCIADAS A ESTA TESIS

Montero DA, Velasco J, Del Canto F, Puente JL, Padola NL, Rasko DA, Farfán M, Salazar JC, Vidal R. Locus of Adhesion and Autoaggregation (LAA), a pathogenicity island present in emerging Shiga Toxin-producing *Escherichia coli* strains. *Sci Rep*. 2017;7: 7011.

Colello R, Vélez MV, González J, **Montero DA**, Bustamante A V, Del Canto F, Etcheverría A, Vidal R, Padola NL. First report of the distribution of Locus of Adhesion and Autoaggregation (LAA) pathogenicity island in LEE-negative Shiga toxin-producing *Escherichia coli* isolates from Argentina. *Microb Pathog*. 2018;123: 259–263.

Montero DA, Del Canto F, Velasco J, Colello R, Padola NL, Salazar JC, San Martin C, Oñate A, Blanco J, Rasko DA, Contreras C, Puente JL, Scheutz F, Franz E, Vidal R. Cumulative acquisition of pathogenicity islands has shaped virulence potential and contributed to the emergence of LEE-negative Shiga toxin-producing *Escherichia coli* strains. *Emerg Microbes Infect*. En revision.

PRESENTACIONES EN CONGRESOS

XXIV CONGRESO LATINOAMERICANO DE MICROBIOLOGÍA. Presentación oral. Contribución de las islas de patogenicidad a la virulencia y emergencia de cepas de *Escherichia coli* productor de shiga toxina LEE-negativo. **Montero DA**. Santiago – Chile; Noviembre de 2018

X INTERNATIONAL SYMPOSIUM ON SHIGA TOXIN (VEROCYTOTOXIN) PRODUCING *Escherichia coli* INFECTIONS. Póster. Curli expression during biofilm formation in Shiga toxin-producing *Escherichia coli* LAA-positive. Velez V., Colello R., **Montero D.**, Etcheverría Al., Vidal R., Padola NL. Florence – Italy; Mayo de 2018.

X INTERNATIONAL SYMPOSIUM ON SHIGA TOXIN (VEROCYTOTOXIN) PRODUCING *Escherichia coli* INFECTIONS. Póster. First report of the novel pathogenicity island of adhesion an autoaggregation (LAA) in Shiga toxin-producing *Escherichia coli* LEE-negative from Argentina. Colello R., Velez V., **Montero D.**, Gonzales J., Bustamante A., Del Canto F., Etcheverría Al., Vidal R., Padola NL. Florence – Italy; Mayo de 2018.

VI CONGRESS OF EUROPEAN MICROBIOLOGISTS. Póster. Identification and distribution of a novel outer membrane antigen related to Shiga Toxin-producing *Escherichia coli* (STEC) strains negative for Locus of Enterocyte Effacement. **Montero DA.**, Del Canto F., Puente JL., Salazar J., Vidal R. Maastricht – Netherlands; Junio de 2015.

XXII CONGRESO LATINOAMERICANO DE MICROBIOLOGÍA y IV CONGRESO COLOMBIANO DE MICROBIOLOGÍA. Póster. Identification and distribution of a novel outer membrane antigen related to Shiga Toxin–producing *Escherichia coli* strains negative for Locus of Enterocyte Effacement. **Montero DA.**, Del Canto F., Salazar J., Vidal R. Cartagena de Indias – Colombia; Noviembre de 2014.

REFERENCIAS

1. UNICEF/WHO. 2009. Diarrhoea: why children are still dying and what can be done.
2. Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, Jha P, Campbell H, Walker CF, Cibulskis R, Eisele T, Liu L, Mathers C. 2010. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* 375:1969–87.
3. Gómez-Duarte OG. 2014. [Acute diarrheal disease caused by enteropathogenic *Escherichia coli* in Colombia]. *Rev Chilena Infectol* 31:577–86.
4. Grimwood K, Forbes D a. 2009. Acute and persistent diarrhea. *Pediatr Clin North Am* 56:1343–61.
5. O’Ryan M, Prado V, Pickering LK. 2005. A millennium update on pediatric diarrheal illness in the developing world. *Semin Pediatr Infect Dis* 16:125–136.
6. Kaper JB, Nataro JP, Mobley HL. 2004. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2:123–40.
7. Nataro JP, Kaper JB. 1998. Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* 11:142–201.
8. Karch H, Denamur E, Dobrindt U, Finlay BB, Hengge R, Johannes L, Ron EZ, Tønsum T, Sansonetti PJ, Vicente M. 2012. The enemy within us: lessons from the 2011 European *Escherichia coli* O104:H4 outbreak. *EMBO Mol Med* 4:841–848.
9. Majowicz SE, Scallan E, Jones-Bitton A, Sargeant JM, Stapleton J, Angulo FJ, Yeung DH, Kirk MD. 2014. Global incidence of human Shiga toxin-producing *Escherichia coli* infections and deaths: A systematic review and knowledge synthesis. *Foodborne Pathog Dis* 11:447–455.
10. Betzen C, Plotnicki K, Fathalizadeh F, Pappan K, Fleming T, Bielaszewska M, Karch H, Tönshoff B, Rafat N. 2016. Shiga toxin 2a-induced endothelial injury in hemolytic uremic syndrome: A metabolomic analysis. *J Infect Dis* 213:1031–1040.
11. Herold S, Paton JC, Paton AW. 2009. Sab, a novel autotransporter of Locus of Enterocyte Effacement-negative Shiga-toxigenic *Escherichia coli* O113:H21, contributes to adherence and biofilm formation. *Infect Immun* 77:3234–3243.
12. Farfan MJ, Torres AG. 2012. Molecular mechanisms that mediate colonization of Shiga toxin-producing *Escherichia coli* strains. *Infect Immun* 80:903–913.
13. Brooks JT, Sowers EG, Wells JG, Greene KD, Griffin PM, Hoekstra RM, Strockbine N a. 2005. Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983-2002. *J Infect Dis* 192:1422–1429.

