



Industria y Comercio

SUPERINTENDENCIA

Delegatura de propiedad Industrial

División de Nuevas Creaciones

PCT SOLICITUD PATENTE DE INVENCION

21. EXPEDIENTE No.

06-52194

54. TÍTULO

FACTORES DE VIRULENCIA BACTERIANA Y
SUS USOS

51. CLASIFICACIÓN INTERNACIONAL

C12N15/31, 1/20, A61K38/16, 39/10
35/74, 39/02, C12Q1/68, E C1237/569, C07K14/24, A61P31/00, 37/02.

71. SOLICITANTE

THE UNIVERSITY OF BRITISH COLUMBIA / UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO

DOMICILIO

BRITISH COLUMBIA, CANADA / MEXICO D.F., MEXICO

74. APODERADO

MONICA WOLF WASSERMAN

22. BOGOTÁ, D. C.,

PETITORIO

Industria y Comercio
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
Rad: 06-052184-00000-0000
Feb. 2006-05-31 16:18:44 Dep. 2020 NUECREACIONES
Trib. 11 PATENTE Y B.
Eve: 375 FASE NACIONAL
Act 411 PRESENTACION Folios: 199

FORMULARIO ÚNICO DE SOLICITUD DE PATENTE PCT

ENTRADA EN FASE NACIONAL

31-05-06

①

SOLICITUD DE :

☒ Patente de Invención.

☐ Patente de Modelo de Utilidad.

☒ Capítulo I.

☐ Capítulo II.

②

SOLICITANTE
(71)

Nombre: VER ANEXO

Dirección: VER ANEXO

Nacionalidad o Domicilio: VER ANEXO

Lugar de Constitución: VER ANEXO

Teléfono: VER ANEXO Fax: VER ANEXO

E-mail:

IDENTIFICACIÓN

C.C. ☐ NT ☐

C.E. ☐ Otro ☐

Cual _____

Número _____

③

REPRESENTANTE
O APODERADO

Nombre: MONICA WOLF WASSERMAN

Dirección: Diagonal 83 A # 23 - 90

Teléfono: 623 58 28 Fax: 617 06 14

E-mail:

IDENTIFICACIÓN

C.C. ☒ NT ☐

C.E. ☐ Otro ☐

Cual _____

Número 51.711.267 TP46.034

④

INVENTOR (ES)
(72)

Nombre:

VER ANEXO

Dirección:

Nacionalidad o Domicilio

⑤ PCT (86)

Solicitud Internacional No. PCT/CA2004/001891 Fecha Octubre 29/2004

Publicación Internacional No. WO 2005/042746 A1 Fecha Mayo 12/2005

⑥ Título (54)

FACTORES DE VIRULENCIA BACTERIANA Y SUS USOS.

⑦

Clasificación Internacional (51) C12N15/31, A61K38/16, 39/108, 35/179, C12W1/29, C12Q1/68, G01N33/569, C07K14/24, A61K39/02, A61P31/00, 37/02

⑧

Prioridad

SI ☒ NO ☐

(33) País de Origen

ESTADOS

UNIDOS

DE

AMERICA

(32) Fecha

Octubre 31

de 2003

(31) Número de Solicitud

60/515,703

⑨

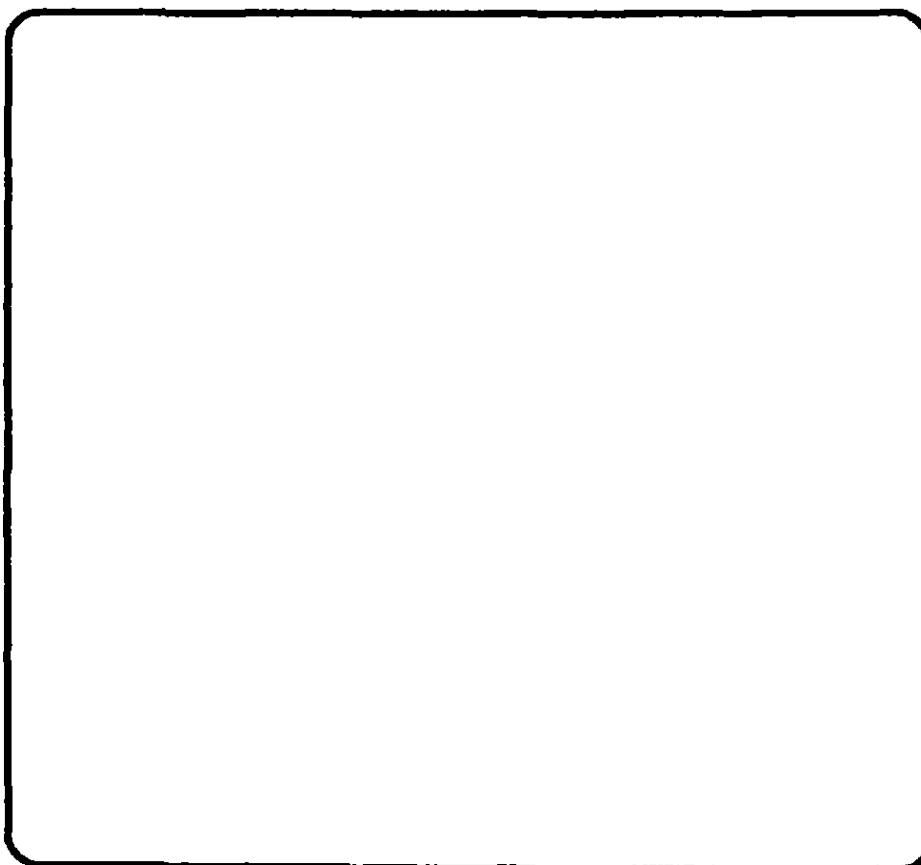
Comprobante de pago No.

Fecha Mayo /2006

Instrucciones para el diligenciamiento del presente formulario, ver reverso de esta página.

- ☒ Comprobante de pago de la tasa de presentación de la solicitud.
- ☐ Documento que acredite la existencia y representación legal cuando el solicitante sea persona jurídica.
- ☐ Poderes, si fuere el caso.
- ☐ Documento de cesión del inventor al solicitante o a su causante.
- ☐ Declaraciones
- ☐ Copia de la solicitud internacional. (Resumen, descripción, reivindicaciones, dibujos)
- ☒ Traducción de la solicitud internacional al castellano. (Resumen, descripción, reivindicaciones, dibujos)
- ☐ Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen Preliminar (IPEA).
- ☐ Traducción del examen preliminar internacional efectuado por la IPEA.
- ☐ De ser el caso, copia del contrato de acceso.
- ☐ De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas.
- ☐ De ser el caso, certificado de depósito del material biológico
- ☐ Arte final 12x12.
- ☐ De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- ☐ De ser el caso, modificaciones efectuadas a la solicitud internacional.
- ☐ Otros

11 FIGURA CARACTERÍSTICA:



12

Solicito la concesión de la patente,

NOMBRE: MONICA HOLE WASSERMAN

FIRMA:

CC: 51.711.762 / TP: 16.03A

ANEXOS

Anexo Solicitante

Nombre: The University of British Columbia (CA/CA)
Dirección: #103 – 6190 Agronomy Road, Vancouver, British Columbia V6T 1Z3, Canadá (CA)
Domicilio: British Columbia, Canadá.

Nombre: Universidad Nacional Autónoma de México, (MX/MX)
Dirección: 9º Piso de la Torre de Rectoría, Ciudad Universitaria, Delegación de Coyoacán, C.P. 04510, México, D.F. México (MX)
Domicilio: México D.F., México.

Anexo Inventores

Nombre: FINLAY, Brett, (CA/CA).
Dirección: 8461 Seafair Drive, Richmond, British Columbia V7C 1X7, (CA) Canada.
Domicilio: British Columbia, Canadá.

Nombre: GRUENHEID, Samantha.
Dirección: (CA/CA) 2316 Balaclava Street, Vancouver, British Columbia V6K 4C7, (CA) Canada.
Domicilio: British Columbia, Canadá.

Nombre: DENG, Wanyin, (CA/CA)
Dirección: 6819 Walker Avenue, Burnaby, British Columbia V5E 3C3, (CA) Canadá.
Domicilio: British Columbia, Canadá.

Nombre: VALLANCE, Bruce, (CA/CA).
Dirección: #603, 1305 West 12th Avenue, Vancouver, British Columbia V6H 1M3, (CA) Canadá.
Domicilio: British Columbia, Canadá.

Nombre: PUENTE, José, L. (MX/MX).
Dirección: Privada de Cananea 8 – 2, Col. Lomas de la Selva, Cuernavaca, Mor. 62250, (MX) México.
Domicilio: Cuernavaca, México.



(43) International Publication Date
12 May 2005 (12.05.2005)

PCT

(10) International Publication Number
WO 2005/042746 A1

(51) International Patent Classification: C12N 15/31,
A61K 38/16, 39/108, 35/74, C12N 1/20, C12Q 1/68,
G01N 33/569, C07K 14/24, A61K 39/02, A61P 31/00

(74) Agents: ROBINSON, Christopher, J. et al.; Smart &
Biggar, Box 11560 Vancouver Centre, Suite 2200, 650
West Georgia Street, Vancouver, British Columbia V6B
4N8 (CA).

(21) International Application Number:
PCT/CA2004/001891

(23) International Filing Date: 29 October 2004 (29.10.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/515,703 31 October 2003 (31.10.2003) US

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AI, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, GR, GU, HT, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,
ZW.

(71) Applicants (for all designated States except US): THE
UNIVERSITY OF BRITISH COLUMBIA [CA/CA];
#103-6190 Agronomy Road, Vancouver, British Columbia
V6T 1Z3 (CA). UNIVERSIDAD NACIONAL AU-
TONOMA DE MEXICO [MX/MX]; 9° Piso de la Torre
De Rectoria, Ciudad Universitaria, Delegación Coyoacán,
C.P. 04510, México, D.F. México (MX).

(72) Inventors; and

(73) Inventors/Applicants (for US only): FINLAY, Brett
[CA/CA]; 8491 Seafair Drive, Richmond, British Colum-
bia V7C 1X7 (CA). GRUNHEID, Samantha [CA/CA];
2316 Balclava Street, Vancouver, British Columbia
V6K 4C7 (CA). DENG, Wanyin [CA/CA]; 6819 Walker
Avenue, Burnaby, British Columbia V5E 3C3 (CA). VAL-
LANCE, Bruce [CA/CA]; #603, 1305 West 12th Avenue,
Vancouver, British Columbia V6H 1M3 (CA). PUENTE,
Jose, L. [MX/MX]; Privada de Cañones 8-2, Col. Lomas
de la Selva, Cuernavaca, Mor. 62250 (MX).

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IL, IT, LU, MC, NL, PL, PT, RO, SE, SI,
SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TO).

Published:

- with international search report
- before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: BACTERIAL VIRULENCE FACTORS AND USES THEREOF

(57) Abstract: The invention relates to, in part, secreted proteins of bacterial pathogens and methods for their use. More specifically, the invention provides in part several new common secreted proteins for A/E pathogens. In some embodiments of the invention, these polypeptides and nucleic acid molecules encoding these polypeptides, or portions thereof, are useful as vaccines, diagnostics, or drug screening tools for A/E pathogenic infections, or as reagents.

WO 2005/042746 A1

FACTORES DE VIRULENCIA BACTERIANA Y SUS USOS

CAMPO DE LA INVENCION

La invención se relaciona a agentes patógenos bacterianos. Más específicamente, la invención se relaciona, en parte, a proteínas de secreción de agentes patógenos bacterianos y a métodos para su uso.

ANTECEDENTES DE LA INVENCION

Escherichia coli es un organismo extremadamente versátil. Además de ser un miembro de la flora intestinal normal, ciertas cepas de *E. coli* también provocan infecciones de la vesícula, meningitis y diarrea. Las *E. coli* diarreagénicas incluyen al menos cinco tipos de *E. coli*, las cuales causan varios síntomas que van desde la diarrea de tipo colérica hasta la colitis extrema. Cada uno de los tipos de *E. coli* diarreagénica posee un conjunto particular de factores de virulencia que incluyen adhesinas, invasinas y/o toxinas, las cuales son responsables de provocar un tipo específico de diarrea.

La *E. coli* enteropatogénica (EPEC) es causa predominante de diarrea infantil en todo el mundo. El mal de EPEC se caracteriza por una diarrea acuosa de severidad variable, con vómitos y fiebre a menudo acompañando la pérdida de fluido. Además de brotes aislados en salas de cuidados diurnos y en guarderías infantiles de países desarrollados, la EPEC plantea una importante amenaza endémica para la salud de niños pequeños (menores de 6 meses) en países en desarrollo. En todo el mundo, la EPEC es la principal causa de diarrea mediada por bacterias en niños pequeños, y se ha estimado que la EPEC mata a varios cientos de miles de niños por año.

La *E. coli* enterohemorrágica (EHEC), también llamada *E. coli* productora de toxina Shiga (STEC), o *E. coli* productora de toxina Vero (VTEC), provoca una diarrea más severa que la EPEC (colitis entérica) y en aproximadamente 10% de los casos esta enfermedad progresa hacia una enfermedad renal a menudo fatal, el síndrome urémico hemolítico (HUS). La EHEC O157:H7 es el serotipo más común en Canadá y en los Estados Unidos, y está asociada al envenenamiento de los alimentos y el agua (3). Otros serotipos de la EHEC también provocan significativos problemas en Asia, Europa,

América del Sur y, en menor grado, en América del Norte. La EHEC coloniza el ganado y causa lesiones A/E, pero no provoca enfermedad en animales adultos, y en cambio propaga organismos en el ambiente. Esto sin embargo causa serios problemas sanitarios, ya que se necesitan relativamente pocas EHEC para infectar a los humanos.

Contrariamente a lo que ocurre con otras diarreas causadas por *E. coli*, tal como la *E. coli* enterotoxigénica, la diarrea causada por EHEC y EPEC no está mediada por una toxina. En cambio, EPEC y EHEC se adhieren a las superficies intestinales (EPEC al intestino delgado, EHEC al intestino grueso) y provocan una lesión histológica característica conocida como lesión de adherencia y destrucción (A/E) (8). Las lesiones A/E se destacan por la disolución de la superficie del "ribete en cepillo" intestinal y la pérdida (destrucción) de microvellosidades epiteliales en los sitios de adherencia de las bacterias. Una vez adheridas, las bacterias residen sobre proyecciones en forma de copa o pedestales. Por debajo de este pedestal, en la célula epitelial hay varios componentes citoesqueléticos, tales como actina y proteínas citoesqueléticas asociadas a la actina. La formación de lesiones A/E y de pedestales ricos en actina por debajo de las bacterias adheridas son el sello distintivo de los agentes patógenos A/E (1, 2). Esta patología puede ser recapitulada en células cultivadas *in vitro*, y la formación de pedestales puede observarse mediante una técnica de tinción fluorescente de la actina (2, 11). La formación de la lesión A/E puede ser responsable de la destrucción del ribete en cepillo y de las microvellosidades, de la secreción de fluido y de la diarrea.

EPEC y EHEC pertenecen a una familia de agentes patógenos A/E que incluye varios agentes patógenos del tipo EPEC de animales, que provocan la enfermedad en conejos (REPEC), cerdos (PEPEC) y ratones (*Citrobacter rodentium*). Estos agentes patógenos contienen islas de patogenicidad (PAIs) que codifican para sistemas de secreción especializados y factores de virulencia secretados que son críticos para la enfermedad. Se cree que los genes requeridos para la formación de lesiones A/E están agrupados en una única isla cromosómica de patogenicidad conocida como locus de destrucción de enterocitos (LEE), el cual incluye elementos reguladores, un sistema de secreción del tipo III (TTSS), proteínas efectoras secretadas, y acompañantes afines. (4-8).

El LEE contiene 41 genes, lo cual lo hace uno de los PAIs más complejos. La función principal del TTSS del LEE es liberar efectores dentro de la célula hospedadora, donde subvierten las funciones de la célula hospedadora e interceden en la enfermedad (9, 10, 34). Se han identificado cinco efectores codificados por el LEE (Tir, EspG, EspF, Map y EspH) (35, 40). El Tir (receptor de la intimina translocada) es translocado dentro de las células hospedadoras, donde se adhiere a proteínas citoesqueléticas y de señalización, y da inicio a la polimerización de la actina en el sitio de adhesión bacteriana (31, 44), dando como resultado la formación de estructuras de actina del pedestal por debajo de las bacterias adheridas, las cuales interactúan directamente con el circuito extracelular del Tir a través de la proteína intimina de la membrana bacteriana externa. El CesT juega un rol de acompañante para la estabilidad y secreción del Tir (18, 19).

Se han caracterizado otros cuatro efectores translocados por el TTSS y codificados por el LEE en agentes patógenos A/E: el EspH promueve el alargamiento de los pedestales de actina (40); el EspF juega un rol en el desarmado de los ajustados empalmes entre las células del epitelio intestinal (38); el EspG está relacionado al efector VirA de unión a microtúbulos Shigella (36, 55); y el Map se localiza en mitocondrias (37) pero también tiene un rol en la dinámica de la actina (48). El Ler (regulador codificado por el LEE) es el único regulador codificado por el LEE que ha sido identificado.

RESUMEN DE LA INVENCION

Esta invención se basa, parcialmente, en la identificación de diversas nuevas proteínas de secreción comunes en agentes patógenos A/E.

La invención provee, en un aspecto, composiciones que incluyen un polipéptido o un fragmento o variante del mismo, o un sobrenadante de cultivo celular que incluye tal polipéptido, donde el polipéptido sustancialmente puro incluye una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera o cualesquiera de las SEQ ID Nos: 22-43, 59 ó 73-84, o a fragmentos o variantes de las mismas. La invención también provee composiciones que incluyen una molécula de ácido nucleico,

donde la molécula de ácido nucleico incluye una secuencia nucleotídica sustancialmente idéntica a la secuencia de cualquiera o cualesquiera de las SEQ ID NO: 22-43, o a fragmentos de las mismas. Las composiciones pueden además incluir un portador fisiológicamente aceptable, o pueden además incluir un adyuvante. Las composiciones también pueden incluir una molécula polipeptídica o de ácido nucleico tal como EspA, EspB, EspD, EspP, Tir, toxina Shiga 1, toxina Shiga 2, o intimina. Los polipéptidos o las moléculas de ácido nucleico pueden ser sustancialmente puras.

La invención también provee, en aspectos alternativos, una bacteria o una preparación de la misma, donde la bacteria incluye una mutación en un gen tal como *nleA*, *nleB*, *nleC*, *nleD*, *nleE*, *nleF*, *nleG*, *nleH*, o un homólogo de los mismos, o incluye una mutación en el genoma bacteriano en una secuencia nucleotídica que es sustancialmente idéntica a las SEQ ID Nos: 1-21 ó 60-72. En algunas realizaciones, la bacteria puede ser un agente patógeno A/E, tal como la *E. coli* enterohemorrágica (EHEC; e.g., EHEC O157:H7 ó EHEC O157:NM), la *E. coli* enteropatogénica (EPEC; e.g., EPEC O127:H6), o *Citrobacter rodentium*. En algunas realizaciones, la mutación puede atenuar la virulencia, o puede estar presente en una secuencia nucleotídica en el genoma del agente patógeno A/E que es sustancialmente idéntica a la secuencia seleccionada a partir del grupo consistente en las SEQ ID NOs: 1-21 ó 60-72. La bacteria puede ser provista como una composición, en combinación con un adyuvante. En algunas realizaciones, la bacteria puede estar viva. En algunas realizaciones, la bacteria puede estar muerta. El modo de administración puede ser oral o parenteral.

La invención también provee, en aspectos alternativos, un método para detectar la presencia de un agente patógeno A/E en una muestra, mediante la provisión de una muestra; y para detectar la presencia de una secuencia nucleotídica sustancialmente idéntica a una secuencia seleccionada a partir de las SEQ ID NOs: 1-21 ó 60-72, ó un fragmento o variante de las mismas; una secuencia nucleotídica codificadora de una secuencia polipeptídica sustancialmente idéntica a una secuencia seleccionada a partir de las SEQ ID NOs: 22-43, 59 ó 73-84, ó un polipéptido que incluya una secuencia de aminoácidos sustancialmente idéntica a una secuencia seleccionada a partir de las SEQ ID NOs: 22-43, 59 ó 73-84, ó un fragmento o variante de las mismas, donde la presencia de la secuencia nucleotídica o la secuencia de aminoácidos indica la

presencia de un agente patógeno A/E en una muestra (e.g., huevo, heces, sangre o intestino). La detección puede incluir contactar la secuencia nucleotídica con una sonda o *primer* sustancialmente idéntico a una secuencia seleccionada a partir del grupo compuesto por las SEQ ID NOs: 1-21 ó 60-72, o una secuencia nucleotídica que codifique para una secuencia polipeptídica sustancialmente idéntica a la secuencia seleccionada a partir del grupo compuesto por las SEQ ID NOs: 22-43, 59 ó 73-84, o una porción de las mismas, o puede incluir contactar la secuencia de aminoácidos con un anticuerpo que se ligue específicamente a una secuencia seleccionada a partir del grupo compuesto por las SEQ ID NOs: 22-43, 59 ó 73-84, ó un fragmento o variante de las mismas.

