



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOMÉDICAS
INSTITUTO DE FISIOLÓGÍA CELULAR
HOSPITAL INFANTIL DE MÉXICO FEDERICO GÓMEZ

Biogénesis y regulación de *longus*: una adhesina de *Escherichia coli* enterotoxigénica

T E S I S

**QUE PARA OPTAR POR EL GRADO DE
DOCTOR EN CIENCIAS BIOMÉDICAS**

**PRESENTA:
ZEUS SALDAÑA AHUACTZI**

**TUTOR PRINCIPAL:
DR. JUAN XICOHTENCATL CORTES
(INSTITUTO DE FISIOLÓGÍA CELULAR, UNAM)**

**MIEMBROS DELCOMITÉ TUTOR:
DRA. BERTHA GONZÁLEZ PEDRAJO
(INSTITUTO DE FISIOLÓGÍA CELULAR, UNAM)
DR. CARLOS ALBERTO ESLAVA CAMPOS
(FACULTAD DE MEDICINA, UNAM)**

CIUDAD UNIVERSITARIA, CD. MX.

Enero, 2017



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



HOSPITAL INFANTIL *de* MÉXICO
FEDERICO GÓMEZ

Instituto Nacional de Salud

LUGAR DE REALIZACIÓN

El presente trabajo se realizó en el Laboratorio de Investigación en Bacteriología Intestinal del Hospital Infantil de México Federico Gómez (HIMFG) bajo la asesoría del Dr. Juan Xicohtencatl Cortes.

FINANCIAMIENTOS Y APOYOS ECONÓMICOS

Este trabajo formo parte del proyecto “Caracterización in vivo de *Longus* una adhesina de *Escherichia coli* enterotoxigénica”, financiado por el consejo Nacional de Ciencia y Tecnología (CONACYT) con número de registro 133451 y fondos federales HIM/2014/016SSA1110.



Agradecimiento especial a CONACYT por la beca otorgada con número de registro 275780 para la realización de este proyecto. Al Posgrado en Ciencias Biomédicas de la Universidad Nacional Autónoma de México (UNAM), por ser el pilar en mi desarrollo profesional y apoyarme para poder asistir a congresos nacionales y por el estímulo otorgado como parte de la mejora a la tasa de graduación.



AGRADECIMIENTOS

- Al Dr. Juan Xicohtencatl Cortes por darme la oportunidad de formarme en su grupo de trabajo, por su orientación y ayuda brindada para la realización de esta tesis, por sus consejos y amistad que me permitieron aprender más que lo estudiado en el proyecto.
- A mi comité tutorial conformado por la Dra. Bertha González Pedrajo y el Dr. Carlos Alberto Eslava Campos por sus asesorías, consejos y por ser más que mi comité tutorial y llegar a ser ejemplos profesionales.
- Al jurado del examen de grado integrado por: Dr. Dimitris Georgellis, Dra. Rosa Laura Camarena, Dr. Juan Pablo Reyes Grajeda y al Dr. Miguel Ángel Cevallos, por dedicar tiempo para la revisión y crítica de este escrito.
- Amigos, compañeros y miembros del laboratorio de Bacteriología Intestinal del HIMFG, gracias por todo.
- Al Dr. Edgar Oliver López Villegas y la Dra. María del Rosario Espinosa Mellado de la Central de Instrumentación de Microscopía de la Escuela Nacional de Ciencias Biológicas del Instituto Politécnico Nacional, por su apoyo técnico en la visualización de las muestras de microscopía electrónica.
- Al MVZ Héctor Alfonso Malagón Rivero del bioterio del Instituto de Fisiología Celular, UNAM y la Dra. Norma Espinosa Sánchez del departamento de Genética Molecular del Instituto de Fisiología Celular, UNAM, por su apoyo y colaboración brindada para la generación de los antisueros usados en este proyecto.
- A los doctores Genaro Patiño López y Ricardo Valle Ríos del Laboratorio de Investigación en Inmunología y Proteómica del HIMFG, por su colaboración

brindada para realización de los estudios de microscopía confocal y citometría de flujo, respectivamente.

DEDICATORIA

- A mis amores Janet, Fany y Melisa por apoyarme siempre y sobre todo por alentarme a seguir preparándome. Este nuevo logro es gracias a ustedes, su comprensión y sus sacrificios valieron la pena, son mi fuente de motivación y esperanza, le doy gracias a Dios por la bendición que me envió con ustedes.
- A mis padres Franco e Hilaria, por su apoyo incansable, valores que me han inculcado, consejos y por haberme dado la oportunidad de tener una excelente educación en el transcurso de mi vida ya que gracias a ello hoy culmino una etapa más de mi vida.
- A mis hermanos Omar y Viridiana porque ser parte importante de mi vida y por estar siempre ahí para apoyarme.
- A mi abuela, abuelo, tíos, tías, primos, primas, sobrino y demás miembros de la gran familia a la que pertenezco, por alentarme a seguir preparándome.



PDCB/grad/135/Jur/2016

C.P. AGUSTIN MERCADO
DIRECTOR DE CERTIFICACIÓN
Y CONTROL DOCUMENTAL, UNAM
Presente.

Nos permitimos informarle que con base al Artículo 31 del RGEF el Comité Académico de **DOCTORADO EN CIENCIAS BIOMÉDICAS**, en su reunión 396 del 05 de octubre de 2016, designó el siguiente jurado para examen grado de **DOCTOR EN CIENCIAS** de **ZEUS SALDAÑA AHUACTZI**, con número de cuenta **513012754**, con la tesis titulada "**BIOGÉNESIS Y REGULACIÓN DE LONGUS: UNA ADHESINA DE ESCHERICHIA COLI ENTEROTOXIGÉNICA**." dirigida por el Dr. Juan Xicohtencatl Cortés

Presidente: Dr. Dimitris Georgellis
Secretario: Dr. Juan Xicohtencatl Cortés
Vocal: Dra. Rosa Laura Camarena Mejía
Vocal: Dr. Juan Pablo Reyes Grajeda
Vocal: Dr. Miguel Ángel Cevallos Gaos

Atentamente

"POR MI RAZA HABLARÁ EL ESPÍRITU"

Ciudad Universitaria, Cd. Mx., a 10 de octubre de 2016

DRA. AUREA OROZCO RIVAS
COORDINADORA

DRA. ELVIRA GALARRAGA PALACIO
RESPONSABLE DE ENTIDAD

*AOR/EGP/SMI

CONTENIDO

1.0	RESUMEN	9
2.0	ABSTRACT	11
3.0	ANTECEDENTES	13
3.1	Las enfermedades diarreicas	13
3.2	Patogrupos de <i>E. coli</i>	14
3.3	Características generales de ETEC	16
3.4	Los <i>pili</i> tipo IV	17
3.5	Características de <i>Longus</i> ó CS21	18
3.6	Biogénesis de los TIVPb	19
3.7	Regulación de los TIVPb	21
4.0	PLANTEAMIENTO DEL PROBLEMA	24
5.0	JUSTIFICACIÓN	25
6.0	OBJETIVO GENERAL	26
6.1	Objetivos particulares	26
7.0	MÉTODOS	27
7.1	Reacción en cadena de la Polimerasa (PCR)	27
7.2	Preparación de células electrocompetentes.....	27
7.3	Transformación de las células electrocompetentes por electroporación.....	28
7.4	Construcción de mutantes isogénicas.....	28
7.5	Eliminación del casete de resistencia por recombinación homóloga	29

7.6	Expresión y purificación de las proteínas recombinantes LngB-His y LngC-His	29
7.7	Purificación de los <i>pili</i> CS21	30
7.8	Producción de sueros anti-CS21, anti-LngB y anti-LngC (preparación de los antígenos y protocolo de inmunización).....	31
7.9	Cuantificación de bacterias adheridas a células intestinales HT-29.....	32
7.10	Western blot (WB)	32
7.11	Ensayo de inmunofluorescencia (IF).....	33
7.12	Microscopía electrónica de transmisión (TEM) y microscopía inmunoelectrónica (IEM)	34
7.13	Ensayo de formación de biopelículas.....	35
7.14	Ensayo de auto-agregación bacteriana	36
7.15	Ensayos de citometría de flujo.....	37
7.16	Análisis estadístico	37
7.17	Análisis Bioinformáticos.....	37
8.0	RESULTADOS	40
8.1	Ensamble del <i>pilus</i> CS21.....	40
8.1.1	Proteínas involucradas en la biogénesis de los <i>pili</i> TCP, CFA/III, BFP y CS21.....	40
8.1.2	Predicción <i>in silico</i> de la localización subcelular de las proteínas Lng	42
8.1.3	Mutagénesis de los genes <i>lngR</i> , <i>lngS</i> , <i>lngA</i> , <i>lngB</i> , <i>lngC</i> , <i>lngD</i> , <i>lngH</i> , <i>lngP</i> y <i>hns</i>	47
8.1.4	Re-introducción del gen silvestre en cada una de las mutantes generadas	50
8.1.5	Mutación de los genes <i>lngA</i> , <i>lngB</i> , <i>lngC</i> , <i>lngD</i> , <i>lngH</i> y <i>lngP</i> en ETEC E9034A afecta la adherencia a las células intestinales HT-29 y la auto-agregación bacteriana	52

8.1.6	El <i>pilus</i> CS21 de ETEC no contribuye a la formación de biopelículas.....	57
8.1.7	LngP es una prepilina peptidasa que procesa a la pilina LngA	58
8.1.8	Las mutantes en los genes <i>lngA</i> , <i>lngB</i> , <i>lngC</i> , <i>lngD</i> , <i>lngH</i> y <i>lngP</i> no ensamblan el <i>pilus</i> CS21	60
8.1.9	El suero anti-CS21 reconoce a las proteínas LngA y LngB.....	63
8.1.10	LngB es una proteína estructural menor del <i>pilus</i> CS21	64
8.1.11	Las proteínas LngC, LngD y LngH se requieren para el ensamble del <i>pilus</i> CS21	68
8.2	Regulación del <i>pilus</i> CS21	75
8.2.1	El gen <i>lngA</i> tiene un promotor corriente arriba de su codón de inicio	75
8.2.2	Regulación postranscripcional de <i>lng</i>	77
8.2.3	LngR favorece la expresión de CS21	78
8.2.4	LngS es un probable activador transcripcional de la expresión del <i>pilus</i> CS21	83
8.2.5	H-NS es un represor de la expresión del <i>pilus</i> CS21.....	85
9.0	DISCUSIÓN	90
10.0	CONCLUSIONES	104
11.0	APÉNDICES	105
11.1	Amortiguadores y soluciones	105
11.2	Cepas bacterianas, medios de cultivo y línea celular	106
12.0	BIBLIOGRAFIA	110
14.0	PUBLICACIONES	117

INDICE DE TABLAS

Tabla 1. Tipos de <i>E. coli</i> patogénicas.....	16
Tabla 2. Mecanismos regulatorios que controlan la expresión de genes que codifican para la maquinaria de ensamble de los TIVPb.	23
Tabla 3. Función de las proteínas del ensamble de los <i>pili</i> CS21, CFA/III y TCP.	43
Tabla 4. Localización celular de las proteínas Lng en la biogénesis del <i>pilus</i> CS21.....	44
Tabla 5. Predicción de cruces transmembranales de las proteínas involucradas en la biogénesis de CS21.....	46
Tabla 6. Predicción de péptido señal de las proteínas Lng involucradas en la biogénesis de CS21..	47
Tabla 7. Cepas y plásmidos utilizados en este estudio.....	107
Tabla 8. Lista de oligonucleótidos usados en este estudio.....	108

INDICE DE FIGURAS

Figura 1. Distribución mundial de muertes causadas por diarrea en niños menores de 5 años de edad en el año 2000.....	14
Figura 2. Estimación de la mortalidad por diarrea entre el periodo de 1975 a 2000. Tomado de Keush et al., 2006.....	15
Figura 3. Divergencia entre las proteínas LngA, CofA, TcpA y BfpA de los diferentes TIVPb, PilA de los TIVPa y el pseudopilus GspG del sistema de secreción tipo II.	41
Figura 4. Organización y comparación de las proteínas estructurales codificadas en los operones de <i>cof</i> (A), <i>lng</i> (B) y <i>tcp</i> (C).	42
Figura 5. Amplificación del casete de resistencia de Km y/o Cm.....	48
Figura 6. Confirmación de la mutación del gen <i>lngA</i> en la cepa E9034A.....	49

Figura 7. Confirmación de la mutación de los genes <i>lngC</i> y <i>lngD</i>	49
Figura 8. Confirmación de la mutación del gen <i>lngB</i> y re-introducción del gen silvestre en la cepa E9034AΔ <i>lngB</i>	51
Figura 9. Curva de crecimiento de las cepas mutantes en los genes <i>lng</i>	52
Figura 10. Mutación de los genes <i>lngA</i> , <i>lngB</i> , <i>lngC</i> , <i>lngD</i> , <i>lngH</i> y <i>lngP</i> en la cepa ETEC E9034A afecta la adherencia a células intestinales HT-29.	53
Figura 11. Adherencia mediada por el <i>pilus</i> CS21 de ETEC a células intestinales HT-29.....	55
Figura 12. La ausencia de los genes <i>lngA</i> , <i>lngB</i> , <i>lngC</i> , <i>lngD</i> , <i>lngH</i> y <i>lngP</i> , afecta la auto-agregación de ETEC.....	56
Figura 13. Cuantificación de la formación de biopelículas de las cepas de ETEC.	57
Figura 14. LngP es una prepilina peptidasa que procesa a la pilina LngA.	59
Figura 15. Presencia o ausencia del <i>pilus</i> CS21 en cepas de ETEC.	60
Figura 16. Visualización del <i>pilus</i> CS21 en la cepa E9034A.	61
Figura 17. Visualización del <i>pilus</i> CS21 por IEM en el conjunto de cepas de ETEC.	62
Figura 18. LngA y LngB son reconocidos por el suero anti-CS21.	64
Figura 19. Expresión del gen <i>lngB</i> en <i>trans</i> restaura la expresión del <i>pilus</i> CS21.	66
Figura 20. Sobreexpresión de la proteína LngA no afecta la longitud del <i>pilus</i> CS21.....	67
Figura 21. Detección de la proteína LngB en la punta de los <i>pili</i> CS21.....	68
Figura 22. Procesamiento de la proteína LngC.	70
Figura 23. La mutación de los genes <i>lngC</i> y <i>lngD</i> afectó el procesamiento de la proteína LngA.	71
Figura 24. Modelo propuesto de la localización celular de las proteínas Lng y posible ensamblaje de CS21.....	74
Figura 25. Regiones intergénicas y codificantes del cluster <i>lng</i> y posibles promotores.	75
Figura 26. Expresión y procesamiento completo de la proteína LngA en una cepa de <i>E. coli</i> K-12 TOP10(pLngA).....	76

Figura 27. Probables terminadores o atenuadores transcripcionales independiente de Rho presentes en el cluster de <i>lng</i> .	78
Figura 28. LngR y PapB son homólogos con una identidad del 35.6%.	79
Figura 29. Confirmación de la mutación del gen <i>lngR</i> en la cepa E9034A.	79
Figura 30. La cepa E9034AΔ <i>lngR</i> :: <i>km</i> expresa más <i>pili</i> CS21 y se adhiere tres veces más a las células intestinales HT-29 que la cepa silvestre E9034A.	80
Figura 31. Caracterización de la cepa mutante E9034AΔ <i>lngR</i> y la cepa E9034AΔ <i>lngR</i> (pUC <i>lngR</i>).	82
Figura 32. Confirmación de la mutación del gen <i>lngS</i> y eliminación del casete de resistencia.	83
Figura 33. Sobreexpresión de LngS incrementa los niveles de expresión de LngA y del <i>pilus</i> CS21 por WB e IF.	84
Figura 34. Mutación del gen <i>hns</i> en la cepa E9034A.	86
Figura 35. Sobreexpresión de LngA en la mutante del gen <i>hns</i> .	86
Figura 37. LngR y LngS son necesarios para la expresión de CS21 y H-NS muestra un efecto negativo.	88
Figura 38. Modelo de la regulación del <i>pilus</i> CS21 en ETEC. H-NS actúa como un regulador negativo de la expresión del <i>pilus</i> CS21.	89

ABREVIATURAS

AAFs	Fimbrias de adherencia agregativa (por sus siglas en inglés: <u>A</u> ggregative <u>A</u> dherence <u>F</u> imbriae).
AIDA	Adhesina involucrada en la adherencia difusa (por sus siglas en inglés: <u>A</u> dhesin <u>I</u> nvolved in <u>D</u> iffuse <u>A</u> dherence)
BFP	<i>Pilus</i> formador de bucles o mechones (por sus siglas en inglés: <u>B</u> undle <u>F</u> orming <u>P</u> ilus).
BPROM	Predictor de promotores bacterianos (por sus siglas en inglés: <u>B</u> acterial <u>P</u> romoter prediction).
CCTOP	Servidor de predicción restringida a la topología consenso (por sus siglas en inglés: <u>C</u> onstrained <u>C</u> onsensus <u>T</u> OPology prediction server)
CELLO	Predictor de la localización subcelular (por sus siglas en inglés: sub <u>C</u> ELLular <u>L</u> ocalization predictor)
CFA/III	Factor antigénico de colonización (por sus siglas en inglés: <u>C</u> olonization <u>F</u> actor <u>A</u> ntigen III)
CFAs	Factores antigénicos de colonización (por sus siglas en inglés: <u>C</u> olonization <u>F</u> actors <u>A</u> ntigens).
C-terminal	Carboxilo terminal
Daa	Adhesina difusa (por sus siglas en inglés Diffuse Adhesin)
DMEM-HG	Medio de Cultivo Eagle Modificado de Dulbecco alto en glucosa (por sus siglas en inglés: <u>D</u> ubelcco's <u>M</u> odified <u>E</u> agle <u>M</u> édium- <u>H</u> igh <u>G</u> lucose).
Dr	Adhesina Dr
DS	Desviación estándar
HRP	Peroxidasa de rábano (por sus siglas en inglés: Horseradish peroxidase)
IEM	Microscopía inmunoelectrónica (por sus siglas en inglés: <u>I</u> mmuno <u>E</u> lectron <u>M</u> icroscopy)
IF	Inmunofluorescencia
IS	Secuencia de Inserción o Elemento de Inserción (por sus siglas en inglés: <u>I</u> nsertion <u>S</u> equences)
Ipa	Antígeno plasmidico de invasión (por sus siglas en inglés: <u>I</u> nvasion plasmid <u>a</u> ntigen)
ITU	Infecciones del tracto urinario
LB	Luria-Bertani
LEE	Locus de esfacelado (borramiento) del enterocito (por sus siglas en inglés: <u>L</u> ocus for <u>E</u> nterocyte <u>E</u> ffacement).
LT	Toxina termo-lábil (por sus siglas en inglés: <u>L</u> abile <u>T</u> oxin)
MOI	Multiplicidad de infección (por sus siglas en inglés:

	<u>M</u> ultiplicity <u>O</u> f <u>I</u> nfection).
NNPP	Predicción de promotores por redes neuronales (por sus siglas en inglés: <u>N</u> eural <u>N</u> etwork <u>P</u> romoter <u>P</u> rediction).
N-terminal	Amino terminal
ORF	Marco de lectura abierto (por sus siglas en inglés: <u>O</u> pen <u>R</u> eading <u>F</u> rame)
PePPER	Predicción de promotores procariontes y regulones (por sus siglas en inglés: <u>P</u> rediction of <u>P</u> ro <u>k</u> aryote <u>P</u> romoter <u>E</u> lements and <u>R</u> egulons).
PSI	Libras por pulgada cuadrada (por sus siglas en inglés <u>P</u> ounds per <u>S</u> quare <u>I</u> nch)
PSLpred	Predictor de la localización subcelular de proteínas procariotas (por sus siglas en inglés: <u>P</u> redicts <u>S</u> ubcellular <u>L</u> ocalization of prokaryotic proteins)
PPLO	Medio de cultivo para organismos similares a los de la pleuroneumonía (por sus siglas en ingles <u>P</u> leuro- <u>P</u> neumonia- <u>L</u> ike <u>O</u> rganisms)
PVDF	Difluoruro de polivinilideno (por sus siglas en inglés: <u>P</u> oly <u>v</u> inylidene <u>d</u> ifluoride)
RMN	Resonancia magnética nuclear
ST	Toxina termo-estable (por sus siglas en inglés: <u>S</u> t <u>a</u> ble <u>T</u> oxin)
SVM	Máquinas de soporte vectorial (por sus siglas en inglés: <u>S</u> upport <u>V</u> ector <u>M</u> achines)
TB	Caldo terrific broth
TCP	<i>Pilus</i> co-regulado con la toxina colérica (por sus siglas en inglés: <u>T</u> oxin <u>C</u> orregulated <u>P</u> ilus)
TEM	Microscopía electrónica de transmisión (por sus siglas en inglés: <u>T</u> ransmission <u>E</u> lectron <u>M</u> icroscopy)
TIVP	<i>Pili</i> tipo IV (por sus siglas en inglés, <u>T</u> ype <u>I</u> V <u>P</u> ili)
UFC	Unidades formadoras de colonias
WB	Western blot

1.0 RESUMEN

Escherichia coli enterotoxigénica (ETEC) es uno de los principales agentes microbianos responsables de la morbilidad en niños menores de 5 años de edad en países de ingresos bajos y medianos, y la principal causa de la diarrea del viajero en todo el mundo. La capacidad de ETEC para colonizar el epitelio intestinal está mediada por adhesinas fimbriales, tales como CS21 (*Longus*). Esta adhesina es un *pilus* tipo IV involucrada en la adherencia a las células intestinales *in vitro* y en la auto-agregación bacteriana. Se han propuesto catorce marcos de lectura abiertos que podrían estar involucrados en el ensamble de CS21. Sin embargo, hasta ahora sólo los genes *lngA* y *lngB*, que codifican para la subunidad estructural mayor (LngA) y menor (LngB), se han caracterizado. En este estudio, se investigó el papel de las proteínas LngA, LngB, LngC, LngD, LngH y LngP en el ensamble de CS21 y LngR, LngS y H-NS en la regulación de *lng* en la cepa de ETEC E9034A. La mutación de los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* ó *lngP*, abolió el ensamble de CS21 en la cepa de ETEC E9034A y la adherencia a células intestinales HT-29 se redujo un 90%, en comparación con la cepa silvestre E9034A. La predicción de la localización subcelular de las proteínas CS21 fue similar a otras proteínas homólogas de fimbrias tipo IV bien conocidas. Se demostró que LngP es la prepilin peptidasa de LngA y que la cepa E9034A tiene otra prepilin peptidasa capaz de procesar LngA, aunque con menor eficiencia. Además, se presentan imágenes de microscopía electrónica para mostrar que la proteína LngB podría localizarse en la punta de CS21. La sobreexpresión de LngB, *in trans*, aumenta la longitud CS21. LngH y LngD son homólogas de las proteínas de unión a nucleótidos y las lipoproteínas de membrana externa (secretina) de otras fimbrias tipo IV,

respectivamente. La mutación de los genes *lngC* y *lngD* no afectó la expresión LngA, pero su procesamiento fue incompleto y el ensamble de CS21 se inhibió. LngR y LngS son los probables reguladores positivos de *lng* que superan la represión mediada por la proteína H-NS. En conclusión, nuestros resultados demuestran que las proteínas LngA, LngB, LngC, LngD, LngH y LngP son esenciales para el ensamblaje del *pilus* CS21 y que LngR y LngS son los probables reguladores positivos de *lng*.

2.0 ABSTRACT

Enterotoxigenic *Escherichia coli* (ETEC) is one of the major microbial agents responsible of morbidity in children under 5 years of age in low- and middle-income countries and a leading cause of traveler's diarrhea worldwide. The ability of ETEC to colonize the intestinal epithelium is mediated by fimbrial adhesins, such as CS21 (*Longus*). This adhesin is a type IVb *pilus* involved in adherence to intestinal cells *in vitro* and bacterial self-aggregation. Fourteen open reading frames have been proposed to be involved in CS21 assembly, hitherto only the *lngA* and *lngB* genes, coding for the major (LngA) and minor (LngB) structural subunit, have been characterized. In this study, we investigated the role of the LngA, LngB, LngC, LngD, LngH, and LngP proteins in the assembly of CS21 and LngR, LngS, and H-NS in CS21 regulation in ETEC strain E9034A. The deletion of the *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, or *lngP* genes, abolished CS21 assembly in ETEC strain E9034A and adherence to HT-29 cells was reduced 90%, compared to wild-type strain. Subcellular localization prediction of CS21 proteins was similar to other well-known type IV *pili* homologues. We showed that LngP is the prepilin peptidase of LngA, and that ETEC strain E9034A has another peptidase capable of processing LngA, although with less efficiency. Additionally, we present immuno-electron microscopy images to show that the LngB protein could be localized at the tip of CS21. Interesting, overexpression of LngB *trans* increased CS21 length. LngH and LngD are homologous with the nucleotide-binding proteins and the multimeric outer membrane lipoprotein (secretin), respectively, related to other type IV *pilus* assembly. The mutation of the *lngC* and *lngD* genes did not affect LngA expression, but the processing of LngA was incomplete and assembly of CS21 was

disrupted. LngR and LngS are the possible positive regulators of *lng* that overcome the repression mediated by H-NS. In conclusion, our results demonstrate that the LngA, LngB, LngC, LngD, LngH, and LngP proteins are essential for CS21 assembly.

3.0 ANTECEDENTES

3.1 Las enfermedades diarreicas

Las enfermedades diarreicas continúan siendo una de las principales causas de muerte prevenibles en los países en vías de desarrollo, especialmente en los niños menores de cinco años de edad (Keusch et al., 2006). El tracto intestinal normal, regula la absorción y la secreción de electrolitos, y de agua para satisfacer las necesidades fisiológicas del cuerpo. Los intestinos de una persona adulta sana, reabsorbe más del 98% de los 10 litros de fluido que pasan por ellos por día (Keusch, 2001). La diarrea es una alteración de la función intestinal normal que se caracteriza por un aumento del volumen, la fluidez y las deposiciones, en comparación con las condiciones fisiológicas normales. La frecuencia de las deposiciones acuosas generalmente son de tres o más, en menos de 24 horas, durante tres días o más y menor a 14 días, considerándose como una enfermedad (Keusch et al., 2006). La mayoría de las personas que fallecen por enfermedades diarreicas, mueren por una grave deshidratación. Los episodios graves y repetidos de diarrea en los recién nacidos y en los niños pequeños puede causar efectos a largo plazo, tales como: retraso en el crecimiento y alteraciones cognitivas (Niehaus et al., 2002). La diarrea infecciosa es causada generalmente por agentes biológicos, como: virus, bacterias, protozoos y helmintos; los cuales, se transmiten por la ruta fecal-oral. Estos organismos se diferencian, entre otras cosas, por sus factores de virulencia específicos, síntomas y en el número de organismos necesarios para causar infección y enfermedad (Keusch et al., 2006). La diarrea infecciosa es una de las principales causas de morbilidad y mortalidad en todo el mundo **(Figura 1)**.

Los episodios de diarrea infecciosa ocurren alrededor de 2,000-4,000 millones de casos cada año y son especialmente frecuentes en niños menores de dos años (Hodges y

Gill, 2010). La mortalidad de la diarrea infecciosa a nivel mundial ha disminuido notoriamente de 2,487,400 casos en 1990 a 1,445,800 casos en 2010, como se muestra en la figura 2 (Lozano et al., 2012). Sin embargo, la morbilidad ha disminuido lentamente de 2.9 mil millones de casos en 1990 a 2.7 mil millones de casos en el 2013 (Vos et al., 2015). La diarrea puede durar varios días o incluso semanas, dependiendo del agente biológico. En países en vías de desarrollo, los niños menores de tres años sufren en promedio tres episodios de diarrea al año asociados principalmente a rotavirus y *E. coli* (WHO, 2013).

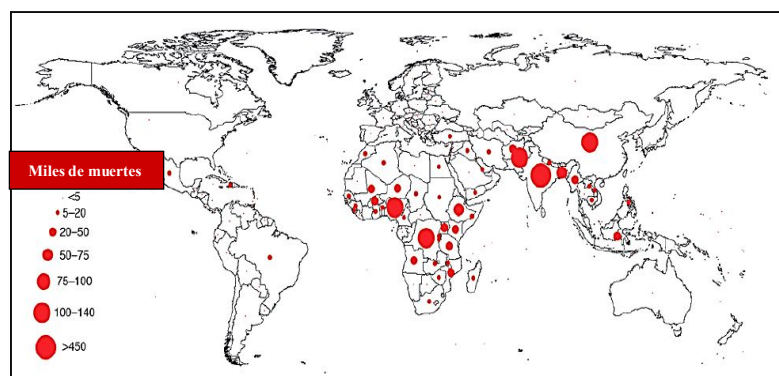


Figura 1. Distribución mundial de muertes causadas por diarrea en niños menores de 5 años de edad en el año 2000. Tomado de Keush et al., 2006.

3.2 Patogrupos de *E. coli*

E. coli es un bacilo Gram negativo de la familia enterobacteriacea que coloniza el intestino y aparece en las heces de los bebés pocos días después de nacer (Hewitt y Rigby, 1976; Feeney et al., 1980). *E. coli* se considera una bacteria comensal que vive en asociación y mutuo beneficio con su hospedero y raramente causa enfermedad; sin embargo, existen otras *E. coli* patógenas que causan una amplia gama de enfermedades en el ser humano y en los animales (Allocati et al., 2013). La capacidad de causar una enfermedad se debe a diversos factores de virulencia específicos de cepas de *E. coli* patógena; las cuales se han

clasificado mediante pruebas serológicas, utilizando anticuerpos específicos contra los antígenos O (lipopolisacárido) y H (flagelar) (Kaper et al., 2004). Además, los factores de virulencia expresados por diferentes cepas de *E. coli* patogénicas, han sido una herramienta importante para clasificar a estas bacterias en diferentes patotipos (**Tabla 1**).

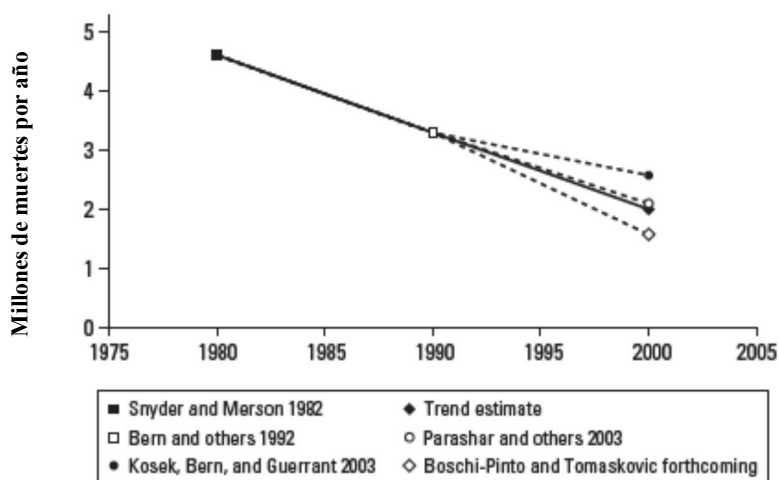


Figura 2. Estimación de la mortalidad por diarrea entre el periodo de 1975 a 2000. Tomado de Keush et al., 2006.

Los patotipos de *E. coli*, se han dividido en dos grupos principalmente: intestinales y extra intestinales. Los patotipos de *E. coli* intestinales o diarrogénicas, son: *E. coli* enteroagregativa (EAEC, por sus siglas en inglés), *E. coli* de adherencia difusa (DAEC, por sus siglas en inglés), *E. coli* adherente invasiva (AIEC, por sus siglas en inglés), *E. coli* enterohemorrágica (EHEC, por sus siglas en inglés), *E. coli* enteroinvasiva (EIEC, por sus siglas en inglés), *E. coli* enteropatógena (EPEC, por sus siglas en inglés) y *E. coli* enterotoxigénica (ETEC, por sus siglas en inglés). Los patotipos extraintestinales, son: *E. coli* uropatógena (UPEC, por sus siglas en inglés) y *E. coli* causante de meningitis neonatal (MAEC, por sus siglas en inglés). Los patogrupos diarrogénicos de *E. coli*, se han

asociado con casos esporádicos y brotes de diarrea en humanos (Levine, 1987; Nataro y Kaper, 1998; Allocati et al., 2013).

Tabla 1. Tipos de *E. coli* patogénicas.

Patotipo	Enfermedad	Síntomas	Factores de virulencia
EPEC	Diarrea en niños	Diarrea acuosa y vomito	BFP, Intimina, LEE
EHEC	Colitis hemorrágica, Síndrome urémico hemolítico	Diarrea con sangre, fiebre, dolor abdominal	Toxina Shiga, Intimina, LEE
ETEC	Diarrea del viajero	Diarrea acuosa y vomito	ST, LT, CFAs
EAEC	Diarrea en niños	Diarrea con moco y vomito	AAFs, citotóxicas
DAEC	Diarrea aguda en niños	Diarrea acuosa, ITU recurrentes	Daa, Dr, AIDA
EIEC	Disentería	Diarrea acuosa o con moco o con sangre, cólicos abdominales.	Toxina Shiga, hemolisina, invasión celular, Ipa
AIEC	Asociada con la enfermedad de Crohn	Inflamación intestinal persistente	<i>Pili</i> tipo 1, Invasión celular
UPEC	ITU e infecciones sistémicas	Cistitis, pielonefritis	<i>Pili</i> tipo 1, fimbria P, AAFs, hemolisina
NMEC	Meningitis neonatal	Meningitis aguda, sepsis	Fimbria S, Capsula K1
APEC	Fuente probable de enfermedades alimentarias	-	<i>Pili</i> tipo 1, fimbria P, capsula K1

BFP, *pilus* formador de mechones; LEE, locus de borramiento del enterocito; ST, toxina termo-estable; LT, toxina termo-lábil; CFAs, factores antigénicos de colonización; Daa, adhesina difusa; AIDA, adhesina involucrada en la adherencia difusa; AAFs, fimbrias de adherencia agregativa. Tomado de Nataro y Kaper, 1998 y Kaper et al., 2004.

3.3 Características generales de ETEC

ETEC es uno de los principales agentes etiológicos bacterianos asociado a enfermedades diarreicas en niños menores de 2 años y es causante de la diarrea del viajero en países en vías de desarrollo, como México (Nataro y Kaper, 1998; Kaper et al., 2004). La mortalidad por ETEC en el 2010, se estimó en 120,800 casos aproximadamente (Lozano et al., 2013). Las infecciones diarreicas producidas por ETEC se manifiestan principalmente en las poblaciones de bajo ingreso económico y con condiciones sanitarias no favorables (Nataro y Kaper, 1998; Kaper et al., 2004). El cuadro clínico de la infección por ETEC se

caracteriza por una diarrea acuosa que conlleva a una hipersecreción masiva de electrolitos. Las enterotoxinas [toxinas termo-lábil (LT) y termo-estable (ST)] son las responsables de los síntomas arriba descritos (Nataro y Kaper, 1998; Kaper et al., 2004). La infección por ETEC, ocurre primeramente por adhesión y colonización de la bacteria a los enterocitos del intestino, a través de la expresión de adhesinas fimbriales antigénicamente distintas, tales como: factores antigénicos de colonización (CFAs), antígenos de superficie de *E. coli* (CSs) y factores de colonización potenciales (PCFs); los cuales, reconocen diferentes receptores celulares aún no descritos (Kaper et al., 2004). Análisis de diferentes estudios epidemiológicos en el mundo, de cepas de ETEC aisladas de enfermedades diarreicas, han revelado la presencia de diferentes CFAs, CSs y PCFs (Gaastra y de Graaf, 1982; Levine, 1987; Gaastra y Svennerholm, 1996). Las cepas típicas de ETEC expresan generalmente de una a tres adhesinas fimbriales. Las adhesinas fimbriales más prevalentes a nivel mundial son los CFA/I, CS1, CS2, CS3, CS5, CS6 y CS21 (Gaastra y Svennerholm, 1996; Isidean et al., 2011).

3.4 Los *pili* tipo IV

Los *pili* tipo IV (TIVP, por sus siglas en inglés) son apéndices filamentosos delgados que emergen de la envoltura bacteriana hacia el medio ambiente extracelular (Roux et al., 2012). Los TIVP están formados por homopolímeros de una misma proteína, llamada comúnmente pilina (Roux et al., 2012). Los TIVP participan en el proceso de adhesión a las células hospedadoras, formación de biopelículas, auto-agregación bacteriana, intercambio de ADN y motilidad tipo *twitching* (Mazariego-Espinosa et al., 2010; Roux et al., 2012). Los *pili* tipo IV tienen la capacidad de producir un mecanismo de retracción y despolimerización para generar un movimiento bacteriano especializado en superficies

semisólidas (como los epitelios de las mucosas); el cual, es llamado *twitching motility* (Pizarro-Cerda y Cossart, 2006).

Los TIVP se clasifican en dos tipos: TIVPa y TIVPb basado en las características estructurales de sus pilinas. Las pilinas de los TIVPb se distinguen por tener un péptido señal largo en su N-terminal de 13-30 residuos y generalmente no se han asociado con la motilidad de tipo twitching; aunque, tienen mecanismos de ensamblaje con características similares a los TIVPa (Roux et al., 2012). Las pilinas de los TIVPb, se caracterizan por estar codificados por un número limitado de genes (11-14) agrupados en operones; los cuales, probablemente se adquirieron por transferencia horizontal. Las pilinas de los TIVPa tienen un péptido señal corto de 5-6 residuos y su biogénesis requiere de más de 40 genes, agrupados en diferentes operones y diferentes locus (Roux et al., 2012).

3.5 Características de *Longus* ó CS21

Longus (long pilus) o CS21, es un TIVPb polar, semi-flexible capaz de formar mechones y agregados. El *pilus* CS21 está constituido por cientos de unidades de la misma proteína “LngA” de 22-kDa, que se ensamblan para formar estructuras poliméricas con una longitud mayor a 20 μm y un diámetro de entre 5 y 7 nm (Girón et al., 1997). La proteína LngA se ha relacionado con la subunidad estructural CofA de CFA/III con una similitud entre ellas del 79%. La región N-terminal de la proteína LngA tiene una identidad del 76% con la subunidad estructural TcpA del *pilus* corregulado de la toxina colérica (TCP, por sus siglas en ingles) de *Vibrio cholerae* (*V. cholerae*) y del 36% con el N-terminal de PilA de *Pseudomonas aeruginosa* (Girón et al., 1997).

El gen *lngA* que codifica para la proteína LngA, se encuentra distribuido ampliamente en diferentes cepas clínicas de ETEC aisladas en el mundo, tales como:

Egipto, Bangladesh, Argentina, Chile y México (Giron et al., 1995; Gutierrez-Cazarez et al., 2000; Pichel et al., 2002; Gomez-Duarte et al., 2007). La prevalencia del gen *LngA* está en función de la región y se ha observado que el 35% de cepas de ETEC aisladas de niños en México, 20.7% de cepas aisladas en Argentina y 8.5% de cepas aisladas de Bangladesh contienen este gen (Gutierrez-Cazarez et al., 2000; Qadri et al., 2000; Pichel et al., 2002).

El *pilus* CS21 participa en la adherencia a células intestinales HT-29, T84 y Caco2; además, es la estructura proteica responsable del movimiento tipo *twitching* de la bacteria (Mazariego-Espinosa et al., 2010). La auto-agregación bacteriana es otra función del *pilus* CS21, un mecanismo que favorece que las cepas de ETEC CS21⁺ puedan resistir agentes antimicrobianos (Clavijo et al., 2010). Recientemente, otros estudios empleando un modelo de ratón neonato, han mostrado que el *pilus* CS21 tiene una función importante en la patogénesis de ETEC (Guevara et al., 2013). Nuestro grupo de trabajo demostró que el *pilus* CS21 de una cepa clínica de ETEC 73332 (aislada de una paciente pediátrica mexicana), realiza una función importante en la adhesión a células intestinales HT-29 y en la auto-agregación bacteriana (Cruz-Cordova et al., 2014). Estudios *in vivo*, demostraron que la proteína LngA y el *pilus* CS21 de ETEC, son estructuras inmunogénicas y con la capacidad de proteger contra la infección de ETEC CS21⁺. La inmunogenicidad y antigenicidad de la proteína LngA y el *pilus* CS21 indican que pueden ser candidatos ideales para generar una vacuna multivalente contra la diarrea provocada por ETEC (Zhang et al., 2016).

3.6 Biogénesis de los TIVPb

La biogénesis de los TIVP involucra la oligomerización y la secreción de las pilinas por un mecanismo que requiere un conjunto de proteínas del aparato de ensamble y exportación.

Los *pili* se ensamblan en la membrana interna y los oligómeros se secretan a través de la membrana externa (Wolfgang et al., 2000). El ensamble requiere de una proteína de unión a nucleótidos “ATPasa” para proveer energía, una proteína de membrana interna politópica “prepilin peptidasa” que procesa a las pilinas antes de su oligomerización y proteínas accesorias de membrana interna y externa (Mattick, 2002). Las subunidades que oligomerizan para formar el *pilus* se sintetizan como prepilinas con un péptido señal en su N-terminal, seguido de una hélice alfa hidrofóbica y una región que se pliega en un dominio globular con un par de residuos de cisteína en el extremo C-terminal (Roux et al., 2012). Las pilinas tipo IVb pueden ser procesadas por el grupo I de las prepilin peptidasas semejantes a PilD, como BfpP (BFP, EPEC), TcpJ (TCP, *V. cholerae*), LngP (CS21, ETEC) y CofP (CFA/III, ETEC) (Roux et al., 2012). Alternativamente, las pilinas tipo IVb podrían ser procesadas por una prepilin peptidasa filogenéticamente distante del grupo II (Roux et al., 2012). El procesamiento incluye la escisión del péptido señal y la metilación del primer aminoácido de la proteína madura, por ejemplo, una metionina para las pilinas TcpA, LngA y CofA, y una lisina para BfpA (Hobbs y Mattick, 1993).

La biogénesis del *pilus* formador de bucles o mechones “BFP” de EPEC, el *pilus* co-regulado con la toxina colérica “TCP” de *V. cholerae* y el factor antigénico de colonización “CFA/III” de ETEC, involucra varias proteínas homologas del *pilus* CS21 codificadas en los genes del operón *lng* (Gomez-Duarte et al., 2007; Roux et al., 2012). Las proteínas involucradas en la biogénesis del *pilus* CS21 se encuentran codificadas por los genes localizados en una región de 14 kb y 14 marcos de lectura (Gomez-Duarte et al., 2007). Estos 14 genes (*lngR*, *lngS*, *lngT*, *lngA*, *lngB*, *lngC*, *lngD*, *lngE*, *lngF*, *lngG*, *lngH*, *lngI*, *lngJ*, y *lngP*) son homólogos y topológicamente idénticos al cluster de los genes de CFA/III (Taniguchi et al., 2001; Gomez-Duarte et al., 2007). La identidad entre los *pili*

CS21 y CFA/III es del 62% a 83% a nivel de secuencia nucleotídica y del 57% a 95% a nivel de secuencia de aminoácidos (Gomez-Duarte et al., 2007). La similitud alta que existe entre los genes involucrados en la biogénesis de los *pili* CS21 y CFA/III, indica que estos son parálogos (Gomez-Duarte et al., 2007). Las proteínas LngA, LngB, LngC, LngD, LngE, LngF, LngG, LngH, LngI, LngJ y LngP tienen más del 50% de identidad con CofA, CofB, CofC, CofD, CofE, CofF, CofG, CofH, CofI, CofJ y CofP, respectivamente (Gomez-Duarte et al., 2007). La proteína CofA es la pilina estructural de CFA/III y CofB es la pilina menor, necesaria para iniciar el ensamblaje del *pilus* CFA/III (Taniguchi et al., 2001; Kolappan et al., 2015). La proteína CofJ, es una molécula soluble secretada a través del sistema CFA/III que se ha propuesto como una adhesina potencial de ETEC (Yuen et al., 2013). La proteína CofP tiene una función de prepilin peptidasa de CofA (Taniguchi et al., 1999). La biogénesis del *pilus* CS21 involucra proteínas como LngA, la pilina estructural del *pilus* CS21 y LngB, la pilina necesaria para iniciar la oligomerización de LngA (Giron et al., 1994; Kolappan et al., 2015). El cluster *lng* contiene dos genes corriente arriba del gen *lngA*, que codifican para proteínas regulatorias LngR y LngS. La proteína LngR, es homóloga a las proteínas de la familia PapB de reguladores fimbriales y afimbriales (Regulador de los *pili* Tipo I y Pap); mientras, la proteína LngS tiene es homóloga a las proteínas de la familia AraC de reguladores transcripcionales; ambos reguladores contienen dominios típicos de unión al ADN (Gomez-Duarte et al., 2007).

3.7 Regulación de los TIVPb

La actividad de la ARN polimerasa, en un promotor, es regulada por la interacción con proteínas accesorias que afectan su capacidad de reconocer el sitio de inicio de la transcripción. Los reguladores positivos (activadores) influyen en el reconocimiento de la

RNA polimerasa hacia el promotor y los reguladores negativos (represores) actúan obstruyendo el paso de la RNA polimerasa esta o impidiendo su reclutamiento. La accesibilidad a regiones del ADN del promotor en procariontes, frecuentemente es regulado por la interacción de proteínas con el operador (región adyacente al promotor donde se unen los activadores y/o represores).

La expresión de genes que codifican para proteínas involucradas en el ensamble de los TIVPb, es controlada por un sistema de regulación complejo (**Tabla 2**). La regulación de los TIVPb, como BFP de EPEC y TCP de *V. cholerae* involucra a reguladores de la familia AraC (Roux et al., 2012). La activación transcripcional de *bfp*, está regulada por el producto del gen *perA* y en *tcp* por el producto de *toxt*. La proteína PerA pertenece específicamente a la familia de reguladores AraC/XyIS, actúa como un regulador positivo en la expresión de *bfp* y controla su propia expresión por un mecanismo de autorregulación que responde a diferentes condiciones ambientales. La expresión génica del operón *bfp* requiere de un *locus* de tres genes, localizado a una distancia de 6 kb del operón de *bfp*. Estos genes son: *bfpT*, *bfpV* y *bfpW*, los cuales codifican a las proteínas BfpT, BfpV y BfpW, respectivamente. La proteína BfpT se une al promotor del gen *bfpA* y requiere de las proteínas BfpV y BfpW, para la máxima transcripción del operón *bfpA* (Tobe et al., 1996). Contrariamente, la expresión del operón *bfp* sólo requiere de BfpT, que ha sido renombrado como PerA, sin requerimientos de BfpV (PerB) o BfpW (PerC) como factores accesorios. El promotor de *bfp* tiene una región curva, centrada en los sitios de unión de PerA, donde H-NS compete con PerA; lo cual, ayuda a contrarrestar el efecto negativo de la proteína H-NS en el promotor de *bfp* y permitir su transcripción (Porter et al., 2004).

Tabla 2. Mecanismos regulatorios que controlan la expresión de genes que codifican para la maquinaria de ensamble de los TIVPb.

Tipo de mecanismo regulatorio	BFP	TCP
Señal	Temperatura, calcio, amonio, N-acetillactosamina	Ácidos grasos insaturados, bilis
Reguladores AraC Terminación transcripcional Otros reguladores	PerA	ToxT Degradación de ToxT IHF OmpR-like TcpP y ToxR TcpH y proteínas asociadas a ToxS VarS/VarA TCS TarB sRNA
Represor H-NS Regulación postranscripcional	H-NS DegP/Cpx	VicH TcpJ-degradación dependiente de MSHA

Modificado de (Roux et al., 2012).

4.0 PLANTEAMIENTO DEL PROBLEMA

Escherichia coli enterotoxigénica (ETEC) es uno de los principales agentes etiológicos bacterianos asociado a enfermedades diarreicas en niños menores de 2 años y es causante de la diarrea del viajero en países en vías de desarrollo como México (Nataro y Kaper, 1998; Kaper et al., 2004). ETEC coloniza al intestino a través de la expresión de adhesinas fimbriales como el *pilus* CS21 y/o otros factores de colonización (CFs).

CS21 de ETEC es un *pilus* tipo IVb que participa en la colonización a células intestinales, tales como: HT-29, T84 y Caco-2. Existen diversas adhesinas fimbriales tipo IVb en las bacterias enteropatógenas que comparten características similares en su biogénesis y para la cual se requieren entre 11 a 14 genes. Aún no se ha descrito la participación de algunas proteínas potenciales codificadas por genes localizados en el operón *lng* y que probablemente están involucradas en el ensamble del *pilus* CS21. Únicamente existen estudios relacionados con análisis *in silico*. Por lo tanto, en este estudio se plantea identificar los genes que codifican a proteínas que son requeridas para el ensamble y regulación del *pilus* CS21.

5.0 JUSTIFICACIÓN

El *pilus* CS21 participa en la colonización de ETEC a células intestinales y en la autoagregación bacteriana. Además, es uno de los CFs más prevalentes en cepas ETEC en todo el mundo. Sin embargo, existe poca información de las proteínas requeridas para el ensamble del *pilus* CS21. Por lo tanto, la identificación de las proteínas involucradas en la biogénesis del *pilus* CS21 de ETEC será de gran importancia para entender el mecanismo de acción de estas estructuras en el proceso de colonización a los enterocitos. Por otra parte, la identificación de proteínas que puedan estar interaccionando con la proteína LngA en el filamento podrían ser utilizadas como candidatos terapéuticos contra ETEC.

6.0 OBJETIVO GENERAL

Definir la función de los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, *lngP* en el ensamble de CS21 y la función de los genes *lngR*, *lngS* y *hns* en la regulación de *lng* en la cepa de ETEC E9034A.

6.1 Objetivos particulares

1. Análisis bioinformático de las proteínas codificadas por los genes *lng*.
2. Mutar los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, *lngP*, *lngR*, *lngS* y *hns* en la cepa de ETEC E9034A.
3. Complementar en *trans* las mutantes generadas.
4. Evaluar si los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* son necesarios para el ensamble del *pilus* CS21.
5. Determinar la función de los genes *lngR*, *lngS* y *hns* en la regulación del operón *lng* en la cepa de ETEC E9034A.

7.0 MÉTODOS

7.1 Reacción en cadena de la Polimerasa (PCR)

La amplificación de los fragmentos génicos se realizó con la técnica de reacción en cadena de la polimerasa PCR (Polymerase Chain Reaction, por sus siglas en inglés), utilizando como molde el ADN extraído de la cepa silvestre, mutantes y complementadas. Las reacciones de amplificación se realizaron en tubos de polipropileno estériles de 0.6 ml de capacidad. La enzima Platinum® Taq ADN polimerasa de alta fidelidad se utilizó para amplificar los fragmentos génicos para la clonación de acuerdo a las indicaciones de la casa comercial de Invitrogen (CA, EE. UU.). La amplificación de los fragmentos génicos utilizados para la mutación y su comprobación se realizó con el PCR Master Mix 2x siguiendo las indicaciones de la casa comercial Promega Corporation (WI, EE. UU.). Los oligonucleótidos utilizados para PCR se diseñaron manualmente y analizaron con los programas OligoAnalyzer v3.1 y Oligo Calc. Los oligonucleótidos se sintetizaron en Integrated DNA Technologies, Inc (IDT, CA, EE. UU.) y se muestran en la **Tabla 8**.

7.2 Preparación de células electrocompetentes

Las células electrocompetentes se prepararon a partir de un cultivo fresco en caldo LB. Las cepas de *E. coli* se sembraron por estría cruzada en agar LB y se incubaron durante toda la noche a 37°C. Las colonias aisladas de cada cepa de *E. coli* se inocularon en 5 ml de caldo LB a 37°C en agitación constante de 250 rpm hasta alcanzar 0.3 de absorbancia a una densidad óptica de 600 nm. A partir de estos cultivos se tomaron 1 ml de cada uno y se transfirieron a tubos falcon de 50 ml con 24 ml de caldo LB, se incubó a 37°C en agitación constante de 250 rpm hasta alcanzar la fase logarítmica (absorbancia de 0.5-0.6). Los cultivos bacterianos se centrifugaron a 4000 x g por 15 min a 4°C. Los sobrenadantes se

decantaron y las pastillas se lavaron de la siguiente manera: Las pastillas se resuspendieron en 1 ml de agua desionizada, fría y estéril; posteriormente, se agregaron 24 ml más de agua desionizada y se centrifugaron nuevamente. Este procedimiento se repitió tres veces. Las pastillas se resuspendieron en 600 µl de glicerol al 20% en agua desionizada fría y estéril. Finalmente, se realizaron alícuotas de 100 µl de células electrocompetentes para su conservación a -70°C hasta su uso.

7.3 Transformación de las células electrocompetentes por electroporación

Cien microlitros de células electrocompetentes de *E. coli* se transformaron por electroporación a 1800 V utilizando 5 µl del plásmido purificado (~100 ng). Las células transformadas se recuperaron en 1 ml de caldo LB a 37°C en agitación constante de 250 rpm durante 1 h. Los plásmidos transformados contienen casetes de resistencia que sirvieron como marcadores de selección. Las colonias transformadas se seleccionaron en placas de agar LB con el antibiótico correspondiente.

7.4 Construcción de mutantes isogénicas

La cepa silvestre de ETEC E9034A se utilizó para generar las mutantes en los genes *hns*, *lngR*, *lngS*, *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, y *lngP*, empleando la técnica de inactivación en un paso (Datsenko y Wanner, 2000). Esta técnica consiste en utilizar un producto de PCR del gen *kan* (kanamicina) o *cat* (cloranfenicol) flanqueado por secuencias homólogas al gen de interés (blanco); además, requiere de las recombinasas (*exo*, *gamma* y *beta*) del fago lambda Red que son expresadas bajo el control de un promotor inducible (P_{araB}) en un plásmido fácilmente curable (temperatura sensible). Los oligonucleótidos para amplificar el producto de PCR, se diseñaron con un tamaño de 39 a 50 bases de longitud de las regiones homólogas a los extremos del gen de interés (**Figura 5**). Los genes de interés fueron

sustituidos por recombinación homóloga con un casete de antibiótico (kanamicina o cloranfenicol). Para la confirmación de las mutantes se diseñaron oligonucleótidos que se alinean corriente arriba y corriente debajo de la mutación del gen de interés y por PCR se confirmó la inserción del casete de resistencia. Los oligonucleótidos utilizados se muestran en la **Tabla 8**.

7.5 Eliminación del casete de resistencia por recombinación homóloga

La expulsión del casete de kanamicina se realizó utilizando el vector pFLP2. La cepa SM10 λ *pir*(pFLP2) se conjugó con las cepas de interés y por recombinación homóloga de uno de los sitios FRT del plásmido pFLP2 se insertó en la región plasmídica que alberga al casete de resistencia (Hoang et al., 1998). Las cepas transformantes se seleccionaron por el marcador de resistencia a ampicilina del vector pFLP2. Posteriormente, el plásmido linearizado e insertado se curó por una subsecuente recombinación y seleccionando colonias que crecieran en agar LB con 5% de sacarosa. Esta selección es posible gracias al gen *sacB* que está presente en el plásmido pFPL2 y que codifica a una enzima llamada levansucrasa. La enzima cataliza la hidrólisis de la sacarosa y la síntesis de polímeros de levano que tienen efectos tóxicos sobre las células bacterianas (Gay et al., 1985). Dos sitios FRT que están flanqueando a los elementos insertados extras del vector en el sitio de la mutación son eliminados gracias a la ayuda de la enzima Flipasa, dando origen a una mutante curada del promotor y gen de resistencia a kanamicina.

7.6 Expresión y purificación de las proteínas recombinantes LngB-His y LngC-His

El diseño de oligonucleótidos, amplificación, clonación y expresión de los genes *lngB* y *lngC* en el plásmido pLATE31 se realizaron de acuerdo a las especificaciones del kit aLICator, sistema de ligación independiente, clonación y expresión (Thermo Scientific,

CA, EE.UU.). Los vectores pLATE31-IngB y pLATE31-IngC se utilizaron para transformar por electroporación a la cepa *E. coli* BL21 (DE3). La expresión y purificación de las proteínas recombinantes IngB-His y IngC-His se realizó a partir de 200 ml de cultivos bacterianos de las cepas de *E. coli* BL21 (DE3) pLATE31-IngB y pLATE31-IngC crecidas en caldo LB a 37°C con agitación constante en fase logarítmica (0.5-0.6 de absorbancia). La inducción de la expresión de las proteínas recombinantes se realizó con 1 mM de IPTG (Thermo Scientific, CA, EE.UU.) durante 5 h. Los cultivos se centrifugaron a 3500 x g y las pastillas obtenidas se resuspendieron en 20 ml de un amortiguador de lisis pH 8.2; las cuales, se sonicaron a 40 kHz, utilizando 10 ciclos con pulsos de 1 minuto y 2 minutos de enfriamiento entre cada pulso. Posteriormente, las muestras sonicadas se centrifugaron a 3500 x g. Los sobrenadantes con las proteínas recombinantes His-tag se purificaron utilizando una columna de cromatografía de afinidad con resina de Ni-NTA (Qiagen, Hilden, Alemania). La resina en la columna se lavó con 60 ml de un amortiguador de lisis (pH 6.5). Las proteínas recombinantes se eluyeron con 10 ml de un amortiguador de lisis (pH 4.0) colectando fracciones de 1 ml y se analizaron por geles de SDS-PAGE al 16%. Las fracciones eluidas que contenían las proteínas recombinantes purificadas se dializaron gradualmente durante 1 h con 6, 5, 4, 3, 2 y 1 M de urea pH 8.0; finalmente, se dializaron durante 1 hora en PBS 1x pH 7.0.

7.7 Purificación de los *pili* CS21

La extracción de los *pili* CS21 de la cepa E9034A Δ IngR::Km, se realizó cosechando la masa bacteriana de 100 placas de agar LB con Km; las cual, se resuspendió en 100 ml de PBS 1X y por agitación mecánica durante 15 minutos se desprendieron los *pili* CS21. La purificación de los *pili* CS21 se realizó por centrifugación diferencial; así mismo, un

gradiente de sulfato de amonio del 10 al 60% su utilizó para asegurar una mejor purificación verificada en geles de SDS-PAGE al 16%. Las fracciones precipitadas se dializaron con PBS y se usaron para la generación del suero anti-CS21 de conejo.

7.8 Producción de sueros anti-CS21, anti-LngB y anti-LngC (preparación de los antígenos y protocolo de inmunización)

Para la obtención de los sueros anti-CS21, anti-LngB y anti-LngC, se emplearon 3 conejos blancos de la raza Nueva Zelanda que fueron inmunizados con los antígenos CS21, LngB-His y LngC-His en el bioterio del Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México, con la colaboración del MVZ Héctor Alfonso Malagón Rivero y la Dra. Norma Espinosa Sánchez. Previo a la inmunización, a los conejos se les extrajo una muestra de sangre de la oreja y el suero obtenido se analizó por WB utilizando los *pili* CS21, las proteínas LngB-His y LngC-His. Los conejos sin reactividad inmunológica con los antígenos de interés se consideraron aptos para su inmunización. El protocolo de inmunización fue vía subcutánea con intervalos de dos semanas entre cada inmunización. Una primera y única dosis de 1 ml con 500 µl de adyuvante completo de Freund más 1 mg de antígeno en 500 µl de PBS. Tres dosis de 500 µl de adyuvante incompleto de Freund más 0.5 mg de antígeno en 500 µl de PBS. Durante la inmunización de los conejos, se realizaron tres sangrías parciales para evaluar la producción de anticuerpos frente a los antígenos de interés por WB. Al finalizar el protocolo de inmunización, se sangró a blanco los conejos y se separó el suero del paquete celular. Los sueros obtenidos se adsorbieron 8 veces contra las cepas E9034AΔ*lngA*, E9034AΔ*lngB* y E9034AΔ*lngC*, respectivamente, para eliminar los anticuerpos no específicos y aumentar la especificidad de los antisueros. Los antisueros se utilizaron en los ensayos de WB, inmunofluorescencia y ensayos de microscopia electrónica con inmuno-marcaje de partículas de oro.

7.9 Cuantificación de bacterias adheridas a células intestinales HT-29

Los ensayos de adherencia se realizaron en placas de 24 pozos (Corning® Costar®, NY, EE.UU.) con una confluencia por pozo de 80% de células HT-29 (~ 1×10^5 células) en 1 ml de medio de cultivo DMEM-HG. Previamente, las bacterias se crecieron en caldo de cultivo PPLO a 37°C durante 18 h a 200 rpm. El inóculo de infección fue de 1×10^7 bacterias en un volumen de 10 μ l y la infección se llevó a cabo durante 6 h a 37°C en una atmósfera de 5% de CO₂. La inducción del promotor *lac* en las cepas complementadas se llevo a cabo con 1 mM de IPTG (Thermo Scientific, CA, EE.UU.) o para el promotor *ara* con 0.02% de arabinosa (Sigma-Aldrich Co. LLC, MO, EE.UU.) tres horas después de transcurrida la infección. Después de la infección el medio de cultivo se removió y se lavaron las células tres veces con 1 ml de PBS 1x. Las bacterias adheridas a las células se despegaron con 1 ml de 0.1 % de Tritón en PBS 1x durante 20 minutos. Las disoluciones de las bacterias despegadas se transfirieron a tubos de microcentrífuga de 1.5 ml, se realizaron diluciones seriadas de 10^{-1} a 10^{-5} de cada muestra y diez microlitros de cada dilución se sembraron en placas de agar LB, de acuerdo al método de gota en placa (Miles et al., 1938). Las placas se incubaron a 37°C durante 24 h para cuantificar las unidades formadoras de colonias (UFC) por ml.

7.10 Western blot (WB)

Los extractos proteicos totales de cada cepa se utilizaron para detectar la expresión de la proteína LngA por WB. Brevemente, las bacterias en 1 ml de PBS 1x se ajustaron a 1.0 de absorbancia a una densidad óptica de 600 nm y 70 μ l de cada suspensión bacteriana se mezcló con 25 μ l de amortiguador de Laemmli 4x y 5 μ l de β -mercaptoetanol. Las muestras se hirvieron durante 5 minutos, se centrifugaron a 5000 xg durante 1 min y 25 μ l

de cada muestra se sometieron a geles de poliacrilamida de SDS-PAGE al 16%. Las proteínas separadas en geles de poliacrilamida se transfirieron a membranas de PVDF de 6x9 cm (Millipore Corporation, MA, EE.UU.) por electrotransferencia utilizando el sistema semi-seco Trans-Blot® SD (Bio-Rad, PA, EE.UU.) con un amortiguador de transferencia. Las membranas se incubaron con 20 ml de una disolución de leche descremada al 5% en PBS-Tween (peso/volumen) en agitación, a temperatura ambiente y durante 1 h. Posteriormente, las membranas se lavaron tres veces con 20 ml de PBS-Tween por 5 min cada uno y se incubaron con suero anti-CS21 de conejo a 1:3,000 en PBS-Tween a temperatura ambiente, en agitación constante y durante 1 h. Las membranas se lavaron nuevamente con PBS-Tween y se incubaron con anticuerpos anti-IgG de conejo conjugados a peroxidasa 1:20,000 en PBS-Tween a temperatura ambiente durante 1 h. Las membranas se lavaron nuevamente con PBS-Tween y se revelaron con 1 ml de sustrato Immobilon quimioluminiscente de HRP para transferencias “Immobilon Western” (Merck Millipore, Darmstadt, Alemania) durante 2 min. La reacción se detectó en un fotodocumentador Bio-Imaging System 2.0. Los anticuerpos monoclonales anti-6His conjugados con peroxidasa se obtuvieron de Abcam (Abcam, MA, EE.UU.).

7.11 Ensayo de inmunofluorescencia (IF)

El ensamble del *pilus* CS21 de la cepa E9034A, cepas mutantes en los genes *lng* y cepas complementadas en *trans*, se visualizó por IF. Brevemente, placas de 24 pozos (Corning® Costar®, NY, EE.UU.) con cubreobjetos redondos de vidrio de 12 mm (Corning®, NY, EE. UU.), se utilizaron para crecer las células HT-29 en 1 ml de medio de cultivo DMEM-HG hasta alcanzar una confluencia de 80% (~ 1×10^5 células). Las monocapas infectadas (descrito en la sección 6.2.9) se lavaron tres veces y se fijaron con formaldehído al 2% en

PBS 1x durante 30 min a temperatura ambiente (22-25°C) ó toda la noche a 4°C. Las muestras fijadas se lavaron tres veces y se incubaron durante 5 min a temperatura ambiente con 0.1% de Tritón X-100 en PBS 1x, seguido de una incubación con RNAsa a una concentración de 20 µg/ml (Sigma-Aldrich Co. LLC, MO, EE.UU.) en PBS 1x con 10% de suero de caballo (PBS-HS) durante 30 min a 37°C. Las muestras fijadas se incubaron con suero anti-CS21 diluido 1:2,000 en PBS-HS, seguido de anticuerpos anti-IgG (H&L) de conejo conjugados con Alexa Fluor® 488 (Abcam, Cambridge, UK) a 1: 2,500 en PBS-HS e incubaron durante 1 h a 37°C. Las muestras se lavaron 3 veces con 1 ml de PBS 1x después de cada etapa. Los cubreobjetos se montaron en portaobjetos de vidrio con 3 µl de DAPI (Merck Millipore, EE.UU.) o yoduro de propidio (Sigma-Aldrich Co. LLC, MO, EE.UU.). Las imágenes de cada una de las muestras se visualizaron y procesaron en un microscopio de fluorescencia Axio Imager 2 (Zeiss, GER) y por microscopía confocal en un microscopio Leica TCS SP8X en el Laboratorio de Investigación en Inmunología y Proteómica del HIMFG, con la colaboración del Dr. Genaro Patiño López. Las imágenes se procesaron con un objetivo de 63x y un zoom digital de 5x; las cuales, se analizaron con el software Leica Application Suite; Fluorescencia Avanzada Lite Versión 2.6.3 build 8173 (LAS AF Lite; Leica Microsystems, GER).

7.12 Microscopía electrónica de transmisión (TEM) y microscopía inmunoelectrónica (IEM)

Las rejillas de cobre de 300 cuadrantes (Electron Microscopy Sciences, PA, EE.UU.) recubiertos con formvar al 1% (Sigma-Aldrich Co. LLC, MO, EE.UU.) se utilizaron para visualizar los *pili* CS21 por TEM e IEM. Diez microlitros de los cultivos bacterianos se colocaron sobre las rejillas durante 5 min hasta adsorberse y el exceso de líquido se

absorbió con papel de filtro. Posteriormente, las muestras se incubaron con 10 µl de ácido fosfotúngstico al 1% pH 7.4 (Electron Microscopy Sciences, PA, EE. UU.) durante 4 min y el exceso se eliminó nuevamente con papel de filtro después de la tinción negativa. Las rejillas se enjuagaron con dos gotas de agua destilada, se secaron al aire y se examinaron por TEM utilizando un microscopio electrónico JEM-1010 (JEOL, Tokio, Japón) en la Central de Instrumentación de Microscopía de la Escuela Nacional de Ciencias Biológicas del Instituto Politécnico Nacional, con la colaboración del Dr. Edgar Oliver López Villegas y la Dra. María del Rosario Espinosa Mellado.

Los ensayos de IEM se realizaron para confirmar la identidad de los *pili* CS21 incluyendo a la proteína LngB en la cepa E9034A, cepas mutantes y cepas complementadas. Las rejillas de cobre cubiertas con formvar y las bacterias adsorbidas se incubaron con suero anti-CS21 o anti-LngB (diluido 1:10 en PBS con 10% de BSA) durante 1 h. Las muestras se enjuagaron con una gota de PBS, tres veces, seguido de 1 h de incubación con anticuerpos anti-IgG de conejo H&L conjugados con partículas de oro de 10 nm (Sigma-Aldrich Co. MO, EE.UU.). Las muestras se lavaron nuevamente con PBS y se tñieron negativamente con ácido fosfotúngstico al 1% pH 7.4.

7.13 Ensayo de formación de biopelículas

Las cepas silvestres ETEC E9034A, EAEC 042 y mutantes en los genes *lng* se examinaron para determinar su capacidad para formar biopelículas sobre placas de 96 pozos (Corning® Costar®, NY, EE.UU.). Las cepas bacterianas se cultivaron en caldo LB durante toda la noche y diez microlitros de estos cultivos se incubaron en cada pozo con 200 µl de DMEM-HG y/o caldo PPLO a 37 °C en una atmósfera de 5% de CO₂ y durante 48 h. Los pozos se lavaron tres veces con PBS 1x y las biopelículas se fijaron con formaldehído al 2% en PBS

1x a 4°C durante 16 h. Para el análisis cualitativo, las muestras se lavaron tres veces con PBS 1x, se tiñeron con una solución de cristal violeta al 0.1% en agua y se visualizaron bajo un microscopio invertido. Para el análisis cuantitativo, el colorante absorbido por las bacterias se eluyó con 200 µl de metanol absoluto y la absorbancia se determinó en un lector de ELISA Multiskan FC (Thermo Scientific, CA, EE.UU.) a una longitud de onda de 620 nm. Los ensayos se realizaron por triplicado en dos días diferentes. La cepa EAEC 042 fue usada como control positivo y la cepa *E. coli* K-12 DH5α como control negativo. Los pozos con medio de cultivo sin infectar se utilizaron como controles de fondo de la tinción y el colorante adsorbido se restó de las muestras problema.