14. Ethelberg S, Olsen KEP, Scheutz F, Jensen C, Schiellerup P, Engberg J, Petersen AM, Olesen B, Gerner-Smidt P, Mølbak K. 2004. Virulence factors for hemolytic uremic syndrome, Denmark. *Emerg Infect Dis* 10:842–847.
15. Wells JG, Shipman LD, Greene KD, Sowers EG, Green JH, Cameron DN, Downes FP, Martin ML, Griffin PM, Ostroff SM. 1991. Isolation of *Escherichia coli* serotype O157:H7 and other Shiga-like-toxin-producing *E. coli* from dairy cattle. *J Clin Microbiol* 29:985–9.
16. Gyles CL. 2007. Shiga toxin-producing *Escherichia coli*: an overview. *J Anim Sci* 85:E45-62.
17. Griffin PM, Tauxe R V. 1991. The epidemiology of infections caused by *Escherichia coli* O157:H7, other enterohemorrhagic *E. coli*, and the associated hemolytic uremic syndrome. *Epidemiol Rev* 13:60–98.
18. Karmali M a, Gannon V, Sargeant JM. 2010. Verocytotoxin-producing *Escherichia coli* (VTEC). *Vet Microbiol* 140:360–70.
19. Delannoy S, Beutin L, Fach P. 2013. Discrimination of enterohemorrhagic *Escherichia coli* (EHEC) from Non-EHEC strains based on detection of various combinations of Type III effector genes. *J Clin Microbiol* 51:3257–3262.
20. Gould LH, Mody RK, Ong KL, Clogher P, Cronquist AB, Garman KN, Lathrop S, Medus C, Spina NL, Webb TH, White PL, Wymore K, Gierke RE, Mahon BE, Griffin PM. 2013. Increased recognition of non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States during 2000-2010: epidemiologic features and comparison with *E. coli* O157 infections. *Foodborne Pathog Dis* 10:453–60.
21. Soborg B, Lassen SG, Müller L, Jensen T, Ethelberg S, Mølbak K, Scheutz F. 2013. A verocytotoxin-producing *E. coli* outbreak with a surprisingly high risk of haemolytic uraemic syndrome, Denmark, September-October 2012. *Euro Surveill* 18:4–6.
22. Instituto de Salud Pública de Chile. 2014. Vigilancia de laboratorio de *E. coli* productora de toxina Shiga. Chile, 2007 - 2013. Boletín Instituto de Salud Pública de Chile.
23. Ferdous M, Friedrich AW, Grundmann H, de Boer RF, Croughs PD, Islam MA, Kluytmans-van den Bergh MFQ, Kooistra-Smid AMD, Rossen JWA. 2016. Molecular characterization and phylogeny of Shiga toxin–producing *Escherichia coli* isolates obtained from two Dutch regions using whole genome sequencing. *Clin Microbiol Infect* 22:642.e1-642.e9.
24. Franz E, van Hoek AH a. M, Wuite M, van der Wal FJ, de Boer AG, Bouw E, Aarts HJM. 2015. Molecular hazard identification of non-O157 Shiga toxin-producing *Escherichia coli* (STEC). *PLoS One* 10:e0120353.
25. Fruth A, Prager R, Tietze E, Rabsch W, Flieger A. 2015. Molecular epidemiological view on Shiga toxin-producing *Escherichia coli* causing human