La invención también provee, en aspectos alternativos, métodos para producir una respuesta inmune contra un agente patógeno A/E o un componente del mismo, o para reducir la colonización o la difusión de un agente patógeno A/E en un animal (e.g., un humano, un rumiante, tal como ovejas –sujeto ovino-, cabras, vacas –sujeto bovino-, etc.; o cualquier otro animal, e.g., cerdos, conejos, aves de corral –e.g., patos, pollos, pavos-, etc.), mediante la identificación de un animal infectado con -o en riesgo de ser infectado por- un agente patógeno A/E; y para administrar al animal una cantidad efectiva de una composición que comprenda un polipéptido que incluya una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera o cualesquiera de las SEQ ID NOs: 22-43, 59 ó 73-8443; una secuencia nucleotídica que codifique para una secuencia polipeptídica sustancialmente idéntica a una secuencia seleccionada a partir de las SEQ ID NOs: 22-43, 59 ó 73-84; una molécula de ácido nucleico que incluya una secuencia nucleotídica sustancialmente idéntica a la secuencia de cualquiera o cualesquiera de las SEQ ID NOs: 1-21 ó 60-72; o un sobrenadante de cultivo celular que incluya tales polipéptidos, por ende produciendo una respuesta inmune, o reduciendo la colonización o la difusión del agente patógeno A/E en el animal.

La invención también provee, en aspectos alternativos, un método para atenuar la virulencia de un agente patógeno A/E mediante la provisión de un agente patógeno A/E; y para mutar un gen tal como *nleA*, *nleB*, *nleC*, *nleD*, *nleE*, *nleF*, *nleG* o *nleH*, o un homólogo de los mismos en el agente patógeno A/E, o mutar una o más secuencias

nucleotídicas en el genoma del agente patógeno A/E, donde la secuencia nucleotídica es seleccionada a partir de las SEQ ID NOs: 1-21 ó 60-72, de tal modo atenuando la virulencia.

La invención también provee, en aspectos alternativos, un método de selección para un compuesto que atenúa la virulencia de un agente patógeno A/E mediante la provisión de un sistema (e.g., una célula, tal como una célula de EHEC, EPEC o *C. rodentium*, un modelo animal, o un sistema *in vitro*) que incluye: un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera o cualesquiera de las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas; una secuencia nucleotídica que codifica para una secuencia polipeptídica sustancialmente idéntica a una secuencia seleccionada a partir de las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas; o una molécula de ácido nucleico que incluye una secuencia nucleotídica sustancialmente idéntica a la secuencia de cualquiera o cualesquiera de las SEQ ID NOs: 1-21 ó 60-72, o un fragmento o variante de las mismas; la provisión de un compuesto de prueba; y la determinación de si el compuesto de prueba modula la expresión, secreción o actividad biológica del polipéptido o de la molécula de ácido nucleico, donde un cambio, e.g., la disminución de la expresión, secreción o actividad biológica del polipéptido o la molécula de ácido nucleico, indica un compuesto que atenúa la virulencia de un agente patógeno A/E.

La invención también provee, en aspectos alternativos, un método para la producción de un factor de virulencia de un agente patógeno A/E mediante la provisión de una célula recombinante que incluye una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas; una secuencia nucleotídica que codifica para una secuencia polipeptídica sustancialmente idéntica a una secuencia seleccionada a partir del grupo compuesto por las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas; o una molécula de ácido nucleico que incluye una secuencia nucleotídica sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 1- 21 ó 60-72, o un fragmento o variante de las mismas; cultivar la célula recombinante en condiciones que permitan la expresión y/o la secreción del polipéptido, y

opcionalmente, aislar el polipéptido. En algunas realizaciones, el polipéptido puede ser secretado a partir de la célula.

La invención también provee, en aspectos alternativos, un método para tratar o prevenir la infección por un agente patógeno A/E mediante la identificación de un mamífero que tenga – o esté en riesgo de tener- una infección por un agente patógeno A/E; y administrar al mamífero una cantidad efectiva de un compuesto que atenúe la virulencia de un agente patógeno A/E, donde el compuesto inhibe la expresión o secreción de un polipéptido que incluya una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas. En algunas realizaciones, el compuesto puede ser una molécula de ácido nucleico antisentido que es complementaria de una secuencia nucleotídica sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 1- 21 ó 60-72, o un fragmento o variante de las mismas, o puede ser un ARNsi.

La invención también provee, en aspectos alternativos, un polipéptido recombinante que incluye una secuencia de aminoácidos sustancialmente idéntica a la secuencia de las SEQ ID NOs: 22-43, 59 ó 73-84, o una molécula aislada de ácido nucleico que incluye una secuencia nucleotídica sustancialmente idéntica a la secuencia de las SEQ ID NOs: 1-21 ó 60-72; y/o un vector que incluye tales secuencias nucleotídicas; y/o una célula hospedadora (e.g., un agente patógeno A/E tal como la *E. coli* enterohemorrágica –EHEC-, la *E. coli* enteropatogénica –EPEC- o *Citrobacter rodentium*, incluidos tales vectores). El vector puede ser capaz o incapaz de integrarse dentro del genoma de un agente patógeno A/E.

En aspectos alternativos, la invención también provee usos de las composiciones, bacterias, polipéptidos o las moléculas de ácido nucleico de acuerdo a la invención, para la preparación de un medicamento para la producción de una respuesta inmune contra un agente patógeno A/E o componentes del mismo, o para la reducción de la propagación o colonización de un agente patógeno A/E en un animal.

En aspectos alternativos, la invención también provee kits que incluyen un reactivo para la detección de un agente patógeno en una muestra, y un prospecto con

instrucciones para la detección del agente patógeno A/E en la muestra. El reactivo puede incluir una sonda o una sonda iniciadora o un *primer* sustancialmente idéntico a: una secuencia nucleotídica seleccionada a partir del grupo que comprende una o más de las SEQ ID NOs: 1-21 ó 60-72, o un fragmento o variante de las mismas, o una secuencia nucleotídica codificadora de un polipéptido sustancialmente idéntico a una o más de las SEQ ID NOs: 22-43, 59, 73-84 ó un fragmento o variante de las mismas, o un anticuerpo que se une específicamente a una secuencia seleccionada a partir del grupo que comprende una o más de las SEQ ID NOs: 22-43, 59, 73-84, ó un fragmento o variante de las mismas.

Un "agente patógeno A/E" es un agente patógeno, por ejemplo una bacteria patogénica *E. coli*, que puede adherirse a las superficies intestinales de un animal, por ejemplo un mamífero, e.g., vacas, ovejas, cabras, cerdos, conejos, perros, gatos, etc., o una especie de ave, e.g., pollos, patos, pavos, etc., y causar una lesión histológica característica llamada lesión de adherencia y destrucción (A/E) (8). En general, una infección patogénica A/E puede dar como resultado diarrea, colitis entérica, una enfermedad renal (tal como el síndrome urémico hemolítico). Sin embargo, una infección con un agente patógeno A/E no necesariamente se manifiesta como síntomas de enfermedad; un mamífero hospedador infectado con un agente patógeno A/E puede ser un portador del agente patógeno y permanecer sano y libre de enfermedad. De esta manera, los mamíferos infectados con –o en riesgo de infección por– un agente patógeno A/E incluyen animales, e.g., animales de granja, tales como aves de corral, e.g., pollos, pavos, patos, o rumiantes, e.g., vacas, ovejas, cabras, etc., u otros animales de granja, e.g., cerdos, que no manifiestan síntomas de enfermedad, y también incluyen humanos, que son susceptibles a enfermedades entéricas severas como resultado de una infección patogénica A/E.

Agentes patógenos A/E que sirven de ejemplo incluyen, sin limitación, a las *E. coli* enterohemorrágicas (EHEC) (también conocidas como *E. coli* productoras de toxina Shiga –STEC– o *E. coli* productoras de toxina Vero –VTEC), por ejemplo los serotipos EHEC 0157 (e.g., EHEC O157:H7, cuya secuencia genómica se describe en los Accesos No. AE005594, AE005595, AP002566, AE 005174, NC_002695 ó NC_002655), ó 0158, 05, 08, 018, 026, 045, 048, 052, 055, 075, 076, 078, 084, 91,

0103, 0104, 0111, 0113, 0114, 0116, 0118, 0119, 0121, 0125, 028, 0145, 0146, 0163, 0165; las *E. coli* enteropatógenicas (EPEC); y también a las *E. coli* patógenicas que infectan a ratones (e.g., *Citrobacter rodentium*); conejos (e.g., las cepas RDEC-1, tales como la O15:H-); cerdos; ovejas; perros; y otros mamíferos. Muchas cepas de agentes patógenos A/E están disponibles comercialmente, por ejemplo, a través de la American Type Culture Collection (ATCC), Manassus, VA, EEUU. Los agentes patógenos A/E también pueden ser aislados a partir de individuos infectados, por ejemplo, mediante cultivo directo en agar MacConkey de sorbitol suplementado con cefixima y telurita o mediante inoculación en el mismo medio luego de enriquecerlo inmunomagnéticamente (72, 107, 108).

Una "proteína", un "péptido" o un "polipéptido" es cualquier cadena de dos o más aminoácidos, comprendiendo aminoácidos o análogos de aminoácidos naturales o no naturales, sin tener en cuenta modificaciones post-traduccionales (e.g., glicosilación o fosforilación). Una "secuencia de aminoácidos", un "polipéptido", un "péptido" o una "proteína" de la invención puede incluir péptidos o proteínas que tienen uniones anormales, enlaces cruzados y capuchones terminales, uniones no peptídicas o grupos modificadores alternativos. Tales péptidos modificados también están dentro del alcance de la invención. El término "grupo modificador" pretende incluir a estructuras que están directamente acopladas a la estructura peptídica (e.g., por acoplamiento covalente) y a aquellas que están acopladas indirectamente a la estructura peptídica (e.g., mediante una asociación no covalente estable o por acoplamiento covalente a residuos aminoácidos adicionales, o a miméticos, análogos o derivados de los mismos, que puedan flanquear la estructura peptídica central). Por ejemplo, el grupo modificador puede estar acoplado al extremo amino o al extremo carboxilo de una estructura peptídica, o a una región peptídica o péptido-mimética que se encuentre flanqueando el dominio central. Alternativamente, el grupo modificador puede estar acoplado a una cadena lateral de al menos un residuo aminoácido de una estructura peptídica, o a una región peptídica o péptido-mimética que se encuentre flanqueando el dominio central (e.g., a través del grupo epsilon amino de uno o más residuos lisilo, a través del grupo carboxilo de uno o más residuos de ácido aspártico o uno o más residuos de ácido glutámico, a través de un grupo hidroxilo de uno o más residuos tirosilo, uno o más residuos de serina, o uno o más residuos de treonina, u otro grupo reactivo

apropiado en una cadena lateral aminoacídica). Los grupos modificadores que están acoplados en forma covalente a la estructura peptídica pueden ser adosados mediante el uso de métodos bien conocidos en el arte para ligar estructuras químicas, incluyendo por ejemplo los enlaces amida, alquilamino, carbamato o urea.

Los términos "ácido nucleico" o "molécula de ácido nucleico" comprenden tanto al ARN (hebras positiva y negativa) como al ADN, incluidos el ADNc, el ADN genómico, y el ADN sintético (e.g., sintetizado químicamente). El ácido nucleico puede ser de hebra doble o hebra única. Si es de hebra única, el ácido nucleico puede ser la hebra con sentido o la hebra antisentido. Una molécula de ácido nucleico puede ser cualquier cadena de dos o más nucleótidos ligados en forma covalente, incluidos los nucleótidos naturales o no naturales, o análogos o derivados de nucleótidos. Por "ARN" se entiende una secuencia de dos o más ribonucleótidos naturales o modificados, ligados en forma covalente. Un ejemplo de ARN modificado incluido dentro de este término es el ARN fosforotioato. Por "ADN" se entiende una secuencia de dos o más desoxirribonucleótidos naturales o modificados, ligados en forma covalente. Por "ADNc" se entiende el ADN copia o complementario producido a partir de una plantilla de ARN por la acción de una ADN polimerasa ARN-dependiente (transcriptasa inversa). De esta manera, un "clon de ADNc" significa una secuencia duplex de ADN complementaria a una molécula de ARN de interés, portada en un vector de clonado. Por "complementaria" se entiende que dos ácidos nucleicos, e.g., ADN o ARN, contienen un número suficiente de nucleótidos capaces de formar pares de bases Watson-Crick para producir una región de hebra doble entre los dos ácidos nucleicos. De esta manera, la adenina en una de las hebras de ADN o ARN se aparea con timina en una hebra complementaria y opuesta de ADN, o con uracilo en una hebra complementaria y opuesta de ARN. Se comprenderá que cada nucleótido en una molécula de ácido nucleico no necesita formar un par correspondiente de bases Watson-Crick con un nucleótido en una hebra complementaria y opuesta para formar un duplex. Una molécula de ácido nucleico es "complementaria" de otra molécula de ácido nucleico si se hibridiza, en condiciones muy estrictas, con la segunda molécula de ácido nucleico.

Un "sobrenadante de cultivo celular", tal como se usa aquí, se refiere generalmente a un sobrenadante derivado del cultivo de una bacteria u otro organismo

(e.g., levadura) o célula (e.g., célula de insecto) con la capacidad de secretar uno o más polipéptidos con una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22-43, 59, 73-84, ó a un fragmento o variante de las mismas, o a una porción inmunogénica de las mismas, en el medio de cultivo celular. En algunas realizaciones, el sobrenadante de cultivo celular es sustancialmente puro, i.e., sustancialmente libre de células bacterianas o del lisato de tales células. En algunas realizaciones, el sobrenadante de cultivo celular también puede contener uno o más de los polipéptidos EspA, EspB, EspD, Tir, intimina, toxina Shiga 1 ó 2, ó EspP, o fragmentos o agregados de los mismos.

La bacteria puede ser un agente patógeno A/E, por ejemplo EHEC, EPEC o *Citrobacter rodentium* que, en algunas realizaciones, puede ser modificado o mutado para expresar o secretar preferencialmente las proteínas descritas aquí, o puede ser alguna otra bacteria, por ejemplo una bacteria no patogénica, e.g., una *E. coli* no patogénica tal como HB101, o un agente patógeno no-A/E, que ha sido modificada o mutada, por ejemplo, por técnica recombinante u otra, de manera tal que secreta una o más proteínas descritas aquí, por ejemplo un polipéptido que comprenda una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22-43, 59, 73-84, ó a un fragmento o variante de las mismas, o a una porción inmunogénica de las mismas, en el medio de cultivo celular. En algunas realizaciones, la bacteria no es EHEC o EPEC. En algunas realizaciones, donde la bacteria es un agente patógeno A/E, ésta también puede portar una modificación adicional que perjudica su capacidad de expresar o secretar un polipéptido (por ejemplo, EspA, EspB, EspD, Tir, intimina, toxina Shiga 1 ó 2, ó EspP) que normalmente secretaría en ausencia de tal modificación. En algunas realizaciones, el otro organismo (e.g., levadura) o célula (e.g., célula de insecto) ha sido modificado o mutado, por ejemplo por técnica recombinante u otra, de manera tal que secreta una o más de las proteínas aquí descritas, por ejemplo un polipéptido que comprenda una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22 a 43, o a una porción inmunogénica de las mismas, en el medio de cultivo celular.

Un compuesto es "sustancialmente puro" o "aislado" cuando éste es separado de los componentes que lo acompañan naturalmente. Típicamente, un compuesto es

sustancialmente puro cuando es al menos 10%, 20%, 30%, 40%, 50% ó 60%, o más generalmente al menos 70%, 75%, 80%, 85%, 90%, 95% ó 99% en peso, del material total en una muestra. De esta manera, por ejemplo, un polipéptido que es sintetizado o producido químicamente mediante tecnología recombinante, generalmente estará sustancialmente libre de sus componentes asociados naturales. Un polipéptido también generalmente será sustancialmente puro si es separado de sus componentes asociados naturales mediante técnicas físicas, tales como centrifugación, precipitación, cromatografía en columna, electroforesis en gel, HPLC, etc.

Una molécula de ácido nucleico generalmente será sustancialmente pura o "aislada" cuando no es inmediatamente contigua con (i.e., ligada en forma covalente a) las secuencias codificadoras con las cuales es normalmente contigua en el genoma natural del organismo a partir del cual fue derivado el ADN de la invención. Por lo tanto, un gen o molécula de ácido nucleico "aislado" pretende significar un gen o molécula de ácido nucleico que no está flanqueado por las moléculas de ácido nucleico que normalmente (en la naturaleza) flanquean al gen o a la molécula de ácido nucleico (tal como en secuencias genómicas) y/o ha sido completa o parcialmente purificado a partir de otras secuencias transcritas (como en una biblioteca de ADNc o ARN). Por ejemplo, un ácido nucleico aislado de la invención puede ser sustancialmente aislado con respecto al complejo ambiente celular en el cual naturalmente se presenta. En algunas ocasiones, el material aislado formará parte de una composición (por ejemplo, un extracto crudo que contiene otras sustancias), un sistema buffer u una mezcla de reactivos. En otras circunstancias, el material puede ser purificado hasta la homogeneidad esencial, por ejemplo como la determinada por PAGE o cromatografía en columna tal como HPLC. El término por lo tanto incluye, e.g., un ácido nucleico recombinante incorporado en un vector, tal como un plásmido o virus replicante en forma autónoma; o en ADN genómico de un procarionte o eucariota, o el cual existe como una molécula separada (e.g., un ADNc o un fragmento de ADN genómico producido por PCR o por tratamiento con endonucleasas de restricción), independiente de otras secuencias. También incluye un ácido nucleico recombinante que es parte de un gen híbrido que codifica para secuencias polipeptídicas adicionales. Preferentemente, un ácido nucleico aislado comprende al menos aproximadamente 50%, 80% ó 90% (sobre una base molar) de todas las especies macromoleculares

presentes. De esta manera, un gen o molécula de ácido nucleico aislado puede incluir un gen o molécula de ácido nucleico que es sintetizado químicamente o por medios recombinantes. Los ADN recombinantes contenidos en un vector están incluidos en la definición de "aislados" tal como se la usa aquí. Además, las moléculas de ácido nucleico aisladas incluyen moléculas recombinantes de ADN en células hospedadoras heterólogas, así como moléculas de ADN parcial o sustancialmente purificadas en solución. Las moléculas de ácido nucleico "aisladas" también comprenden a las transcripciones de ARN *in vivo* e *in vitro* de las moléculas de ADN de la presente invención. Tales moléculas de ácido nucleico aisladas son útiles para la fabricación del polipéptido codificado, como sondas para aislar secuencias homólogas (e.g., a partir de otras especies de mamíferos), para el mapeo de genes (e.g., por hibridación *in situ* con cromosomas), o para detectar la expresión del gen en tejidos (e.g., tejido humano, tal como sangre periférica), tal como mediante análisis Northern Blot.

Un compuesto sustancialmente puro puede obtenerse, por ejemplo, por extracción a partir de una fuente natural; por expresión de una molécula recombinante de ácido nucleico que codifique para un compuesto polipeptídico; o por síntesis química. La pureza puede medirse usando cualquier método apropiado, tal como cromatografía en columna, electroforesis en gel, HPLC, etc. Una preparación sustancialmente pura de una célula, por ejemplo una célula bacteriana, es una preparación de células en la cual las células contaminantes que no tienen el genotipo mutante deseado o que no expresan o secretan el polipéptido deseado en cantidades suficientes, constituye menos del 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ó 90% del número total de células en la preparación.