7.14 Ensayo de auto-agregación bacteriana

La cepa silvestre E9034A, cepas mutantes y cepas complementadas en *trans* se cultivaron en 5 ml de caldo TB a 37°C en agitación a 250 rpm y durante 18 h. Los cultivos bacterianos se diluyeron 1:100 en 1 ml de medio de cultivo TB y se cultivaron en estático en placas de cultivo de 24 pozos de fondo plano (Corning, NY, EE.UU.) a 37°C durante 4 h. Los agregados bacterianos se visualizaron directamente sobre las placas de cultivo utilizando un microscopio de luz invertida (Olympus, PA, EE.UU.). El fenotipo de la auto-agregación bacteriana se observó como agregados bacterianos sedimentados en el fondo de la placa de cultivo y aclaramiento del medio de cultivo. Tres ensayos por triplicados se realizaron en tres momentos diferentes. La cepa de ETEC E9034A y la cepa K-12 DH5α se utilizaron como controles positivos y negativos, respectivamente. El análisis cuantitativo de los sobrenadantes sin perturbar los agregados bacterianos, se realizó midiendo la absorbancia a 600 nm utilizando un espectrofotómetro 2802 UV/VIS Spectrophotometer (Unico Leyvitec Laboratorios, D.F. México).

7.15 Ensayos de citometría de flujo

La citometría de flujo se utilizó para evaluar cuantitativamente la producción de los *pili* CS21 en las cepas de ETEC. Los sobrenadantes de la infección de las células HT-29 se recolectaron y alícuotas de 100 μ l se centrifugaron a 3500 x g durante 5 min. Las pastillas generadas se resuspendieron con 50 μ l del suero anti-CS21 (1:500 en PBS 1x) durante 1 h en hielo. Las muestras se lavaron con 1 ml de PBS 1x frío y se centrifugaron a 3500 x g durante 5 min, el sobrenadante se decantó y los paquetes bacterianos se lavaron dos veces más. Los paquetes bacterianos se incubaron con 10 μ l de los anticuerpos anti-IgG (H&L) de conejo conjugado con Alexa Fluor® 488 (Abcam, Cambridge, UK) 1:500 en PBS 1x y se incubaron en hielo durante 1 h. Las muestras se lavaron nuevamente, se resuspendieron en 1 ml de PBS 1x y se leyeron en un citómetro de flujo MoFlo® Astrios™ EQ. (Beckman Coulter, Life Sciences, IN, EE.UU.) en el Laboratorio de Investigación en Inmunología y Proteómica del HIMFG, con la colaboración del Dr. Ricardo Valle Ríos. La excitación fue 488 nm y la emisión de fluorescencia isotiocianato de fluoresceína se registró a través de un filtro de paso de 513/26 nm con 50,000 eventos.

7.16 Análisis estadístico

Los datos generados de los ensayos de adherencia, formación de biopelículas y auto-agregación bacteriana se analizaron mediante la prueba *t* de Student (prueba *t* no pareada) utilizando el Software GraphPad Inc. (La Jolla, CA, EE.UU.). El valor de *p* utilizada en el estudio fue ≤ 0.005 como punto de referencia de una diferencia estadística significativa.

7.17 Análisis Bioinformáticos

Los análisis bioinformáticos se realizaron utilizando las siguientes páginas WEB:

- BLAST, programa de alineamiento de secuencias de tipo local.
<http://blast.ncbi.nlm.nih.gov/Blast.cgi>
- OligoAnalyzer v3.1, análisis de oligonucleótidos.
<https://www.idtdna.com/calc/analyzer>
- Oligo Calc, calculador de propiedades de los oligonucleótidos.
<http://biotools.nubic.northwestern.edu/OligoCalc.html>
- Melting Temperature TM calculation, calculador de la T_m de los oligonucleótidos (6-50 bases)
http://www.biophp.org/minitools/melting_temperature/demo.php?formula=basic
- ORFfinder, buscador de marcos de lectura abiertos.
<https://www.ncbi.nlm.nih.gov/orffinder/>
- Clustal Omega, programa de alineamiento de secuencias múltiple.
<http://www.ebi.ac.uk/Tools/msa/clustalo/>
- MUSCLE, programa de alineamiento de secuencias múltiple.
<http://www.ebi.ac.uk/Tools/msa/muscle/>
- CELLO v2.5, predictor de localización subcelular.
<http://cello.life.nctu.edu.tw/>
- PSLpred, predictor de localización subcelular.
<http://www.imtech.res.in/raghava/pslpred/>
- PSORTb v3.0.2, predictor de localización subcelular.
<http://www.psort.org/psortb/>
- PROSITE, base de datos de dominios, familias y función de proteínas.
<http://prosite.expasy.org/prosite.html>
- Conserved Domain Database (CDD), base de datos de dominios conservados.
<http://www.ncbi.nlm.nih.gov/cdd>
- TMHMM v2.0, predictor de hélices transmembranales en proteínas.
<http://www.cbs.dtu.dk/services/TMHMM/>
- SOSUI, predictor para discriminar proteínas de membrana de las solubles junto con la predicción de hélices transmembranales.
http://harrier.nagahama-i-bio.ac.jp/sosui/sosui_submit.html

- CCTOP, predictor de hélices transmembranales, localización y orientación.
http://cctop.enzim.ttk.mta.hu/?_=
- Jpred4, predictor de estructura secundaria.
<http://www.compbio.dundee.ac.uk/jpred/>
- PORTER, predictor de estructura secundaria.
<http://distill.ucd.ie/porter/>
- SignalP 4.1, predictor de péptido señal y de sitio de corte.
<http://www.cbs.dtu.dk/services/SignalP/>
- InterPro, predictor de péptido señal y de sitio de corte.
<http://www.ebi.ac.uk/interpro/>
- PrediSi, predictor de péptido señal.
<http://www.predisi.de/>
- Phobius, predictor de péptido señal y de sitio de corte.
<http://phobius.Sbc.su.se/>
- ARNold, predictor de terminadores transcripcionales independientes de Rho.
<http://rna.igmors.u-psud.fr/>
- UNAFold, predictor de la estructura secundaria y energía libre de Gibbs de las secuencias de ARN y ADN.
<http://unafold.rna.albany.edu/>
- NEBcutter v2.0, análisis de marcos de lectura y sitios de restricción enzimática
<http://nc2.neb.com/NEBcutter2/>

8.0 RESULTADOS

8.1 Ensamble del *pilus* CS21

8.1.1 Proteínas involucradas en la biogénesis de los *pili* TCP, CFA/III, BFP y CS21

La biogénesis de los *pili* TCP, CFA/III, BFP y CS21 requiere entre 11 y 14 genes, localizados dentro de los operones *tcp*, *cof*, *bfp* y *lng*, respectivamente (Roux et al., 2012). La similitud entre las proteínas que participan en el ensamble de los *pili* CS21 y CFA/III, sugiere que ambos apéndices estructurales son parálogos, y por lo tanto, tienen un ancestro en común (Gomez-Duarte et al., 2007). Las secuencias de las proteínas LngA de diferentes cepas clínicas de ETEC, se alinearon con las proteínas CofA, TcpA, BfpA y PilA de otros TIVP y el pseudopilus GspG del sistema de secreción tipo II, para generar un árbol filogenético de máxima parsimonia, utilizando el programa MEGA 6.06 (**Figura 3**). Para determinar que los clados fueran confiables se realizó un análisis de bootstrap de 500 replicas. El árbol consenso con análisis bootstrap inferido a partir de 500 repeticiones se tomo para representar la historia evolutiva de los taxones analizados (**Figura 3**). La distribución de la proteína LngA en el árbol filogenético, mostró una relación más cercana con las proteínas CofA y TcpA; mientras, con las proteínas BfpA y PilA mostró una relación más lejana.

El ensamble de los *pili* TCP, CFA/III y BFP involucra varias proteínas; las cuales, son homólogas a proteínas codificadas en el operón de *lng*, como se describe en la **tabla 3**. Los operones que contienen genes que codifican para proteínas involucradas en el ensamble de CS21, CFA/III y TCP, mostraron una organización genética en común, tamaño y todos ellos son sinténicos y se transcriben en la misma dirección; así mismo, un porcentaje alto de identidad se observó entre productos génicos de CS21 con CFA/III y CS21 con TCP

(Figura 4). Las proteínas LngA, CofA y TcpA son las moléculas estructurales principales de los *pili* CS21, CFA/III y TCP, respectivamente (Roux et al., 2012). La proteína LngA tiene una identidad a nivel de aminoácidos del 79% y 35% con las proteínas CofA y TcpA, respectivamente. La proteína LngB es una molécula requerida para iniciar el ensamble del *pilus* CS21 y el 78% de su secuencia primaria es idéntica a la proteína CofB. La región C-terminal de CofB se requiere para iniciar el ensamblaje del *pilus* CFA/III de ETEC (Kolappan et al., 2015).

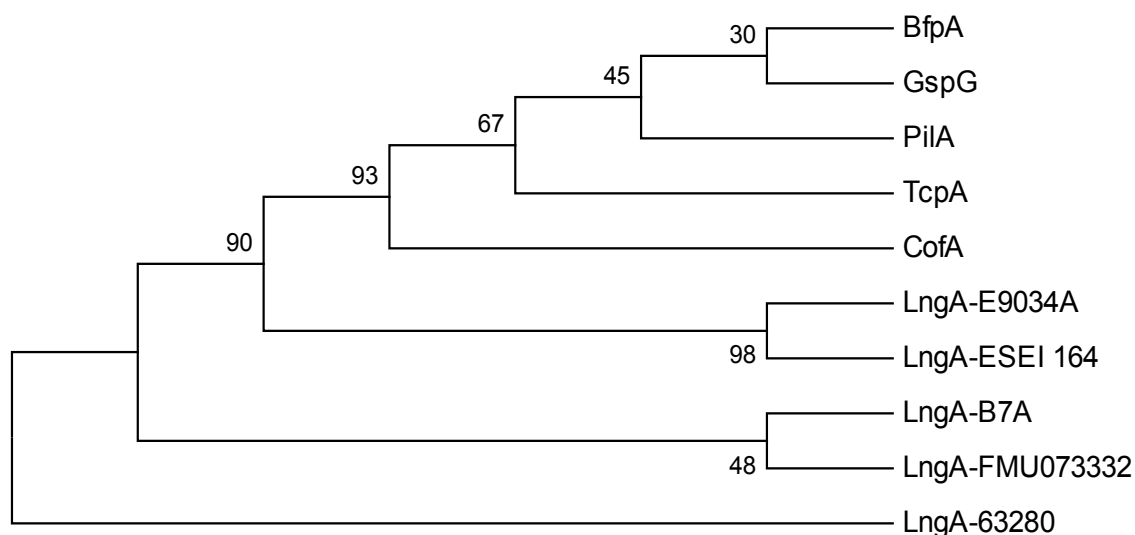


Figura 3. Divergencia entre las proteínas LngA, CofA, TcpA y BfpA de los diferentes TIVPb, PilA de los TIVPa y el pseudopilus GspG del sistema de secreción tipo II. La historia evolutiva fue inferida utilizando el método de máxima parsimonia. El árbol consenso del análisis bootstrap inferido a partir de 500 repeticiones se tomó para representar la historia evolutiva de los taxones analizados. El árbol de máxima parsimonia se obtuvo utilizando el algoritmo Subtree-Pruning-Regrafting (SPR) con el nivel de búsqueda 1 en el que los árboles iniciales se obtuvieron mediante la adición aleatoria de secuencias (10 repeticiones). El análisis involucró 10 secuencias de aminoácidos. Se eliminaron todas las posiciones que contenían vacíos y datos perdidos. Había un total de 62 posiciones en el conjunto de datos final. El porcentaje de árboles replicados en el que los taxones asociados agrupados de la prueba de bootstrap (500 repeticiones), se muestran junto a las ramas.

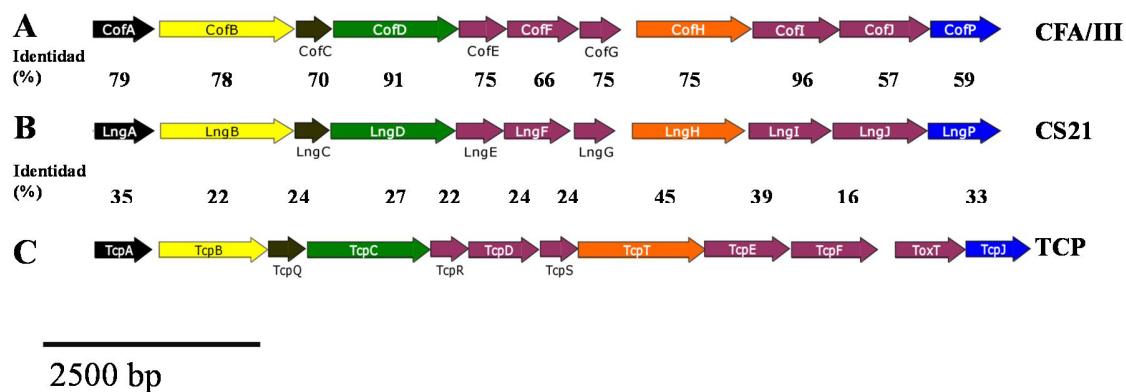


Figura 4. Organización y comparación de las proteínas estructurales codificadas en los operones de *cof* (A), *lng* (B) y *tcp* (C). La orientación de las flechas indica la dirección de la transcripción de los genes codificados en los operones y los porcentajes de identidad de la secuencias aminoacídicas entre: CS21 con CFA/III y CS21 con TCP.

8.1.2 Predicción *in silico* de la localización subcelular de las proteínas Lng

La búsqueda de similitud entre las secuencias reportadas en las bases de datos, es una estrategia empleada para asignar una función a las proteínas, incluyendo la localización subcelular e identificación de motivos y dominios conservados. Los numerosos algoritmos desarrollados para la predicción de localización subcelular, se emplean comúnmente por los diferentes programas bioinformáticos (Reinhardt y Hubbard, 1998; Boden y Hawkins, 2005; Matsuda et al., 2005). Los programas CELLO v.2.5, PSLpred y PSORTb v3.0.2, emplean diferentes algoritmos de aprendizaje supervisado, denominados “máquinas de soporte vectorial” (SVM, por sus siglas en inglés) para predecir la localización subcelular de proteínas en bacterias Gram negativas (Yu et al., 2004). Los programas predicen cinco principales sitios de localización subcelular en bacterias Gram negativas, tales como: citoplasma, membrana interna, membrana externa, periplasma y espacio extracelular. Además, incluyen varios módulos SVM basados en diferentes rasgos de las proteínas, tales

como: composición de aminoácidos, composición de dipéptidos, propiedades fisicoquímicas e información evolutiva de PSI-Blast.

Tabla 3. Función de las proteínas del ensamble de los *pili* CS21, CFA/III y TCP.

CS21			CFA/III			TCP		
Proteína	Función	Referencia	Proteína	Función	Referencia	Proteína	Función	Referencia
LngA	Proteína estructural mayor	(Giron et al., 1994)	CofA	Proteína estructural mayor	(Taniguchi et al., 1994; Taniguchi et al., 1995)	TcpA	Proteína estructural mayor	(Taylor et al., 1987)
LngB	Proteína estructural menor	(Kolappan et al., 2015)	CofB	Proteína estructural menor, Inicia ensamble del <i>pilus</i>	(Kolappan et al., 2015),	TcpB	Desconocida	
LngC	Desconocida		CofC	Desconocida		TcpQ	Requerida para la localización de TcpC a la membrana externa	(Bose y Taylor, 2005)
LngD	Lipoproteína de membrana externa*	(Gomez-Duarte et al., 2007)	CofD	Lipoproteína de membrana externa *	(Taniguchi et al., 2001)	TcpC	Lipoproteína de membrana externa	(Bose y Taylor, 2005)
LngE	Desconocida		CofE	Desconocida		TcpR	Proteína de membrana interna, bitópica *	(Tripathi y Taylor, 2007)
LngF	Desconocida		CofF	Desconocida		TcpD	Proteína de membrana interna, bitópica *	(Roux et al., 2012)
LngG	Desconocida		CofG	Desconocida		TcpS	Desconocida	
LngH	Proteína de unión a nucleótido, ATPasa *	(Gomez-Duarte et al., 2007)	CofH	Proteína de unión a nucleótido, ATPasa *	(Taniguchi et al., 2001)	TcpT	ATPasa	(Iredell y Manning, 1997)
LngI	Desconocida		CofI	Desconocida		TcpE	Proteína de membrana interna Politópica*	(Roux et al., 2012)
LngJ	ATPasa, factor de colonización *	(Gomez-Duarte et al., 2007)	CofJ	Factor de colonización potencial	(Yuen et al., 2013)	TcpF	Factor de colonización, extracelular	(Kirn et al., 2003)
LngP	Prepilina peptidasa*	(Gomez-Duarte et al., 2007)	CofP	Prepilina peptidasa	(Taniguchi et al., 1999)	TcpJ	Prepilina peptidasa	(Kaufman et al., 1991; LaPointe y Taylor, 2000)

*Predicción de la función de las proteínas a través de análisis bioinformáticos.

La localización subcelular de las proteínas LngR, LngS, LngT, LngA, LngB, LngC, LngD, LngE, LngF, LngG, LngH, LngI, LngJ y LngP se muestra en la **tabla 4**. Las escalas de certidumbre de la predicción fueron de 0-5, 0-100 y 0-10 para los programas CELLO, PSLpred y PSORTb, respectivamente. Los datos obtenidos de las predicciones correlacionan parcialmente con la localización de las proteínas homólogas involucradas en el ensamble de los *pili* CFA/III y TCP (**Tabla 3**).

Tabla 4. Localización celular de las proteínas Lng en la biogénesis del *pilus* CS21.

Proteína	Localización subcelular					
	CELLO	Escala 0-5	PSLpred (Método híbrido)	Escala 0-100%	PSORTb	Escala 0-10
LngR	Citoplasma	3.631	Citoplasma	68.1	Citoplasma	9.26
LngS	Citoplasma	3.577	Citoplasma	90.2	Citoplasma	8.96
LngT	Membrana interna/ Periplasma/ Citoplasma	1.568/ 1.315/ 1.310	Citoplasma	68.1	Periplasma	9.44
LngA	Extracelular	2.641	Extracelular	53.1	Extracelular	9.71
LngB	Extracelular	2.973	Extracelular	71.1	Desconocida	
LngC	Membrana externa	2.329	Membrana externa	53.1	Desconocida	
LngD	Membrana externa	4.611	Membrana externa	98.1	Membrana externa	9.92
LngE	Citoplasma	2.544	Citoplasma	68.1	Desconocida	
LngF	Membrana externa	3.421	Extracelular	68.1	Desconocida	
LngG	Membrana externa/ Periplasma	1.843/1.638	Proteína de membrana interna	53.1	Desconocida	
LngH	Citoplasma	3.610	Citoplasma	98.1	Citoplasma	9.97
LngI	Membrana interna	4.125	Membrana interna	98.1	Membrana interna	10
LngJ	Membrana externa	3.561	Extracelular	53.1	Extracelular	9.65
LngP	Membrana interna	4.744	Membrana interna	71.1	Membrana interna	10

La arquitectura global de los TIVP involucra la organización de diversas proteínas que forman la maquinaria de su ensamblaje, dividida en 4 subcomplejos: subcomplejo de membrana interna (proteínas bitópicas y politópicas), subcomplejo de membrana externa (secretina y proteínas accesorias), subcomplejo motor (ATPasa) y el propio *pilus* (pilinas mayor y menor) (Burrows, 2012). Las proteínas de membrana (interna y externa), tienen una función importante durante el ensamble de los TIVP, pueden formar hélices transmembranales en la bicapa lipídica y poros en la membrana externa. La presencia, número y localización exacta de las hélices transmembranales es importante para dirigir el análisis funcional de las proteínas bitópicas y politópicas; las cuales, están formadas por extensiones largas de residuos hidrofóbicos (Sonnhammer et al., 1998). Las secuencias de las proteínas Lng se analizaron con los programas TMHMM, SOSUI y CCTOP a través de un análisis de hidrofobicidad. Los programas predijeron hélices transmembranales en las proteínas LngA, LngB, LngE, LngF y LngP localizadas en la membrana interna (**Tabla 5**).

Los predictores ocasionalmente están limitados en discriminar entre el péptido señal y las hélices transmembranales debido a la alta similaridad de sus propiedades físico-químicas. Las secuencias de las proteínas Lng se sometieron a análisis bioinformáticos utilizando los predictores: SignalP4.1, PrediSi y Phobius. Los resultados mostraron varios posibles péptidos señal en las proteínas LngC, LngD, LngG y LngJ sin la presencia de cruces transmembranales (**Tabla 6**).

Tabla 5. Predicción de cruces transmembranales de las proteínas involucradas en la biogénesis de CS21.

Proteína (longitud, aa)	Programas Cruces transmembranales					
	TMHMM (#)	Posición	SOSUI (#)	Posición	CCTOP (#)	Posición
LngR(101)	0	-	0	-	0	-
LngS (298)	0	-	0	-	1	166-188
LngT (147)	0	-	0	-	0	-
LngA (236)	1	32-54	1	33-55	1	32-53
LngB (523)	1	7-26	1	5-27	1	6-26
LngC (137)	0	-	1	4-26	0	-
LngD (984)	0	-	0	-	5	79-99 177-197 275-295 669-689 767-787
LngE (186)	1	161-180	1	161-183	1	161-180
LngF (257)	1	20-42	1	19-40	1	20-40
LngG (161)	0	-	1	2-24	1	102-122
LngH (444)	0	-	0	-	2	180-195 285-300
LngI (322)	3	90-112 143-165 291-313	3	95-117 146-168 288-310	3	99-115 145-165 294-313
LngJ (373)	0	-	0	-	0	-
LngP (284)	7	7-29 86-105 112-131 141-163 168-187 207-229 245-267	7	7-29 88-109 113-135 138-160 165-187 203-225 244-266	8	5-26 86-105 111-131 141-160 166-187 194-213 216-231 248-269

aa; aminoácidos

Tabla 6. Predicción de péptido señal de las proteínas Lng involucradas en la biogénesis de CS21.

Proteína (Longitud en aa)	Péptido señal					
	SignalP 4.1 Escala de 0 a 1	Posición	PrediSi Escala de 0 a 1	Posición	Phobius	Posición
LngR (101)	No	-	No	-	No	-
LngS (298)	No	-	No	-	No	-
LngT (147)	Si (0.787)	1-20	No	-	Si	1-20
LngA (236)	No	-	No	-	No	-
LngB (523)	No	-	No	-	No	-
LngC (137)	Si (0.910)	1-20	Si (0.9017)	1-20	Si	1-20
LngD (984)	Si (0.570)	1-20	No	-	Si	1-26
LngE (186)	No	-	No	-	No	-
LngF (257)	No	-	No	-	No	-
LngG (161)	Si (0.570)	1-25	Si (0.8113)	1-25	Si	1-25
LngH (444)	No	-	No	-	No	-
LngI (322)	No	-	No	-	No	-
LngJ (373)	Si (0.850)	1-22	Si (0.7276)	1-22	Si	1-27
LngP (284)	No	-	No	-	No	-

aa; aminoácidos

8.1.3 Mutagénesis de los genes *lngR*, *lngS*, *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, *lngP* y *hns*

Los genes *lngR*, *lngS*, *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, *lngP* y *hns* se mutaron mediante la técnica de inactivación en un paso, utilizando productos de PCR de los genes *kan* (kanamicina) o *cat* (cloranfenicol); los cuales, se flanquean por secuencias homólogas al gen seleccionado para la mutación (Datsenko y Wanner, 2000). Los oligonucleótidos específicos para amplificar el producto de PCR, se diseñaron manualmente con un tamaño de 39 a 60 nucleótidos de longitud de las regiones homólogas a los extremos de los genes

de interés de la cepa E9034A incluyendo los sitios P1 y P2 del casete de resistencia (**Figura 5**) (**Tabla 8**). La eliminación y la sustitución de los genes *lngR*, *lngS*, *lngA*, *lngB*, *lngC*, *lngD*, *lngH* *lngP* y *hns*, se confirmó utilizando oligonucleótidos específicos a través de ensayos de PCR punto final (**Tabla 8**). Las cepas mutantes se complementaron en *trans* con las distintas construcciones que contenía el gen mutando bajo el control de un promotor *lacZ* (**Tabla 7**). Los plásmidos pUC18, pUC19 ó pBAD-Topo sin los genes *lng*, se transformaron en cada una de las mutantes y se utilizaron como controles negativos de las construcciones. Además, permitieron verificar que las complementaciones de las cepas mutantes con las construcciones conteniendo a los genes de interés hayan restaurado el fenotipo observado en la cepa E9034A.

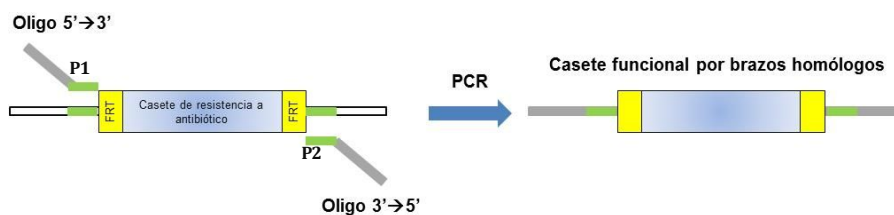


Figura 5. Amplificación del casete de resistencia de Km y/o Cm. Los oligonucleótidos específicos para la mutagénesis se diseñaron de 50-70 pb de longitud, incluyendo las regiones homólogas del gen de interés (recuadro gris) y la región de 20 pb del plásmido pKD4 con los sitios P1 y P2 (recuadro verde). El casete de resistencia de Km y/o Cm contiene las secuencias FRT (sitios blancos de reconocimiento de la proteína recombinasa FLP) en ambos extremos.

Los genes mutados *lngS*, *lngR*, *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, *lngP* y *hns*, se confirmaron mediante PCR punto final utilizando oligonucleótidos específicos que alinearon, corriente arriba y corriente abajo del sitio de recombinación (**Figura 6 y 7**).

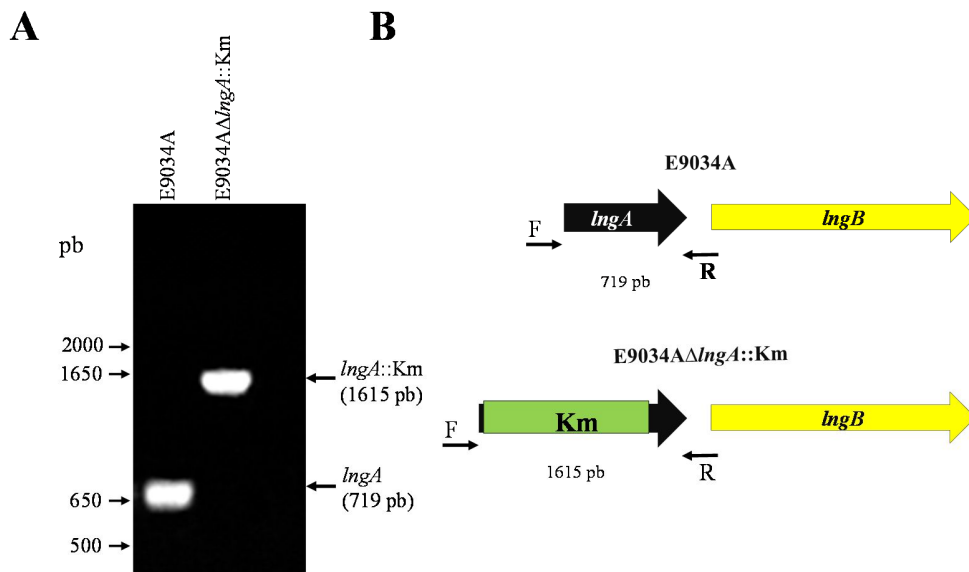


Figura 6. Confirmación de la mutación del gen *lngA* en la cepa E9034A. A) Fotografía de un gel de agarosa al 1% mostrando los amplicones obtenidos por PCR punto final, confirmando la mutación del gen *lngA* en la cepa E9034AΔ*lngA*::Km. B) Diagramas de la inserción del casete de Km en el gen *lngA*.

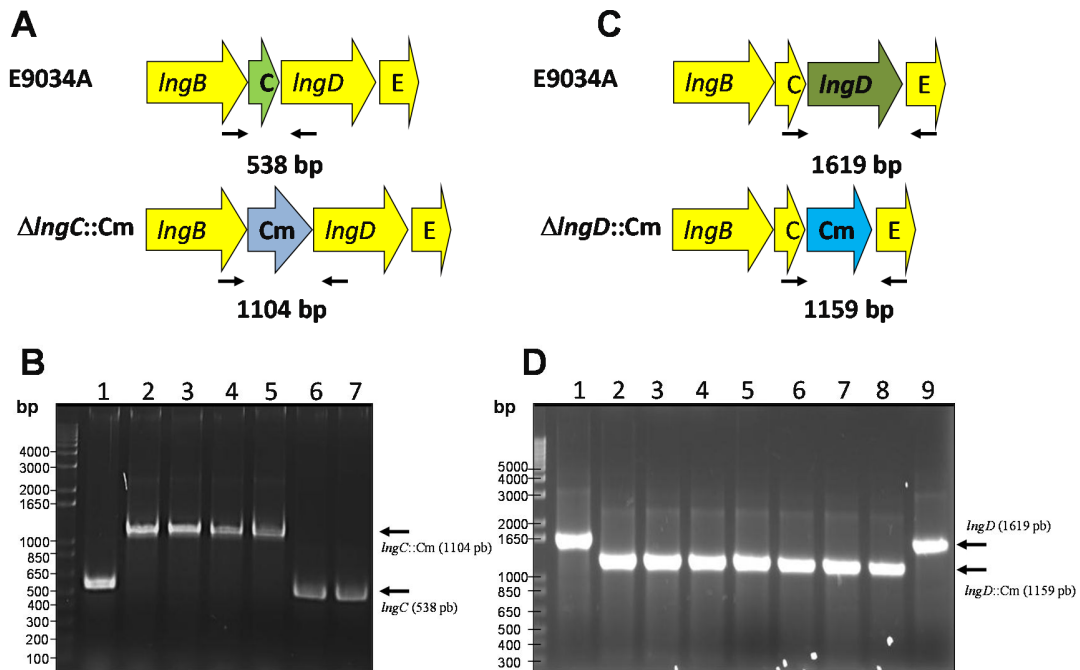


Figura 7. Confirmación de la mutación de los genes *lngC* y *lngD*. Diagramas de la inserción del casete de Cm en los genes *lngC* (A) y *lngD* (C). B y D) Las fotografías de los gels de agarosa mostraron amplicones de 538 pb, 1104 pb, 1619 pb y 1159 pb que corresponden a los genes *lngC*, *lngD*::cm, *lngD*, *lngD*::cm, respectivamente.

8.1.4 Re-introducción del gen silvestre en cada una de las mutantes generadas

La re-introducción del gen silvestre deletado en las cepas mutantes ($\Delta lngR::Km$, $\Delta lngS::Km$, $\Delta lngA::Km$, $\Delta lngB::Cm$, $\Delta lngC::Cm$, $\Delta lngD::Cm$, $\Delta lngH::Cmy$ $\Delta lngP::Cm$), se realizó en *trans* utilizando oligonucleótidos específicos, para asegurar los fenotipos en cada una de las cepas (**Tabla 8**). Los genes *lngS*, *lngR*, *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* se amplificaron con la enzima *Platinum® Taq DNA Polymerase, High Fidelity*. Los oligonucleótidos empleados para la reacción de amplificación se sintetizaron con los sitios de restricción mostrados en la **tabla 8**. Los amplicones se digirieron con las enzimas de restricción correspondientes (EcoRI, BamHI, HindIII, SacI, SalI y SmaI) y se ligaron en los vectores de clonación (pUC18, pUC19 y pBAD-Topo). Las construcciones generadas se analizaron por PCR punto final para confirmar la clonación y orientación de los amplicones en los vectores, utilizando oligonucleótidos que alinean en el inserto y en el vector de clonación. A las cepas mutantes generadas en este estudio se les introdujo exitosamente por electroporación los plásmidos generados. Por ejemplo, la mutación correspondiente a la inserción del casete de Cm y la presencia del gen *lngB* silvestre, se muestra en **figura 8**. El diagrama de la inserción del casete Cm en el gen *lngB*, junto con los amplicones esperados se muestra panel A de la **figura 8**. Adicionalmente, en las mutantes del gen *lngB* se obtuvieron amplicones de 1206 pb por PCR punto final que confirman la mutación del gen *lngB* en estas clonas. La cepa E9034A mostró un amplicon de 1768 pb que corresponde al gen *lngB* más la inclusión de las regiones adyacentes (**Figura 8B**). El mapa del vector pUClngB utilizado para introducir el gen silvestre *lngB* a la cepa E9034A $\Delta lngB::Cm$ se muestra en la **figura 8C**. La presencia de los alelos mutante y silvestre en la cepa complementada E9034A $\Delta lngB$ (pUClngB) se confirmaron mediante PCR punto final y se

observo la presencia de dos amplicones de 1768 y 1206 pb que corresponden al gen *lngB* silvestre (presente en el plásmido) y a la inserción del casete de Cm (*lngB*::Cm), respectivamente (Figura 8D).

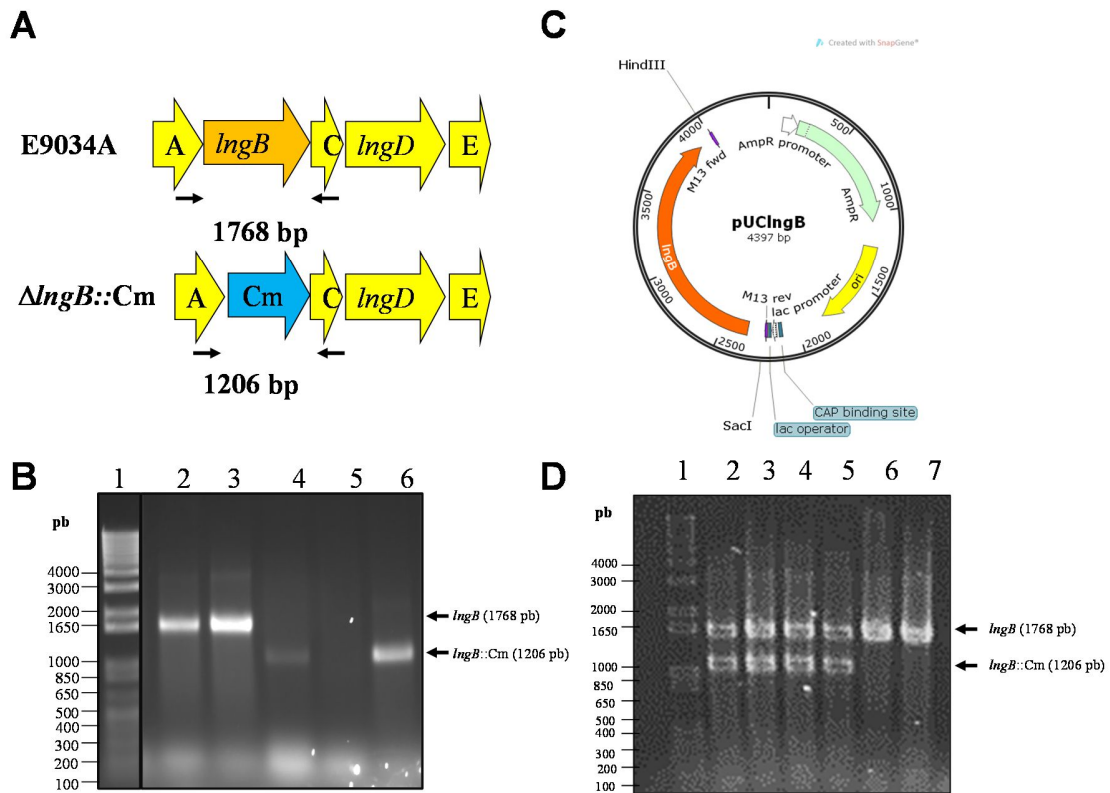


Figura 8. Confirmación de la mutación del gen *lngB* y re-introducción del gen silvestre en la cepa E9034A Δ *lngB*. A) Diagrama de la inserción del casete de Cm en el gen *lngB*. B) Gel agarosa al 1% teñido con bromuro de etidio mostrando los perfiles encontrados en 5 clones después de la mutación del gen *lngB*. En los carriles 2 y 3, se observa una banda de 1768 pb que corresponde al gen *lngB* incluyendo las secuencias que flanquean al gen. En los carriles 4 y 6 se observa una banda de 1206 pb que corresponde a la inserción del casete de Cm y de la mutación del gen *lngB*. C) Mapa del vector pUC18 incluyendo el gen *lngB* (pUClngB). D) Gel de agarosa mostrando los perfiles encontrados en 4 clones complementados: bandas de 1206 y 1768 pb que corresponden a los amplicones del gen *lngB* deletado y del gen *lngB* presentes en el plásmido (carriles 2 a 5), respectivamente. El vector pUClngB que se muestra en el carril 6 y la cepa silvestre E9034A en el carril 7 se utilizaron como controles positivos.

8.1.5 Mutación de los genes *IngA*, *IngB*, *IngC*, *IngD*, *IngH* y *IngP* en ETEC E9034A afecta la adherencia a las células intestinales HT-29 y la auto-agregación bacteriana

La producción de los *pili* CS21 en ETEC E9034A se favorece cuando se cultiva en caldo PPLO a 37°C, indicado que su expresión está regulado bajo condiciones nutricionales (Mazariego-Espinosa et al., 2010). Las cepas mutantes generadas en este estudio se cultivaron en caldo PPLO a 37°C a diferentes tiempos y sus absorbancias se representaron en una curva de crecimiento; la cual, no mostró efectos negativos en la tasa de crecimiento bacteriano. Es importante hacer énfasis que la todas las cepas mutadas incluyendo la cepa E9034A mostraron la misma tasa de crecimiento (**Figura 9**).

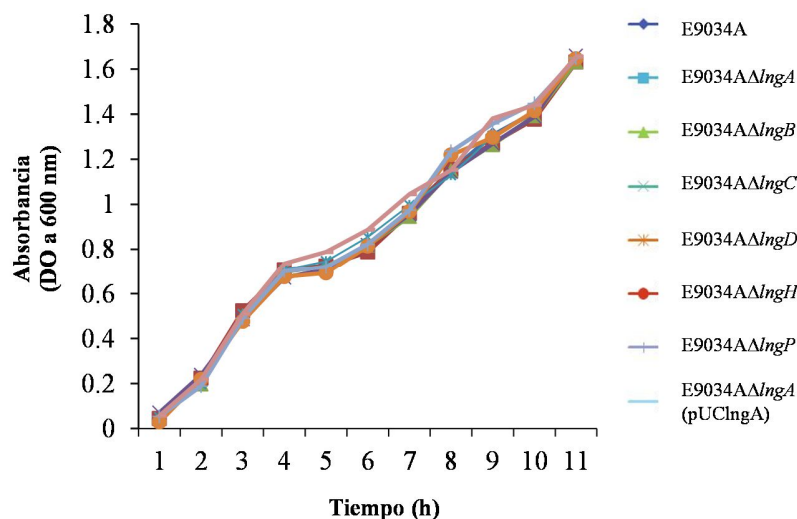


Figura 9. Curva de crecimiento de las cepas mutantes en los genes *Ing*.

El *pilus* CS21 de ETEC contribuye significativamente a la adherencia a células intestinales HT-29 (Mazariego-Espinosa et al., 2010). Los análisis bioinformáticos realizados en este estudio, sugieren que las proteínas LngA, LngB, LngC, LngD, LngH y LngP son importantes en el ensamble del *pilus* CS21. Las mutantes generadas se caracterizaron fenotípicamente mediante ensayos de adherencia a las células intestinales

HT-29, agregación bacteriana y formación de biopelículas. Los ensayos de adherencia se realizaron infectando a las monocapas de células intestinales HT-29, con la cepa silvestre y con las mutantes generadas con una MOI de 1:100. Los tipos y porcentajes de adherencia se analizaron cuali- y cuantitativamente.

Análisis cuantitativos. La determinación de las UFC por el método de Miles y Misra, mostró una reducción significativa (valor de $p < 0.005$, datos analizados por la prueba *t student*) en la adherencia entre el 85% a 91% en todas la cepas de ETEC mutadas (en los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP*), comparado con la cepa silvestre E9034A (**Figura 10**).

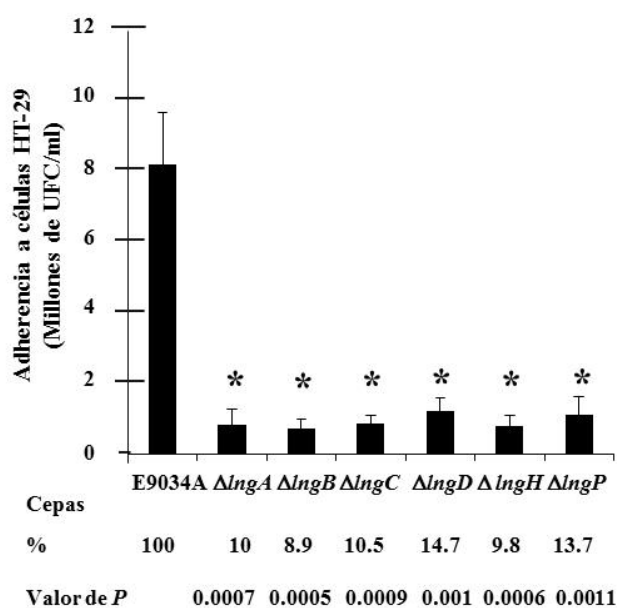


Figura 10. Mutación de los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* en la cepa ETEC E9034A afecta la adherencia a células intestinales HT-29. Las bacterias adheridas a las células HT-29 se cuantificaron y se graficaron los promedios de 6 determinaciones de las UFC \pm DS por ml. Los asteriscos indican las diferencias significativas ($p < 0.005$) entre la cepa silvestre E9034A con las cepas mutantes.

Las cepas mutantes en *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP*, se complementaron con estos mismos genes clonados en los vectores de expresión, para restaurar el fenotipo observado en la cepa silvestre E9034A. El fenotipo de cada una de ellas se determinó realizando ensayos de adherencia; los cuales, se analizaron cualitativamente por microscopía de luz y cuantitativamente por el conteo de las UFC.

Análisis cualitativos. Las bacterias adheridas (cepas de ETEC incluidas en este estudio), se fijaron a las monocapas de células presentes en cubreobjetos, se lavaron con PBS 1x, se tiñeron con Giemsa al 0.1% y se visualizaron por microscopía de luz. Las imágenes obtenidas por microscopia de luz de las mutantes, mostraron un número reducido de bacterias adheridas a las células HT-29, comparadas con las bacterias de la cepa silvestre E9034A; además, se observaron patrones de adherencia de tipo agregativo en todas las cepas de ETEC (**Figura 11**). El patrón de adherencia observado en cada una de las cepas examinadas, sugiere que es independiente de la expresión del *pilus* CS21 de ETEC E9034A. Los análisis cuantitativos de los ensayos de adherencia de las cepas de ETEC (cepa silvestre y cepas mutantes), correlacionan con los resultados obtenidos de los análisis cualitativos (**Figura 10 y 11**).

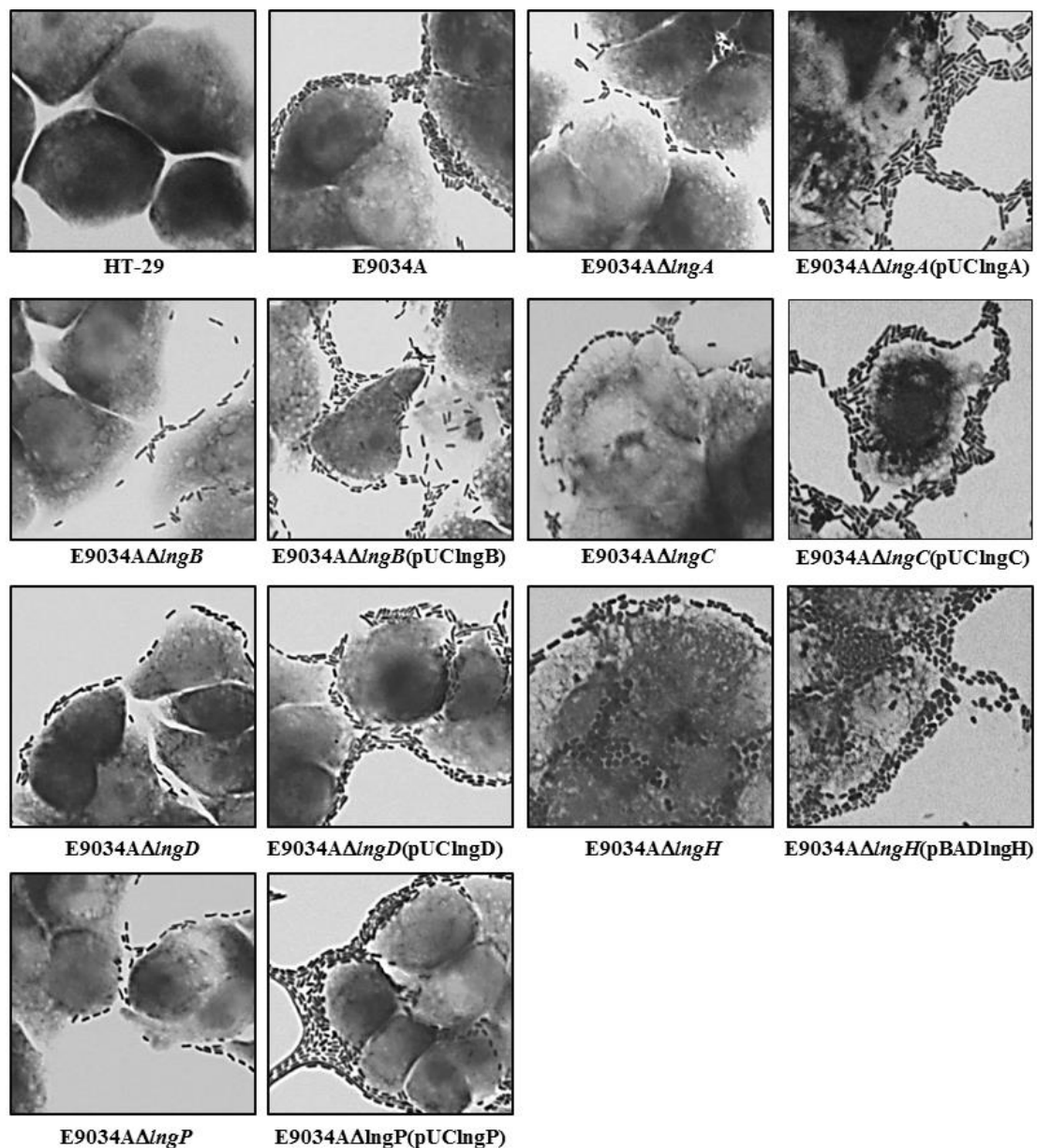


Figura 11. Adherencia mediada por el *pilus* CS21 de ETEC a células intestinales HT-29. Las mutaciones (*lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP*) y las cepas complementadas de ETEC, mostraron diferentes niveles de bacterias adheridas a las monocapas de las células.

El *pilus* CS21 se ha asociado a la auto-agregación bacteriana, una característica que confiere a ETEC ventajas en la colonización del enterocito (Clavijo et al., 2010). La auto-agregación bacteriana observada en la cepas de ETEC E9034A y cepas complementadas,

mostraron aglomerados bacterianos que sedimentan y clarifican el medio de cultivo (TB); la cual, se favorece por la expresión de los *pili* CS21 (**Figura 12A**). Las cepas mutantes en los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* no mostraron este fenotipo debido a la ausencia de la expresión del *pilus* CS21 (**Figura 12A**).

La auto-agregación de cada una de las cepas de ETEC incluidas en este estudio se cuantificó, midiendo la absorbancia a 600 nm de los sobrenadantes de los cultivos bacterianos. El análisis cuantitativo mostró un incremento significativo de la absorbancia en cada una de las cepa mutantes de ETEC, cuando se compararon con la absorbancia de la cepa de ETEC E9034A y cepas complementadas (**Figura 12B**). Adicionalmente, las cepas complementadas restablecieron el fenotipo auto-agregativo, mostrando valores similares de absorbancia a la cepa de ETEC E9034A (**Figura 12B**).

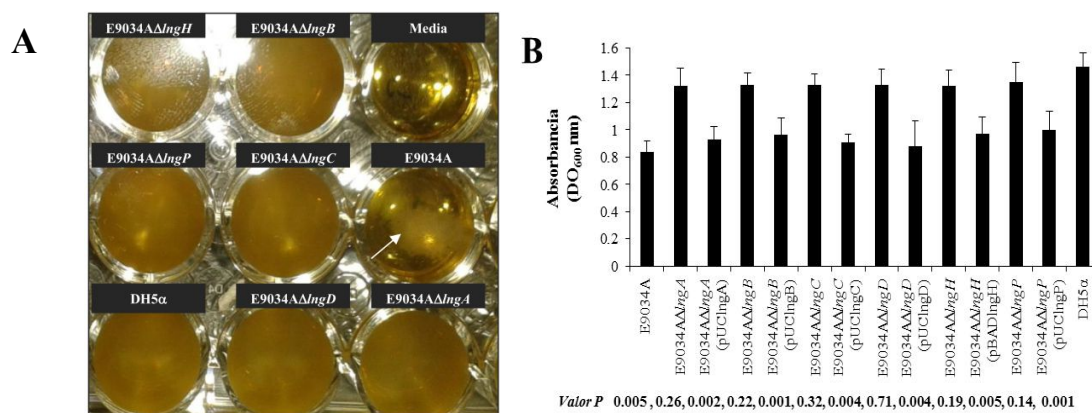


Figura 12. La ausencia de los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP*, afecta la auto-agregación de ETEC. A) La auto-agregación bacteriana en caldo TB, provoca la aglomeración bacteriana y sedimentación en la cepa E9034A (flecha), pero no en las mutantes carentes de la expresión del *pilus* CS21. Análisis cuantitativo del ensayo de auto-agregación bacteriana. B) Los sobrenadantes de las cepas mutantes de ETEC, mostraron una disminución significativa en la absorbancia a 600 nm, cuando comparamos con los valores de los sobrenadantes de la cepa ETEC E9034A y cepas complementadas. Los valores graficados son los promedios (\pm DS) de dos experimentos por triplicados realizados en dos diferentes días. El valor de $p \leq 0.005$ indica una diferencia significativa con respecto a la cepa silvestre E9034A.

8.1.6 El *pilus* CS21 de ETEC no contribuye a la formación de biopelículas

Los TIVP de bacterias Gram negativas contribuyen generalmente a la formación de biopelículas en superficies bióticas y abióticas (Roux et al., 2012). La formación de biopelículas se realizó utilizando las cepas de ETEC (cepa silvestre, cepas mutantes y cepas complementadas), cultivadas en placas de 96 pozos utilizando los medios de cultivo PPLO y DMEM-HG. El análisis cuantitativo, no mostró diferencias en la formación de biopelículas entre la cepa silvestre E9034A con las cepas mutantes y las complementadas (Figura 13). La cepa de *E. coli* enteroagregativa 042 se utilizó como control positivo y la cepa *E. coli* K12 DH5 α como control negativo.

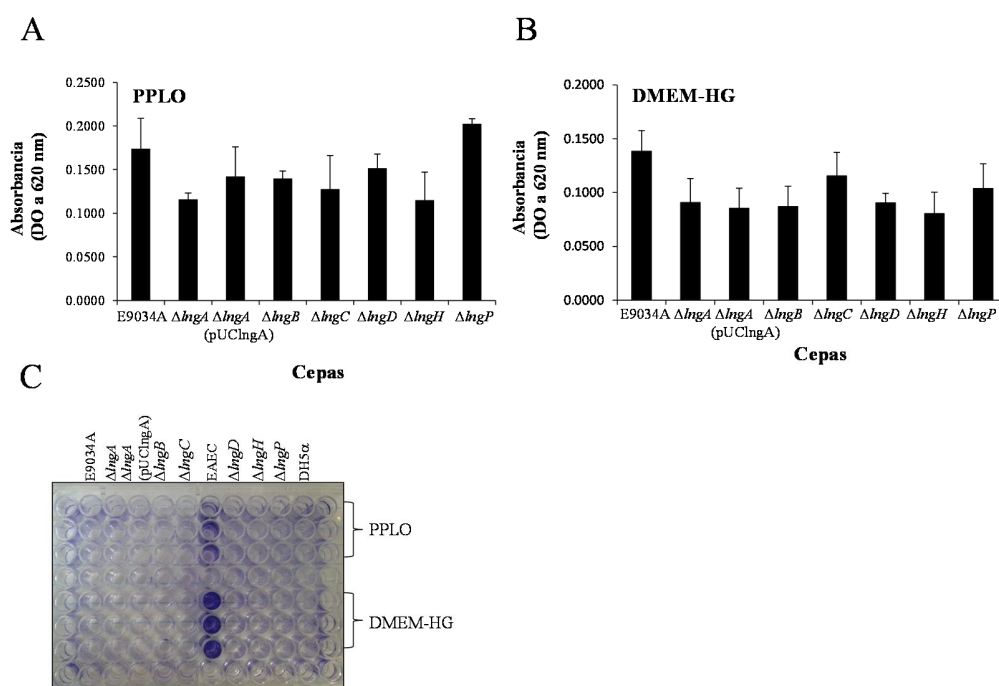


Figura 13. Cuantificación de la formación de biopelículas de las cepas de ETEC. Las biopelículas se cuantificaron por la retención de cristal violeta después de 48 horas de incubación a 37°C y en una atmósfera de 5% de CO₂. Los colorantes absorbidos por las bacterias se eluyeron con metanol absoluto y se cuantificaron espectrofotométricamente a una densidad óptica (DO) a 620 nm. A) Absorbancias promedio de la cuantificación de las biopelículas en medio PPLO. B) Absorbancias promedio de la cuantificación de las biopelículas en medio DMEM-HG. C) Visualización de la formación de las biopelículas teñidas con cristal violeta al 0.1%.

8.1.7 LngP es una prepilina peptidasa que procesa a la pilina LngA

Las prepilinas peptidasas del grupo I “PilD” procesan a las pilinas de los TIVPb. Este procesamiento comprende la escisión del péptido señal y la metilación del primer aminoácido de la pilina madura (Hobbs y Mattick, 1993; Ayers et al., 2010). Una prepilina peptidasa aún no descrita es quién procesa el péptido señal del amino terminal de la pilina LngA del *pilus* CS21 (Girón et al., 1997). Sin embargo, los análisis bioinformáticos del operón *lng* mostraron que el gen *lngP* codifica para LngP, una proteína con dominios de prepilina peptidasa. El dominio N-terminal de la proteína LngP contiene los motivos CXXC y GXCXXC que podrían participar en la actividad metiltransferasa; mientras, el dominio C-terminal contiene dos residuos de aspartato con una posible actividad de peptidasa (Roux et al., 2012). Estos datos teóricos se corroboraron experimentalmente a través de ensayos de WB, utilizando anticuerpos específicos [anti-LngA reportado por (Mazariego-Espinosa et al., 2010) y anti-CS21] y lisados bacterianos de la cepas E9034A, E9034A Δ *lngA*, E9034A Δ *lngP* y E9034A Δ *lngP*(pUCLngP) (**Figura 12**). Las cepas E9034A y E9034A Δ *lngP*(pUCLngP) produjeron una proteína de 22-kDa que corresponde a la proteína LngA madura “LngA”; mientras, la cepa E9034A Δ *lngP* produjo dos proteínas, una de 22-kDa y otra de 25-kDa que corresponden a la proteína LngA madura y a la proteína LngA sin procesar “pLngA”, respectivamente (**Figura 14**).

Los análisis bioinformáticos (motivos, cruces transmembranales y localización celular), los ensayos de adherencia y auto-agregación bacteriana, sugieren que LngP es una proteína con funciones de una prepilina péptidasa. Sin embargo, los datos de WB sugieren que existe otra prepilina péptidasa, debido al procesamiento débil de pLngA observada en la cepa E9034A Δ *lngP*.

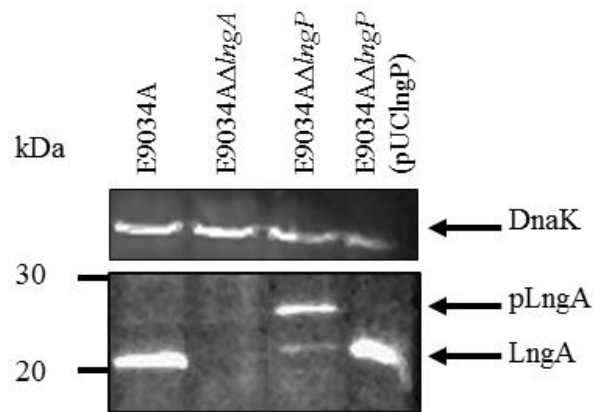


Figura 14. LngP es una prepilina peptidasa que procesa a la pilina LngA. Los extractos proteicos separados por SDS-PAGE y transferidos a membranas de PVDF se incubaron con los anticuerpos anti-LngA. La cepa E9034A expresa una proteína de 22-kDa que corresponde a la pilina LngA y la cepa E9034AΔlngP una proteína de ~25-kDa que corresponde a la pilina LngA sin procesar “pLngA”. La cepa E9034AΔlngA fue utilizada como control negativo y la detección de la proteína DnaK se utilizó como control de carga.

La señal débil de la pilina LngA en la cepa E9034AΔlngP identificada por WB, sugiere que este procesamiento mínimo podría ser suficiente para ensamblar el *pilus* CS21. Para probar esta hipótesis, se analizaron las cepas silvestre E9034A, E9034AΔlngA, E9034AΔlngP y E9034AΔlngP(pUClngP) mediante ensayos de IF. Las imágenes de los ensayos de IF mostraron la presencia de estructuras largas, fibrilares, delgadas y polares características del *pilus* CS21 en las cepas E9034A y E9034AΔlngP(pUClngP); mientras, la cepa E9034AΔlngA no mostró estas estructuras (**Figura 15**). En la imagen de la cepa E9034AΔlngP, se observó una reacción débil por IF; sin embargo, las estructuras que definen al *pilus* CS21 no fueron claras (**Figura 15**).

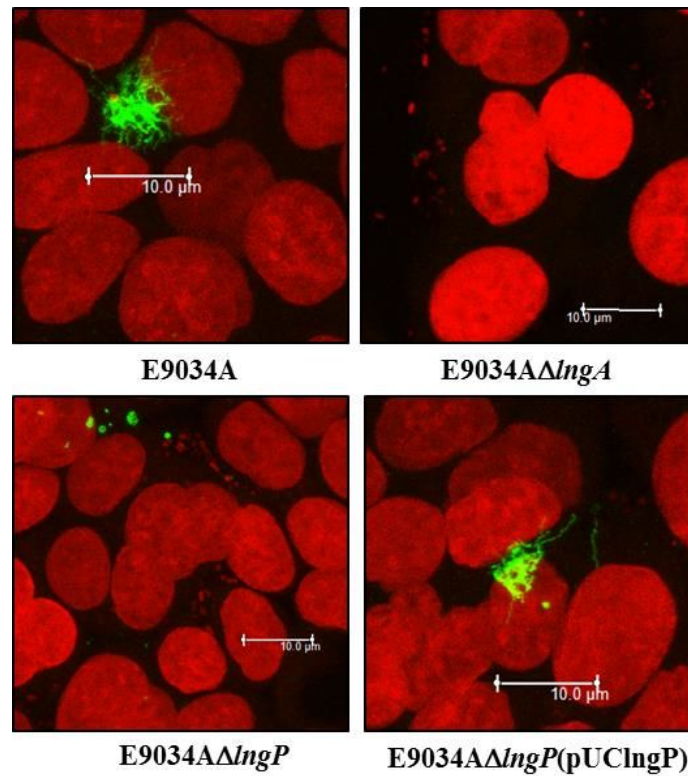


Figura 15. Presencia o ausencia del *pilus* CS21 en cepas de ETEC. La visualización de la expresión del *pilus* CS21 en las cepas de ETEC E9034A, E9034AΔ*lngA*, E9034AΔ*lngP* y E9034AΔ*lngP*(pUC*lngP*), se realizó después de la infección a células HT-29 a través de ensayos de inmunofluorescencia. Las bacterias y los núcleos de las células se tiñeron con yoduro de propidio (rojo) y la expresión del *pilus* CS21 se identificó con suero anti-CS21 seguido de un segundo anticuerpo anti-IgG de conejo conjugado con Alexa fluor 488 (verde). El ensamble del *pilus* CS21 únicamente se observó en las cepas E9034A y E9034AΔ*lngP*(pUC*lngP*).

8.1.8 Las mutantes en los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* no ensamblan el *pilus* CS21

Los sobrenadantes de las cepas de ETEC [E9034A, cepas mutantes (*lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP*) y cepas con el gen silvestre en plásmido] obtenidas después de la infección, se utilizaron para evaluar la expresión del *pilus* CS21 por microscopía electrónica de transmisión con y sin marcaje de oro. En las electromicrográficas de MET se visualizaron estructuras largas entrelazadas (mechones), fibrilares y polares que corresponden al *pilus* CS21 de la cepa E9034A; en contraste, en la cepa E9034AΔ*lngA*

observan únicamente estructuras largas, gruesas y onduladas características de flagelo (Figura 16).

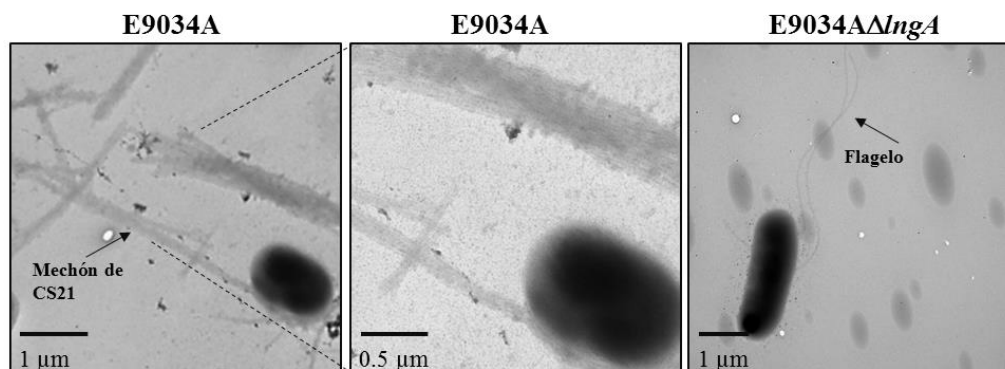


Figura 16. Visualización del *pilus* CS21 en la cepa E9034A. Electromicrografías de MET mostrando las estructuras típicas de CS21 (flecha) en la cepa E9034A y del flagelo en la cepa E9034AΔ*lngA*. Las muestras fueron contrastadas con ácido fosfotúngstico al 1% pH 7.2.

La identificación del *pilus* CS21 en la cepa de ETEC E9034A y cepas complementadas se realizó por ensayos de IEM utilizando suero anti-CS21. Las partículas de oro de 10 nm confirman la presencia de los *pili* CS21 expresada por la cepa E9034A y las cepas que son complementadas por la expresión del gen silvestre en *trans* (Figura 17). La reacción antígeno (*pilus* CS21)-anticuerpo (anti-CS21 y anti-IgG de conejo acoplados a partículas de oro de 10 nm) confirma que las estructuras marcadas con oro son mechones entrelazados de varios *pili* CS21. En las mutantes no se observaron estructuras marcadas con oro que indique la expresión del *pilus* CS21. En las electromicrografías de IEM de las cepas E9034AΔ*lngA*, E9034AΔ*lngC*, E9034AΔ*lngD*, E9034AΔ*lngH* y E9034AΔ*lngP*, se observaron aglomerados de partículas de oro en la membrana externa; sin embargo, en la cepa E9034AΔ*lngB* no se observó este fenotipo (Figura 17). Estos datos, sugieren que *LngB* podría ser una proteína estructural menor del *pilus* CS21; la cual, el suero anti-CS21 es capaz de reconocer.

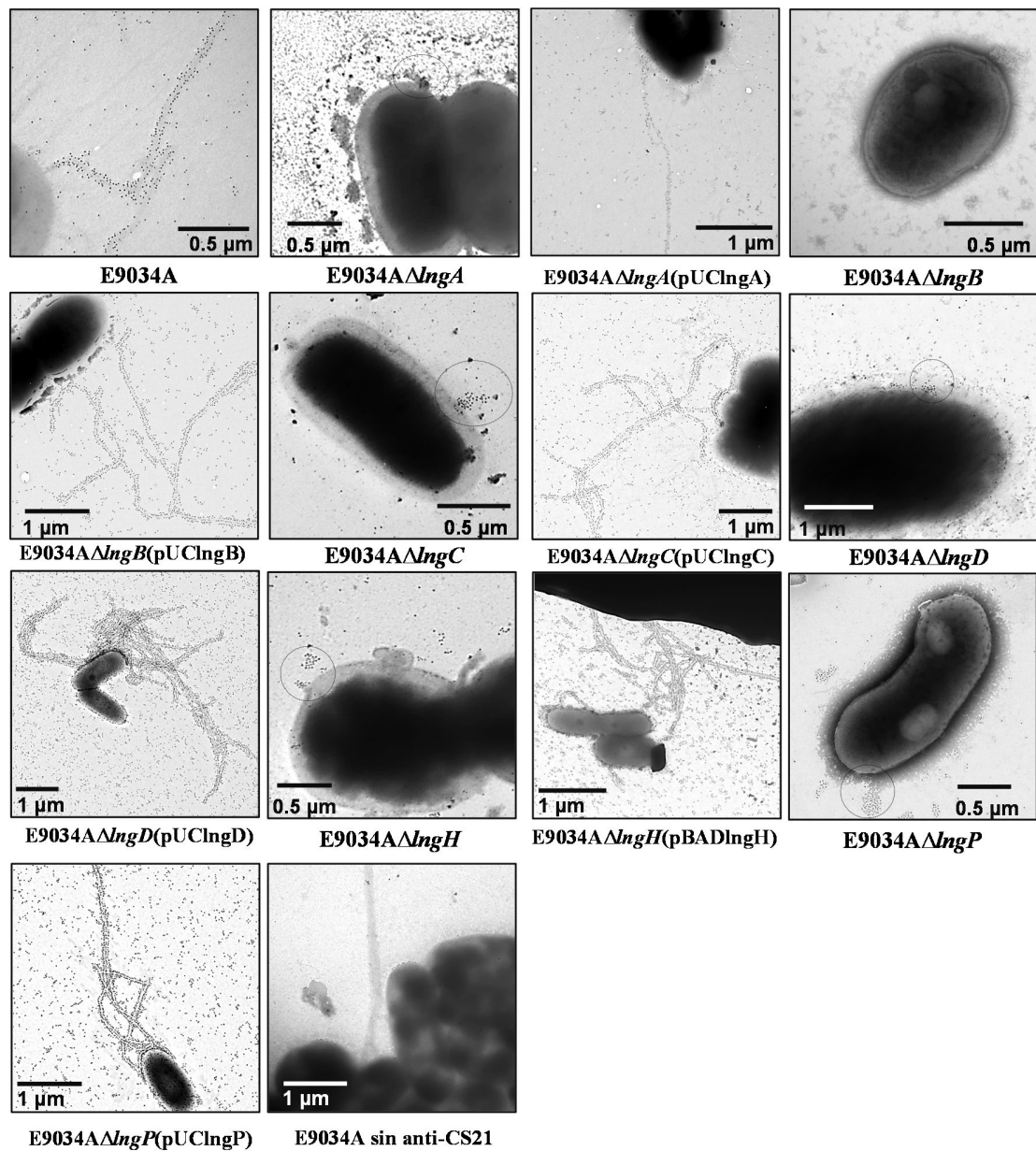


Figura 17. Visualización del pilus CS21 por IEM en el conjunto de cepas de ETEC. El ensamble del pilus CS21, se identificó utilizando suero de conejo anti-CS21 y anticuerpos anti-IgG de conejo conjugado con partículas de oro de 10 nm. Las estructuras largas entrelazadas, fibrilares y polares marcadas con partículas de oro confirmaron la naturaleza del pilus CS21 en la cepa E9034A y cepas complementadas. Las cepas mutantes en los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* carecieron de este fenotipo. La cepa E9034A incubada sin suero anti-CS21 y con anti-IgG de conejo conjugado con partículas de oro de 10 nm se utilizó como control negativo. Las escalas en las electromicrografías son indicadas con barras.