- disease in Germany: Diversity, prevalence, and outbreaks. *Int J Med Microbiol* 305:697–704.
26. Cundon C, Marey E, Roldan F, Canosa C, Navarro A, Gadea P, Blanco X, Babich J, Rocchi D, Kieernicki M, Binotti S, Calzetta A, Betancour A. 2015. Preliminary detection and characterization of *Escherichia coli* O174 shiga toxin-producing. *Senasa* 8:52–63.
 27. Lee JB, Han D, Lee HT, Wi SM, Park JH, Jo J, Cho Y, Hahn T, Lee S, Kang B, Kwak HS, Kim J, Yoon JW. 2018. Pathogenic and phylogenetic characteristics of non-O157 Shiga toxin-producing *Escherichia coli* isolates from retail meats in South Korea. *J Vet Sci* 19:251.
 28. Messens W, Bolton D, Frankel G, Liebana E, McLauchlin J, Morabito S, Oswald E, Threlfall EJ. 2015. Defining pathogenic verocytotoxin-producing *Escherichia coli* (VTEC) from cases of human infection in the European Union, 2007-2010. *Epidemiol Infect* 143:1652–1661.
 29. Rauw K De, Jacobs S, Pie D. 2018. Twenty-seven years of screening for Shiga toxin-producing *Escherichia coli* in a university hospital. Brussels, Belgium, 1987-2014. *PLoS One* 1–15.
 30. Nüesch-Inderbinnen M, Morach M, Cernela N, Althaus D, Jost M, Mäusezahl M, Bloomberg G, Stephan R. 2018. Serotypes and virulence profiles of Shiga toxin-producing *Escherichia coli* strains isolated during 2017 from human infections in Switzerland. *Int J Med Microbiol* 0–1.
 31. Galli L, Miliwebsky E, Irino K, Leotta G, Rivas M. 2010. Virulence profile comparison between LEE-negative Shiga toxin-producing *Escherichia coli* (STEC) strains isolated from cattle and humans. *Vet Microbiol* 143:307–313.
 32. EFSA (European Food Safety Authority) and ECDC (European Centre for Disease Prevention and Control). 2014. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2012. *EFSA J* 12.
 33. Bettelheim K a. 2007. The non-O157 shiga-toxigenic (verocytotoxigenic) *Escherichia coli*; under-rated pathogens. *Crit Rev Microbiol* 33:67–87.
 34. Doughty S, Sloan J, Bennett-Wood V, Robertson M, Robins-Browne RM, Hartland EL. 2002. Identification of a novel fimbrial gene cluster related to long polar fimbriae in locus of enterocyte effacement-negative strains of enterohemorrhagic *Escherichia coli*. *Infect Immun* 70:6761–6769.
 35. Montero D. 2013. Identificación de proteínas inmunogénicas de membrana externa de *Escherichia coli* productor de shiga toxina (STEC). Universidad de Chile.
 36. Li Z, Elliott E, Payne J, Isaacs J, Gunning P, O'loughlin E V. 1999. Shiga toxin-producing *Escherichia coli* can impair T84 cell structure and function without inducing attaching/effacing lesions. *Infect Immun* 67:5938–45.

37. Phillips AD, Navabpour S, Hicks S, Dougan G, Wallis T, Frankel G. 2000. Enterohaemorrhagic *Escherichia coli* O157:H7 target Peyer's patches in humans and cause attaching/effacing lesions in both human and bovine intestine. *Gut* 47:377–381.
38. Jacewicz M, Clausen H, Nudelman E, Donohue-Rolfe A, Keusch GT. 1986. Pathogenesis of *Shigella* diarrhea. XI. Isolation of a shigella toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. *J Exp Med* 163:1391–404.
39. Sandvig K, van Deurs B. 1994. Endocytosis and intracellular sorting of ricin and Shiga toxin. *FEBS Lett* 346:99–102.
40. Saxena SK, O'Brien AD, Ackerman EJ. 1989. Shiga toxin, Shiga-like toxin II variant, and ricin are all single-site RNA N-glycosidases of 28 S RNA when microinjected into *Xenopus* oocytes. *J Biol Chem* 264:596–601.
41. Proulx F, Seidman EG, Karpman D. 2001. Pathogenesis of Shiga toxin-associated hemolytic uremic syndrome. *Pediatr Res* 50:163–71.
42. Brigotti M. 2012. The interactions of human neutrophils with shiga toxins and related plant toxins: danger or safety? *Toxins (Basel)* 4:157–90.
43. Prado J V, Cavagnaro S M F. 2008. Hemolytic uremic syndrome associated to shigatoxin producing *Escherichia coli* in Chilean children: clinical and epidemiological aspects. *Rev Chilena Infectol* 25:435–44.
44. Gasser C, Gautier E, Steck A, Siebenmann RE, Oechslin R. 1955. [Hemolytic-uremic syndrome: bilateral necrosis of the renal cortex in acute acquired hemolytic anemia]. *Schweiz Med Wochenschr* 85:905–909.
45. Scheutz F, Teel LD, Beutin L, Pierard D, Buvens G, Karch H, Mellmann A, Caprioli A, Tozzoli R, Morabito S, Strockbine NA, Melton-Celsa AR, Sanchez M, Persson S, O'Brien AD. 2012. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J Clin Microbiol* 50:2951–2963.
46. Food and Agriculture Organization of the United Nations (FAO) and World Health Organization (WHO). 2018. Shiga toxin-producing *Escherichia coli* (STEC) and food: attribution, characterization, and monitoring. Roma.
47. Melton-Celsa AR. 2014. Shiga Toxin (Stx) classification, structure, and function. *Microbiol Spectr* 2:1–13.
48. Boerlin P, McEwen SA, Boerlin-Petzold F, Wilson JB, Johnson RP, Gyles CL. 1999. Associations between virulence factors of Shiga toxin-producing *Escherichia coli* and disease in humans. *J Clin Microbiol* 37:497–503.
49. Nishikawa Y, Zhou Z, Hase A, Ogasawara J, Cheasty T, Haruki K. 2000. Relationship of genetic type of Shiga toxin to manifestation of bloody diarrhea due to enterohemorrhagic *Escherichia coli* serogroup O157 isolates in Osaka City, Japan. *J Clin Microbiol* 38:2440–2.