Varios genes y secuencias de ácido nucleico de la invención pueden ser secuencias recombinantes. El término "recombinante" significa que algo ha sido recombinado, de manera que cuando se lo aplica en referencia a una construcción de ácido nucleico, el término se refiere a una molécula que está compuesta por secuencias de ácido nucleico las cuales son unidas o producidas mediante técnicas de biología molecular. El término "recombinante", cuando es aplicado en referencia a una proteína o un polipéptido, se refiere a una proteína o molécula polipeptídica que se expresa usando una construcción de ácido nucleico recombinante creada mediante técnicas de

biología molecular. El término "recombinante", cuando es aplicado en referencia a la composición genética, se refiere a una gameta o progenie con nuevas combinaciones de alelos que no estaban presentes en los genomas paternos. Las construcciones recombinantes de ácido nucleico pueden incluir una secuencia nucleotídica ligada a –o manipulada para ser ligada a– una secuencia de ácido nucleico a la cual no se encuentra ligada en la naturaleza, o a la cual está ligada en un sitio diferente en la naturaleza. Referirse a una construcción de ácido nucleico como "recombinante" indica por lo tanto que la molécula de ácido nucleico ha sido manipulada por medio de ingeniería genética, i.e., con intervención humana.

Las construcciones recombinantes de ácido nucleico pueden por ejemplo ser introducidas en una célula hospedadora mediante transformación. Tales construcciones recombinantes de ácido nucleico pueden incluir secuencias derivadas a partir de la misma especie de célula hospedadora o a partir de diferentes especies de célula hospedadora, las cuales han sido aisladas y reintroducidas en células de la especie hospedadora. Las secuencias de las construcciones recombinantes de ácido nucleico pueden ser integradas al genoma de la célula hospedadora, ya sea como resultado de la transformación original de la célula hospedadora, o como el resultado de la subsiguiente recombinación y/o eventos de reparación.

Tal como se emplea aquí, "heterólogo" en referencia a un ácido nucleico o proteína es una molécula que ha sido manipulada con intervención humana de manera que está localizada en un lugar distinto del lugar en el cual se encuentra naturalmente. Por ejemplo, una secuencia de ácido nucleico de una especie puede ser introducida en el genoma de otra especie, o una secuencia de ácido nucleico de un locus genómico puede ser desplazada a otro locus genómico o extracromosómico en la misma especie. Una proteína heteróloga incluye, por ejemplo, una proteína expresada a partir de una secuencia codificadora heteróloga, o una proteína expresada a partir de un gen recombinante en una célula que naturalmente no expresaría la proteína.

Una secuencia "sustancialmente idéntica" es una secuencia de aminoácidos o nucleótidos que difiere de una secuencia de referencia sólo por una o más sustituciones conservativas, tal como se discute aquí, o por una o más sustituciones, deleciones o

inserciones no conservativas localizadas en posiciones de la secuencia que no destruyen la función biológica del aminoácido o de la molécula de ácido nucleico. Tal secuencia puede ser cualquier número entero entre 10% y 99%, o más generalmente al menos 10%, 20%, 30%, 40%, 50, 55% ó 60%, ó al menos 65%, 75%, 80%, 85%, 90% ó 95%, ó tanto como 96%, 97%, 98% ó 99% idéntica, al nivel de aminoácidos o nucleótidos, a la secuencia usada para comparación, usando por ejemplo el programa Align (96) o FASTA. Para polipéptidos, la longitud de las secuencias de comparación puede ser al menos de 2, 5, 10 ó 15 aminoácidos, o al menos de 20, 25 ó 30 aminoácidos. En realizaciones alternativas, la longitud de las secuencias de comparación puede ser al menos de 35, 40 ó 50 aminoácidos, o superior a 60, 80 ó 100 aminoácidos. Para moléculas de ácido nucleico, la longitud de las secuencias de comparación puede ser al menos de 5, 10, 15, 20 ó 25 nucleótidos, o al menos de 30, 40 ó 50 nucleótidos. En realizaciones alternativas, la longitud de las secuencias de comparación puede ser al menos de 60, 70, 80 ó 90 nucleótidos, o superior a 100, 200 ó 500 nucleótidos. La identidad de la secuencia puede ser fácilmente medida usando software de análisis de secuencias disponible al público (e.g., Sequence Analysis Software Package del Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705, o el software BLAST de la National Library of Medicine, o como se describe aquí). Los ejemplos de software útil incluyen los programas Pile-up y PrettyBox. Tal software establece correspondencias entre secuencias similares mediante la asignación de grados de homología a varias sustituciones, deletaciones y otras modificaciones.

Alternativamente, o adicionalmente, dos secuencias de ácido nucleico pueden ser "sustancialmente idénticas" si se hibridizan en condiciones muy estrictas. En algunas realizaciones, condiciones muy estrictas son, por ejemplo, condiciones que permiten una hibridación comparable con la hibridación que ocurre al usarse una sonda de ADN de al menos 500 nucleótidos de longitud, en un buffer que contiene NaHPO_4 0,5 M, pH 7,2; SDS al 7%, EDTA 1mM y BSA al 1% (fracción V) a una temperatura de 65°C, o un buffer que contiene formamida al 48%, 4,8x SSC, Tris-Cl 0,2 M, pH 7,6; 1x solución de Denhardt, sulfato de dextrano al 10% y SDS al 0,1%, a una temperatura de 42°C. (Estas son condiciones típicas para hibridaciones Northern y Southern de elevada exactitud.) Las hibridaciones pueden llevarse a cabo durante un periodo de alrededor

de 20 a 30 minutos, o alrededor de 2 a 6 horas, o alrededor de 10 a 15 horas, o superior a 24 horas o más. De la hibridación en elevada exactitud depende también el éxito de numerosas técnicas ejecutadas en forma rutinaria por biólogos moleculares, tales como PCR de elevada exactitud, secuenciación de ADN, análisis conformacional de polimorfismo de hebra única, e hibridación in situ. Contrariamente a las hibridaciones Northern y Southern estas técnicas usualmente se realizan con sondas relativamente cortas (e.g., usualmente alrededor de 16 nucleótidos o más larga para PCR o secuenciación, y alrededor de 40 nucleótidos o más larga para hibridación in situ). Las condiciones de elevada exactitud usadas en estas técnicas son bien conocidas para aquéllos diestros en el arte de la biología molecular (61).

Una secuencia sustancialmente idéntica puede ser por ejemplo una secuencia de aminoácidos que es sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas, o una secuencia nucleotídica sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 1-21 ó 60-72, o un fragmento o variante de las mismas. En algunas realizaciones, una secuencia sustancialmente idéntica puede ser por ejemplo una secuencia nucleotídica que es complementaria a –o se hibridiza con– la secuencia de cualquiera de las SEQ ID NOs: 1-21 ó 60-72, o un fragmento o variante de las mismas. En algunas realizaciones, una secuencia sustancialmente idéntica puede ser derivada a partir de un agente patógeno A/E.

Una "sonda" o "*primer*" es una molécula de ADN o ARN de secuencia definida y con una sola hebra, que puede aparearse por sus bases a una segunda molécula de ADN o ARN que contiene una secuencia complementaria (el blanco). La estabilidad de la molécula híbrida resultante depende del grado de apareamiento de bases que tiene lugar, y es afectada por parámetros tales como el grado de complementariedad entre la sonda y la molécula blanco, y el grado de exactitud de las condiciones de hibridación. El grado de exactitud de hibridación es afectado por parámetros tales como la temperatura, la concentración de sales, y la concentración de moléculas orgánicas tales como formamida, y se determina por métodos que son conocidos a aquéllos diestros en el arte. Las sondas o primers específicos para las secuencias de ácido nucleico descritas aquí, o porciones de las mismas, pueden variar en longitud por cualquier

número entero desde al menos 8 nucleótidos hasta más de 500 nucleótidos, incluyendo cualquier valor entre éstos, dependiendo del propósito para el cual —y las condiciones en las cuales— la sonda o primer es usado. Por ejemplo, una sonda o primer puede tener al menos 8, 10, 15, 20 ó 25 nucleótidos de longitud, o puede tener al menos 30, 40, 50 ó 60 nucleótidos de longitud, o puede tener más de 100, 200, 500 ó 1000 nucleótidos de longitud. Las sondas o primers específicos para las moléculas de ácido nucleico descritas aquí pueden ser cualquier número entero desde 10% hasta 99%, o más generalmente al menos 10%, 20%, 30%, 40%, 50%, 55% ó 60%, o al menos 65%, 75%, 80%, 85%, 90% ó 95%, o tanto como 96%, 97%, 98% ó 99% idénticas a las secuencias de ácido nucleico descritas aquí usando por ejemplo el programa Align (96).

Las sondas y los primers pueden marcarse detectablemente, ya sea radiactivamente o no radiactivamente, mediante métodos que son conocidos para aquéllos diestros en el arte. Las sondas o primers pueden usarse para métodos que implican la hibridación de ácido nucleico, tales como la secuenciación de ácido nucleico, amplificación de ácido nucleico por reacción en cadena con polimerasa, análisis de polimorfismo conformacional de hebra única (SSCP), análisis de polimorfismo de fragmentos por restricción (RFLP), hibridación Southern, hibridación Northern, hibridación in situ, ensayo de desplazamiento por movilidad electroforética (EMSA), y otros métodos que son conocidos para aquéllos diestros en el arte. Las sondas o primers pueden ser derivadas a partir de ADNc o ADN genómico, por ejemplo por amplificación, o a partir de segmentos de ADN clonados, o pueden ser sintetizados químicamente.

Una "mutación" incluye cualquier alteración en la secuencia de ADN, i.e., genoma, de un organismo cuando se lo compara con la cepa parental. Las alteraciones pueden surgir espontáneamente o por exposición del organismo a un estímulo mutagénico tal como un compuesto químico mutagénico, energía, radiación, técnicas recombinantes, apareamiento o cualquier otra técnica empleada para alterar ADN. Una mutación puede incluir una alteración en cualquiera de las secuencias nucleotídicas aquí descritas, o puede incluir una alteración en una secuencia nucleotídica que codifique para cualquiera de los polipéptidos descritos aquí.

Una mutación puede "atenuar virulencia" si, como resultado de la mutación, el nivel de virulencia de la célula mutante disminuye al menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ó 100% respecto de la cepa parental. La disminución de la virulencia también puede medirse por una disminución de al menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ó 100% en la expresión de un polipéptido, por ejemplo un polipéptido que incluya una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas, en la cepa mutante cuando se la compara con la cepa parental. La virulencia de un agente patógeno A/E puede medirse según se describe aquí o tal como es conocido para el arte. Una disminución de la virulencia también puede medirse por un cambio de al menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ó 100% en la actividad biológica de un polipéptido, por ejemplo un polipéptido que incluya una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas.

"Modular" o "modula" significa cambiar, ya sea por aumento o disminución. El aumento o la disminución puede ser un cambio de cualquier valor entero entre 10% y 90%, o de cualquier valor entero entre 30% y 60%, o puede ser superior a 100%, cuando se la compara con un control o una muestra o compuesto de referencia.

Un "compuesto de prueba" es cualquier compuesto químico natural o derivado artificialmente. Los compuestos de prueba pueden incluir, sin limitación, a péptidos, polipéptidos, moléculas orgánicas sintetizadas, moléculas orgánicas naturales, moléculas de ácido nucleico. Un compuesto de prueba puede "competir" con un compuesto conocido tal como cualquiera de los polipéptidos o moléculas de ácido nucleico descritas aquí, por ejemplo, interfiriendo con la virulencia o interfiriendo con cualquier respuesta biológica inducida por el compuesto conocido.

Generalmente, un compuesto de prueba puede exhibir cualquier valor entre 10% y 200%, o superior a 500% de modulación cuando se lo compara con un compuesto de referencia. Por ejemplo, un compuesto de prueba puede exhibir al menos cualquier valor entero positivo o negativo entre 10% y 200% de modulación, o al menos cualquier

valor entero positivo o negativo entre 30% y 150% de modulación, o al menos cualquier valor entero positivo o negativo entre 60% y 100% de modulación, o cualquier valor entero positivo o negativo superior a 100% de modulación. Un compuesto que es un modulador negativo en general disminuirá la modulación relativa a un compuesto conocido, mientras que un compuesto que es un modulador positivo en general aumentará la modulación relativa a un compuesto conocido.

Un "vector" es una molécula de ADN derivada, por ejemplo, de un plásmido, bacteriófago, virus de mamífero o de insecto, o cromosoma artificial, en la cual puede insertarse una molécula de ácido nucleico, por ejemplo una secuencia nucleotídica substancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 1-21 ó 60-72 ó un fragmento o variante de las mismas. Un vector puede contener uno o más sitios únicos de restricción y puede ser capaz de replicación autónoma en un definido organismo hospedador o vehículo, de manera tal que la secuencia clonada sea reproducible. Un vector puede ser un vector de expresión de ADN, i.e., cualquier elemento autónomo capaz de dirigir la síntesis de un polipéptido recombinante, y puede por lo tanto ser usado para expresar un polipéptido, por ejemplo un polipéptido que comprenda una secuencia de aminoácidos substancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas, en una células hospedadora. Los vectores de expresión de ADN incluyen plásmidos y fagos bacterianos y plásmidos y virus de mamíferos e insectos. Un vector puede ser capaz de integrarse al genoma de la célula hospedadora, de manera tal que cualquier modificación introducida en el genoma de la célula hospedadora por el vector deviene parte del genoma de la célula hospedadora. Un vector puede ser incapaz de integrarse al genoma de la célula hospedadora, y por lo tanto permanecerá como una unidad autónomamente replicante, tal como un plásmido.

Un anticuerpo se "une específicamente" a un antígeno cuando lo reconoce y se une al antígeno, por ejemplo, un polipéptido que incluya una secuencia de aminoácidos substancialmente idéntica a la secuencia de cualquiera de las SEQ ID NOs: 22-43, 59 ó 73-84, o un fragmento o variante de las mismas, pero substancialmente no reconoce y no se une a moléculas en una muestra. Tal anticuerpo tiene, por ejemplo, una afinidad por

el antígeno que es al menos 10, 100, 1000 ó 10000 veces mayor que la afinidad del anticuerpo por otra molécula de referencia en una muestra.

Una "muestra" puede ser cualquier órgano, tejido, célula o extracto de células aislado de un sujeto, tal como una muestra aislada de un animal infectado con un agente patógeno A/E, o un animal al cual se le ha administrado uno o más de los polipéptidos o moléculas de ácido nucleico de la invención, o fragmentos inmunogénicos de las mismas. Por ejemplo, una muestra puede incluir, sin limitación, tejido (e.g., de una biopsia o autopsia), células, sangre, suero, leche, orina, deposiciones, saliva, heces, huevos, cultivo de células de mamífero o medio de cultivo, o cualquier otro espécimen, o cualquier extracto de los mismos, obtenido de un paciente (humano o animal), sujeto de prueba o animal experimental. Una muestra también puede incluir, sin limitación, productos producidos en un cultivo celular por células normales o transformadas (e.g., mediante ADN recombinante o tecnología de anticuerpos monoclonales). Una "muestra" también puede ser una célula o línea celular creada en condiciones experimentales, que no son aisladas directamente de un sujeto. Una muestra también puede ser libre de células, derivada artificialmente o sintetizada.

La muestra puede ser analizada para detectar la presencia de un gen, genoma, polipéptido, o molécula de ácido nucleico derivada de un agente patógeno A/E, o para detectar una mutación en un gen derivado de un agente patógeno A/E, para detectar niveles de expresión de un gen o polipéptido derivado de un agente patógeno A/E, o para determinar la función biológica de un gen o polipéptido derivado de un agente patógeno A/E, usando métodos que son conocidos en el arte y/o descritos aquí. Por ejemplo, métodos tales como secuenciación, análisis de polimorfismo conformacional de hebra única (SSCP), o análisis de restricción del polimorfismo de la longitud de fragmentos (RFLP) de los productos de PCR derivados de una muestra pueden usarse para detectar una mutación en un gen; para medir los niveles de afinidad de los polipéptidos o anticuerpos puede usarse ELISA o Western Blot; para medir los niveles de ARNm puede usarse Northern Blot, o puede usarse PCR para medir el nivel de una molécula de ácido nucleico.

Una "respuesta inmune" incluye, pero no está limitada, a una o más de las siguientes respuestas en un mamífero: inducción de anticuerpos, células B, células T (incluidas células T ayudantes, células T supresoras, células T citotóxicas, células T $\gamma\delta$) dirigidos específicamente a el/los antígeno/s en una composición o vacuna, seguido de la administración de la composición o vacuna. Una respuesta inmune a una composición o vacuna, por lo tanto, generalmente implica el desarrollo, en el mamífero hospedador, de una respuesta celular y/o mediada por anticuerpos hacia la composición o vacuna de interés. En general, la respuesta inmune dará como resultado la prevención o la reducción de la infección por un agente patógeno A/E; la resistencia del intestino a la colonización por un agente patógeno A/E; o la reducción de la difusión del agente patógeno A/E.

Un "fragmento inmunogénico" de un polipéptido o de una molécula de ácido nucleico se refiere a una secuencia de aminoácidos o de nucleótidos que produce una respuesta inmune. Por lo tanto, un fragmento inmunogénico puede incluir, sin limitación, cualquier porción de cualquiera de las secuencias descritas aquí, o una secuencia sustancialmente idéntica a las mismas, que incluyan uno o más epítopes (el sitio reconocido por una célula específica del sistema inmunitario, tal como una célula T o una célula B). Por ejemplo, un fragmento inmunogénico puede incluir, sin limitación, péptidos con un número cualquiera entre 6 y 60, o superior a 60, de aminoácidos de longitud, e.g., péptidos con un número cualquiera entre 10 y 20 aminoácidos de longitud, o entre 20 y 40 aminoácidos de longitud, derivados de cualquiera o cualesquiera de las secuencias descritas aquí. Tales fragmentos pueden identificarse usando métodos corrientes conocidos para aquéllos diestros en el arte, tales como técnicas de mapeo de epítopes o antigenicidad o diagramas de hidropatía usando, por ejemplo, la versión 1.0 del programa Omega del Oxford Molecular Group (ver, por ejemplo, Patente US No. 4.708.871) (76, 77, 81, 92, 73).

BREVE DESCRIPCION DE LAS FIGURAS

Las Figuras 1A-F muestran las secuencias de nucleótidos y aminoácidos del NleA de *C. rodentium*, EPEC, y EHEC (SEQ ID NOS: 1-3 y 22-24).

Las Figuras 2A-J muestran las secuencias de nucleótidos y aminoácidos del NieB y NieB2 de *C. rodentium*, EPEC, y EHEC (SEQ ID NOs: 4-7 & 25-29 y 60).

Las Figuras 3A-F muestran las secuencias de nucleótidos y aminoácidos del NieC de *C. rodentium*, EPEC, y EHEC (SEQ ID NOs: 8-10 y 30-32).

Las Figuras 4A-F muestran las secuencias de nucleótidos y aminoácidos del NieD de *C. rodentium*, EPEC, y EHEC (SEQ ID NOs: 11-13 y 33-35).

Las Figuras 5A-H muestran las secuencias de nucleótidos y aminoácidos del NieE de *C. rodentium*, EPEC, y EHEC (SEQ ID NOs: 14-17 y 36-39).

Las Figuras 6A-H muestran las secuencias de nucleótidos y aminoácidos del NieF de *C. rodentium*, EPEC, y EHEC (SEQ ID NOs: 18-21 y 40-43).

La Figura 7 muestra un alineamiento secuencial de aminoácidos de *C. rodentium* Orf11/GrlA (SEQ ID NO: 56) con un regulador transcripcional positivo, CaiF (SEQ ID

NO: 57), y con la secuencia de aminoácidos deducida de una proteína no caracterizada de *Salmonella* (SEQ ID NO: 58). La estructura hélice-giro-hélice esperada, característica de las proteínas de unión del ADN, está subrayada. Los residuos aminoacídicos idénticos están señalados con *, mientras que los cambios conservados está marcados con +.

La Figura 8 muestra la complementación del $\Delta orf11$ de *C. rodentium* por el *orf11* de *C. rodentium* (pCRorf11), EHEC (pEHorf11), o EPEC (pEPorf11), y la complementación del mutante doble $\Delta ler - \Delta orf11$ de *C. rodentium* por el *ler* o el *orf11* de *C. rodentium*.

La Figura 9 es un diagrama esquemático que muestra las localizaciones relativas de las islas con los 6 genes efectores recientemente identificados en el genoma EHEC O157:H7. También se muestran las localizaciones de los genes de la toxina Shiga (*stx*), el LEE, y el TTSS *inv-spa*. Adviértase la asociación de muchos de estos genes con profagos (CP-933 y BP-933).