8.1.9 El suero anti-CS21 reconoce a las proteínas LngA y LngB

La predicción de la localización extracelular de LngB y su homología con la proteína CofB, sugiere que LngB pudiera ser una proteína estructural menor del *pilus* CS21. La generación de la proteína recombinante LngB-His fue una herramienta esencial para confirmar la localización extracelular de la proteína LngB como parte estructural del *pilus* CS21. El suero anti-CS21 reconoció principalmente dos proteínas de 22-kDa y 57-kDa en los extractos proteicos de la cepa E9034A por WB, que corresponden al peso molecular teórico de LngA y LngB (**Figura 18A**). El suero anti-CS21 reconoció a la proteína LngA, LngB y LngB-His, pero no a la proteína recombinante LngC-His de 16-kDa; la cual, se predijo como una proteína de membrana externa y que se utilizó como un control negativo para la localización de LngB (**Figura 18C**). Las proteínas recombinantes LngB-His de 57-kDa y LngC-His de 16-kDa, fueron detectados por anticuerpos monoclonales anti-His y se utilizaron como controles positivos (**Figura 18D**). Los de ensayos de IEM apoyaron los resultados obtenido por WB, observado que el suero anti-CS21 reconoce todo el filamento (LngA) y de manera notable se observa un incremento evidente de las partículas de oro en la punta del *pilus* (LngB) CS21 (**Figura 18E y F**). Estos datos confirmaron que el suero anti-CS21 reconoce mayoritariamente a las proteínas LngA y LngB del *pilus* CS21.

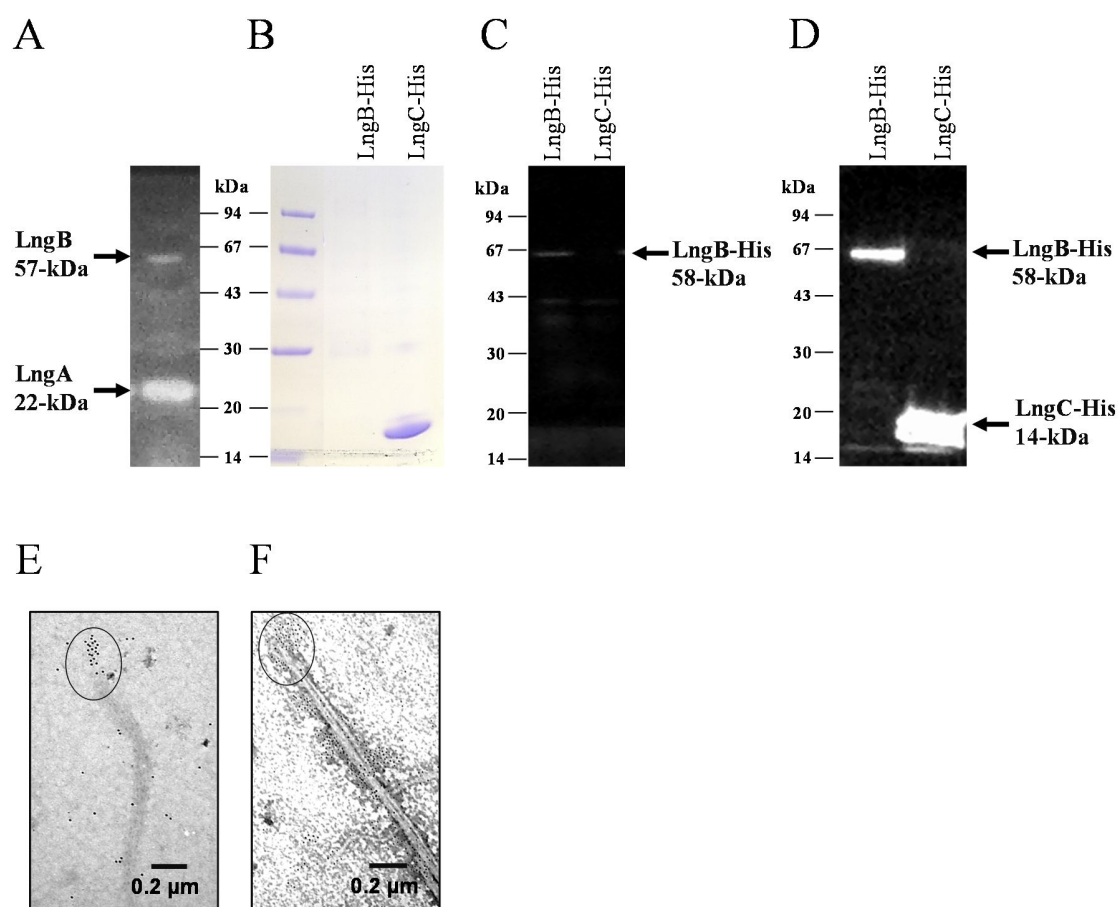
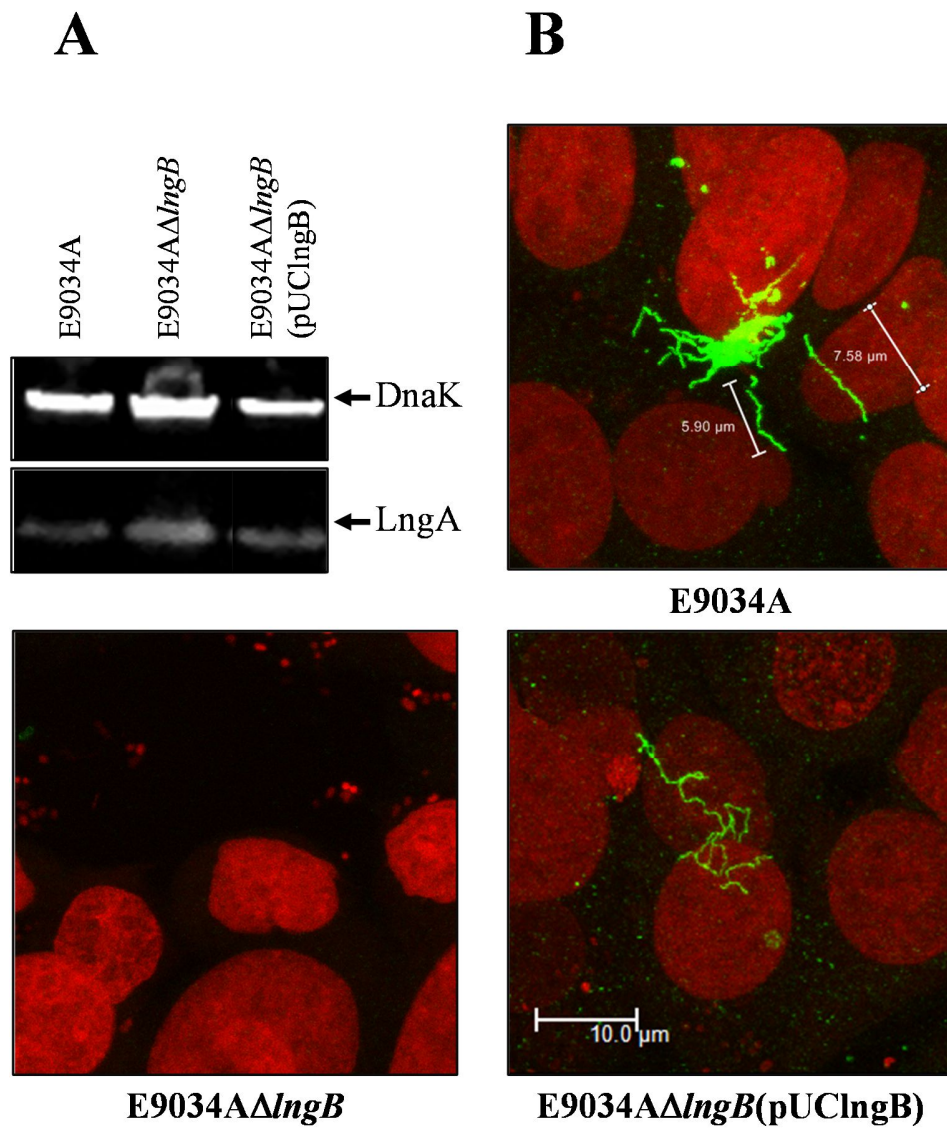


Figura 18. LngA y LngB son reconocidos por el suero anti-CS21. A) WB de los extractos proteicos totales de la cepa E9034A muestra dos bandas de 22-kDa y 57-kDa cuando se reaccionaron con el suero anti-CS21. B) Geles de SDS-PAGE mostrando dos bandas de 57-kDa y 16-kDa que corresponden a las proteínas LngB-His y LngC-His, respectivamente. C y D) WB mostrado a las proteínas recombinantes LngB-His y LngC-His con anti-CS21 y anticuerpos monoclonales anti-His, respectivamente. E y F) Acumulación de partículas de oro en un extremo de los filamentos del *pilus* CS21 identificadas con anti-CS21 y anti-IgG de conejo acoplado a partículas de oro.

8.1.10 LngB es una proteína estructural menor del *pilus* CS21

La proteína CofB de CFA/III tiene una identidad del 78% con la proteína LngB del *pilus* CS21 y se ha demostrado que su región amino terminal se requiere para iniciar el ensamble de CFA/III. Además, CofB se ha propuesto como una proteína extracelular cuya localización podría ser en la punta del *pilus* CFA/III (Kolappan et al., 2015). Bajo este contexto, consideramos que LngB, es una proteína menor requerida para el ensamble del

pilus CS21. La mutación del gen *lngB* abolió el ensamble del *pilus* CS21; aunque, no se afectó la expresión y el procesamiento de la proteína LngA (**Figura 19A y B**). La complementación en *trans* de la cepa E9034A Δ *lngB* con el vector pUC*lngB* restauró el ensamble del *pilus* CS21 pero con filamentos más largos comparados con los de la cepa E9034A (**Figura 19**).



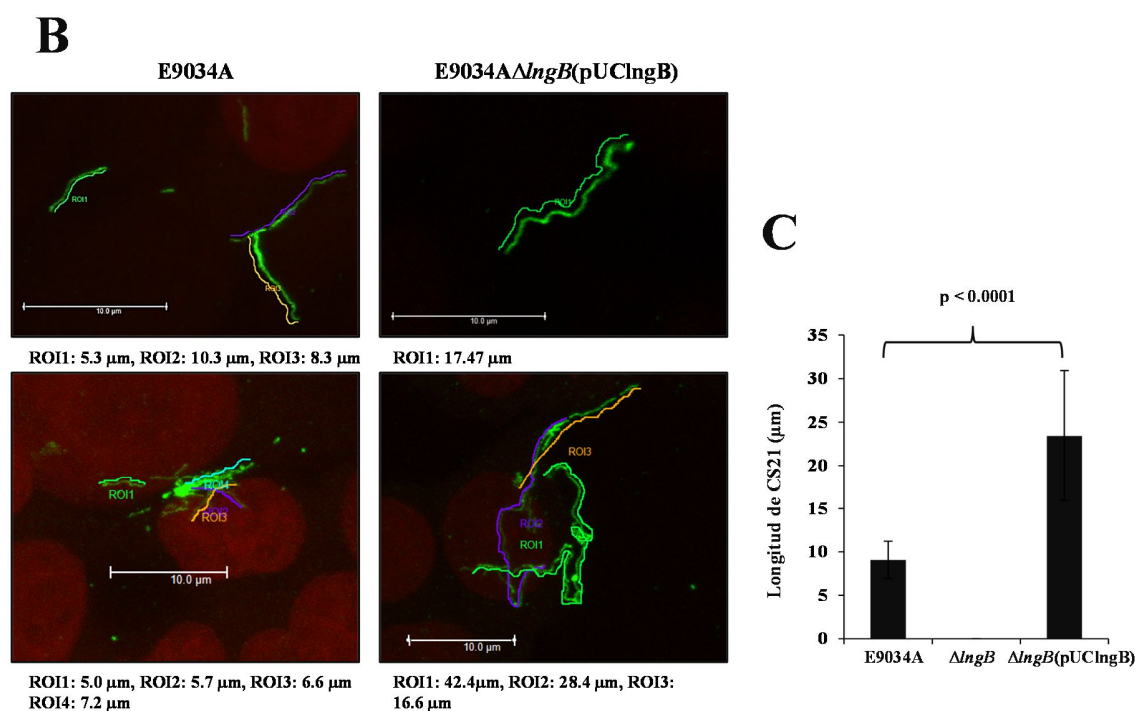


Figura 19. Expresión del gen *lngB* en *trans* restaura la expresión del *pilus* CS21. A) WB mostrando la misma expresión de la proteína LngA con el suero anti-CS21 y de la proteína DnaK con anticuerpos monoclonales anti-DnaK. B) Micrografías de inmunofluorescencia mostrando los filamentos de CS21 después de márcalos con suero anti-CS21 y anticuerpos anti-IgG de conejo conjugados con Alexa-fluor 488 (verde). El ADN eucariota y bacteriano se tiñó con yoduro de propidio (rojo). Las micrografías de microscopía confocal se tomaron con un aumento de 63x. C) Representación gráfica de los promedios y desviación estándar de la medición de la longitud de los *pili* CS21 en las cepas E9034A, E9034AΔ*lngB* y E9034AΔ*lngB*(pUClngB). El valor de $p < 0.0001$ indica una diferencia significativa con respecto a la cepa silvestre E9034A

La longitud de los *pili* CS21 de la cepa E9034A y de la cepa E9034AΔ*lngB*(pUClngB), se realizó mediante microscopía confocal utilizando un microscopio Leica TCS SP8X, con un objetivo de inmersión en aceite de 63x y un zoom digital de 5x. Las imágenes se analizaron con el software Leica Application Suite, Fluorescencia Avanzada Lite Versión 2.6.3 build 8173. La longitud de los *pili* CS21 en la cepa E9034A fueron de $9.01 \pm 2.05 \mu\text{m}$ y en la cepa E9034AΔ*lngB*(pUClngB) de $22.62 \pm 6.77 \mu\text{m}$ (**Figura 19C**). La cepa E9034Δ*lngA*(pUClngA), no mostró un incremento en la longitud del *pilus* CS21 como se observó en la cepa E9034AΔ*lngB*(pUClngB) a pesar de

utilizar el mismo vector multicopia de complementación (**Figura 20**). El incremento en la longitud de los *pili* CS21 en la cepa E9034AΔ*IngB*(pUC*IngB*), se debe al uso de un vector de alto número de copias y por lo tanto a una sobreexpresión de la proteína *IngB*; no obstante, este fenómeno no se observó con la sobreexpresión de la proteína *IngA* (**Figura 20**).

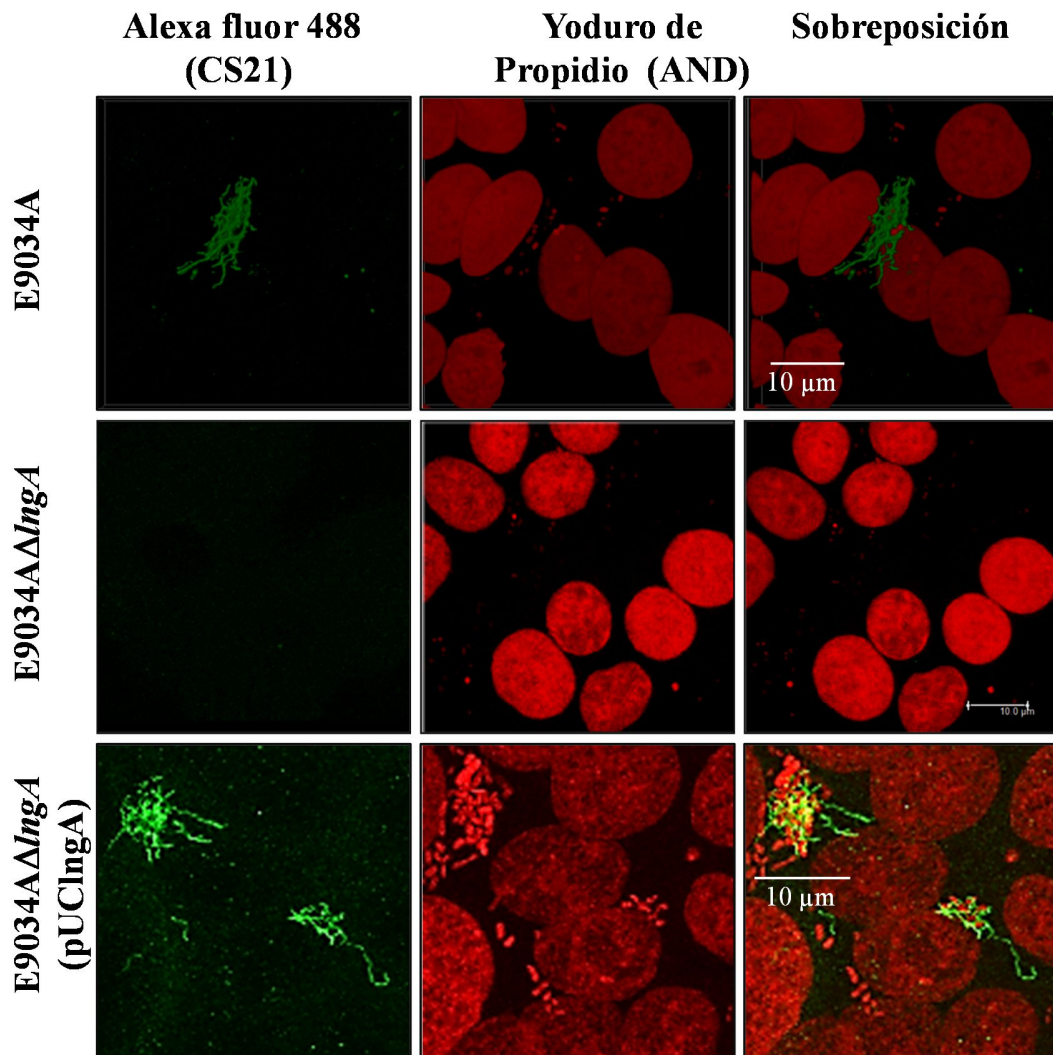


Figura 20. Sobreexpresión de la proteína *IngA* no afecta la longitud del *pilus* CS21. Imágenes de inmunofluorescencia mostrando los filamentos de los *pili* CS21 marcados con suero anti-CS21 y anticuerpos anti-IgG de conejo conjugados con Alexa Fluor 488 (verde). El ADN eucariota y bacteriano se tiñó con yoduro de propidio (rojo). Las micrografías de microscopía confocal se tomaron con un aumento de 63x.

Recientemente, ha sido propuesto que la proteína LngB es una pilina menor que forma parte del *pilus* CS21; sin embargo, estos datos no se corroboraron experimentalmente (Kolappan et al., 2015). En este trabajo, generamos anticuerpos anti-LngB a partir de la proteína recombinante LngB-His, para demostrar la localización específica de la proteína LngB en el *pilus* CS21. En las imágenes de IEM, se observan cúmulos de partículas de oro en la punta del *pilus* CS21 de la cepa E9034A y una ausencia de este fenotipo en la cepa E9034AΔ*lngB* (Figura 21). La presencia de numerosas partículas de oro sugiere que ETEC produce varios filamentos entrelazados, constituidos por la proteína LngA y la punta asociada con la proteína LngB. Estos resultados apoyan que la proteína LngB, sea una subunidad menor anclada en la punta del *pilus* CS21.

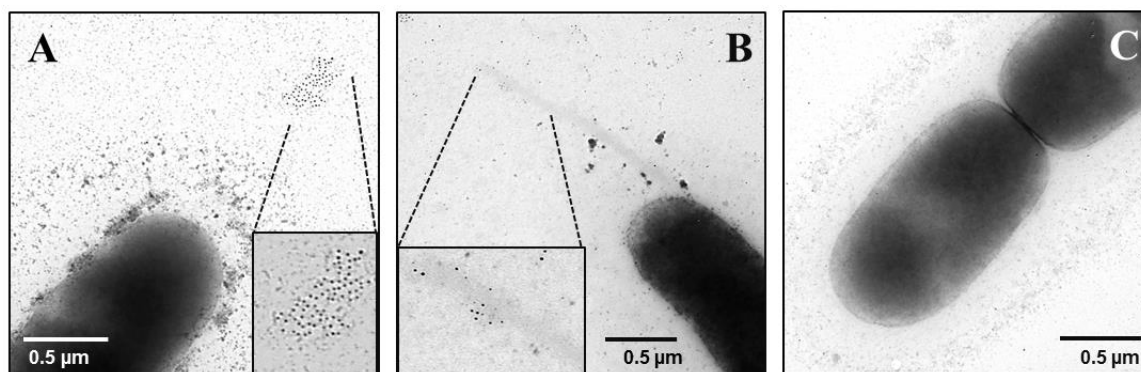


Figura 21. Detección de la proteína LngB en la punta de los *pili* CS21. A y B) Imágenes de IEM mostrando partículas de oro de 10 nm que indican la presencia de la proteína LngB en la punta de los *pili* CS21. Los recuadros muestran un aumento de la localización de la proteína LngB. C) En la cepa E9034AΔ*lngB* (control negativo) no se observaron cúmulos de partículas de oro de 10 nm.

8.1.11 Las proteínas LngC, LngD y LngH se requieren para el ensamble del *pilus* CS21

La proteína LngC tiene una identidad del 24%, 20.3% y 70% con las proteínas TcpQ, BfpG y CofC, respectivamente; mientras, la región C-terminal tiene una alta identidad con un

dominio conservado de la proteína TcpQ (**Figura 4**). La proteína TcpQ de *V. cholerae* y BfpG de EPEC contienen un dominio transmembranal que favorecen la interacción con las secretinas de la membrana externa, TcpC y BfpB, respectivamente (Schmidt et al., 2001; Bose y Taylor, 2005). Los monómeros de BfpB son estables en ausencia de BfpG y los monómeros de TcpC son degradados completamente en ausencia de TcpQ; por lo tanto, TcpQ es requerida para la formación y/o estabilidad de la secretina TcpC, mientras la secretina BfpB es estable en ausencia de BfpG (Schmidt et al., 2001; Bose y Taylor, 2005). La predicción bioinformática indicó que la proteína LngC se puede localizar en la membrana externa, coincidiendo con la localización de sus homólogos TcpQ y BfpG; sin embargo, no se detectaron cruces transmembranales y si un péptido señal en la secuencia de amino ácido de la proteína LngC (**Tablas 5 y 6**). La presencia de un péptido señal en la secuencia de LngC es indicativo de que puede ser translocada a periplasma. Aunado a esto, la homología que tiene con proteínas que experimentalmente se han demostrado su localización celular en la membrana externa, sugiere que LngC es una proteína de membrana externa. El péptido señal identificado por los tres predictores (SignalP4.1, PrediSi y Phobius) fue de 20 aminoácidos en la secuencia de la proteína LngC (**Tabla 6 y Figura 22A**). Estos datos teóricos se corroboraron experimentalmente con la proteína recombinante LngC-His a través de un corrimiento electroforético en geles de SDS-PAGE al 16%, en donde se detectaron dos bandas de aproximadamente 14-kDa y 16-kDa que corresponden a la proteína madura e inmadura de LngC, respectivamente (**Figura 22**).

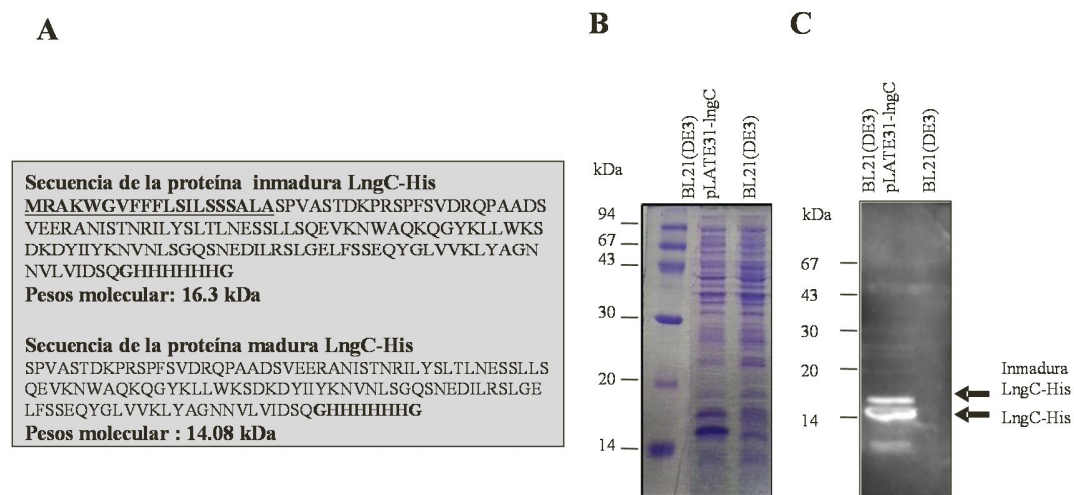


Figura 22. Procesamiento de la proteína LngC. A) Secuencia de aminoácidos de la proteína LngC-His inmadura (incluyendo el péptido señal) y madura. B) Gel de poliacrilamida al 16% mostrando el perfil proteico de los lisados bacterianos de la cepa BL21(DE3) y BL21 (DE3) con el pLATE31-lngC, inducidas con 1 mM de IPTG durante 5 h. C) Identificación de la proteína LngC-His (14-kDa y 16-kDa) por WB utilizando anticuerpos monoclonales anti-His.

La mutación del gen *lngC* en la cepa E9034A no afectó la expresión y si el procesamiento de la proteína LngA. Dos bandas de 22-kDa y 25-kDa se identificaron por WB con extractos proteicos de la cepa E9034AΔ*lngC*, que corresponden a los pesos moleculares de la prepilina y pilina LngA (**Figura 23B**). Además, la mutación del gen *lngC* en la cepa E9034A afectó el ensamble del *pilus* CS21, como se observa en las imágenes de IEM e IF (**Figura 17 y 23C**). La complementación de la cepa E9034AΔ*lngC* a través de la incorporación del gen *lngC* en *trans* restauró el fenotipo observado en la cepa silvestre (**Figura 23C**).

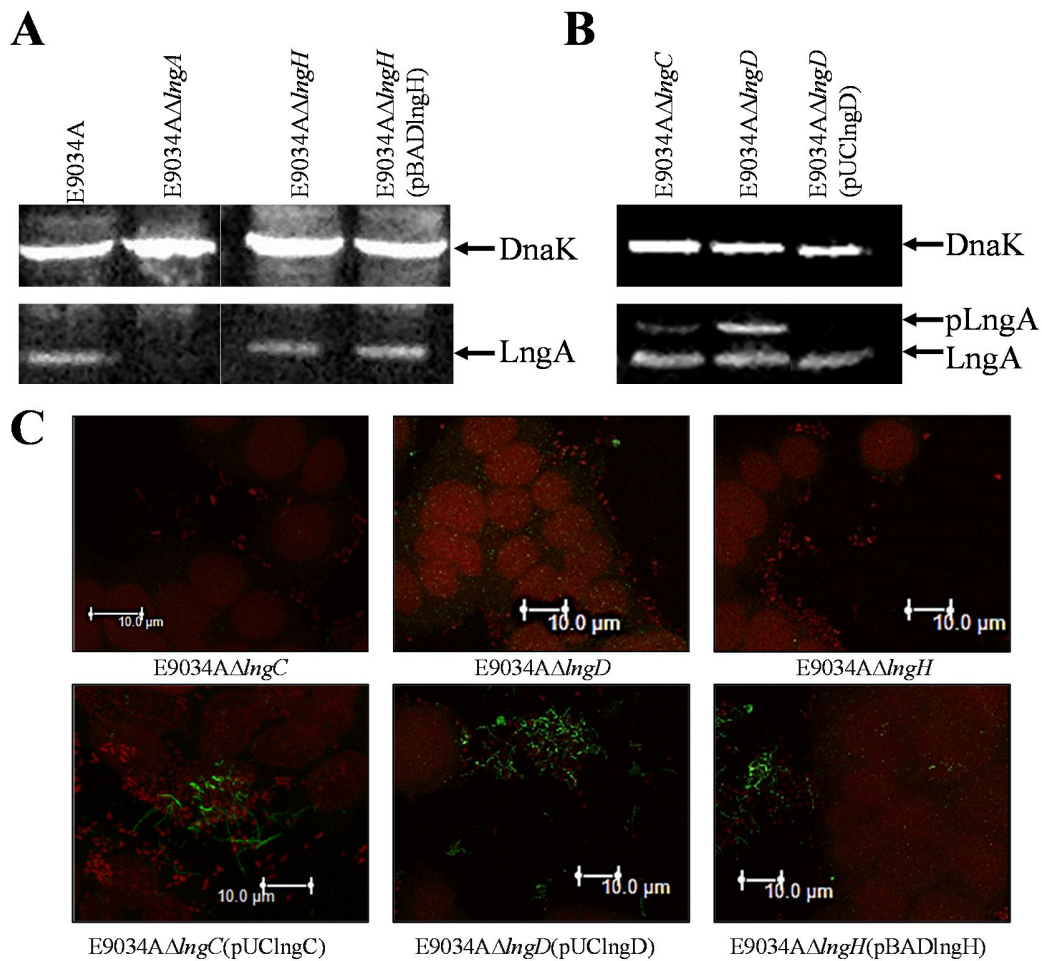


Figura 23. La mutación de los genes *lngC* y *lngD* afectó el procesamiento de la proteína LngA. A) Identificación por WB de la proteína LngA de 22-kDa utilizando anti-CS21. B) Identificación por WB de las prepilinas de 25-kDa en las cepas E9034AΔ*lngC* y E9034AΔ*lngD* utilizando el suero anti-CS21. En ambos casos, la detección de la proteína DnaK se utilizó como control de carga. C) Imágenes de inmunofluorescencia mostrando la presencia de los *pili* CS21 en la cepas E9034AΔ*lngC*(pUC*lngC*) y E9034AΔ*lngD*(pUC*lngD*) y E9034AΔ*lngH*(pBAD*lngH*); mientras, la mutantes en los genes *lngC*, *lngD* y *lngH* carecieron de la presencia del *pilus* CS21. El ADN fue teñido con yoduro de propidio (rojo) y las estructuras de CS21 se detectaron con suero de conejo anti-CS21 y anticuerpos policlonales anti-IgG de conejo conjugado con Alexa-fluor 488 (verde). Las fotografías fueron tomadas a 63x.

LngD y LngH son proteínas homólogas a la lipoproteína de membrana externa y a la proteína de unión a nucleótido, respectivamente, de CFA/III, TCP y BFP. LngD tiene una identidad del 91%, 17.6% y 19.7% con CofD, TcpC y BfpB, respectivamente. El C-terminal de LngD tiene un dominio conservado en la super-familia de proteínas secretinas *pilus_B_mal_scr*. La super-familia de proteínas secretinas son predichas a ser ricas en

estructura secundaria de tipo hojas β -plegadas (Koebnik et al., 2000). La predicción de la estructura secundaria de la secuencia LngD con el programa Jpred4 y PORTER revelo que hojas β -plegadas están predominantemente presentes en la region C-terminal (región amino acidica de 333-492). LngD fue predicho a tener un péptido señal, lo que nos sugiere que LngD es probablemente translocado al espacio periplasmico. La predicción de la localización celular de LngD fue en la membrana externa, su homólogo TcpC se ha demostrado experimentalmente esta misma localización (Bose y Taylor, 2005). La mutación del gen *lngD* en la cepa E9034A redujo la adherencia a las células intestinales HT-29 en un 85% y la complementación en *trans* restaura el fenotipo silvestre (**Figura 10 y 11**). El procesamiento de LngA en la cepa E9034A Δ *lngD* mostró dos bandas por WB, que corresponden a la prepilina (pLngA) y pilina (LngA), muy parecido a las observadas en la cepa E9034A Δ *lngC* (**Figura 23B**). El ensamble del *pilus* CS21 se abolió en la cepa E9034A Δ *lngD*, aunque LngA se expresa y se procesa parcialmente, pero al no haber un poro en la membrana externa (secretina) las subunidades de LngA no se ensamblan en las estructuras poliméricas de CS21 y por tanto la bacteria carece de esta estructura para adherirse a las células HT-29 (**Figura 23C**).

LngH tiene una identidad del 75%, 44.7% y 21.9% con CofH, TcpT y BfpD, respectivamente. TcpT y BfpD son proteínas de unión a nucleótido encargadas de proveer energía (ATP) para polimerizar TCP y BFP, respectivamente (Anantha et al., 2000; Crowther et al., 2005). LngH pertenece a la super-familia P-asa NTPasa, caracterizada por un motivo conservado de unión a ATP, también denominado como el motivo Walker A [GxxxxGK(ST)] (Iyer et al., 2004). La predicción de la localización celular de LngH fue en citoplasma lo que coincide con su homólogo TcpT (Iredell y Manning, 1997). La mutación del gen *lngH* en la cepa E9034A mostró una reducción en la adherencia del 91%

comparado con la cepa silvestre E9034A (**Figura 10**). La expresión y procesamiento de LngA en la cepa E9034A Δ LngH no se vio afectada (**Figura 23A**), pero si el ensamble del *pilus* CS21 (**Figura 23C**), esto sugiere que LngH pudiera ser la ATPasa encargada de proveer la energía necesaria para el ensamble del *pilus* CS21.

Basados en los resultados obtenidos de adherencia a células intestinales, auto-agregación bacteriana, W.B, IF, IEM, predicción de la localización celular, cruces transmembranales, péptido señal, motivos y análisis de identidad con proteínas homólogas en CFA/III, TCP y BFP se puede comenzar a inferir sobre la posible función y localización de las proteínas Lng en el ensamble de CS21 y proponer un posible modelo como el que se muestra en la **figura 24**.

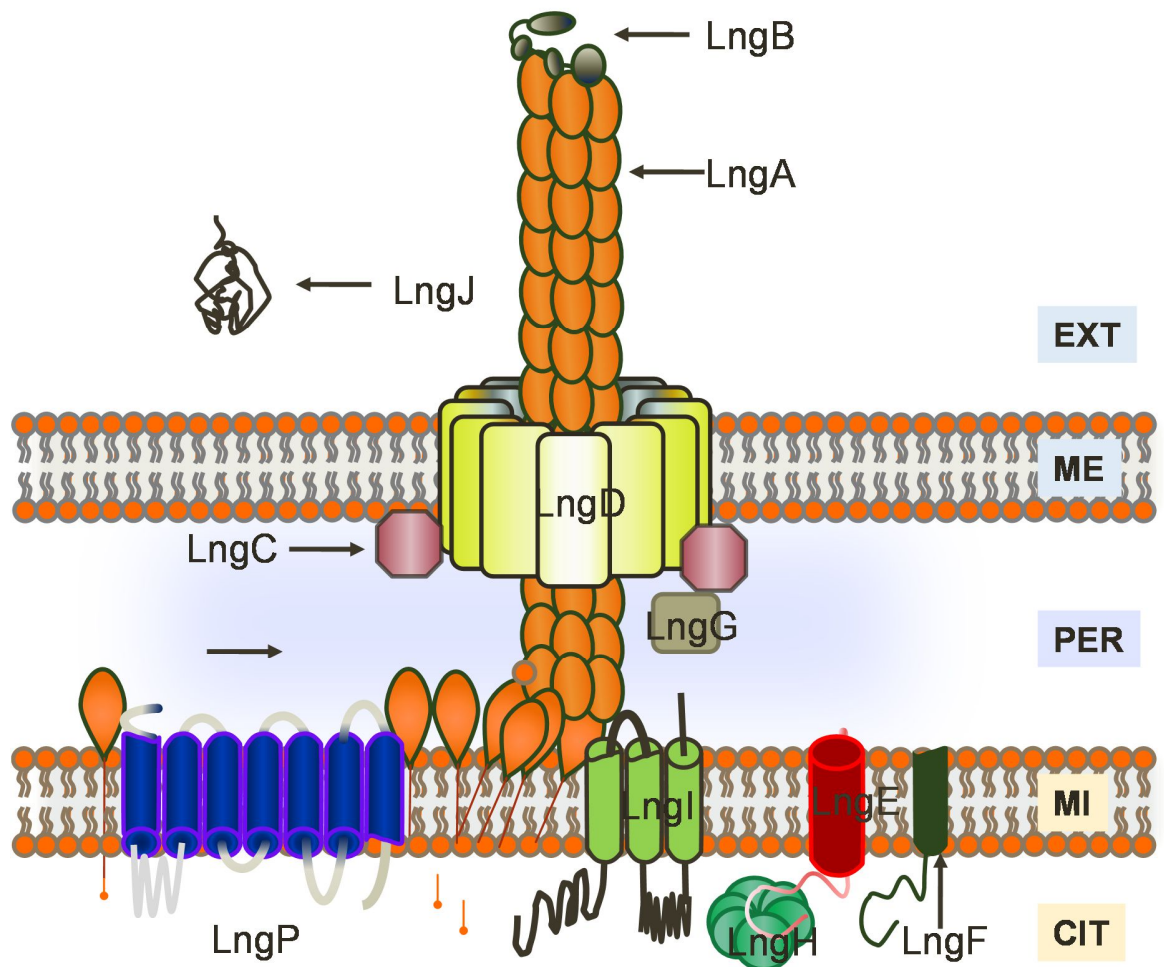


Figura 24. Modelo propuesto de la localización celular de las proteínas Lng y posible ensamblaje de CS21. LngA es la subunidad estructural mayor, LngB es una pilina que se localiza en la punta del *pilus* CS21 y controla la longitud de este. LngC es una proteína accesoria de membrana externa. LngD es la probable secretina de membrana externa. LngH es la ATPasa de citoplasma y LngP es la prepilina peptidasa. LngE, LngG, LngI y LngJ fueron incluidas en el modelo con base a la predicción de sus homologos en TCP y BFP y a los análisis bioinformaticos realizados. ME; membrana externa, MI; membrana interna, EXT; espacio extracelular, CIT; citoplasma y PER; periplasma.

8.2 Regulación del *pilus* CS21

8.2.1 El gen *IngA* tiene un promotor corriente arriba de su codón de inicio

Para determinar si los 14 marcos de lectura forman parte del mismo operón, se realizó un análisis bioinformático con los programas NNPP versión 2.2, PePPER y BPROM (Solovyev et al., 2010), los cuales predicen promotores bacterianos. BPROM predice promotores de tipo sigma 70 (σ^{70}) con un 80% de precisión y especificidad en las regiones intergénicas de *E. coli*.

En base a estos análisis se diseñaron oligonucleótidos para clonar las posibles regiones que contienen el promotor de *Ing* en el vector pJET1.2, estas regiones se encuentran corriente arriba de los genes *IngR*, *IngS* y *IngA* (Figura 25).

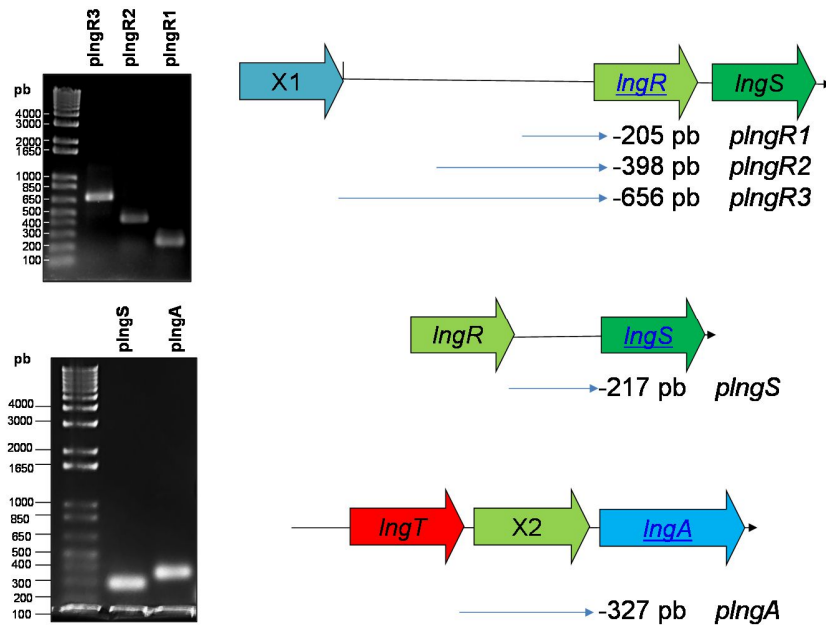


Figura 25. Regiones intergénicas y codificantes del cluster *Ing* y posibles promotores. Geles de agarosa al 1% mostrando los amplicones obtenidos por PCR donde se predice puede haber una región promotora. Representación esquemática de las regiones intergénicas que se encuentran corriente arriba de los genes *IngR*, *IngS* y *IngA*. Las flechas por debajo de los esquemas representan los productos esperados por PCR en la cepa E9034A.

El gen *lngA* codifica a la subunidad estructural LngA del *pilus* CS21 y corriente arriba de su ATG se encuentra un marco de lectura X2 de 222 pb que se predice codifica a una proteína hipotética. La región intergenica corriente arriba del marco de lectura X2 y el mismo X2 tiene una longitud de 300 pb. Los programas NNPP versión 2.2, PePPER y BPRM predijeron un posible promotor a -250 pb del ATG del gen *lngA*. Para confirmar la predicción de la región promotora se clonó el gen *lngA* con y sin la región de -327 pb en el vector pJET1.2, dando origen a los vectores *plngA* y *lngA*, respectivamente. La construcciones se transformaron en la cepa *E. coli* K-12 TOP10 y examinaron por su capacidad de expresar y procesar la proteína LngA por WB. Como se puede apreciar en la **figura 26**, la cepa TOP10(*plngA*), que alberga al gen *lngA* en *trans* y la región de -327 pb es capaz de expresar y procesar a la proteína LngA. Sin embargo la cepa TOP10(*lngA*) que únicamente contiene al gen *lngA* en *trans* no expresó la proteína LngA. Como controles negativos se transformó la cepa TOP10 con el vector solo (pJET1.2) y como control positivo se usó la cepa silvestre E9034A.

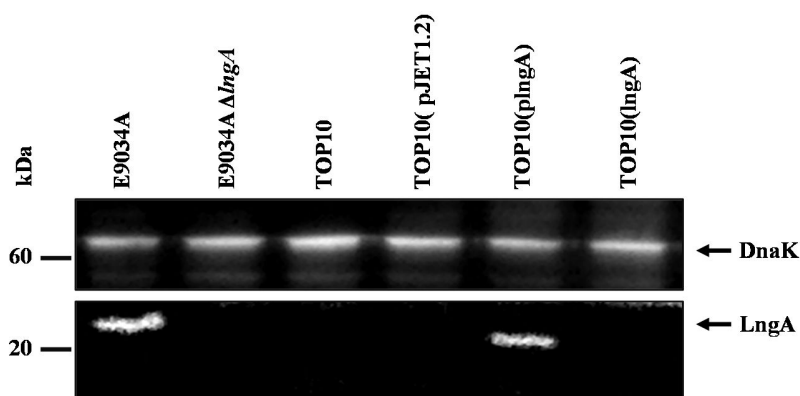


Figura 26. Expresión y procesamiento completo de la proteína LngA en una cepa de *E. coli* K-12 TOP10(*plngA*). Extractos proteicos de las cepas indicadas arriba se sometieron a SDS-PAGE, se transfirieron a membranas de PVDF y se reaccionaron con suero anti-CS21. Las cepas E9034AΔ*lngA*, TOP10 y TOP10(pJET1.2) fueron usadas como control negativo de la expresión de LngA. En la cepa silvestre E9034A y cepa TOP10(*plngA*) se observa una banda de aproximadamente 22-kDa que corresponde a la pilina LngA. Las flechas indican la proteína LngA de 22-kDa y DnaK de 70-kDa, este última, se utilizó como control de carga.

8.2.2 Regulación postranscripcional de *lng*

La terminación Rho-independiente es un mecanismo importante de paro transcripcional en bacterias que controla tanto la terminación 3' normal como una amplia gama de eventos de atenuación de la transcripción. Los terminadores independientes de Rho contienen regiones palindrómicas que forman horquillas cuya longitud varía de 7 a 20 pb. A la estructura de tallo y lazo le sigue una serie de residuos U (Lewin y Lopez, 1996). La predicción de terminadores Rho-independiente es un paso importante en la anotación de operones bacterianos. Varios algoritmos eficientes están disponibles para este propósito. En *E. coli* los terminadores independientes de Rho se caracterizan por tener una estructura de tallo en el RNA con un alto contenido de G+C, seguido por una serie de residuos de uracilo, pero pueden ser sólo parcialmente predichas por la estabilidad de su estructura o por su secuencia primaria (d'Aubenton Carafa et al., 1990). El servidor ARNold se usó para la identificación de terminadores transcripcionales independientes de Rho y el servidor UNAFold se usó para predecir la estructura secundaria y la energía libre de Gibbs de las secuencias de RNA y ADN. El servidor ARNold predijo tres terminadores transcripcionales independientes de Rho en el cluster de los 14 genes de *lng* (13.93 kb): uno a 40 pb del ATG del gen *lngS*, el segundo a 384 pb del ATG del gen *lngA* y el tercero a 994 pb del ATG del gen *lngD*. Las estructuras secundarias y la energía libre de Gibbs se muestran en la **figura 27**.

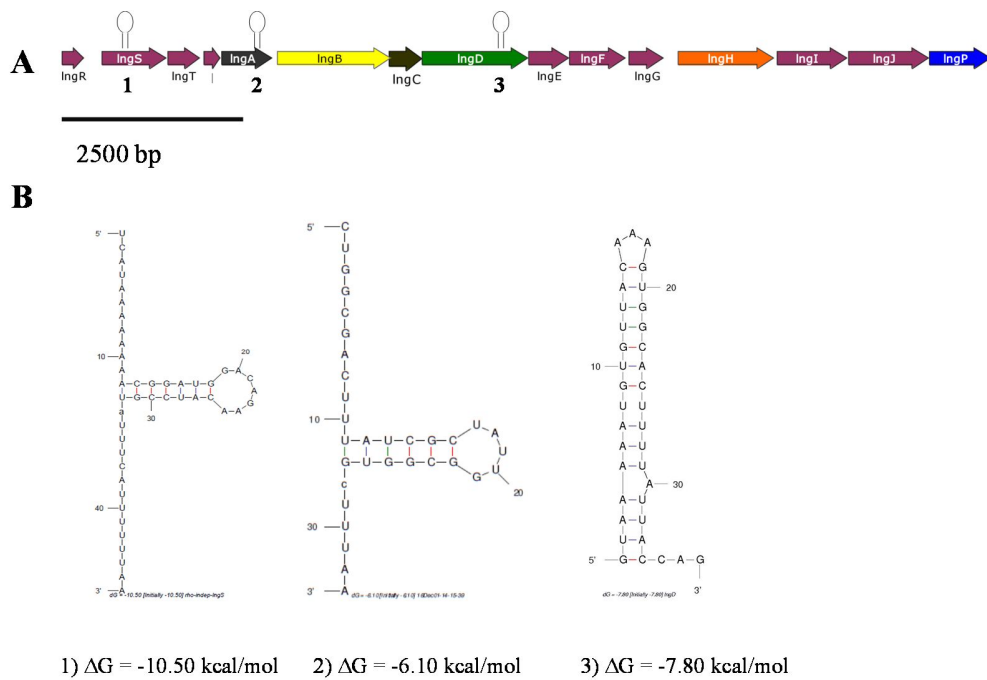


Figura 27. Probables terminadores o atenuadores transcripcionales independiente de Rho presentes en el cluster de *lng*. A) Representación esquemática de la organización del cluster de genes *lng* con los tres probables terminadores de la transcripción Rho independientes señalados (Ω). B) Estructuras secundarias tallo-lazo y la energía libre de Gibbs de las secuencias de RNA con su respectiva energía libres (ΔG).

8.2.3 LngR favorece la expresión de CS21

Dentro de los 14 marcos de lectura involucrados en la biogénesis del *pilus* CS21, existen dos genes que codifican para proteínas con dominios regulatorios que se localizan corriente arriba del gen *lngA*: *lngR* y *lngS*. LngR tiene homología con la familia PapB, la cual incluye reguladores fimbriales y afimbriales, los cuales activan o reprimen la transcripción mediante unión al ADN (Xia y Uhlin, 1999). LngR tiene una identidad del 35.6% con PapB y ambos presentan los sitios de unión a ADN y los sitios de oligomerización, como se muestra en la **figura 28**.



Figura 28. LngR y PapB son homólogos con una identidad del 35.6%. Alineamiento múltiple usando Clustal Omega de las secuencias de aminoácidos de LngR y PapB mostrando los sitios de unión a ADN (#) y oligomerización (*).

En la cepa de ETEC E9034A es posible inferir que las proteínas LngR y LngS estén relacionadas con una regulación negativa o positiva en la biogénesis del *pilus* CS21; por lo tanto, las proteínas de ensamblaje propuestas están codificadas “río abajo” de los genes regulatorios (*lngR* y *lngS*). Con el fin de elucidar el papel de la proteína LngR en la regulación del *pilus* CS21, se mutó el gen *lngR* por el método de inactivación en un paso (Datsenko y Wanner, 2000) (Figura 29).

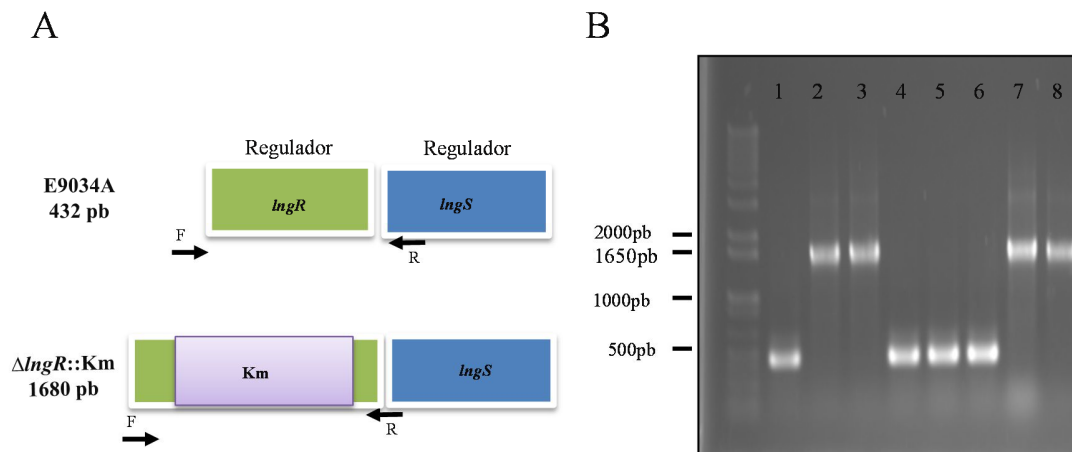


Figura 29. Confirmación de la mutación del gen *lngR* en la cepa E9034A. A) Esquema de la inserción del casete de Km y eliminación del gen *lngR*. B) Gel agarosa al 1% teñido con bromuro de etidio mostrando los perfiles encontrados en 8 clonas después de la mutación del gen *lngR*. En los carriles 1, 4, 5 y 6 se observa una banda de 432 pb, la cual corresponde a un genotipo de tipo silvestre y en los carriles 2, 3, 7 y 8 una banda de 1680 pb correspondiente al genotipo de una cepa mutada en el gen *lngR* con inserción del casete de Km.

La cepa generada E9034A Δ IngR::Km se caracterizó mediante ensayos de adherencia a células intestinales HT-29, IF, tinción de Giemsa y auto-agregación bacteriana. Los resultados obtenidos mostraron un incremento de 3.6 veces en la adherencia a células HT-29 en la cepa E9034A Δ IngR::Km y un incremento notable en la auto-agregación bacteriana, cuando se comparan con la cepa silvestre E9034A. La sobre-expresión de los *pili* CS21 (verde) observada en las imágenes de IF en la cepa E9034A Δ IngR::Km confirmó que el incremento en la adherencia se debió a una mayor producción de los *pili* CS21 (**Figura 30**).

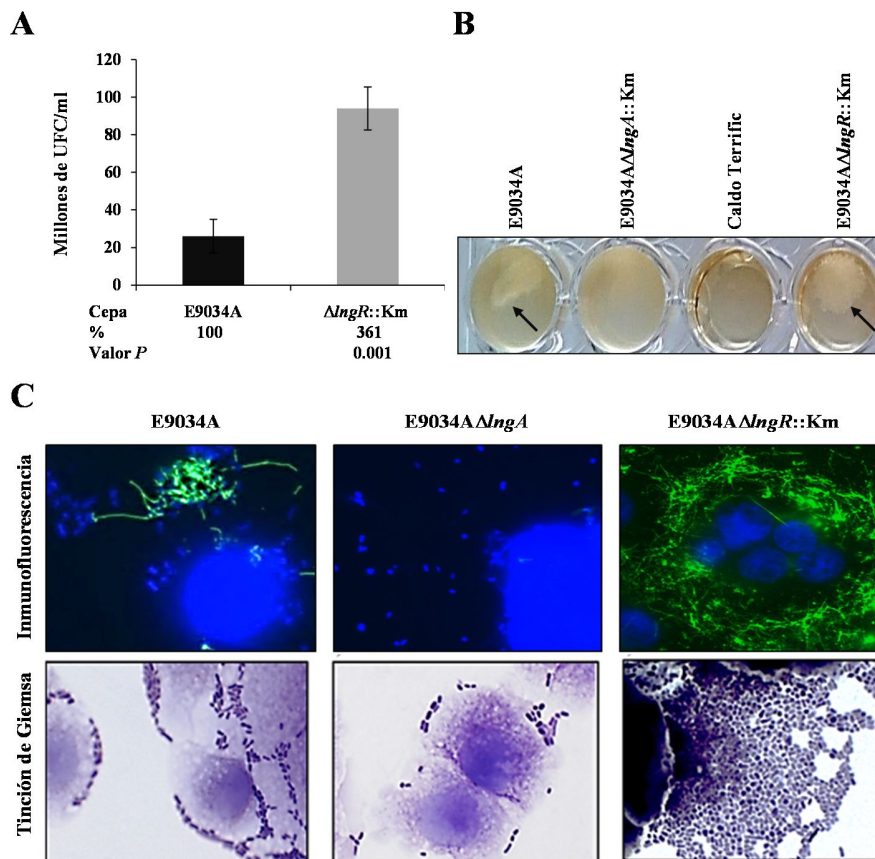


Figura 30. La cepa E9034A Δ IngR::Km expresa más *pili* CS21 y se adhiere tres veces más a las células intestinales HT-29 que la cepa silvestre E9034A. A) Cuantificación de las UFC del ensayo de adherencia a células intestinales HT-29. Estos valores fueron estadísticamente significantes con un valor de $p < 0.001$. B) Ensayo de auto-agregación bacteriana en medio de cultivo TB. El *pilus* CS21 favorece la interacción bacteria-

bacteria formando agregados bacterianos. C) Inmunofluorescencia y tinción de Giemsa de la infección de las células intestinales HT-29 con las cepas indicadas en la parte superior.

Con base en estos resultados, se propuso a *LngR* como un regulador negativo de la formación de CS21 y para confirmar que el fenotipo observado fue debido a la mutación del gen *lngR*, se procedió a la complementación de la cepa mutante con el gen *lngR* en *trans* (pUC*lngR*). Para la complementación del gen *lngR* se amplificó éste junto con -103 pb corriente arriba de su ATG, ya que esta secuencia contiene al posible promotor de *lngR* en el vector pUC19, dando origen al vector pUC*lngR*. Este vector se utilizó para complementar la cepa mutante E9034AΔ*lngR*::Km. Posteriormente se crecieron las cepas silvestre y mutante en presencia de células intestinales HT-29 por 6 h y se recolectó el sobrenadante de la infección para realizar ensayos de WB con suero anti-CS21. La cepa complementada con el vector pUC*lngR* no restauró el fenotipo de la cepa silvestre ya que seguía expresando más *LngA* que la cepa silvestre E9034A. Este resultado nos alertó sobre la posibilidad de que el vector no estuviera complementando la mutación, debido a un efecto polar del casete de kanamicina sobre los genes adyacentes. La mutación del gen *lngR* se generó con la inserción de un casete de kanamicina orientado en la misma dirección de los 13 marcos de lectura que se encuentran corriente abajo. Este hecho pudo haber provocado que el control de la transcripción de dichos marcos de lectura quedara bajo el promotor del gen de kanamicina. Para eliminar este obstáculo se procedió a la expulsión del casete de Km usando el vector pFLP2. Mediante la conjugación de la cepa SM10λ*pir*(pFLP2) (Hoang et al., 1998) con la cepa E9034AΔ*lngR*::Km y por recombinación homóloga de uno de los sitios FRT se insertó el plásmido pFLP2 en el plásmido que alberga al cluster *lng* de la cepa E9034AΔ*lngR*::Km. Dos sitios FRT que están flanqueando a los elementos insertados extras del vector en el sitio de la mutación

(*lngR*) se eliminan gracias a la ayuda de la enzima Flipasa, dando origen a una mutante curada sin casete de kanamicina denominada E9034A Δ *lngR*.

Con la mutante generada ya curada del marcador de selección se procedió a realizar los ensayos de WB y adherencia a células intestinales HT-29. Los resultados obtenidos mostraron una reducción tanto en la expresión de LngA por WB, como en la adherencia a células HT-29 de la cepa E9034A Δ *lngR* en comparación a la cepa silvestre E9034A (**Figura 31**). Estos resultados sugieren que LngR es un posible regulador positivo de CS21, al contrario de los resultados obtenidos previamente, la complementación de esta cepa confirmó el fenotipo real de la mutación, ya que la cepa E9034A Δ *lngR*(pUC*lngR*) se comportaba similar a la cepa silvestre.

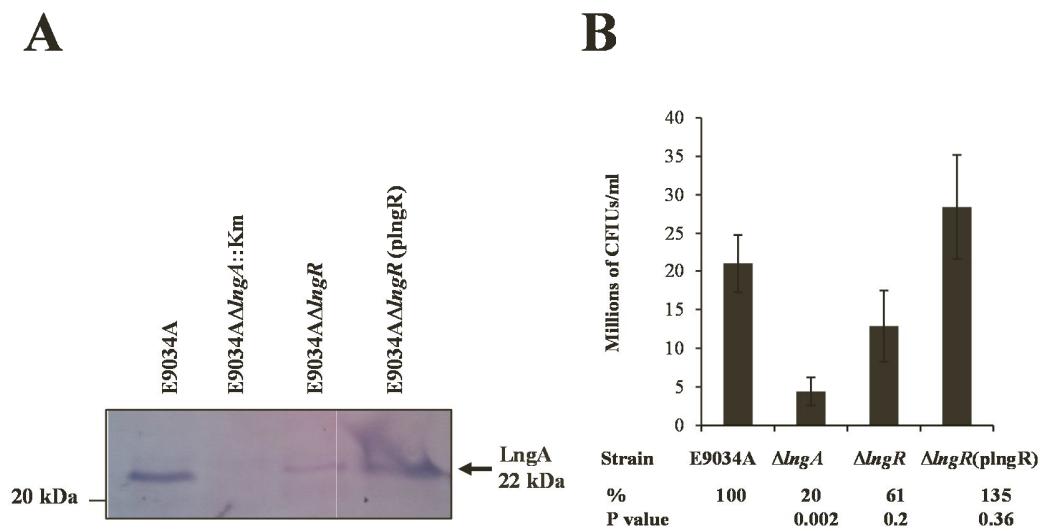


Figura 31. Caracterización de la cepa mutante E9034A Δ *lngR* y la cepa E9034A Δ *lngR*(pUC*lngR*). A) Western blot realizado con extractos proteico de las cepas indicadas y suero anti-CS21. B) Cuantificación de las UFC después de 6 h de infección a células HT-29. El porcentaje de adherencia de cada una de las cepas con respecto a la cepa silvestre E9034A se muestra debajo de la gráfica junto con el valor de *p*, se considera menor a 0.05 como estadísticamente significativo.

8.2.4 LngS es un probable activador transcripcional de la expresión del *pilus CS21*

LngS tiene homología con la familia de los reguladores transcripcionales AraC, con dominios típicos de unión hélice-giro-hélice que es la región que se une al surco menor del ADN de un promotor. Con el fin de elucidar el papel de LngS en la regulación de CS21 se diseñaron oligonucleótidos para la mutación y confirmación de la mutación del gen *lngS*. Se usó como molde el casete de Km para inactivar al gen *lngS*. Después de 5 intentos se pudo generar la mutante y con los antecedentes que se tenían con la mutación del gen *lngR* se procedió a la eliminación del casete de resistencia, previo a su caracterización (**Figura 32**). La caracterización de la mutante en el gen *lngS* sin el casete de resistencia se muestra en la página 87.

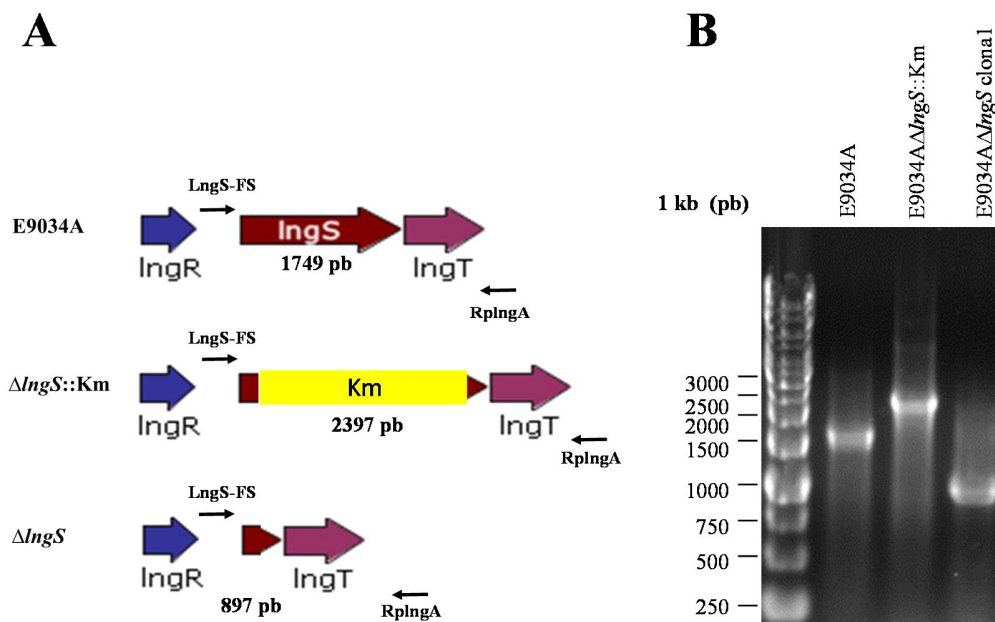


Figura 32. Confirmación de la mutación del gen *lngS* y eliminación del casete de resistencia. A) Esquema de la mutación del gen *lngS* y eliminación del gen de resistencia a kanamicina. B) Gel agarosa al 1% teñido con bromuro de etidio mostrando la mutación del gen *lngS* con y sin el casete de resistencia a Km

Otra estrategia empleada para demostrar el papel de LngS, fue sobre-expresar LngS. El gen *lngS* se clonó en el vector de expresión pBAD-Topo dando origen al vector pBADlngS, el cual se transformó en la cepa silvestre E9034A. La expresión del *pilus* CS21 se determinó a través de ensayos de WB e IF. Los resultados indican que la inducción del promotor *ara* con 0.02% de arabinosa en la cepa E9034A(pBADlngS) tres horas después de haber iniciado la infección incrementa los niveles de expresión de la proteína LngA y por lo tanto, el ensamble del *pilus* CS21 (**Figura 33**). Lo que sugiere que LngS es probablemente un regulador transcripcional positivo de la expresión del *pilus* CS21.

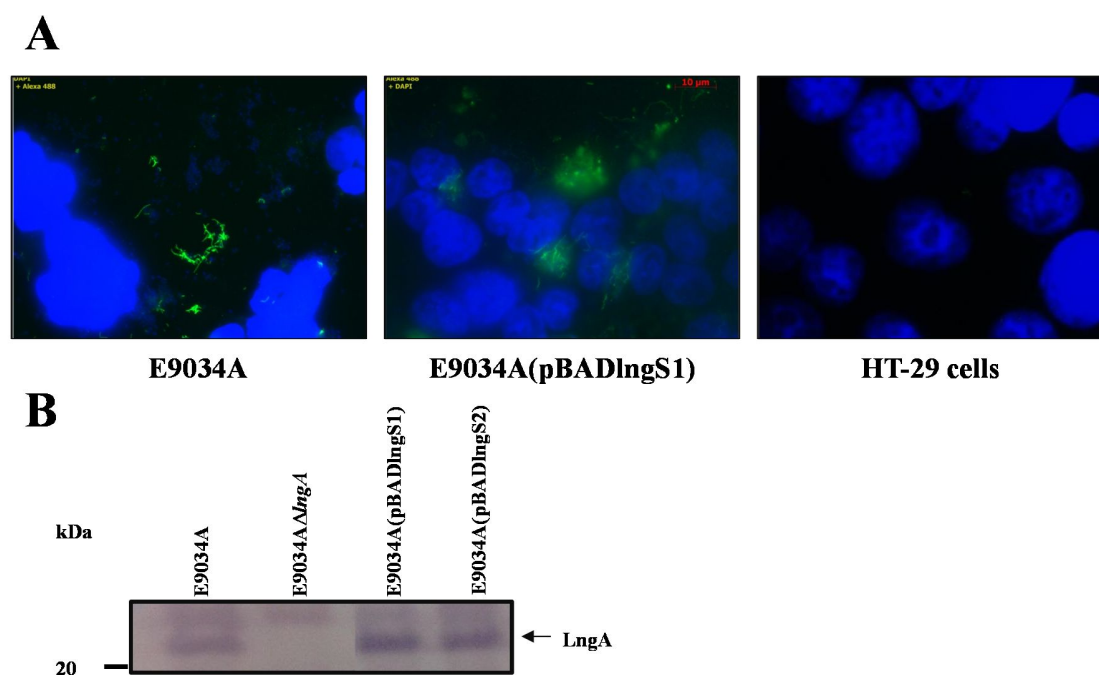


Figura 33. Sobreexpresión de LngS incrementa los niveles de expresión de LngA y del *pilus* CS21 por WB e IF. Los ensayos de inmunofluorescencia se realizaron con suero anti-CS21. Células HT-29 fueron infectadas con las cepas E9034A y dos clones E9034A(pBADlngS) por 6 h. Las bacterias y los núcleos de las células fueron tenidas con DAPI (azul) y la identificación de los anticuerpos anti-CS21 se llevó a cabo con la ayuda de un segundo anticuerpo anti-IgG de conejo conjugado con Alexa-fluor 488 (verde). Los sobrenadantes de la infección fueron ajustados y procesados para WB. Un incremento en la expresión de LngA se muestra en las clones E9034A(pBADlngS1) y E9034A(pBADlngS2).

8.2.5 H-NS es un represor de la expresión del *pilus* CS21

La regulación de los TIVPb, como BFP de EPEC y TCP de *V. cholerae* comparte a reguladores de la familia AraC (Roux et al., 2012). El producto del gen *perA* es el activador transcripcional de BFP y en el de TCP por el producto del gen *toxt* (Tobe et al., 1996; Yu y DiRita, 2002). En ambos casos el represor transcripcional es un miembro de la familia H-NS (Roux et al., 2012).

La proteína similar a histona H-NS, es un importante componente de la cromatina bacteriana que influye en la estructura del ADN y la expresión génica (Navarre et al., 2007). Con la finalidad de elucidar el papel de H-NS en la regulación del *pilus* CS21, se generó una mutación del gen *hns* en la cepa E9034A. El genoma de la cepa E9034A no se ha secuenciado, por lo tanto, para el diseño de los oligonucleótidos para mutar el gen *hns* se siguió la siguiente estrategia. Se diseñaron oligonucleótidos basados en la secuencia del gen *hns* de dos cepas de ETEC: B7A (CP005998.1) y H10407A (NC_017633.1) y una cepa de *E. coli* K-12 MG1655 (NC_000913.3) secuenciadas. Sitios idénticos dentro del gen *hns*, en ambas cepas se usaron para el diseño de los oligonucleótidos tanto para mutar como para confirmar la mutación. Con los oligonucleótidos diseñados se procedió a la amplificación del casete de Km para la mutación del gen *hns*. Se obtuvieron diez clonas con la inserción correcta del casete de Km en el gen *hns* (**Figura. 34**). Tres clonas se analizaron por WB para la expresión de LngA. La expresión de LngA aumentó en las mutantes del gen *hns* en comparación con la cepa silvestre E9034A (**Figura 35**). El ensamble del *pilus* CS21 en la mutante del gen *hns* se identificó mediante microscopia electrónica e inmuno-oro usando suero específico anti-CS21 (**Figura 35**).