50. Lorenz SC, Son I, Maounounen-Laasri A, Lin A, Fischer M, Kase JA. 2013. Prevalence of hemolysin genes and comparison of ehxA subtype patterns in Shiga toxin-producing *Escherichia coli* (STEC) and Non-STEC strains from clinical, food, and animal sources. *Appl Environ Microbiol* 79:6301–6311.
51. Taneike I, Zhang HM, Wakisaka-Saito N, Yamamoto T. 2002. Enterohemolysin operon of Shiga toxin-producing *Escherichia coli*: a virulence function of inflammatory cytokine production from human monocytes. *FEBS Lett* 524:219–224.
52. Law D, Kelly J. 1995. Use of heme and hemoglobin by *Escherichia coli* O157 and other Shiga-like-toxin-producing *E. coli* serogroups. *Infect Immun* 63:700–702.
53. Bielaszewska M, Sinha B, Kuczylus T, Karch H, Kuczius T. 2005. Cytolethal distending toxin from Shiga toxin-producing *Escherichia coli* O157 causes irreversible G2/M arrest, inhibition of proliferation, and death of human endothelial cells. *Infect Immun* 73:552–562.
54. Orth D, Grif K, Dierich MP, Würzner R. 2006. Cytolethal distending toxins in Shiga toxin-producing *Escherichia coli*: Alleles, serotype distribution and biological effects. *J Med Microbiol* 55:1487–1492.
55. Smith JL, Bayles DO. 2006. The contribution of cytolethal distending toxin to bacterial pathogenesis. *Crit Rev Microbiol* 32:227–248.
56. Allué-Guardia A, García-Aljaro C, Muniesa M. 2011. Bacteriophage-encoding cytolethal distending toxin type V gene induced from nonclinical *Escherichia coli* isolates. *Infect Immun* 79:3262–3272.
57. Bielaszewska M, Stoewe F, Fruth A, Zhang W, Prager R, Brockmeyer J, Mellmann A, Karch H, Friedrich AW. 2009. Shiga toxin, cytolethal distending toxin, and hemolysin repertoires in clinical *Escherichia coli* O91 isolates. *J Clin Microbiol* 47:2061–2066.
58. Paton AW, Srimanote P, Talbot UM, Wang H, Paton JC. 2004. A new family of potent AB(5) cytotoxins produced by Shiga toxigenic *Escherichia coli*. *J Exp Med* 200:35–46.
59. Michelacci V, Tozzoli R, Caprioli A, Martínez R, Scheutz F, Grande L, Sánchez S, Morabito S. 2013. A new pathogenicity island carrying an allelic variant of the Subtilase cytotoxin is common among Shiga toxin producing *Escherichia coli* of human and ovine origin. *Clin Microbiol Infect* 19:E149–E156.
60. Yahiro K, Nagasawa S, Ichimura K, Takeuchi H, Ogura K, Tsutsuki H, Shimizu T, Iyoda S, Ohnishi M, Iwase H, Moss J, Noda M. 2018. Mechanism of inhibition of Shiga-toxigenic *Escherichia coli* SubAB cytotoxicity by steroids and diacylglycerol analogues. *Cell Death Discov*.
61. Funk J, Stoeber H, Hauser E, Schmidt H. 2013. Molecular analysis of subtilase cytotoxin genes of food-borne Shiga toxin-producing *Escherichia coli* reveals a new allelic subAB variant. *BMC Microbiol* 13:230.

62. Nüesch-Inderbilen MT, Funk J, Cernela N, Tasara T, Klumpp J, Schmidt H, Stephan R. 2015. Prevalence of subtilase cytotoxin-encoding subAB variants among Shiga toxin-producing *Escherichia coli* strains isolated from wild ruminants and sheep differs from that of cattle and pigs and is predominated by the new allelic variant subAB2-2. *Int J Med Microbiol* 305:124–128.
63. Jerse a E, Yu J, Tall BD, Kaper JB. 1990. A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. *Proc Natl Acad Sci U S A* 87:7839–43.
64. Moon HW, Whipp SC, Argenzio R a, Levine MM, Giannella R a. 1983. Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. *Infect Immun* 41:1340–51.
65. Paton AW, Srimanote P, Woodrow MC, Paton JC. 2001. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. *Infect Immun* 69:6999–7009.
66. Leyton DL, Sloan J, Hill RE, Doughty S, Hartland EL. 2003. Transfer region of pO113 from enterohemorrhagic *Escherichia coli*: similarity with R64 and identification of a novel plasmid-encoded autotransporter, EpeA. *Infect Immun* 71:6307–6319.
67. Lu Y, Iyoda S, Satou H, Satou H, Itoh K, Saitoh T, Watanabe H. 2006. A new immunoglobulin-binding protein, EibG, is responsible for the chain-like adhesion phenotype of locus of enterocyte effacement-negative, shiga toxin-producing *Escherichia coli*. *Infect Immun* 74:5747–5755.
68. Tarr PI, Bilge SS, Vary JC, Jelacic S, Habeeb RL, Ward TR, Baylor MR, Besser TE. 2000. Iha: a novel *Escherichia coli* O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. *Infect Immun* 68:1400–1407.
69. Fleckenstein JM, Kopecko DJ, Warren RL, Elsinghorst EA. 1996. Molecular characterization of the tia invasion locus from enterotoxigenic *Escherichia coli*. *Infect Immun* 64:2256–65.
70. van der Woude MW, Henderson IR. 2008. Regulation and function of Ag43 (flu). *Annu Rev Microbiol* 62:153–69.
71. Brunder W, Schmidt H, Karch H. 1997. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol Microbiol* 24:767–778.
72. Schmidt H, Zhang W-L, Hemmrich U, Jelacic S, Brunder W, Tarr PI, Dobrindt U, Hacker J, Karch H. 2001. Identification and characterization of a novel genomic island integrated at selC in Locus of Enterocyte Effacement-negative, Shiga toxin-producing *Escherichia coli*. *Infect Immun* 69:6863–6873.
73. Karch H, Schubert S, Zhang D, Zhang W, Schmidt H, Ölschläger T, Hacker J.