Las Figuras 10A-B muestran el análisis proteómico de las proteínas EHEC secretadas. A. Gel SDS-PAGE unidimensional de proteínas secretadas totales de EHEC de tipo salvaje (wt) y del mutante de secreción del tipo III (EscN-). La migración de los marcadores de peso molecular (en kDa) está indicada a la izquierda del gel. B. Gel bidimensional de proteínas secretadas totales de EHEC de tipo salvaje. La

migración de los marcadores de peso molecular (en kDa) está indicada a la izquierda, y los valores aproximados de pI se muestran arriba del gel. Las manchas de proteína analizadas por espectroscopia de masa (ver Tabla I) están encerradas en círculos y numeradas.

Las Figuras 11A-C muestran la organización genómica, distribución y conservación del NleA. A. Representación gráfica de la región alrededor de nleA en el genoma EHEC. La dirección transcripcional de cada ORF está indicada con una flecha. La anotación de los ORF está modificada de (3). El NleA está destacado en negrita. B. Análisis por Southern Blot del ADN genómico de EPEC, EHEC, REPEC, *Citrobacter rodentium* (Citro.), y la cepa no patogénica HB101 de *E. coli*. Cada una de las muestras de ADN genómico fue digerida con BamHI (líneas 1, 4, 7, 10,13), EcoRI (líneas 2, 5, 8, 11, 14), y PstI (líneas 3, 6, 9, 12, 15). C. Alineamiento de secuencias de proteínas múltiples de NleA de EHEC, el profago de una cepa EHEC intimina-positiva y no-0157 (O84:H4), EPEC y *Citrobacter rodentium*. Los residuos idénticos están representados por un punto (.), los aminoácidos ausentes en una secuencia en particular están representados con un guión (-). Dos tramos hidrofóbicos que podrían ser dominios putativos de transmembrana están destacados en negrita en la secuencia de EHEC.

La Figura 12 muestra un análisis Western Blot de proteínas secretadas (líneas a la izquierda) y *pellets* bacterianos (líneas a la derecha) de EHEC de tipo salvaje (wt) y el mutante de secreción de tipo III (EscN-) que expresa pNleA-HA y los controles no transformados. Los blots fueron sondeados con anti-HA (A.), anti-DnaK (B.), anti-Tir (C.).

Las Figuras 13A-B muestran secreción de tipo III y translocación en la EHEC Δ nleA. A. Gel SDS-PAGE de las proteínas secretadas totales de la EHEC salvaje (wt) y el mutante Δ nleA. La migración de marcadores de peso molecular (en kDa) está indicada a la izquierda del gel. B. El análisis Western Blot de proteínas secretadas de la EHEC de tipo salvaje (wt) y el mutante Δ nleA con antisuero anti-NleA. La migración de los marcadores de peso molecular (en kDa) está indicada a la izquierda del gel.

Las Figuras 14A-B muestran en análisis Western Blot de fracciones de células hospedadoras infectadas. A. Las células HeLa fueron infectadas con EHEC de tipo salvaje (wt) o escN- que expresan NleA marcado con HA y sometido a fraccionamiento subcelular por centrifugación diferencial. Las fracciones analizadas fueron: bacterias, células y citoesqueletos sanos (pellet de baja velocidad), citosol de la células hospedadora (citosol hospedador), y membranas de células hospedadoras (membrana hospedadora). Las fracciones fueron analizadas por Western Blot usando anticuerpos anti-HA, anti-DnaK, anti-Calnexina, y antitubulina. B. Se aislaron fracciones de membrana de células infectadas con EHEC de tipo salvaje que expresa NleA marcado con HA. Las fracciones de membrana luego fueron extraídas sobre hielo en condiciones d alto nivel de sal (NaCl 1M), alto pH (pH 11,4), pH neutral y sal isotónica (control), o pH neutral y sal isotónica conteniendo 1% de triton x100 (Triton X100), y centrifugadas para obtener fracciones de membrana solubles (S) e insolubles (P). Estas fracciones fueron sometidas a análisis Western Blot con anticuerpos anti-HA (panel superior), anticalnexina (panel medio) y anti-calreticulina (panel inferior).

Las Figuras 15A-D muestra estudios de virulencia de *Citrobacter rodentium* en ratones. A. Análisis Western Blot de extractos bacterianos totales de *C. rodentium* de tipo salvaje (wt) y el mutante $\Delta nleA$, sondeados con antisuero anti-NleA. La migración de los marcadores de peso molecular (en kDa) está indicada a la izquierda del gel. B. gráficos de supervivencia para ratones C3H/HeJ infectados con *C. rodentium* de tipo salvaje (cuadrados negros), *C. rodentium* $\Delta nleA$ (círculos blancos), y ratones previamente infectados con el mutante $\Delta nleA$ y subsiguientemente expuestos a *C. rodentium* de tipo salvaje (barras verticales). Los ratones fueron monitoreados diariamente durante el curso de la infección y cuando se tomaban moribundos eran sacrificados inmediatamente. Se indica el porcentaje del número inicial de ratones en cada grupo que eran viables cada día. C. Titulaciones de *C. rodentium* del colon de ratones suizos NIH infectados. Los ratones fueron infectados con *C. rodentium* de tipo salvaje (círculos negros) o la cepa $\Delta nleA$ (círculos blancos) y sacrificados al día 10 post-infección. El tejido del colon y los pellets fecales fueron homogeneizados y cultivados en agar MacConkey para determinar la carga total de *C. rodentium* en el colon del ratón al momento del sacrificio. Cada ratón en el experimento está representado por un único punto. La media de cada grupo está indicada en el gráfico mediante barras

horizontales. D. Pesos del colon y el bazo de los ratones suizos NIH infectados. Los ratones fueron infectados con *C. rodentium* de tipo salvaje (cuadrados y triángulos negros) o la cepa $\Delta nleA$ (cuadrados y triángulos blancos), y sacrificados al día 10 post-infección. Los cólores (cuadrados) y bazos (triángulos) fueron disecados y pesados. Cada ratón en el experimento está representado por un único punto. La media de cada grupo está indicada en el gráfico mediante barras horizontales.

Las Figuras 16A-B muestran las secuencias de nucleótidos y aminoácidos del homólogo de NleG de EHEC (SEQ ID NOs: 61 y 73).

Las Figuras 17A-B muestran las secuencias de nucleótidos y aminoácidos del homólogo de NleH1 de EHEC (SEQ ID NOs: 62 y 74).

Las Figuras 18A-B muestran las secuencias de nucleótidos y aminoácidos del homólogo de NleH2 de EHEC (SEQ ID NOs: 63 y 75).

Las Figuras 19A-B muestran las secuencias de nucleótidos y aminoácidos del Z2076 de EHEC (SEQ ID NOs: 64 y 76).

Las Figuras 20A-B muestran las secuencias de nucleótidos y aminoácidos del Z2149 de EHEC (SEQ ID NOs: 65 y 77).

Las Figuras 21A-B muestran las secuencias de nucleótidos y aminoácidos del Z2150 de EHEC (SEQ ID NOs: 66 y 78).

Las Figuras 22A-B muestran las secuencias de nucleótidos y aminoácidos del Z2151 de EHEC (SEQ ID NOs: 67 y 79).

Las Figuras 23A-B muestran las secuencias de nucleótidos y aminoácidos del Z2337 de EHEC (SEQ ID NOs: 68 y 80).

Las Figuras 24A-B muestran las secuencias de nucleótidos y aminoácidos del Z2338 de EHEC (SEQ ID NOs: 69 y 81).

Las Figuras 25A-B muestran las secuencias de nucleótidos y aminoácidos del Z2339 de EHEC (SEQ ID NOs: 70 y 82).

Las Figuras 26A-B muestran las secuencias de nucleótidos y aminoácidos del Z2560 de EHEC (SEQ ID NOs: 71 y 83).

Las Figuras 27A-B muestran las secuencias de nucleótidos y aminoácidos del Z2976 de EHEC (SEQ ID NOs: 72 y 84).

DESCRIPCION DETALLADA DE LA INVENCION

Hemos identificado varias nuevas comunes proteínas de secreción para agentes patógenos A/E (Tabla 2) usando un regulador LEE positivo (Regulador Global del Activador LEE, o GrlA) las cuales pueden usarse para aumentar significativamente la secreción, y nos ha permitido seleccionar funcionalmente las proteínas secretadas via el TTSS codificado por LEE mediante un acercamiento basado en la proteómica. Estas nuevas proteínas, llamadas Nle (efector no codificado por LEE) desde A hasta H, están presentes en agentes patógenos que contienen LEE, y están ausentes en cepas no patogénicas de *E. coli* y en agentes patógenos que no son LEE; están codificadas fuera del LEE por 3 PAIs presentes en los agentes patógenos A/E y han coevolucionado con el LEE (3, 8). La identificación de estas proteínas ha permitido, en algunos casos, la asignación de funciones a los ORFs con funciones previamente desconocidas. Una proteína que sirve de ejemplo, NleA (p54), es un efector del tipo III en agentes patógenos A/E, incluso en *C. rodentium*, EPEC y EHEC, y juega un rol crítico en la virulencia. La NleA esta codificada en una isla de patogenicidad asociada a un fago, dentro del genoma de EHEC, separado del LEE. El TTSS codificado por LEE dirige la translocación de la NleA hacia dentro de la célula hospedadora, donde se localiza al aparato de Golgi. La NleA está presente en agentes patógenos con LEE, y está ausente en cepas no patogénicas de *E. coli* y en agentes patógenos sin LEE.

En algunas realizaciones de la invención, estos polipéptidos y las moléculas de ácido nucleico que codifican para estos polipéptidos, o porciones de las mismas,

pueden ser útiles como vacunas, terapéuticas, diagnósticos, o herramientas para la selección de drogas para infecciones por agentes patógenos A/E, o como reactivos.

Polipéptidos y compuestos de prueba

Los compuestos de acuerdo a la invención incluyen, sin limitación, a los polipéptidos y moléculas de ácido nucleico descritas en, por ejemplo, las SEQ ID NOs: 1-56, 59-84, y a los fragmentos, análogos y variantes de las mismas. Los compuestos de acuerdo a la invención también incluyen a los productos de los genes *orf11/grlA*, *nleA*, *nleB*, *nleB2*, *nleC*, *nleD*, *nleE*, *nleF*, *nleG*, *nleH* (*nle H1* y/o *nle H2*), o los homólogos de los mismos. Los compuestos de acuerdo a la invención también incluyen a los polipéptidos y moléculas de ácido nucleico descritas en, por ejemplo, la secuencia genómica de EHEC (e.g., AE005174) como los números Z0985 (*NleB2*), Z0986 (*NleC*), Z0990 (*NleD*), Z6020 (*NleF*), Z6024 (*NleA*), Z4328 (*NleB*), Z4329 (*NleE*), Z6025 (*NleG* homolog), Z6021 (*NleH1*), Z0989 (*NleH2*), Z2076, Z2149, Z2150, Z2151, Z2337, Z2338, Z2339, Z2560, Z2976, ó L0043 (*Orf11/GrlA*)(Acceso No. AF071034), y a los fragmentos, análogos y variantes de las mismas.

Los compuestos pueden prepararse, por ejemplo, remplazando, eliminando o insertando un residuo aminoacídico en cualquier posición de un polipéptido aquí descrito, con otros residuos aminoacídicos conservativos, i.e., residuos con similares propiedades físicas, biológicas o químicas, y seleccionando, por ejemplo, por la capacidad del compuesto para atenuar la virulencia. En algunas realizaciones de la invención, los compuestos de la invención incluyen anticuerpos que se unen específicamente a los polipéptidos aquí descritos, por ejemplo, las SEQ ID NOs: 22-43, 59 ó 73-84.

Es bien conocido para el arte el hecho de que algunas modificaciones y cambios pueden hacerse en la estructura de un polipéptido sin alterar sustancialmente la función biológica de ese péptido, para obtener un polipéptido biológicamente equivalente. En un aspecto de la invención, los polipéptidos de la presente invención se extienden también a péptidos o "variantes" biológicamente equivalentes que difieren de una porción de la secuencia de los polipéptidos de la presente invención por sustituciones conservativas

de aminoácidos, o difieren por sustituciones no conservativas que no afectan la función biológica, e.g., la virulencia. Tal como se lo emplea aquí, el término "sustituciones conservadas de aminoácidos" se refiere a la sustitución de un aminoácido por otro en una posición dada del péptido, donde la sustitución puede hacerse sin pérdida sustancial de la función relevante. En la realización de tales cambios, las sustituciones de residuos aminoacídicos parecidos puede hacerse sobre la base de la similitud relativa de los sustituyentes en la cadena lateral, por ejemplo, su tamaño, carga, hidrofobicidad, hidrofiliidad, y demás, y tales sustituciones pueden ser sometidas a ensayos respecto de su efecto sobre la función del péptido mediante pruebas de rutina.

Tal como se lo emplea aquí, el término "aminoácidos" se refiere a aquellos L-aminoácidos comúnmente hallados en proteínas presentes en la naturaleza, D-aminoácidos y tales aminoácidos cuando han sido modificados. De acuerdo a esto, los aminoácidos de la invención pueden incluir, por ejemplo: Acido 2-Aminoadípico; Acido 3-Aminoadípico; beta-Alanina; Acido beta-Aminopropiónico; Acido 2-Aminobutírico; Acido 4-Aminobutírico; Acido piperidínico; Acido 6-Aminocaproico; Acido 2-Aminoheptanoico; Acido 2-Aminoisobutírico; Acido 3-Aminoisobutírico; Acido 2-Aminopimélico; Acido 2,4 Diaminobutírico; Desmosina; Acido 2,2'-Diaminopimélico; Acido 2,3-Diaminopropiónico; N-Etilglicina; N-Etilasparagina; Hidroxilisina; alo-Hidroxilisina; 3-Hidroxiprolina; 4-Hidroxiprolina; Isodesmosina; alo-Isoleucina; N-Metilglicina; sarcosina; N-Metilisoleucina; 6-N-metil-lisina; N-Metilvalina; Norvalina; Norleucina; y Ornitina.

En algunas realizaciones, las sustituciones conservadas de aminoácidos pueden hacerse donde un residuo aminoacídico es sustituido por otro con un similar valor de hidrofiliidad (e.g., dentro de un valor de más o menos 2,0; o más o menos 1,5; o más o menos 1,0; o más o menos 0,5), donde los siguientes, que pueden ser un aminoácido con un índice hidropático de aproximadamente -1,6 tal como Tyr (-1,3) o Pro (-1,6), son asignados a residuos aminoacídicos (tal como se detalla en la Pat. US No. 4.554.101, incorporada aquí como referencia): Arg (+3,0); Lys (+3,0); Asp (+3,0); Glu (+3,0); Ser (+0,3); Asn (+0,2); Gln (+0,2); Gly (0); Pro (-0,5); Thr (-0,4); Ala (-0,5); His (-0,5); Cys (-1,0); Met (-1,3); Val (-1,5); Leu (-1,8); Ile (-1,8); Tyr (-2,3); Phe (-2,5); y Trp (-3,4).

En realizaciones alternativas, las sustituciones conservativas de aminoácidos pueden ser hechas donde un residuo aminoacídico es sustituido por otro con un similar índice hidropático (e.g., dentro de un valor de más o menos 2,0; o más o menos 1,5; o más o menos 1,0; o más o menos 0,5). En tales realizaciones, a cada residuo aminoacídico se le puede asignar un índice hidropático sobre la base de su hidrofobicidad y carga características, tal como sigue: Ile (+4,5); Val (+4,2); Leu (+3,8); Phe (+2,8); Cys (+2,5); Met (+1,9); Ala (+1,8); Gly (-0,4); Thr (-0,7); Ser (-0,8); Trp (-0,9); Tyr (-1,3); Pro (-1,6); His (-3,2); Glu (-3,5); Gln (-3,5); Asp (-3,5); Asn (-3,5); Lys (-3,9); y Arg (-4,5).

En realizaciones alternativas, las sustituciones conservativas de aminoácidos pueden ser hechas usando familias de matrices de similitud disponibles al público (60, 70, 102, 103, 94, 104, 86). La matriz PAM está basada sobre recuentos derivados de un modelo evolutivo, mientras que la matriz Blosum emplea recuentos derivados de bloques altamente conservados dentro de un alineamiento. Un valor de similitud superior a cero tanto en la matriz PAM como en la matriz Blosum puede usarse para realizar sustituciones conservativas de aminoácidos.

En realizaciones alternativas, las sustituciones conservativas de aminoácidos pueden ser realizadas donde un residuo aminoacídico es sustituido por otro de la misma clase, donde los aminoácidos están divididos en no polares, ácidos, básicos y neutros, de la siguiente manera: no polares: Ala, Val, Leu, Ile, Phe, Trp, Pro, Met; ácidos: Asp, Glu; básicos: Lys, Arg, His; neutros: Gly, Ser, Thr, Cys, Asn, Gln, Tyr.

Los cambios conservativos de aminoácidos pueden incluir la sustitución de un L-aminoácido por el correspondiente D-aminoácido, por un D-aminoácido conservativo, o por una forma natural de aminoácido codificada no genéticamente, así como una sustitución conservativa de un L-aminoácido. Los aminoácidos naturales no genéticamente codificados incluyen: beta-alanina, ácido 3-amino-propiónico, ácido 2,3-diaminopropiónico, ácido alfa-aminoisobutírico, ácido 4-amino-butírico, N-metilglicina (sarcosina), hidroxiprolina, ornitina, citrulina, t-butilalanina, t-butilglicina, N-metilisoleucina, fenilglicina, ciclohexilalanina, norleucina, norvalina, 2-naftilalanina, piridilalanine, 3-benzotienil alanine, 4-clorofenilalanina, 2-fluorofenilalanina, 3-

fluorofenilalanina, 4-fluorofenilalanine, penicilamina, ácido 1,2,3,4-tetrahydro-isoquinolina-3-carboxílico, beta-2-tienilalanina, sulfoxido de metionina, homoarginina, N-acetil lisina, ácido 2-amino butírico, ácido 2-amino butírico, ácido 2,4,-diamino butírico, p-aminofenilalanina, N-metilvalina, homocisteína, homoserina, ácido cisteico, ácido epsilon-amino hexanoico, ácido delta-amino valérico, o ácido 2,3-diaminobutírico.

En realizaciones alternativas, los cambios conservativos de aminoácidos incluyen cambios basados en consideraciones de hidrofiliidad o hidrofobicidad, tamaño o volumen, o carga. Los aminoácidos pueden caracterizarse generalmente como hidrofóbicos o hidrofílicos, dependiendo primariamente de las propiedades de la cadena lateral del aminoácido. Un aminoácido hidrofóbico exhibe una hidrofobicidad superior a cero, y un aminoácido hidrofílico exhibe una hidrofiliidad inferior a cero, basado en la escala de hidrofobicidad consensuada normalizada de Eisenberg *et al.* (71). Los aminoácidos hidrofóbicos genéticamente codificados incluyen: Gly, Ala, Phe, Val, Leu, Ile, Pro, Met y Trp, y los aminoácidos hidrofílicos genéticamente codificados incluyen: Thr, His, Glu, Gln, Asp, Arg, Ser y Lys. Los aminoácidos hidrofóbicos no genéticamente codificados incluyen a la t-butilalanina, mientras que los aminoácidos hidrofílicos no genéticamente codificados incluyen a la citrulina y la homocisteína.

Los aminoácidos hidrofóbicos o hidrofílicos pueden además subdividirse según las características de sus cadenas laterales. Por ejemplo, un aminoácido aromático es un aminoácido hidrofóbico con una cadena lateral que contiene al menos un anillo aromático o heteroaromático, el cual puede contener uno o más sustituyentes tales como -OH, -SH, -CN, -F, -Cl, -Br, -I, -NO₂, -NO, -NH₂, -NHR, -NRR, -C(O)R, -C(O)OH, -C(O)OR, -C(O)NH₂, -C(O)NHR, -C(O)NRR, etc., donde R es independientemente (C1-C6) alkil, (C1-C6) alkil sustituido, (C1-C6) alkenil, (C1-C6) alkenil sustituido, (C1-C6) alkinil, (C1-C6) alkinil sustituido, (C5-C20) aril, (C5-C20) aril sustituido, (C6-C26) alkaril, (C6-C26) alkaril sustituido, heteroaril de 5-20 miembros, heteroaril de 5-20 miembros sustituido, alkheteroaril de 6-26 miembros o alkheteroaril de 6-26 miembros sustituido. Los aminoácidos aromáticos genéticamente codificados incluyen: Phe, Tyr y Trp, mientras que los aminoácidos aromáticos no genéticamente codificados incluyen: fenilglicina, 2-naftilalanina, beta-2-tienilalanina, ácido 1,2,3,4-tetrahydro-

isoquinolina-3-carboxílico, 4-clorofenilalanina, 2-fluorofenilalanina, 3-fluorofenilalanina, y 4-fluorofenilalanina.