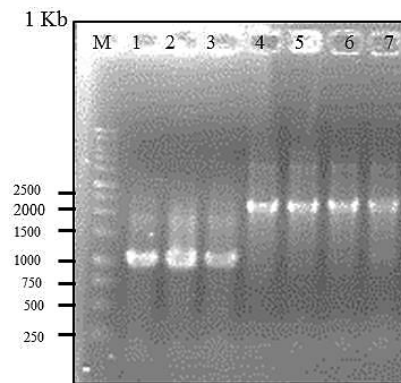


Figura 34. Mutación del gen *hns* en la cepa E9034A. La mutación del gen *hns* se llevó a cabo mediante la inserción del casete de Km y remplazo del gen *hns*. El carril 1 corresponde al amplicon de *hns* en la cepa silvestre E9034A (991 pb). Del carril 2 al 7 son amplicones de 6 posibles mutantes. Del carril 4 al 7, es el perfil de la inserción del casete de Km en el gen *hns* con un amplicon de 1818 pb.

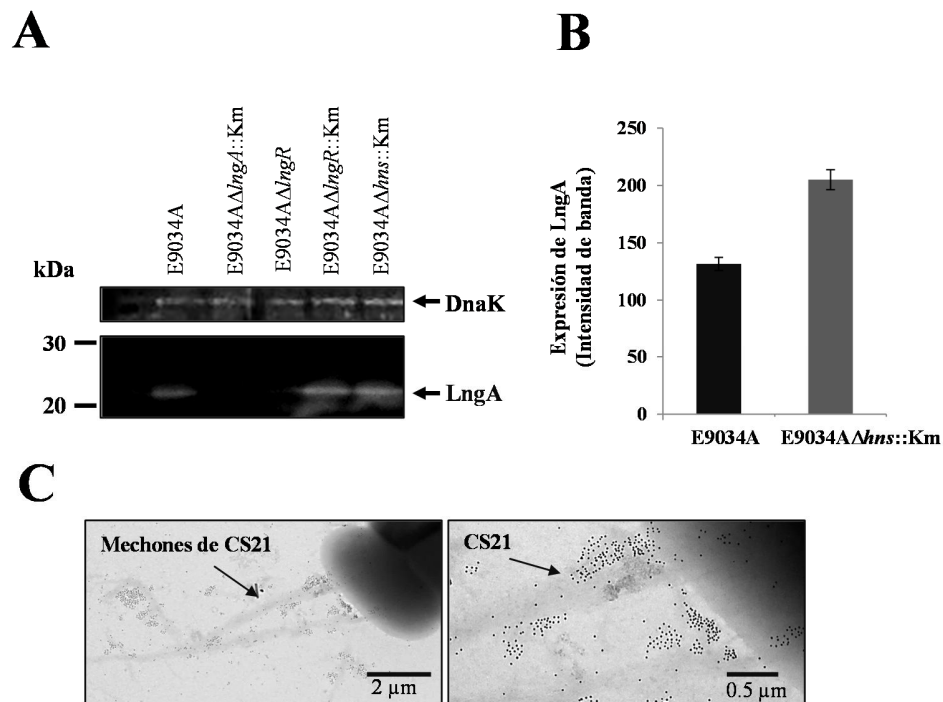


Figura 35. Sobreexpresión de LngA en la mutante del gen *hns*. A) Extractos proteicos de las bacterias recobradas de la infección de células HT-29 fueron procesados para Western blot. La unión del anticuerpo a la proteína LngA se detectó por quimioluminiscencia usando anticuerpos específicos anti-CS21. La identificación de la proteína DnaK fue usada como control de carga. En la cepa silvestre E9034A y mutante del gen *hns* se observó una inmunodetección con una banda de 22-kDa que corresponde a la pilina LngA. Interesantemente la cepa E9034AΔ*hns*::Km se observó una sobreexpresión de LngA en comparación con la cepa silvestre. B) La gráfica representa la cuantificación de la intensidad de banda del Western blot. El análisis se realizó con el programa Quantity One versión 4.2.3 de Bio-Rad. C) Sobreexpresión del *pilus* CS21 en la mutante del gen *hns* observada mediante IEM. Las barras en las imágenes de microscopía electrónica representan la escala de referencia.

Con las mutantes en los genes *lngR* y *lngS*, curadas del casete de Km y la mutante del gen *hns*, se procedió a la caracterización de estas mediante ensayos citometría de flujo y WB usando suero anti-CS21 (**Figura 36 y 37**). Células HT-29 fueron infectadas por 6 h con la cepas E9034A, E9034A Δ *lngA*, E9034A Δ *lngR*, E9034A Δ *lngR*(pUC*lngR*), E9034A Δ *lngS* y E9034A Δ *hns*::Km. Los sobrenadantes de los ensayos de adherencia se recobraron y se procesaron para citometría de flujo con suero anti-CS21. La expresión del *pilus* CS21 en la cepa E9034A Δ *hns*::Km aumentó con respecto a la cepa silvestre, al contrario de lo observado en la cepa E9034A Δ *lngR* donde la expresión de CS21 se abatió (**Figura 36**). La complementación de la cepa E9034A Δ *lngR* con el plásmido pUC*lngR*, incrementó la expresión de CS21 aún más que la cepa silvestre, probablemente debido al vector multicopia usado.

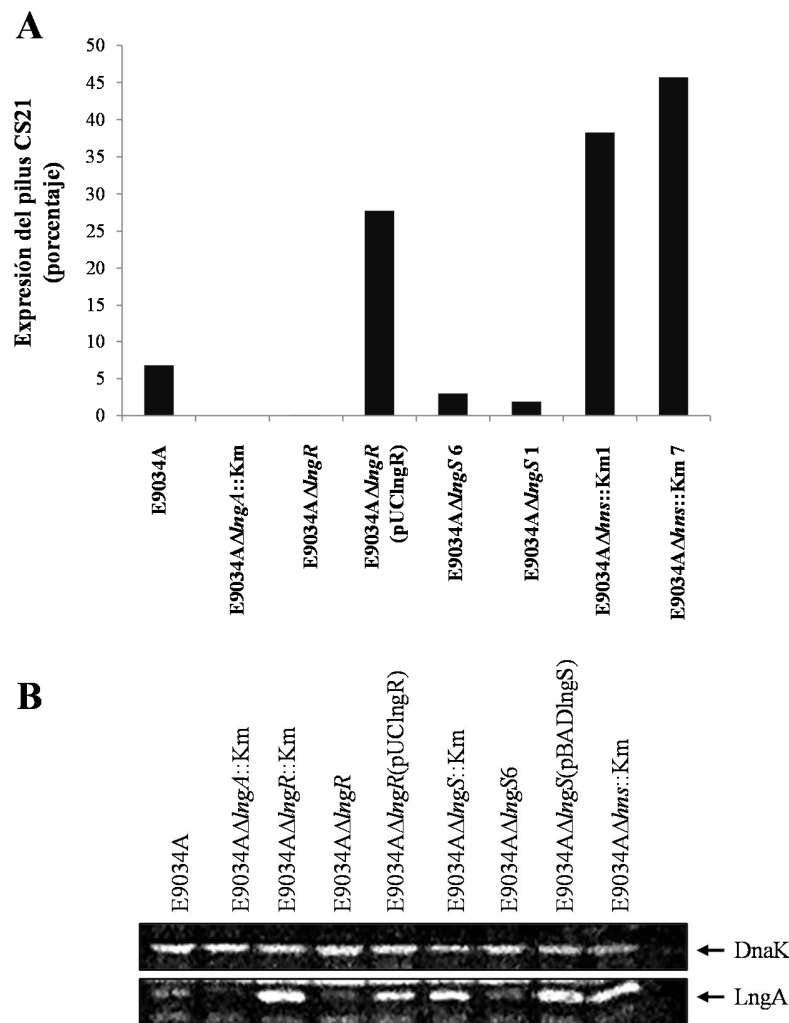


Figura 36. LngR y LngS son necesarios para la expresión de CS21 y H-NS muestra un efecto negativo. A) Cuantificación de la expresión del *pilus* CS21 en la cepa silvestre E9034A y mutantes de los genes *lngR*, *lngS* y *hns* por citometría de flujo. Sobrenadantes de los ensayos de adherencia a células intestinales HT-29 fueron recobrados y reaccionados con suero de conejo anti-CS21 y un conjugado anti-IgG de conejo conjugado con Alexa-fluor 488. Los porcentajes de expresión están representados en la gráfica. B) Extractos proteicos de las bacterias crecidas en medio PPLO fueron procesados para WB usando anticuerpos específicos anti-CS21. La identificación de la proteína DnaK fue usada como control de carga.

De los resultados obtenidos es posible suponer que LngR y LngS estén actuando como activadores transcripcionales de *lng* y que H-NS es el represor de *lng*. Un modelo de la regulación de CS21 se presenta en la **figura 38**.

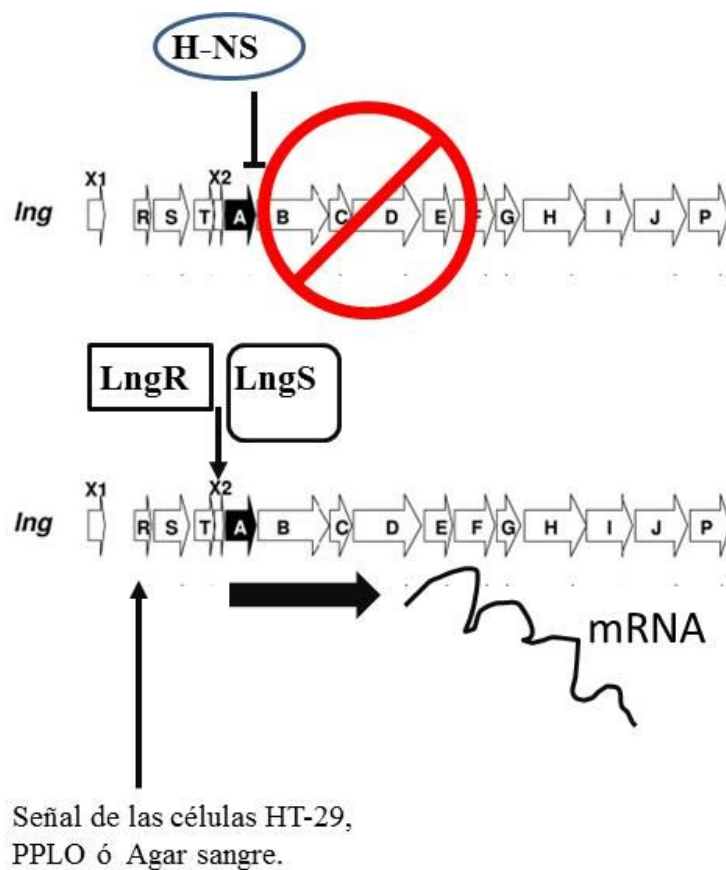


Figura 37. Modelo de la regulación del pilus CS21 en ETEC. H-NS actúa como un regulador negativo de la expresión del pilus CS21. En condiciones donde no se expresa el pilus CS21, se propone que H-NS pudiera estar unido al promotor o cerca del promotor de *lngA* y reprimir la transcripción de este. Una señal o señales de las células HT-29, de los medios de cultivo PPLO o agar sangre pudiera favorecer la transcripción de los genes *lngR* y *lngS*. La represión mediada por H-NS sería aliviada por los reguladores positivos LngR y LngS, permitiendo así la transcripción del gen *lngA*.

9.0 DISCUSIÓN

E. coli enterotoxigénica es el segundo agente etiológico responsable de las infecciones intestinales en niños menores de 5 años y es productora de dos TIVPb; CS21 y CFA/III (Kaper et al., 2004). Los *pili* CS21 y CFA/III de ETEC se encuentran evolutivamente relacionados; aunque, los genes de *lng* y *cof* son parálogos, las fimbrias CS21 y CFA/III son morfológicamente diferentes. El *pilus* CS21 es polar, delgado, largo y puede formar mechones; mientras, el *pilus* CFA/III es peritrico, delgado y corto (Girón et al., 1997). Las diferencias estructurales entre ambos *pili* pudieran estar generando diferencias funcionales, por ejemplo, el *pilus* CS21 es el único TIVPb con capacidad de mediar una movilidad tipo *twitching* (Mazariego-Espinosa et al., 2010). Además, el *pilus* CS21 de ETEC media la adherencia a células intestinales (HT-29, Caco-2 y T84) y favorece la formación de agregados bacterianos como un mecanismo de protección a los agentes antimicrobianos (Mazariego-Espinosa et al., 2010); (Clavijo et al., 2010). La interacción entre los *pili* favorece la formación de micro-colonias en *V. cholerae* por TCP y en EPEC por BFP, un paso crítico en la colonización intestinal (Girón et al., 1991; Kim et al., 2000). Bajo este contexto, observamos cambios en los fenotipos de adherencia a células intestinales y auto-agregación bacteriana en cada una de las cepas mutantes generadas en los genes *lng*.

El ensamble de los TIVP es un proceso que requiere la interacción de varias proteínas que son parte del sistema de ensamblaje y de exportación. Algunas proteínas que se ensamblan en la membrana interna se exportan a través de la membrana externa para formar la estructura oligomérica (Wolfgang et al., 2000; Roux et al., 2012). Las proteínas LngA y LngB son las únicas moléculas que se han descrito que participan en el ensamble del *pilus* CS21; aunque, existen otras proteínas que aún no se han estudiado. Por lo tanto,

resulta relevante caracterizar a otras proteínas que participan en la biogénesis del *pilus* CS21. Estas proteínas son codificadas por genes continuos (*lngR*, *lngS*, *lngT*, *lngA*, *lngB*, *lngC*, *lngD*, *lngE*, *lngF*, *lngG*, *lngH*, *lngI*, *lngJ*, y *lngP*) en el cluster *lng* localizado dentro de un megaplásmido (Gomez-Duarte et al., 2007). La homología que existe entre las proteínas Lng con las proteínas que participan en la biogénesis de CFA/III, TCP y BFP, fue una herramienta importante para sugerir la función de cada una de ellas. Además, se identificaron motivos, dominios funcionales, localización subcelular, cruces transmembranales y péptido señal de las proteínas Lng.

Recientemente, se han desarrollado algunos algoritmos para la predicción de la localización subcelular, basados en redes neuronales (Reinhardt y Hubbard, 1998; Boden y Hawkins, 2005) y en sistemas relacionados con el aprendizaje supervisado, empleados en clasificaciones y regresiones SVM (Matsuda et al., 2005). Los datos bioinformáticos y experimentales corroboraron la localización extracelular de las proteínas LngA, BfpA y TcpA (Taylor et al., 1987; Girón et al., 1991; Giron et al., 1994). La localización subcelular de LngB se ha predicho como una proteína extracelular; sin embargo, no existen datos experimentales acerca de su localización subcelular y tampoco de sus homólogos CofB y TcpB. La función de LngC aún no se ha descrito y únicamente se ha predicho como una proteína de membrana externa por predicción de localización subcelular. Este resultado concuerda con la localización experimental de sus homólogos BfpG y TcpQ descritos como proteínas de membrana externa (Bose y Taylor, 2005; Daniel et al., 2006).

La localización subcelular de la proteína LngD (proteína de membrana externa) y LngP (proteína de membrana interna) corresponde con la localización de sus homólogos BfpB/TcpC y BfpP/TcpJ. LngH y TcpT son homólogas y su localización subcelular se ha predicho como proteínas localizadas en el citoplasma. La proteína LngH tiene motivos típicos Walker A y B, que están asociados con la unión a fosfato y la hidrólisis de ATP (Iyer et al., 2004). Análisis bioinformáticos han predicho que las proteínas LngH y TcpT carecen de dominios transmembranales como se ha descrito para otras ATPasas (Tripathi y Taylor, 2007). Por lo tanto, se sugiere que la proteína LngH pudiera estar localizada en la membrana interna, favorecida por interacciones de otras proteínas ancladas en la membrana interna y que participan en el ensamble del *pilus* CS21. La determinación estructural de las proteínas de membrana, es complejo debido a las dificultades que existen para su cristalización. Por lo tanto, la predicción de la estructura secundaria y terciaria de proteínas transmembranales es de gran interés para elucidar como estas participan en el ensamble de los TPIV. LngD, LngC, LngI y LngP fueron predichas como proteínas de membrana cuya localización es consistente con otras proteínas homólogas de otros TIVP (Roux et al., 2012). La predicción de cruces transmembranles para las proteínas LngA, LngB, LngE, LngF, LngI y LngP utilizando tres predictores mostraron datos consistentes. Las proteínas LngA, LngB, LngE y LngF presentaron un solo cruce transmembranal (bitópicas); mientras, las proteínas LngI y LngP presentaron 3 y 7 cruces transmembranales, respectivamente (politópicas).

Basados en otros modelos de la biogénesis de otros TIVP se sabe que las pilinas son translocadas del citoplasma al espacio periplasmico por el sistema Sec y que esta translocación es dependiente del péptido señal de las proteínas. Las proteínas en el periplasma se mantienen unidas a la membrana interna favorecidas por un cruce

transmembranal, hasta su oligomerización. LngE fue predicha como una proteína citoplasmática con un cruce transmembranal, dato que corresponde con su homólogo TcpR del *pilus* TCP (Tripathi y Taylor, 2007). Sin embargo, la predicción de la localización subcelular de la proteína LngF no fue congruente entre los tres predictores; por tal motivo, su localización en el modelo fue con base a su homólogo TcpD del *pilus* TCP. Las proteínas LngE, LngF y LngI fueron predichas para estar localizadas en la membrana interna, de acuerdo a sus proteínas homólogas en TCP y BFP (Roux et al., 2012). La proteína LngI al igual que sus homólogos TcpE y BfpE son moléculas politópicas de membrana interna, con tres y cuatro cruces transmembranales, respectivamente (Blank y Donnenberg, 2001; Crowther et al., 2004; Roux et al., 2012). LngP es una proteína politópica con 7 cruces transmembranales como se ha descrito para sus homologos TcpJ y BfpP (Roux et al., 2012).

Los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* se mutaron en la cepa E9034A para elucidar la función potencial de cada una de sus proteínas en el ensamble del *pilus* CS21. Los resultados demostraron que los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* se requieren esencialmente para el ensamble del *pilus* CS21. Las cepas mutantes generadas mostraron la ausencia del *pilus* CS21 cuando se visualizaron por IF e IEM. Es importante hacer énfasis, que las mutaciones no generaron efectos polares sobre los genes adyacentes, ya que la complementación en *trans* en las cepas mutantes con el gen silvestre de cada una de ellas, restauró el ensamble del *pilus* CS21. La adherencia a células intestinales HT-29 y auto-agregación bacteriana son dos fenotipos atribuidos al *pilus* CS21. Por lo tanto, las mutaciones de los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* que afectaron el correcto ensamble del *pilus* CS21, afectaron también los fenotipos mediados por CS21.

Las pilinas de los TIVP se procesan por prepilinas peptidasas que cortan el péptido señal y metilan el primer aminoácido de la pilina madura. La pilina madura se transloca al espacio periplasmico donde comienza a oligomerizar para formar la estructura fimbrial; la cual, crece hacia el espacio extracelular a través de la secretina de membrana externa. La pilina LngA tiene un péptido señal en el N-terminal que es procesado por una prepilina péptidasa y hemos propuesto que el gen *lngP* codifica a LngP, una proteína con dominios de prepilina peptidasa.

BfpP de EPEC y TcpJ de *V. cholerae* son las prepilina peptidasas de BfpA y TcpA, respectivamente. Ambas tienen una identidad con LngP del 31.9% y 33%, respectivamente. Los análisis filogenéticos con las proteínas LngP, BfpP y TcpJ demostraron que son ortólogas y que probablemente pudieran descender de un ancestro común (Gomez-Duarte et al., 2007). Los análisis bioinformáticos de la proteína LngP indican la presencia de dos motivos (CXXC y GXCXXC) en la región N-terminal que pudieran participar en la actividad metiltransferasa y de dos residuos de aspartato altamente conservadas en el dominio C-terminal que pudieran estar involucrados con la actividad péptidasa de la pilina (Roux et al., 2012). La mutante en el gen *lngP* se examinó mediante ensayos de adherencia a células intestinales y por IF. Las imágenes de IF de la cepa silvestre E9034A mostraron filamentos largos y polares característicos del *pilus* CS21; sin embargo, la cepa mutante del gen *lngP* no observaron estructuras características del *pilus* CS21. Estos resultados demostraron que la proteína LngP codificada por el gen *lngP* es requerida para el ensamble del *pilus* CS21.

Para demostrar que LngP es la prepilina peptidasa que procesa a LngA se analizaron los extractos proteicos de la cepa E9034A y la cepa mutante en el gen *lngP*, para WB usando suero de conejo anti-LngA. En la imagen de WB se observó una reacción positiva

con una banda de peso molecular de 22-kDa en la cepa silvestre E9034A y dos bandas de peso molecular de 22-kDa y 25-kDa en la mutante del gen *lngP*. La banda de 22-kDa corresponde al peso molecular de la pilina madura y la de 25-kDa al peso molecular estimado de la pilina inmadura. En la cepa mutante del gen *lngP* la intensidad de la banda de 25-kDa fue mayor que la banda de 22-kDa. Este resultado sugiere que LngP no es la única prepilina peptidasa que procesa a LngA y que probablemente otro gen que codifica a otra prepilina peptidasa pudiera estar presente en el genoma de la cepa E9034A. Más de 30 genes en el genoma de *E. coli* K-12 son homólogos a genes de la vía secretoria, involucrada en la secreción o el ensamble de los TIVP, incluyendo dos genes homólogos que codifican a prepilin peptidasas *gspO* y *pppA* (Francetic y Pugsley, 1996; Francetic et al., 1998; Pugsley y Francetic, 1998). El gen *gspO* codifica a una enzima funcional; sin embargo, su transcripción es muy baja en condiciones de laboratorio (Francetic y Pugsley, 1996). Por otro lado, el gen *pppA* codifica a una prepilin peptidasa funcional capaz de procesar sustratos como: prePulG (pre-pseudopilina de *Klebsiella oxytoca*), prePile (pre-pilina tipo IV de *Neisseria gonorrhoeae*) y prePpdD o preHcpA (pre-pilina tipo IV de *Escherichia coli* enterohemorrágica) (Whitchurch y Mattick, 1994; Francetic et al., 1998; Xicohtencatl-Cortes et al., 2007).

La proteína LngA se procesó completamente en la cepa de *E. coli* TOP10 transformada con el gen *lngA* en *trans*. Estos datos sugirieron que la cepa *E. coli* TOP10 codifica para una prepilin peptidasa funcional y capaz de cortar sustratos como la prepilina LngA. El gen *pppA* es un gen conservado en el genoma de las *E. coli* y probablemente sea este el que codifique a la prepilin peptidasa responsable de procesar a LngA en la cepa *E. coli* TOP10. El procesamiento parcial de LngA observado en la cepa E9034AΔ*lngP*, sugiere que las proteínas PppA y/o GspO pueden estar implicadas en el procesamiento de

LngA. Sin embargo, estas proteínas podrían estar siendo expresadas a bajos niveles en la cepa E9034A en comparación con la cepa *E. coli* TOP10 o la afinidad de estas por la proteína LngA ser menores que LngP. Curiosamente, el procesamiento parcial de LngA en la cepa E9034A Δ LngP no fue suficiente para el ensamble del *pilus* CS21, como se demuestra en las imágenes de IF e IEM.

Un fragmento de 5-kb obtenido por restricción enzimática con la enzima BamHI a partir del plásmido pE9034A fue suficiente para el ensamblaje del *pilus* CS21 y el fenotipo de auto-agregación bacteriana en una cepa de *E. coli* K-12 (DH5 α F'IQ) (Giron et al., 1994). Análisis de la secuencia parcial del plásmido pE9034A (número de acceso GenBank: EF595770.1), evidenció que el fragmento de 5-kb alberga solamente a los genes *lngR*, *lngS*, *lngT*, *lngA* y *lngB* (secuencia truncada con 1106 pb). Estos datos sugieren que otros genes en el cromosoma de *E. coli* K-12 codifican para proteínas involucradas en el ensamble de TIVP y podrían reemplazar las funciones de las proteínas Lng restantes. Además, la cepa *E. coli* K-12 contiene 16 genes en siete diferentes *locus* de su cromosoma que codifican para proteínas del ensamble de los TIVP (Sauvonnet et al., 2000). Estudios posteriores son necesarios para poder elucidar las diferencia entre los genes de la cepa *E. coli* K-12 y la cepa ETEC E9034A, así como su expresión. Los resultados aquí presentados demuestran que LngP participa en el ensamble del *pilus* CS21 actuando como una de las prepilina peptidasas que procesa a LngA.

Las imágenes de IEM de las cepas mutantes en los genes *lngA*, *lngC*, *lngD*, *lngH* y *lngP*, evidenciaron el marcaje con partículas de oro de una estructura sobre la membrana externa, pero no estructuras de tipo CS21. Sin embargo, en la mutante del gen *lngB* este marcaje no se observó. En la cepa E9034A y en las cepas complementadas se observó marcaje con partículas de oro sobre estructuras largas, polares, delgadas y en forma de

mechones, características del *pilus* CS21. Estos resultados permitieron hipotetizar que el suero anti-CS21 podría estar reconociendo a las proteínas LngA y LngB. La cepa E9034A Δ lngP presentó más marcaje, lo que sugiere que probablemente tanto LngA como LngB estén siendo reconocidas por el suero anti-CS21. El WB de la cepa silvestre E9034A con el suero anti-CS21 mostró una reacción positiva con bandas de 22-kDa y 57-kDa, pesos moleculares que corresponden al peso molecular teórico de las proteínas LngA y LngB.

La mutación del gen *lngB* no afectaba la expresión y el procesamiento de LngA, sin embargo, el ensamble del *pilus* CS21 si se afectó. Estos datos sugieren que la proteína LngB podría ser necesaria para la translocación u oligomerización de LngA y por lo tanto para el correcto ensamble del *pilus* CS21. La complementación de la cepa E9034A Δ lngB con el vector pUClngB no mostró algún efecto en la expresión y procesamiento de LngA, sin embargo, las imágenes de IF mostraron *pilus* CS21 de mayor longitud que los de la cepa silvestre E9034A. Las mediciones mediante microscopia confocal de la longitud de los *pili* CS21, demostraron que los mechones de la cepa E9034A Δ lngB(pUClngB) fueron 2.5 veces más largos que los de la cepa silvestre E9034A. Por el contrario, Kolappan et al., 2015 encontraron que la sobreexpresión de LngB en *trans* no afectó la expresión ó longitud de CS21, sin embargo, en este estudio no se llevaron a cabo mediciones de estas fimbrias como en este estudio. Recientemente, la estructura tridimensional de la proteína CofB (homólogo de LngB con 78% de identidad) se resolvió recientemente y permitió proponer un modelo en el que CofB inicia el ensamble del filamento de CFA/III (Kolappan et al., 2015). En este modelo se propone que CofB está localizado en la punta del filamento y es consistente con el papel de CofB como iniciador de la piliación de CFA/III (Kolappan et al., 2015). LngB es una subunidad pilina menor de CS21, cuya localización subcelular predicha es extracelular y experimentalmente por IEM se demuestra que es extracelular y

específicamente en la punta del *pilus* CS21. Se llevarán a cabo experimentos adicionales para determinar si LngB es una adhesina del *pilus* CS21.

Las proteínas codificadas por los genes *lngC* y *lngD* tienen un papel importante en el ensamble del *pilus* CS21. La mutación de estos genes no afectó la expresión de LngA, pero sí el procesamiento y ensamble del *pilus* CS21. Se propone en este estudio que la prepilina LngA observada por WB en la cepa E9034AΔ*lngD* sea resultado de una inactivación alostérica de LngP (prepilin peptidasa), debido a la acumulación de LngA en el periplasma por la ausencia de la secretina de membrana externa LngD. La presencia de la prepilina LngA en la cepa E9034AΔ*lngC*, sugiere que LngC pudiera requerirse para la estabilidad y/o localización de LngD en la membrana externa. Esto debido a que TcpQ, el homólogo de LngC con 24% de identidad, se requiere para la estabilidad y la localización en la membrana externa de TcpC (secretina de TCP) (Bose y Taylor, 2005). LngC tiene 20.3% de identidad con BfpG y al contrario de TCP la localización a la membrana externa de BfpG es dependiente de BfpB (secretina de BFP) (Daniel et al., 2006). La mutación de los homólogos de *lngC* y *lngD* en BFP “*bfpB* y *bfpG*” no afectan la expresión ó el procesamiento de BfpA, pero inhibe el ensamble de BFP (Anantha et al., 2000).

En este estudio, también se demostró que una peptidasa, probablemente LngP corta la proteína recombinante LngC-His, ya que se observaron en los geles de poliacrilamida teñidos con azul de Coomassie dos bandas que corresponden a LngC sin procesar y procesada. El sitio de corte del péptido señal de LngC predicho daría como productos dos proteínas de pesos moleculares de 14-kDa y 16-kDa que concuerdan con los pesos moleculares de las bandas observadas en el gel de poliacrilamida. Similarmente, la proteína purificada BfpG migra en geles de poliacrilamida como un doblete. Los autores proponen que la proteína BfpG madura no es una lipoproteína y muestran que la proteína BfpG existe

en dos formas dependiendo del sitio de corte por la peptidasa I, esto se demostró por secuenciación del amino terminal (Daniel et al., 2006). Aunque en este estudio no se secuenció la proteína procesada y sin procesar de LngC, los pesos moleculares concuerdan perfectamente con los predichos para la proteína completa y la proteína madura procesada.

Los TIVPb tienen una organización genética muy similar de entre 12-14 genes agrupados en operones (Roux et al., 2012). Las proteínas codificadas por los genes del operon *lng* tienen homología con las proteínas involucradas en el ensamble de CFA/III, TCP y BFP. Esta homología, los análisis bioinformáticos y los resultados experimentales fueron de gran ayuda para proponer un modelo de ensamble del *pilus* CS21. En este modelo, LngA es la subunidad estructural pilina que oligomeriza para formar el filamento de CS21. La hélice hidrofóbica del amino N-terminal (péptido señal) de las subunidades pilina está anclada a la membrana interna y las subunidades pilinas se incorporan en el *pilus* creciente. LngP es una prepilin peptidasa de la membrana interna que procesa a la prepilina LngA antes de ser oligomerizada. La proteína LngB es una subunidad de pilina menor que se localiza en la punta y que es requerida para iniciar la oligerización de LngA. LngC se predijo a estar localizada en la membrana externa y con base a su homología con TcpQ de TCP, podría necesitarse para la estabilidad LngD y/o localización en la membrana externa. El C-terminal de LngD tiene un dominio conservado en la superfamilia de las secretinas. Las secretinas son proteínas de membrana externa que se utilizan por los TIVP para translocar las subunidades a través de la membrana externa. El ensamble también requiere de una proteína de unión a nucleótidos para proporcionar energía para el ensamble. El gen *lngH*, codifica una probable ATPasa, “LngH” que podría proveer la energía necesaria para el ensamble del *pilus* CS21. Los genes restantes *lngT*, *lngE*, *lngF*, *lngG*, *lngI* y *lngJ* se incluyeron en este modelo, aunque no fueron mutados en este estudio. En conclusión, el

modelo del ensamblaje del *pilus* CS21 propuesto en este estudio es muy similar a los modelos de ensamblaje de otros TIVP como TCP y CFA/III.

Regulación del *pilus* CS21

La terminación Rho-independiente es un mecanismo importante de paro transcripcional en bacterias que controla tanto la terminación 3' normal como una amplia gama de eventos de atenuación transcripcional. En el cluster de los genes *lng* se encontraron tres terminadores Rho-independientes: uno a 40 pb del ATG del gen *lngS*, el segundo a 384 pb del ATG del gen *lngA* y el tercero a 994 pb del ATG del gen *lngD*. La posible terminación Rho-independiente en la secuencia del gen *lngS* podría servir de paro de la transcripción de los reguladores transcripcionales *lngR* y *lngS*. Dentro de la secuencia que codifica al gen *lngA* se encontró una región predicha a ser una terminación Rho-independiente, sin embargo, esta región podría servir de atenuación de la transcripción de los genes corriente abajo del gen *lngA*.

Dentro de los 14 marcos de lectura involucrados en la biogénesis de CS21, existen dos genes con dominios regulatorios que se localizan corriente arriba de *lngA*: *lngR* y *lngS*. LngR tiene homología con la familia PapB, la cual incluye reguladores fimbriales y afimbriales, los cuales activan o reprimen la transcripción mediante unión al ADN. Los sitios de unión al ADN y de oligomerización de PapB, se conocen (Xia y Uhlin, 1999) y están presentes en la secuencia de amino ácidos de la proteína LngR. PapB en UPEC, es un regulador transcripcional del promotor de *pap* que a bajos niveles estimula la expresión del operón *pap*, pero a niveles altos lo reprime (Forsman et al., 1989). LngS tiene homología con la familia de los reguladores transcripcionales AraC, ambos con dominios típicos de unión hélice-giro-hélice que es la región que se une al surco menor del ADN de un promotor. Entre los reguladores de los TIVP clase B de la familia AraC, se encuentran

CofS, ToxT y PerA, los cuales tienen una identidad con LngS del 69%, 17% y 20%. La proteína LngR inicialmente se presumía que era un regulador negativo del operón *lng*, ya que la mutante del gen *lngR* con el casete de resistencia, expresaba más LngA y ensamblaba más CS21 que la cepa silvestre E9034A. Sin embargo, la introducción del gen silvestre en *trans* no restauró el fenotipo de la cepa parental. Este resultado nos hizo pensar que probablemente el casete de kanamicina podría estar teniendo efectos polares sobre los genes que se encuentran corriente abajo. La cepa curada del casete de resistencia “E9034AΔ*lngR*” al contrario del fenotipo previo, expresaba menos proteína LngA y por lo tanto se adhería menos que la cepa silvestre. La cepa complementada restauró los fenotipos de adherencia a células intestinales y expresión de CS21. Estos resultados mostraron que el papel de LngR es el de favorecer la expresión del *pilus* CS21. El gen *lngS* se mutó en la cepa E9034A y se procedió a eliminar el casete de resistencia para evitar los mismos problemas observados con la cepa E9034AΔ*lngR*::Km. Las cepas E9034Δ*lngS* y E9034Δ*lngS*(pBAD*lngS*) se examinaron mediante IF, citometría de flujo y WB, la mutante expresó menos *pilus* CS21 que la cepa silvestre y la cepa complementada restableció el fenotipo silvestre. Estos resultados sugieren que en la regulación del *pilus* CS21, tanto LngR como LngS podrían actuar como reguladores positivos de *lng*. En otros TIVP clase b como BFP y TCP, un miembro de la familia AraC actúa como activador de la transcripción de *bfp* y *tcp*, respectivamente y la proteína H-NS actúa como represor (Roux et al., 2012). La proteína asociada a nucleóide, H-NS participa en la regulación de la expresión de varios genes de virulencia de enteropatógenos como EPEC, ETEC, EIEC, EHEC, *Shigella flexneri* y *V. cholerae*. En estas bacterias la función principal es como regulador negativo, el cual reprime la transcripción en respuesta a cambios en el medio, como temperatura y osmolaridad (Navarre et al., 2007). La represión ejercida por H-NS se retira cuando otras

proteínas la desplazan del ADN, como en el caso de ToxT en los promotores *ctx* y *tcpA* de *V. cholerae* (Yu y DiRita, 2002). En el caso del *pilus* CS21 se propone a H-NS como el represor de la transcripción de *lng* y que LngR y/o LngS podrían desplazarlo y favorecer la transcripción del operon *lng* en la cepa E9034A.

El mecanismo de acción de H-NS depende de tres componentes estructurales; un dominio N-terminal con capacidad de dimerización/oligomerización (Renzoni et al., 2001), un dominio C-terminal con capacidad de unión a ácidos nucleicos (Shindo et al., 1995) y una región de unión flexible entre ambos que permite la actuación de forma independiente de ambos dominios. Se han propuesto dos mecanismos mediante los cuales se modula la transcripción por parte de la proteína H-NS. Estos dos mecanismos se conocen como silenciamiento transcripcional (Goransson et al., 1990) y represión vía topología del ADN (Hulton et al., 1990; Tupper et al., 1994) y no son excluyentes. Estos mecanismos dependen de la presencia de regiones de ADN con curvatura intrínseca que flanquean a la secuencia promotora del gen y que favorecen la unión de la proteína H-NS. Bajo este contexto H-NS podría estar formando una cremallera en el ADN que flanquea al promotor de *lngA*, produciendo un complejo H-NS-ADN que impide el avance de la RNA polimerasa en la fase de iniciación o elongación de la transcripción. Se sabe que H-NS puede reconocer y silenciar la expresión de ADN foráneo con alto contenido de adenina y timina comparado con el genoma residente (Navarre et al., 2007). El cluster que alberga los 14 genes involucrados en la biogénesis del *pilus* CS21 se encuentra en un plásmido de 54 kb y en el caso de una cepa clínica de ETEC 73332 específicamente en el plásmido pEcETEC73332b (datos sin publicar). Este cluster tiene un contenido de GC del 37% y está flanqueado por secuencias de inserción (IS) y transposasas lo que indica que es ADN foráneo (datos sin publicar). El contenido de GC de la cepa de ETEC 73332 y H10407 es de 50.5% y 50.8%,

respectivamente (Crossman et al., 2010). La región intergénica corriente arriba del gen *IngA* tiene un contenido de GC del 33%. Estos datos nos hacen hipotetizar que H-NS podría estar uniéndose a esta región y reprimir la transcripción de *Ing*. Se requieren experimentos de EMSA y ADN *footprinting* con la proteína H-NS y la región intergenica para confirmar esta hipótesis.

10.0 CONCLUSIONES

Para la biogénesis de TIVP b se requiere de 11 a 14 genes (Roux et al., 2012). En el presente estudio se mostró que los genes *lngA*, *lngB*, *lngC*, *lngD*, *lngH* y *lngP* se necesitan para el correcto ensamble del *pilus* CS21 debido a que las mutantes en dichos genes no formaron el *pili* CS21. Además se observó una reducción en adherencia a células intestinales HT-29 y en la auto-agregación bacteriana lo cual es dependiente de la presencia de dicha estructura. LngA es la subunidad estructural del *pilus* CS21 y LngB es una subunidad menor que se asocia a la punta y que se requiere para iniciar el ensamble del *pilus* CS21. LngP es la prepilin peptidasa responsable de procesar a LngA y probablemente LngC. LngC y LngD son proteínas necesarias para el correcto procesamiento de la prepilina LngA y en ensamble del *pilus* CS21. LngH se predice que tienen dominios de ATPasa que probablemente estén involucrados en proveer energía para el ensamble del *pilus* CS21 y para el movimiento tipo *twitching*. Dos marcos de lectura corriente arriba del gen *lngA*, codifican para los probables reguladores positivos de la transcripción de *lng*; LngR y LngS podrían aliviar la represión mediada por H-NS. La sobreexpresión de LngR o LngS favorece la sobreexpresión del *pilus* CS21 y la adherencia de ETEC a células intestinales HT-29.

11.0 APÉNDICES

11.1 Amortiguadores y soluciones

1. Amortiguador de fosfatos (PBS 1x): para la preparación de 1 litro se pesaron las siguientes sales: 8 g de NaCl, 0.2 g de KCl, 1.44 g de Na₂HPO₄ y 0.24 g de KH₂PO₄. Las sales se disolvieron en 800 ml de agua destilada y se ajustó la disolución a un pH de 7.4 con HCl 1M. Se agregó agua destilada a un volumen total de 1 l. La disolución se esterilizó por calor húmedo durante 30 min a 121°C y 20-22 PSI.
2. PBS-Tween: para 1 litro de solución se agregó 5 ml de Tween-20 (Amresco LLC, OH, EE. UU.) a 9995 ml de PBS 1x.
3. Solución de Tritón al 0.1%: para 100 ml de solución se tomaron 100 µl de Tritón X-100 (Sigma-Aldrich Co. LLC, MO, EE.UU.) y se mezclaron con 9.900 ml de PBS 1x. La solución se esterilizó por calor húmedo.
4. Disolución amortiguadora de Laemmli 4x: para 100 ml se pesaron 3.028 g de Tris-base, 8 g de SDS, 0.1 g de azul de bromofenol y se disolvieron en 60 ml de agua destilada. Se agregó 40 ml de 100% glicerol, se mezcló por inversión y se ajustó el pH a 6.8 con HCl 1M. Alícuotas de 1 ml se guardaron a -20°C hasta su uso.
5. Disolución amortiguadora de transferencia: para 1 litro se pesaron 5.82 g de Tris-base y 2.95 g de glicina y se disolvieron en 800 ml de agua destilada y 200 ml de metanol.
6. Disolución amortiguadora de corrida: para 1 litro se pesaron 6.06 g de Tris-base, 28.22 g de glicina y 1 g de SDS y se disolvieron en 1 litro de agua destilada.
7. Disolución amortiguadora de lisis pH 8.2: para preparar 300 ml se pesaron 144.15 g de urea, 0.3 g de Tris-HCl, 4.14 g de NaH₂PO₄ y 0.204 g de imidazol. Se disolvieron en 200 ml de agua destilada y se ajustó el pH a 8.0 con 1M de NaOH y se aforó a 300 ml.
8. Disolución de azul de Coomassie: para preparar 500 ml se pesaron 1.25 g de azul de Coomassie Blue G250 y se disolvieron en 375 ml de agua destilada, más 75 ml de etanol y 50 ml de ácido acético.

9. Solución desteñidora de azul de Coomassie: para 1 litro se tomaron 500 ml de metanol más 100 ml de ácido acético y 400 ml de agua destilada.

11.2 Cepas bacterianas, medios de cultivo y línea celular

Las cepas bacterianas y plásmidos utilizados en este estudio se describen en la **Tabla 7**. Las cepas se conservaron a -70°C en caldo Luria-Bertani (LB) (BD Difco, NJ, EE.UU.) suplementado al 20% de glicerol. Las cepas se cultivaron en agar LB o en caldo LB a 30, 37, o 42°C de acuerdo al tipo de ensayo requerido para este estudio. Las cepas de ETEC se cultivaron en caldo PPLO (Pleuro-Pneumonia-Like Organisms, por sus siglas en inglés) (BD Difco, NJ, EE.UU.) o en medio TB (Terrific Broth, por sus siglas en inglés) (Amresco LLC, OH, EE.UU.) para la expresión del *pilus* CS21. Los medios de cultivos se suplementaron con kanamicina (50 $\mu\text{g/ml}$), ampicilina (100 $\mu\text{g/ml}$) y cloranfenicol (30 $\mu\text{g/ml}$) de acuerdo al tipo de ensayo requerido (por ejemplo: generación de cepas mutantes y complementadas *en trans*).

Los ensayos de adherencia a células intestinales se realizaron utilizando la línea celular HT-29 ATCC® HTB-38™ (células de adenocarcinoma de colon). Las células HT-29 se mantuvieron en medio DMEM-HG (Dulbecco's Modified Eagle Medium High Glucose, por sus siglas en inglés) (Gibco No.10567-014) suplementado con 10% de suero fetal bovino a 37°C y en una atmósfera de 5% de CO_2 .

Tabla 7. Cepas y plásmidos utilizados en este estudio.

Cepas	Descripción	Referencia o fuente
E9034A	Cepa silvestre ETEC E9034A, serotipo O8:H9, <i>st⁺/lt⁺</i>	(Levine et al., 1984)
DH5 α	<i>E. coli</i> K-12 <i>supE44lacU169 F80 lacZ ΔM15 hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i>	Colección del laboratorio
TOP10	One Shot® TOP10 <i>E. coli</i> , F- <i>mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ</i>	Colección del laboratorio
BL21DE3	<i>fhuA2 [lon], ompT gal (λ DE3), [dcm], ΔhsdS, λDE3=λ, sBamHI, ΔEcoRI-B, int::(<i>lacI::PlacUV5::T7 gene1</i>), i21 Δnin5</i>	Colección del laboratorio
EAEC 042	<i>E. coli</i> enteroagregativa 042, serotipo O44:H18	Colección del laboratorio
E9034A Δ lngA::Km	E9034A con deleción del gen <i>lngA</i> e inserción del gen de resistencia a kanamicina	(Cruz-Cordova et al., 2014)
E9034A Δ lngB::Cm	E9034A con deleción del gen <i>lngB</i> e inserción del gen de resistencia a cloranfenicol.	Este estudio
E9034A Δ lngC::Cm	E9034A con deleción del gen <i>lngC</i> e inserción del gen de resistencia a cloranfenicol.	Este estudio
E9034A Δ lngD::Cm	E9034A con deleción del gen <i>lngD</i> e inserción del gen de resistencia a cloranfenicol.	Este estudio
E9034A Δ lngH::Cm	E9034A con deleción del gen <i>lngH</i> e inserción del gen de resistencia a cloranfenicol.	Este estudio
E9034A Δ lngP::Cm	E9034A con deleción del gen <i>lngP</i> e inserción del gen de resistencia a cloranfenicol.	Este estudio
E9034A Δ lngR::Km	E9034A con deleción del gen <i>lngR</i> e inserción del gen de resistencia kanamicina.	Este estudio
E9034A Δ lngR	E9034A con deleción del gen <i>lngR</i>	Este estudio
E9034A Δ lngS::Km	E9034A con deleción del gen <i>lngS</i> e inserción del gen de resistencia kanamicina.	Este estudio
E9034A Δ lngS	E9034A con deleción del gen <i>lngS</i>	Este estudio
E9034A Δ hms::Km	E9034A con deleción del gen <i>hms</i> e inserción del gen de resistencia kanamicina.	Este estudio
Plásmidos		
pKD46	Plásmido del sistema recombinación λ Red	(Datsenko y Wanner, 2000)
pKD3	Plásmido con el casete de resistencia a cloranfenicol	(Datsenko y Wanner, 2000)
pKD4	Plásmido con el casete de resistencia a kanamicina	(Datsenko y Wanner, 2000)
pUC18	Vector de alto número de copias	Colección del laboratorio
pUC19	Vector de alto número de copias	Colección del laboratorio
pBAD-Topo	Vector de bajo número de copias	Invitrogen
pJET1.2/blunt	Vector de bajo número de copias	ThermoFisher
pLATE31	Vector de expresión bacteriano (Amp ^r) His-tag en C-terminal	ThermoFisher
pUClngA	Gen <i>lngA</i> en pUC18 (Amp ^r)	Este estudio
pUClngB	Gen <i>lngB</i> en pUC18 (Amp ^r)	Este estudio
pUClngC	El gen <i>lngC</i> en pUC19 (Amp ^r)	Este estudio
pUClngD	El gen <i>lngD</i> en pUC19 (Amp ^r)	Este estudio
pBADlngH	El gen <i>lngH</i> en pBAD-Topo (Amp ^r)	Este estudio
pUClngP	El gen <i>lngP</i> en pUC19 (Amp ^r)	Este estudio
pUClngR	El gen <i>lngR</i> en pUC19 (Amp ^r)	Este estudio
pBADlngS	El gen <i>lngS</i> en pBAD-Topo (Amp ^r)	Este estudio
pLATE31- <i>lngC</i>	El gen <i>lngC</i> en pLATE31	Este estudio
pLATE31- <i>lngB</i>	El gen <i>lngB</i> en pLATE31	Este estudio

Km^r, resistencia a kanamicina; Amp^r, resistencia a ampicilina; Cm^r, resistencia a cloranfenicol.

Tabla 8. Lista de oligonucleótidos usados en este estudio.

Nombre	Secuencia 5'→3'	Sitio de restricción	Uso
LngR-FM	ATATTGCTCTGAAGTTATATGTGTCAAATAATGC	-	Mutagénesis
LngR-RM	AAATGTGTAGGCTGGAGCTGCTTC	-	
LngR-FS	CATAATATTATAAACAGTTCATTGCCTGTCTTGT	HindIII BamHI	Detección y clonación del gen <i>lngR</i> en el vector pUC19
LngR-RS	AATAATGATATGAATATCCTCCTTAG CTGAAGCTTAGTATTCATCCACACTGTAA GTTGGATCCGGTACATAATATTATAAAC		
LngS-FM	AATGAAACTGGAAACAAAATGAGTGGTTATTCGT	-	Mutagénesis
LngS-RM	TAAAAGTGTAGGCTGGAGCTGCTTC	-	
LngS-FS	TACGCATTATGCCTTCGTTGTTCACTGATTCACTG	HindIII BamHI	Detección y clonación del gen <i>lngS</i> en el vector pBAD-Topo
LngS-RS	ATTCATAAAAATATGAATATCCTCCTTAG GAAAAGCTTGC GGACATAAGCCT TGCGGATCCGTTTCTGCAGAC		
HNS-FM	CAATATAAGTTTGAGATTACTACAATGAGCGAAG	-	Mutagénesis
HNS-RM	CACTTAAAATTGTGTAGGCTGGAGCTGCTTC	-	
HNS-FS	GCAATCTACAAAAGATTATTGCTTGATCAGGAAA	HindIII BamHI	Detección y clonación del gen <i>hns</i> en el vector pUC19
HNS-RS	TCGTCGAGATATGAATATCCTCCTTAG TGAAAGCTTGTA ACTATTCACAATCTTTAACC TGGATAACACCGATACGGATCCTAC		
LngA-F	GGGGGGGATCCTAACGGCTACCTAAAGTAA	HindIII	Confirmación de la mutación del gen <i>lngA</i>
LngA-R	TGAAGCTTTTAAACGGCTACCTAAAGTAATTG		
LngA-FP	TGCGGATCCGTGATCTGAAGAAAAAGTAA	BamHI	Clonación de <i>lngA</i> con promotor en pJET1.2/blunt
LngA-RP	TGTGAGAAGGTACTAGCCTATCATATT	-	
LngA-FC	TGCGAATTCGTGATCTGAAGAAAAAGTAAAAG	EcoRI	Clonación del gen <i>lngA</i> con promotor en el vector pUC18
LngA-RC	TGAAGCTTTTAAACGGCTACCTAAAGTAATTG	HindIII	
LngB-FM	CACAGAGACAATTTTATGAAAATGAGAGGCTTCG	-	Mutagénesis
LngB-RM	TGTAGGCTGGAGCTGCTTC	-	
LngB-FS	CATTTTATCCTCTCCATTACATTAGGTTTGTGGTTC	Sacl HindIII	Detección y clonación del gen <i>lngB</i> en el vector pUC18
LngB-RS	TGACTATATGAATATCCTCCTTAG GTAGAGCTCAGGTAAACTCAATTACTTTAGGTAG CCGT TGCAAGCTTCAGGAGATGCCAGTGC ACTCG		
FlnGBALL	AGAAGGAGATATAACTATGAAAATGAGAGGCTTC	-	Clonación del gen <i>lngB</i> en el vector pLATE31
RlnGBALL	ACACTTCTGGAGATGATTATCACTCTCGC GTGGTGGTGATGGTGATGGCCGGTTTGTGGTCTG TACTGCACCAGGTTGTGATA	-	
LngC-FM	AATGGAGAGGATAAAAATGAGAGCAAAAATGGGTG	-	Mutagénesis
LngC-RM	TAGGCTGGAGCTGCTTC	-	
LngC-FS	ACTTCATATTCTTATACCTGACATTCATTGAGAGT	NcoI y BamHI EcoRI	Detección y clonación del gen <i>lngC</i> en el vector
LngC-RS	CGATAACCAGTACATATGAATATCCTCCTTAG CAACCATGGGGATCCCAGAACCACAAACCTAATG TGATGAGAATTCCTGCAATGCCAATAAAACTAAC ATC		

FlnGCALL	AGAAGGAGATATAACTATGAGAGCAAAATGGGG TGTTTTTCTTCCTTAGTATCTTCTTCGAG	-	pUC19 Clonación del gen <i>lngC</i> en el vector pLATE31
RlnGCALL	GTGGTGGTGATGGTGATGGCCTTGAGAGTCGATA ACCAGTACGTTGTTTC	-	
LngD-FM	ACTGGTTATCGACTCTCAATGAATGTCAGGTATA AGAATGTGTAGGCTGGAGCTGCTTC	-	Mutagénesis
LngD-RM	CGATAATAGTCAATACCTACAACCATTGCGATGG TATTATCAATATGAATATCCTCCTTAG	-	
LngD-FS LngD-RS	GTCCCATGGGGATCCGTTTTCTTCAGAACAATAT CCATAAGAGCTCCAGCGCAATTTTTTCATC	NcoI y BamHI SacI	Detección y clonación del gen <i>lngD</i> en el vector pUC19
LngH-FM	CTTGACATAAAGTTAAAATGACGGGCCTTGACAG TCGTTGTGTAGGCTGGAGCTGCTTC	-	Mutagénesis
LngH-RM	CAAACTGATAGAGATAAATACGTTGTTTTTATT GAATTTTTCAAGTAATACCTCATATGAATATCCT CCTTAG	-	
LngH-FS LngH-RS	AGAGAATCCCGGGAAAGTACAGGCTG GAGTCATAGATCGGTAATCCTGAAAGCTTCAT	EcoRI HindIII	Detección y clonación del gen <i>lngH</i> en el vector pBAD-Topo
LngP-FM	ATGTATGTTGAAATCGGCGTTTTCTTTTTTTTATTC ATTACAGTGTAGGCTGGAGCTGCTTC	-	Mutagénesis
LngP-RM	ATCTCTATGCATTTCTCATAGAAATAAGAGAAAG TATCATAAACATATGAATATCCTCCTTAG	-	
LngP-FS LngP-RS	GTTGGATCCAGATTTAGTGGGCCT GATGTCGACGAGCTCTATAGATATTAATCTCTAT	BamHI SacI y Sall	Detección y clonación del gen <i>lngP</i> en el vector pUC19
Puc/M13F Puc/M13R	CGCCAGGGTTTTCCAGTCACGAC CAATTTACACAGGAAACAGCTATGACC		Analizar orientación de insertos en los vectores pUC18 y 19
pFlngA pRlnG A pFlngR1	TGCGAATTCGTGATCTGAAGAAAAAGTAAAAG GTTGGATCCGGATAGCATAAGTATTCCTCC CTTGAATTCGCTGTTGTTTATTATCCTG	EcoRI BamHI EcoRI	Clonar promotor de <i>lngA</i> en pMLB1034 Clonar promotor de <i>lngR</i> (656 pb) en pMLB1034
pFlngR2	TCAGAATTCCTTAATCCTCCATATTGTTGTCC	EcoRI	Clonar promotor de <i>lngR</i> (398 pb) en pMLB1034
pFlngR3	GCAGAATTCAGCACTATGGCGTGATTG	EcoRI	Clonar promotor de <i>lngR</i> (205 pb) en pMLB1034
pRlnG R	GCAGGATCCTGACACATATAACTTCAGAGC	BamHI	Reverso usado para clonar promotor de <i>lngR</i>
pFlngS pRlnG S	AAGCCCGGGCAATGAACTGTTTATAATATTATG CGAGGATCCACTCATTTTGTTCAGTTTCA	SmaI BamHI	Clonar promotor de <i>lngS</i> en pMLB1034

Subrayado se muestran los sitios de restricción enzimática

12.0 BIBLIOGRAFIA

1. Allocati, N., Masulli, M., Alexeyev, M.F., and Di Ilio, C. (2013) *Escherichia coli* in Europe: an overview. *Int J Environ Res Public Health* **10**: 6235-6254.
2. Anantha, R.P., Stone, K.D., and Donnenberg, M.S. (2000) Effects of *bfp* mutations on biogenesis of functional enteropathogenic *Escherichia coli* type IV pili. *J Bacteriol* **182**: 2498-2506.
3. Ayers, M., Howell, P.L., and Burrows, L.L. (2010) Architecture of the type II secretion and type IV pilus machineries. *Future Microbiol* **5**: 1203-1218.
4. Blank, T.E., and Donnenberg, M.S. (2001) Novel topology of BfpE, a cytoplasmic membrane protein required for type IV fimbrial biogenesis in enteropathogenic *Escherichia coli*. *J Bacteriol* **183**: 4435-4450.
5. Boden, M., and Hawkins, J. (2005) Prediction of subcellular localization using sequence-biased recurrent networks. *Bioinformatics* **21**: 2279-2286.
6. Bose, N., and Taylor, R.K. (2005) Identification of a TcpC-TcpQ outer membrane complex involved in the biogenesis of the toxin-coregulated pilus of *Vibrio cholerae*. *J Bacteriol* **187**: 2225-2232.
7. Burrows, L.L. (2012) *Pseudomonas aeruginosa* twitching motility: type IV pili in action. *Annu Rev Microbiol* **66**: 493-520.
8. Clavijo, A.P., Bai, J., and Gomez-Duarte, O.G. (2010) The Longus type IV pilus of enterotoxigenic *Escherichia coli* (ETEC) mediates bacterial self-aggregation and protection from antimicrobial agents. *Microb Pathog* **48**: 230-238.
9. Crossman, L.C., Chaudhuri, R.R., Beatson, S.A., Wells, T.J., Desvaux, M., Cunningham, A.F. et al. (2010) A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. *J Bacteriol* **192**: 5822-5831.
10. Crowther, L.J., Anantha, R.P., and Donnenberg, M.S. (2004) The inner membrane subassembly of the enteropathogenic *Escherichia coli* bundle-forming pilus machine. *Mol Microbiol* **52**: 67-79.
11. Crowther, L.J., Yamagata, A., Craig, L., Tainer, J.A., and Donnenberg, M.S. (2005) The ATPase activity of BfpD is greatly enhanced by zinc and allosteric interactions with other Bfp proteins. *J Biol Chem* **280**: 24839-24848.
12. Cruz-Cordova, A., Espinosa-Mazariego, K., Ochoa, S.A., Saldana, Z., Rodea, G.E., Cazares-Dominguez, V. et al. (2014) CS21 positive multidrug-resistant ETEC clinical isolates from children with diarrhea are associated with self-aggregation, and adherence. *Front Microbiol* **5**: 709.
13. d'Aubenton Carafa, Y., Brody, E., and Thermes, C. (1990) Prediction of rho-independent *Escherichia coli* transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol* **216**: 835-858.

14. Daniel, A., Singh, A., Crowther, L.J., Fernandes, P.J., Schreiber, W., and Donnenberg, M.S. (2006) Interaction and localization studies of enteropathogenic *Escherichia coli* type IV bundle-forming pilus outer membrane components. *Microbiology* **152**: 2405-2420.
15. Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640-6645.
16. Feeney, A.R., Cooke, E.M., and Shinebaum, R. (1980) A comparative study of gram-negative aerobic bacilli in the faeces of babies born in hospital and at home. *J Hyg (Lond)* **84**: 91-96.
17. Forsman, K., Goransson, M., and Uhlin, B.E. (1989) Autoregulation and multiple DNA interactions by a transcriptional regulatory protein in *E. coli* pili biogenesis. *EMBO J* **8**: 1271-1277.
18. Francetic, O., and Pugsley, A.P. (1996) The cryptic general secretory pathway (*gsp*) operon of *Escherichia coli* K-12 encodes functional proteins. *J Bacteriol* **178**: 3544-3549.
19. Francetic, O., Lory, S., and Pugsley, A.P. (1998) A second prepilin peptidase gene in *Escherichia coli* K-12. *Mol Microbiol* **27**: 763-775.
20. Gaastra, W., and de Graaf, F.K. (1982) Host-specific fimbrial adhesins of noninvasive enterotoxigenic *Escherichia coli* strains. *Microbiol Rev* **46**: 129-161.
21. Gaastra, W., and Svennerholm, A.M. (1996) Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends Microbiol* **4**: 444-452.
22. Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T., and Kado, C.I. (1985) Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *J Bacteriol* **164**: 918-921.
23. Giron, J.A., Levine, M.M., and Kaper, J.B. (1994) Longus: a long pilus ultrastructure produced by human enterotoxigenic *Escherichia coli*. *Mol Microbiol* **12**: 71-82.
24. Giron, J.A., Viboud, G.I., Sperandio, V., Gomez-Duarte, O.G., Maneval, D.R., Albert, M.J. et al. (1995) Prevalence and association of the longus pilus structural gene (*lngA*) with colonization factor antigens, enterotoxin types, and serotypes of enterotoxigenic *Escherichia coli*. *Infect Immun* **63**: 4195-4198.
25. Girón, J.A., Ho, A.S., and Schoolnik, G.K. (1991) An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. *Science* **254**: 710-713.
26. Girón, J.A., Gomez-Duarte, O.G., Jarvis, K.G., and Kaper, J.B. (1997) Longus pilus of enterotoxigenic *Escherichia coli* and its relatedness to other type-4 pili--a minireview. *Gene* **192**: 39-43.
27. Gomez-Duarte, O.G., Chattopadhyay, S., Weissman, S.J., Giron, J.A., Kaper, J.B., and Sokurenko, E.V. (2007) Genetic diversity of the gene cluster encoding longus, a type IV pilus of enterotoxigenic *Escherichia coli*. *J Bacteriol* **189**: 9145-9149.

28. Goransson, M., Sonden, B., Nilsson, P., Dagberg, B., Forsman, K., Emanuelsson, K., and Uhlin, B.E. (1990) Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli*. *Nature* **344**: 682-685.
29. Guevara, C.P., Luiz, W.B., Sierra, A., Cruz, C., Qadri, F., Kaushik, R.S. et al. (2013) Enterotoxigenic *Escherichia coli* CS21 pilus contributes to adhesion to intestinal cells and to pathogenesis under in vivo conditions. *Microbiology*.
30. Gutierrez-Cazarez, Z., Qadri, F., Albert, M.J., and Giron, J.A. (2000) Identification of enterotoxigenic *Escherichia coli* harboring longus type IV pilus gene by DNA amplification. *J Clin Microbiol* **38**: 1767-1771.
31. Hewitt, J.H., and Rigby, J. (1976) Effect of various milk feeds on numbers of *Escherichia coli* and *Bidifobacterium* in the stools of new-born infants. *J Hyg (Lond)* **77**: 129-139.
32. Hoang, T.T., Karkhoff-Schweizer, R.R., Kutchma, A.J., and Schweizer, H.P. (1998) A broad-host-range F₁-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene* **212**: 77-86.
33. Hobbs, M., and Mattick, J.S. (1993) Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. *Mol Microbiol* **10**: 233-243.
34. Hodges, K., and Gill, R. (2010) Infectious diarrhea: Cellular and molecular mechanisms. *Gut Microbes* **1**: 4-21.
35. Hulton, C.S., Seirafi, A., Hinton, J.C., Sidebotham, J.M., Waddell, L., Pavitt, G.D. et al. (1990) Histone-like protein H1 (H-NS), DNA supercoiling, and gene expression in bacteria. *Cell* **63**: 631-642.
36. Iredell, J.R., and Manning, P.A. (1997) Translocation failure in a type-4 pilin operon: *rfb* and *tcpT* mutants in *Vibrio cholerae*. *Gene* **192**: 71-77.
37. Isidean, S.D., Riddle, M.S., Savarino, S.J., and Porter, C.K. (2011) A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression. *Vaccine* **29**: 6167-6178.
38. Iyer, L.M., Leipe, D.D., Koonin, E.V., and Aravind, L. (2004) Evolutionary history and higher order classification of AAA+ ATPases. *J Struct Biol* **146**: 11-31.
39. Kaper, J.B., Nataro, J.P., and Mobley, H.L. (2004) Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2**: 123-140.
40. Kaufman, M.R., Seyer, J.M., and Taylor, R.K. (1991) Processing of TCP pilin by TcpJ typifies a common step intrinsic to a newly recognized pathway of extracellular protein secretion by gram-negative bacteria. *Genes Dev* **5**: 1834-1846.
41. Keusch, G.T. (2001) Toxin-Associated Gastrointestinal Disease: A Clinical Overview. In *Molecular Medical Microbiology*. Sussman, M. (ed). New York: Academic Press., pp. 1083-1088.

42. Keusch, G.T., Fontaine, O., Bhargava, A., and al., e. (2006) Diarrheal Diseases. In *Disease Control Priorities in Developing Countries* D.T., J., J.G., B., A.R., M., and et al. (eds). Washington (DC).The International Bank for Reconstruction and Development / The World Bank.
43. Kirn, T.J., Bose, N., and Taylor, R.K. (2003) Secretion of a soluble colonization factor by the TCP type 4 pilus biogenesis pathway in *Vibrio cholerae*. *Mol Microbiol* **49**: 81-92.
44. Kirn, T.J., Lafferty, M.J., Sandoe, C.M., and Taylor, R.K. (2000) Delineation of pilin domains required for bacterial association into microcolonies and intestinal colonization by *Vibrio cholerae*. *Mol Microbiol* **35**: 896-910.
45. Koebnik, R., Locher, K.P., and Van Gelder, P. (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol* **37**: 239-253.
46. Kolappan, S., Ng, D., Yang, G., Harn, T., and Craig, L. (2015) Crystal Structure of the Minor Pilin CofB, the Initiator of CFA/III Pilus Assembly in Enterotoxigenic *Escherichia coli*. *J Biol Chem* **290**: 25805-25818.
47. LaPointe, C.F., and Taylor, R.K. (2000) The type 4 prepilin peptidases comprise a novel family of aspartic acid proteases. *J Biol Chem* **275**: 1502-1510.
48. Levine, M.M. (1987) *Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. *J Infect Dis* **155**: 377-389.
49. Levine, M.M., Ristaino, P., Marley, G., Smyth, C., Knutton, S., Boedeker, E. et al. (1984) Coli surface antigens 1 and 3 of colonization factor antigen II-positive enterotoxigenic *Escherichia coli*: morphology, purification, and immune responses in humans. *Infect Immun* **44**: 409-420.
50. Lewin, B., and Lopez, A.A. (1996) Genes. Volumen 1. In *Genes*: Reverte, pp. 377-390.
51. Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V. et al. (2012) Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**: 2095-2128.
52. Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V. et al. (2013) Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**: 2095-2128.
53. Matsuda, S., Vert, J.P., Saigo, H., Ueda, N., Toh, H., and Akutsu, T. (2005) A novel representation of protein sequences for prediction of subcellular location using support vector machines. *Protein Sci* **14**: 2804-2813.
54. Mattick, J.S. (2002) Type IV pili and twitching motility. *Annu Rev Microbiol* **56**: 289-314.

55. Mazariego-Espinosa, K., Cruz, A., Ledesma, M.A., Ochoa, S.A., and Xicohtencatl-Cortes, J. (2010) Longus, a type IV pilus of enterotoxigenic *Escherichia coli*, is involved in adherence to intestinal epithelial cells. *J Bacteriol* **192**: 2791-2800.
56. Miles, A.A., Misra, S.S., and Irwin, J.O. (1938) The estimation of the bactericidal power of the blood. *J Hyg (Lond)* **38**: 732-749.
57. Nataro, J.P., and Kaper, J.B. (1998) Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**: 142-201.
58. Navarre, W.W., McClelland, M., Libby, S.J., and Fang, F.C. (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev* **21**: 1456-1471.
59. Niehaus, M.D., Moore, S.R., Patrick, P.D., Derr, L.L., Lorntz, B., Lima, A.A., and Guerrant, R.L. (2002) Early childhood diarrhea is associated with diminished cognitive function 4 to 7 years later in children in a northeast Brazilian shantytown. *Am J Trop Med Hyg* **66**: 590-593.
60. Pichel, M.G., Binsztein, N., Qadri, F., and Giron, J.A. (2002) Type IV longus pilus of enterotoxigenic *Escherichia coli*: occurrence and association with toxin types and colonization factors among strains isolated in Argentina. *J Clin Microbiol* **40**: 694-697.
61. Pizarro-Cerda, J., and Cossart, P. (2006) Bacterial adhesion and entry into host cells. *Cell* **124**: 715-727.
62. Porter, M.E., Mitchell, P., Roe, A.J., Free, A., Smith, D.G., and Gally, D.L. (2004) Direct and indirect transcriptional activation of virulence genes by an AraC-like protein, PerA from enteropathogenic *Escherichia coli*. *Mol Microbiol* **54**: 1117-1133.
63. Pugsley, A.P., and Francetic, O. (1998) Protein secretion in *Escherichia coli* K-12: dead or alive? *Cell Mol Life Sci* **54**: 347-352.
64. Qadri, F., Giron, J.A., Helander, A., Begum, Y.A., Asaduzzaman, M., Xicohtencatl-Cortes, J. et al. (2000) Human antibody response to longus type IV pilus and study of its prevalence among enterotoxigenic *Escherichia coli* in Bangladesh by using monoclonal antibodies. *J Infect Dis* **181**: 2071-2074.
65. Reinhardt, A., and Hubbard, T. (1998) Using neural networks for prediction of the subcellular location of proteins. *Nucleic Acids Res* **26**: 2230-2236.
66. Renzoni, D., Esposito, D., Pfuhl, M., Hinton, J.C., Higgins, C.F., Driscoll, P.C., and Ladbury, J.E. (2001) Structural characterization of the N-terminal oligomerization domain of the bacterial chromatin-structuring protein, H-NS. *J Mol Biol* **306**: 1127-1137.
67. Roux, N., Spagnolo, J., and de Bentzmann, S. (2012) Neglected but amazingly diverse type IVb pili. *Res Microbiol* **163**: 659-673.
68. Sauvonnnet, N., Gounon, P., and Pugsley, A.P. (2000) PpdD type IV pilin of *Escherichia coli* K-12 can be assembled into pili in *Pseudomonas aeruginosa*. *J Bacteriol* **182**: 848-854.