1999. A genomic island, termed high-pathogenicity island, is present in certain non-O157 Shiga toxin-producing *Escherichia coli* clonal lineages. *Infect Immun* 67:5994–6001.
74. Friedrich AW, Borell J, Bielaszewska M, Fruth A, Tschäpe H, Karch H. 2003. Shiga toxin 1c-producing *Escherichia coli* strains: phenotypic and genetic characterization and association with human disease. *J Clin Microbiol* 41:2448–53.
 75. Bondi R, Chiani P, Michelacci V, Minelli F, Caprioli A, Morabito S. 2017. The gene *tia*, harbored by the Subtilase-Encoding Pathogenicity Island, Is Involved in the ability of Locus of Enterocyte Effacement-negative Shiga toxin-producing *Escherichia coli* strains to invade monolayers of epithelial cells. *Infect Immun* 85:e00613-17.
 76. Schubert S, Rakin A, Heesemann J. 2004. The *Yersinia* high-pathogenicity island (HPI): Evolutionary and functional aspects. *Int J Med Microbiol* 294:83–94.
 77. Schubert S, Rakin A, Karch H, Carniel E, Heesemann J. 1998. Prevalence of the “high-pathogenicity island” of *Yersinia* species among *Escherichia coli* strains that are pathogenic to humans. *Infect Immun* 66:480–5.
 78. Montero D, Orellana P, Gutiérrez D, Araya D, Salazar JC, Prado V, Oñate A, Del Canto F, Vidal R. 2014. Immunoproteomic analysis to identify Shiga toxin-producing *Escherichia coli* uuter membrane proteins expressed during human infection. *Infect Immun* 82:4767–77.
 79. Mancini J, Weckselblatt B, Chung YK, Durante JC, Andelman S, Glaubman J, Dorff JD, Bhargava S, Lijek RS, Unger KP, Okeke IN. 2011. The Heat-resistant agglutinin family includes a novel adhesin from enteroaggregative *Escherichia coli* strain 60A. *J Bacteriol* 193:4813–4820.
 80. Coulouris Y, Zaretskaya I, Cutcutache I, Rozen S, Madden T. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13:134.
 81. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–77.
 82. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–5.
 83. Datsenko K a, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97:6640–5.
 84. Cherepanov PP, Wackernagel W. 1995. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-

resistance determinant. *Gene* 158:9–14.

85. Girardeau JP, Bertin Y, Martin C. 2009. Genomic analysis of the PAI ICL3 locus in pathogenic LEE-negative Shiga toxin-producing *Escherichia coli* and *Citrobacter rodentium*. *Microbiology* 155:1016–1027.
86. Gutiérrez D, Pardo M, Montero D, Oñate A, Farfán MJ, Ruiz-Pérez F, Del Canto F, Vidal R. 2015. TleA, a Tsh-Like autotransporter identified in a human enterotoxigenic *Escherichia coli* strain. *Infect Immun* 83:1893–1903.
87. Bhargava S, Johnson BB, Hwang J, Harris T a, George AS, Muir A, Dorff J, Okeke IN. 2009. Heat-resistant agglutinin 1 is an accessory enteroaggregative *Escherichia coli* colonization factor. *J Bacteriol* 191:4934–4942.
88. Rivas L, Fegan N, Dykes G a. 2008. Expression and putative roles in attachment of outer membrane proteins of *Escherichia coli* O157 from planktonic and sessile culture. *Foodborne Pathog Dis* 5:155–164.
89. Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri GM, Carnemolla B, Orecchia P, Zardi L, Righetti PG. 2004. Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* 25:1327–33.
90. Lindgren SW, Melton AR, O'Brien AD. 1993. Virulence of enterohemorrhagic *Escherichia coli* O91:H21 clinical isolates in an orally infected mouse model. *Infect Immun* 61:3832–42.
91. Joensen KG, Tetzschner AMM, Iguchi A, Aarestrup FM, Scheutz F. 2015. Rapid and easy in silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. *J Clin Microbiol* 53:2410–2426.
92. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *J Clin Microbiol* 52:1501–1510.
93. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MCJ, Ochman H, Achtman M. 2006. Sex and virulence in *Escherichia coli*: An evolutionary perspective. *Mol Microbiol* 60:1136–1151.
94. Larsen M V., Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T, Ussery DW, Aarestrup FM, Lund O. 2012. Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol* 50:1355–1361.
95. Beghain J, Bridier-Nahmias A, Le Nagard H, Denamur E, Clermont O. 2018. ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. *Microb Genomics* 1–8.
96. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147.
97. Kurtz S, Choudhuri J V, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R.