Un aminoácido apolar es un aminoácido hidrofóbico con una cadena lateral que está libre de carga a pH fisiológico y la cual tiene uniones en las cuales un par de electrones compartidos en común por dos átomos está sostenido igualmente por cada uno de los átomos (i.e., la cadena lateral no es polar). Los aminoácidos apolares codificados genéticamente incluyen Gly, Leu, Val, Ile, Ala y Met, mientras que los aminoácidos apolares no genéticamente codificados incluyen a la ciclohexilalanina. Los aminoácidos apolares pueden además ser subdivididos para incluir aminoácidos alifáticos, los cuales son aminoácidos hidrofóbicos con una cadena lateral de hidratos de carbono alifáticos. Los aminoácidos alifáticos genéticamente codificados incluyen Ala, Leu, Val e Ile, mientras que los aminoácidos alifáticos no genéticamente codificados incluyen a la norleucina.

Un aminoácido polar es un aminoácido hidrofílico con una cadena lateral que está libre de carga a pH fisiológico, pero que tiene una unión en la cual el par de electrones compartidos en común por dos átomos está sostenido más próximamente por uno de los átomos. Los aminoácidos polares codificados genéticamente incluyen Ser, Thr, Asn y Gln, mientras que los aminoácidos polares no genéticamente codificados incluyen citrulina, N-acetil lisina y sulfóxido de metionina.

Un aminoácido ácido es un aminoácido hidrofílico con una cadena lateral con un valor de pKa inferior a 7. Los aminoácidos ácidos típicamente tienen cadenas laterales cargadas negativamente a pH fisiológico a causa de la pérdida de un ión hidrógeno. Los aminoácidos ácidos codificados genéticamente incluyen Asp y Glu. Un aminoácido básico es un aminoácido hidrofílico con una cadena lateral con un valor de pKa superior a 7. Los aminoácidos básicos típicamente tienen cadenas laterales cargadas positivamente a pH fisiológico debido a la asociación con un ión hidronio. Los aminoácidos básicos codificados genéticamente incluyen Arg, Lys y His, mientras que los aminoácidos básicos no genéticamente codificados incluyen los aminoácidos no cíclicos ornitina, ácido 2,3-diaminopropiónico, ácido 2,4-diaminobutírico, y homoarginina.

Quien sea diestro en el arte apreciará que las clasificaciones de más arriba no son absolutas y que un aminoácido puede ser clasificado en más de una categoría. Además, los aminoácidos pueden ser clasificados sobre la base de su comportamiento conocido y/o sus propiedades químicas, físicas o biológicas características basadas en ensayos específicos o en comparación con aminoácidos previamente identificados. Los aminoácidos también pueden incluir porciones bifuncionales con cadenas laterales similares a aminoácidos.

Los cambios conservativos también pueden incluir la sustitución de una porción químicamente derivatizada por un residuo no derivatizado, por ejemplo, por reacción de un grupo lateral funcional de un aminoácido. De esta manera, estas sustituciones pueden incluir compuestos cuyos grupos amino libres han sido derivatizados a amino hidroclouros, grupos p-tolueno sulfonilo, grupos carbobenzoí, grupos t-butiloxycarbonilo, grupos cloroacetilo o grupos formilo. En forma similar, los grupos carboxilo libres pueden ser derivatizados para formar sales, metil y etil ésteres u otros tipos de ésteres o hidrazidas, y las cadenas laterales pueden ser derivatizadas para formar derivados O-acilo u O-alkilo para grupos hidroxilo libres o N-im-bencilhistidina para el nitrógeno del imidazol de la histidina. Los análogos de péptidos también incluyen aminoácidos que han sido químicamente alterados, por ejemplo, por mutilación, por amidación del aminoácido C-terminal por una alquilamina tal como la etilamina, etanolamina o etilen diamina, o por acilación o metilación de una cadena lateral del aminoácido (tal como acilación del grupo epsilon amino de la lisina). Los análogos de péptidos también pueden incluir el reemplazo del enlace amida del péptido por una mida sustituida (por ejemplo, grupos de la fórmula $-C(O)-NR$, donde R es (C1-C6) alquil, (C1-C6) alkenil, (C1-C6) alkinil, (C1-C6) alfil sustituido, (C1-C6) alkenil sustituido, o (C1-C6) alkinil sustituido) o un isómero de un enlace amida (por ejemplo, $-CH_2NH-$, $-CH_2S-$, $-CH_2CH_2-$, $-CH=CH-$ (cis y trans), $-C(O)CH_2-$, $-CH(OH)CH_2-$ ó $-CH_2SO-$).

El compuesto puede estar ligado covalentemente, por ejemplo, por polimerización o conjugación, para formar homopolímeros o heteropolímeros. Pueden emplearse espaciadores y enlazadores, típicamente compuestos por pequeñas

moléculas neutras tales como aminoácidos que están libres de carga en condiciones fisiológicas. Los enlaces pueden lograrse de varias maneras. Por ejemplo, los residuos de cisteína pueden agregarse a los extremos del péptido, y varios péptidos pueden ligarse covalentemente por oxidación controlada. Alternativamente, pueden usarse agentes heterobifuncionales tales como los agentes formadores de disulfuro/amida o los agentes formadores de tioéter/amida. El compuesto también puede ligarse a otro compuesto que puede, por ejemplo, modular una respuesta inmunogénica. El compuesto también puede ser restringido, por ejemplo, por tener porciones cíclicas.

Los péptido o análogos de péptidos pueden sintetizarse por técnicas químicas regulares, por ejemplo, por síntesis automática mediante la metodología de síntesis de soluciones o de fase sólida. Los sintetizadores automáticos de péptidos pueden adquirirse comercialmente y emplean técnicas bien conocidas para el arte. Los péptidos y análogos de péptidos también pueden prepararse mediante tecnología de ADN recombinante usando métodos regulares tales como los descritos en, por ejemplo, Sambrook *et al.* (110) o Ausubel *et al.* (111). En general, los compuestos candidatos son identificados a partir de extensas bibliotecas de productos naturales como de extractos sintéticos (o semi-sintéticos) o bibliotecas químicas según métodos conocidos para el arte. Aquéllos diestros en el campo del descubrimiento y desarrollo de drogas comprenderán que la fuente precisa de extractos o compuestos de prueba no es crítica para el/los método/s de la invención. Por lo tanto, puede examinarse virtualmente cualquier número de extractos o compuestos químicos mediante los métodos ejemplificados y descritos aquí. Los ejemplos de tales extractos o compuestos incluyen –pero no se limitan– a extractos basados en plantas, hongos, procariontes o animales, caldos de fermentación, y compuestos sintéticos, así como modificaciones de compuestos existentes. También existen numerosos métodos para generar síntesis al azar o dirigida (e.g., semi-síntesis o síntesis total) de cualquier número de compuestos químicos, incluyendo –pero no limitado– a compuestos basados en sacáridos, lípidos, péptidos y ácidos nucleicos. Se encuentran comercialmente disponibles bibliotecas de compuestos sintéticos. Alternativamente, se encuentran comercialmente disponibles bibliotecas de compuestos naturales en forma de extractos bacterianos, fúngicos, vegetales y animales, a partir de varias fuentes, incluyendo Biotics (Sussex, Reino Unido), Xenova (Slough, Reino Unido), Harbor Branco Oceanographic Institute (Ft.

Pierce, FL, EEUU) y PharmaMar, MA, EEUU. Además pueden generarse, si se desea, bibliotecas de polipéptidos de agentes patógenos A/E producidos sintéticamente según métodos conocidos para el arte, e.g., por métodos regulares de extracción y fraccionamiento. Más aún, si se desea, cualquier biblioteca o compuesto puede modificarse fácilmente mediante métodos químicos, físicos o biológicos regulares.

Cuando se verifica que un extracto crudo modula la virulencia, se requiere el fraccionamiento adicional del extracto principal positivo para aislar los constituyentes químicos responsables del efecto observado. Por lo tanto, el objetivo del proceso de extracción, fraccionamiento y purificación es la cuidadosa caracterización e identificación de una entidad química dentro del extracto crudo, con propiedades moduladoras de la virulencia. Los mismos ensayos descritos aquí para la detección de actividades en mezclas de compuestos pueden emplearse para purificar el componente activo y para someter a prueba sus derivados. Los métodos de fraccionamiento y purificación de tales extractos heterogéneos son conocidos para el arte. Si se desea, los compuestos verificados como agentes útiles para el tratamiento se modifican químicamente de acuerdo a métodos conocidos para el arte. Los compuestos identificados como terapéuticos, profilácticos, diagnósticos, o con otro valor, pueden ser subsecuentemente analizados mediante un modelo de *Citrobacter* o bovino para infecciones patogénicas A/E, o cualquier otro modelo animal para infecciones patogénicas A/E.

Vacunas

Una vacuna es una composición que incluye materiales que producen una respuesta inmune deseada. Una vacuna puede seleccionar, activar o expandir células B y T de memoria del sistema inmune para, por ejemplo, permitir la eliminación de agentes infecciosos tales como patógenos A/E o componentes de los mismos. En algunas realizaciones, una vacuna incluye un portador adecuado, tal como un adyuvante, que es un agente que actúa de una manera no específica para aumentar la respuesta inmune a un antígeno específico o a un grupo de antígenos, permitiendo la reducción de la cantidad de antígeno en cualquier dosis dada de vacuna, o la reducción de la frecuencia de la dosificación requerida para generar la respuesta inmune

deseada. Una respuesta inmune deseada puede incluir la protección total o parcial contra la propagación (presencia en las heces de un animal infectado, e.g., un mamífero) o la colonización (presencia en el intestino de un animal infectado, e.g., un mamífero) por un agente patógeno A/E. Por ejemplo, una respuesta inmune deseada puede incluir cualquier valor desde 10% a 100%, e.g., 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% de protección contra la propagación o colonización por un agente patógeno A/E en un animal vacunado, en comparación a un animal no vacunado.

Las vacunas de acuerdo a la invención pueden incluir a los polipéptidos y moléculas de ácido nucleico aquí descritas, o a fragmentos inmunogénicos de las mismas, y pueden ser administradas mediante cualquier forma de administración conocida para el arte o descrita aquí. En algunas realizaciones de la invención, la vacuna puede incluir un agente patógeno A/E vivo, un agente patógeno A/E muerto, o componentes de los mismos. Los agentes patógenos vivos, que pueden ser administrados en la forma de una vacuna oral, pueden contener alteraciones genéticas no revertibles que afectan la virulencia del agente patógeno A/E, pero no su inducción de una respuesta inmune. Una vacuna viva puede ser capaz de colonizar el intestino de un animal inoculado, e.g., un mamífero.

En algunas realizaciones, los polipéptidos y las moléculas de ácido nucleico aquí descritas, o fragmentos inmunogénicos de los mismos, o las bacterias mutadas (e.g., bacterias atenuadas) descritas aquí pueden ser administradas a aves de corral, e.g., pollos, patos, pavos, etc., de forma tal que produzcan una respuesta inmune, e.g., aumento de anticuerpos, en las aves de corral. Los huevos, o productos de los mismos, obtenidos de tales aves de corral, que exhiban una respuesta inmune contra los polipéptidos o moléculas de ácido nucleico descritas aquí, o contra fragmentos inmunogénicos de las mismas, pueden ser administrados a un animal, e.g., humanos, vacas, cabras, ovejas, etc., para producir una respuesta inmune a los polipéptidos y moléculas de ácido nucleico descritas aquí, o a fragmentos inmunogénicos de las mismas, en el animal. Métodos para cultivar anticuerpos en aves de corral, y para administrar tales anticuerpos, se describen en, por ejemplo: Patente US 5.750.113

emitida a Cook el 12 de mayo de 1998; Patente US 6.730.822 emitida a Ivarie *et al.* el 4 de mayo de 2004; y las publicaciones 113-117 citadas aquí.

Las vacunas de acuerdo a la invención pueden ser suplementadas adicionalmente mediante al agregado de antígenos recombinantes o purificados tales como EspA, EspB, EspD, EspP, Tir, toxina Shiga 1 ó 2, y/o intimina, usando técnicas regulares conocidas para el arte. Por ejemplo, se ha descrito la producción de recombinantes y el uso de las proteínas de la EHEC O157:H7 tales como EspA, Tir, EspB e intimina (Publicación PCT No. WO 97/40063; Publicación PCT No. WO 99/24576; 51).

Cultivo celular

Los agentes patógenos A/E pueden ser cultivados según cualesquiera de los métodos conocidos para el arte o descritos aquí. Por ejemplo, los agentes patógenos A/E pueden ser cultivados primero en medio Luria-Bertani (LB) por un periodo de aproximadamente 8 a 48 horas, o aproximadamente 12 a 24 horas, y luego diluidos aproximadamente entre 1:5 y 1:100, e.g., 1:67, o aproximadamente entre 1:5 y 1:25, o aproximadamente 1:10, en medio mínimo M-9 suplementado con NaHCO_2 ó NaHCO_3 20-100 mM, o aproximadamente 30-50 mM, o NaHCO_2 ó NaHCO_3 aproximadamente 44 mM; MgSO_4 4-20 mM, o aproximadamente 5-10 mM; o MgSO_4 aproximadamente 0,8 mM a 8 mM, glucosa 0,1 a 1,5%, o aproximadamente 0,2 a 1%, o glucosa aproximadamente al 0,4%, casaminoácidos aproximadamente entre 0,05 y 0,5%, o aproximadamente entre 0,07 y 0,2%; o casaminoácidos aproximadamente al 0,1%. Los cultivos generalmente se mantienen aproximadamente a 37 grados C, opcionalmente en 2-10% de CO_2 , u opcionalmente en alrededor de 5% de CO_2 , y cultivados hasta una densidad óptica de aproximadamente 600 nm de 0,7 a 0,8. Luego se sacan las células enteras mediante cualquier medio apropiado, e.g., microfiltración o centrifugación, y el sobrenadante puede concentrarse, e.g., 10-1000 veces o más, por ejemplo 100 veces, usando diálisis, ultrafiltración y métodos similares. La proteína total puede determinarse fácilmente mediante métodos bien conocidos para el arte.

Los sobrenadantes de los cultivos celulares pueden ser producidos con cultivos de cualquier agente patógeno A/E, por ejemplo, EHEC, como se describe aquí o como es conocido para aquéllos diestros en el arte, incluso agentes patógenos A/E de tipo salvaje o mutantes. Generalmente, el agente patógeno A/E se cultiva en un medio apropiado, en condiciones que favorecen la secreción del antígeno de tipo III (Patentes US No. 6.136.554 y 6.165.743) (51, 74).

Aislamiento e identificación de genes adicionales

Sobre la base de las secuencias de nucleótidos y aminoácidos descritas aquí, por ejemplo, en SEQ ID NOs:1- 56 ó 59-84, es posible el aislamiento y la identificación de genes adicionales usando técnicas regulares. Cualquier agente patógeno A/E puede servir como fuente de tales genes.

En algunas realizaciones, las secuencias de ácido nucleico aquí descritas pueden usarse para diseñar sondas o *primers*, incluso sondas o *primers* de oligonucleótidos degenerados, sobre la base de la secuencia de cualquiera de las hebras de ADN. Las sondas o *primers* pueden luego usarse para examinar bibliotecas genómicas o de ADNc en busca de genes de otros agentes patógenos A/E, usando técnicas regulares de amplificación o hibridación.

En algunas realizaciones, las secuencias de aminoácidos descritas aquí pueden usarse para generar anticuerpos u otros reactivos que pueden usarse para seleccionar polipéptidos de agentes patógenos A/E que se unan a estos anticuerpos.

En algunas realizaciones, pueden identificarse socios de enlace mediante la marcación de los polipéptidos de la invención (e.g., aquéllos sustancialmente idénticos a las SEQ ID NOs; 22-43, 59 ó 73-84) con una secuencia epítipo (e.g., FLAG ó 2HA), y la introducción en células hospedadoras por transfección con un vector apropiado que contiene una secuencia de ácido nucleico codificadora de un polipéptido de la invención, o bien por emisión bacteriana endógena de tipo III, seguida de inmunoprecipitación e identificación del socio de enlace. Las células HeLa pueden infectarse con cepas que expresan las fusiones FLAG o 2HA, seguido por lisis e

inmunoprecipitación con anticuerpos anti-FLAG y anti-2HA. Los socios de enlace pueden identificarse por espectroscopia de masa. Si el polipéptido de la invención no es producido en cantidades suficientes, tal método puede no emitir suficiente proteína marcada como para identificar a su socio. Como parte de un acercamiento complementario, cada polipéptido de la invención puede ser clonado en un vector de transfección de mamíferos fusionado a, por ejemplo, 2HA, GFP y/o FLAG. Luego de la transfección, las células HeLa pueden ser lisadas y el polipéptido marcado puede ser inmunoprecipitado. El socio de enlace puede identificarse por SDS PAGE seguido de espectroscopía de masa.

En algunas realizaciones, los polipéptidos de la invención pueden marcarse, sobreproducirse, y usarse en columnas de afinidad y en inmunoprecipitaciones para identificar compuestos blanco y/o confirmar aquéllos identificados. Pueden usarse proteínas marcadas con FLAG, HA y/o His para que tales columnas de afinidad extraigan los factores de las células hospedadoras a partir de los extractos celulares, y cualesquiera de los aciertos pueden ser validados por ensayos regulares de enlace, curvas de saturación, y otros métodos como los que se describen aquí o son conocidos para aquéllos diestros en el arte.

En algunas realizaciones, puede usarse un sistema bacteriano de dos híbridos para estudiar las interacciones proteína-proteína. Las secuencias de ácido nucleico aquí descritas, o las secuencias sustancialmente idénticas a las mismas, pueden clonarse en el plásmido cebo pBT del sistema de dos híbridos, y puede usarse una biblioteca de bazo de murdo, con 5×10^9 clones independientes y disponible comercialmente, como biblioteca blanco para los cebos. Los aciertos potenciales pueden caracterizarse adicionalmente recuperando los plásmidos y retransformándolos para reducir los falsos positivos que resultan de las variantes del cebo clonal y clones blanco de la biblioteca que activan a los genes reportadores independientes del cebo clonado. Los aciertos reproducibles pueden estudiarse más en profundidad como se describe aquí.

En algunas realizaciones, una cepa de patógeno A/E que expresa GrIA, por ejemplo una cepa EHEC O157 que expresa un GrIA clonado, puede usarse para el análisis proteómico de los proteomas A/E de tipo III secretados, mediante, por ejemplo,

el análisis en gel 2D de los sobrenadantes. Además, se pueden usar agentes patógenos A/E completos (e.g., series de EHEC) para definir cuáles genes son regulados por GrIA. La virulencia puede ser comprobada según lo descrito aquí o según la forma conocida por aquéllos diestros en el arte.

Una vez que las secuencias codificadoras han sido identificadas, éstas pueden aislarse mediante técnicas regulares de clonación, e insertarse en cualquier vector o replicón apropiado para, por ejemplo, la producción de polipéptidos. Tales vectores y replicones incluyen, sin limitación: bacteriófago X (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pKT230 (bacterias gram-negativas), pGV1 106 (bacterias gram-negativas), pLAFRI (bacterias gram-negativas), pME290 (bacterias gram-negativas no-*E. coli*), pHV14 (*E. coli* y *Bacillus subtilis*), pBD9 (*Bacillus*), pIJ61 (*Streptomyces*), pUC6 (*Streptomyces*), Ylp5 (*Saccharomyces*), YCpl9 (*Saccharomyces*) o virus del papiloma bovino (células de mamífero).