69. Schmidt, S.A., Bieber, D., Ramer, S.W., Hwang, J., Wu, C.Y., and Schoolnik, G. (2001) Structure-function analysis of BfpB, a secretin-like protein encoded by the bundle-forming-pilus operon of enteropathogenic *Escherichia coli*. *J Bacteriol* **183**: 4848-4859.
70. Shindo, H., Iwaki, T., Ieda, R., Kurumizaka, H., Ueguchi, C., Mizuno, T. et al. (1995) Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from *Escherichia coli*. *FEBS Lett* **360**: 125-131.
71. Solovyev, V.V., Shahmuradov, I.A., and Salamov, A.A. (2010) Identification of promoter regions and regulatory sites. *Methods Mol Biol* **674**: 57-83.
72. Sonnhammer, E.L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**: 175-182.
73. Taniguchi, T., Fujino, Y., Yamamoto, K., Miwatani, T., and Honda, T. (1995) Sequencing of the gene encoding the major pilin of pilus colonization factor antigen III (CFA/III) of human enterotoxigenic *Escherichia coli* and evidence that CFA/III is related to type IV pili. *Infect Immun* **63**: 724-728.
74. Taniguchi, T., Yasuda, Y., Tochikubo, K., Yamamoto, K., and Honda, T. (1999) The gene encoding the prepilin peptidase involved in biosynthesis of pilus colonization factor antigen III (CFA/III) of human enterotoxigenic *Escherichia coli*. *Microbiol Immunol* **43**: 853-861.
75. Taniguchi, T., Arita, M., Sato, M., Yamamoto, K., Miwatani, T., and Honda, T. (1994) Evidence that the N-terminal amino acid sequence of pilus colonization factor antigen III produced by human enterotoxigenic *Escherichia coli* is similar to that of TcpA pilin of *Vibrio cholerae*. *J Infect Dis* **170**: 1049-1050.
76. Taniguchi, T., Akeda, Y., Haba, A., Yasuda, Y., Yamamoto, K., Honda, T., and Tochikubo, K. (2001) Gene cluster for assembly of pilus colonization factor antigen III of enterotoxigenic *Escherichia coli*. *Infect Immun* **69**: 5864-5873.
77. Taylor, R.K., Miller, V.L., Furlong, D.B., and Mekalanos, J.J. (1987) Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc Natl Acad Sci U S A* **84**: 2833-2837.
78. Tobe, T., Schoolnik, G.K., Sohel, I., Bustamante, V.H., and Puente, J.L. (1996) Cloning and characterization of *bfpTVW*, genes required for the transcriptional activation of *bfpA* in enteropathogenic *Escherichia coli*. *Mol Microbiol* **21**: 963-975.
79. Tripathi, S.A., and Taylor, R.K. (2007) Membrane association and multimerization of TcpT, the cognate ATPase ortholog of the *Vibrio cholerae* toxin-coregulated-pilus biogenesis apparatus. *J Bacteriol* **189**: 4401-4409.
80. Tupper, A.E., Owen-Hughes, T.A., Ussery, D.W., Santos, D.S., Ferguson, D.J., Sidebotham, J.M. et al. (1994) The chromatin-associated protein H-NS alters DNA topology in vitro. *EMBO J* **13**: 258-268.

81. Vos, T., Barber, R.M., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I. et al. (2015) Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013;2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* **386**: 743-800.
82. Whitchurch, C.B., and Mattick, J.S. (1994) *Escherichia coli* contains a set of genes homologous to those involved in protein secretion, DNA uptake and the assembly of type-4 fimbriae in other bacteria. *Gene* **150**: 9-15.
83. WHO (2013). Diarrhoeal disease [WWW document]. URL <http://www.who.int/mediacentre/factsheets/fs330/en/>.
84. Wolfgang, M., van Putten, J.P., Hayes, S.F., Dorward, D., and Koomey, M. (2000) Components and dynamics of fiber formation define a ubiquitous biogenesis pathway for bacterial pili. *EMBO J* **19**: 6408-6418.
85. Xia, Y., and Uhlin, B.E. (1999) Mutational analysis of the PapB transcriptional regulator in *Escherichia coli*. Regions important for DNA binding and oligomerization. *J Biol Chem* **274**: 19723-19730.
86. Xicohtencatl-Cortes, J., Monteiro-Neto, V., Ledesma, M.A., Jordan, D.M., Francetic, O., Kaper, J.B. et al. (2007) Intestinal adherence associated with type IV pili of enterohemorrhagic *Escherichia coli* O157:H7. *J Clin Invest* **117**: 3519-3529.
87. Yu, C.S., Lin, C.J., and Hwang, J.K. (2004) Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci* **13**: 1402-1406.
88. Yu, R.R., and DiRita, V.J. (2002) Regulation of gene expression in *Vibrio cholerae* by ToxT involves both antirepression and RNA polymerase stimulation. *Mol Microbiol* **43**: 119-134.
89. Yuen, A.S., Kolappan, S., Ng, D., and Craig, L. (2013) Structure and secretion of CofJ, a putative colonization factor of enterotoxigenic *Escherichia coli*. *Mol Microbiol* **90**: 898-918.
90. Zhang, C., Iqbal, J., and Gomez-Duarte, O.G. (2016) Murine immunization with CS21 pili or LngA major subunit of enterotoxigenic *Escherichia coli* (ETEC) elicits systemic and mucosal immune responses and inhibits ETEC gut colonization. *Vet Microbiol*.

14. PUBLICACIONES

1. Effects of *lng* mutations on LngA expression, processing and CS21 assembly in enterotoxigenic *Escherichia coli* E9034A. **Saldaña-Ahuactzi Zeus**, Rodea Gerardo E., Cruz-Córdova Ariadna, Rodríguez-Ramírez Viridiana, Espinosa-Mazariego Karina, González-Montalvo Martín A., Ochoa Sara A., González-Pedrajo Bertha, Eslava-Campos Carlos A., López-Villegas Edgar O., Hernández-Castro Rigoberto, Arellano-Galindo José, Patiño-López Genaro, Xicohtencatl-Cortes Juan. J. Front Microbiol. 2016 Jul 7: PMID: 27536289

Coautoría en publicaciones:

1. CS21 positive multidrug-resistant ETEC clinical isolates from children with diarrhea are associated with self-aggregation, and adherence. Cruz-Córdova A, Espinosa-Mazariego K, Ochoa SA, **Saldaña Z**, Rodea GE, Cázares-Domínguez V, Rodríguez-Ramírez V, Eslava-Campos CA, Navarro-Ocaña A, Arellano-Galindo J, Hernández-Castro R, Gómez-Duarte OG, Qadri F, Xicohtencatl-Cortes J. Front Microbiol. 2014 Dec 17;5:709. PMID: 25646093
2. Dimeric and Trimeric Fusion Proteins Generated with Fimbrial Adhesins of Uropathogenic *Escherichia coli*. Luna-Pineda VM, Reyes-Grajeda JP, Cruz-Córdova A, **Saldaña-Ahuactzi Z**, Ochoa SA, Maldonado-Bernal C, Cázares-Domínguez V, Moreno-Fierros L, Arellano-Galindo J, Hernández-Castro R, Xicohtencatl-Cortes J. Front Cell Infect Microbiol. 2016 Oct 31;6:135. PMID: 27843814



Effects of *Ing* Mutations on LngA Expression, Processing, and CS21 Assembly in Enterotoxigenic *Escherichia coli* E9034A

Zeus Saldaña-Ahuactzi^{1,2}, Gerardo E. Rodea^{1,2}, Ariadna Cruz-Córdova¹, Viridiana Rodríguez-Ramírez¹, Karina Espinosa-Mazariego¹, Martín A. González-Montalvo¹, Sara A. Ochoa¹, Bertha González-Pedrajo³, Carlos A. Eslava-Campos⁴, Edgar O. López-Villegas⁵, Rigoberto Hernández-Castro⁶, José Arellano-Galindo⁷, Genaro Patiño-López⁸ and Juan Xicohtencatl-Cortes^{1*}

OPEN ACCESS

Edited by:

Jorge Blanco,
University of Santiago de Compostela,
Spain

Reviewed by:

Lisa Craig,
Simon Fraser University, Canada
Felipe Del Canto,
University of Chile, Chile

*Correspondence:

Juan Xicohtencatl-Cortes
juanxico@yahoo.com

Specialty section:

This article was submitted to
Infectious Diseases,
a section of the journal
Frontiers in Microbiology

Received: 08 March 2016

Accepted: 19 July 2016

Published: 03 August 2016

Citation:

Saldaña-Ahuactzi Z, Rodea GE, Cruz-Córdova A, Rodríguez-Ramírez V, Espinosa-Mazariego K, González-Montalvo MA, Ochoa SA, González-Pedrajo B, Eslava-Campos CA, López-Villegas EO, Hernández-Castro R, Arellano-Galindo J, Patiño-López G and Xicohtencatl-Cortes J (2016) Effects of *Ing* Mutations on LngA Expression, Processing, and CS21 Assembly in Enterotoxigenic *Escherichia coli* E9034A. *Front. Microbiol.* 7:1201. doi: 10.3389/fmicb.2016.01201

¹ Laboratorio de Investigación en Bacteriología Intestinal, Hospital Infantil de México Federico Gómez, Ciudad de México, Mexico, ² Instituto de Fisiología Celular at the Universidad Nacional Autónoma de México, Ciudad de México, Mexico, ³ Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Ciudad de México, Mexico, ⁴ Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional Autónoma de México, Ciudad de México, Mexico, ⁵ Laboratorio Central de Microscopía, Departamento de Investigación-SEPI, Instituto Politécnico Nacional, Ciudad de México, Mexico, ⁶ Departamento de Ecología de Agentes Patógenos, Hospital General "Dr. Manuel Gea González," Ciudad de México, Mexico, ⁷ Departamento de Infectología, Hospital Infantil de México Federico Gómez, Ciudad de México, Mexico, ⁸ Laboratorio de Investigación en Inmunología y Proteómica, Hospital Infantil de México Federico Gómez, Ciudad de México, Mexico

Enterotoxigenic *Escherichia coli* (ETEC) is a major cause of morbidity in children under 5 years of age in low- and middle-income countries and a leading cause of traveler's diarrhea worldwide. The ability of ETEC to colonize the intestinal epithelium is mediated by fimbrial adhesins, such as CS21 (Longus). This adhesin is a type IVb pilus involved in adherence to intestinal cells *in vitro* and bacterial self-aggregation. Fourteen open reading frames have been proposed to be involved in CS21 assembly, hitherto only the *IngA* and *IngB* genes, coding for the major (LngA) and minor (LngB) structural subunit, have been characterized. In this study, we investigated the role of the LngA, LngB, LngC, LngD, LngH, and LngP proteins in the assembly of CS21 in ETEC strain E9034A. The deletion of the *IngA*, *IngB*, *IngC*, *IngD*, *IngH*, or *IngP* genes, abolished CS21 assembly in ETEC strain E9034A and the adherence to HT-29 cells was reduced 90%, compared to wild-type strain. Subcellular localization prediction of CS21 proteins was similar to other well-known type IV pili homologs. We showed that LngP is the prepilin peptidase of LngA, and that ETEC strain E9034A has another peptidase capable of processing LngA, although with less efficiency. Additionally, we present immuno-electron microscopy images to show that the LngB protein could be localized at the tip of CS21. In conclusion, our results demonstrate that the LngA, LngB, LngC, LngD, LngH, and LngP proteins are essential for CS21 assembly, as well as for bacterial aggregation and adherence to HT-29 cells.

Keywords: ETEC, CS21, biogenesis, type IV pilus, adherence to intestinal cells, pilus

INTRODUCTION

Enterotoxigenic *Escherichia coli* (ETEC) remains as a major bacterial pathogen associated to high morbidity, mainly in children less than 5 years of age in low- and middle-income countries. The mortality of ETEC in 2010 was estimated to be of 120,800 deaths (Lozano et al., 2013). ETEC is also a leading cause of diarrheal illness in travelers to endemic countries (Nataro and Kaper, 1998; Kaper et al., 2004). ETEC possess two enterotoxins: thermo-labile (LT) and thermo-stable (ST), which are responsible for the secretory diarrhea (Nataro and Kaper, 1998). The colonization factors (CFs) are indispensable for ETEC virulence, and are responsible for intestinal epithelium adherence and colonization (Gaastra and Svennerholm, 1996). More than 20 CFs have been identified, however, only eight (CFA/I, CS1, CS2, CS3, CS4, CS5, CS6, and CS21) are the most prevalent worldwide (Gaastra and Svennerholm, 1996; Isidean et al., 2011). CS21, also called longus, is a class b type IV pilus which is encoded in a 14 kb *lng* gene cluster located in a large virulence plasmid (50–90 kb) (Girón et al., 1994; Gomez-Duarte et al., 2007). Type IV pilus have been classified into two classes: IVa and IVb. Type IVa pilus assembly requires around 40 genes products encoded in different parts of the genome; whereas type IVb system comprises about a dozen genes (between 11 and 14), encoded within the pilus operon (Strom and Lory, 1993; Roux et al., 2012). Girón et al. (1994) reported for the first time that CS21 structural gene (*lngA*) is encoded in a large plasmid, and that a 5-kb Bam HI restriction fragment from this plasmid was sufficient for pilus assembly in an *Escherichia coli* (*E. coli*) strain K-12 (DH5 α F'IQ). Gomez-Duarte et al. (2007) reported a cluster of 14 genes (*lngR*, *lngS*, *lngT*, *lngA*, *lngB*, *lngC*, *lngD*, *lngE*, *lngF*, *lngG*, *lngH*, *lngI*, *lngJ*, and *lngP*) of 13.93 kb located in tandem that might be required for expression of CS21.

The CS21 filament is composed of a 22-kDa monomer (LngA) which shares considerable N-terminal sequence similarity with CofA of ETEC, TcpA of *Vibrio cholerae* and BfpA of enteropathogenic *E. coli* (EPEC) (Taylor et al., 1987; Girón et al., 1997). The *lngA* gene codes for an immature LngA protein (prepilin), and requires processing by a prepilin peptidase, before of its translocation and oligomerization. The processing includes signal sequence removal and first residue (methionine) methylation, leading to a mature pilin (Hobbs and Mattick, 1993; Strom and Lory, 1993).

The biogenesis of type IV pilus involves oligomerization and secretion of the pilin subunit by a mechanism that requires a set of proteins, which are part of an assembly and export apparatus. The assembly of type IV pili also requires a nucleotide-binding protein that provides energy, a polytopic inner membrane protein, a prepilin peptidase, and accessory proteins (Mattick, 2002). Several proteins involved in pilus biogenesis of the Bundle-Forming Pilus (BFP) of EPEC, Toxin Coreregulated Pilus (TCP) of *V. cholerae*, and Colonization Factor Antigen III (CFA/III) of ETEC, are homologous with proteins encoded in the *lng* cluster (Gomez-Duarte et al., 2007; Roux et al., 2012). For example, LngD has two homologs, BfpB and TcpC, that are multimeric outer membrane lipoproteins through which the pilin subunits BfpA and TcpA are extruded (Ogierman and Manning, 1992;

Ramer et al., 1996). The high similarity between CS21 and CFA/III biosynthetic genes indicates that CS21 and CFA/III are paralogous (Gomez-Duarte et al., 2007).

LngA, LngB, LngJ, and LngP have more than 60% of identity with CofA, CofB, CofJ, and CofP, respectively (Gomez-Duarte et al., 2007). The latter proteins have been previously characterized. CofA is the major pilin subunit of CFA/III (Taniguchi et al., 2001), and CofB the minor pilin subunit, which has been proposed to initiate pilus assembly (Kolappan et al., 2015). CofJ is a soluble protein secreted via the CFA/III system and has been proposed to be a putative adhesin of ETEC (Yuen et al., 2013). CofP has a prepilin peptidase function (Taniguchi et al., 1999).

CS21 induces bacterial self-aggregation, which protects ETEC against antimicrobial agents *in vitro* (Clavijo et al., 2010). The role of CS21 in human colonic cells colonization has been studied *in vitro* by Mazariego-Espinosa et al. (2010); and recently, Guevara et al. (2013) demonstrated the role of CS21 in the pathogenesis of ETEC *in vivo* using a neonatal mice challenge infection model. However, the proteins involved in CS21 assembly have not been fully described. The purpose of this study was to determine the effect of the deletion of six *lng* (*lngA*, *lngB*, *lngC*, *lngD*, *lngH*, and *lngP*) genes in the ETEC strain E9034A on CS21 assembly. These genes were chosen based on its putative function with homologous proteins of other type IV pili. Here we report that deletions of the *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, and *lngP* genes affect CS21 assembly, bacterial self-aggregation, and adherence to HT-29 cells.

MATERIALS AND METHODS

Bacterial Strains and Culture Media

The strains and plasmids used in this study are listed in **Table 1**. All the strains were stored in Luria-Bertani (LB) broth (Difco, NJ, USA) with 20% glycerol at -70°C . Strains were grown on LB agar or in LB broth at 30, 37, or 42°C depending on the assay. For CS21 expression, Pleuropneumonia-Like Organisms (PPLO) broth (BD, Difco, NJ, USA) or terrific broth (TB) (Amresco LLC, OH, USA) media was used. When required, kanamycin (50 $\mu\text{g/ml}$), ampicillin (100 $\mu\text{g/ml}$) and/or chloramphenicol (50 $\mu\text{g/ml}$) antibiotics were added to the media.

Construction of Isogenic Mutants

Non-polar deletion mutants in *lngB*, *lngC*, *lngD*, *lngH*, and *lngP* genes were generated by the λ -Red recombinase method (Datsenko and Wanner, 2000). The primers employed for DNA amplification are listed in **Table 2**. Primers flanking the *lngB*, *lngC*, *lngD*, *lngH*, and *lngP* genes were used to confirm the gene deletion and replacement by PCR (**Table 2**). The mutant strains were complemented by the addition of a wild-type copy of the mutant gene in *trans* on a high-copy-number vector (**Table 1**). Empty plasmids were transformed into each mutant to verify that complementation was due to the presence of the wild-type gene.

Molecular and Genetic Techniques

All oligonucleotides used for PCR were obtained from Integrated DNA Technologies, Inc (IDT, CA, USA) and are listed in

TABLE 1 | Strains and plasmids used in this study.

Strain or plasmid	Description	References or source
E9034A	Wild-type ETEC (O8:H9, SL ⁺ /LT ⁺)	Levine et al., 1984
DH5 α	<i>E. coli</i> K-12 supE44 Δ lacU169 F80 lacZ Δ M15 hsdR17 recA1 endA1 gyrA96 thi-1 relA1	Lab collection
TOP10	One Shot [®] TOP10 <i>E. coli</i> , F- mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80lacZ Δ M15 Δ lacX74 nupG recA1 araD139 Δ (ara-leu) 7697 galE15 galK16 rpsL (Str ^R) endA1 λ -	Lab collection
BL21 (DE3)	<i>E. coli</i> K-12 fhuA2 [lon] ompT gal (λ , DE3) [dcm] Δ hsdS λ . DE3 = λ . sBamHI Δ EcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δ nin5	Lab collection
E9034A Δ IngA::km	E9034A with a non-polar insertional mutation in <i>IngA</i> and kanamycin resistance	Cruz-Córdova et al., 2014
E9034A Δ IngB::cm	E9034A with a non-polar insertional mutation in <i>IngB</i> and chloramphenicol resistance	This study
E9034A Δ IngC::cm	E9034A with a non-polar insertional mutation in <i>IngC</i> and chloramphenicol resistance	This study
E9034A Δ IngD::cm	E9034A with a non-polar insertional mutation in <i>IngD</i> and chloramphenicol resistance	This study
E9034A Δ IngH::cm	E9034A with a non-polar insertional mutation in <i>IngH</i> and chloramphenicol resistance	This study
E9034A Δ IngP::cm	E9034A with a non-polar insertional mutation in <i>IngP</i> and chloramphenicol resistance	This study
E9034A Δ IngR::km	E9034A with a non-polar insertional mutation in <i>IngR</i> and kanamycin resistance	Unpublished data
Plasmids		
pKD46	λ Red recombinase system plasmid	Datsenko and Wanner, 2000
pKD3	Cm ^r cassette template plasmid	Datsenko and Wanner, 2000
pKD4	Km ^r cassette template plasmid	Datsenko and Wanner, 2000
pUC18	High-copy-number cloning vector	Lab collection
pUC19	High-copy-number cloning vector	Lab collection
pBAD-Topo	Low-copy-number expression vector	Invitrogen
pJET1.2/blunt	Low-copy number cloning vector	ThermoFisher
pLATE31	Bacterial expression vectors (Amp ^r)	ThermoFisher Scientific
	His-tag in C-terminal	
pUCingA	<i>IngA</i> in pUC18 (Amp ^r)	This study
pUCingB	<i>IngB</i> in pUC19 (Amp ^r)	This study
pUCingC	<i>IngC</i> in pUC19 (Amp ^r)	This study
pUCingD	<i>IngD</i> in pUC19 (Amp ^r)	This study
pBADIngH	<i>IngH</i> in pBAD-Topo (Amp ^r)	This study
pUCingP	<i>IngP</i> in pUC19 (Amp ^r)	This study
pLATE31-IngC	<i>IngC</i> in pLATE31	This study
pLATE31-IngB	<i>IngB</i> in pLATE31	This study

Km^r, kanamycin resistance; Amp^r, ampicillin resistance; Cm^r, chloramphenicol resistance.

Table 2. PCR, restriction endonuclease digestion, ligation, transformation, plasmid DNA preparations, and DNA electrophoresis were performed using standard techniques (Sambrook and Russell, 2001). PCR amplifications were performed with PCR Master Mix 2x (Promega Corporation, WI, USA) and Platinum[®] Taq DNA Polymerase High Fidelity (Invitrogen, USA). Restriction enzymes were purchased from Promega (Promega Corporation, WI, USA) and were used according to the manufacturer's recommendations.

Cloning of the *IngB* and *IngC* Genes in the pLATE31 Expression Plasmid

Primer design, amplification, cloning and expression of *IngB* and *IngC* genes in the pLATE31 plasmid were conducted according to the aLICator Ligation Independent Cloning and Expression System (Thermo Scientific, CA, USA). pLATE31 and the recombinant expression plasmid pLATE31-*IngB* and pLATE31-*IngC* were transformed into *E. coli* BL21 (DE3). For the expression and detection of *IngB* and *IngC*, *E. coli* was cultured in LB medium at 37°C, and induced with 1 mM IPTG for 5 h. *IngB* and *IngC* His-Tag recombinant proteins were purified from 200 ml of induced cultures media. The cultures were centrifuged at 3500 \times g, and the pellets were resuspended in 20 ml of pH 8.2 lysis buffer (8M Urea, 100 mM NaH₂PO₄, 10 mM Tris-HCl, and 10 mM imidazole) and sonicated for 30 min at 40 kHz, using 10 cycles of 1 min pulse and 2 min of cooldown. The sonicated samples were centrifuged at 3500 \times g, and the His-tag recombinant proteins from the clear supernatants were captured in an affinity chromatography column of Ni-NTA Agarose (Qiagen, Hilden, Germany), washed with 60 ml of lysis buffer (pH 6.5), and eluted with 10 ml of lysis buffer (pH 4.0). The eluted fractions containing the recombinant proteins were refolded by dialysis and maintained in phosphate-buffered saline (PBS) pH 7.0. Anti His-Tag monoclonal antibodies were used to detect the expression of recombinant proteins by immunoblotting.

Antibodies Production

Rabbit anti-CS21, anti-*IngB*, and anti-*IngC* sera were produced by immunization of New Zealand white rabbits with either purified CS21 obtained from the ETEC strain E9034A Δ IngR::km (Saldaña et al., unpublished data) or C-terminal His-Tag *IngB* and *IngC* purified proteins. Rabbits were immunized every 2 weeks, and emulsions in complete Freund's adjuvant (1 dose with 1 μ g of antigen in 500 μ l of PBS + 500 μ l of adjuvant) or incomplete Freund's adjuvant (3 doses with 0.5 μ g of antigen in 500 μ l of PBS + 500 μ l of adjuvant) were subcutaneously administered. The antisera obtained were adsorbed 8 times against E9034A Δ IngA, E9034A Δ IngB, and E9034A Δ IngC strains respectively, to remove nonspecific antibodies and increase the specificity. The antisera were used in Western blot (WB), immunofluorescence (IF), and immunogold assays, as described below.

Adherence Assays to HT-29 Cells

Adherence assays were carried out in 24-well tissue culture plates (Corning, NJ, USA), with or without 12 mm round coverslips

TABLE 2 | List of primers used in this study.

Primer	Sequence	Use
LngA-FC	5' →TCGCCATGGGGATCCAATTACGTAAACAACGT-3'	Cloning of <i>IngA</i> in pUC18
LngA-RC	5' →TTGAAGCTTTTAAACGGCTACCTAAAGTAATTG-3'	Cloning of <i>IngA</i> in pUC18
LngA-FP	5' →TGCGGATCCGTGATCTGAAGAAAAATAA-3'	Cloning of <i>IngA</i> in pJET1.2/blunt
LngA-RP	5' →TGTGAGAAGGTACTAGCCTATCATATT-3'	Cloning of <i>IngA</i> in pJET1.2/blunt
LngB-FM	5' →CACAGAGACAATTTATGAAAATGAGAGGCTTCGTGATAGGCTGGAGCTGCTTC-3'	Mutagenesis
LngB-RM	5' →CATTTTATCCTCTCCATTACATTAGGTTTGTGGTTCTGTACTATATGAATATCCTCCTTAG-3'	Mutagenesis
LngB-FS	5' →GTAGAGCTCAGGTAACCTCAATTACTTTAGGTAGCCGT-3'	Screening mutants
LngB-RS	5' →TGCAAGCTTCAGGAGATGCCAGTGCACCTCG-3'	Screening mutants
FIngBALL	5' →AGAAGGAGATATACTATGAAAATGAGAGGCTTCACACTTCTGGAGATGATTACTACTCTCGC-3'	Cloning of <i>IngB</i> in pLATE31
RIngBALL	5' →GTGGTGGTGATGGTGATGGCCGGTTTGTGGTTCTGTACTGCACCAGGTTGTGATA-3'	Cloning of <i>IngB</i> in pLATE31
LngC-FM	5' →AATGGAGAGGATAAAAATGAGAGCAAAATGGGTGATAGGCTGGAGCTGCTTC-3'	Mutagenesis
LngC-RM	5' →ACTTCATATTCTTATACCTGACATTCATTGAGAGTCGATAACCAGTACATATGAATATCCTCCTTAG-3'	Mutagenesis
LngC-FS	5' →CAACCATGGGGATCCCAGAACCACAAACCTAATG-3'	Screening mutants
LngC-RS	5' →TGATGAGAATTCCTGCAATGCCAATAAACTAACATC-3'	Screening mutants
FIngCALL	5' →AGAAGGAGATATACTATGAGAGCAAAATGGGGTGTTTTTTCTTCTTAGTATTCTTTCTTCGAG-3'	Cloning of <i>IngC</i> in pLATE31
RIngCALL	5' →GTGGTGGTGATGGTGATGGCCTTGAGAGTCGATAACCAGTACGTTGTTTC-3'	Cloning of <i>IngC</i> in pLATE31
LngD-FM	5' →ACTGGTTATCGACTCTCAATGAATGTCAGGTATAAGAAATGTGATAGGCTGGAGCTGCTTC-3'	Mutagenesis
LngD-RM	5' →CGATAATAGTCAATACCTACAACCATTGCGATGGTATTATCAATATGAATATCCTCCTTAG-3'	Mutagenesis
LngD-FS	5' →GTCCCATGGGGATCCGTTTTCTTCAGAACAAATAT-3'	Screening mutants
LngD-RS	5' →CCATAAGAGCTCCAGCGCAATTTTTTCATC-3'	Screening mutants
LngH-FM	5' →CTTGTACATAAAGTTAAATGACGGCCTTGCAGTCGTTGTGATAGGCTGGAGCTGCTTC-3'	Mutagenesis
LngH-RM	5' →CAAACTGATAGAGATAAATACGTTGTTTTTATTGAATTTTTTCAAGTAATACCTCATATGAATATCCTCCTTAG-3'	Mutagenesis
LngH-FS	5' →AGAGAATTCGGGAAAGTACAGGCTG-3'	Screening mutants
LngH-RS	5' →GAGTCATAGATCGGTAATCCTGAAAGCTTCAT-3'	Screening mutants
LngP-FM	5' →ATGTATGTTGAAATCGGCGTTTTCTTTTTTATTATTACAGTGTAGGCTGGAGCTGCTTC-3'	Mutagenesis
LngP-RM	5' →ATCTCTATGCATTTCTCATAGAAATAAGAGAAAGTATTCATAACATATGAATATCCTCCTTAG-3'	Mutagenesis
LngP-FS	5' →GTTGGATCCAGATTAGTGGCCCT-3'	Screening mutants
LngP-RS	5' →GATGTCGACGACTCTATAGATATTAATCTCTAT-3'	Screening mutants

(Bellco Glass, NJ, USA). Colon adenocarcinome HT-29 (ATCC HTB-38, VA, USA) cell monolayers were used as previously described (Saldana et al., 2009). Briefly, approximately 5×10^6 colony forming units (CFU) grown overnight in PPLO were added to HT-29 monolayers at 80% confluence (1×10^5 cells), which were then incubated at 37°C for 6 h in a 5% CO₂ atmosphere. Infected monolayers were washed with PBS 1X and then lysed with 0.1% Triton X-100 (Amresco Bioscience, OH, USA) in PBS 1X, and the bacteria were quantified by plate counts. Adhesion assays were performed in triplicate on three different days to obtain an average of the data expressed as CFU/ml. The standard deviations were calculated from all the results and represented as error bars in the graphs. Samples with coverslips were fixed with 2% formaldehyde in PBS 1X for Giemsa staining or for immunofluorescence microscopy (IFM) as previously described by Saldana et al. (2009).

Immunofluorescence Assay

The assay was carried out to visualize CS21 assembly by E9034A ETEC strain and isogenic mutants attached to HT-29 cells. Samples were incubated for 5 min at room temperature with 0.1% Triton X-100 in PBS, followed by a 30 min incubation

with RNase (Sigma-Aldrich-Co. LLC, MO, USA) diluted 1:5000 in PBS 1X with 10% horse serum (PBS-HS) at 37°C. Cells were then incubated with anti-CS21 serum diluted 1:2000 in PBS-HS, followed by goat anti-rabbit IgG antibodies conjugated with Alexa Fluor® 488 (Abcam, Cambridge, UK) diluted 1:2500 in PBS-HS and incubated during 1 h at 37°C. Cells were washed 3 times with 1 ml of PBS after each step. Coverslips were mounted on glass slides with 3 µl of DAPI (Merck Millipore, USA) or propidium iodide (Sigma-Aldrich-Co. LLC, MO, USA). Samples were visualized under an Axio Imager 2 Research Microscope (Zeiss, GER). Confocal microscopy images were taken in a Leica TCS SP8x microscope with a 63x oil immersion objective and digital zoom of 5x. The samples were analyzed with the Leica Application Suite software; Advanced Fluorescence Lite version 2.6.3 build 8173 (LAS AF Lite; Leica Microsystems, GER).

Transmission Electron Microscopy (TEM) and Immuno-Electron Microscopy (IEM)

Bacterial cultures drops (10 µl) were placed onto formvar-coated 300 mesh copper grids and adsorbed for 5 min, fluid excess was wiped with filter paper. Subsequently, 4 min negative staining was performed with 6 µl of 1% phosphotungstic acid (pH 7.4)

(EMS, PA, USA); any excess was removed using filter paper. The grids were rinsed with two drops of distilled water and finally were air-dried. The samples were examined by TEM using a JEM-1010 microscope (JEOL, Tokyo, Japan). IEM studies were performed to confirm the presence or absence of CS21 and LngB on E9034 strains and the isogenic mutants. The copper grids covered with the bacterial cultures were incubated for 1 h with anti-CS21 or anti-LngB antibodies (diluted 1:10 in PBS containing 10% BSA). The samples were rinsed with a drop of PBS three times, followed by 1 h incubation with a goat anti-rabbit IgG H&L conjugated to 10 nm gold particles (Sigma-Aldrich-Co, MO, USA) and rinsed again as previously described. Subsequently, a negative staining was performed as previously described.

Protein Electrophoresis and Western-Blot

For the analysis of whole-cell extracts, 5 ml of PPLO cultures were grown overnight with shaking at 37°C or supernatants from adherence assay were adjusted to 1 ml of culture with an optical density (OD₆₀₀) of 1.0, harvested by centrifugation and lysed by heating at 95°C for 5 min in 200 µl of loading buffer (Laemmli, 1970). Whole-cell extracts from equivalent cell numbers were resolved by SDS-PAGE and endogen chaperone DnaK was used as a loading control. Twenty-five microliter aliquots were loaded into SDS-PAGE gels. Gels were run at 100 V for 2 h at room temperature. For the WB, proteins separated by electrophoresis were transferred to PDVF membranes at 21 V for 1 h. The blots were blocked with PBS containing 0.1% (vol/vol) Tween 20 and 5% milk. Blocked membranes were reacted for 1 h with anti-CS21 or anti-DnaK antibodies (MBL International, MA, USA) in PBS-Tween 0.1%, washed 3 times with PBS-Tween 0.1% and incubated for 1 h with goat anti-rabbit IgG conjugated with horseradish peroxidase (Sigma-Aldrich-Co, MO, USA). The membranes were washed and revealed by chemiluminescence (ECL) (Amersham Life Science, Ill, USA).

Self-Aggregation Assay

Wild-type and isogenic mutants were grown overnight in 5 ml of TB at 37°C. A 1:100 dilution of each culture was incubated in flat-bottom 24-well tissue culture plates (Corning, NY, USA) at 37°C for 4 h. The bacterial self-aggregation phenotype consists of bacterial clumping. The clumps were visualized directly on the culture plates using an inverted light microscope (Olympus, Center Valley, PA; Clavijo et al., 2010). Triplicates of the assays were performed at three different times. ETEC E9034A and *E. coli* DH5α strains were used as positive and negative controls, respectively. Bacterial-bacterial aggregation mediated by CS21 causes sedimentation and media clearance. Quantitative analysis was accomplished by measuring the absorbance at 600 nm of the supernatant cultures without disturbing the aggregates.

Bioinformatic Analysis

Bacterial protein subcellular localization was predicted using the following servers: CELLO v.2.5, from the Molecular Bioinformatics Center (<http://cello.life.nctu.edu.tw/>), which is an amino acid composition-based method; PSLpred from

the Bioinformatics center Institute of Microbial Technology (<http://www.imtech.res.in/raghava/pslpred/>); and PSORTb version 3.0.2 from the Brinkman Laboratory, Simon Fraser University (<http://www.psort.org/psortb/>). Bioinformatics programs PSORTb and PSLpred integrate various protein characteristics, such as evolutionary information of PSI-Blast, amino acid, and dipeptide composition, as well as 33 physicochemical properties. CELLO uses 4 types of sequence coding schemes: the amino acid composition, the di-peptide composition, the partitioned amino acid composition, and sequence composition based on the physicochemical properties of amino acids. The analysis of protein domains was performed using the following servers: PROSITE from the Swiss Institute of Bioinformatics (SIB) (<http://prosite.expasy.org/prosite.html>), and Conserved Domain Database (CDD) from the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/cdd/>). The signal peptide in the amino acid sequences was predicted using these servers: InterPro from the European Bioinformatics Institute (<http://www.ebi.ac.uk/interpro/>); SignalP4.1 from the Center for Biological Sequence Analysis (<http://www.cbs.dtu.dk/services/SignalP/>), and Phobius from the Stockholm Bioinformatic Center (<http://phobius.sbc.su.se/>). Prediction of transmembrane helices was done using the TMHMM Server v. 2.0 from the Center for Biological Sequence Analysis (<http://www.cbs.dtu.dk/services/TMHMM/>), HMMTOP v 2.0 from the Research Centre for Natural Sciences (<http://www.enzim.hu/hmmtop/html/submit.html>) and Phobius programs.

Statistical Analysis

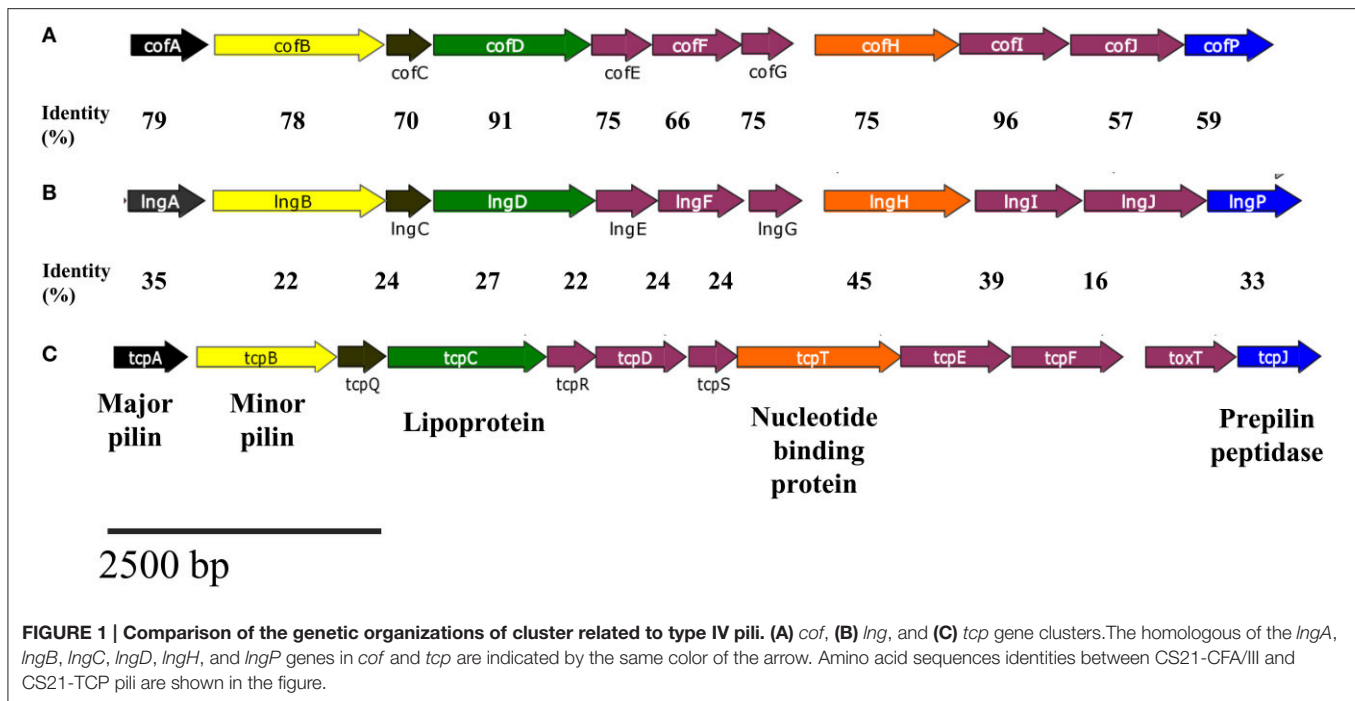
Data corresponding to adherence assays, and biofilm formation were analyzed using the Unpaired Student's *t*-test from the GraphPad software Inc. (La Jolla, CA, USA). The *p*-value used in the study was ≤0.005 as point of statistical significance.

RESULTS

TCP, CFA/III, BFP, and CS21 Share Homologous Proteins among Them

TCP, CFA/III, BFP, and CS21 are type IVb pili, which biogenesis requires between 11 and 14 genes, encoded within the pilus operon (Roux et al., 2012). Several proteins involved in TCP, CFA/III, and BFP assembly are homologous with proteins encoded in the *lng* cluster (Gomez-Duarte et al., 2007; Roux et al., 2012). Comparison of the genetic organizations of *cof*, *lng*, and *tcp* gene clusters show a common genetic organization, and a similar size (Figure 1). The percentage of identity between CS21-CFA/III and CS21-TCP genes products are shown in Figure 1. LngA, CofA, and TcpA are the major pilin subunits and share a high percentage of identity to the amino acid level (Roux et al., 2012). LngB and CofB proteins have been proposed to be the minor pilin subunits of CS21 and CFA/III, respectively (Kolappan et al., 2015).

The gene and amino acid sequence similarity between CS21 components and those of CFA/III and TCP suggested the functions of the CS21 proteins. Additionally, we analyzed the



amino acid sequences using predictive tools to gain further support for these assignments. Automatized prediction of bacterial protein subcellular localization is an important step to elucidate its function. Prediction analysis is based on the presence of a specific motif, protein characteristics, and homology to proteins with known localization (Yu et al., 2004, 2010; Bhasin et al., 2005). CELLO, PSLpred, and PSORTb were used to analyze the amino acid sequences of *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, and *lngP* (Table 4). *lngA*, *cofA*, and *tcpA* proteins are the major pilin subunits of CS21, CFA/III, and TCP, respectively; their localization was predicted to be extracellular and experimentally has been demonstrated (Tables 3, 4). *lngC* and *cofC* are proteins with unknown function whose subcellular localization was predicted to be in the outer membrane (Table 4). On the other hand, *tcpQ* (the homolog of *lngC* with 24% of identity) is required for *tcpC* (secretin of TCP) stability and outer membrane localization (Bose and Taylor, 2005). *lngD* has a putative function as a multimeric outer membrane lipoprotein, and *lngP* as prepilin peptidase. *lngD* and *lngP* were predicted to be in the outer and inner membrane, respectively (Table 4). *lngD* homologs *tcpC* and *bfpB* were demonstrated experimentally to be lipoproteins and are located in the outer membrane (Ramer et al., 1996; Bose and Taylor, 2005). *cofJ*, *tcpJ*, and *bfpP* are homologous of *lngP*, with 59, 33, and 32% of identity, respectively. The homologous of *lngP* have been demonstrated to be prepilin peptidases of *cofA*, *tcpA*, and *bfpA*, respectively (Kaufman et al., 1991; Zhang et al., 1994; Taniguchi et al., 1999). Membrane protein topology prediction of *lngP* predicts inner membrane localization with 6 internal helices. *lngH* is predicted to be a nucleotide binding protein; its homolog *tcpT* (45% of identity) is the ATPase of TCP (Iredell and Manning, 1997). The subcellular localization of *lngH* was predicted into the

cytoplasm. On the other hand, *tcpT* was localized in the inner membrane in a *tcpR*-dependent manner (Tripathi and Taylor, 2007).

The *lngB*, *lngC*, *lngD*, *lngH*, and *lngP* Proteins are Essential for CS21-Mediated Adherence to HT-29 Cells and Bacterial Self-Aggregation

CS21 is required to mediate human colonic cells colonization by ETEC CS21⁺ (Mazariego-Espinosa et al., 2010). We made deletions of the *lngB*, *lngC*, *lngD*, *lngH*, and *lngP* genes in the E9034A strain to evaluate their role in CS21 adherence. The growth rate was not affected in the mutants when compared with the wild-type strain (Supplemental Figure 1). Mutant strains showed 85–90% reduction on adherence to HT-29 cells compared to the wild-type strain (Figure 2). Bacterial self-aggregation mediated by CS21 causes bacterial clumping and media clearance. The mutants examined did not form clumps and sediment as the wild-type strain (Figure 2B). Complementation in *trans* of each mutant restored the adhesive properties and bacterial self-aggregation to wild type levels (Figures 2C, 3). All proteins tested in this study were involved in CS21 assembly and indirectly in adherence to HT-29 cells and bacterial self-aggregation.

lngP Is the Prepilin Peptidase of *lngA* in the E9034A Strain

Type IV class b prepilins are processed by group I PilD-like prepilin peptidases (Ayers et al., 2010), resulting in the methylation of the first residue at the N-terminus of the mature

TABLE 3 | Function of CS21, CFA/III, and TCP proteins.

Protein	Function	References	Protein	Function	References	Protein	Function	References
LngA	Major pilin subunit	Girón et al., 1994	CofA	Major pilin subunit	Taniguchi et al., 1994, 1995	TcpA	Major pilin subunit	Taylor et al., 1987
LngB	Minor pilin subunit	Kolappan et al., 2015	CofB	Minor pilin subunit	Kawahara et al., 2015; Kolappan et al., 2015; Kawahara et al., 2016	TcpB	Unknown	
LngC	Unknown		CofC	Initiate pilus assembly Unknown		TcpQ	Required for proper localization of TcpC to the outer membrane	Bose and Taylor, 2005
LngD	Lipoprotein*	Gomez-Duarte et al., 2007	CofD	Lipoprotein*	Taniguchi et al., 2001	TcpC	Lipoprotein	Bose and Taylor, 2005
LngE	Unknown		CofE	Unknown		TcpR	Bitopic IM protein*	Tripathi and Taylor, 2007
LngF	Unknown		CofF	Unknown		TcpD	Bitopic IM protein*	Roux et al., 2012
LngG	Unknown		CofG	Unknown		TcpS	Unknown	
LngH	Nucleotide binding, twitching motility*	Gomez-Duarte et al., 2007	CofH	Nucleotide binding protein*	Taniguchi et al., 2001	TcpT	ATPase	Iredell and Manning, 1997
LngI	Unknown		CofI	Unknown		TcpE	Polytopic IM protein*	Roux et al., 2012
LngJ	ATPase, twitching motility*	Gomez-Duarte et al., 2007	CofJ	Putative colonization factor	Yuen et al., 2013	TcpF	Colonization factor	Kirn et al., 2003
LngP	Prepilin peptidase*	Gomez-Duarte et al., 2007	CofP	Prepilin peptidase	Taniguchi et al., 1999	TcpJ	Prepilin peptidase	Kaufman et al., 1991; LaPointe and Taylor, 2000

*Predicted function.

TABLE 4 | Subcellular localization of CS21 proteins.

Protein	Putative function	Localization					
		CELLO	Scale 0–5	PSLPred (Hybrid approach)	Scale 0–100%	PSORTb	Scale 0–10
LngA	Major pilin subunit	Extracellular	2.641	Extracellular	53.1	Extracellular	9.71
LngB	Minor pilin subunit	Extracellular	2.973	Extracellular	71.1	Unknown	
LngC	Unknown	Outer membrane	2.329	Outer membrane	53.1	Unknown	
LngD	Outer membrane lipoprotein	Outer membrane	4.611	Outer membrane	98.1	Outer membrane	9.92
LngH	ATPase	Cytoplasmic	3.610	Cytoplasmic	98.1	Cytoplasmic	9.97
LngP	Prepilin peptidase	Inner membrane	4.744	Inner membrane	71.1	Inner membrane	10

protein and cleavage of the signal peptide (Hobbs and Mattick, 1993).

LngP protein is a putative prepilin peptidase with a conserved N-domain involved in the methyltransferase activity (CXXC and GXCXXC residues) and two highly conserved aspartate residues in the C-terminal domain that could be involved with the peptidase activity (Roux et al., 2012). LngA prepilin processing was affected in the E9034AΔ*lngP* strain, as demonstrated by an incomplete cleavage of LngA (Figure 4). The immunoblot showed two bands in the E9034AΔ*lngP* strain, a faint band

of 22-kDa (pilin) and a stronger band of 25-kDa (prepilin), unlike the E9034A strain, which showed a single band of 22-kDa (pilin) (Figure 4). The deletion of the *lngP* gene showed a reduction in adherence of 86% compared to the wild-type strain (Figure 2). The complementation of E9034AΔ*lngP* strain in *trans* with the plasmid pUC*lngP* restored the adherence levels and LngA cleavage (Figures 3, 4). This result suggests that LngP is the prepilin peptidase of LngA, and furthermore, that the E9034A strain has another prepilin peptidase capable of processing LngA prepilin, although with less efficiency. A minor

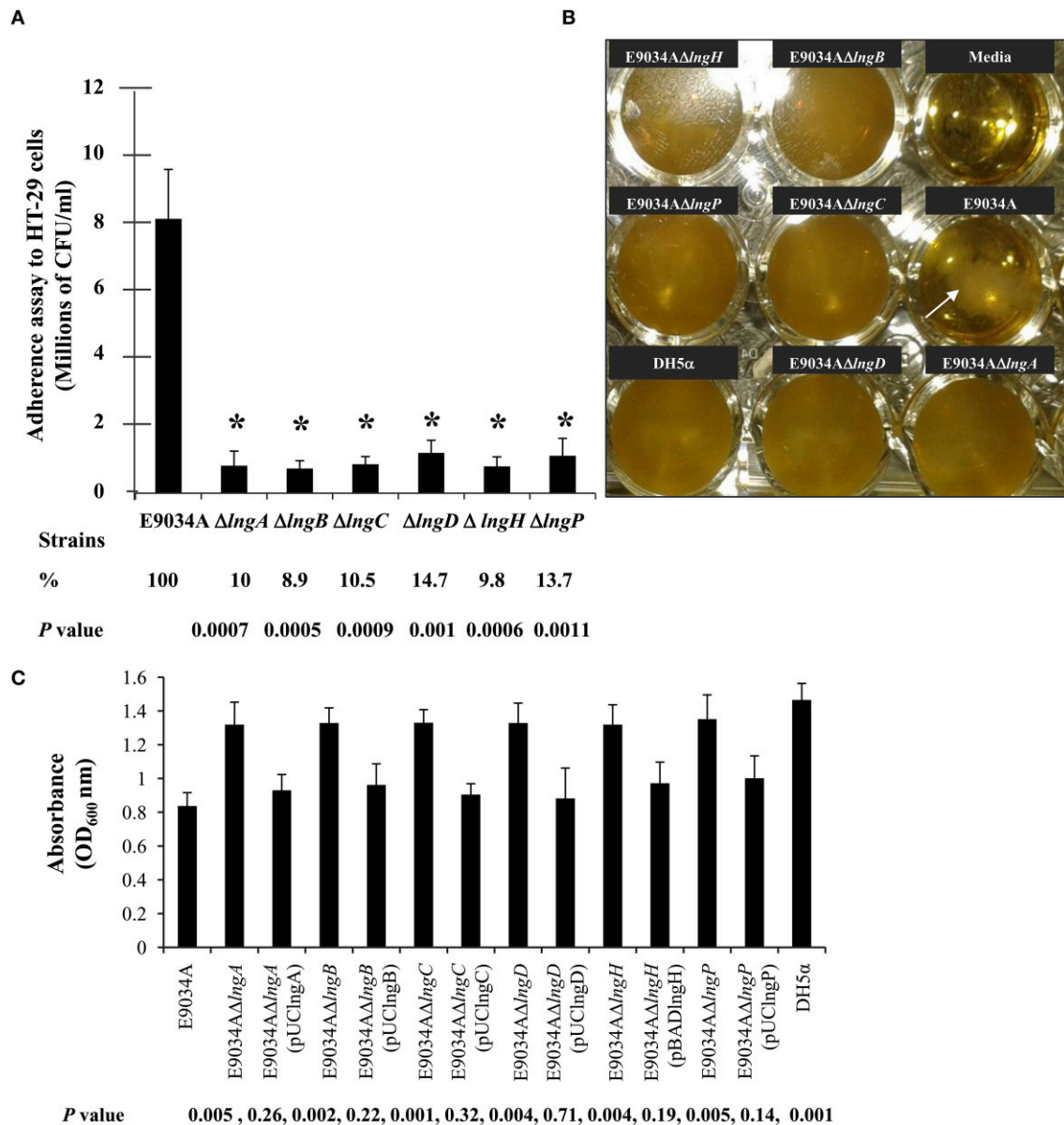
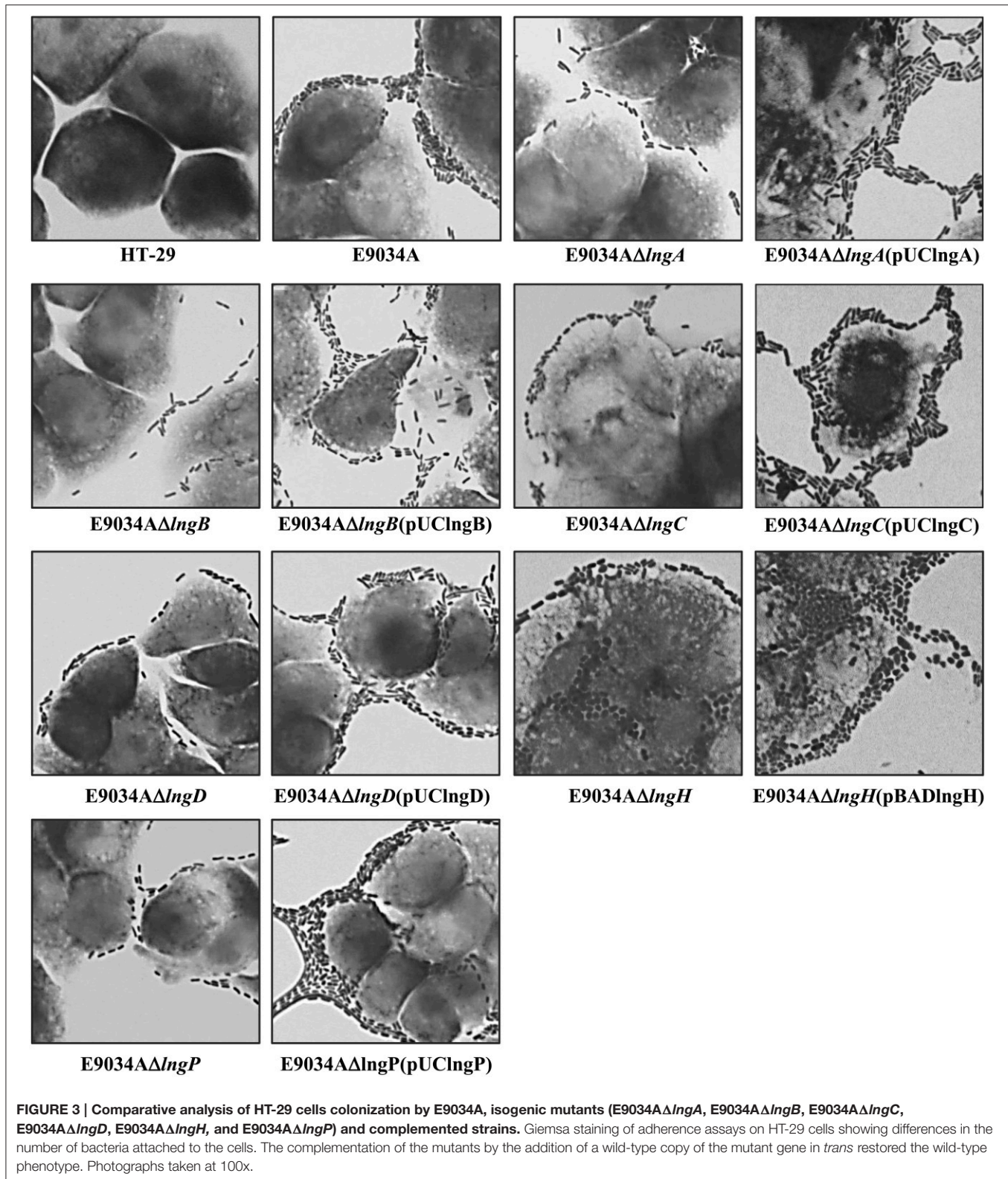


FIGURE 2 | Mutagenesis of *IngA*, *IngB*, *IngC*, *IngD*, *IngH*, and *IngP* genes reduce CS21 adherence to HT-29 cells and bacterial self-aggregation. (A) Quantitative analysis of HT-29 cells adherence by E9034A and isogenic mutants. **(B)** Self-aggregation assay in TB showing bacterial clumping in the E9034A strains (arrow) but not in the mutants examined. Non-inoculated media and the *E. coli* DH5 α strain were used as negative controls. **(C)** Quantitative analysis of the self-aggregation assay. Absorbance at 600 nm from the supernatant of TB media with each mutant and complemented strains was plotted. These data represent the average of two experiments repeated on different days in triplicate. * $p \leq 0.005$ statistically significant difference compared to the wild-type strain.

proportion of LngA pilin identified in the E9034A Δ IngP strain by WB, suggests that this strain could be able to assemble CS21. To test this hypothesis we examined the E9034A Δ IngP and wild-type strain by immunofluorescence assay. Immunofluorescence assay demonstrated the presence of CS21 characteristic structures in the E9034A and E9034A Δ IngP (pUCIngP) strains, but not in the E9034A Δ IngA strain (Figures 4B, 5). In the E9034A Δ IngP strain a faint reaction was observed by IF assays; however, the presence of CS21 structures was not evident by IEM (Figures 4B, 5).

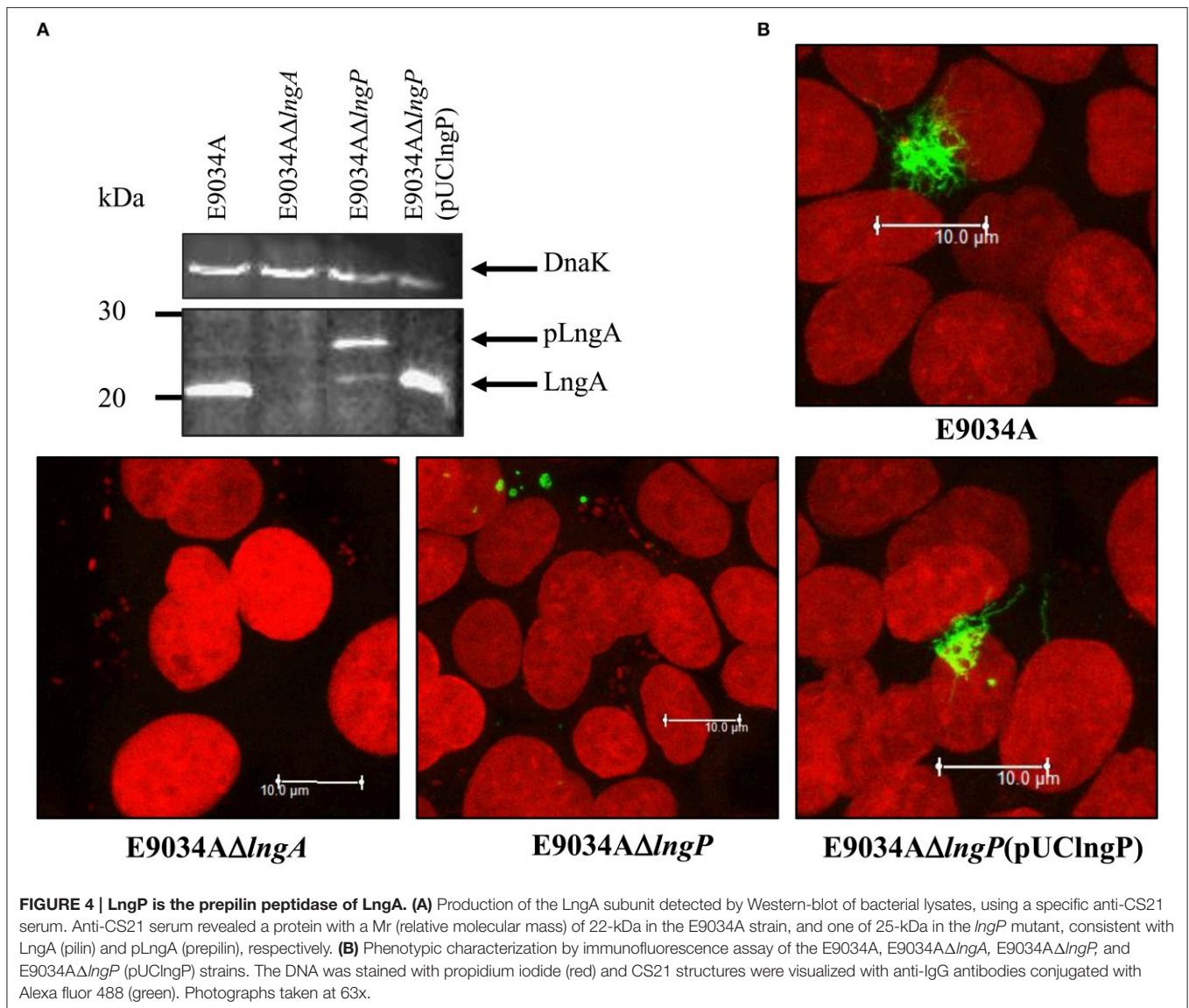
The Anti-CS21 Serum Recognizes LngA and LngB Proteins

To confirm CS21 assembly on the wild-type strain and mutants generated, immunogold-labeling TEM was carried out using specific anti-CS21 serum. Gold particles were found attached to long, polar, fimbrial structures on E9034A strain, typical features of CS21 pilus (Figure 5). In contrast, all the mutants tested by IEM no CS21 structures were observed (Figure 5). Clusters of gold particles on the cell surface were observed in the *lngA*, *lngC*, *lngD*, *lngH*, and *lngP* mutants instead,



but not in the E9034AΔIngB strain. The absence of gold particles in the E9034AΔIngB strain was indicative of another antigen recognized by the anti-CS21 serum and also it showed that the secondary antibody did not label unspecifically the

bacteria. These results suggest that the anti-CS21 serum could be recognizing both LngA and LngB proteins. Whole cell lysate of the E9034A strain and recombinant his-tag proteins of LngB and LngC were subjected to WB assay. The anti-CS21

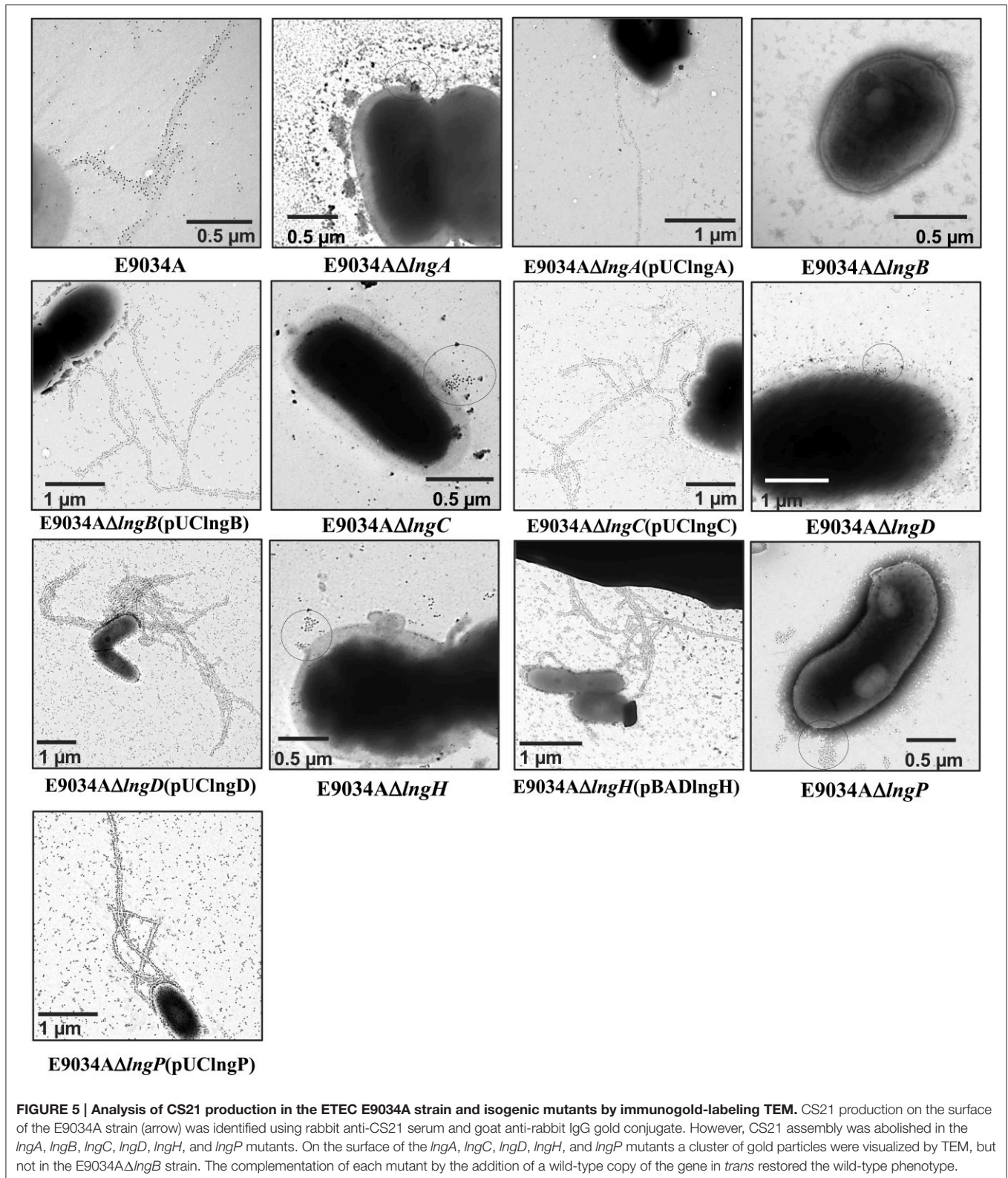


serum recognized two proteins, one of ~22-kDa and other of ~60-kDa in the E9034A strain (Supplemental Figure 2A). The calculated molecular weight of the LngB protein is 57.04-kDa and the recombinant His-tag protein of 57.98-kDa. LngB-His and LngC-His recombinant proteins resolved by SDS-PAGE showed a molecular weight that correspond with the calculated molecular weight (Supplemental Figure 2B). Interestingly, the anti-CS21 serum recognizes the LngB-His protein, but not LngC-His protein; these results confirm that the anti-CS21 serum recognized LngA and LngB proteins (Supplemental Figures 2C,D). Notably, anti-CS21 serum by IEM recognizes both CS21 filament and tip as shown by the gold particles associated to the end of the filament (Supplemental Figures 2E,F).

LngB Is a Minor Pilin Subunit of CS21

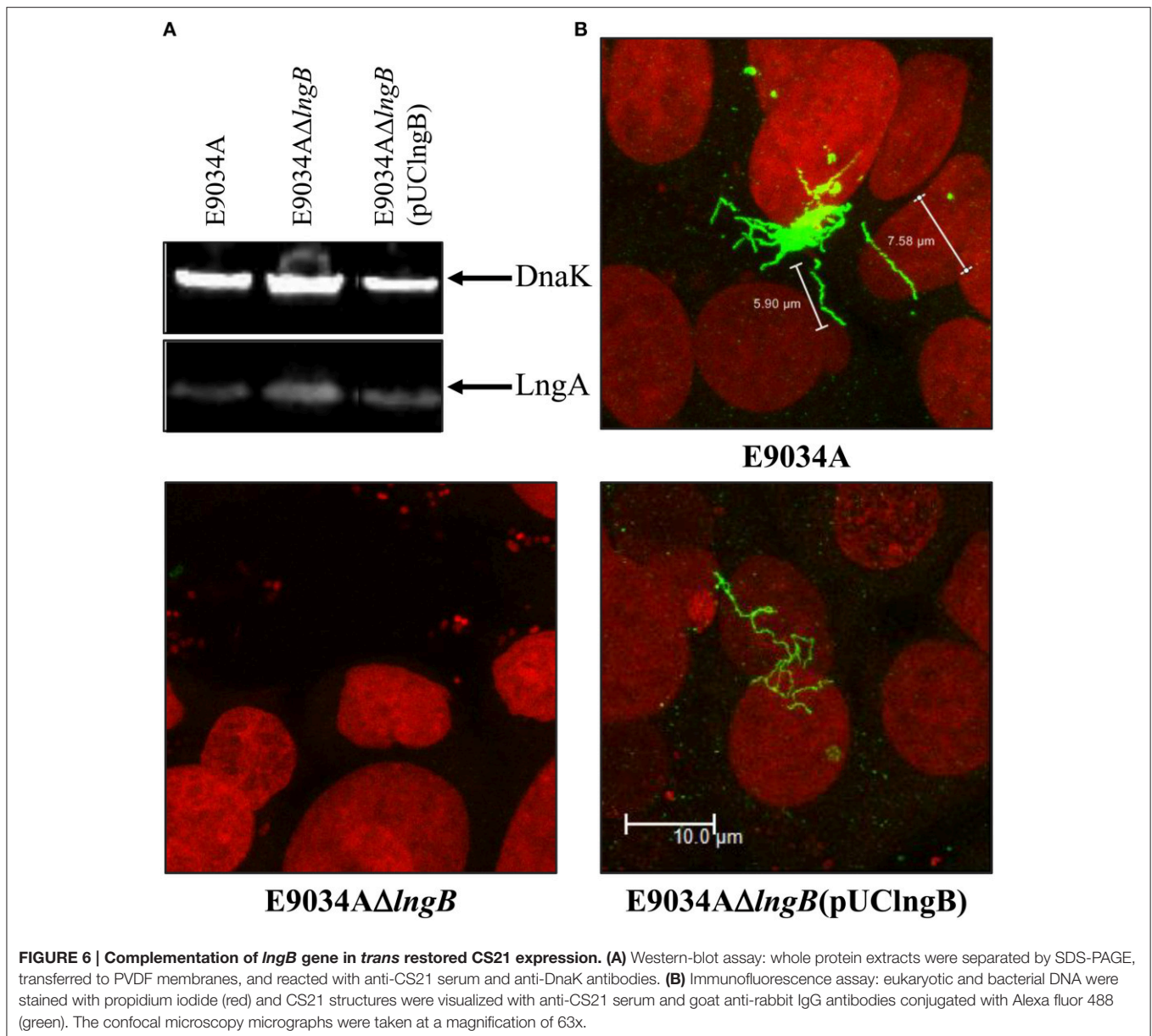
LngB protein is a minor pilin subunit, necessary to initiate the assembly of the CS21 pilus. CofB protein is the homologous of LngB protein in CFA/III and has been demonstrated that the C-terminal region of this protein is required to initiate

the assembly of CFA/III pilus (Kolappan et al., 2015). In this study, we reproduced the results previously demonstrated by Kolappan et al. (2015). The *lngB* deletion did not affect LngA expression and processing as demonstrated by WB (Figure 6A). However, it disrupts CS21 assembly as shown by IEM and IF assay (Figures 5, 6B). These data suggest that LngB plays a role in LngA assembly, but not in LngA expression or processing. The subcellular localization of LngB was predicted to be extracellular, and this prediction correlates with the localization of the gold particles observed in the *lngA*, *lngC*, *lngD*, and *lngH* mutant strains, but not in the E9034AΔ*lngB* strain (Figure 7C). Immunogold assays using specific anti-LngB serum suggested that LngB is localized at the tip of CS21 in the E9034A strain (Figures 7A,B). ETEC E9034A strain produces more than one filament, the filaments intertwined forming hair like or bundles as described by Girón et al. (1991). We propose that the presence of the gold particles indicate the presence of various filaments associated with LngB at CS21 tip.