2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res* 29:4633–4642.
98. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 34:D32-6.
 99. Lowe TM, Eddy SR. 1996. TRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964.
 100. Hudson CM, Lau BY, Williams KP. 2015. Islander: a database of precisely mapped genomic islands in tRNA and tmRNA genes. *Nucleic Acids Res* 43:D48-53.
 101. Yoon SH, Park YK, Kim JF. 2015. PAIDB v2.0: Exploration and analysis of pathogenicity and resistance islands. *Nucleic Acids Res* 43:D624–D630.
 102. Bi D, Xu Z, Harrison EM, Tai C, Wei Y, He X, Jia S, Deng Z, Rajakumar K, Ou HY. 2012. ICEberg: A web-based resource for integrative and conjugative elements found in Bacteria. *Nucleic Acids Res* 40:621–626.
 103. Aziz RK, Bartels D, Best A, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* 9:1–15.
 104. Sullivan MJ, Petty NK, Beatson S a. 2011. Easyfig: A genome comparison visualizer. *Bioinformatics* 27:1009–1010.
 105. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B. 2016. R Package 'gplots'.
 106. R Core Team. 2014. R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing. Vienna, Austria.
 107. Csardi G. 2015. R Package 'igraph'.
 108. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. 2014. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS One* 9:e104984.
 109. Gardner SN, Slezak T, Hall BG. 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 31:2877–2878.
 110. Letunic I, Bork P. 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* 39:W475–W478.
 111. Fagan RP, Smith SGJ. 2007. The Hek outer membrane protein of *Escherichia coli* is an auto-aggregating adhesin and invasin. *FEMS Microbiol Lett* 269:248–255.

112. Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
113. Okonechnikov K, Golosova O, Fursov M, Varlamov A, Vaskin Y, Efremov I, German Grehov OG, Kandrov D, Rasputin K, Syabro M, Tleukenov T. 2012. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* 28:1166–1167.
114. Omasits U, Ahrens CH, Müller S, Wollscheid B. 2014. Protter: Interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 30:884–886.
115. Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, Hönigschmid P, Schafferhans A, Roos M, Bernhofer M, Richter L, Ashkenazy H, Punta M, Schlessinger A, Bromberg Y, Schneider R, Vriend G, Sander C, Ben-Tal N, Rost B. 2014. PredictProtein-an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res* 42:W337-43.
116. Kelly LA, Mezulis S, Yates C, Wass M, Sternberg M. 2015. The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat Protoc* 10:845–858.
117. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
118. Reidl S, Lehmann A, Schiller R, Salam Khan a, Dobrindt U. 2009. Impact of O-glycosylation on the molecular and cellular adhesion properties of the *Escherichia coli* autotransporter protein Ag43. *Int J Med Microbiol* 299:389–401.
119. Lloyd AL, Smith SN, Eaton K a, Mobley HLT. 2009. Uropathogenic *Escherichia coli* suppresses the host inflammatory response via pathogenicity island genes *sisA* and *sisB*. *Infect Immun* 77:5322–5333.
120. Ruan L, Pleitner A, Gänzle MG, McMullen LM. 2011. Solute transport proteins and the outer membrane protein NmpC contribute to heat resistance of *Escherichia coli* AW1.7. *Appl Environ Microbiol* 77:2961–2967.
121. Nishio M, Okada N, Miki T, Haneda T, Danbara H. 2005. Identification of the outer-membrane protein PagC required for the serum resistance phenotype in *Salmonella enterica* serovar Choleraesuis. *Microbiology* 151:863–873.
122. Jacob-Dubuisson F, Guérin J, Baelen S, Clantin B. 2013. Two-partner secretion: as simple as it sounds? *Res Microbiol* 164:583–595.
123. Hacker J, Kaper JB. 2000. Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* 54:641–679.
124. Lioliou EE, Mimitou EP, Grigoroudis AI, Panagiotidis CH, Panagiotidis C a, Kyriakidis D a. 2005. Phosphorylation activity of the response regulator of the two-component signal transduction system AtoS–AtoC in *E. coli*. *Biochim Biophys Acta - Gen Subj* 1725:257–268.
125. Shen S, Mascarenhas M, Rahn K, Kaper JB, Karmal MA. 2004. Evidence for a