En general, los polipéptidos de la invención pueden producirse en cualquier célula hospedadora transformada o transfectada con un vector apropiado (69). El método de transformación o transfección y la elección del vehículo de expresión dependerá del sistema hospedador seleccionado. Puede usarse una gran variedad de sistemas de expresión, y la célula hospedadora exacta a ser usada no es crítica para la invención. Por ejemplo, un polipéptido de acuerdo a la invención puede producirse en un hospedador procariótico (e.g., *E. coli*) o en un hospedador eucariótico (e.g., *Saccharomyces cerevisiae*, células de insecto, e.g., células Sf21, o células de mamífero, e.g., células NIH 3T3, HeLa o COS). Tales células pueden conseguirse de una amplia variedad de orígenes (e.g., la American Type Culture Collection, Manassus, VA.). Los sistemas bacterianos de expresión para la producción de polipéptidos incluyen al sistema de expresión pET de *E. coli* (Novagen, Inc., Madison, Wis.) y al sistema de expresión pGEX (Pharmacia).

Ensayos

Los compuestos candidatos, incluyendo polipéptidos, moléculas de ácido nucleico y pequeñas moléculas, pueden ser seleccionados y sometidos a prueba

mediante una variedad de técnicas descritas aquí o conocidas para aquéllos diestros en el arte. Un compuesto que reduce el nivel de expresión de cualquiera de los polipéptidos o moléculas de ácido nucleico de la invención puede ser útil, por ejemplo, como un terapéutico contra una infección patogénica A/E.

Los ensayos de selección pueden realizarse, por ejemplo, midiendo la expresión génica por análisis Western Blot, usando cualquier fragmento apropiado de ácido nucleico de acuerdo a la invención como una sonda de hibridación, donde el nivel de expresión génica en presencia del compuesto candidato es comparado con el nivel de expresión génica en ausencia del compuesto candidato. Alternativamente, o adicionalmente, el efecto de un compuesto candidato puede determinarse por el nivel de expresión o secreción del polipéptido, usando por ejemplo inmunoprecipitación o Western Blot.

Otros ensayos pueden ser realizados de la siguiente manera:

Confirmación de Secreción de Tipo III y Translocación hacia Células Hospedadoras

Para determinar cuáles de los compuestos candidatos requieren el TTSS para secreción, cada gen o porción de los mismos puede fusionarse a una FLAG por su extremo C, y los sobrenadantes pueden colectarse a partir de los WT y de los mutantes TTSS, por ejemplo la EHEC WT y el mutante isogénico *EscN* tipo III, tal como se describe aquí o como es conocido para el arte. Métodos alternativos para determinar secreción incluyen el examen de sobrenadantes por pérdida de producto secretado en la cepa mutante, o el desarrollo de anticuerpos para la proteína y el uso de análisis Western para los sobrenadantes de WT y de tipo III. Para confirmar que ninguna de las proteínas de interés es un componente TTSS, los sobrenadantes de cada uno de los compuestos candidatos, cultivados en condiciones de inducción de tipo III, pueden ser examinados en busca de secreción de tipo III. El TTSS LEE secreta dos clases de proteínas: el translocón (RspA, B y D), el cual está montado sobre la superficie de la célula bacteriana, y los efectores, que son translocados directamente dentro de la célula hospedadora. Los compuestos candidatos pueden ser sometidos a prueba para

determinar si son efectores o translocadores. Por ejemplo, los efectores putativos marcados con FLAG en EHEC o en otros patógenos A/E pueden usarse para infectar células epiteliales HeLa cultivadas, y examinarse por microscopía de inmunofluorescencia luego de colorear con anticuerpos anti-FLAG. Tal visualización usualmente demuestra la emisión bacteriana dentro de la célula hospedadora, y a menudo indica a cuál organela está dirigido el efector (e.g., Tir a membrana, NleA a Golgi). Pueden usarse anticuerpos a varios compartimientos celulares para confirmar la localización. Para complementar la visualización, las células HeLa infectadas pueden fraccionarse en fracciones citosol, insoluble y membrana, usando métodos conocidos de fraccionamiento (30), y análisis Western realizado con anticuerpos anti-FLAG para definir a cuál de las fracciones celulares está dirigido el efector. Como control, las células pueden ser infectadas con el efector marcado expresado en una cepa TTSS defectuosa. Si está dirigido a la fracción de membrana, puede usarse un tratamiento con alto nivel de sales o pH alcalino para determinar si es una proteína integral de membrana. Si el compuesto candidato se expresa a un nivel bajo, y la detección de translocaciones por inmunofluorescencia es dificultosa, se pueden hacer fusiones genéticas a la adenilato ciclasa, una enzima que requiere un cofactor citoplasmático de mamífero (calmodulina) para su actividad (87).

Efectos sobre la formación del pedestal y la captación

Dado que la condensación de actina y la formación del pedestal son sellos distintivos de los agentes patógenos A/E, los compuestos candidatos pueden ser examinados en cuanto a acumulación de actina y formación de pedestales en, por ejemplo, células epiteliales HeLa cultivadas.

La invasión de EPEC y EHEC es otro fenotipo de cultivo celular que se mide fácilmente y brinda una indicación de interacciones con células epiteliales cultivadas y de la capacidad para alterar el citoesqueleto del hospedador (los mutantes del tipo III no invaden, como tampoco lo hacen las cepas sin Tir ni intimina). Los niveles de invasión varios compuestos candidatos pueden ser comparados en patógenos A/E wt y de tipo salvaje, por ejemplo, EHEC WT y mutantes TTSS en células HeLa, usando un ensayo de protección de gentamicina.

Además, la capacidad de un compuesto candidato para bloquear la fagocitosis de macrófagos cultivados puede someterse a prueba, ya que EPEC y EHEC inhiben la fagocitosis en macrófagos cultivados por inhibición de la actividad de la PI-3 kinasa en el hospedador de una manera tipo III-dependiente (Celli, J., M. Olivier y B. B. Finlay, 2001, Enteropathogenic *Escherichia coli* mediates antiphagocytosis through the inhibition of PI 3-kinase-dependent pathways. *Embo. J.*, 20: 1245-58). Si cualesquiera compuestos candidatos son incapaces de inhibir la fagocitosis, puede realizarse un Segundo ensayo de inhibición de la PI-3 kinasa.

Efectos sobre monocapas epiteliales polarizadas

La integridad de los empalmes entre células juega un rol preponderante en la diarrea. Además de la formación de pedestales, los agentes patógenos A/E causan otros efectos LEE de tipo III sobre células epiteliales polarizadas, incluyendo pérdida de microvellosidades (destrucción de microvellosidades) y pérdida de resistencia eléctrica transmonocapa, una medida de firmes empalmes. Usando monocapas intestinales humanas polarizadas de células Caco-2, puede realizarse microscopía electrónica de barrido de alta resolución sobre monocapas infectadas con patógenos A/E WT (e.g., EHEC WT), mutantes TTSS de patógenos A/E, e.g., EHEC escN, y cada uno de los compuestos candidatos. Las monocapas pueden ser infectadas por diversos tiempos, lavadas y procesadas para SEM usando técnicas regulares (66) y examinadas en busca de pérdida de resistencia eléctrica luego de infectar células Caco-2 polarizadas.

Efectos sobre la inmunidad innata y la inflamación

Un tema que crece rápidamente con respecto a los agentes patógenos es la capacidad para inhibir la inmunidad innata y las respuestas inflamatorias. Tales efectos han sido reportados para patógenos A/E tales como EHEC y EPEC, y estos ensayos pueden usarse para examinar a los compuestos candidatos en cepas de patógenos A/E WT y mutantes TTSS. Por ejemplo, EHEC provoca la inhibición de NF- κ B, resultando en la supresión de varias citoquinas tales como IL-8, IL-6 e IL-1 α en células HeLa (80), y este proceso requiere del TTSS LEE. Los compuestos candidatos pueden ser sometidos a

ensayos de la inhibición de estos factores seguidos de, por ejemplo, infección de células HeLa, usando métodos regulares tales como RT-PCR (PCR en tiempo real), y ensayos de ELISA disponibles comercialmente.

Estudios funcionales basados en información de localización

Además de ensayos fenotípicos, los compuestos candidatos pueden ser sometidos a prueba según su localización en una célula hospedadora. Por ejemplo, si un compuesto candidato se localiza en el Golgi, puede ser ensayado para determinar si afecta la función del Golgi, incluyendo estudios bioquímicos para examinar la glicosilación, y ensayos funcionales del Golgi en levaduras que expresan al compuesto candidato. Si el compuesto candidato se localiza en mitocondria, pueden utilizarse ensayos sobre apoptosis y otras funciones mitocondriales. Si el compuesto candidato tiene como blanco al retículo endoplasmático, pueden diseñarse ensayos de síntesis y secreción de proteínas. Si el blanco es el núcleo, pueden llevarse a cabo estudios transcripcionales.

Rol en la virulencia

Los índices competitivos (CI) se han usado extensivamente para determinar el rol de factores de virulencia menores, y si dos factores de virulencia pertenecen al mismo "curso" de virulencia (63). En síntesis, dos cepas caracterizadas por diferentes resistencias a antibióticos, son coinfectadas en un animal y luego de un tiempo de incubación apropiado las bacterias son cosechadas y se determina la proporción entre las dos cepas. Un valor de 1 indica igual virulencia. Si los compuestos identificados tienen un efecto sobre la virulencia, pueden determinarse su CI en comparación a WT. También pueden usarse el CI para determinar a cuáles cursos de virulencia pertenece el compuesto candidato. Por ejemplo, los CI pueden determinarse comparando cada uno contra WT. Cuando se comparan los mutantes simples y contra un mutante doble, una proporción CI de 1 indica que éstos pertenecen al mismo curso general de virulencia, mientras que cualquier valor distinto de 1 indica que están en diferentes "cursos" de virulencia.

Estudios de microscopía

Para caracterizar enfermedades causadas por agentes patógenos A/E pueden usarse técnicas de microscopía, incluidas microscopía electrónica de transmisión (TEM) y microscopía electrónica de barrido (SEM), para examinar folículos linfoides en placas de Peyer ileales distales de conejos infectados con REPEC para confirmar la formación de lesiones A/E (63), microscopía confocal para mostrar la emisión hacia vellosidades múridas, y análisis histológico del proceso de la enfermedad en ratones infectados.

Estas técnicas pueden emplearse para examinar compuestos candidatos en modelos animales apropiados, por ejemplo, ratones infectados con varias cepas de *Citrobacter* portadoras de mutaciones en las secuencias candidatas. Para cada compuesto candidato puede emprenderse un estudio detallado para seguir la progresión (o la falta de progresión) de la enfermedad en este sistema. Además, puede determinarse el nivel de colonización de estos mutantes en las superficies del intestino. Pueden usarse cepas marcadas con antibióticos para infectar animales, seguido de la cosecha del tejido intestinal. El tejido puede homogeneizarse y las bacterias cuantificadas por recuento en placas selectivas.

Una de las principales características de la infección por agentes patógenos A/E, e.g., infección por *Citrobacter*, es una extensa inflamación. Los eventos celulares de la inflamación pueden seguirse por microscopía confocal para varios mutantes en compuestos candidatos. Etiquetando tejidos con anticuerpos o lectinas, seguido por microscopía confocal, las células inflamatorias que reclutadas hacia el lugar de la infección por los mutantes pueden ser definidas en comparación con la cepa parental. Existen anticuerpos para varios factores innatos de respuesta que también pueden usarse para analizar los fenotipos mutantes durante la infección, examinando factores innatos de respuesta y células tales como macrófagos, neutrófilos, la producción de iNOS, células dendríticas, etc.

Estudios histológicos

Pueden realizarse estudios histológicos de tejido coloreado. Por ejemplo, pueden estudiarse secciones ileales de conejo infectado con REPEC y coloreadas con hematoxilina y eosina, y cepas con delecciones en compuestos candidatos para comparar la inflamación y el daño tisular, y caracterizar las infecciones por patógenos A/E (58). Una coloración similar puede realizarse con cepas de *Citrobacter* carentes de compuestos candidatos en ratones. Existen varias otras coloraciones histológicas que pueden definir más ampliamente la inflamación asociada con *Citrobacter* y mutantes isogénicos, incluyendo Giemsa y Azul de Toluidina O (para morfología general), Acido Schiff Periódico (colorea hidratos de carbono, permitiendo el examen de la capa de mucus intestinal y células en copa), coloración Gram, esterasa de cloracetato (una coloración de células inflamatorias), y un ensayo de caspasa (para apoptosis). La inmunohistoquímica permite la utilización de anticuerpos dirigidos contra antígenos de células bacterianas y de mamíferos.

Anticuerpos

Los compuestos de la invención pueden usarse para preparar anticuerpos para los polipéptidos de la invención, mediante técnicas regulares de preparación (45) o conocidas para aquéllos diestros en el arte.

Por ejemplo, una secuencia codificadora de un polipéptido de la invención puede purificarse hasta el grado necesario para la inmunización de conejos. Para intentar minimizar los problemas potenciales de baja afinidad o especificidad de los antisueros, pueden generarse dos o tres contrucciones de polipéptidos para cada proteína, y cada construcción se inyecta en por lo menos dos conejos. Los antisueros pueden cultivarse por inyecciones en serie, preferentemente incluyendo al menos tres inyecciones de refuerzo. Las inmunizaciones primarias pueden llevarse a cabo con el adyuvante completo de Freund, y las inmunizaciones subsiguientes con el adyuvante incompleto de Freund. Las titulaciones de anticuerpos pueden ser monitoreadas por análisis de Western Blot e inmunoprecipitación, usando la proteína purificada. Los sueros inmunes pueden purificarse por afinidad mediante la Proteína Acoplada a CNBr-Sefarosa. La especificidad del antisuero puede determinarse mediante un panel d proteínas no relacionadas. Alternativamente o adicionalmente, los péptidos correspondientes a

regiones inmunogénicas relativamente únicas de un polipéptido de la invención pueden generarse y acoplarse a la hemocianina cerradura de lapa (KLH) a través de una lisina C-terminal introducida. El antisuero para cada uno de estos péptidos puede purificarse por afinidad sobre péptidos conjugados a BSA, y la especificidad puede ser examinada por ELISA y Western Blot usando conjugados de péptidos, y por Western Blot e inmunoprecipitación.

Alternativamente, se preparan anticuerpos monoclonales que se unen específicamente a cualquiera de los polipéptidos de la invención, de acuerdo a la tecnología regular de hibridomas (91, 90, 89, 78). Una vez producidos, los anticuerpos monoclonales también pueden ser analizados en cuanto a su reconocimiento específico por Western Blot o inmunoprecipitación. Los anticuerpos que se unen específicamente al polipéptido de la invención son considerados como útiles; tales anticuerpos pueden emplearse, e.g., en un inmunoensayo. Alternativamente, los anticuerpos monoclonales pueden prepararse usando el polipéptido de la invención descrito más arriba y una biblioteca de exposición de fagos (112).

En algunas realizaciones, los anticuerpos pueden producirse usando fragmentos de polipéptidos que parecen ser inmunogénicos, por criterios tales como alta frecuencia de residuos cargados. Los anticuerpos pueden ser adaptados para minimizar una respuesta inmune adversa por parte del hospedador mediante, por ejemplo, el uso de anticuerpos quiméricos que contienen un dominio de unión al antígeno de una especie, y la porción Fc de otra especie, o mediante el uso de anticuerpos hechos a partir de hibridomas de la especie apropiada.

En algunas realizaciones, los anticuerpos contra cualquiera de los polipéptidos aquí descritos pueden emplearse para tratar o prevenir la infección por un agente patógeno A/E.

Modelos Animales

Los compuestos pueden ser fácilmente examinados en varios modelos animales de infección patógena A/E.

El modelo de infección m rda por *Citrobacter* es un modelo de una enfermedad natural, y el modelo de propagaci n de EHEC en bovinos adultos es un modelo de transporte de una no-enfermedad natural (las vacas no est n enfermas, sin embargo EHEC no es parte de la flora normal). Para el modelo de *Citrobacter*, los mutantes pueden ser examinados en cuanto a su virulencia en ratones, tal como se describe m s arriba. Pueden usarse ratones *Knock out* (KO) para estudiar el rol de las secuencias candidatas en la infecci n. Se han desarrollado varias l neas de ratones KO, incluso definiendo el rol de iNOS (20), c lulas T y C (108), Tlr-4 y estableciendo un intervalo de variaci n de la susceptibilidad del hospedador (54), tanto como Nck.

Para el modelo bovino, los mutantes EHEC pueden investigarse en vacas a aies (ver, por ejemplo, la publicaci n PCT No. WO 02/053181). En breve, se administran 10⁸ CFU de mutante o WT O157 a las vacas via intubaci n oral-g strica. 14 d as post-inoculaci n se monitorea la propagaci n fecal mediante el cultivo selectivo en placas, y las colonias se verifican por PCR multiplex, niveles de propagaci n y an lisis histol gico y microsc pico de la regi n perianal donde se concentra EHEC (97). Otro modelo es el modelo de asa intestinal de bovino, en el cual las asas intestinales son inyectadas con EHEC y examinadas histol gicamente y microsc picamente en distintos momentos, y tambi n se cuantifican las bacterias adherentes mediante cultivo en placas.

Un modelo de infecci n natural en conejos con RDEC-1 puede realizarse de la siguiente manera: Se colectan por centrifugaci n los cultivos bacterianos desarrollados durante una noche, y se resuspenden en un ml de soluci n salina con buffer de fosfato. Se someten a ayuno conejos blancos de Nueva Zelanda (peso: 1,0 a 1,6 kg) y luego se les inoculan dentro del est mago, mediante tubos orog stricos, cinco ml de bicarbonato de sodio est ril al 2,5% y un ml de RDEC-1   cepas mutantes de secuencias candidatas (2:-5x10¹⁰). La misma dosificaci n de bacterias se inocula en cada conejo al d a siguiente. Cada conejo es pesado diariamente y se colecta el contenido fecal de bacterias con hisopos rectales y a partir de pellets fecales. Los hisopos rectales se hacen rodar en mitad de la superficie de placas MacConkey con  cido nalid xico. Cinco pellets fecales la misma cantidad de deposiciones l quidas se colectan de cada conejo y se resuspenden en tres ml de soluci n salina con buffer de fosfato y 0,1 ml de cada

suspensión fecal se cultiva en placa MacConkey con ácido nalidíxico. El crecimiento de colonias nalidíxico-resistentes se clasifica de la siguiente manera: 0, sin crecimiento; 1, colonias ampliamente espaciadas; 2, colonias estrechamente espaciadas; 3, crecimiento confluyente de colonias. Los tejidos se extirpan inmediatamente después del sacrificio por inyección intravenosa de ketamina y sobredosificando con fenobarbital de sodio. La proporción de la colonización bacteriana en tejidos intestinales se verifica de la siguiente manera: los segmentos intestinales (10 cm), a excepción del ciego, son doblemente ligados en sus extremos proximal y distal, y disecados entre las partes doblemente ligadas, luego lavados con 10 ml de solución salina con buffer de fosfato helada. Se agrega un gramo del contenido viscoso del ciego a 9 ml de solución salina con buffer de fosfato. Las suspensiones de solución salina con buffer de fosfato resultantes se diluyen y se cultivan en placas MacConkey con ácido nalidíxico. Las muestras de tejido se extirpan con un sacabocados de 9 mm de diámetro, se lavan tres veces con solución salina con buffer de fosfato, se agregan a dos ml de solución salina con buffer de fosfato helada, y se homogeneizan con un homogeneizador, luego las muestras seriales diluidas se cultivan en placas MacConkey. El número de bacterias adheridas a cada tejido por centímetro cuadrado se calcula de la siguiente manera: $CFU/cm^2 = \text{el número de bacterias/placa} \times \text{factor de dilución} \times 2 \text{ ml} / 0,452$.

Terapéutica y Diagnóstico

Los polipéptidos y moléculas de ácido nucleico descritas aquí pueden usarse como terapéuticos, por ejemplo, para la preparación de vacunas o composiciones terapéuticas, o la construcción de patógenos A/E que están atenuados en su virulencia. Tales patógenos A/E pueden construirse como se describe aquí, por ejemplo, designando *primers* basados en las secuencias de ácido nucleico aquí descritas, y usando una técnica de intercambio alélico basada en el gen *sac-B* (29). Los polipéptidos y moléculas de ácido nucleico pueden usarse solas o combinadas entre ellas o con otras moléculas apropiadas, tales como EspA, EspB, EspD, EspP, Tir, toxina Shiga 1, toxina Shiga 2, ó moléculas de intimina.