Genetic restoration of the *IngB* gene carried on pUClngB into the E9034AΔ*IngB* strain restored CS21 assembly as demonstrated by the presence of long, polar, and

flexible structures (Figure 6B, Supplemental Figure 3). Immunofluorescence assays using anti-CS21 antibodies showed that the length of CS21 in the E9034AΔ*IngB* (pUClngB) strain



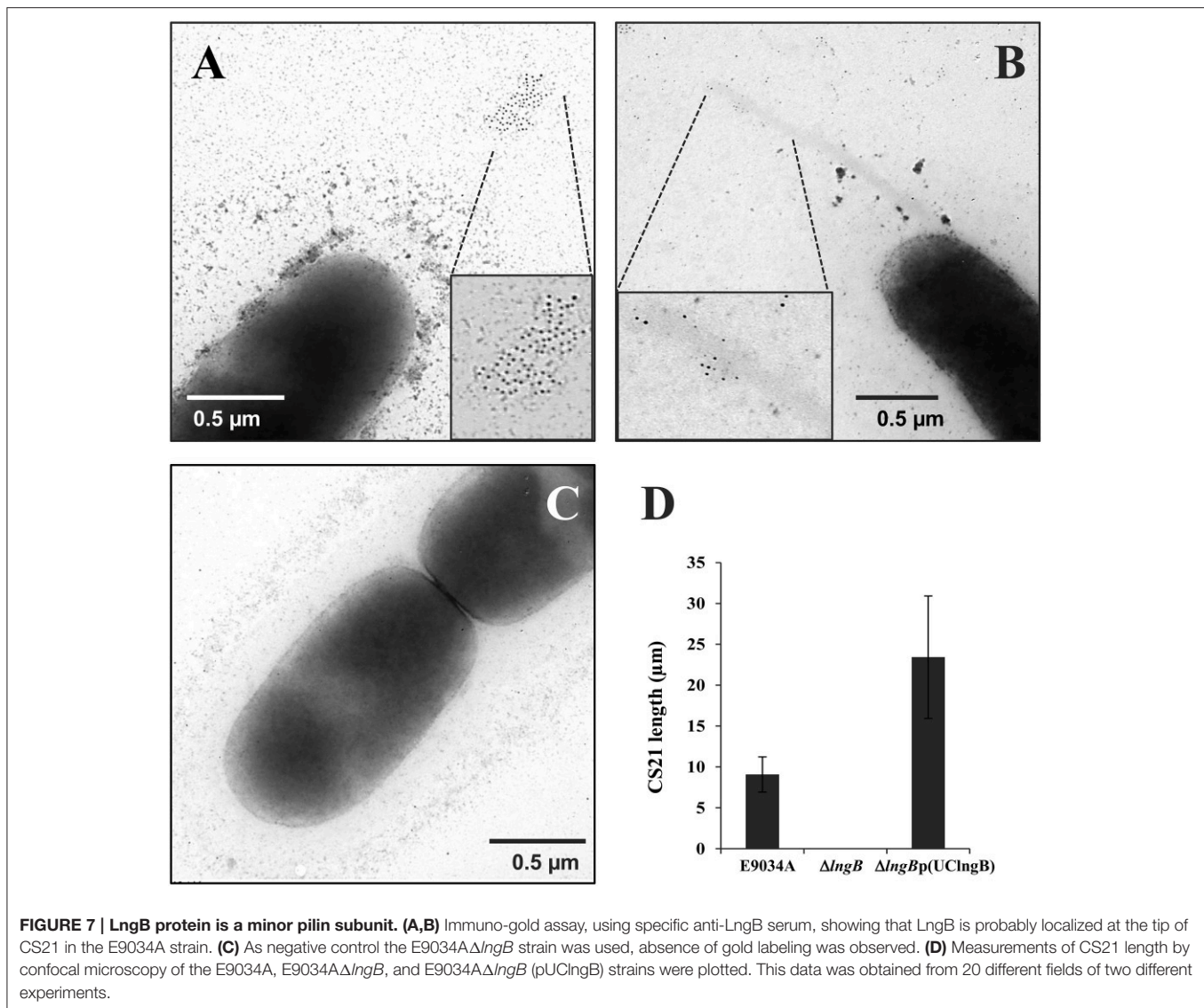
was longer than the structures visualized in the wild-type strain (**Figure 6B**). Quantitative analysis of confocal images allowed us to measure CS21 length; additionally, we found that the average length of CS21 from the E9034A strain was $9.01 \pm 2.05 \mu\text{m}$ and the E9034AΔ*lngB* (pUCLngB) was $22.62 \pm 6.77 \mu\text{m}$ (**Figure 7D**, Supplemental Figure 3). On the other hand, the E9034AΔ*lngA* (pUCLngA) strain that was complemented using the same vector as the E9034AΔ*lngB* (pUCLngB), did not show an CS21 length increase as the one observed for the E9034AΔ*lngB* (pUCLngB) strain (**Figure 8**).

LngC, LngD, and LngH Proteins are Required for CS21 Assembly

LngC has an identity of 24, 20.3, and 70% with TcpQ, BfpG, and CofC, respectively (**Figure 1**), and is predominantly at the

C-terminus, in which a conserved protein domain of the TcpQ family was found. The TcpQ protein is required for proper localization of TcpC protein (secretin) in the outer membrane of the bacteria (Bose and Taylor, 2005). On the other hand, the BfpG protein is required for the formation and/or stability of the multimer protein but not for the localization of BfpB protein (secretin) in the outer-membrane of the bacteria (Schmidt et al., 2001).

LngD and LngH are homologous to the outer membrane lipoprotein and nucleotide-binding protein, respectively, related to other type IV pili biogenesis apparatus (Roux et al., 2012). LngD has an identity of 17.6, 19.7, and 91% with TcpC, BfpB, and CofD respectively, and its C-terminus is homologous with the conserved domain of *pilus_B_mal_scr*, member of the secretin protein superfamily. On the other hand, LngH has an identity



of 21.9, 44.7, and 75% with BfpD, TcpT, and CofH, respectively. LngH belongs to the P-loop NTPase superfamily, characterized by a conserved nucleotide phosphate-binding motif, also referred to as the Walker A motif [GxxxxGK(ST)] (Iyer et al., 2004).

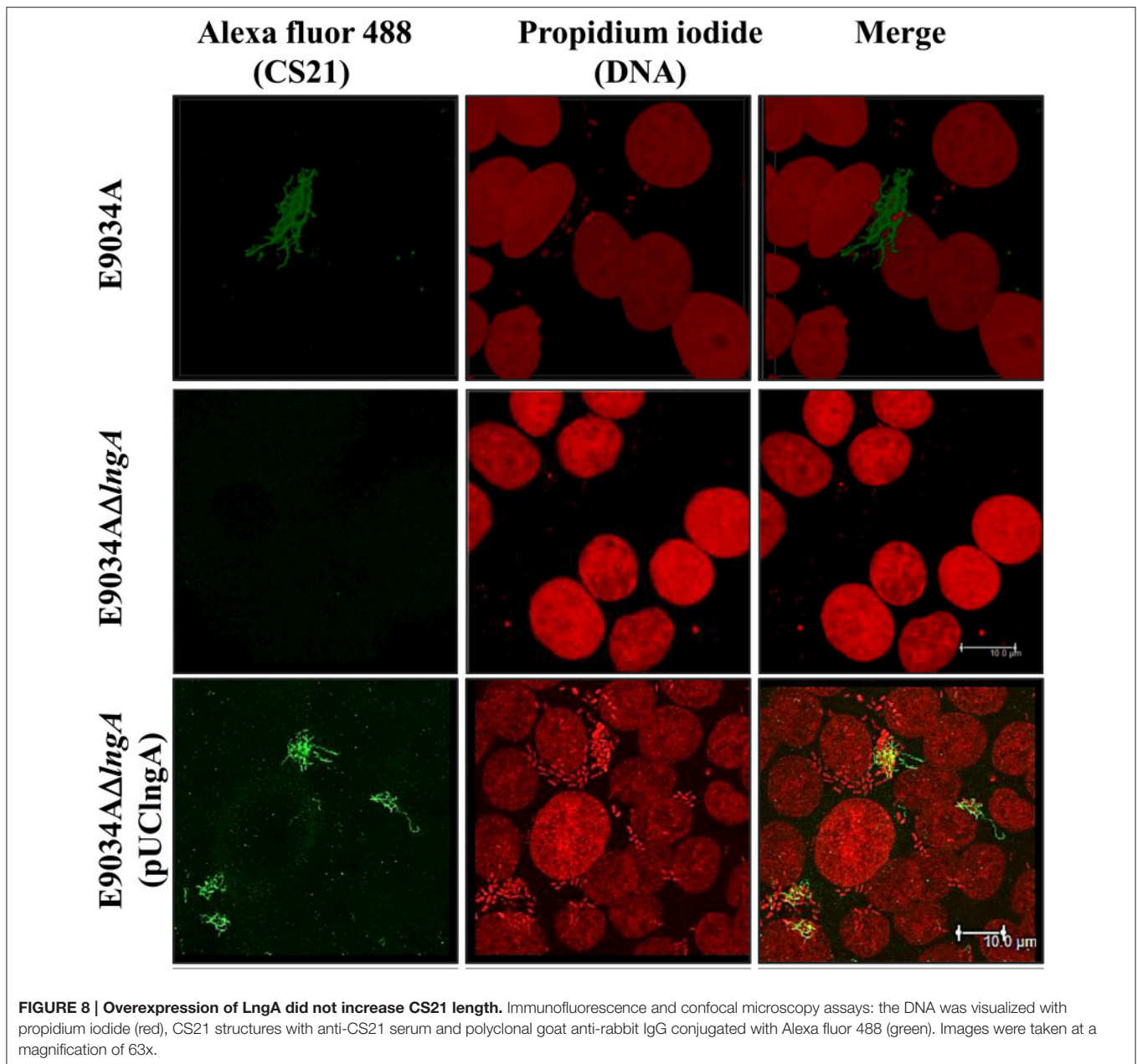
The deletion of *lngC* in the E9034A strain did not affect LngA expression, but the processing of LngA was incomplete as demonstrated by WB, with the recognition of two bands (prepilin and pilin), being the lower band (pilin) the one with higher intensity (Figure 9B). CS21 assembly was abolished in the E9034A Δ lngC strain as demonstrated by IEM and IF (Figures 5, 9C). Complementation of the *lngC* mutant by the addition of a wild-type copy gene in *trans* restores the wild-type phenotype (Figures 3, 5, 9). LngC has a signal peptide of 20 amino acids as predicted using the following softwares: SignalP4.1, Inter Pro and Phobius (Supplemental Figure 4A). The recombinant LngC-His protein was purified and subjected to SDS-PAGE. A double band of approximately 14 and 16 kDa was observed by Coomassie staining (Supplemental Figure 4B).

Western blot analysis demonstrated a positive reaction with both bands using anti-His antibodies showing that LngC is processed (Supplemental Figure 4C).

Interestingly, the processing of LngA in the E9034A Δ lngD strain showed two bands by WB that correspond to prepilin (pLngA) and pilin (LngA); as a result of *lngD* gene deletion, CS21 pilus assembly was abolished (Figures 9B,C). In contrast, the E9034A Δ lngH strain did not affect the expression and processing of the LngA protein, but it affects the CS21 assembly (Figures 9A,C).

DISCUSSION

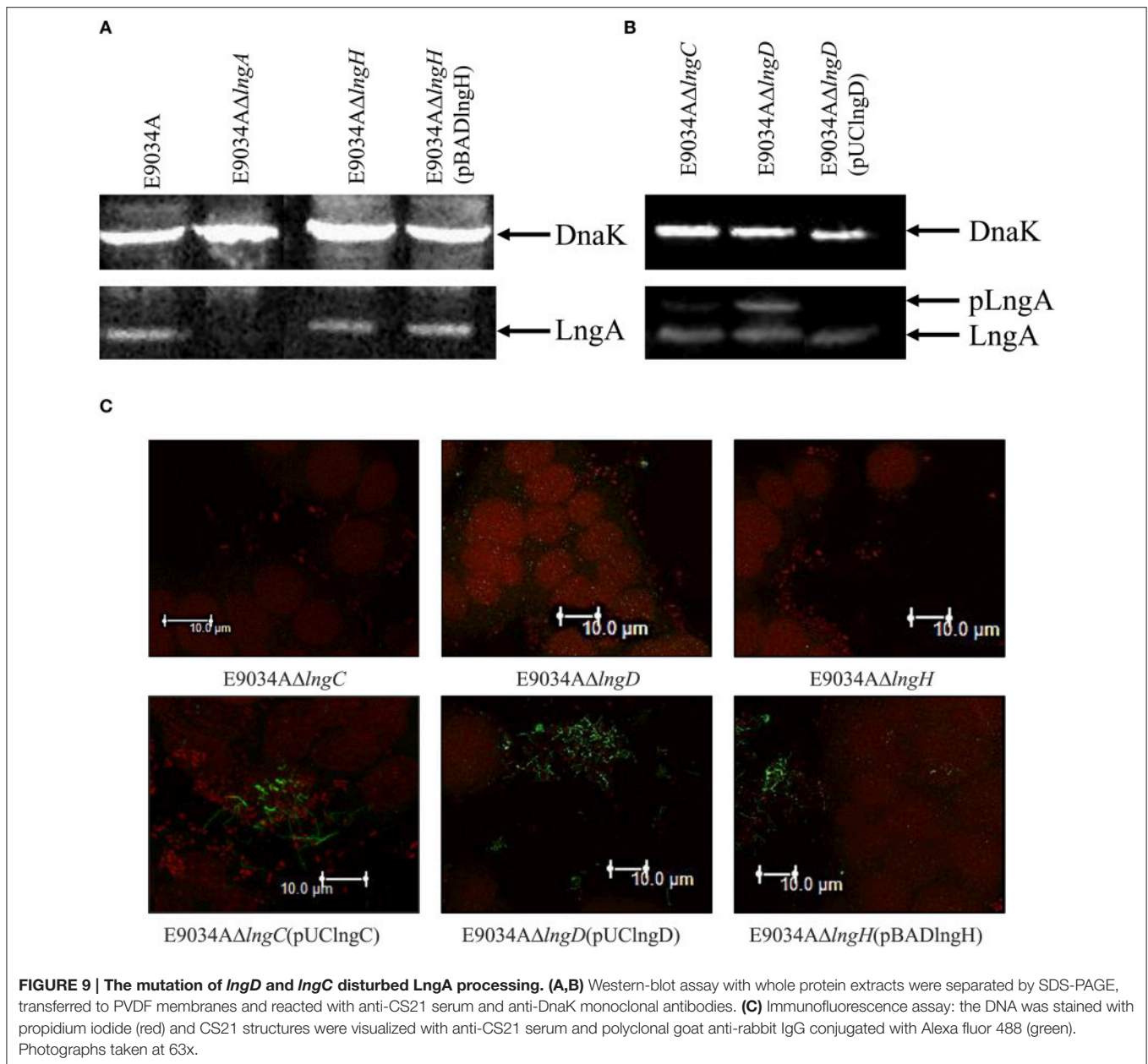
Type IV pili assembly is a process that requires the interaction of a complex biogenesis apparatus as described for the TCP, and BFP pili (Roux et al., 2012). Despite the homology that keep the different proteins involved in the biogenesis of CS21, TCP, BFP, and CFA/III, there are differences in the process of their assembly



(Roux et al., 2012). Although, CS21 is related to TCP and BFP, the proteins involved in the assembly process have not yet been explored and the efforts only focused in elucidating how the LngA protein contributes in the colonization of ETEC to intestinal cells (Mazariego-Espinosa et al., 2010; Guevara et al., 2013).

A cluster of 14 genes in tandem (*lngR*, *lngS*, *lngT*, *lngA*, *lngB*, *lngC*, *lngD*, *lngE*, *lngF*, *lngG*, *lngH*, *lngI*, *lngJ*, and *lngP*) may be required for CS21 assembly (Gomez-Duarte et al., 2007). Additionally, proteins coded in the *lng* cluster are homologous to proteins involved in the biogenesis of TCP, BFP, and CFA/III (Gomez-Duarte et al., 2007; Roux et al., 2012). This homology was essential to propose the hypothesis about the protein's function coded in the *lng* cluster and required for CS21 assembly.

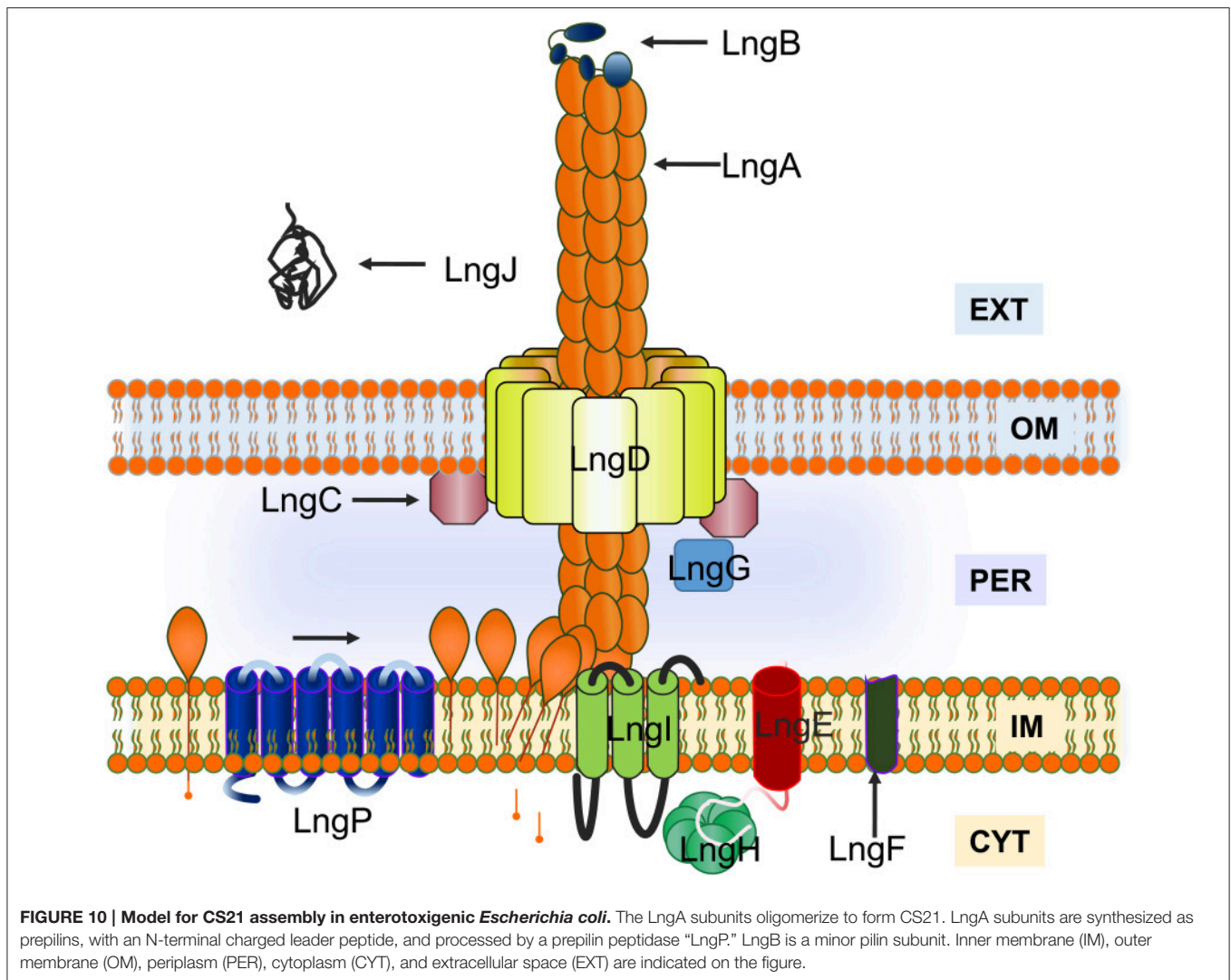
Subcellular localization of CS21 proteins provides an important clue to their function. Protein subcellular localization by conventional techniques is costly and time consuming. Recently, many algorithms have been developed for subcellular localization prediction, based on neural networks and the supervised learning used in support vector machines (Reinhardt and Hubbard, 1998; Bodén and Hawkins, 2005; Matsuda et al., 2005). Analyses of subcellular localization and experimental data showed that LngA, BfpA and TcpA are extracellular proteins anchored to the inner membrane cell (Taylor et al., 1987; Girón et al., 1991, 1994). Despite the unknown function, LngC has been predicted to be an outer membrane protein according to bioinformatic analysis. We support this result based on the fact that the homologs proteins



BfpG and TcpQ were designated as outer membrane proteins (Bose and Taylor, 2005; Daniel et al., 2006). Localization analysis of LngD (outer membrane protein) and LngP (inner membrane protein) correspond to BfpB/TcpC and BfpP/Tcp), respectively. LngH and TcpT are homologous proteins, and their subcellular localization prediction was in the cytoplasm. TcpT and LngH lack any predicted transmembrane domain as it has been reported for other ATPases (Tripathi and Taylor, 2007). LngH has the typical Walker A (also known as the phosphate-binding loop) and B motif, which are associated with phosphate binding and ATP hydrolysis, respectively (Iyer et al., 2004). We hypothesize, that LngH is associated with CS21 assembly apparatus favored by the interactions with unknown proteins that are anchored

to the inner membrane. LngB protein is a minor pilin subunit, whose subcellular localization was predicted to be extracellular protein by CELLO and PSLpred predictors; however, there is no experimental data about the subcellular localization of the LngB.

The prepilin proteins require processing by a prepilin peptidase, before translocation and oligomerization (Hobbs and Mattick, 1993; Strom and Lory, 1993). In this study, we demonstrated that the *IngP* gene coded for a prepilin peptidase and its absence shows a reduction in the LngA processing; however, deletion of the *IngP* gene did not completely abolish LngA processing. This data suggested the presence of another prepilin peptidase capable of processing the LngA protein in the E9034A strain. More than 30 genes in the *E. coli* K-12



genome sequence are homologous to genes of the general secretory pathway, involved in either secretion or type IV piliation including two with homology to genes coding for the prepilin peptidases *gspO* and *pppA* (Francetic and Pugsley, 1996; Francetic et al., 1998; Pugsley and Francetic, 1998). The *gspO* gene encodes a functional enzyme; however, its transcription levels are very low under laboratory conditions (Francetic and Pugsley, 1996). On the other hand, the *pppA* gene has been shown to code for a functional prepilin peptidase capable of processing typical prepilin peptidase substrates such as: prePulG (*Klebsiella oxytoca* type IV pre-pseudopilin), prePile (*Neisseria gonorrhoeae* type IV prepilin), and prePpdD or preHcpA (enterohemorrhagic *E. coli* type IV prepilin “HCP”; Whitchurch and Mattick, 1994; Francetic et al., 1998; Xicohtencatl-Cortes et al., 2007). Interestingly, the LngA protein was completely processed in the *E. coli* TOP10 strain harboring the *lngA* gene in *trans* (Supplemental Figure 5). This data suggested that *E. coli* TOP10 prepilin peptidase (probably PppA) is functional and capable of cleaving typical prepilin peptidase substrates. In contrast,

LngA partial processing observed in the E9034AΔ*lngP* strain, suggests the presence of another prepilin peptidase, probably PppA. The PppA and/or GspO proteins in ETEC strain E9034A may be involved in the maturation of LngA; however, they could be expressed at low levels compared to *E. coli* TOP10. Interestingly, partial processing of LngA was not enough for assembly of CS21, as demonstrated by IF and IEM using anti-CS21 serum. A previous report mentioned that a 5-kb Bam HI restriction fragment from the pE9034A mega plasmid was sufficient for pilus assembly and bacterial aggregation in an *E. coli* strain K-12 (DH5αF'1Q) (Girón et al., 1994). The 5-kb Bam HI restriction fragment from the partial sequences of the pE9034A plasmid was analyzed and we found that the *lngR*, *lngS*, *lngT*, *lngA*, and *lngB* (truncated sequence with 1106 bp) genes are contained in this fragment. *E. coli* K-12 contains 16 genes at seven different loci in its chromosome that encode components of type IV piliation machinery (Sauvonnet et al., 2000). These studies suggest that other genes in the *E. coli* K-12 chromosome are coding for proteins of type IV pilus and

could replace the functions of the remaining CS21 proteins in the *E. coli* K-12 (DH5 α F'IQ) but not in the *E. coli* E9034A strain.

The *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, and *lngP* deleted strains showed a significant reduction in adherence to HT-29 cells, due to CS21 assembly failure. Gold particles evidenced the presence of long, polar, and fimbrial structures emerging from the E9034A strain surface, when reacted against anti-CS21 serum. However, *lngA*, *lngC*, *lngD*, *lngH*, and *lngP* deleted strains, showed a cluster of gold particles on the bacterial surface, but no CS21 labeled structures. Interestingly, a total lack of gold particles in the *lngB* gene deleted strain suggests that the anti-CS21 serum recognizes LngA and LngB proteins. Detection of LngA and LngB by the anti-CS21 serum could be the explanation for denser clusters in the *lngP* mutant strain than the rest of the mutants examined. To identify LngB localization, we generated a specific anti-LngB serum and showed for first time an extracellular localization for the LngB protein.

In this study, the mutation of *lngB* gene did not affect LngA expression and processing, but abolished CS21 assembly, and overexpression of LngB increased CS21 length in the E9034A Δ *lngB* (pUC Δ lngB) strain. In contrast, Kolappan et al. (2015) found that LngB overexpression did not affect CS21 assembly. Additionally, the three-dimensional structure of the CofB protein (homologous to LngB) determined by crystallography allowed to propose a model where CofB initiate and control the filament growth of CFA/III (Kolappan et al., 2015). Docking of CofB into CFA/III pilus filament model suggests a tip localization and is consistent with CofB's role as an initiator of pilus assembly (Kolappan et al., 2015). LngB protein is a minor pilin subunit, which subcellular localization by informatics analysis, homology with CofB and IEM images suggested is extracellular and probably at the tip of CS21. Future experiments will be conducted to determine if LngB is an adhesin of CS21.

The proteins coded by *lngC* or *lngD* genes might have an important role on pili basal apparatus assembly. We propose that the LngA prepilin, observed in the E9034A Δ *lngD* strain, causes an allosteric inactivation of LngP (prepilin peptidase) due to accumulation of LngA in the periplasm by the absence of the LngD (outer membrane secretin). The presence of unprocessed LngA in the E9034A Δ *lngC* strain suggests that LngC could be required for LngD stability and/or outer membrane localization. TcpQ, the homolog of LngC is required for TcpC (secretin of TCP) stability and outer membrane localization (Bose and Taylor, 2005). Deletion of *lngC* or *lngD* genes affected LngA processing, and thus, CS21 assembly, this effect is similar to their homologs in TCP. In this study, we showed that purified LngC is cleaved by a peptidase, probably LngP, as determined by a double band observed by Coomassie staining. LngC has 20.3% identity with BfpG, and the localization of BfpG to the outer membrane is dependent on BfpB (homolog of LngD; Daniel et al., 2006). The mutation of *bfpB* or *bfpG* genes does not affect BfpA expression or processing, but inhibits BFP biogenesis (Anantha et al., 2000). Interestingly, the purified BfpG-His protein migrates on acrylamide gels as a doublet. The authors propose that the mature BfpG is not a lipoprotein, and showed that purified

BfpG exists in two forms after cleavage at either of two typical signal peptidase I sites, as determined by amino-terminal acid sequencing (Daniel et al., 2006). Similarly, the purified LngC-His protein migrates on acrylamide gels as a doublet; however, it was predicted that the LngC has a single cleavage site and the molecular weight from upper and lower band matches the molecular weight of the immature (16.29-kDa) and mature (13.99-kDa) protein, respectively.

Type IVb pili have a genetic organization that required proteins coded by 12–14 genes clustered into operons (Roux et al., 2012). The proteins coded by genes in the *lng* operon have homology with proteins involved in TCP and BFP assembly. This homology was a tool to understand the assembly process of the CS21 pilus. In this study, it was demonstrated that LngA, LngB, LngC, LngD, LngH, and LngP proteins are essential to CS21 pilus assembly and the deletion of the genes that code for these proteins affect the bacterial self-aggregation and adherence phenotypes of E9034A strain. The assembly of type IVb pili requires various proteins forming the piliation assembly machinery into inner and outer membrane subcomplexes. A model for CS21 assembly was proposed based on proteins localization prediction, homology to other characterized proteins, identification of motifs, and *lngA*, *lngB*, *lngC*, *lngD*, *lngH*, and *lngP* genes deletions (Figure 10). LngA is the major pilin subunit that oligomerizes forming the filament. The N-terminal helices for the pilin subunits are bound to the inner membrane and are incorporate into the growing pilus from the inner membrane. LngP is an inner membrane prepilin peptidase that process LngA prepilin. LngB protein is a minor pilin subunit. LngC was predicted to be localized in the outer membrane, and based on its homology with TcpQ of TCP, could be required for LngD stability and/or outer membrane localization. The C-terminal of LngD has a region with homology to the conserved protein domain of the secretin superfamily, which is used by type IV pilus to translocate pilin subunits and macromolecules across the outer membrane. A nucleotide-binding protein is required to provide energy to the assembly of CS21. The *lngH* gene, encodes a putative ATPase, "LngH." The remaining *lngR*, *lngS*, *lngT*, *lngE*, *lngF*, *lngG*, *lngI*, and *lngJ* genes were not included in this model, but are currently being studied. In conclusion, the model proposed for the CS21 assembly in this study is similar to the models for other type IV pili as TCP and CFA/III.

AUTHOR CONTRIBUTIONS

Designed and conceived the experiments: ZS, JX. Performed the experiments: ZS, KE, and VR. Analyzed the data: ZS, BG, and JX. Contributed reagents/materials/analysis tools: BG, AC, GP, EL, JA, SO, RH, CE, and JX. Wrote and reviewed the manuscript: ZS, GR, AC, VR, KE, MG, BG and JX.

ACKNOWLEDGMENTS

We acknowledge Dra. Norma Espinosa Sánchez and Dra. María del Rosario Espinosa Mellado for technical assistance. Dr. Oscar G. Gómez-Duarte and Francisco Xavier Montiel Infante for its

comments and suggestions to the article. ZS had a CONACyT doctoral fellowship with the number 275780 and data in this work is part of his doctoral dissertation in the Programa de Doctorado en Ciencias Biomédicas, Instituto de Fisiología Celular at the Universidad Nacional Autónoma de México (UNAM).

REFERENCES

- Anantha, R. P., Stone, K. D., and Donnenberg, M. S. (2000). Effects of *bfp* mutations on biogenesis of functional enteropathogenic *Escherichia coli* type IV pili. *J. Bacteriol.* 182, 2498–2506. doi: 10.1128/JB.182.9.2498-2506.2000
- Ayers, M., Howell, P. L., and Burrows, L. L. (2010). Architecture of the type II secretion and type IV pilus machineries. *Future Microbiol.* 5, 1203–1218. doi: 10.2217/fmb.10.76
- Bhasin, M., Garg, A., and Raghava, G. P. (2005). PSLpred: prediction of subcellular localization of bacterial proteins. *Bioinformatics* 21, 2522–2524. doi: 10.1093/bioinformatics/bti309
- Bodén, M., and Hawkins, J. (2005). Prediction of subcellular localization using sequence-biased recurrent networks. *Bioinformatics* 21, 2279–2286. doi: 10.1093/bioinformatics/bti372
- Bose, N., and Taylor, R. K. (2005). Identification of a TcpC-TcpQ outer membrane complex involved in the biogenesis of the toxin-coregulated pilus of *Vibrio cholerae*. *J. Bacteriol.* 187, 2225–2232. doi: 10.1128/JB.187.7.2225-2232.2005
- Clavijo, A. P., Bai, J., and Gómez-Duarte, O. G. (2010). The Longus type IV pilus of enterotoxigenic *Escherichia coli* (ETEC) mediates bacterial self-aggregation and protection from antimicrobial agents. *Microb. Pathog.* 48, 230–238. doi: 10.1016/j.micpath.2010.03.006
- Cruz-Córdova, A., Espinosa-Mazariego, K., Ochoa, S. A., Saldaña, Z., Rodea, G. E., Cázares-Domínguez, V., et al. (2014). CS21 positive multidrug-resistant ETEC clinical isolates from children with diarrhea are associated with self-aggregation, and adherence. *Front. Microbiol.* 5:709. doi: 10.3389/fmicb.2014.00709
- Daniel, A., Singh, A., Crowther, L. J., Fernandes, P. J., Schreiber, W., and Donnenberg, M. S. (2006). Interaction and localization studies of enteropathogenic *Escherichia coli* type IV bundle-forming pilus outer membrane components. *Microbiology* 152, 2405–2420. doi: 10.1099/mic.0.28860-0
- Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645. doi: 10.1073/pnas.120163297
- Francetic, O., Lory, S., and Pugsley, A. P. (1998). A second prepilin peptidase gene in *Escherichia coli* K-12. *Mol. Microbiol.* 27, 763–775. doi: 10.1046/j.1365-2958.1998.00723.x
- Francetic, O., and Pugsley, A. P. (1996). The cryptic general secretory pathway (*gsp*) operon of *Escherichia coli* K-12 encodes functional proteins. *J. Bacteriol.* 178, 3544–3549.
- Gaastera, W., and Svennerholm, A. M. (1996). Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends Microbiol.* 4, 444–452. doi: 10.1016/0966-842X(96)10068-8
- Girón, J. A., Gomez-Duarte, O. G., Jarvis, K. G., and Kaper, J. B. (1997). Longus pilus of enterotoxigenic *Escherichia coli* and its relatedness to other type-4 pili—a minireview. *Gene* 192, 39–43. doi: 10.1016/S0378-1119(97)00039-5
- Girón, J. A., Ho, A. S., and Schoolnik, G. K. (1991). An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. *Science* 254, 710–713. doi: 10.1126/science.1683004
- Girón, J. A., Levine, M. M., and Kaper, J. B. (1994). Longus: a long pilus ultrastructure produced by human enterotoxigenic *Escherichia coli*. *Mol. Microbiol.* 12, 71–82. doi: 10.1111/j.1365-2958.1994.tb00996.x
- Gomez-Duarte, O. G., Chattopadhyay, S., Weissman, S. J., Giron, J. A., Kaper, J. B., and Sokurenko, E. V. (2007). Genetic diversity of the gene cluster encoding longus, a type IV pilus of enterotoxigenic *Escherichia coli*. *J. Bacteriol.* 189, 9145–9149. doi: 10.1128/JB.00722-07
- Guevara, C. P., Luiz, W. B., Sierra, A., Cruz, C., Qadri, F., Kaushik, R. S., et al. (2013). Enterotoxigenic *Escherichia coli* (ETEC) CS21 pilus contributes to adhesion to intestinal cells and to pathogenesis under *in vivo* conditions. *Microbiology* 159, 1725–1735. doi: 10.1099/mic.0.065532-0
- Hobbs, M., and Mattick, J. S. (1993). Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. *Mol. Microbiol.* 10, 233–243. doi: 10.1111/j.1365-2958.1993.tb01949.x
- Iredell, J. R., and Manning, P. A. (1997). Translocation failure in a type-4 pilin operon: *rfb* and *tcpT* mutants in *Vibrio cholerae*. *Gene* 192, 71–77. doi: 10.1016/S0378-1119(97)00040-1
- Isidean, S. D., Riddle, M. S., Savarino, S. J., and Porter, C. K. (2011). A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression. *Vaccine* 29, 6167–6178. doi: 10.1016/j.vaccine.2011.06.084
- Iyer, L. M., Leipe, D. D., Koonin, E. V., and Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* 146, 11–31. doi: 10.1016/j.jsb.2003.10.010
- Kaper, J. B., Nataro, J. P., and Mobley, H. L. (2004). Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* 2, 123–140. doi: 10.1038/nrmicro818
- Kaufman, M. R., Seyer, J. M., and Taylor, R. K. (1991). Processing of TCP pilin by TcpJ typifies a common step intrinsic to a newly recognized pathway of extracellular protein secretion by gram-negative bacteria. *Genes Dev.* 5, 1834–1846. doi: 10.1101/gad.5.10.1834
- Kawahara, K., Oki, H., Fukakusa, S., Maruno, T., Kobayashi, Y., Motooka, D., et al. (2015). Cloning, expression, purification, crystallization and X-ray crystallographic analysis of CofB, the minor pilin subunit of CFA/III from human enterotoxigenic *Escherichia coli*. *Acta Crystallogr. F Struct. Biol. Commun.* 71, 663–667. doi: 10.1107/S2053230X15005890
- Kawahara, K., Oki, H., Fukakusa, S., Yoshida, T., Imai, T., Maruno, T., et al. (2016). Homo-trimeric Structure of the Type IVb Minor Pilin CofB Suggests Mechanism of CFA/III Pilus Assembly in Human Enterotoxigenic *Escherichia coli*. *J. Mol. Biol.* 428, 1209–1226. doi: 10.1016/j.jmb.2016.02.003
- Kirn, T. J., Bose, N., and Taylor, R. K. (2003). Secretion of a soluble colonization factor by the TCP type 4 pilus biogenesis pathway in *Vibrio cholerae*. *Mol. Microbiol.* 49, 81–92. doi: 10.1046/j.1365-2958.2003.03546.x
- Kolappan, S., Ng, D., Yang, G., Harn, T., and Craig, L. (2015). Crystal Structure of the Minor Pilin CofB, the Initiator of CFA/III Pilus Assembly in Enterotoxigenic *Escherichia coli*. *J. Biol. Chem.* 290, 25805–25818. doi: 10.1074/jbc.M115.676106
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0
- LaPointe, C. F., and Taylor, R. K. (2000). The type 4 prepilin peptidases comprise a novel family of aspartic acid proteases. *J. Biol. Chem.* 275, 1502–1510. doi: 10.1074/jbc.275.2.1502
- Levine, M. M., Ristaino, P., Marley, G., Smyth, C., Knutton, S., Boedeker, E., et al. (1984). Coli surface antigens 1 and 3 of colonization factor antigen II-positive enterotoxigenic *Escherichia coli*: morphology, purification, and immune responses in humans. *Infect. Immun.* 44, 409–420.
- Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., et al. (2013). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380, 2095–2128. doi: 10.1016/S0140-6736(12)61728-0
- Matsuda, S., Vert, J. P., Saigo, H., Ueda, N., Toh, H., and Akutsu, T. (2005). A novel representation of protein sequences for prediction of subcellular location using support vector machines. *Protein Sci.* 14, 2804–2813. doi: 10.1110/ps.051597405
- Mattick, J. S. (2002). Type IV pili and twitching motility. *Annu. Rev. Microbiol.* 56, 289–314. doi: 10.1146/annurev.micro.56.012302.160938
- Mazariego-Espinosa, K., Cruz, A., Ledesma, M. A., Ochoa, S. A., and Xicohtencatl-Cortes, J. (2010). Longus, a type IV pilus of enterotoxigenic *Escherichia coli*, is

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2016.01201>

- involved in adherence to intestinal epithelial cells. *J. Bacteriol.* 192, 2791–2800. doi: 10.1128/JB.01595-09
- Nataro, J. P., and Kaper, J. B. (1998). Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* 11, 142–201.
- Ogierman, M. A., and Manning, P. A. (1992). TCP pilus biosynthesis in *Vibrio cholerae* O1: gene sequence of tcpC encoding an outer membrane lipoprotein. *FEMS Microbiol. Lett.* 76, 179–184. doi: 10.1111/j.1574-6968.1992.tb05459.x
- Pugsley, A. P., and Francetic, O. (1998). Protein secretion in *Escherichia coli* K-12: dead or alive? *Cell. Mol. Life Sci.* 54, 347–352. doi: 10.1007/s000180050162
- Ramer, S. W., Bieber, D., and Schoolnik, G. K. (1996). BfpB, an outer membrane lipoprotein required for the biogenesis of bundle-forming pili in enteropathogenic *Escherichia coli*. *J. Bacteriol.* 178, 6555–6563.
- Reinhardt, A., and Hubbard, T. (1998). Using neural networks for prediction of the subcellular location of proteins. *Nucleic Acids Res.* 26, 2230–2236. doi: 10.1093/nar/26.9.2230
- Roux, N., Spagnolo, J., and de Bentzmann, S. (2012). Neglected but amazingly diverse type IVB pili. *Res. Microbiol.* 163, 659–673. doi: 10.1016/j.resmic.2012.10.015
- Saldana, Z., Erdem, A. L., Schuller, S., Okeke, I. N., Lucas, M., Sivananthan, A., et al. (2009). The *Escherichia coli* common pilus and the bundle-forming pilus act in concert during the formation of localized adherence by enteropathogenic *E. coli*. *J. Bacteriol.* 191, 3451–3461. doi: 10.1128/JB.01539-08
- Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Sauvonnnet, N., Gounon, P., and Pugsley, A. P. (2000). PpD Type IV Pilin of *Escherichia coli* K-12 Can Be Assembled into Pili in *Pseudomonas aeruginosa*. *J. Bacteriol.* 182, 848–854. doi: 10.1128/JB.182.3.848-854.2000
- Schmidt, S. A., Bieber, D., Ramer, S. W., Hwang, J., Wu, C. Y., and Schoolnik, G. (2001). Structure-function analysis of BfpB, a secretin-like protein encoded by the bundle-forming-pilus operon of enteropathogenic *Escherichia coli*. *J. Bacteriol.* 183, 4848–4859. doi: 10.1128/JB.183.16.4848-4859.2001
- Strom, M. S., and Lory, S. (1993). Structure-function and biogenesis of the type IV pili. *Annu. Rev. Microbiol.* 47, 565–596. doi: 10.1146/annurev.mi.47.100193.003025
- Taniguchi, T., Akeda, Y., Haba, A., Yasuda, Y., Yamamoto, K., Honda, T., et al. (2001). Gene cluster for assembly of pilus colonization factor antigen III of enterotoxigenic *Escherichia coli*. *Infect. Immun.* 69, 5864–5873. doi: 10.1128/IAI.69.9.5864-5873.2001
- Taniguchi, T., Arita, M., Sato, M., Yamamoto, K., Miwatani, T., and Honda, T. (1994). Evidence that the N-terminal amino acid sequence of pilus colonization factor antigen III produced by human enterotoxigenic *Escherichia coli* is similar to that of TcpA pilin of *Vibrio cholerae*. *J. Infect. Dis.* 170, 1049–1050. doi: 10.1093/infdis/170.4.1049
- Taniguchi, T., Fujino, Y., Yamamoto, K., Miwatani, T., and Honda, T. (1995). Sequencing of the gene encoding the major pilin of pilus colonization factor antigen III (CFA/III) of human enterotoxigenic *Escherichia coli* and evidence that CFA/III is related to type IV pili. *Infect. Immun.* 63, 724–728.
- Taniguchi, T., Yasuda, Y., Tochikubo, K., Yamamoto, K., and Honda, T. (1999). The gene encoding the prepilin peptidase involved in biosynthesis of pilus colonization factor antigen III (CFA/III) of human enterotoxigenic *Escherichia coli*. *Microbiol. Immunol.* 43, 853–861. doi: 10.1111/j.1348-0421.1999.tb01220.x
- Taylor, R. K., Miller, V. L., Furlong, D. B., and Mekalanos, J. J. (1987). Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci. U.S.A.* 84, 2833–2837. doi: 10.1073/pnas.84.9.2833
- Tripathi, S. A., and Taylor, R. K. (2007). Membrane association and multimerization of TcpT, the cognate ATPase ortholog of the *Vibrio cholerae* toxin-coregulated-pilus biogenesis apparatus. *J. Bacteriol.* 189, 4401–4409. doi: 10.1128/JB.00008-07
- Whitchurch, C. B., and Mattick, J. S. (1994). *Escherichia coli* contains a set of genes homologous to those involved in protein secretion, DNA uptake and the assembly of type-4 fimbriae in other bacteria. *Gene* 150, 9–15. doi: 10.1016/0378-1119(94)90851-6
- Xicohtencatl-Cortes, J., Monteiro-Neto, V., Ledesma, M. A., Jordan, D. M., Francetic, O., Kaper, J. B., et al. (2007). Intestinal adherence associated with type IV pili of enterohemorrhagic *Escherichia coli* O157:H7. *J. Clin. Invest.* 117, 3519–3529. doi: 10.1172/JCI30727
- Yu, C. S., Lin, C. J., and Hwang, J. K. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci.* 13, 1402–1406. doi: 10.1110/ps.03479604
- Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., et al. (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26, 1608–1615. doi: 10.1093/bioinformatics/btq249
- Yuen, A. S., Kolappan, S., Ng, D., and Craig, L. (2013). Structure and secretion of CofJ, a putative colonization factor of enterotoxigenic *Escherichia coli*. *Mol. Microbiol.* 90, 898–918. doi: 10.1111/mmi.12407
- Zhang, H. Z., Lory, S., and Donnemberg, M. S. (1994). A plasmid-encoded prepilin peptidase gene from enteropathogenic *Escherichia coli*. *J. Bacteriol.* 176, 6885–6891.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Saldaña-Ahuactzi, Rodea, Cruz-Córdova, Rodríguez-Ramírez, Espinosa-Mazariago, González-Montalvo, Ochoa, González-Pedrajo, Eslava-Campos, López-Villegas, Hernández-Castro, Arellano-Galindo, Patiño-López and Xicohtencatl-Cortes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



CS21 positive multidrug-resistant ETEC clinical isolates from children with diarrhea are associated with self-aggregation, and adherence

Ariadna Cruz-Córdova¹, Karina Espinosa-Mazariego¹, Sara A. Ochoa¹, Zeus Saldaña^{1,2}, Gerardo E. Rodea^{1,2}, Vicenta Cázares-Domínguez¹, Viridiana Rodríguez-Ramírez¹, Carlos A. Eslava-Campos^{3,4}, Armando Navarro-Ocaña⁴, José Arrellano-Galindo⁵, Rigoberto Hernández-Castro⁶, Oscar G. Gómez-Duarte⁷, Firdausi Qadri⁸ and Juan Xicohtencatl-Cortes^{1*}

¹ Unidad de Hemato-Oncología e Investigación, Laboratorio de Investigación en Bacteriología Intestinal, Hospital Infantil de México Federico Gómez, Mexico City, Mexico

² Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Mexico City, Mexico

³ Unidad de Hemato-Oncología e Investigación, Laboratorio de Patogenicidad Bacteriana, Hospital Infantil de México Federico Gómez, Mexico City, Mexico

⁴ Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional Autónoma de México, Mexico City, Mexico

⁵ Laboratorio de Infectología, Departamento de Infectología, Hospital Infantil de México Federico Gómez, Mexico City, Mexico

⁶ Departamento de Ecología de Agentes Patógenos, Hospital General "Dr. Manuel Gea González," Mexico City, Mexico

⁷ Department of Pediatrics, Division of Pediatric Infectious Diseases, Vanderbilt University School of Medicine, Nashville, TN, USA

⁸ Centre for Vaccine Sciences, International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh

Edited by:

Christina Maria Joseph Elisabeth
Vandenbroucke-Grauls, VU
University Medical Center,
Netherlands

Reviewed by:

Elizabeth B. Norton, Tulane
University, USA
Magdalena Chirila, Iuliu Hațieganu
University of Medicine and
Pharmacy, Romania

*Correspondence:

Juan Xicohtencatl-Cortes,
Laboratorio de Investigación en
Bacteriología Intestinal, Unidad de
Hemato-Oncología e Investigación,
Hospital Infantil de México Federico
Gómez, Dr. Márquez 162, Col.
Doctores, Delegación Cuauhtémoc,
Mexico City 06720, Mexico
e-mail: juanxico@yahoo.com

Background: Enterotoxigenic *Escherichia coli* (ETEC) colonize the human intestinal mucosa using pili and non-pili colonization factors (CFs). CS21 (also designated Longus) is one of the most prevalent CFs encoded by a 14 kb *Ing* DNA cluster located in a virulence plasmid of ETEC; yet limited information is available on the prevalence of CS21 positive ETEC isolates in different countries. The aim of this study was to evaluate the prevalence of CS21 among ETEC clinical isolates from Mexican and Bangladeshi children under 5 years old with diarrhea and to determine the phenotypic and genotypic features of these isolates.

Methods: ETEC clinical isolates positive to *IngA* gene were characterized by genotype, multidrug-resistance, self-aggregation, biofilm formation, and adherence to HT-29 cell line.

Results: A collection of 303 *E. coli* clinical isolates were analyzed, the 81.51% (247/303) were identified as ETEC, 30.76% (76/247) were *st*⁺/*lt*⁺, and 25.10% (62/247) were positive for the *IngA* gene. Among the *IngA*⁺ ETECs identified, 50% of isolates (31/62) were positive for *IngA* protein. The most frequent serotype was O128ac:H12 found in 19.35% (12/62) of *IngA*⁺ ETEC studied. Multidrug-resistance (MDR) *IngA*⁺ ETEC isolates was identified in 65% (39/60), self-aggregation in 48.38% (30/62), and biofilm formation in 83.87% (52/62). ETEC *IngA*⁺ isolates were able to adhere to HT-29 cells at different levels. Two *IngA* isogenic mutants were constructed in the ETEC E9034A and ETEC73332 clinical isolate, showing a 77% and 98% reduction in adherence, respectively with respect to the wild type.

Conclusion: ETEC isolates that have the *IngA* gene showed features associated with self-aggregation, and adherence to HT-29 cells, important characteristics in the human gut colonization process and pathogenesis.

Keywords: ETEC, CS21, multidrug-resistance, biofilm, self-aggregation and adherence

INTRODUCTION

Bacterial adherence is the first and most important step in bacterial infection. Diverse pathogenic mechanisms allow the bacteria to adhere to host tissues and aggregate at the infection site with the subsequent colonization and dissemination to other anatomical sites within the host (Nataro and Kaper, 1998; Craig et al.,

2004). Enterotoxigenic *Escherichia coli* (ETEC), a leading traveler's diarrhea etiologic agent, also cause 300,000–500,000 deaths annually in children under the age of 5 years living in developing countries (WHO, 2006). Profuse watery diarrhea is the disease's hallmark, and similarly to cholera, it is mediated by the secretion of a cholera-like heat-labile (LT) and/or heat-stable toxins

(ST) (Nataro and Kaper, 1998). These enterotoxins are believed to be encoded by large virulence plasmids and co-regulated with pili and non-pili surface proteins (Brinton, 1965; Beachey, 1981). Colonization of the small bowel mucosa by ETEC isolates is mediated by a broad variety of fimbrial and non-fimbrial surface structures called colonization factors (CFs). CFs includes colonization factor antigens (CFAs), coli surface antigens (CSs), and putative colonization factors (PCFs) (Gaastra and Svennerholm, 1996; Nataro and Kaper, 1998). More than 25 CFs have been described in human ETEC isolates (Qadri et al., 2005) and they are strongly associated with ETEC infection. Evidence of the role of CFs in pathogenesis include a study reporting that pili-less ETECs rarely induce diarrhea (Ahren and Svennerholm, 1985). CFs are highly prevalent in different geographic regions of the world and recognized by specific receptors not yet described (Gaastra and Svennerholm, 1996; Nataro and Kaper, 1998; Isidean et al., 2011). Epidemiological studies have shown that CFs produced by ETEC have immunogenic properties, and also have a protective immunity, which can be achieved through multiple infections by the host (Ahren and Svennerholm, 1985; Nataro and Kaper, 1998).

Human ETEC isolates express CS21 (Longus), a type IV pilus (T4P), with $>20 \mu\text{m}$ in length. CS21 is expressed at 37°C when grown on blood agar plates and Pleuropneumoniae-Like Organisms (PPLO) agar (Giron et al., 1994; Mazariego-Espinosa et al., 2010). CS21 is composed of a single repeating structural protein of 22-kDa called LngA. The *lngA* gene, encoding the LngA major subunit, is contained in a large virulence plasmid of 90 kb (Giron et al., 1994). The N-terminal amino acid sequence of the LngA protein shares homology with the CofA (CFA/III) pilin subunit ETEC, TcpA of *V. cholerae*, and BfpA of enteropathogenic *E. coli* (Taylor et al., 1987; Girón et al., 1991, 1997; Giron et al., 1994). The gene cluster organization and amino acid sequence identity, suggest that these genetically diverse T4P may have a common ancestor (Taniguchi et al., 2001).

CS21 plays an important role in the adherence to intestinal epithelial cells (HT-29, Caco-2, and T84 cells) and display twitching motility activity; likewise, it was shown that CS21 assembly is also influenced by nutritional growth conditions that phenotypically affect the expression of LngA (Mazariego-Espinosa et al., 2010). In addition, CS21 is associated with bacterial self-aggregation, protection against environmental stress, biofilm formation, and adherence to primary intestinal epithelial cells and recently (Guevara et al., 2013) demonstrated the role of CS21 in the pathogenesis of ETEC *in vivo* using a neonatal mice challenge infection model (Clavijo et al., 2010; Guevara et al., 2013). Human-ETEC isolates encoding CFAs isolates from different geographic regions (Argentina, Bangladesh, Chile, Brazil, Egypt, and Mexico), have shown a wide distribution of the *lngA* gene (Girón et al., 1995a; Gutierrez-Cazarez et al., 2000; Pichel et al., 2002). In addition to CS21, ETEC has another T4P called CFA/III for which *cofA* is the major subunit, both are paralogous in nature and belong to evolutionarily distinct types of fimbriae (Gomez-Duarte et al., 2007); however, CFA/III has lower prevalence than CS21 (Honda et al., 1984, 1985; Isidean et al., 2011). The aim of this work was to genotypically and phenotypically characterize a collection of clinical *lngA*⁺ ETEC isolates through PCR, resistance profiles, self-aggregation properties, biofilm formation, and

adherence to HT-29 cells to further understand if these attributes are related to the pathogenesis of CS21 positive clinical isolates from children with diarrhea.

MATERIALS AND METHODS

BACTERIAL ISOLATES

Previously, clinical isolates identified as ETEC by serotyping were obtained from an *E. coli* culture collection repository at the Facultad de Medicina from Universidad Nacional Autónoma de México (UNAM) for the last three decades. All clinical ETEC were isolated from stool samples of children under 5 years old with diarrhea from México and Bangladesh (Table S1). Isolates were kept at -70°C in Brain Heart Infusion broth (BHI, Difco, New Jersey, USA) with 15% glycerol, until their use. The ETEC isolates were cultured on 5% sheep blood agar plates (BBL, Franklin Lakes, NJ) at 37°C and on MacConkey agar (Difco, New Jersey, USA).

IDENTIFICATION OF VIRULENCE GENES BY PCR

Genomic DNA was purified using a commercial Wizard kit (Promega, USA) from a bacterial culture grown overnight in Luria Bertani broth (LB, Difco, New Jersey, USA) at 37°C . PCR assays were performed using a commercial PCR kit (Promega, USA) with the following reaction mixtures: $5 \mu\text{l}$ of reaction buffer [MgCl_2 (50 mM); *Taq* polymerase (0.1 U), and dNTPs (2 mM)], $0.5 \mu\text{l}$ of the forward primer (10 μM), $0.5 \mu\text{l}$ of the reverse primer (10 μM), $3 \mu\text{l}$ of water, and $1 \mu\text{l}$ of DNA (100 ng) from each isolate. PCR reactions were conducted in a thermal cycler (Applied Biosystems GeneAmp PCR system 9700, New York, USA), with the specific melting temperature for each primer (Table 1). PCR products were separated on a 1.5% agarose gel in TAE (Tris-Acetate-EDTA) at 100 v, stained with ethidium bromide (5 $\mu\text{g}/\text{ml}$), and visualized by UV transilluminator (Bio-Imaging Systems, AccesoLab, México, D.F., México). To identify the *st*, *lt*, *lngA*, and *cs3* and *cfal* genes in clinical isolates, the genomic DNAs of ETEC E9034A (O8:H9, *st*⁺, *lt*⁺, *lngA*⁺, and *cs3*⁺ genes) and ETEC H10467 (*cfal*⁺ gene) were used as positive controls; whereas, the genomic DNA of *E. coli* HB101 was used as negative control. Based on PCR results, ETEC isolates positive for the *lngA* gene were selected for further characterization.

IMMUNOBLOT

Whole cell extracts obtained from ETEC isolates positive for *lngA* were harvested from PPLO (Difco, New Jersey, USA), adjusted to the same concentration ($\text{OD}_{600} = 1.0$), boiled for 5 min with sample buffer and subjected to sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970; Xicohtencatl-Cortes et al., 2007). Proteins separated in a 16% SDS-PAGE were electroblotted onto nitrocellulose membranes (Bio-Rad, Hercules, CA, USA) and then incubated with rabbit anti-CS21 antibodies (1:3,000 dilution) followed by incubation with a secondary goat anti-rabbit immunoglobulin G (IgG) antibody conjugated to alkaline phosphatase (diluted 1:30,000). Blots were developed with BCIP/NTM alkaline phosphatase antibody detection reagent substrate (Millipore, Darmstadt, Germany).

Table 1 | Primer sequences, amplicon size, and melting temperatures used in this study.

Gene	Primer	Sequence 5'-3'	Size (bp)	T _m (°C)	References
<i>IngA</i>	F (J5)	ATG AGC CTG CTG GAA GTT ATC ATT G	608	62	Mazariego-Espinosa et al., 2010
	R (J6)	TTA ACG GCT ACC TAA AGT AAT TGA GTT			
<i>st</i>	F	ATT TTT CTT TCT GTA TTG TCT T	190	50	Perez et al., 2010
	R	CAC CCG GTA CAA GCA GGA TT			
<i>lt</i>	F	GGC CAC AGA TTA TAC CGT GC	450	50	Levy, 2002
	R	CGG TCT CTA TAT TCC CTG TT			
<i>cfal</i>	F	GGT GCA ATG GCT CTG ACC ACA	478	55.7	This study
	R	AGT AGT ATC TCT TGT AAT GAC			
<i>cooA (csi)</i>	F	GTC CAC ACC ATC AAC ACC GTT	320	62	This study
	R	ATT ATC CTG ACT AAG TCA ACG			
<i>cs3</i>	F	GGG CCC ACT CTA ACC AAA GAA	401	55.7	This study
	R	TTT AGT TTC AGG TAA TTA CCG			
<i>cofA (cfalIII)</i>	F	ATC CTT TCG GTT TAT AACAGA ACG G	713	45	This study
	R	CGG CTC GCC AAA GTA ATA GAG			

st: gene coding for a heat-stable enterotoxin; *lt*: gene coding for a heat-labile enterotoxin; *IngA*: gene coding for the major subunit of longus; *cfal*: gene coding for the colonization factor I; *cooA*: gene coding for the major subunit of CSI; *cs3*: gene coding for the CS3; *cofA*: gene coding for the major subunit of CFA/III.

ANTIBIOTIC SUSCEPTIBILITY

Antibiotic susceptibility test was determined for ETEC isolates positive for the *IngA* gene by the Kirby-Bauer disk-diffusion method as recommended by the Clinical and Laboratory Standards Institute 2013 (CLSI-2013). Five colonies from each isolate were grown in Mueller Hinton (MH) broth (Becton Dickinson, Maryland, USA) at 37°C with constant shaking for 2–5 h until reaching an optical density at 600 nm (OD_{600nm}) equivalent to 0.5 on the McFarland scale. MH agar plates were massively seeded with bacterial suspension using a sterile swab. Discs with the appropriate antibiotics were placed on the inoculated plates and incubated at 37°C for 18–24 h. For susceptibility testing a total of 12 antibiotic categories were used: cephalosporins I/II: 30 µg cephalothin (Becton Dickinson, Maryland, USA) and 30 µg cefaclor (Becton Dickinson, Maryland, USA); carbapenems: 10 µg meropenem (Oxoid Sunnyvale, California, USA) and 10 µg imipenem (Becton Dickinson, Maryland, USA); quinolones: 30 µg nalidixic acid (Becton Dickinson, Maryland, USA); penicillins: 10 µg ampicillin (Becton Dickinson, Maryland, USA); β-Lactam/β-lactamase inhibitor combination: 20/10 µg amoxicillin-clavulanate (Becton Dickinson, Maryland, USA), 100/10 µg piperacillin-tazobactam (Oxoid Sunnyvale, California, USA), and 75/10 µg ticarcillin-clavulanate (Oxoid Sunnyvale, California, USA); tetracyclines: 30 µg tetracycline (Becton Dickinson, Maryland, USA); folate pathway inhibitors: 1.25/23.75 µg trimethoprim-sulfamethoxazole (Becton Dickinson, Maryland, USA); aminoglycosides: 10 µg gentamicin (Becton Dickinson, Maryland, USA) and 30 µg amikacin (Becton Dickinson, Maryland, USA); phenicols: 30 µg chloramphenicol (Becton Dickinson, Maryland, USA); fluoroquinolones: 5 µg ciprofloxacin (Becton Dickinson, Maryland, USA); nitrofurans: 300 µg nitrofurantoin (Becton Dickinson, Maryland, USA); monobactams: 30 µg aztreonam (Becton Dickinson, Maryland, USA); and cephalosporin III/IV: 30 µg ceftriaxone (Becton Dickinson, Maryland, USA), 30 µg ceftazidime

(Becton Dickinson, Maryland, USA), and 30 µg cefepime (Becton Dickinson, Maryland, USA). Inhibition zones were determined and interpreted according to the recommendations of the CLSI-2013. *E. coli* ATCC (American Type Culture Collection) 25922 was used as quality control.

SELF-AGGREGATION ASSAY

The bacterial self-aggregation phenotype was analyzed using flat-bottom 24-well tissue culture plates (Corning, NY, USA); the assay consists of bacterial clumping during growth in liquid media. Bacteria grown overnight in Terrific broth (TB) (Amresco LLC, Ohio, USA) at 37°C were subcultured using a 1:100 dilution and incubated at 37°C in TB for 4 h. Self-aggregates were visualized directly on a bacterial suspension by an inverted light microscope (Olympus, Center Valley, PA) (Clavijo et al., 2010). Triplicates of the assays were performed at three different times. ETEC isolates: E9034A and *IngA*⁻ clinical isolate 114246 (background: O132:H25, *st*⁺, *lt*⁺, *ecpA*⁺) were used as positive and negative controls, respectively. To establish the role of *IngA* in self-aggregation, a comparison between isogenic mutants E9034A vs. E9034Δ*IngA::km* and 73332 vs. 73332Δ*IngA::km* was made.

BIOFILM FORMATION

ETEC clinical isolates *IngA* positive were quantitatively analyzed according to the protocol described by Saldaña et al. (2009). In brief, 24-well plates containing 1 ml of PPLO were inoculated with 10 µl (1.5 × 10⁸ bacteria/ml) of bacterial suspensions and incubated at 37°C for 24 h. Biofilms on the surface of the wells were slowly washed three times with phosphate buffer saline 1x (PBS) (pH 7.4) and fixed with 2% formalin at 4°C overnight. Wells with fixed biofilms were decanted, washed three times with PBS and stained with 1 ml of 1% crystal violet for 20 min. The excess of crystal violet was removed and the plates washed with water twice. Subsequently, 1 ml of 70% methanol was added for biofilm quantification by measuring optical density at 600 nm. Assays were made in Triplicates and repeated three consecutive

times. Enteroaggregative *E. coli* 042 and ETEC *IngA*⁻ clinical isolate 114246 were used as positive and negative controls, respectively. ETEC *IngA*⁻ absorbance was used to classify the biofilm formation into four categories as described by Saxena et al. (2014): non-biofilm, weakly, moderately, or strongly biofilm formers based upon the ODs of bacterial films (Saxena et al., 2014). The cut-off OD_{600 nm} for a 24 wells-plates is defined as two standard deviations above the mean OD_{600 nm} of the negative control. To establish the role of *IngA* in biofilm formation, a comparison between isogenic mutants E9034A vs. E9034Δ*IngA::km* and 73332 vs. 73332Δ*IngA::km* was performed.

BACTERIAL ADHERENCE TO HT-29 CELLS

Colon adenocarcinoma HT-29 cells (ATCC HTB-38) were employed for the adherence assays, as described by Mazariego-Espinosa et al. (2010). Clinical ETEC *IngA* positive isolates were grown in PPLO overnight at 37°C. Cell monolayers contained in 24-well polystyrene plates (Corning, NY, USA), were culture in Dulbecco's minimal Eagle medium (DMEM) (Invitrogen, Carlsbad, CA) supplemented with 10% FBS (Fetal Bovine Serum) (Gibco, USA) at 37°C under 5% CO₂ until reaching an 80% confluence. The monolayers were infected with 10 μl of each bacterial suspension to an OD_{600 nm} of 0.5 in the McFarland scale and incubated for 4 h at 37°C and 5% CO₂. Infected monolayers were washed three times with 1x PBS (pH 7.4) to remove unbound bacteria and incubated of 0.1% Triton-100x (Amresco Bioscience, Sweden) for 5 min; serial dilutions were plated on LB, incubated at 37°C for 24 h, and colony forming units (CFU) counted to determine CFU/ml. Adhesion assays were performed in triplicate and on three different days. Adherence values correspond to averages of triplicate assay and the corresponding standard deviation. ETEC isolates adherence was classified with cut-off CFU/ml for Enteroaggregative *E. coli* 042 is defined as two standard deviations above the mean of the positive control.

CONSTRUCTION OF AN *IngA* ISOGENIC MUTANT

The lambda red system was used to construct an isogenic mutant of the *IngA* gene in clinical isolate 73332 and the E9034A as described by Datsenko and Wanner (2000), Mazariego-Espinosa et al. (2010), and Mazariego-Espinosa et al. (2010). Briefly, the kanamycin resistance (*km*^r) cassette from plasmid pKD4 was amplified using the following primers *IngA* F (5' ttgagttacctgagcagctacaggtacttagaattttgtctgcagtcagggtgaggctggagctgct tc 3') and *IngA* R: (5' tgctatccgtgtataaccggacacagaaatttaagaagaggcaagaaacatgatgatctccttag 3'). The PCR-assay was performed in a 200 μl reaction mixture containing 10 mM Tris-HCl, 3 mM MgCl₂, 200 μM deoxynucleoside triphosphates (dNTPs), 2 U of DNA polymerase, 0.25 μM each primer, and 10 μl of template DNA. The PCR products obtained were electroporated into ETEC clinical isolate 73332 and E9034A, harboring the pKD46 plasmid, and transformants were selected for kanamycin resistance.

STATISTICAL ANALYSIS

The results showed are the mean of three experiments performed in triplicate on different days. Statistical analysis was done using the Student's *t*-test.