- hybrid genomic island in Verocytotoxin-producing *Escherichia coli* CL3 (Serotype O113:H21) containing segments of EDL933 (Serotype O157:H7) O islands 122 and 48. *Infect Immun* 72:1496–1503.
126. Lalioui L, Le Bouguéne C. 2001. afa-8 Gene cluster is carried by a pathogenicity island inserted into the tRNA(Phe) of human and bovine pathogenic *Escherichia coli* isolates. *Infect Immun* 69:937–48.
 127. Colello R, Vélez MV, González J, Montero DA, Bustamante A V, Canto F Del, Etcheverría AI, Vidal R, Padola NL. 2018. First report of the distribution of Locus of Adhesion and Autoaggregation (LAA) pathogenicity island in LEE-negative Shiga toxin-producing *Escherichia coli* isolates from Argentina. *Microb Pathog* 123:259–263.
 128. Yan X, Fratamico PM, Bono JL, Baranzoni GM, Chen C-Y. 2015. Genome sequencing and comparative genomics provides insights on the evolutionary dynamics and pathogenic potential of different H-serotypes of Shiga toxin-producing *Escherichia coli* O104. *BMC Microbiol* 15:83.
 129. Zhang W, Nadirk J, Kossow A, Bielaszewska M, Leopold SR, Witten A, Fruth A, Karch H, Ammon A, Mellmann A. 2014. Phylogeny and phenotypes of clinical and environmental Shiga toxin-producing *Escherichia coli* O174. *Environ Microbiol* 16:963–76.
 130. Mora A, Blanco M, Blanco JE, Dahbi G, López C, Justel P, Alonso MP, Echeita A, Bernárdez MI, González EA, Blanco J. 2007. Serotypes, virulence genes and intimin types of Shiga toxin (verocytotoxin)-producing *Escherichia coli* isolates from minced beef in Lugo (Spain) from 1995 through 2003. *BMC Microbiol* 7:13.
 131. Baliere C, Rince A, Delannoy S, Fach P, Gourmelon M. 2016. Molecular profiling of Shiga toxin-producing *Escherichia coli* and Enteropathogenic *E. coli* strains isolated from French coastal environments. *Appl Environ Microbiol* 82:3913–3927.
 132. Budič M, Rijavec M, Petkovšek Ž, Žgur-Bertok D. 2011. *Escherichia coli* bacteriocins: Antimicrobial efficacy and prevalence among isolates from patients with bacteraemia. *PLoS One* 6:e28769.
 133. Ruhe ZC, Low DA, Hayes CS. 2013. Bacterial contact-dependent growth inhibition. *Trends Microbiol* 21:230–237.
 134. de Been M, Lanza VF, de Toro M, Scharringa J, Dohmen W, Du Y, Hu J, Lei Y, Li N, Tooming-Klunderud A, Heederik DJJ, Fluit AC, Bonten MJM, Willems RJL, de la Cruz F, van Schaik W. 2014. Dissemination of cephalosporin resistance genes between *Escherichia coli* strains from farm animals and humans by specific plasmid lineages. *PLoS Genet* 10:e1004776.
 135. Martínez-Jéhanne V, Pichon C, du Merle L, Poupel O, Cayet N, Bouchier C, Le Bouguéne C. 2012. Role of the vpe carbohydrate permease in *Escherichia coli* urovirulence and fitness in vivo. *Infect Immun* 80:2655–2666.

136. Martinez-Jéhanne V, Merle L Du, Bernier-Fébreau C, Usein C, Gassama-Sow A, Wane AA, Gouali M, Damian M, Aidara-Kane A, Germani Y, Fontanet A, Coddeville B, Guérardel Y, Bouguéneq C Le. 2009. Role of deoxyribose catabolism in colonization of the murine intestine by pathogenic *Escherichia coli* strains. *Infect Immun* 77:1442–1450.
137. Perna NT, Glasner JD, Burland V, Plunkett G. 2002. The genomes of *Escherichia coli* K-12 and pathogenic *E. coli*, p. 3–53. *In* Michel S. Donnenberg (ed.), *Escherichia Coli*. Elsevier.
138. Al-Hasani K, Henderson IR, Sakellaris H, Rajakumar K, Grant T, Nataro JP, Robins-Browne R, Adler B. 2000. The *sigA* gene which is borne on the she pathogenicity island of *Shigella flexneri* 2a encodes an exported cytopathic protease involved in intestinal fluid accumulation. *Infect Immun* 68:2457–2463.
139. Munera D, Ritchie JM, Hatzios SK, Bronson R, Fang G, Schadt EE, Davis BM, Waldor MK. 2014. Autotransporters but not pAA are critical for rabbit colonization by Shiga toxin-producing *Escherichia coli* O104:H4. *Nat Commun* 5:3080.
140. Johnson CM, Grossman AD. 2015. Integrative and Conjugative Elements (ICEs): What they do and how they work. *Annu Rev Genet* 49:577–601.
141. Burrus V, Pavlovic G, Decaris B, Guédon G. 2002. Conjugative transposons: The tip of the iceberg. *Mol Microbiol* 46:601–610.
142. Manning SD, Motiwala AS, Springman AC, Qi W, Lacher DW, Ouellette LM, Mladonicky JM, Somsel P, Rudrik JT, Dietrich SE, Zhang W, Swaminathan B, Alland D, Whittam TS. 2008. Variation in virulence among clades of *Escherichia coli* O157:H7 associated with disease outbreaks. *Proc Natl Acad Sci* 105:4868–4873.
143. Ogura Y, Gotoh Y, Itoh T, Sato MP, Seto K, Yoshino S, Isobe J, Etoh Y, Kurogi M, Kimata K, Maeda E, Piérard D, Kusumoto M, Akiba M, Tominaga K, Kirino Y, Kato Y, Shirahige K, Ooka T, Ishijima N, Lee K, Iyoda S, Mainil JG, Hayashi T. 2017. Population structure of *Escherichia coli* O26 : H11 with recent and repeated *stx2* acquisition in multiple lineages. *Microb genomics* 3.
144. Mellor GE, Sim EM, Barlow RS, D’Astek BA, Galli L, Chinen I, Rivas M, Gobius KS. 2012. Phylogenetically related Argentinean and Australian *Escherichia coli* O157 isolates are distinguished by virulence clades and alternative Shiga toxin 1 and 2 prophages. *Appl Environ Microbiol* 78:4724–31.
145. Coombes BK, Wickham ME, Mascarenhas M, Gruenheid S, Finlay BB, Karmali M a. 2008. Molecular analysis as an aid to assess the public health risk of non-O157 shiga toxin-producing *Escherichia coli* strains. *Appl Environ Microbiol* 74:2153–2160.
146. Ulett GC, Webb RI, Schembri MA. 2006. Antigen-43-mediated autoaggregation impairs motility in *Escherichia coli*. *Microbiology* 152:2101–2110.
147. Wu X Bin, Tian LH, Zou HJ, Wang CY, Yu ZQ, Tang CH, Zhao FK, Pan JY. 2013.