En algunas realizaciones, las moléculas de ácido nucleico aquí descritas pueden usarse en técnicas antisentido. Por "antisentido", tal como se usa aquí en referencia a

ácidos nucleicos, se quiere significar una secuencia de ácido nucleico que es complementaria a la hebra codificadora de una molécula de ácido nucleico, por ejemplo, un gen, tal como los genes *nleA*, *nleB*, *nleC*, *nleD*, *nleE* o *nle*, o que es complementario a una secuencia de nucleótidos sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NO: 1-21 ó 60-72 ó un fragmento o variante de las mismas. En algunas realizaciones, una molécula de ácido nucleico antisentido es una con la capacidad de disminuir el nivel del polipéptido codificado por el gen complementario cuando ambos están expresados en una célula. En algunas realizaciones, el nivel del polipéptido es disminuido en cualquier número desde al menos 10% a por lo menos 25%, o en cualquier número desde al menos 25% a por lo menos 50%, o en cualquier número desde al menos 50% a por lo menos 75%, o en cualquier número desde al menos 75% a 100%, cuando se lo compara al nivel del polipéptido en una célula que expresa sólo el gen, y no la molécula de ácido nucleico antisentido complementaria.

En algunas realizaciones, la expresión de un gen o región codificadora o no codificadora de interés puede estar inhibida o anulada mediante tecnología basada en ARN de interferencia (ARNi), un tipo de silenciamiento génico post-transcripcional. El ARNi puede usarse para crear un "knockout" funcional, i.e., un sistema en el cual la expresión de un gen o región codificadora o no codificadora de interés está reducida, resultando en una reducción general del producto codificado. Como tal, el ARNi puede realizarse apuntando a un ácido nucleico de interés o fragmento o variante del mismo, para, a cambio, reducir su expresión y el nivel de actividad del producto que codifica. Tal sistema puede usarse para estudios funcionales del producto, tanto como para tratar infecciones por patógenos A/E. El ARNi está descrito, por ejemplo, en las solicitudes de patentes US publicadas No. 20020173478 (Gewirtz, publicada el 21 de noviembre de 2002) y No. 20020132788 (Lewis et al.; publicada el 7 de noviembre de 2002) (79, 67). Los reactivos y los kits para realizar ARNi pueden conseguirse comercialmente de, por ejemplo, Ambion Inc. (Austin, TX, EEUU) y New England Biolabs Inc. (Beverly, MA, EEUU).

El agente inicial para ARNi en algunos sistemas parece ser la molécula ARNd correspondiente a un ácido nucleico "blanco". Se cree que el ARNd es luego clivado en breves ARN de interferencia (ARNsi) con una longitud de 21-23 nucleótidos (duplex de

19-21bp, cada uno con dos proyecciones nucleotídicas 3'). La enzima que parece afectar este primer paso de clivaje a sido llamada "Cortadora" y es categorizada como miembro de la familia ARNasa III de ribonucleasas ARNs-específicas. Alternativamente, el ARNi puede ser realizado mediante la introducción directa dentro de la célula -o la generación dentro de la célula mediante la introducción dentro de la célula- de un precursor apropiado (e.g., un vector, etc.) de tal ARNsi o molécula similar a un ARNsi. Un ARNsi puede estar asociado con otros componentes intracelulares para formar un complejo silenciador inducido por ARN (RISC). El RISC así formado puede subsecuentemente transformar en "blanco" a una transcripción de interés a través de las interacciones de apareamiento de bases entre su componente ARNsi y la transcripción "blanco" por virtud de la homología, resultando en el clivaje de la transcripción "blanco" aproximadamente a los 12 nucleótidos a partir del extremo 3' del ARNsi. Luego el ARNm "blanco" es clivado y se reduce el nivel del producto proteico que éste codifica.

El ARNi puede ser realizado mediante la introducción en las células de apropiados ARNsi -o moléculas similares a ARNsi- sintetizados *in vitro*. El ARNi puede ser realizado, por ejemplo, usando ARN sintetizado químicamente (64), para lo cual las moléculas apropiadas de ARN pueden sintetizarse químicamente mediante métodos conocidos. Alternativamente, pueden usarse vectores de expresión apropiados para transcribir tal ARN *in vitro* o bien *in vivo*. La transcripción *in vitro* de hebras antisentido (codificadas por secuencias que están presentes en el mismo vector o en vectores separados) puede realizarse usando por ejemplo la ARN polimerasa T7, en cuyo caso el vector puede comprender una secuencia de codificación apropiada ligada operativamente a un promotor T7. En algunas realizaciones, el ARN transcripto *in vivo* puede ser procesado (e.g., usando ARNasa III de *E. coli*) *in vitro* hasta un tamaño conductivo para ARNi. Las transcripciones con sentido y antisentido combinadas para formar un duplex de ARN que es introducido en una célula "blanco" de interés. Pueden usarse otros vectores que expresan pequeños ARN en horquilla (ARNsh) que pueden procesarse en moléculas similares a ARNsi. Se conocen para el arte varios métodos basados en vectores (65, 93, 95, 98, 99, 105, 109). Se conocen para el arte varios métodos para introducir tales vectores dentro de células, ya sea *in vitro* o *in vivo* (e.g., terapia génica).

Por lo tanto, en una realización, la expresión de un polipéptido con una secuencia de aminoácidos sustancialmente idéntica a la secuencia de cualquiera de las SEQ ID NO: 22 a 43 puede ser inhibida por la introducción o la generación dentro de una célula de un ARNsi o molécula similar al ARNsi correspondiente a una molécula de ácido nucleico codificadora del polipéptido o fragmento del mismo, o a un homólogo de ácido nucleico del mismo. En varias realizaciones, tal método puede implicar la administración directa del ARNsi o molécula similar al ARNsi dentro de la célula, o el uso de los métodos basados en vectores arriba descritos. En una realización, el ARNsi o molécula similar al ARNsi tiene una longitud menor a aproximadamente 30 nucleótidos. En otra realización, los ARNsi o moléculas similares al ARNsi tienen una longitud de aproximadamente 21-23 nucleótidos. En una realización, los ARNsi o moléculas similares al ARNsi comprenden una porción duplex de 19-21 bp, y cada hebra tiene una proyección 3' de dos nucleótidos. En algunas realizaciones, el ARNsi o molécula similar al ARNsi es sustancialmente idéntica a un ácido nucleico codificador del polipéptido o un fragmento o variante (o un fragmento de una variante) del mismo. Tal variante es capaz de codificar una proteína que tiene la actividad de un polipéptido codificada por las SEQ ID NO: 22-43. En algunas realizaciones, la hebra con sentido del ARNsi o molécula similar al ARNsi es sustancialmente idéntica a las SEQ ID NO: 1-21 o a un fragmento de las mismas (el ARN teniendo residuos U en lugar de T de la secuencia del ADN).

En algunas realizaciones, los anticuerpos desarrollados contra los polipéptidos de la invención pueden usarse como terapia contra la infección por agentes patógenos A/E.

En algunas realizaciones, las moléculas polipetídicas y de ácido nucleico de la invención pueden usarse para detectar la presencia de un agente patógeno A/E en una muestra. Las moléculas de ácido nucleico pueden usarse para diseñar sondas o *primers* que podrían, por ejemplo, hibridizarse al ADN de un patógeno A/E en una muestra, o podrían usarse para amplificar el ADN de un patógeno A/E en una muestra mediante, por ejemplo, técnicas de reacción en cadena para polimerasas. Los polipéptidos podrían usarse por ejemplo para desarrollar anticuerpos que se unan

específicamente a un polipéptido expresado por un agente patógeno A/E. Tales sondas o primers o anticuerpos pueden ser etiquetados en forma detectable. Se entiende por "etiquetados en forma detectable" cualesquiera medios para marcar e identificar la presencia de una molécula, e.g., una sonda o *primer* oligonucleotídico, un gen o fragmento del mismo, o una molécula de ADNc. Los métodos para etiquetar en forma detectable una molécula son bien conocidos para el arte e incluyen, sin limitación, el etiquetamiento radiactivo (e.g., con un isótopo tal como ^{32}P ó ^{35}S) y etiquetamiento no radiactivo tal como el etiquetamiento enzimático (por ejemplo, usando peroxidasa de rábano o fosfatasa alcalina), etiquetamiento quimioluminiscente, etiquetamiento fluorescente (por ejemplo, usando fluoresceína), etiquetamiento bioluminiscente, o detección por anticuerpos de un ligando adherido a una sonda. Se incluye también en esta definición a una molécula que es etiquetada en forma detectable por medios indirectos, por ejemplo, una molécula que está unida con una primera porción (tal como biotina) que está, a su vez, unida a una segunda porción que puede ser observada o sometida a ensayo (tal como la estreptavidina etiquetada con fluoresceína). Las etiquetas incluyen también a la digoxigenina, luciferasas, y a la acurina.

Composiciones, Dosificaciones y Administración de Uso Farmacéutico y Veterinario

Los compuestos incluyen las moléculas polipeptídicas y de ácido nucleico descritas aquí, así como compuestos identificados según los métodos de la invención. Los compuestos de acuerdo a la invención pueden proveerse solos o en combinación con otros compuestos (por ejemplo, moléculas de ácido nucleico, pequeñas moléculas, polipéptidos, péptidos o análogos de péptidos), en presencia de un liposoma, un adyuvante, o cualquier portador farmacéuticamente aceptable, en una forma apropiada para ser administrados a mamíferos, por ejemplo, humanos, vacas, ovejas, etc. Si se desea, el tratamiento con un compuesto de acuerdo a la invención puede combinarse con más terapias tradicionales y existentes para una infección por patógenos A/E.

Puede emplearse la práctica farmacéutica convencional para proveer formulaciones o composiciones apropiadas para administrar los compuestos a sujetos infectados por un agente patógeno A/E. Puede emplearse cualquier ruta de

administración adecuada, por ejemplo, parenteral, intravenosa, subcutánea, intramuscular, intracraneal, intraorbital, oftálmica, intraventricular, intracapsular, intraespinal, intracisternal, intraperitoneal, intranasal, aerosol o administración oral. Las formulaciones pueden ser en forma de suspensiones o soluciones líquidas; tabletas o cápsulas; polvos, gotas nasales o aerosoles.

Los métodos para hacer formulaciones son bien conocidos para el arte (57). Las formulaciones para administración parenteral pueden, por ejemplo, contener excipientes, agua estéril, o solución salina, polialquilen-glicoles tales como polietilenglicol, aceites de origen vegetal, o naftalenos hidrogenados. Para controlar la liberación de los compuestos pueden usarse polímeros de láctido, copolímeros de láctido/glicólido, o copolímeros de polioxietileno-polioxipropileno biocompatibles y biodegradables. Otros sistemas de administración parenteral potencialmente útiles incluyen a las partículas de copolímeros de etilen-vinil-acetato, bombas osmóticas, sistemas de infusión implantables, y liposomas. Las formulaciones para inhalación pueden contener excipientes, por ejemplo lactosa, o pueden ser soluciones acuosas que contengan, por ejemplo, éter de polioxietilén-9-laurilo, glicolato y deoxicolato, o pueden ser soluciones aceitosas para administración en la forma de gotas nasales, o como un gel. Para composiciones terapéuticas o profilácticas, los compuestos se administran a un individuo en cantidades suficientes para detener o reducir una infección por agentes patógenos A/E.

"Cantidades suficientes" de un compuesto de acuerdo a la invención implican cantidades terapéuticamente efectivas o cantidades profilácticamente efectivas. Las "cantidades terapéuticamente efectivas" se refieren a cantidades efectivas, a dosificaciones y por periodos de tiempo necesarios, como para lograr el resultado terapéutico deseado, tal como la reducción de la colonización o la diseminación de un patógeno A/E. Una cantidad terapéuticamente efectiva de un compuesto puede variar según factores tales como el estado de la enfermedad, la edad, el sexo y el peso del sujeto, y la capacidad del compuesto para producir la respuesta deseada en el sujeto. Los regímenes de dosificación pueden ajustarse para proveer la óptima respuesta terapéutica. Una cantidad terapéuticamente efectiva también es una en la cual los efectos terapéuticamente beneficiosos pesan más que cualquier efecto tóxico o

perjudicial del compuesto. Una "cantidad profilácticamente efectiva" se refiere a una cantidad efectiva, a dosificaciones y por periodos de tiempo necesarios, para lograr el resultado profiláctico deseado, tal como la prevención de la colonización por un agente patógeno A/E. Típicamente, una dosis profiláctica se usa en sujetos antes –o durante un estadio temprano de- la enfermedad, de manera que una cantidad profilácticamente efectiva puede ser menor que una cantidad terapéuticamente efectiva. Un valor apropiado para cantidades terapéuticamente u profilácticamente efectivas de un compuesto puede ser cualquier número entre 0,1 nM – 0,1 M; 0,1 nM – 0,05 M; 0,05 nM – 15 µM; ó 0,01 nM – 10 µM.

Debe advertirse que los valores de dosificación pueden variar según la severidad de la condición a ser aliviada. Para cualquier sujeto en particular, pueden ajustarse regímenes de dosificación específicos a lo largo del tiempo de acuerdo con los requerimientos individuales y el juicio profesional de la persona que administra o supervisa la administración de las composiciones. Los intervalos de dosificación aquí expuestos son sólo a modo de ejemplo y no limitan los intervalos de dosificación que pueden seleccionarse por profesionales de la medicina. La cantidad de compuesto activo en la composición puede variar de acuerdo a factores tales como el estado de la enfermedad, la edad, el sexo y el peso del individuo. Los regímenes de dosificación pueden ajustarse para proveer la óptima respuesta terapéutica. Por ejemplo, puede administrarse un único bolo, varias dosis divididas pueden administrarse a lo largo del tiempo o la dosis puede ser proporcionalmente reducida o aumentada según lo indicado por las exigencias de la situación terapéutica. Puede resultar ventajoso formular composiciones parenterales en la forma de unidades de dosificación para facilitar la administración y la uniformidad de las dosis.

En el caso de formulaciones para vacunas, una cantidad efectiva de un compuesto de la invención puede proveerse sola o en combinación con otros compuestos, y con uno o más adyuvantes inmunológicos para inducir una respuesta inmune. Los adyuvantes de acuerdo a la invención pueden incluir emulsificadores, dipéptidos de muramilo, avridina, adyuvantes acuosos tales como hidróxido de aluminio, adyuvantes basados en quitosano, saponinas, aceites, Amphigen®, LPS, extractos de paredes celulares bacterianas, ADN bacteriano, complejos bacterianos, oligonucleótidos

sintéticos y combinaciones de los mismos (100). El adyuvante puede incluir un extracto de paredes celulares de *Mycobacterium phlei* (*M. phlei*, MCWE) (Patente US No. 4.744.984), un complejo de paredes celulares de ADN de *M. phlei* (M-DNA) o de micobacterias (MCC). Los emulsionantes incluyen agentes emulsificadores naturales o sintéticos. Los agentes emulsificadores sintéticos incluyen agentes iónicos (e.g., sales de potasio, sodio y amonio de ácido láurico y oleico, las sales de calcio, magnesio y aluminio de ácidos grasos, i.e., jabones de metales, y sulfonatos orgánicos tales como el sulfato de lauril sodio), agentes catiónicos (e.g., bromuro de cetiltrimetilamonio), y agentes no iónicos (e.g., ésteres de glicerol tales como el monoestearato de glicerol, ésteres y éteres de polioxietilenglicol, y los ésteres de ácidos grasos de sorbitano tales como el monopalmitato de sorbitano y sus derivados de polioxietileno, e.g., monopalmitato de sorbitan polioxietileno). Los agentes emulsificantes naturales incluyen incluyen: acacia, gelatina, lecitina y colesterol. Otros adyuvantes adecuados incluyen un aceite que puede ser un aceite simple, una mezcla de aceites, una emulsión de aceite en agua (e.g., EMULSIGENTM, EMULSIGEN PLUSTM o VSA3), o una emulsión no-aceite en agua (e.g., una emulsión de aceite, una emulsión de agua en aceite, o una emulsión de agua-en aceite-en agua). El aceite puede ser un aceite mineral, un aceite vegetal o un aceite animal. Los aceites animales adecuados incluyen, por ejemplo, aceite de hígado de bacalao, aceite de hipogloso, aceite de lacha, aceite de "reloj anaranjado", o aceite de hígado de tiburón, todos los cuales están disponibles comercialmente. Los aceites vegetales adecuados incluyen, sin limitación, aceite de canola, aceite de almendra, aceite de semilla de algodón, aceite de maíz, aceite de oliva, aceite de maní, aceite de cártamo, aceite de sésamo o aceite de soja. Alternativamente, pueden usarse varias bases nitrogenadas alifáticas como adyuvantes con las formulaciones de vacunas. Por ejemplo, los adyuvantes inmunológicos conocidos incluyen aminas, compuestos de amonio cuaternarios, guanidinas, benzamidinas y tiouroniums (75). Los compuestos adyuvantes específicos incluyen el bromuro dimetildioctadecil-amonio (DDA) (Patente US No. 5.951.988) (88, 59, 85, 68, 84, 82, 83) y el N, N-dioctadecil-N, N-bis (2-hidroxietil) propanodiamina ("avridina") (Patente US No. 4.310.550, Patente US No. 5.151.267) (62). Un adyuvante de acuerdo a la invención puede incluir por ejemplo un aceite mineral y bromuro dimetildioctadecil-amonio.

Las composiciones de vacunas pueden prepararse usando técnicas corrientes que incluyen -aunque no están limitadas a- mezcla, sonicación y microfluidificación. El adyuvante puede comprender aproximadamente 10 a 50% (v/v) de la vacuna, o aproximadamente 20 a 40% (v/v) o aproximadamente 20 a 30% ó 35% (v/v), o cualquier valor dentro de estos intervalos. El compuesto puede también estar ligado a una molécula portadora, tal como albúmina de suero bovino o hemocianina cerradura de lapa para aumentar la inmunogenicidad.

En general, los compuestos de la invención deberían usarse sin causar una sustancial toxicidad. La toxicidad de los compuestos de la invención puede determinarse usando técnicas corrientes, por ejemplo, mediante pruebas en tejidos celulares o en animales experimentales y determinando el índice terapéutico, i.e., la relación entre el LD50 (la dosis letal para el 50% de la población) y el LD100 (la dosis letal para el 100% de la población). En algunas circunstancias, sin embargo, tal como en severas condiciones de enfermedad, puede ser necesario administrar sustanciales excesos de las composiciones.

Los siguientes ejemplos pretenden ilustrar realizaciones de la invención, y no debieran ser considerados como limitaciones del alcance de la invención.

EJEMPLO 1
Generación de mutantes

Las cepas de bacterias usadas son las siguientes: EHEC O157:H7 cepa 86/24 (32), EHEC EDL933 (producto de aislamiento O127:H7 de *E. coli* de tipo salvaje; 3); EHEC *escN*- (47), EPEC E2348/69 de tipo salvaje (producto de aislamiento O127:H6 de *E. coli* de tipo salvaje; 50), DBS100 de *C. rodentium* (ATCC 51459; 53), REPEC 0103:K:H2 85/150 (52), *E. coli* HB101.

Los plásmidos usados fueron los siguientes:

Identificación	Descripción y contenido Relevante	Referencia
pRE118	Vector suicida basado en <i>secB</i> para intercambio génico Kan ^r	4

pKD46	Recombinasa λ roja de expresión del plásmido, AMP ^r	5
pACYC184	Vector de clonación. CM ^r Tetr	NEB
pCR2.1-TOPO	Vector de clonación de PCR, Amp ^r Kan ^r	Invitrogen
pTOPO-2HA	Cassette de expresión basado en pCR2.1-TOPO y dirigido por <i>P_{lac}</i> para la marcación del 2HA C-terminal, Amp ^r Kan ^r	
pCReepG-2HA/BglII	Cassette de expresión basado en pACYC184 y dirigido por el promotor <i>espG</i> de <i>Citrobacter</i> para el 2HA C-terminal que marca a Cm ^r	
pCR1er	<i>C. rodentium</i> <i>ler</i> en pCR2.1-TOPO	
pCRorf11	<i>C. rodentium</i> <i>orf11/gr1A</i> en pCR2.1-TOPO	
pEHorf11	EHEC <i>orf11/gr1A</i> en pCR2.1-TOPO	
pEPorf11	EPEC <i>orf11/gr1A</i> en pCR2.1-TOPO	
pKK232-8	Derivado de pBR322 que contiene un gen <i>cat</i> sin promotor	Pharmacia
pLEE1-CAT	Derivado de pKK232-8 que porta la fusión transcripcional <i>LEE1</i> (<i>Ler</i>)- <i>cat</i> de <i>C. rodentium</i> desde los nucleótido -162 al +216.	
pLEE5-CAT	Derivado de pKK232-8 que porta la fusión transcripcional <i>LEE5</i> (<i>Tir</i>)- <i>cat</i> de <i>C. rodentium</i> desde los nucleótido -262 al +201.	