RESULTS

IDENTIFICATION OF CS21 AMONG ETEC CLINICAL ISOLATES FROM CHILDREN WITH DIARRHEA

A total of 303 clinical isolates previously identified as ETEC by serotyping, were obtained from Mexican and Bangladeshi children stool and screened by PCR for the following genes: *st*, *lt*, *IngA*, *cs1*, *cs3*, *cfa/I*, and *cofA*. Eighty one point fifty one percent (247/303) of the isolates amplified at least one gene (Table 2). The most frequent genotypes identified for the ETEC isolates were: *st*⁺/*lt*⁺ with 30.76% (76/247); *lt*⁺ 30.36% (75/247); *st*⁺ 11.33% (28/247); *lt*⁺/*IngA*⁺ 7.69% (19/247), and *st*⁺/*lt*⁺/*IngA*⁺ with 5.26% (13/247). Differences between geographical regions were observed: in Bangladeshi isolates the *st*⁺/*lt*⁺ genotype was the most prevalent with 37% (75/201); in contrast Mexican isolates showed a prevalence of *lt*⁺ genotype of 48% (22/46) (Table 2).

CS21 ETEC ISOLATES CARRY VARIABLE ENTEROTOXIGENIC VIRULENCE GENES

According to Isidean et al. (2011), CS21 is perhaps one of the most prevalent CF worldwide (Isidean et al., 2011). The present study was addressed toward ETEC isolates, which possess the *IngA* gene coding for the *IngA* protein, the structural subunit of Longus. The results showed that the *IngA* gene was amplified in 25.10% (62/247) of the isolates. The *IngA*⁺ gene was associated to *lt*⁺ 33.87% (21/62), *st*⁺/*lt*⁺ 33.87% (21/62), *st*⁺ 19.35% (12/62), *cfaI*⁺ 17.74% (11/62), *cs3*⁺ 6.45%

Table 2 | Genotype detection frequency of ETEC isolates from children with diarrhea.

Genes	Percentage (positive isolates)	Mexican percentage (positive isolates)	Bangladesh percentage (positive isolates)
<i>st</i> ⁺ / <i>lt</i> ⁺	30.76 (76)	2.17 (1)	37.31 (75)
<i>lt</i> ⁺	30.36 (75)	47.83 (22)	26.37 (53)
<i>st</i> ⁺	11.33 (28)	8.70 (4)	11.94 (24)
<i>lt</i> ⁺ / <i>IngA</i> ⁺	7.69 (19)	8.70 (4)	7.46 (15)
<i>st</i> ⁺ / <i>lt</i> ⁺ / <i>IngA</i> ⁺	5.26 (13)	2.17 (1)	5.97 (12)
<i>IngA</i> ⁺	2.83 (7)	10.87 (5)	0.99 (2)
<i>st</i> ⁺ / <i>IngA</i> ⁺	2.83 (7)	2.17 (1)	2.99 (6)
<i>st</i> ⁺ / <i>IngA</i> ⁺ / <i>cfaI</i> ⁺	2.42 (6)	4.35 (2)	1.99 (4)
<i>st</i> ⁺ / <i>lt</i> ⁺ / <i>IngA</i> ⁺ / <i>cfaI</i> ⁺	2.02 (5)	0	2.49 (5)
<i>st</i> ⁺ / <i>lt</i> ⁺ / <i>IngA</i> ⁺ / <i>cs3</i> ⁺	1.21 (3)	6.52 (3)	0
<i>st</i> ⁺ / <i>lt</i> ⁺ / <i>cs3</i> ⁺	0.8 (2)	2.17 (1)	0.50 (1)
<i>lt</i> ⁺ / <i>cs3</i> ⁺	0.8 (2)	0	0.99 (2)
<i>lt</i> ⁺ / <i>csI</i> ⁺	0.4 (1)	2.17 (1)	0
<i>st</i> ⁺ / <i>csI</i> ⁺	0.4 (1)	0	0.50 (1)
<i>st</i> ⁺ / <i>lt</i> ⁺ / <i>IngA</i> ⁺ / <i>csI</i> ⁺ / <i>cs3</i> ⁺	0.4 (1)	2.17 (1)	0
<i>lt</i> ⁺ / <i>IngA</i> ⁺ / <i>cofA</i> ⁺	0.4 (1)	0	0.50 (1)
Total	100 (247)	99.99 (46)	100 (201)

st, gene coding for heat-stable enterotoxin; *lt*, gene coding for heat-labile enterotoxin; *IngA*, gene coding for the major subunit of longus; *cfaI*, gene coding for the colonization factor I; *cooA*, gene coding for the major subunit of CS1; *cs3*, gene coding for the CS3; *cofA*, gene coding for the major subunit of CFA/III.

(4/62), *cs1*⁺ 1.61% (1/62), and *cofA*⁺ 1.61% (1/62) genes. Association of *lngA* with *lt*⁺ Bangladeshi isolates was 36% (15/42) and, just the *lngA*⁺ genotype alone with a 25% (5/20) was found in Mexican isolates. Expression of *LngA* in *lngA*⁺ ETEC isolates was determined by immunoblot. Fifty percent (31/62) of the isolates analyzed produced *LngA*, protein production distribution was similar in both geographic regions: Mexican 65% (13/20) and Bangladesh 42.9% (18/42). These data suggest that even though the isolates encode the *LngA* protein, specific regulatory signals may be necessary to trigger *LngA* expression in the remaining 50% of the clinical isolates.

SEROTYPING OF ETEC ISOLATES

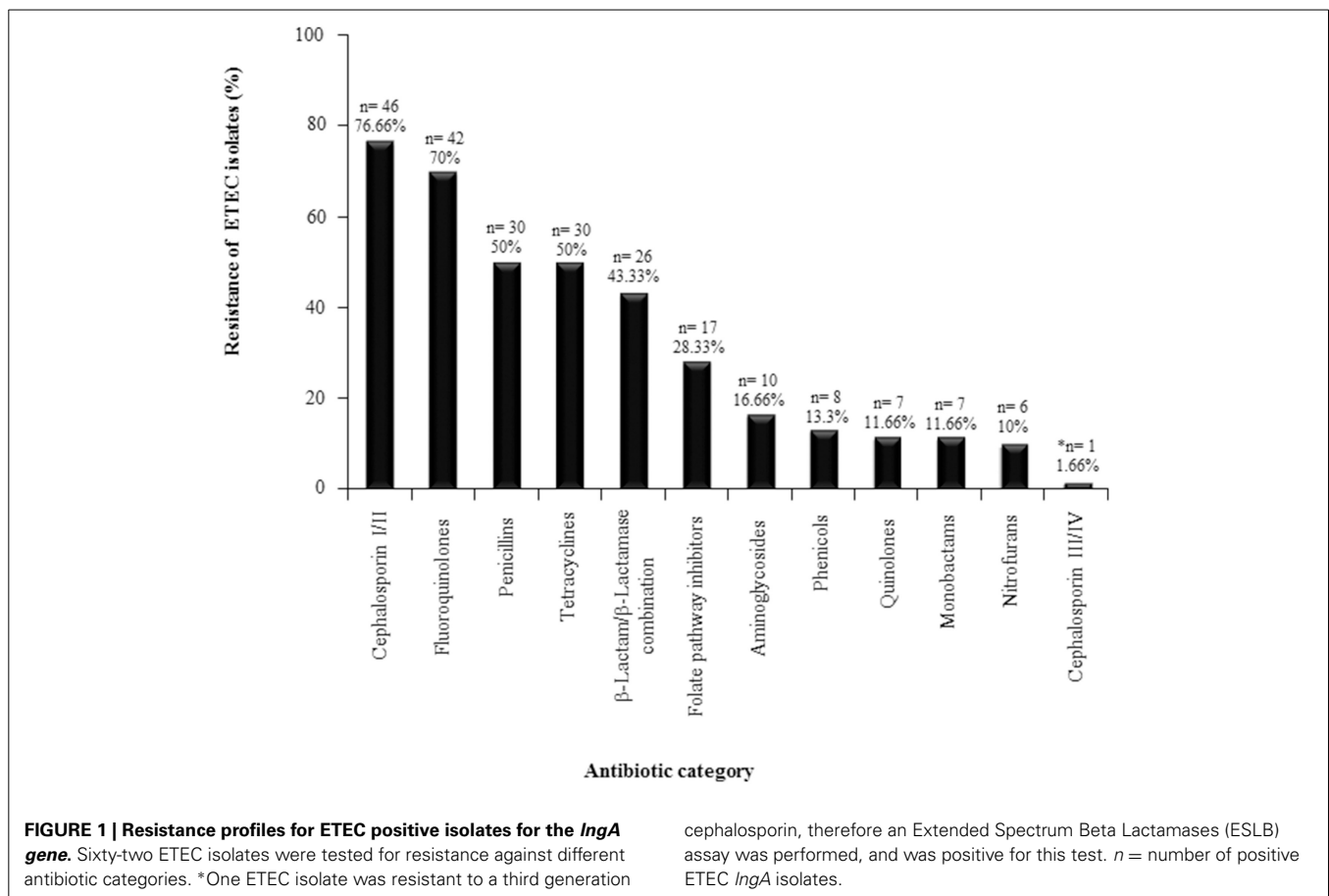
Thirty-one serotypes were identified in *lngA*⁺ ETEC isolates (Table S1). Three main serotypes prevailed: O128ac:H12 in 19.35% (12/62), O78:H12 and O6:H16 with 11.29% (7/62). These serotypes showed a strong association between the *lngA*⁺ gene with the *lt*⁺ and *st*^{+/lt}⁺ genes (Table S1).

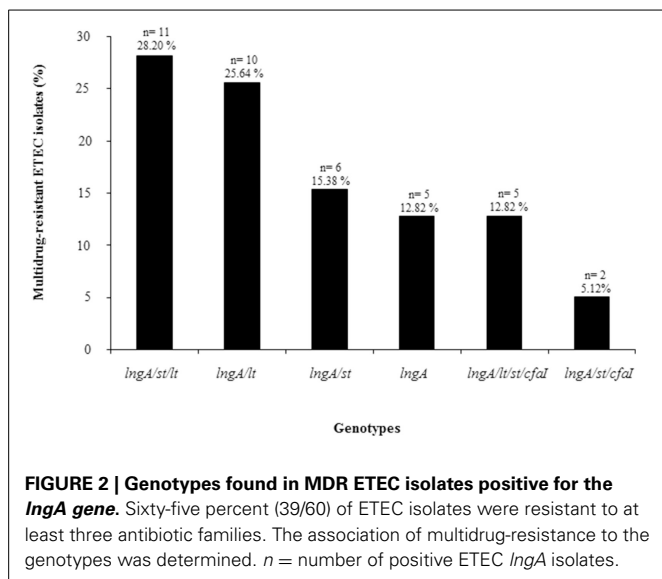
ETEC ISOLATES WERE MULTIDRUG-RESISTANT (MDR)

Multidrug-resistance defined as antibiotic resistance to at least three antibiotic categories is widespread among foodborne and waterborne enteric pathogens including ETEC. While ETEC treatment is not based on antibiotics in the majority of cases,

evaluating antibiotic susceptibility is important in cases where antibiotic use is clinically indicated. In this study, 62 *lngA*⁺ ETEC isolates were tested for resistance to different antibiotic categories. Interestingly, 96.77% (60/62) of the isolates were resistant to at least one antibiotic category. Resistance to first and second generation cephalosporins was observed in 76.66% (46/60) of the isolates; to fluoroquinolones in 70% (42/60); to penicillin and tetracyclines in 50% (30/60); to β -Lactam/ β -lactamase inhibitor combination in 43.33% (26/60); and to folate pathway inhibitors 28.33% (17/60). ETEC antibiotic resistance to the remaining antibiotic categories was less than 17% (Figure 1). Only one ETEC isolate was resistant to a third generation cephalosporin; therefore, an Extended Spectrum Beta Lactamases (ESBL) assay was performed, and resulted positive for this test (Table S1).

Multidrug-resistance was identified among 65% (39/60) of the *lngA*⁺ ETEC isolates, based on geographic regions, distribution of MDR clinical isolates were from Bangladesh with 76.2% (32/42), compared with Mexican isolates with 35% (7/20) (Table S1). The association of multidrug-resistance to some genotypes was determined after observing the following distribution: 28.20% (11/39) to *lngA*⁺/*st*^{+/lt}⁺; 25.64% (10/39) to *lngA*⁺/*lt*⁺; 15.38% (6/39) to *lngA*⁺/*st*⁺; 12.82% (5/39) to *lngA*⁺; 12.82% (5/39) to *lngA*⁺/*lt*⁺/*st*^{+/cfaI}⁺; and 5.12% (2/39) to *lngA*⁺/*st*^{+/cfaI}⁺ genotypes, respectively (Figure 2).





CS21 POSITIVE ETEC ISOLATES SELF-AGGREGATE

Self-aggregation was visualized in 48.38% (30/62) of the ETEC isolates analyzed. Self-aggregation distribution in the geographic regions was as follows: Mexican 60% (12/20) and Bangladesh 42.9% (18/42). This phenotype was classified as weak (+), moderate (++), and strong (+++) according to the levels of aggregation observed in ETEC strain E9034A, used as positive control. **Figure 3** showed the self-aggregation patterns of E9034A (weak), 115340 (moderate), and 45162 (strong) ETEC isolates. Accordingly, 23.33% (7/30), 46.66% (14/30), and 30% (9/30) of the ETEC isolates showed a weakly, moderately, and strongly self-aggregation, respectively (**Figure 3**). Interestingly, 76.66% (23/30) of ETEC isolates expressing LngA exhibited self-aggregation, suggesting that CS21 among ETEC clinical isolates is strongly associated to this phenotype. The remaining clinical isolates 21.87% (7/32) were unable to express LngA and showed a weak or moderate self-aggregation phenotype, suggesting that CS21 contribute to the aggregate formation. To corroborate the role of *lngA* in self-aggregation, a comparison between isogenic mutants E9034A vs. E9034Δ*lngA::km* and 73332 vs. 73332Δ*lngA::km* was made, in which the loss of the phenotype was observed in both mutants (Figure S1).

CS21 ETEC ISOLATES FORM BIOFILM

Biofilm formation is a critical event for bacterial pathogenesis. ETEC *lngA*⁻ clinical isolate 114246 was used as negative control and its absorbance was used to classify as weakly biofilm former (0.129–0.258); moderately biofilm former (0.259–0.516); and strongly biofilm former (≥ 0.517) (**Table 3**). Among *lngA*⁺ ETEC, 83.87% (52/62) formed biofilm, grouping them as: 40.38% (21/52) weakly biofilm formers; 46.15% (24/52) moderately biofilm formers, and 13.46% (7/52) strongly biofilm formers (**Figure 4**). Interestingly, 48.07% (25/52) of ETEC clinical isolates producing the LngA protein detected by immunoblot were associated to different levels in biofilm formation. Of those, 32% (8/25), 48% (12/25), and 20% (5/25) ETEC isolates were weakly, moderately, and strongly biofilm formers,

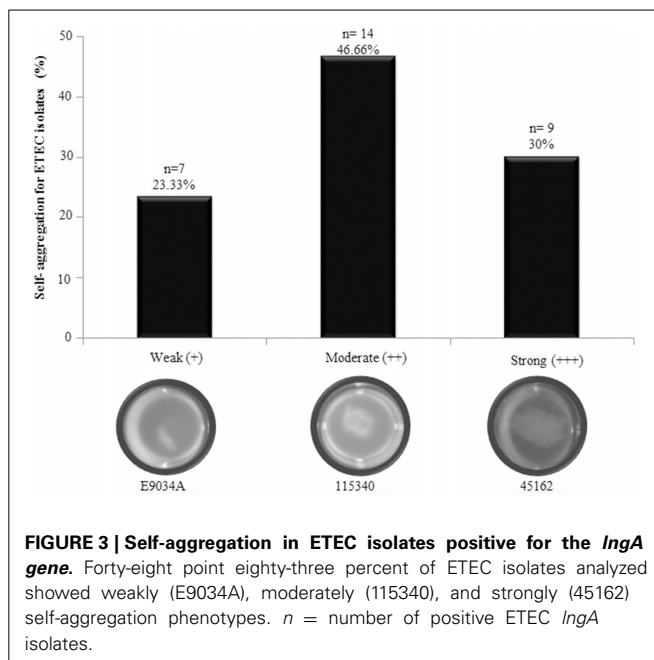


Table 3 | Classification of *lngA*⁺ ETEC isolates for biofilm formation.

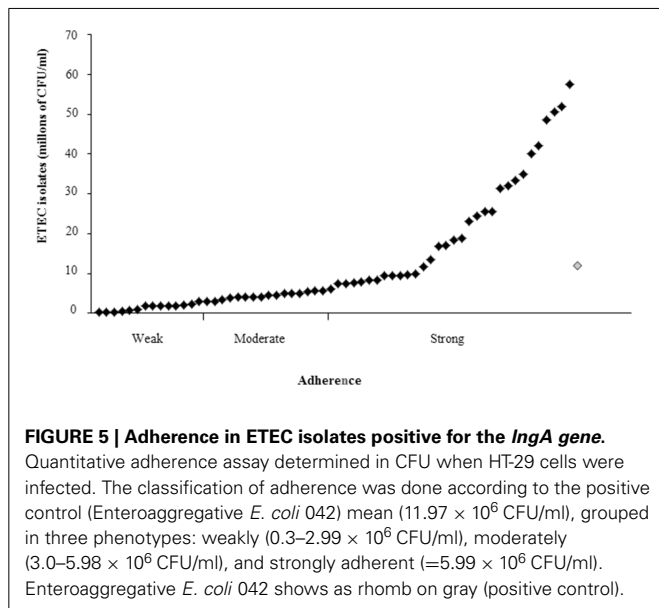
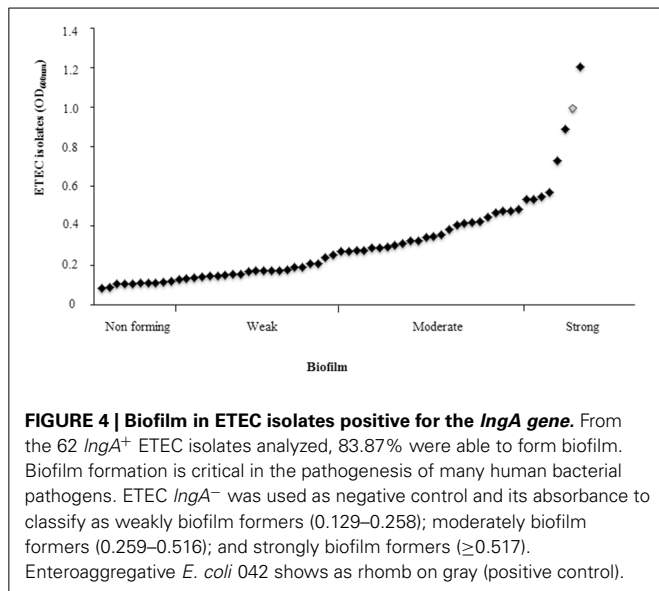
	<i>OD</i> _{600 nm} value	Biofilm forming
$OD \leq OD_c$	<0.129	Non
$OD_c < OD \leq 2 OD_c$	0.129–0.258	Weak
$2 \times OD_c < OD \leq 4 \times OD_c$	0.259–0.516	Moderate
$4 \times OD_c < OD$	≥ 0.517	Strong

OD_c, Optical density cut off value: average OD of negative control.

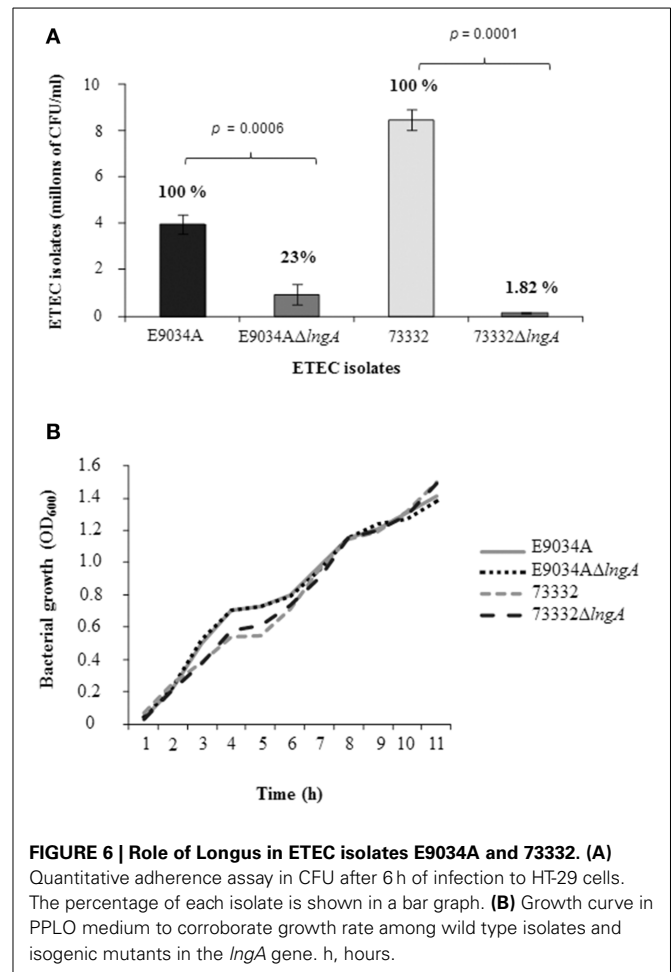
respectively. To corroborate the role of *lngA* in biofilm formation, a comparison between isogenic mutants E9034A (0.1824 *OD*_{600 nm}) vs. E9034Δ*lngA::km* (0.1214 *OD*_{600 nm}) and 73332 (0.1459 *OD*_{600 nm}) vs. 73332Δ*lngA::km* (0.1262 *OD*_{600 nm}) was made, where the statistical analysis showed no statistical significance ($p = 0.27$) and ($p = 0.46$), respectively.

ETEC ISOLATES WERE ABLE TO STRONGLY ADHERE TO HT-29 CELLS

The HT-29 cells were infected with ETEC isolates positive for the *lngA* gene and the results showed different adherence values after the quantification. HT-29 cell adherence distribution was similar in both geographic regions. The adherence levels ranged from 0.3×10^6 to 57.5×10^6 CFU/ml (**Figure 5**). The classification of adherence was done according to the positive control (Enteroaggregative *E. coli* 042) mean (11.97×10^6 CFU/ml), grouped in three phenotypes: weakly (0.3 – 2.99×10^6 CFU/ml), moderately (3.0 – 5.98×10^6 CFU/ml), and strongly adherent ($\geq 5.99 \times 10^6$ CFU/ml). Briefly, 22.58% (14/62), 25.80% (16/62), and 51.62% (32/62) of *lngA*⁺ ETEC isolates were weak, moderate, and strong adherent. The relationship between LngA expression and adherence was as follows: 29.03% (9/31) for weak and moderate adherent, and 41.94% (13/31) for strong adherent.



Clinical isolate 73332 serotype O6:H6 carrying *lt*⁺, *st*⁺, *cs3*⁺, and *lngA*⁺ was selected to study the role of CS21 in the adherence to HT-29 cells. This isolate expresses *LngA* and was classified as highly adherent (8.48×10^6 CFU/ml). The adherence levels were compared among ETEC wild-type (E9034A and 73332) and isogenic mutants (E9034A Δ *lngA* and 73332 Δ *lngA*). Reduction of 98% (1.55×10^5 CFU/ml) cell adherence with ETEC 73332 isolate (8.48×10^6 CFU/ml) compared to 77% (9.33×10^5 CFU/ml) reduction in ETEC E9034A strain (3.97×10^6 CFU/ml) demonstrates that the reduction in the adherence was attributed to the *lngA* gene mutation (Figure 6A). A statistical difference in cell adherence reduction was observed when the *lngA* gene was deleted in ETEC clinical isolate 73332 and ETEC E9034A, with *p*-values of 0.0001 and 0.0006, respectively.



A growth kinetic curve with wild type ETEC and *lngA* mutants was performed in PPLO media (Figure 6B) showing that there was no difference in growth among the wild-type and isogenic mutants ETEC isolates.

DISCUSSION

ETEC isolates are a frequent cause of diarrhea worldwide; ETEC causes nearly 400 million diarrhea episodes every year affecting children from developing countries as well as travelers (Wenneras and Erling, 2004; Riddle et al., 2006). ETEC colonizes the small bowel epithelium through the expression of fimbrial and non-fimbrial CFs and producing LT and/or ST enterotoxins causing massive excretion of water and electrolytes (Qadri et al., 2005; Turner et al., 2006; Guttman and Finlay, 2008; Croxen and Finlay, 2010). After screening CFs in ETEC isolates from different geographical origins, CFA/I and Longus are the most prevalent CFs (Giron et al., 1994; Girón et al., 1995a; Gutierrez-Cazarez et al., 2000; Pichel et al., 2002). A strategy to control enterotoxigenic *E. coli* infections has been the development of a vaccine that triggers protective immunity against the most prevalent CFs (Giron et al., 1994; Girón et al., 1995a; Gutierrez-Cazarez et al., 2000; Pichel et al., 2002). In this study we analyzed a total of 303 clinical isolates obtained from Mexican and Bangladeshi children stools, classified as ETEC by serotyping. The *st*⁺/*lt*⁺ (30.76%) was the

most frequent genotype amplified for ETEC isolates analyzed in this study and the frequency was higher than previously reported by Guerra et al. (2014), in which only 15% of ETEC isolates from Colombia were positive for both enterotoxin genes (Guerra et al., 2014). However, a systematic review by Isidean et al. (2011), reported a global distribution of *st*⁺/*lt*⁺ of 33% in ETEC clinical isolates, similar to the findings in this study. The frequency of the *lngA* gene among ETEC clinical isolates in several countries may vary between 8.5 and 50% (Girón et al., 1995a; Gutierrez-Cazarez et al., 2000; Nishimura et al., 2002; Pichel et al., 2002; Guerra et al., 2014). Meanwhile, in this study the *lngA* gene was amplified in 25.10% of the isolates. A lower frequency of the *lngA*⁺ gene associated to *st*⁺, *st*⁺/*lt*⁺, and *lt*⁺ genes was observed when compared with data described by Nishimura et al. (2002); which reported a frequency of 56%, 31%, and 10%, respectively (Nishimura et al., 2002). Likewise, *lngA*-positive ETEC isolates (25.10%) also harbor genes coding for other CFs. In other reports, *lngA*⁺ was most highly associated to isolates coding for CFA/II (Girón et al., 1995b; Nishimura et al., 2002) and CFA/I that has been described as the most prevalent CF in developing countries. However, in this study it appeared in low frequency 6.45%, for *cfbII* (*csI*, *cs3*), and 17.74% for *cfbI* (Pichel et al., 2002; Qadri et al., 2005). Interestingly, *lngA*⁺ isolates were found alone in 11.29% of the ETEC isolates. This result suggest that these isolates could have lost the plasmid that code for enterotoxins, leaving only the megaplasmid coding for *lngA*⁺; therefore, *lngA* is proposed as a molecular marker together with *lt* and *st* genes, for identification of the ETEC pathotype. The most prevalent serotypes found in ETEC *lngA*⁺ isolates were O128ac:H12, O78:H12, and O6:H6. The *lngA* gene was also described in a considerable number of ETEC serotypes (O2:NT, O6:NM, O6:H2, O6:H16, O6:H48, O8:H2, O8:H6, O8:H21, O8:H29, O20:H2, O25:NM, O25:H42, O49:NM, O128:NM, O139:H28, O148:H28, O153:H2, O153:H45, O158:H10, and O159:H4) (Girón et al., 1995a). These data provide further evidence of the wide distribution of CS21 among ETEC clinical isolates.

The frequency up to date reported for MDR ETEC isolates is 46.4% (Zeighami et al., 2014). In this study, 65% of ETEC isolates positive for *lngA*, were multidrug-resistant to different antibiotic categories. In particular, resistance to cephalosporins I and II (76.66%), fluoroquinolones (70%), penicillins and tetracycline (50%), β -lactam/ β -lactamase combination (43.33%), folate pathway inhibitors (28.33%); and for the rest of the antibiotic categories (aminoglycosides, phenicols, quinolones, nitrofurans, monobactams, and cephalosporin III and IV) we found less than 17% of resistance. In contrast to the reports from other studies in which ampicillin, trimethoprim-sulfamethoxazole, cefazolin, and amoxicillin-clavulanate resistance was detected among 67.5%, 50%, 15%, and 5% isolates, respectively. In addition, no resistance to ceftriaxone, ceftazidime, cefepime, ciprofloxacin, and piperacillin-tazobactam was detected among ETEC isolates (Guerra et al., 2014). In other study, 47% of ETEC isolates were resistant to ampicillin, a third (37%) of the isolates was resistant to trimethoprim-sulfamethoxazole, and 24% of the isolates were tetracycline-resistant (El-Gendy et al., 2013).

Although the treatment of diarrhea caused by ETEC with antibiotic agents is generally restricted to severe and moderate

hospitalized cases or immunocompromised patients, there is an increased number of clinical isolates of MDR ETEC, which could be a potential public health problem; for this reason, the study of antibiotic resistance should be carried out in combination with the frequencies of CFs and enterotoxins. We found that 46.15% of MDR ETEC isolates was able to produce the *LngA* protein. In this study ETEC isolates were MDR and producers of the *LngA* protein providing the bacteria colonization and pathogenic advantages. CS21, a type IV pilus and a common CF in ETEC, induce microcolony formation, twitching motility, adherence to cells, and self-aggregation (Clavijo et al., 2010; Mazariego-Espinosa et al., 2010). Self-aggregation has been associated to bacterial protection against antibiotics in a similar manner as bacterial biofilms avoid contact with lumen intestinal bactericidal agents such as lactoferrin, lysozyme, and secretory antibodies. In this study, 48.38% of the ETEC isolates were able to self-aggregate and 76.66% of the ETEC isolates produced the *LngA* protein.

Biofilms are matrix-enclosed communities of bacteria that show increased antibiotic resistance and the ability to evade the immune system (MacFarlane and Dillon, 2007). They can cause recalcitrant infections which cannot be cured with classical antibiotic therapy (Frier et al., 2014). Biofilm production and adherence mechanisms of ETEC isolates are associated with CFs and multiple genes coding for CFs located on a plasmid (CFA/I and CFA/II), as well as on its chromosome (CS2). In our study, 83.87% of the ETEC isolates tested were able to form biofilm at 24 h. Liaqat and Sakellaris (2012), described that all ETEC isolates analyzed were able to form biofilm after 60 h (Liaqat and Sakellaris, 2012). Outstandingly, most of the ETEC isolates were weak and moderate biofilm producers and a relationship was not found.

Many pathogenic bacteria have the ability to adhere to specific host tissues prior to colonization. Thus, bacterial adhesion is the initial step that precedes colonization and in some cases invasion of the epithelial surfaces (Pizarro-Cerda and Cossart, 2006). All ETEC isolates were able to adhere to HT-29 cells, and the isolates that produce the *LngA* protein were associated with high adherence values. Mazariego-Espinosa et al. (2010), demonstrated that CS21 produced by ETEC E9034A is involved in the phenomenon of adherence as a requirement for mediating the interaction between the bacteria and cultured epithelial cells, specifically intestinal cells (Mazariego-Espinosa et al., 2010). In this study, we demonstrate the role of CS21 for adhesion, using the E9034A prototype strain that has been widely used in other studies and a clinical isolate 73332. The aim was to analyze through quantitative studies the potential of CS21 as an adhesin in clinical isolates of ETEC during the colonization to their host.

The growth curve kinetic was performed to guarantee that the adherence phenotype was not affected by bacterial growth. Isogenic mutants in the *lngA* gene showed a reduction of adherence from 77% to 98.8%. However, Mazariego-Espinosa et al. (2010) showed a reduction of 32% in adherence levels in the E9034A Δ *lngA* when compared to ETEC E9034A (Mazariego-Espinosa et al., 2010). The difference observed in adherence levels in both mutants, could be attributed to the generation of a second mutant in the *lngA* gene employed in this study from a highly adherent colony of the E9034A wild-type strain; which,

was previously selected from multiple subcultures in HT-29 cells in different days.

Antibiotic categories such as sulfonamides and beta-lactam containing compounds can be effective prophylaxis against traveler's diarrhea by deficiency of effective vaccine treatments (Brown et al., 2009) this can be explained by the different geographical origins of the Mexican and Bangladeshi isolates. However, the emergence of antibiotic resistant isolates and the genes encoding resistance proteins are known to reside on horizontal transmissible elements such as conjugative plasmids and integrons promoting the spread of resistance (Levy, 2002). The MDR ETEC isolates able to produce LngA showed features associated with self-aggregation, and adherence to intestinal cell lines. These features are important in the colonization process. Taking into consideration the frequencies of occurrence of the widely spread CFs such as CFA/I and CS21 in other geographic regions, the frequencies of occurrence of the *lngA* pilin gene found here or described by others confirm the importance of the CS21 pilus as a widely distributed antigen among ETEC isolates.

AUTHOR CONTRIBUTIONS

Designed and conceived the experiments: Ariadna Cruz-Córdova, Karina Espinosa-Mazariego, Sara A. Ochoa, Zeus Saldaña. Performed the experiments: Ariadna Cruz-Córdova, Karina Espinosa-Mazariego, Zeus Saldaña, Gerardo E. Rodea, Viridiana Rodríguez-Ramírez. Analyzed the data: Ariadna Cruz-Córdova, Karina Espinosa-Mazariego, Zeus Saldaña, Gerardo E. Rodea. Contributed reagents/materials/analysis tools: Sara A. Ochoa, Carlos A. Eslava-Campos, Firdausi Qadri, Armando Navarro-Ocaña, José Arrellano-Galindo. Reviewed the manuscript: Firdausi Qadri, Rigoberto Hernández-Castro, Oscar G. Gómez-Duarte. Wrote the manuscript Juan Xicohtencatl-Cortes.

ETHICS STATEMENT

The study was reviewed and approved by the Research (Dr. Onofre Muñoz Hernández), Ethics (Dr. Amparo Faure Fontenla), and Biosecurity (Dr. Herlinda Vera Hermosillo) Committee of Hospital Infantil de México Federico Gómez under permit numbers HIM/2014/016, HIM/2013/008, and HIM/2011/017.

ACKNOWLEDGMENTS

We thank Gerardo Escalona Venegas for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2014.00709/abstract>

REFERENCES

- Ahren, C. M., and Svennerholm, A. M. (1985). Experimental enterotoxin-induced *Escherichia coli* diarrhea and protection induced by previous infection with bacteria of the same adhesin or enterotoxin type. *Infect. Immun.* 50, 255–261.
- Beachey, E. H. (1981). Bacterial adherence: adhesin-receptor interactions mediating the attachment of bacteria to mucosal surface. *J. Infect. Dis.* 143, 325–345. doi: 10.1093/infdis/143.3.325
- Brinton, C. C. Jr. (1965). The structure, function, synthesis and genetic control of bacterial pili and a molecular model for DNA and RNA transport in gram negative bacteria. *Trans. N.Y. Acad. Sci.* 27, 1003–1054. doi: 10.1111/j.2164-0947.1965.tb02342.x
- Brown, J. A., Riddle, M. S., Putnam, S. D., Schlett, C. D., Armstrong, A. W., Jones, J. J., et al. (2009). Outcomes of diarrhea management in operations Iraqi Freedom and Enduring Freedom. *Travel Med. Infect. Dis.* 7, 337–343. doi: 10.1016/j.tmaid.2009.09.001
- Clavijo, A. P., Bai, J., and Gomez-Duarte, O. G. (2010). The Longus type IV pilus of enterotoxigenic *Escherichia coli* (ETEC) mediates bacterial self-aggregation and protection from antimicrobial agents. *Microb. Pathog.* 48, 230–238. doi: 10.1016/j.micpath.2010.03.006
- Craig, L., Pique, M. E., and Tainer, J. A. (2004). Type IV pilus structure and bacterial pathogenicity. *Nat. Rev. Microbiol.* 2, 363–378. doi: 10.1038/nrmicro885
- Croxen, M. A., and Finlay, B. B. (2010). Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat. Rev. Microbiol.* 8, 26–38. doi: 10.1038/nrmicro2265
- Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645. doi: 10.1073/pnas.120163297
- El-Gendy, A. M., Mansour, A., Shaheen, H. I., Monteville, M. R., Armstrong, A. W., El-Sayed, N., et al. (2013). Genotypic characterization of Egypt enterotoxigenic *Escherichia coli* isolates expressing coli surface antigen 6. *J. Infect. Dev. Ctries.* 7, 90–100. doi: 10.3855/jidc.2454
- Foier, K., Raemdonck, K., De Smedt, S. C., Demeester, J., Coenye, T., and Braeckmans, K. (2014). Lipid and polymer nanoparticles for drug delivery to bacterial biofilms. *J. Control. Release* 190, 607–623. doi: 10.1016/j.jconrel.2014.03.055
- Gastra, W., and Svennerholm, A. M. (1996). Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends Microbiol.* 4, 444–452. doi: 10.1016/0966-842X(96)10068-8
- Girón, J. A., Gomez-Duarte, O. G., Jarvis, K. G., and Kaper, J. B. (1997). Longus pilus of enterotoxigenic *Escherichia coli* and its relatedness to other type-4 pili—a minireview. *Gene* 192, 39–43. doi: 10.1016/S0378-1119(97)00039-5
- Girón, J. A., Ho, A. S., and Schoolnik, G. K. (1991). An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. *Science* 254, 710–713. doi: 10.1126/science.1683004
- Giron, J. A., Levine, M. M., and Kaper, J. B. (1994). Longus: a long pilus ultrastructure produced by human enterotoxigenic *Escherichia coli*. *Mol. Microbiol.* 12, 71–82. doi: 10.1111/j.1365-2958.1994.tb00996.x
- Girón, J. A., Viboud, G. I., Sperandio, V., Gomez-Duarte, O. G., Maneval, D. R., Albert, M. J., et al. (1995a). Prevalence and association of the longus pilus structural gene (*lngA*) with colonization factor antigens, enterotoxin types, and serotypes of enterotoxigenic *Escherichia coli*. *Infect. Immun.* 63, 4195–4198.
- Girón, J. A., Xu, J. G., Gonzalez, C. R., Hone, D., Kaper, J. B., and Levine, M. M. (1995b). Simultaneous expression of CFA/I and CS3 colonization factor antigens of enterotoxigenic *Escherichia coli* by delta *aroC*, delta *aroD* *Salmonella typhi* vaccine strain CVD 908. *Vaccine* 13, 939–946. doi: 10.1016/0264-410X(95)00003-J
- Gomez-Duarte, O. G., Chattopadhyay, S., Weissman, S. J., Giron, J. A., Kaper, J. B., and Sokurenko, E. V. (2007). Genetic diversity of the gene cluster encoding longus, a type IV pilus of enterotoxigenic *Escherichia coli*. *J. Bacteriol.* 189, 9145–9149. doi: 10.1128/JB.00722-07
- Guerra, J. A., Romero-Herazo, Y. C., Arzuza, O., and Gomez-Duarte, O. G. (2014). Phenotypic and genotypic characterization of enterotoxigenic *Escherichia coli* clinical isolates from Northern Colombia, South America. *Biomed Res. Int.* 2014:236260. doi: 10.1155/2014/236260
- Guevara, C. P., Luiz, W. B., Sierra, A., Cruz, C., Qadri, F., Kaushik, R. S., et al. (2013). Enterotoxigenic *Escherichia coli* (ETEC) CS21 pilus contributes to adhesion to intestinal cells and to pathogenesis under *in vivo* conditions. *Microbiology* 159(Pt 8), 1725–1735. doi: 10.1099/mic.0.065532-0
- Gutierrez-Cazarez, Z., Qadri, F., Albert, M. J., and Giron, J. A. (2000). Identification of enterotoxigenic *Escherichia coli* harboring longus type IV pilus gene by DNA amplification. *J. Clin. Microbiol.* 38, 1767–1771.
- Guttman, J. A., and Finlay, B. B. (2008). Subcellular alterations that lead to diarrhea during bacterial pathogenesis. *Trends Microbiol.* 16, 535–542. doi: 10.1016/j.tim.2008.08.004
- Honda, T., Arita, M., and Miwatani, T. (1984). Characterization of new hydrophobic pili of human enterotoxigenic *Escherichia coli*: a possible new colonization factor. *Infect. Immun.* 43, 959–965.
- Honda, T., Miwatani, T., Mizuno, S., and Maki, S. (1985). Demonstration of antibodies reacting with *Escherichia coli* heat-labile enterotoxin in sera of patients suffering from *Campylobacter jejuni*/coli. *N.Z. Med. J.* 98, 965.

- Isidean, S. D., Riddle, M. S., Savarino, S. J., and Porter, C. K. (2011). A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression. *Vaccine* 29, 6167–6178. doi: 10.1016/j.vaccine.2011.06.084
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0
- Levy, S. B. (2002). The 2000 Garrod lecture. Factors impacting on the problem of antibiotic resistance. *J. Antimicrob. Chemother.* 49, 25–30. doi: 10.1093/jac/49.1.25
- Liaquat, L., and Sakellaris, H. (2012). Biofilm formation and binding specificities of CFA/I, CFA/II and CS2 adhesions of enterotoxigenic *Escherichia coli* and Cfae-R181A mutant. *Braz. J. Microbiol.* 43, 969–980. doi: 10.1590/S1517-83822012000300018
- MacFarlane, S., and Dillon, J. F. (2007). Microbial biofilms in the human gastrointestinal tract. *J. Appl. Microbiol.* 102, 1187–1196. doi: 10.1111/j.1365-2672.2007.03287.x
- Mazariego-Espinosa, K., Cruz, A., Ledesma, M. A., Ochoa, S. A., and Xicohtencatl-Cortes, J. (2010). Longus, a type IV pilus of enterotoxigenic *Escherichia coli*, is involved in adherence to intestinal epithelial cells. *J. Bacteriol.* 192, 2791–2800. doi: 10.1128/JB.01595-09
- Nataro, J. P., and Kaper, J. B. (1998). Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* 11, 142–201.
- Nishimura, L. S., Giron, J. A., Nunes, S. L., and Guth, B. E. (2002). Prevalence of enterotoxigenic *Escherichia coli* strains harboring the longus pilus gene in Brazil. *J. Clin. Microbiol.* 40, 2606–2608. doi: 10.1128/JCM.40.7.2606-2608.2002
- Perez, C., Gomez-Duarte, O. G., and Arias, M. L. (2010). Diarrheagenic *Escherichia coli* in children from Costa Rica. *Am. J. Trop. Med. Hyg.* 83, 292–297. doi: 10.4269/ajtmh.2010.09-0412
- Pichel, M. G., Binsztejn, N., Qadri, F., and Giron, J. A. (2002). Type IV longus pilus of enterotoxigenic *Escherichia coli*: occurrence and association with toxin types and colonization factors among strains isolated in Argentina. *J. Clin. Microbiol.* 40, 694–697. doi: 10.1128/JCM.40.2.694-697.2002
- Pizarro-Cerda, J., and Cossart, P. (2006). Bacterial adhesion and entry into host cells. *Cell* 124, 715–727. doi: 10.1016/j.cell.2006.02.012
- Qadri, F., Svennerholm, A. M., Faruque, A. S., and Sack, R. B. (2005). Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clin. Microbiol. Rev.* 18, 465–483. doi: 10.1128/CMR.18.3.465-483.2005
- Riddle, M. S., Sanders, J. W., Putnam, S. D., and Tribble, D. R. (2006). Incidence, etiology, and impact of diarrhea among long-term travelers (US military and similar populations): a systematic review. *Am. J. Trop. Med. Hyg.* 74, 891–900.
- Saldaña, Z., Xicohtencatl-Cortes, J., Avelino, F., Phillips, A. D., Kaper, J. B., Puente, J. L., et al. (2009). Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli. *Environ. Microbiol.* 11, 992–1006. doi: 10.1111/j.1462-2920.2008.01824.x
- Saxena, S., Banerjee, G., Garg, R., and Singh, M. (2014). Comparative study of biofilm formation in *Pseudomonas aeruginosa* isolates from patients of lower respiratory tract infection. *J. Clin. Diagn. Res.* 8, DC09–DC11. doi: 10.7860/JCDR/2014/7808.4330
- Taniguchi, T., Akeda, Y., Haba, A., Yasuda, Y., Yamamoto, K., Honda, T., et al. (2001). Gene cluster for assembly of pilus colonization factor antigen III of enterotoxigenic *Escherichia coli*. *Infect. Immun.* 69, 5864–5873. doi: 10.1128/IAI.69.9.5864-5873.2001
- Taylor, R. K., Miller, V. L., Furlong, D. B., and Mekalanos, J. J. (1987). Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci. U.S.A.* 84, 2833–2837. doi: 10.1073/pnas.84.9.2833
- Turner, S. M., Scott-Tucker, A., Cooper, L. M., and Henderson, I. R. (2006). Weapons of mass destruction: virulence factors of the global killer enterotoxigenic *Escherichia coli*. *FEMS Microbiol. Lett.* 263, 10–20. doi: 10.1111/j.1574-6968.2006.00401.x
- Wenneras, C., and Erling, V. (2004). Prevalence of enterotoxigenic *Escherichia coli*-associated diarrhoea and carrier state in the developing world. *J. Health Popul. Nutr.* 22, 370–382.
- WHO. (2006). Future directions for research on enterotoxigenic *Escherichia coli* vaccines for developing countries. *Wkly. Epidemiol. Rec.* 81, 97–104. Available online at: www.who.int/wer/2006/wer/811/eng
- Xicohtencatl-Cortes, J., Monteiro-Neto, V., Ledesma, M. A., Jordan, D. M., Francetic, O., Kaper, J. B., et al. (2007). Intestinal adherence associated with type IV pili of enterohemorrhagic *Escherichia coli* O157:H7. *J. Clin. Invest.* 117, 3519–3529. doi: 10.1172/JCI30727
- Zeighami, H., Haghi, F., Hajiahmadi, F., Kashefieh, M., and Memariani, M. (2014). Multi-drug-resistant enterotoxigenic and enterohemorrhagic *Escherichia coli* isolated from children with diarrhea. *J. Chemother.* doi: 10.1179/1973947813Y.0000000161. [Epub ahead of print].

Conflict of Interest Statement: This work was supported by grant number CONACYT 133451 and public Federal Funds HIM/2014/016, HIM/2013/008, and HIM/2011/017 from Hospital Infantil de México Federico Gómez. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 October 2014; accepted: 28 November 2014; published online: 17 December 2014.

Citation: Cruz-Córdova A, Espinosa-Mazariego K, Ochoa SA, Saldaña Z, Rodea GE, Cázares-Domínguez V, Rodríguez-Ramírez V, Eslava-Campos CA, Navarro-Ocaña A, Arrellano-Galindo J, Hernández-Castro R, Gómez-Duarte OG, Qadri F and Xicohtencatl-Cortes J (2014) CS21 positive multidrug-resistant ETEC clinical isolates from children with diarrhea are associated with self-aggregation, and adherence. *Front. Microbiol.* 5:709. doi: 10.3389/fmicb.2014.00709

This article was submitted to Infectious Diseases, a section of the journal *Frontiers in Microbiology*.

Copyright © 2014 Cruz-Córdova, Espinosa-Mazariego, Ochoa, Saldaña, Rodea, Cázares-Domínguez, Rodríguez-Ramírez, Eslava-Campos, Navarro-Ocaña, Arrellano-Galindo, Hernández-Castro, Gómez-Duarte, Qadri and Xicohtencatl-Cortes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Dimeric and Trimeric Fusion Proteins Generated with Fimbrial Adhesins of Uropathogenic *Escherichia coli*

Víctor M. Luna-Pineda^{1,2}, Juan Pablo Reyes-Grajeda³, Ariadna Cruz-Córdova¹, Zeus Saldaña-Ahuactzi^{1,2}, Sara A. Ochoa¹, Carmen Maldonado-Bernal⁴, Vicenta Cázares-Domínguez¹, Leticia Moreno-Fierros⁵, José Arellano-Galindo⁶, Rigoberto Hernández-Castro⁷ and Juan Xicohtencatl-Cortes^{1*}

¹Laboratorio de Investigación en Bacteriología Intestinal, Hospital Infantil de México "Federico Gómez", Ciudad de México, Mexico, ²Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Ciudad de México, Mexico, ³Instituto Nacional de Medicina Genómica, Ciudad de México, Mexico, ⁴Laboratorio de Investigación de Inmunología y Proteómica, Hospital Infantil de México "Federico Gómez", Dirección De Investigación, Ciudad de México, Mexico, ⁵Unidad de Biomedicina, Laboratorio de Inmunidad en Mucosas, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Tlalnepantla, Mexico, ⁶Departamento de Infectología, Hospital Infantil de México "Federico Gómez", Ciudad de México, Mexico, ⁷Departamento de Ecología de Agentes Patógenos, Hospital General "Dr. Manuel Gea González", Ciudad de México, Mexico

OPEN ACCESS

Edited by:

Alfredo G. Torres,
University of Texas Medical Branch,
USA

Reviewed by:

Sheryl S. Justice,
Ohio State University, USA
Charles Martin Dozois,
Institut National de la Recherche
Scientifique, Canada

*Correspondence:

Juan Xicohtencatl-Cortes
juanxico@yahoo.com

Received: 28 July 2016

Accepted: 04 October 2016

Published: 31 October 2016

Citation:

Luna-Pineda VM, Reyes-Grajeda JP, Cruz-Córdova A, Saldaña-Ahuactzi Z, Ochoa SA, Maldonado-Bernal C, Cázares-Domínguez V, Moreno-Fierros L, Arellano-Galindo J, Hernández-Castro R and Xicohtencatl-Cortes J (2016) Dimeric and Trimeric Fusion Proteins Generated with Fimbrial Adhesins of Uropathogenic *Escherichia coli*. *Front. Cell. Infect. Microbiol.* 6:135. doi: 10.3389/fcimb.2016.00135

Urinary tract infections (UTIs) are associated with high rates of morbidity and mortality worldwide, and uropathogenic *Escherichia coli* (UPEC) is the main etiologic agent. Fimbriae assembled on the bacterial surface are essential for adhesion to the urinary tract epithelium. In this study, the FimH, CsgA, and PapG adhesins were fused to generate biomolecules for use as potential target vaccines against UTIs. The fusion protein design was generated using bioinformatics tools, and template fusion gene sequences were synthesized by GenScript in the following order *fimH-csgA-papG-fimH-csgA* (*fcpcf*) linked to the nucleotide sequence encoding the [EAAAK]₅ peptide. Monomeric (*fimH*, *csgA*, and *papG*), dimeric (*fimH-csgA*), and trimeric (*fimH-csgA-papG*) genes were cloned into the pLATE31 expression vector and generated products of 1040, 539, 1139, 1442, and 2444 bp, respectively. Fusion protein expression in BL21 *E. coli* was induced with 1 mM IPTG, and His-tagged proteins were purified under denaturing conditions and refolded by dialysis using C-buffer. Coomassie blue-stained SDS-PAGE gels and Western blot analysis revealed bands of 29.5, 11.9, 33.9, 44.9, and 82.1 kDa, corresponding to FimH, CsgA, PapG, FC, and FCP proteins, respectively. Mass spectrometry analysis by MALDI-TOF/TOF revealed specific peptides that confirmed the fusion protein structures. Dynamic light scattering analysis revealed the polydispersed state of the fusion proteins. FimH, CsgA, and PapG stimulated the release of 372–398 pg/mL IL-6; interestingly, FC and FCP stimulated the release of 464.79 pg/mL ($p \leq 0.018$) and 521.24 pg/mL ($p \leq 0.002$) IL-6, respectively. In addition, FC and FCP stimulated the release of 398.52 pg/mL ($p \leq 0.001$) and 450.40 pg/mL ($p \leq 0.002$) IL-8, respectively. High levels of IgA and IgG antibodies in human sera reacted against the fusion proteins, and under identical conditions, low levels of IgA and IgG antibodies were detected in human urine.

Rabbit polyclonal antibodies generated against FimH, CsgA, PapG, FC, and FCP blocked the adhesion of *E. coli* strain CFT073 to HTB5 bladder cells. In conclusion, the FC and FCP proteins were highly stable, demonstrated antigenic properties, and induced cytokine release (IL-6 and IL-8); furthermore, antibodies generated against these proteins showed protection against bacterial adhesion.

Keywords: UPEC, fimbrial adhesion, fusion protein, UTIs, cytokines

INTRODUCTION

Urinary tract infections (UTIs) are associated with high rates of morbidity and mortality worldwide (Foxman, 2010; Flores-Mireles et al., 2015; Kumar et al., 2015). In Mexico, UTIs are a public health problem and are the third leading cause of morbidity, with ~4 million cases per year (Secretaria de Salud, 2009). Populations with a high risk of acquiring UTIs include newborns, preschool girls, women who are sexually active, and both sexes at advanced ages (Foxman, 2002, 2014).

Uropathogenic *Escherichia coli* (UPEC) is the primary etiologic agent responsible for UTIs, which are classified according to the site of infection: urine (asymptomatic bacteriuria), bladder (cystitis), kidney (pyelonephritis), and blood (urosepsis and bacteremia; Foxman, 2002). The pathogenic mechanism of UPEC begins with adherence via fimbrial adhesins (FimH, PapG, SfaS, and FocH), which are assembled on the distal tip of type 1, P, S, and F1C fimbriae, respectively. Additionally, CsgA (Curli fimbriae) and DrA (Dr fimbriae) proteins have been implicated in epithelial cell adhesion (Antão et al., 2009). These adhesins interact with different receptors (α -D-mannosylated proteins, glycosphingolipids, neuraminic acid, lactosylceramide, decay accelerating factor, and matrix proteins) located on the membrane of cells of the urinary tract (Antão et al., 2009; Lüthje and Brauner, 2014). The FimH adhesin of type 1 fimbriae interacts with uroplakin proteins in the bladder, resulting in an invasion process that allows UPEC to avoid urine flow, antibodies, bactericidal molecules, and antibiotic activity in the urinary tract (Mulvey et al., 1998, 2000; Zhou et al., 2001). UPEC produces biofilm-like structures called intracellular-bacterial communities (IBCs) within the cytoplasm of urothelial cells, conferring protection to the bacteria and facilitating their egress to promote a new cycle of infection through bladder cell lysis (Scott et al., 2015). During infection cycles, UPEC enter a quiescent state for long periods of time, and this quiescence constitutes a mechanism for bacterial persistence (Leatham-Jensen et al., 2016). UPEC then exit the quiescent state by promoting exocytosis from bladder cells and infecting new cells, resulting in recurrent UTIs (rUTIs, Leatham-Jensen et al., 2016). Three percent of women with three or more rUTIs annually are at risk for developing pyelonephritis and urosepsis (Foxman, 2002, 2010). UTIs are typically treated with several broad-spectrum antibiotics (ampicillin, trimethoprim/sulfamethoxazole, fluoroquinolones, and cephalosporin), resulting in increased resistance rates among clinical UPEC strains. This resistance complicates treatment, increases costs, and decreases the efficiency of antibiotics against infection (Biedenbach et al., 2016). The indiscriminate

use of antibiotics modifies the commensal microbiota of patients and generates secondary infections (candida-vaginal and gastrointestinal infections) during and after prophylactic treatment (Flores-Mireles et al., 2015).

The FimH adhesin of UPEC type 1 fimbriae has been used as a biomolecule to induce protection in murine models (Langermann et al., 1997, 2000; Langermann and Ballou, 2001). During infection, type 1 fimbrial expression is regulated by environmental conditions (temperature, osmolality, pH, and nutrients) as well as the specific anatomic site of infection in the urinary tract (bladder, ureters, and kidney). These conditions also dictate the expression of other fimbriae that promote urinary tract colonization by UPEC (Snyder et al., 2005; Greene et al., 2015).

P fimbriae of UPEC have been associated with colonization and inflammation in the kidneys (pyelonephritis) through binding of PapG adhesion to Gal(α 1-4)Gal-containing glycolipids and Toll-like receptor (TLR) 4 in kidney cells (Frendéus et al., 2001; Lane and Mobley, 2007). Curli fimbriae are primarily composed of CsgA protein monomers and are associated with cystitis, pyelonephritis, and bacteremia (Norinder et al., 2012; Hung et al., 2014; Lim et al., 2014). These fimbriae are widely distributed in clinical strains of UPEC and are finely regulated by a regulatory system of on-to-off and/or off-to-on switching (Snyder et al., 2005). Recombinant proteins generated using fusion technology to contain one or more antigens derived from diverse pathogens can induce immune and protective responses against UTIs in *in vivo* models (Asadi Karam et al., 2013; Habibi et al., 2015b). Vaccines designed based on fused proteins of UPEC clinical strains with high variability in their virulence factors might represent potential tools for preventing UTIs. The aim of this study was to generate recombinant fusion proteins (combination of FimH, PapG, and CsgA proteins) as viable biomolecules with vaccine properties to significantly reduce UTIs in Mexico.

MATERIALS AND METHODS

Primary, Secondary, and Tertiary Structures

FimH, CsgA, and PapG protein sequences from the *E. coli* strain CFT073 were obtained from the NCBI database (<http://www.ncbi.nlm.nih.gov/protein>) under accession numbers AAN83822.1, AAN79779.1, and AAN82031.1, respectively. Signal peptide prediction was performed for each protein using the SignalP 4.1 server, and five repetitions of the EAAAK sequence were employed to fuse the FimH, CsgA, and PapG

proteins based on the linker library construction proposed by Li et al. (2016).

In addition, the molecular weight, theoretical isoelectric point (pI), amino acid composition, estimated lifetime, aliphatic index, and grand average of hydropathicity (GRAVY) of the fusion proteins were determined using the ProtParam program of ExPASy (Wilkins et al., 1999). The Codon Adaptation Index (CAI) and guanine and cytosine (GC) content of the genes were determined using the OPTIMIZER program (Puigbò et al., 2007). The secondary structure of the fusion proteins was predicted using the GOR IV program (Sen et al., 2005). Fusion protein modeling was performed using the hierarchical modeling approach (I-TASSER) and visualized with PyMOL software (Yang et al., 2015). Tridimensional (3D) structures were refined and minimized with KoBaMIN (<http://csb.stanford.edu/kobamin/>) and VegaZZ (NAM; Pedretti et al., 2004). The 3D models were then validated by Protein Structure Analysis (ProSA) to determine Z-scores and Ramachandran plots using PROCHECK (Laskowski et al., 1998; Wiederstein and Sippl, 2007). Merging of the 3D fusion proteins with the crystal structures of the mannose-binding domain FimH (Protein Data Bank; 1TR7) and lectin domain PapG (PDB; 1J8R) were calculated with the root mean square deviation (RMSD) using the TM-align program (Zhang and Skolnick, 2005).

Epitope Prediction and Antigen Presentation

Immune response was theoretically determined to establish variants of fusion proteins with enhanced ability to generate an immune response. The primary and secondary structures of the fusion proteins were employed to determine the lineal antigenic epitopes of B with the AbcPred server (<http://www.imtech.res.in/raghava/abcpred/>), and peptides with affinities for major histocompatibility complex (MHC) class II were identified with the NetMHCII program (<http://www.cbs.dtu.dk/services/NetMHCII/>). The 3D structures of the fusion proteins were employed to determine the conformational antigenic epitopes using the DiscoTope server (<http://www.cbs.dtu.dk/services/DiscoTope/>).

Analysis and Synthesis of a Template Fusion Gene

The gene sequences of *fimH*, *csgA*, and *papG* from the *E. coli* strain CFT073 were obtained from GenBank through NCBI (<http://www.ncbi.nlm.nih.gov/genbank/>) under the accession numbers GQ487191.1, NC_004431.1, and AF447814.1, respectively. The conserved sequences of the *fimH*, *csgA*, and *papG* genes from various UPEC strains (UTI89, ABU83972, NA114, UPEC 26-1, CF-088, CF-468, isolates IA2, and AD110) were determined using the BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) programs. Consensus nucleotide sequences were fused with the appropriate codons from the sequence EAAAK to generate a template fusion gene. The sequence of the template fusion gene was optimized using the OPTIMIZER program (<http://genomes.urv.es/OPTIMIZER/>), and messenger RNA

secondary structures of monomeric, dimeric, and trimeric genes were predicted using the Mfold program (<http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form>). The optimized template fusion gene, containing a 5' *Bam*HI site and 3' *Sac*I site, was synthesized by GenScript (Piscataway, NJ, USA) and cloned into the pUC57 vector to amplify and clone monomeric, dimeric, and trimeric genes.

Cloning of Monomeric and Fusion Genes into the pLATE Expression Vector

Primer designs for monomeric, dimeric, and trimeric genes from the synthesized template fusion gene were generated manually according to the aLICator LIC Cloning and Expression handbook (Thermo Fisher Scientific, Waltham, MA, USA) and synthesized by IDT Technologies (Coralville, Iowa, USA). Gene amplification was performed by polymerase chain reaction (PCR) using *Pfu* DNA polymerase. Conditions for amplification and cloning were obtained from the aLICator LIC Cloning and Expression handbook. The pLATE31 expression vector was used to clone the fusion genes and transformed by electroporation into BL21 (DE3) *E. coli*.

Sequencing and GenBank Submission of the Fusion Genes

Fusion genes in recombinant vectors were verified by next-generation sequencing on a NexSeq 500 system (Illumina, San Diego, CA, USA). Specific sequencing primers were obtained from the aLICator LIC Cloning and Expression Kit, and the gene sequences were analyzed by BLAST and submitted to GenBank at NCBI.

Expression and Purification of Fusion Proteins

A BL21 *E. coli* strain carrying the pLATE31 expression vector (fusion proteins) was plated on Luria Bertani (LB; Becton-Dickinson, Franklin Lakes, New Jersey, USA) agar and incubated for 16 h at 30°C. The putative transformant colonies were selected via the colony-blotting method using an anti-6X His (C-Term) HRP antibody (Abcam; Cambridge, MA, USA). Solubility assays were performed following the protocol described in the QIAexpressionist manual from Qiagen (Jacques-Schiesser-Str, Stockach, Germany). Fusion protein expression was performed in 500 mL of LB medium supplemented with 1 mM IPTG for 5 h at 37°C. The cell pellets were resuspended in phosphate buffer (10 mM K₂HPO₄, pH 7.4, 150 mM NaCl, and 1 mM EDTA) with phenylmethylsulfonyl fluoride (PMSF; Sigma-Aldrich Corp., St. Louis, MO, USA), lysed by sonication, and centrifuged at 26,116 g for 20 min. The supernatants were discarded, and the pellets were resuspended in denaturing buffer (8 M guanidine, 100 mM NaCl, and 100 mM K₂HPO₄, pH 8). After 3 days of incubation at room temperature, the lysates were centrifuged at 58,762 g for 20 min. The supernatants were incubated on a column containing nickel-nitrilotriacetic acid-agarose (Qiagen) at 4°C for 1 h, washed with A-buffer (8.5 M urea, 20 mM Tris, pH 7.5, 160 mM NaCl, and 20 mM imidazole) and eluted with B-buffer (8 M urea, 50 mM Na₂HPO₄, pH 8,

100 mM NaCl, and 500 mM imidazole). The collected proteins underwent refolding via dialysis using a urea gradient from 7 to 1 M in C-buffer (25 mM Tris, pH 7.5, 100 mM NaCl, and 0.5 mM EDTA); incubation was performed at 4°C for 24 h. Refolded fusion proteins were stored in C-buffer at -70°C for all assays.

Characterization of Fusion Proteins

Fusion proteins were quantified according to the protocol for the 2D-Quant kit (GE Healthcare Bio-Sciences AB, Björkgatan, Uppsala, Sweden), separated by performing 14% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), visualized by Coomassie staining and identified by mass spectroscopy using a 4800 MALDI TOF/TOF™ Analyzer (Applied Biosystems/MDS SCIEX, Waltham, MA, USA) after excision of the proteins spots. The CsgA protein was treated with 88% formic acid (Sigma-Aldrich Corp., St. Louis, MO, USA) prior to SDS-PAGE (Saldaña et al., 2009). Molecular weights were estimated using Image Lab software version 5.2 from Bio-Rad (Hercules, California, USA). The aggregation state of fusion proteins was determined by performing dynamic light scattering (DLS) using a Zetasizer Helix (Malvern Instruments Ltd, Grovewood Road, Worcestershire, United Kingdom). Fusion proteins with histidine tags were transferred onto polyvinylidene difluoride (PVDF) membranes and confirmed by performing Western blot assays using anti-6His (C-Terminal) HRP antibodies (Abcam; Cambridge, MA, USA) as described by Ledesma et al. (2010). The endotoxin (LPS) levels of the purified fusion proteins were determined using the Pierce™ LAL Chromogenic Endotoxin Quantitation kit according to the manufacturer's protocol (Thermo Fisher Scientific; Waltham, MA, USA). Additionally, fusion proteins were treated with 50 µg/mL polymyxin B (Sigma-Aldrich Corp., St. Louis, MO, USA) for 12 h at 4°C prior to bioactivity assays.

Bioactivity Assays Using the Fusion Proteins

TLR2 and TLR4 expression in HTB5 bladder cells [American Type Culture Collection (ATCC), Manassas, VA, USA] was analyzed by flow cytometry using a human TLR2 fluorescein-conjugated antibody (R&D Systems, Inc., Minneapolis, USA) and a human TLR4/MD-2 complex phycoerythrin-conjugated antibody (Santa Cruz Biotechnology Inc., Texas, USA). Activation of TLR2 and TLR4 by the fusion proteins was assessed by quantifying the release of IL-6 and IL-8. HTB5 cells were cultured in 24 well-plates (Greiner, Germany) at a density of 10⁵ cells/well incubated with 1 mL of fresh Eagle's minimum essential medium (EMEM; ATCC® 30-2003™) supplemented with 10% fetal bovine serum (FBS) from Gibco (Thermo Fisher Scientific; Waltham, MA, USA). Cytokine induction in HTB5 cells was detected after 6 h of incubation with 10 µg/mL FimH, CsgA, PapG, FC, and FCP proteins by performing enzyme-linked immunosorbent assays (ELISAs) following the protocol established by BD Biosciences (San Jose, CA, USA). In addition, 100 ng/mL lipopolysaccharide (LPS; Sigma-Aldrich Corp., St. Louis, MO, USA) from *E. coli* 0111:B4 and 100 ng/mL

lipoteichoic acid (LTA; Sigma-Aldrich Corp., St. Louis, MO, USA) from *S. aureus* were used as controls for the induction of TLR4 and TLR2, respectively.

Antigenicity of Fusion Proteins

The study was approved by the Research (Dr. Onofre Muñoz Hernández), Ethics (Dr. Amparo Faure Fontenla), and Biosecurity (Dr. Herlinda Vera Hermosillo) Committees of Hospital Infantil de México Federico Gómez (HIMFG) under numbers HIM/2014/022 and HIM/2016/027. Physicians from the Infectology Department obtained permission and consent from patients to use the urine and serum samples employed in this study. Serum and urine samples were collected from 14 UTI patients (UP); samples from 14 healthy patients (HP) recruited from "Laboratorio Central del HIMFG" were used as negative controls. Patients were selected based on the following criteria: UTI symptoms, urine culture with ≥100,000 CFU/mL *E. coli*, leukocyte esterase and/or nitrites, and urothelial cells in urine. Urine samples were centrifuged at 7835 g for 5 min and filtered through a 0.22-µm Durapore membrane (Merck Millipore; Darmstadt, Germany). IgG and IgA antibody titers against the FimH, CsgA, PapG, FC, and FCP proteins were determined by ELISA using serum and urine samples diluted 1:50 and 1:10, respectively.

Generation of Polyclonal Rabbit Antibodies

Five female 6-month-old New Zealand rabbits were obtained from the "Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México" for use in this study. Initially, the rabbits were immunized subcutaneously with 200 µg of FimH, CsgA, PapG, FC, and FCP proteins in complete Freund's adjuvant. The rabbits were re-immunized three times (days 21, 28, and 37) with 100 µg of each protein in incomplete Freund's adjuvant and bled via cardiac puncture on day 40. The collected blood was centrifuged at 7835 g for 5 min, and separated sera were stored at -70°C until use. The sera were absorbed using the CFT073 *csgA::km + fimH::cm* mutant strain generated in this study via a one-step inactivation method (Datsenko and Wanner, 2000). Anti-PapG and anti-FCP polyclonal rabbit sera were adsorbed using the CFT073 *csgA::km + fimH::cm* mutant strain under conditions that did not permit PapG expression, as confirmed by RT-PCR. Additionally, sera were heat-inactivated at 56°C for 30 min and titrated by ELISA using serial dilutions from 1:10 to 1:100,000 against the appropriate proteins.

Adherence Inhibition Assay with HTB5 Bladder Cells

HTB5 bladder cells were cultured in 24 well-plates at a density of 10⁵ cells/well in 1 mL of fresh EMEM supplemented with 10% FBS until 80% confluence was reached. Previously, the strain CFT073 was grown overnight in LB broth at 37°C, and an aliquot (1:100 dilution) of the culture was incubated until it reached an OD₆₀₀ of 1.0. The bacterial culture was then mixed with pre-immune or/and immune serum at a 1:1 ratio and maintained 2 h at 37°C with constant agitation. Cell monolayers were infected with 10⁷ bacterial cells [multiplicity

of infection (MOI): 100] and incubated for 2 h at 37°C in 5% CO₂. Infected cell monolayers were washed three times with PBS, and 200 μL of PBS containing 0.1% Triton X-100 was added. The bacteria that adhered to the cell monolayers were homogenized, and colony-forming units were determined using the Spotting Method described by Hannan and Hunstad (2016).