- Outer membrane protein OmpW of *Escherichia coli* is required for resistance to phagocytosis. *Res Microbiol* 164:848–855.
148. Beketskaia MS, Bay DC, Turner RJ. 2014. Outer membrane protein OmpW participates with small multidrug resistance protein member EmrE in quaternary cationic compound efflux. *J Bacteriol* 196:1908–1914.
 149. Melton-Celsa AR, Rogers JE, Schmitt CK, Darnell SC, O'Brien AD. 1998. Virulence of Shiga toxin-producing *Escherichia coli* (STEC) in orally-infected mice correlates with the type of toxin produced by the infecting strain. *Jpn J Med Sci Biol* 51 Suppl:S108-14.
 150. Franz E, Delaquis P, Morabito S, Beutin L, Gobius K, Rasko D a., Bono J, French N, Osek J, Lindstedt B-A, Muniesa M, Manning S, LeJeune J, Callaway T, Beatson S, Eppinger M, Dallman T, Forbes KJ, Aarts H, Pearl DL, Gannon VPJ, Laing CR, Strachan NJC. 2014. Exploiting the explosion of information associated with whole genome sequencing to tackle Shiga toxin-producing *Escherichia coli* (STEC) in global food production systems. *Int J Food Microbiol* 187:57–72.
 151. Scheutz F. 2014. Taxonomy meets public health: The case of Shiga toxin-producing *Escherichia coli*. *Microbiol Spectr* 2:1–15.
 152. EFSA Panel of Biological Hazards. 2013. Scientific Opinion on VTEC-serovars and scientific criteria regarding pathogenicity assessment. *J Efsa* 11:3138.
 153. Beutin L, Miko A, Krause G, Pries K, Haby S, Steege K, Albrecht N. 2007. Identification of human-pathogenic strains of shiga toxin-producing *Escherichia coli* from food by a combination of serotyping and molecular typing of Shiga toxin genes. *Appl Environ Microbiol* 73:4769–4775.
 154. Hazen TH, Michalski J, Luo Q, Shetty AC, Daugherty SC, Fleckenstein JM, Rasko DA. 2017. Comparative genomics and transcriptomics of *Escherichia coli* isolates carrying virulence factors of both enteropathogenic and enterotoxigenic *E. coli*. *Sci Rep* 7:3513.
 155. Baker KS, Dallman TJ, Thomson NR, Jenkins C. 2018. An outbreak of a rare Shiga-toxin-producing *Escherichia coli* serotype (O117:H7) among men who have sex with men. *Microb genomics* 1–9.
 156. Naylor SW, Roe AJ, Nart P, Spears K, Smith DGE, Low JC, Gally DL. 2005. *Escherichia coli* O157:H7 forms attaching and effacing lesions at the terminal rectum of cattle and colonization requires the LEE4 operon. *Microbiology* 151:2773–2781.
 157. Reid SD, Herbelin CJ, Bumbaugh a C, Selander RK, Whittam TS. 2000. Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* 406:64–7.
 158. Benjelloun-Touimi Z, Sansonetti PJ, Parsot C. 1995. SepA, the major extracellular protein of *Shigella flexneri*: autonomous secretion and involvement

in tissue invasion. *Mol Microbiol* 17:123–135.

159. Henderson IR, Czeczulin J, Eslava C, Noriega F, Nataro JP. 1999. Characterization of Pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect Immun* 67:5587–5596.
160. Patel SK, Dotson J, Allen KP, Fleckenstein JM. 2004. Identification and molecular characterization of EatA, an autotransporter protein of Enterotoxigenic *Escherichia coli*. *Infect Immun* 72:1786–1794.
161. Stathopoulos C, Provence DL, Curtiss R. 1999. Characterization of the avian pathogenic *Escherichia coli* hemagglutinin Tsh, a member of the immunoglobulin a protease-type family of autotransporters. *Infect Immun* 67:772–781.