Statistical Analysis

One-way ANOVA and Student's *t*-test were used to compare differences between the mean and median values of groups using GraphPad Prism software (version 6). A *p* < 0.05 was considered statistically significant for all results.

RESULTS

Structural Components of the Fusion Proteins

Fusion proteins were generated using the sequences of the *fimH*, *papG*, and *csgA* genes, which exhibit identity of ≥90% among clinical UPEC strains. In addition, a signal peptide was identified between residues 1 and 27 of the sequences of the FimH and PapG proteins and residues 1–21 of the CsgA protein. Additionally, the CsgA protein contained a nucleation region between residues 22 and 41. These signal peptides and the nucleation region were eliminated from the sequences used to design and synthesize of the fusion proteins. The fusion of the FimH, PapG, and CsgA proteins was generated by employing the five repeats of the EAAAK linker ([EAAAK]₅), which forms an alpha helix with a length of 39.95 Å stabilized by five salt bridges between the Glu⁻ and Lys⁺ residues.

Generation of Dimeric and Trimeric Genes

A template fusion gene containing (in order) *fimH-csgA-papG-fimH-csgA* (*fcpcf*) linked to the nucleotide sequence of the [EAAAK]₅ peptide was designed (Figure 1A). Six specific primers localized in the 5' (forward) and 3' (reverse) regions of each gene were designed and synthesized as described in the aLICator LIC Cloning and Expression Handbook (Table 1). Monomeric, dimeric, and trimeric genes were generated using combinations of the six primers in PCR amplification and cloned into the pLATE31 expression plasmid (Figure 1B).

Synthesis of the Template Fusion Gene

Increased expression levels of the fusion genes were obtained by preferential codon optimization of the *fcpcf* template fusion gene, which exhibited a CAI of 0.433–1.0 and a GC content of 48.9–55.4%. The *fcpcf*-optimized gene was synthesized chemically and cloned into the pUC57 vector to produce pF₂C₂P. Sequencing and digestion of the pF₂C₂P were consistent with the theoretical profile (data not shown).

Primary and Secondary Structures

The physical and chemical parameters of the monomeric (FimH, CsgA, and PapG), dimeric [FimH-CsgA (FC), CsgA-PapG (CP), and PapG-FimH (PF)], and trimeric [FimH-PapG-CsgA (FCP), PapG-CsgA-FimH (PCF), and CsgA-FimH-PapG (CFP)] proteins are shown in Table 2. Primary sequence analysis of these proteins revealed a high number of linear epitopes, between 18 and 33, as well as a high number of peptides binding to MHC-II, between 68 and 159 (Table 3). The secondary structure predictions for the monomeric, dimeric, and trimeric proteins indicated a high percentage of random coil and beta-sheet

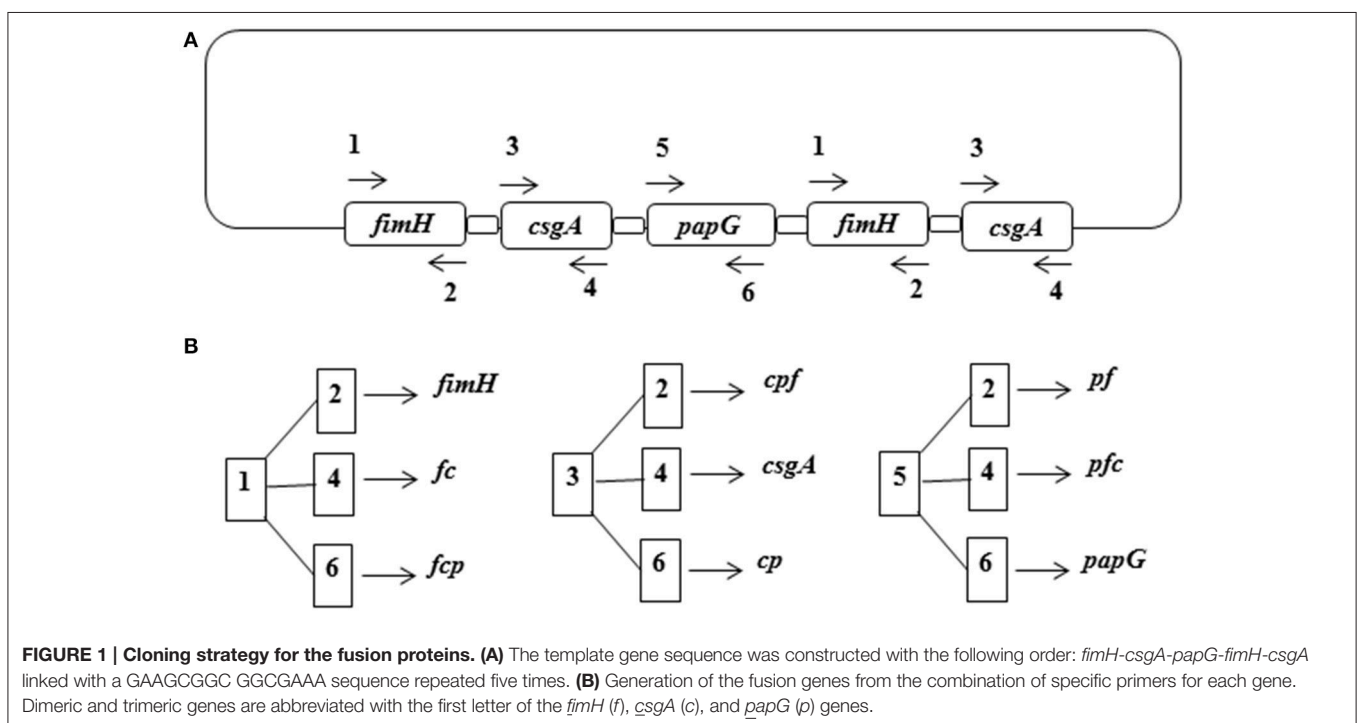


TABLE 1 | List of primers used in this study.

Number	Primer	Sequence (5'-3')	References
1	FimH F	AGAAGGAGATATAACTATGAAAACCGCGAACGGTACCGCGATCCCGATCGGTGGTGGT	This study
2	FimH R	GTGGTGGTGGTGGTGGTGGCCCTGGTAAACGAAGGTAACACCGATGATAGACTGAAC	This study
3	CsgA F	AGAAGGAGATATAACTATGTCTGAACTGAACATCTACCAGTACGGTGGTGGTA	This study
4	CsgA R	GTGGTGGTGGTGGTGGTGGCCGCTACTGGTGC GCGGTGCGTTGTTACCGAA	This study
5	PapG F	AGAAGGAGATATAACTATGTCTCTGGGTAACGTTAACTCTTACCAGGGTGGTAA	This study
6	PapG R	GTGGTGGTGGTGGTGGTGGCCCGGCAGGATCATCAGCAGGGTTCGCGAACCAG	This study

TABLE 2 | Analysis of the physicochemical parameters of the fusion proteins.

Protein	No. of amino acids	Molecular weight (kDa)	Theoretical pI	Extinction coefficient	Estimated half-life	Instability index	Aliphatic index	GRAVY
FimH	283	29.64	6.63	1.179	10	25.59	85.72	0.029
CsgA	116	12.06	5.58	0.826	10	16.23	51.38	-0.791
PapG	316	35.38	8.89	1.882	10	25.89	74.05	-0.364
FC	417	43.09	5.66	1.043	10	21.74	76.07	-0.183
CP	450	48.83	6.92	1.563	10	21.37	68.58	-0.443
PF	617	66.4	8.44	1.529	10	25.58	79.68	-0.161
FCP	751	79.85	6.74	1.397	10	22.37	75.39	-0.245
CPF	751	79.85	6.74	1.397	10	22.37	75.39	-0.245
PFC	751	79.85	6.74	1.397	10	22.37	75.39	-0.245

TABLE 3 | B-cell epitopes and MHC class II binding peptide predictions for the fusion proteins.

Fusion protein	Lineal epitopes	Conformational epitopes	MHC class II binding peptide
FC	18	13	86
CP	27	12	68
PF	32	8	147
FCP	32	17	149
CPF	33	14	152
PFC	33	13	159

structures; however, lower percentages of alpha helices were observed (Figure 2). The conserved region of the alpha helix structure conferred by the [EAAAK]₅ linker was identified in all fusion proteins.

3D Structures of the Fusion Proteins

The 3D structure models of the dimeric and trimeric proteins were refined 10 times to obtain the optimal minimization energy (Figure 2 and Table 4). Interestingly, the FC and FCP proteins exhibited models with optimal theoretical data that indicated a stable structural conformation. Gibbs free energy (ΔG) data indicated high stability for the FC protein, with $\Delta G = -7040.78$ kcal/mol, and the FCP protein, with $\Delta G = -13,343.83$ kcal/mol (Table 4). The 3D model validation of the FC and FCP proteins indicated that 94.7 and 93.1% of the amino acid residues were within permitted regions in the Ramachandran plot, with Z-scores of -6.95 (FC) and -9.53 (FCP).

The 3D models of the FC and FCP proteins exhibited RMSDs of 1.81 and 1.67 Å, respectively, indicating conformations very similar to that of the mannose-binding domain of FimH. The FCP protein also exhibited an RMSD of 1.62 Å with the lectin domain of PapG (Table 4). Analysis of the 3D models using the Discotope program revealed 13 and 17 conformational epitopes in the FC and FCP proteins, respectively (Table 3).

Generation of Fusion Proteins

Bioinformatic analysis was an essential tool for the selection of monomeric (*fimH*, *csgA*, and *papG*), dimeric (*fc*), and trimeric (*fcp*) genes, which were cloned in the pLATE31 expression vector. PCR assays were performed to verify the cloned expression vectors using specific sequencing primers as described in the aLICator LIC Cloning and Expression Handbook (Figure 3A). The sequences of the fusion genes were determined by sequencing and will be submitted to the GenBank database (unpublished data). The BL21 (DE3) *E. coli* strain carrying the expression vectors was induced with 1 mM IPTG, and the His-tagged fusion proteins were purified under denaturing conditions by affinity chromatography. Coomassie blue-stained SDS-PAGE gels and Western blot analysis revealed 26.5, 11.9, 33.9, 44.9, and 82.1 kDa bands corresponding to FimH, CsgA, PapG, FC, and FCP, respectively (Figures 3B,C). The protein concentrations of FimH, CsgA, PapG, FC, and FCP were 2.7, 0.697, 2.63, 1.03, 0.998 mg/mL, respectively.

Mass spectrometry analysis by MALDI-TOF/TOF revealed the presence of specific peptides corresponding to the fusion proteins, and DLS analysis of the fusion proteins indicated a polydispersed state (Table 5).

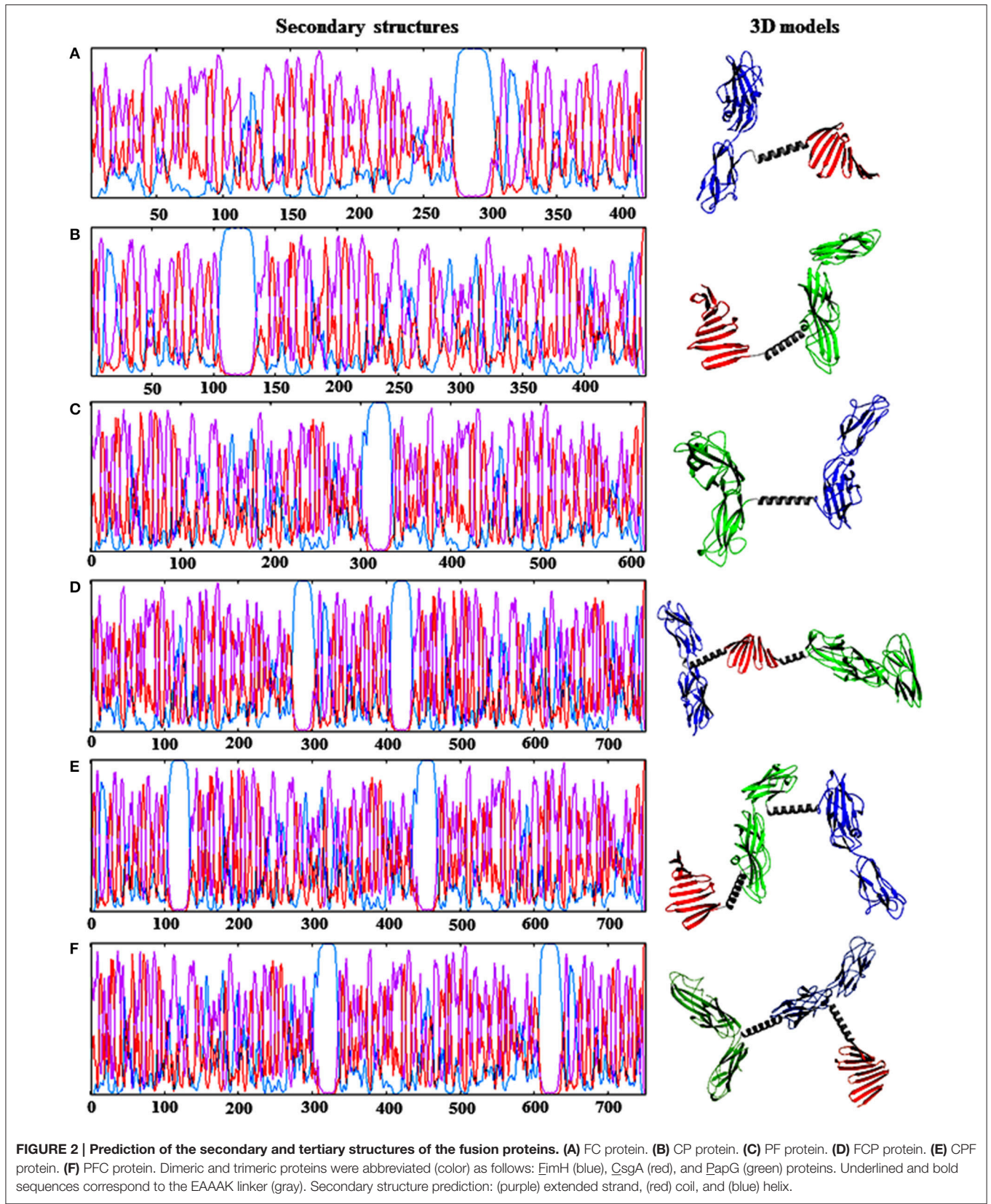
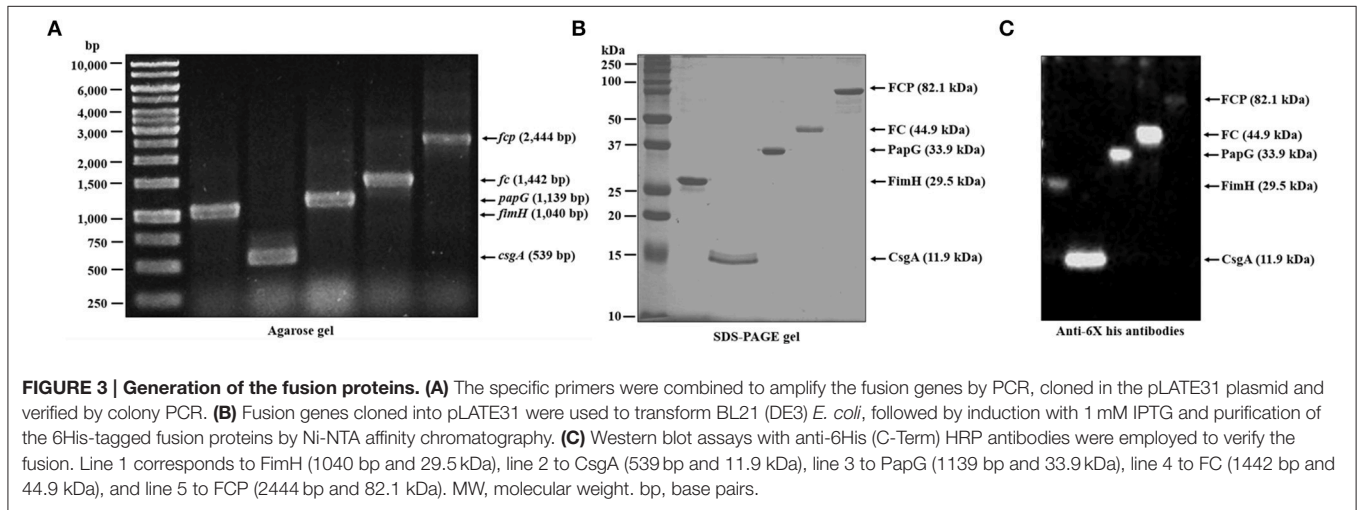


TABLE 4 | Parameters for 3D modeling, refinement, and validation of the fusion proteins.

Fusion protein	Refinement cycle	Minimization energy (ΔG) (kcal/mol)	Ramachandran plot (%)	Z-score	RMSD (FimH-PapG)
FC	10	-11,912.82	94.70	-6.97	1.81 Å-NA
CP	10	-7040.78	92.30	-7.7	NA-2.02 Å
PF	10	-7140.54	92.40	-7.7	1.83 Å-1.37 Å
FCP	10	-13,866.68	93.10	-9.53	1.67 Å-1.62 Å
CPF	10	-13,343.83	92.50	-10.72	1.99 Å-2.02 Å
PFC	10	-13,577.65	91.90	-10.68	1.88 Å-1.38 Å

Calculations for RMSD were generated with the FimH-lectin domain (PDB 1TR7) and PapG-lectin domain (PDB 1J8R). NA, not applicable.



Induction of Proinflammatory Cytokine Release by the Fusion Proteins

Flow cytometry analysis of HTB5 bladder cells revealed an expression of 50.3% for TLR2 and 31.7% for TLR4 (**Figure 4**). Endotoxin levels ≤ 0.012 EU/mL were observed when the FimH, CsgA, PapG, FC, and FCP proteins were treated with 50 μ g/mL polymyxin B. HTB5 bladder cells expressing TLR2 and TLR4 were incubated with 10 μ g/mL FimH, CsgA, PapG, FC, or FCP, which induced different percentages of IL-6 and IL-8 cytokine release. Briefly, the FimH, CsgA, and PapG proteins stimulated the release of 372–398 pg/mL IL-6, and no significant differences in IL-6 release were observed among these proteins. However, the FC and FCP proteins induced the release of 464.79 and 521.24 pg/mL IL-6, respectively, which were significantly different ($p \leq 0.018$ and $p \leq 0.002$, respectively) from the values obtained for the monomeric proteins (**Figure 5**). In addition, 398.52 pg/mL IL-8 was released by the FC protein ($p \leq 0.003$ compared to the monomeric protein), and 450.40 pg/mL IL-8 was released by the FCP protein ($p \leq 0.037$ compared to the monomeric protein; **Figure 6**).

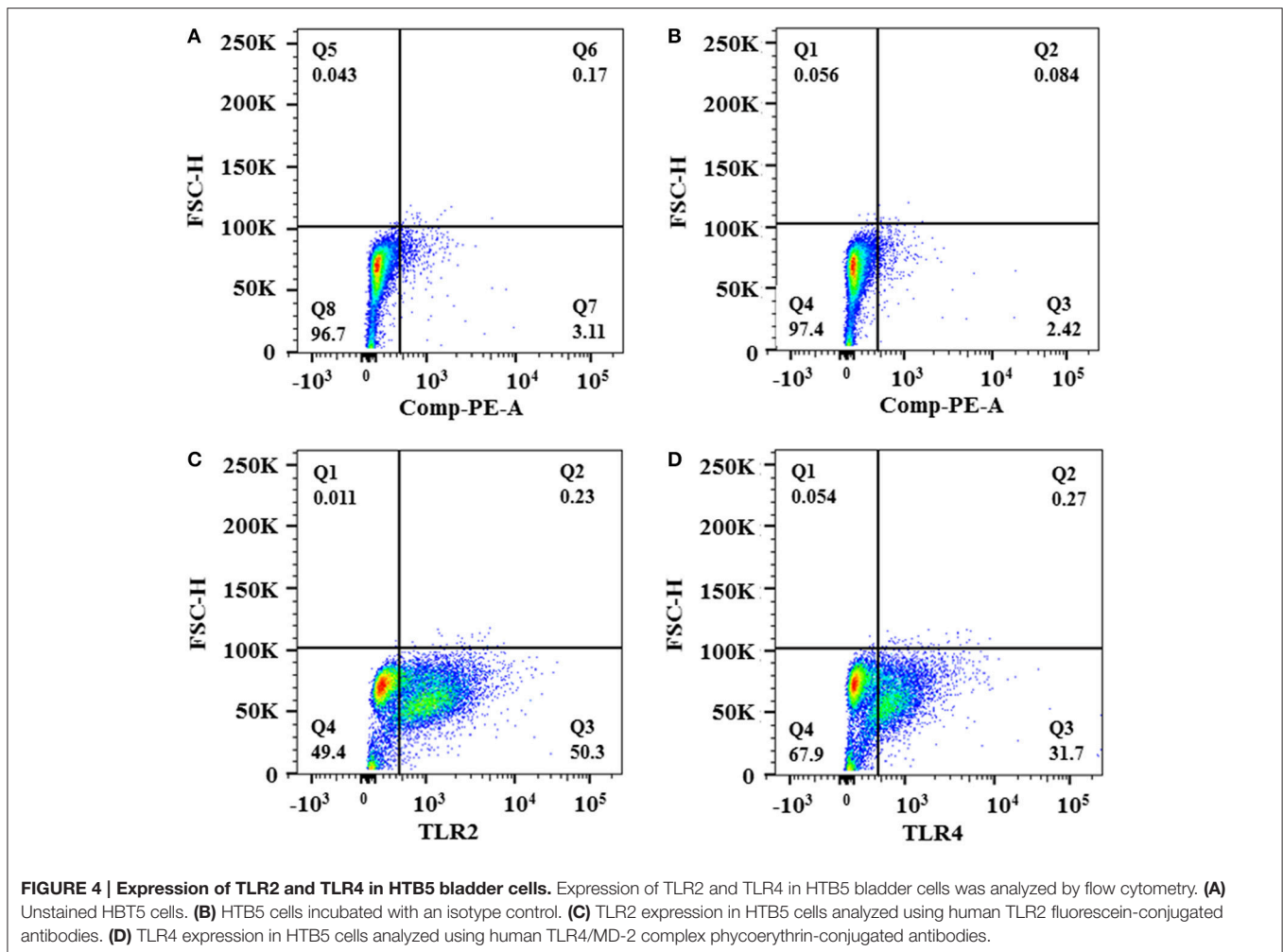
Antigenicity of the Fusion Proteins

ELISA assays were performed to detect antibodies against the FimH, CsgA, PapG, FC, and FCP proteins in the sera and urine of HP and UP (**Figures 7, 8**). Antibodies in the sera reacted with all proteins evaluated. High levels of UP-IgA antibodies

in sera were observed, with a median OD₄₅₀ of 1.628–2.216. Interestingly, high levels of UP-IgA antibodies in sera, with a median OD₄₅₀ of 2.216 were detected when ELISA was performed using the FC protein as an antigen (**Figure 7B**). However, under identical conditions, low levels of UP-IgG (median values of 0.514–0.837) were detected. The FC protein exhibited greater specificity, with sera detecting UP-IgG with a median value of 0.837 (**Figure 7A**). Additionally, all fusion proteins reacted with high levels of UP-IgA antibodies in urine, with a significant difference of $p = 0.0003$ compared to the values for HP-IgA antibodies. Furthermore, 1 μ g/mL fusion protein detected high levels of UP-IgA antibodies in urine, with median values ranging from 0.6189 to 0.950 (**Figure 8B**). In contrast, low levels of UP-IgG were detected in urine, with median values ranging from 0.444 to 0.619 (**Figure 8A**). The difference in the values between UP-IgA and UP-IgG in urine was significant ($p < 0.05$) compared to HP-IgA and HP-IgG (**Figure 8**).

Polyclonal Rabbit Antibodies Block Bacterial Adherence

Polyclonal rabbit antibodies against the FimH, CsgA, PapG, FC, and FCP proteins were used to block bacterial adherence to HTB5 bladder cells. The strain CFT073 showed adherence to 6.3×10^6 HTB5 bladder cells, and this level of adherence was set at 100%. Similar adherence by the strain CFT073 was observed

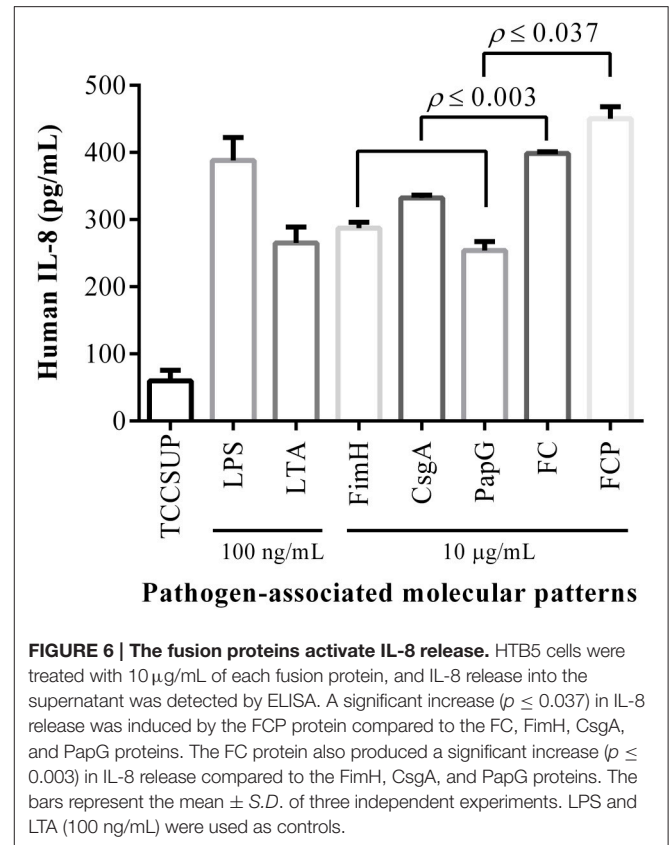
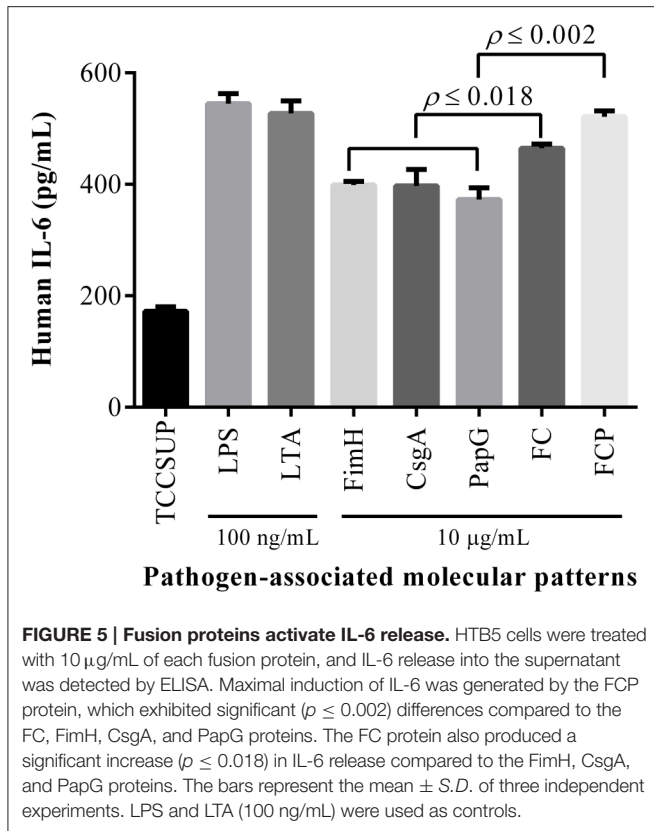


with pooled pre-immune sera (Figure 9). Bacterial adherence in the presence of polyclonal rabbit antibodies showed a significant reduction to 32% for anti-FimH ($p = 0.002$), 21% for anti-CsgA ($p = 0.0011$), 60% for anti-PapG ($p < 0.0001$), 73% for anti-FC ($p < 0.0001$), and 46% for anti-FCP ($p < 0.0001$) compared with the CFT073 and pre-immune serum controls (Figure 9).

DISCUSSION

UTIs are the second most common infection in humans and are associated with high annual economic costs due to patient morbidity. UPEC is the most common etiologic agent of UTIs (Foxman, 2014). Increased antibiotic resistance among UPEC strains has reduced treatment options and significantly increased UTI-associated morbidity and mortality (Melzer and Petersen, 2007; Molina-López et al., 2011). Research to date has focused on more effective and less costly options for antibiotic therapy to treat UTIs and vaccination as a viable alternative due to increasing resistance among UPEC strains (Serino et al., 2010; Sivick and Mobley, 2010; Brumbaugh and Mobley, 2012; Moriel and Schembri, 2013; Mobley and Alteri, 2015).

Fusion technology has been utilized to design new vaccines due to its ability to induce a rapid cellular and humoral immunity response against UPEC (Huleatt et al., 2007). Fimbrial adhesin plays a critical role in UPEC binding, colonization, and bladder and kidney cell invasion (Spaulding and Hultgren, 2016). We hypothesized that blocking adhesion using anti-fimbrial adhesin antibodies would inhibit the initial step in the UPEC pathogenic mechanism. Fusion proteins comprising the FimH, PapG, and CsgA adhesins and EAAAK rigid linkers were designed and analyzed using bioinformatics tools. Based on bioinformatics, we selected the [EAAAK]₅ peptide, which contains a longer helix that results in appropriate separation among fusion proteins. Small-angle X-ray scattering data have revealed that multimerization is caused by short helical linkers ($n = 2-3$), whereas longer linkers ($n = 4-5$) solvate monomeric fusion proteins (Arai et al., 2004; Zagrovic et al., 2005). The monomeric, dimeric, and trimeric fusion genes generated from the *fcpcf* optimized-template exhibited a CAI value of 1.0 and a high GC content (55.4%). Thus, the designed sequence appears to be optimal for expressing all fusion proteins, as described in previous studies (Brinkmann et al., 1989; Andrews et al., 1996; Pedersen-Lane et al., 1997). Three different variants of the



dimeric (*fc*, *cp*, and *pf*) and trimeric (*fc*, *cpf*, and *pcf*) genes were designed to construct and express the fusion proteins. However, based on theoretical analysis of the fusion proteins, FC, and FCP were selected for their structural stability and folding. *In silico* studies of fusion proteins based on FimH from UPEC and MrpH from *Proteus mirabilis* have revealed that the sequence order is critical for stability and folding (Habibi et al., 2015a). Bioinformatics tools were recently used to propose potential vaccine designs for *Helicobacter pylori*, enterotoxigenic *E. coli*, *Brucella*, influenza viruses, and UPEC (Nazarian et al., 2012; Savar et al., 2014; Golshani et al., 2015; Muñoz-Medina et al., 2015; Habibi et al., 2015b; Mohammad et al., 2016). Additionally, vaccines have been designed to contain T-cell and B-cell immune epitopes, and a class II MHC peptide was also identified using bioinformatics tools (De Groot et al., 2002; De Groot and Moise, 2007). Based on these studies, we evaluated the physicochemical properties of the amino acid residues, linear epitopes and peptides binding MHC-II in the primary sequences of the fusion proteins. Prediction of the physicochemical properties of the fusion proteins by our “*in silico*” analysis suggested high amounts of epitopes and peptides that bound MHC-II; furthermore, there appeared to be a relationship between the number of epitopes and the physicochemical properties of the fusion proteins (Sim et al., 1996; Huang et al., 2016).

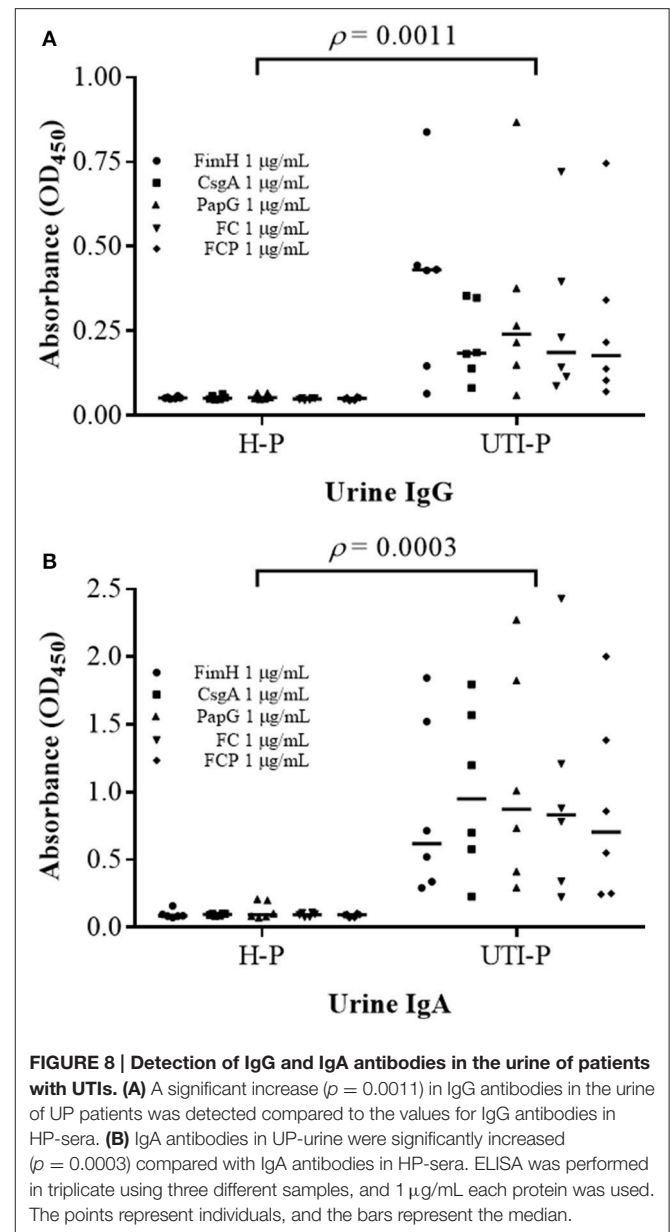
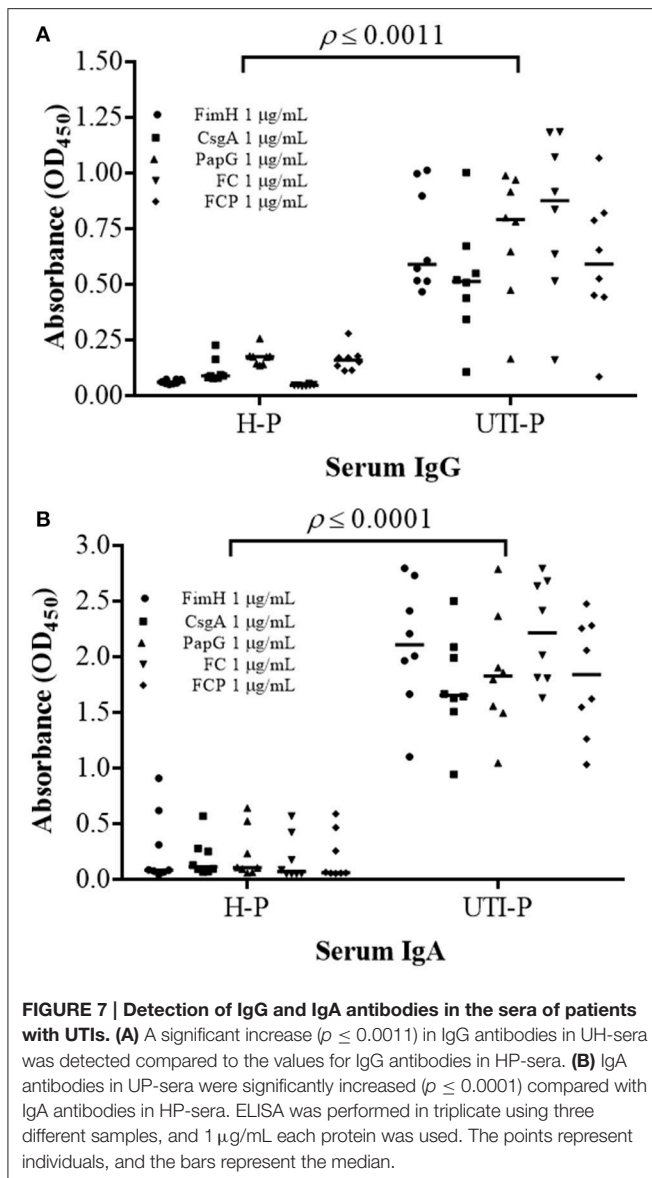
The FimH protein contains a mannose-binding lectin domain between residues 1 and 150 and a pilin domain between residues 159 and 279, with eight amino acids connecting the two domains (Hung et al., 2002). We identified epitopes and peptides in the

TABLE 5 | Characterization of fusion proteins by DLS.

Sample name	T (°C)	Z-Ave (d.nm)	Pdl
FimH	24.9	132.8	0.276
CsgA	24.9	2.97E + 04	0.948
PapG	25	264.9	0.253
FC	24.9	381.2	0.665
FCP	25	113.8	0.451

Pdl ≤ 0.04 , monodispersed; *Pdl* ≥ 0.04 , polydispersed.

minimal mannose-binding region of FimH (residues 1–25), a region that has been reported to confer protection against UTIs caused by UPEC (Thankavel et al., 1997). Similarly, the PapG protein contains a lectin domain between residues 1 and 198 and a pilin domain between residues 206 and 336 (Sung et al., 2001). Our data revealed the presence of both epitopes and peptides of the PapG protein in its Gal(α 1-4)Gal-binding region, between residues 1 and 63 and residues 155 and 173 (Sung et al., 2001). Moreover, CsgA protein prediction revealed two epitopes localized in repetitions 4 and 5 that are protein-binding, functional regions characterized by several beta-turn structures and exposed domains exhibiting a hydrophilic nature in other amyloid proteins (Ikai, 1980; Olsén et al., 2002). Interestingly, the [EAAAK]₅ linker used in the fusion proteins does not contain B-cell epitopes and MHC class II-binding peptides. Moreover, the fusion proteins exhibited a high percentage of random coil structure, which have been reported in other models as principal

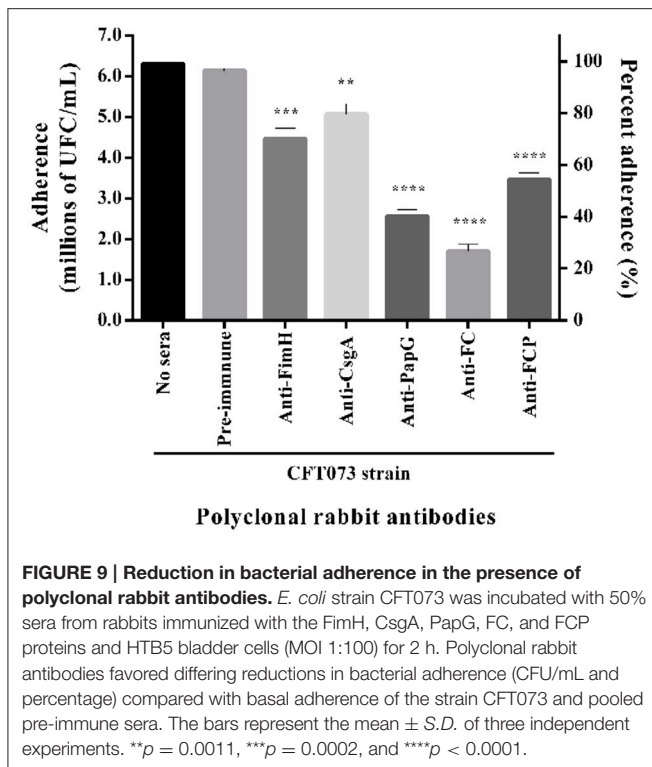


components of antigens with flexible regions that allow protein-protein interactions (Janin and Chothia, 1990; Chen et al., 2007).

3D structure prediction of the fusion proteins confirmed their stability and correct conformation based on refinement, energy minimization, and validation analyses. An RMSD value of 1–2 Å was calculated based on the superimposition of low-resolution X-ray structures, thus indicating that the fusion proteins maintained their structures and revealing the mannose-binding and lectin domains of FimH and PapG in the fusion proteins. Remarkably, stability and structure are properties that are related to protein function. In addition, the CsgA protein has not yet been resolved by crystallography, though circular dichroism, X-ray diffraction, and NMR data for Curli fibers suggest the presence of parallel β -sheet structures. 3D models of CsgA, dimeric (FC and CP), and trimeric proteins proposed in this study reveal the presence of parallel β -sheet structures

(Shewmaker et al., 2009). Interestingly, the FC and FCP proteins exhibited high stability and a greater number of epitopes than the other fusion proteins. FimH, CsgA, and PapG were used as controls and compared to the fusion proteins. Theoretical data obtained by bioinformatics tools reduced the risk of failure of the experimental approaches and optimized resources for generating on the FC and FCP proteins (Luscombe et al., 2001).

The FimH, CsgA, PapG, FC, and FCP proteins in inclusion bodies were purified under denaturing conditions. Several proteins in inclusion bodies have been purified using chaotropic agents, such as urea and GdnHCl (Wang et al., 2007; Seras-Franzoso et al., 2015). Purified CsgA protein was visualized on SDS-PAGE gels only after formic acid treatment, indicating the formation of amyloid functional aggregates. DLS analysis revealed a polydispersity index (PdI) of 0.948 and a Z-Average



size of 2.97×10^4 nm, which confirmed the aggregation rate of the CsgA-CsgA interaction. The PdI and Z-Average values of the FC protein were greater than those of the FCP protein, which indicates that CsgA linked to FimH and/or PapG promotes low aggregate formation by the fusion proteins because the domains of the CsgA proteins were not exposed to interactions with other CsgA monomers. Therefore, these fusion proteins do not require pretreatment with formic acid for visualization in SDS-PAGE gels.

Fusion proteins with bioactivity that depends on separation (linker), stability, and folding have been described (Chen et al., 2013). Fimbriae are involved in cytokine release by pathogen-associated molecular patterns (PAMPs), which are capable of recognizing pattern recognition receptors (PRRs) such as TLR4 in the mucosa of the urinary tract (Sirard et al., 2006). We suggest a dependence of the fusion protein-PPR interaction with regard to the correct structural conformation to induce a pro-inflammatory response, as described in other models (Akira et al., 2001). A recent study demonstrated that Curli fibers (CsgA) bind to the TLR2-TLR1 dimer to produce IL-6 release from marrow-derived macrophages (Rapsinski et al., 2015). Based on these results, the presence of TLR2 and TLR4 in HTB5 bladder cells from anaplastic transitional cell carcinoma was described for the first time in this study (Nayak et al., 1977). In addition, our data showed that the FimH and PapG proteins could induce release of IL-6 and IL-8, which bind specifically to TLR4 present on immune cells and the surface of uroepithelial cells (Frendéus et al., 2001; Fischer et al., 2006). Compared to the values obtained for FimH, CsgA, and PapG, the FC and FCP proteins generated in this study induced significant increases in IL-6 and IL-8 release.

Our data showed that the functionality of the fusion proteins is involved in the release of high levels of IL-6 and IL-8 when compared to monomeric proteins, probably due to the synergistic signaling of TLR2 and TLR4, an example UPEC strains which have been linked to the activation of TLR4 and TLR5 pathways, thereby initiating innate immune responses against UTIs (Cheng et al., 2016). Conversely, fusion proteins may be considered PAMPs that are capable of recognizing PRRs as proteins with adjuvant-type characteristics (Huleatt et al., 2007).

Potential antigens for ideal vaccines should have features such as surface exposure, expression during infection and immunogenic properties (Sivick and Mobley, 2010; Brumbaugh and Mobley, 2012; Mobley and Alteri, 2015). FimH, CsgA, and PapG proteins are adhesins located on the UPEC surface (Antão et al., 2009). An immunoproteomic approach has been used to identify target antigens and viable targets for vaccine development (Hagan and Mobley, 2007). An immunoproteomic test based on ELISA analysis of sera and urine from UTI patients using the fusion proteins was used to determine the antigenicity and *in vivo* expression of FimH, PapG, and CsgA. Using the FimH, CsgA, PapG, FC, and FCP proteins, higher levels of UP-IgA antibodies were detected in both sera and urine compared to HP-IgA and UP-IgG antibodies. These data are consistent with studies using other UPEC proteins (Rene et al., 1982; Rene and Silverblatt, 1982). Additionally, the antibodies detected in serum and urine also confirmed FimH, CsgA, and PapG protein expression *in vivo* during UTIs caused by UPEC. FC protein contains exposed epitopes that were recognized by antibodies in sera and urine, as confirmed by bioinformatics analysis. However, less robust antigenicity was observed for FCP than FC protein, likely due to the structural complexity of the former, which hides potential epitopes. The titration curves for polyclonal rabbit antibodies against the FCP protein revealed titers $> 1:100,000$; however, the anti-CsgA antibodies had titers $< 1:1000$, indicating that CsgA may be hidden by 3D folding of the FimH and PapG proteins (data not shown). A significant reduction in bacterial adherence by strain CFT073 to HTB5 bladder cells was observed with the anti-PapG, anti-FC, and anti-FCP polyclonal rabbit antibodies compared to pre-immune sera and without antibodies.

In conclusion, recombinant fusion proteins designated FC and FCP were designed and selected for bioinformatics analysis and were observed to significantly induce the release of IL-6 and IL-8. In addition, these proteins exhibited antigenicity based on the detection of IgG and IgA antibodies in the sera and urine of UTI patients and the protective functions of antibodies targeting the fusion proteins *in vitro*. These data suggest that FC and FCP proteins are potential biomolecules for generating vaccines against UTIs caused by UPEC. However, immunogenicity studies of the FC and FCP proteins are needed to achieve the goal of developing a vaccine.

AUTHOR CONTRIBUTIONS

Designed and conceived the experiments: VL, JR, and JX. Performed the experiments: VL, ZS, CM,

and VC. Analyzed the data: VL, AC, LM, and JX. Contributed reagents/materials/analysis tools: SO, AC, JA, RH, and JX. Wrote and reviewed the manuscript: VL and JX.

FUNDING

Public Federal Funds grant number HIM/2014/022 and HIM/2016/027 from the HIMFG supported this work.

REFERENCES

- Akira, S., Takeda, K., and Kaisho, T. (2001). Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat. Immunol.* 2, 675–680. doi: 10.1038/90609
- Andrews, B., Adari, H., Hannig, G., Lahue, E., Gosselin, M., Martin, S., et al. (1996). A tightly regulated high level expression vector that utilizes a thermosensitive lac repressor: production of the human T cell receptor V beta 5.3 in *Escherichia coli*. *Gene* 182, 101–109. doi: 10.1016/S0378-1119(96)00523-9
- Antão, E.-M., Wieler, L. H., and Ewers, C. (2009). Adhesive threads of extraintestinal pathogenic *Escherichia coli*. *Gut Pathog.* 1:22. doi: 10.1186/1757-4749-1-22
- Arai, R., Wriggers, W., Nishikawa, Y., Nagamune, T., and Fujisawa, T. (2004). Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering. *Proteins* 57, 829–838. doi: 10.1002/prot.20244
- Asadi Karam, M. R., Oloomi, M., Mahdavi, M., Habibi, M., and Bouzari, S. (2013). Vaccination with recombinant FimH fused with flagellin enhances cellular and humoral immunity against urinary tract infection in mice. *Vaccine* 31, 1210–1216. doi: 10.1016/j.vaccine.2012.12.059
- Biedenbach, D. J., Badal, R. E., Huang, M.-Y., Motyl, M., Singhal, P. K., Kozlov, R. S., et al. (2016). *In Vitro* activity of oral antimicrobial agents against pathogens associated with community-acquired upper respiratory tract and urinary tract infections: a five country surveillance study. *Infect. Dis. Ther.* 5, 139–153. doi: 10.1007/s40121-016-0112-3
- Brinkmann, U., Mattes, R. E., and Buckel, P. (1989). High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product. *Gene* 85, 109–114. doi: 10.1016/0378-1119(89)90470-8
- Brumbaugh, A. R., and Mobley, H. L. T. (2012). Preventing urinary tract infection: progress toward an effective *Escherichia coli* vaccine. *Expert Rev. Vaccines* 11, 663–676. doi: 10.1586/erv.12.36
- Chen, J., Liu, H., Yang, J., and Chou, K.-C. (2007). Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. *Amino Acids* 33, 423–428. doi: 10.1007/s00726-006-0485-9
- Chen, X., Zaro, J. L., and Shen, W.-C. (2013). Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.* 65, 1357–1369. doi: 10.1016/j.addr.2012.09.039
- Cheng, L., Chen, Q., Zhu, W., Wu, H., Wang, Q., Shi, L., et al. (2016). Toll-like receptors 4 and 5 cooperatively initiate the innate immune responses to uropathogenic *E. coli* infection in mouse epididymal epithelial cells. *Biol. Reprod.* 94:58. doi: 10.1095/biolreprod.115.136580
- Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645. doi: 10.1073/pnas.120163297
- De Groot, A. S., and Moise, L. (2007). New tools, new approaches and new ideas for vaccine development. *Expert Rev. Vaccines* 6, 125–127. doi: 10.1586/14760584.6.2.125
- De Groot, A. S., Sbai, H., Aubin, C. S., McMurry, J., and Martin, W. (2002). Immuno-informatics: mining genomes for vaccine components. *Immunol. Cell Biol.* 80, 255–269. doi: 10.1046/j.1440-1711.2002.01092.x
- Fischer, H., Yamamoto, M., Akira, S., Beutler, B., and Svanborg, C. (2006). Mechanism of pathogen-specific TLR4 activation in the mucosa: fimbriae, recognition receptors and adaptor protein selection. *Eur. J. Immunol.* 36, 267–277. doi: 10.1002/eji.200535149

ACKNOWLEDGMENTS

We thank Armando Cruz Rangel and Julio Cesar Reyes Cruz for technical assistance and Norma Espinosa Sánchez, Héctor Malagón Rivero, and Bertha Gonzalez Pedrajo for their help with the generation of polyclonal rabbit antibodies. VL was supported by CONACyT doctoral fellowship number 261764, and the data in this study are part of his doctoral dissertation in the Programa de Doctorado en Ciencias Biomédicas at the Universidad Nacional Autónoma de México (UNAM).

- Flores-Mireles, A. L., Walker, J. N., Caparon, M., and Hultgren, S. J. (2015). Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat. Rev. Microbiol.* 13, 269–284. doi: 10.1038/nrmicro3432
- Foxman, B. (2002). Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *Am. J. Med.* 113(Suppl. 1A), 5S–13S. doi: 10.1016/s0002-9343(02)01054-9
- Foxman, B. (2010). The epidemiology of urinary tract infection. *Nat. Rev. Urol.* 7, 653–660. doi: 10.1038/nrurol.2010.190
- Foxman, B. (2014). Urinary tract infection syndromes: occurrence, recurrence, bacteriology, risk factors, and disease burden. *Infect. Dis. Clin. North Am.* 28, 1–13. doi: 10.1016/j.idc.2013.09.003
- Frendéus, B., Wachtler, C., Hedlund, M., Fischer, H., Samuelsson, P., Svensson, M., et al. (2001). *Escherichia coli* P fimbriae utilize the Toll-like receptor 4 pathway for cell activation. *Mol. Microbiol.* 40, 37–51. doi: 10.1046/j.1365-2958.2001.02361.x
- Golshani, M., Rafati, S., Jahanian-Najafabadi, A., Nejati-Moheimani, M., Siadat, S. D., Shahcheraghi, F., et al. (2015). *In silico* design, cloning and high level expression of L7/L12-TOMP31 fusion protein of *Brucella* antigens. *Res. Pharm. Sci.* 10, 436–445.
- Greene, S. E., Hibbing, M. E., Janetka, J., Chen, S. L., and Hultgren, S. J. (2015). Human urine decreases function and expression of type 1 pili in uropathogenic *Escherichia coli*. *MBio* 6:e00820. doi: 10.1128/mBio.00820-15
- Habibi, M., Asadi Karam, M. R., and Bouzari, S. (2015a). *In silico* design of fusion protein of FimH from uropathogenic *Escherichia coli* and MrpH from *Proteus mirabilis* against urinary tract infections. *Adv. Biomed. Res.* 4, 217. doi: 10.4103/2277-9175.166164
- Habibi, M., Asadi Karam, M. R., Shokrgozar, M. A., Oloomi, M., Jafari, A., and Bouzari, S. (2015b). Intranasal immunization with fusion protein MrpH-FimH and MPL adjuvant confers protection against urinary tract infections caused by uropathogenic *Escherichia coli* and *Proteus mirabilis*. *Mol. Immunol.* 64, 285–294. doi: 10.1016/j.molimm.2014.12.008
- Hagan, E. C., and Mobley, H. L. T. (2007). Uropathogenic *Escherichia coli* outer membrane antigens expressed during urinary tract infection. *Infect. Immun.* 75, 3941–3949. doi: 10.1128/IAI.00337-07
- Hannan, T. J., and Hunstad, D. A. (2016). A murine model for *Escherichia coli* urinary tract infection. *Methods Mol. Biol.* 1333, 159–175. doi: 10.1007/978-1-4939-2854-5_14
- Huang, Z., Li, G., Zhang, C., and Xing, X.-H. (2016). A study on the effects of linker flexibility on acid phosphatase PhoC-GFP fusion protein using a novel linker library. *Enzyme Microb. Technol.* 83, 1–6. doi: 10.1016/j.enzmictec.2015.11.002
- Huleatt, J. W., Jacobs, A. R., Tang, J., Desai, P., Kopp, E. B., Huang, Y., et al. (2007). Vaccination with recombinant fusion proteins incorporating Toll-like receptor ligands induces rapid cellular and humoral immunity. *Vaccine* 25, 763–775. doi: 10.1016/j.vaccine.2006.08.013
- Hung, C., Marschall, J., Burnham, C.-A. D., Byun, A. S., and Henderson, J. P. (2014). The bacterial amyloid curli is associated with urinary source bloodstream infection. *PLoS ONE* 9:e86009. doi: 10.1371/journal.pone.0086009
- Hung, C.-S., Bouckaert, J., Hung, D., Pinkner, J., Widberg, C., DeFusco, A., et al. (2002). Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol. Microbiol.* 44, 903–915. doi: 10.1046/j.1365-2958.2002.02915.x
- Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. *J. Biochem.* 88, 1895–1898.

- Janin, J., and Chothia, C. (1990). The structure of protein-protein recognition sites. *J. Biol. Chem.* 265, 16027–16030.
- Kumar, S., Dave, A., Wolf, B., and Lerma, E. V. (2015). Urinary tract infections. *Dis. Mon.* 61, 45–59. doi: 10.1016/j.disamonth.2014.12.002
- Lane, M. C., and Mobley, H. L. T. (2007). Role of P-fimbrial-mediated adherence in pyelonephritis and persistence of uropathogenic *Escherichia coli* (UPEC) in the mammalian kidney. *Kidney Int.* 72, 19–25. doi: 10.1038/sj.ki.5002230
- Langermann, S., and Ballou, W. R. (2001). Vaccination utilizing the FimCH complex as a strategy to prevent *Escherichia coli* urinary tract infections. *J. Infect. Dis.* 183(Suppl. 1), S84–S86. doi: 10.1086/318857
- Langermann, S., Möllby, R., Burlein, J. E., Palaszynski, S. R., Auguste, C. G., DeFusco, A., et al. (2000). Vaccination with FimH adhesin protects cynomolgus monkeys from colonization and infection by uropathogenic *Escherichia coli*. *J. Infect. Dis.* 181, 774–778. doi: 10.1086/315258
- Langermann, S., Palaszynski, S., Barnhart, M., Auguste, G., Pinkner, J. S., Burlein, J., et al. (1997). Prevention of mucosal *Escherichia coli* infection by FimH-adhesin-based systemic vaccination. *Science* 276, 607–611. doi: 10.1126/science.276.5312.607
- Laskowski, R. A., MacArthur, M. W., and Thornton, J. M. (1998). Validation of protein models derived from experiment. *Curr. Opin. Struct. Biol.* 8, 631–639. doi: 10.1016/S0959-440X(98)80156-5
- Leatham-Jensen, M. P., Mokszycki, M. E., Rowley, D. C., Deering, R., Camberg, J. L., Sokurenko, E. V., et al. (2016). Uropathogenic *Escherichia coli* metabolite-dependent quiescence and persistence may explain antibiotic tolerance during urinary tract infection. *mSphere* 1:e00055-15. doi: 10.1128/mSphere.00055-15
- Ledesma, M. A., Ochoa, S. A., Cruz, A., Rocha-Ramírez, L. M., Mas-Oliva, J., Eslava, C. A., et al. (2010). The hemorrhagic coli pilus (HCP) of *Escherichia coli* O157:H7 is an inducer of proinflammatory cytokine secretion in intestinal epithelial cells. *PLoS ONE* 5:e12127. doi: 10.1371/journal.pone.0012127
- Li, G., Huang, Z., Zhang, C., Dong, B.-J., Guo, R.-H., Yue, H.-W., et al. (2016). Construction of a linker library with widely controllable flexibility for fusion protein design. *Appl. Microbiol. Biotechnol.* 100, 215–225. doi: 10.1007/s00253-015-6985-3
- Lim, J. Y., Pinkner, J. S., and Cegelski, L. (2014). Community behavior and amyloid-associated phenotypes among a panel of uropathogenic *E. coli*. *Biochem. Biophys. Res. Commun.* 443, 345–350. doi: 10.1016/j.bbrc.2013.11.026
- Luscombe, N. M., Greenbaum, D., and Gerstein, M. (2001). What is bioinformatics? A proposed definition and overview of the field. *Methods Inform. Med.* 40, 346–358.
- Lüthje, P., and Brauner, A. (2014). Virulence factors of uropathogenic *E. coli* and their interaction with the host. *Adv. Microb. Physiol.* 65, 337–372. doi: 10.1016/bbs.2014.08.006
- Melzer, M., and Petersen, I. (2007). Mortality following bacteraemic infection caused by extended spectrum beta-lactamase (ESBL) producing *E. coli* compared to non-ESBL producing *E. coli*. *J. Infect.* 55, 254–259. doi: 10.1016/j.jinf.2007.04.007
- Mobley, H. L. T., and Alteri, C. J. (2015). Development of a vaccine against *Escherichia coli* urinary tract infections. *Pathogens* 5:E1. doi: 10.3390/pathogens5010001
- Mohammad, N., Karsabet, M. T., Amani, J., Ardjmand, A., Zadeh, M. R., Gholi, M. K., et al. (2016). *In silico* design of a chimeric protein containing antigenic fragments of *Helicobacter pylori*; a bioinformatic approach. *Open Microbiol. J.* 10, 97–112. doi: 10.2174/1874285801610010097
- Molina-López, J., Aparicio-Ozores, G., Ribas-Aparicio, R. M., Gavilanes-Parra, S., Chávez-Berocal, M. E., Hernández-Castro, R., et al. (2011). Drug resistance, serotypes, and phylogenetic groups among uropathogenic *Escherichia coli* including O25-ST131 in Mexico City. *J. Infect. Dev. Ctries.* 5, 840–849. doi: 10.3855/jidc.1703
- Moriel, D. G., and Schembri, M. A. (2013). Vaccination approaches for the prevention of urinary tract infection. *Curr. Pharm. Biotechnol.* 14, 967–974. doi: 10.2174/1389201014666131226144824
- Mulvey, M. A., Lopez-Boado, Y. S., Wilson, C. L., Roth, R., Parks, W. C., Heuser, J., et al. (1998). Induction and evasion of host defenses by type 1-piliated uropathogenic *Escherichia coli*. *Science* 282, 1494–1497. doi: 10.1126/science.282.5393.1494
- Mulvey, M. A., Schilling, J. D., Martinez, J. J., and Hultgren, S. J. (2000). Bad bugs and beleaguered bladders: interplay between uropathogenic *Escherichia coli* and innate host defenses. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8829–8835. doi: 10.1073/pnas.97.16.8829
- Muñoz-Medina, J. E., Sánchez-Vallejo, C. J., Méndez-Tenorio, A., Monroy-Muñoz, I. E., Angeles-Martínez, J., Santos Coy-Arechavaleta, A., et al. (2015). *In silico* identification of highly conserved epitopes of influenza A H1N1, H2N2, H3N2, and H5N1 with diagnostic and vaccination potential. *Biomed Res. Int.* 2015:813047. doi: 10.1155/2015/813047
- Nayak, S. K., O'Toole, C., and Price, Z. H. (1977). A cell line from an anaplastic transitional cell carcinoma of human urinary bladder. *Br. J. Cancer* 35, 142–151.
- Nazarian, S., Mousavi Gargari, S. L., Rasooli, I., Amani, J., Bagheri, S., and Alerasool, M. (2012). An *in silico* chimeric multi subunit vaccine targeting virulence factors of enterotoxigenic *Escherichia coli* (ETEC) with its bacterial inbuilt adjuvant. *J. Microbiol. Methods* 90, 36–45. doi: 10.1016/j.mimet.2012.04.001
- Norinder, B. S., Köves, B., Yadav, M., Brauner, A., and Svanborg, C. (2012). Do *Escherichia coli* strains causing acute cystitis have a distinct virulence repertoire? *Microb. Pathog.* 52, 10–16. doi: 10.1016/j.micpath.2011.08.005
- Olsén, A., Herwald, H., Wikström, M., Persson, K., Mattsson, E., and Björck, L. (2002). Identification of two protein-binding and functional regions of curli, a surface organelle and virulence determinant of *Escherichia coli*. *J. Biol. Chem.* 277, 34568–34572. doi: 10.1074/jbc.M206353200
- Pedersen-Lane, J., Maley, G. F., Chu, E., and Maley, F. (1997). High-level expression of human thymidylate synthase. *Protein Expr. Purif.* 10, 256–262. doi: 10.1006/prep.1997.0750
- Pedretti, A., Villa, L., and Vistoli, G. (2004). VEGA—an open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. *J. Comput. Aided Mol. Des.* 18, 167–173. doi: 10.1023/B:JCAM.0000035186.90683.f2
- Puigbò, P., Guzmán, E., Romeu, A., and Garcia-Vallvé, S. (2007). OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res.* 35, W126–W131. doi: 10.1093/nar/gkm219
- Rapsinski, G. J., Wynosky-Dolfi, M. A., Oppong, G. O., Tursi, S. A., Wilson, R. P., Brodsky, I. E., et al. (2015). Toll-like receptor 2 and NLRP3 cooperate to recognize a functional bacterial amyloid, curli. *Infect. Immun.* 83, 693–701. doi: 10.1128/IAI.02370-14
- Rene, P., Dinolfo, M., and Silverblatt, F. J. (1982). Serum and urogenital antibody responses to *Escherichia coli* pili in cystitis. *Infect. Immun.* 38, 542–547.
- Rene, P., and Silverblatt, F. J. (1982). Serological response to *Escherichia coli* pili in pyelonephritis. *Infect. Immun.* 37, 749–754.
- Saldaña, Z., Xicohtencatl-Cortes, J., Avelino, F., Phillips, A. D., Kaper, J. B., Puente, J. L., et al. (2009). Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli. *Environ. Microbiol.* 11, 992–1006. doi: 10.1111/j.1462-2920.2008.01824.x
- Savar, N. S., Jahanian-Najafabadi, A., Mahdavi, M., Shokrgozar, M. A., Jafari, A., and Bouzari, S. (2014). *In silico* and *in vivo* studies of truncated forms of flagellin (FliC) of enteroaggregative *Escherichia coli* fused to FimH from uropathogenic *Escherichia coli* as a vaccine candidate against urinary tract infections. *J. Biotechnol.* 175, 31–37. doi: 10.1016/j.jbiotec.2014.01.037
- Scott, V. C. S., Haake, D. A., Churchill, B. M., Justice, S. S., and Kim, J.-H. (2015). Intracellular bacterial communities: a potential etiology for chronic lower urinary tract symptoms. *Urology* 86, 425–431. doi: 10.1016/j.urology.2015.04.002
- Secretaria de Salud (2009). *Panorama Epidemiológico de las Infecciones de Vías Urinarias en México 2003-2008*. Available online at: <http://www.epidemiologia.salud.gob.mx/doctos/boletin/2009/sem51.pdf>
- Sen, T. Z., Jernigan, R. L., Garnier, J., and Kloczkowski, A. (2005). GOR V server for protein secondary structure prediction. *Bioinformatics* 21, 2787–2788. doi: 10.1093/bioinformatics/bti408
- Seras-Franzoso, J., Peternel, S., Cano-Garrido, O., Villaverde, A., and García-Fruitós, E. (2015). Bacterial inclusion body purification. *Methods Mol. Biol.* 1258, 293–305. doi: 10.1007/978-1-4939-2205-5_16
- Serino, L., Moriel, D. G., Rappuoli, R., and Pizza, M. (2010). Towards a vaccine against *Escherichia coli*-associated urinary tract infections. *Future Microbiol.* 5, 351–354. doi: 10.2217/fmb.10.6
- Shewmaker, F., McGlinchey, R. P., Thurber, K. R., McPhie, P., Dyda, F., Tycko, R., et al. (2009). The functional curli amyloid is not based on

- in-register parallel beta-sheet structure. *J. Biol. Chem.* 284, 25065–25076. doi: 10.1074/jbc.M109.007054
- Sim, B. C., Zerva, L., Greene, M. I., and Gascoigne, N. R. (1996). Control of MHC restriction by TCR Valpha CDR1 and CDR2. *Science* 273, 963–966. doi: 10.1126/science.273.5277.963
- Sirard, J.-C., Bayardo, M., and Didierlaurent, A. (2006). Pathogen-specific TLR signaling in mucosa: mutual contribution of microbial TLR agonists and virulence factors. *Eur. J. Immunol.* 36, 260–263. doi: 10.1002/eji.200535777
- Sivick, K. E., and Mobley, H. L. T. (2010). Waging war against uropathogenic *Escherichia coli*: winning back the urinary tract. *Infect. Immun.* 78, 568–585. doi: 10.1128/IAI.01000-09
- Snyder, J. A., Haugen, B. J., Lockatell, C. V., Maroncle, N., Hagan, E. C., Johnson, D. E., et al. (2005). Coordinate expression of fimbriae in uropathogenic *Escherichia coli*. *Infect. Immun.* 73, 7588–7596. doi: 10.1128/IAI.73.11.7588-7596.2005
- Spaulding, C. N., and Hultgren, S. J. (2016). Adhesive pili in UTI pathogenesis and drug development. *Pathogens* 5:E30. doi: 10.3390/pathogens5010030
- Sung, M.-A., Fleming, K., Chen, H. A., and Matthews, S. (2001). The solution structure of PapGII from uropathogenic *Escherichia coli* and its recognition of glycolipid receptors. *EMBO Rep.* 2, 621–627. doi: 10.1093/embo-reports/kve133
- Thankavel, K., Madison, B., Ikeda, T., Malaviya, R., Shah, A. H., Arumugam, P. M., et al. (1997). Localization of a domain in the FimH adhesin of *Escherichia coli* type 1 fimbriae capable of receptor recognition and use of a domain-specific antibody to confer protection against experimental urinary tract infection. *J. Clin. Invest.* 100, 1123–1136. doi: 10.1172/JCI119623
- Wang, X., Smith, D. R., Jones, J. W., and Chapman, M. R. (2007). *In vitro* polymerization of a functional *Escherichia coli* amyloid protein. *J. Biol. Chem.* 282, 3713–3719. doi: 10.1074/jbc.M609228200
- Wiederstein, M., and Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35, W407–W410. doi: 10.1093/nar/gkm290
- Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., et al. (1999). Protein identification and analysis tools in the ExpASY server. *Methods Mol. Biol.* 112, 531–552.
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* 12, 7–8. doi: 10.1038/nmeth.3213
- Zagrovic, B., Jayachandran, G., Millett, I. S., Doniach, S., and Pande, V. S. (2005). How large is an alpha-helix? Studies of the radii of gyration of helical peptides by small-angle X-ray scattering and molecular dynamics. *J. Mol. Biol.* 353, 232–241. doi: 10.1016/j.jmb.2005.08.053
- Zhang, Y., and Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* 33, 2302–2309. doi: 10.1093/nar/gki524
- Zhou, G., Mo, W. J., Sebbel, P., Min, G., Neubert, T. A., Glockshuber, R., et al. (2001). Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: evidence from *in vitro* FimH binding. *J. Cell Sci.* 114(Pt 22), 4095–4103.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Luna-Pineda, Reyes-Grajeda, Cruz-Córdova, Saldaña-Ahuactzi, Ochoa, Maldonado-Bernal, Cázares-Domínguez, Moreno-Fierros, Arellano-Galindo, Hernández-Castro and Xicohtencatl-Cortés. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.