Para PCR y PCR inversa, se usó como rutina el sistema de lectura de pruebas ELONGASA Amplification System (GIBCO BRL/Life Technologies) para minimizar la tasa de error del PCR. Los productos de PCR fueron clonados utilizando el kit de clonación Topo TA de Invitrogen con pCR2 I TOPO o pCR II TOPO. La secuencia del DNA se determinó mediante el método terminador Taq Dye y un secuenciador automatizado de DNA 373A (Applied Biosystems). Para la clonación, transformación e infecciones de rutina, las bacterias se cultivaron en agar Luria-bertani (LB) o en caldo LB suplementados con los antibióticos apropiados a 37°C. Se utilizaron varios antibióticos con las siguientes concentraciones: Ampicilina 100 µg/ml, carbenicilina 100 µg/ml, kanamicina 50 µg/ml y cloranfenicol 30 µg/ml. El crecimiento en el Medio de Eagle modificado por Dulbecco (MEMD) se usó para la inducción de la expresión génica del LEE y la secreción de la proteína de tipo III.

Para generar los mutantes se usaron el intercambio alélico basado en genes *sacB* (14) y el sistema de recombinasa Rojo lambda (17). En general, se eliminó el 75% o más de la porción interna de cada gen para asegurar la interrupción de la función del gen. El primer gen se mutó mediante el método de la recombinasa rojo lambda y los genes *ler*, *orf11/gr1A*, y *cesT* se mutaron mediante el método de intercambio alélico basado en genes *sacB*. El mutante doble de *ler/orf11* se generó también utilizando el

método *sacB*. La generación de los mutantes *tir* y *escD* fue descrita en otras partes (20,56). Los mutantes eran mutantes de delección "en marco", con la introducción de un sitio de enzima de restricción (tanto *Nhe* I, *Bam* HI como *Sal* I) en el lugar de la delección, como se indica a continuación:

Nombre del mutante LEE	Longitud de la proteína (aa)	Codón eliminado (de # a #)	Elementos introducidos en el sitio de la delección.	Métodos de delección empleados
<i>Δtir</i>	129	23-97 ó 9-122	Cassette <i>Nhe</i> I o <i>aph</i> T	<i>SacB</i> o Rojo Lambda
<i>Δorf11/grfA</i>	135	23-115	<i>Nhe</i> I	<i>SacB</i>
<i>ΔescT</i>	156	25-147	<i>Nhe</i> I	<i>SacB</i>

Estos mutantes se verificaron con múltiples reacciones PCR. La complementación se sometió a prueba para los mutantes *Δtir*, *Δtir* y *Δorf11* suministrando los genes respectivos en un plásmido. Todos estos mutantes pueden ser complementados, confirmando que las mutaciones generadas por ambos métodos de intercambio alélico no afectaron la función de los genes corriente abajo y son por lo tanto no-polares.

Para hacer EHEC-pNleA-HA, la región codificadora de *nleA* fue amplificada mediante el sistema de prueba de lectura ELONGASE Amplification System (Invitrogen) y los siguientes *primers*: Z6024F:

5'AGATCTGAAGGAGATATTATGAACATTCAACCGACCATAC (SEQ ID NO: 44); Z6024R: 5'CTCGAGGACTCTTGTTTCTTCGATTATATCAAAG (SEQ ID NO: 45). Los productos PCR fueron clonados utilizando el kit de clonado TOPO TA (Invitrogen) y la secuencia de DNA se verificó usando el método del terminador Taq Dye y un secuenciador automatizado de DNA 373A (Applied Biosystems). El producto fue luego subclonado en pCRespG-2HA/BgIII, un plásmido derivado de pACYC desarrollado para conducir una expresión proteínica desde un promotor *EspG* de *C. rodentium* y para agregar dos hemaglutininas de influenza (HA) al extremo C de la proteína expresada. Las construcciones de plásmidos se introdujeron luego en EHEC y EHEC *escN* de tipo salvaje electroporación.

Se creó un mutante de delección en *nleA* en una cepa de EHEC resistente al ácido nalidíxico mediante un intercambio alélico basado en el gen *sacB* (29). Dos fragmentos de DNA que flanquean al *nleA* fueron amplificados usando DNA cromosómico de EHEC como plantilla.

El fragmento A fue amplificado por PCR usando el *primer* NT10 5'CCGGTACCTCTAACCATTGACGCACTCG (SEQ ID NO:46) y un *primer* NT11 5'AACCTGCAGAACTAGGTATCTCTAATGCC (SEQ ID NO:47) para generar un producto de 1,3 Kb.

El fragmento B fue amplificado usando el *primer* NT12 5'AACCTGCAGCTGACTATCCTCGTATATGG (SEQ ID NO:48) y un *primer* NT13 5'CCGAGCTCAGGTAATGAGACTGTCAGC (SEQ ID NO:49) para generar un producto de 1,3 Kb.

Los fragmentos A y B fueron luego digeridos con *Pst*I por 1 hora y luego la enzima fue inactivada por calor durante 20 minutos a 65°C. Aproximadamente 50 ng de cada fragmento digerido se ligó con ligasa ADN T4 por 1 hora a temperatura ambiente. La reacción de ligamiento se diluyó 1/10 y 1 µl fue agregado a un PCR utilizando *primers* NT10 y NT13. El producto PCR de 2.3 Kb resultante fue luego digerido con *Sac*I y *Kpn*I, ligado a los lugares correspondientes de pRE112 (14) y transformados en DH5αpir para generar pNT225. El pNT225 fue transformado en la cepa conjugativa SM10λpir que funcionó como cepa donante en una conjugación con el EHEC de tipo salvaje. Se seleccionaron ácidos nalidíxicos y exoconjugantes resistentes al cloranfenicol en agar LB. Los exoconjugantes se sembraron luego en agar LB con sacarosa al 5% (w/v) y sin NaCl. Las colonias resultantes fueron luego examinadas en cuanto a su sensibilidad al cloranfenicol, seguido de PCR para identificar productos de aislamiento con la delección *nleA* y con pérdida de secuencias de plásmidos. Un mutante de delección *nleA* fue luego aislado y usado para trabajo ulterior. Se generó un mutante *nleA* de *C. rodentium* usando PCR para crear, en pRE118 (14), un vector suicida con una delección interna de *nleA*.

Se usaron los siguientes *primers*: del1F: 5'GGTACCACCACACAGAATAATC (SEQ ID NO:50); del1R: 5'CGCTAGCCTATATACTGCTGTTGGTT (SEQ ID NO:51); del2F:

5'GCTAGCTGACAGGCAACTCTTGGACTGG (SEQ ID NO:52); del2R: 5'GAGCTCAACATAATTTGATGGATTATGAT (SEQ ID NO:53). El plásmido resultante fue introducido en *C. rodentium* por electroporación para crear una cepa merodiploide resistente a los antibióticos. La pérdida de secuencias del plásmido a través de un segundo evento de recombinación seleccionada como se describió más arriba. Se examinaron por PCR las colonias resistentes a la sacarosa y sensibles a los antibióticos en busca del evento de recombinación apropiado, utilizando *primers* flanqueando la región eliminada. La ausencia de NleA se verificó aplicando Western Blot a los lisados de células enteras con antisuero policlonal anti-NleA.

Evaluación de proteínas totales y secretadas / proteómica

Las proteínas secretadas fueron preparadas como se describió más arriba (51).

Las cepas de *C. rodentium* se cultivaron durante una noche en un agitador en 4 ml de caldo Luria (LB) a 37°C. Los cultivos fueron subcultivados 1 a 50 en 4 ml de medio de Eagle modificado por Dulbecco (DMEM), el cual fue pre-incubado en un incubador de cultivos de tejido (con 5% de CO₂) durante una noche, y cultivados durante 6 horas en el mismo incubador para inducir la expresión génica del LEE. Los cultivos fueron centrifugados dos veces a 13.000 rpm durante 10 min. para extraer las bacterias, y el sobrenadante fue precipitado con ácido tricloroacético (TCA) para concentrar las proteínas secretadas en el medio de cultivo. El pellet bacteriano se disolvió en buffer 2X SDS-PAGE y fue designado como proteínas bacterianas totales. Las proteínas secretadas precipitadas del sobrenadante también fueron disueltas en buffer 2X SDS-PAGE y el TCA residual fue neutralizado con 1 µl de Tris saturado. Los volúmenes de buffer usados para re-suspender el pellet bacteriano y las proteínas secretadas fueron normalizados al OD600 del cultivo para asegurar la igual carga de las muestras. Las proteínas secretadas fueron analizadas en SDSPAGE al 12% ó 17%, y coloreadas con Azul Coomassie G250 al 0,1%. Para el análisis Western Blot, las proteínas totales o secretadas fueron separadas en SDSPAGE al 10%, y transferidas a

membrana de celulosa (Bio-Rad). Los anticuerpos usados fueron anticuerpos policlonales de rata contra Tir de *Citrobacter* marcado con His, y anticuerpo monoclonal de ratón contra EspB de EPEC. Se siguieron los protocolos regulares de Western Blot ECL (Amersham Life Science).

EHEC de tipo salvaje y EHEC CVD451 se cultivaron durante la noche en medio LB. Los cultivos luego se disolvieron 1:100 en medio mínimo M-9 suplementado con NaHCO_3 44 mM, MgSO_4 8 mM, glucosa 0,4% y casaminoácidos 0.1%, y se cultivaron a 37°C en 5% de CO_2 hasta un OD600 de 0,6 a 0,8.

Las proteínas secretadas se cosecharon centrifugando los cultivos a 8000g durante 30 minutos, luego separando los sobrenadantes de los pellets. Los sobrenadantes se filtraron a través de filtros de 0,45 micrones y se determinó la concentración de proteína mediante el ensayo de BCA (Sigma). Las proteínas se prepararon para electroforesis mediante precipitación con 1/9 del volumen de TCA frío al 100%, sobre hielo durante 45 a 120 min, seguido por centrifugación a 17600g durante 30 min. Los pellets se enjuagaron en acetona fría al 100% y solubilizada en buffer laemmli 1X (para geles SDS-PAGE unidimensionales) o buffer de muestra 2D para geles 2D (urea 8M, tiourea 2M, CHAPS al 4%, Tris 20 mM, azul bromofenol al 0,002%). Para los geles 2D, se agregó DTT hasta 6 mg/ml, y buffer IPG (pH 3-10: Amersham Biosciences) hasta 0,5% antes de la carga. Se rehidrataron tiras secas Immobiline (Immobiline Dry Strips) de 18 cm (pH 3-10: Amersham) en la muestra durante la noche a 20°C. Luego las muestras fueron enfocadas a 15°C por 65.000 Vh. Luego de ser enfocadas, las tiras se equilibraron en EB + 10 mg/ml de DTT durante 15 minutos, y luego EB + 25 mg/ml de iodoacetamida durante 15 minutos (EB es Tris 50 mM, urea 6M, glicerol al 30%, SDS al 2%, pH 8,8). Las tiras equilibradas se sellaron a la parte superior geles SDS-PAGE de gran formato (acrilamida al 12% ó 14%) usando agarosa al 0,5% en buffer de corrida SDS-PAGE + azul bromofenol al 0,002%, y los geles se corrieron hasta que el frente de la coloración corriera fuera del gel.

Los geles se colorearon con Sypro Ruby siguiendo las instrucciones de los fabricantes (BioRad) y se visualizaron en un iluminador UV o por MS/MS en un espectrómetro de masa LCQ Deca Ion Trap (Thermo Finnigan) equipado con un

sistema de cromatografía líquida Nanoflow (LC Packings–Dionex). En geles visualizados con un iluminador UV, las manchas de interés se extrajeron manualmente. La digestión intra-gel de las proteínas fue realizada en el Investigator ProGest Robot (Genomic Solutions, Ann Arbor, MI), tal como es descrito (23). Las muestras con elevada abundancia de proteínas se analizaron mediante un sistema LC-MS consistente en un sistema de cromatografía líquida Nanoflow equipado con un automuestreador FAMOS (LC Packings –Dionex, San Francisco, CA), y un espectrómetro de masa LCQ Deca Ion Trap (Thermo Finnigan, San Jose, CA)(24).

Se usaron columnas PicoFrit de fase invertida PFC7515-PP18-5 (New Objective, Woburn, MA, EEUU) para la separación de péptidos, y el efluente de la columna fue rociado directamente dentro del espectrómetro de masa. Se usó un flujo de 200 nL/min, y el tiempo total de adquisición fue igual a 45 min. por muestra.

Las muestra con bajas concentraciones de proteína se analizaron en un espectrómetro de masa API QSTAR Pulsar Hybrid MS/MS (Applied Biosystems/MDS SCIEX, Concord, Canada) (12) equipado con una fuente de iones Nanospray (Proxeon, Odense, Denmark). Previo al análisis, las muestra se purificaron y concentraron en ZipTips (Millipore, Billerica, MA). El API QSTAR Pulsar también se usó para la secuenciación peptídica *de novo*.

Los espectros se analizaron contra la base de datos NCBI (Bethesda, MD) con los buscadores Mascot (Matrixscience) o Sonar (Proteometrics Canada Ltd.).

Construcción de las fusiones transcripcionales *cat* y el ensayo CAT

Los fragmentos PCR que portaban los promotores y todos los elementos reguladores corriente-arriba del *ler* (*LEE1*) y el *tir* (*LEE5*) de *Citrobacter* fueron digeridos con *Bam* HI y *Hind* III, y clonados en el plásmido pKK232-8, el cual contiene un gen *cat* carente de promotor. Las posiciones ocupadas por los fragmentos clonados son como se indican aquí. La actividad de CAT dirigida por estas fusiones en diferentes cepas de *Citrobacter* fue determinada como se describió previamente (21, 22). Las muestras se

colectaron en diferentes puntos de tiempo a partir de cultivos bacterianos desarrollados en DMEM como se describió más arriba.

Análisis de secuencias y herramientas bioinformáticas

El análisis de secuencias y la búsqueda de homologías del ADN y las proteínas con BLASTN, TBLASTN y BLASTP se llevaron a cabo usando programas disponibles en el sitio web de NCBI (<http://www.ncbi.nih.org/>). Las bases de datos usadas incluyen las del sitio de NCBI, el Centro de Genoma Sanger (<http://www.sanger.ac.uk/Projects/Microbes>), y el SwissProt (<http://www.expasy.org/sprot/>). Las posiciones del LEE PAI así como los profagos en el genoma de EHEC fueron obtenidas a partir de datos generados por el programa IslandPath (<http://www.pathogenomics.sfu.ca/islandpath/>) (13).

Análisis Southern Blot

Las muestras de ADN genómico para el análisis Southern Blot se prepararon usando el DNeasy Tissue kit (Qiagen). La sonda se preparó mediante digestión del pNleA-HA con las enzimas *SalI* y *BglII* para obtener un fragmento de 500 bp que fue etiquetado con el *kit* de etiquetado no isotópico BrightStar Psoralen-Biotin (Ambion).

Se digirieron completamente cinco microorganismos de cada muestra de ADN genómico con 25 unidades de *BamHI*, *EcoRI* y *PstI* durante una noche. Las muestras se resolvieron por electroforesis en gel de agarosa al 1%, y se transfirieron durante una noche a una membrana de nylon BrightStar-Plus (Ambion) mediante levigación pasiva y levemente alcalina. El ADN fue ligado en forma cruzada a la membrana exponiendo a ésta última a luz UV durante 2 min, seguido por 30 min de calentamiento a 80°C.

La membrana fue prehibridizada lavándola en 10 ml de buffer de hibridación ultrasensible ULTRAhyb (Ambion) a 42°C durante 30 min. Diez microlitros de la sonda preparada fueron luego agregados a la membrana prehibridizada en buffer y la sonda fue hibridizada a la membrana durante una noche a 42°C. La membrana se lavó dos veces 5 min en buffer de lavado de baja exactitud (Ambion) y 2 veces 15 min en buffer

de lavado de alta exactitud (Ambion) a temperatura ambiente. La sonda hibridizada fue detectada con el kit de detección no isotópica BrightStar BioDetect (Ambion), seguido por exposición a una película Kodak.

Generación de antisuero anti-NleA

La porción codificadora de *nleA* fue amplificada a partir del ADN genómico de EHEC y clonada en un vector de expresión marcado con his (pET28a, Novagen) usando los siguientes *primers*: 5'TTCCATATGAACATTCAACCGACC (SEQ ID NO: 54) y 5'GGAATTCAATAATAGCTGCCATCC (SEQ ID NO: 55). Este plásmido fue introducido en BL21 (λ DE3), desarrollado hasta una densidad óptica (A600) de 0,8 e inducido con IPTG 0.5 mM durante 16 horas a 20°C. La proteína marcada con his fue purificada en una columna Ni-NTA siguiendo las instrucciones de fabricante (Qiagen). Las fracciones con NleA fueron agrupadas y se agregó trombina (500:1), y la proteína se dializó durante la noche contra tris 20 mM pH 8,2; NaCl 50 mM. Al día siguiente la proteína fue cargada en una columna monoQ de FPLC y la columna fue desarrollada con un gradiente lineal de NaCl desde 50 a 500 mM. Las fracciones con NleA fueron agrupadas. Después de este paso, la proteína fue >90% pura. La proteína purificada se usó para inmunizar dos ratas macho Sprague Dawley, 300 μ g de proteína/rata, usando el adyuvante completo de Freund (Sigma), y el antisuero resultante fue purificado en su afinidad usando el soporte activado de inmunoadfinidad Affi-Gel 15 según las instrucciones del fabricante (BioRad). Para los experimentos de inmunofluorescencia, el antisuero fue ulteriormente purificado por absorción contra polvos de acetona preparados a partir de células HeLa y a partir de EHEC. NleA tal como se describe en (45). La especificidad del antisuero fue confirmada mediante el análisis Western Blot de extractos celulares de EHEC de tipo salvaje y NleA de EHEC.

Análisis inmunoblot

Las muestras para análisis Western Blot fueron resueltas por SDS-PAGE (9% a 12% de poliacrilamida). Las proteínas se transfirieron a nitrocelulosa y los *inmunoblots* se bloquearon en leche seca no grasa (NFDM) al 5% en TBS con pH 7,2, y con Tween 20 (TBST) al 0,1%, durante la noche a 4°C, y luego incubadas con anticuerpo primario

en TBST NFDM al 1% durante 1 hora a temperatura ambiente (RT). Las membranas se lavaron 6 veces en TBST y luego se incubaron con una dilución 1:5000 de peroxidasa de rábano – anticuerpo anti-ratón de cabra (H+L) (Jackson ImmunoResearch Laboratories Inc.) durante una hora a RT. Las membranas luego fueron lavadas como se describiera más arriba. Los complejos antígeno-anticuerpo se visualizaron con el kit mejorado de detección por quimioluminiscencia (Amersham), seguido por exposición a película Kodak (Perkin Elmer). Se utilizaron los siguientes anticuerpos primarios: anti-HA.11 (Covance), anti-DnaK (Stressgen), anti-EHEC Tir, anti-NleA (este estudio), anti-Calnexina (Stressgen), anti-Calreticulina (Affinity Bioreagents), anti-tubulina (Sigma).

Inmunofluorescencia

Las células HeLa se cultivaron sobre cubreobjetos de vidrio en placas para cultivo de tejidos de 24 pocillos e infectadas durante 6 horas con 1 μ l (EHEC) de cultivo del día anterior de OD~0,4. A las 6 horas postinfección las células se lavaron 3 veces en PBS con Ca^{2+} y Mg^{2+} , y fijadas en paraformaldehído al 2,5% en PBS durante 15 minutos a temperatura ambiente. Las células fueron permeabilizadas en saponina al 0,1% en PBS, bloqueadas en suero de cabra al 5% ó BSA al 5% en PBS + saponina al 0,1%, e incubadas con los siguientes anticuerpos primarios diluidos en solución bloqueante durante 1 hora a temperatura ambiente: anti-EHEC Tir, 1:1000; anti-*E. coli* O157 (Difco), 1:200; anti-NleA policlonal de rata purificado en su afinidad (este estudio), 1:100; anti-manosidasa II (gentilmente provista por la Dra. Marilyn Farquhar, UCSD), 1:1000. Luego de 3 lavados en PBS/saponina, las células se incubaron en anticuerpos secundarios (Anti-ratón, conejo, rata conjugado con Alexa-488 ó -568 [Molecular Probes], 1:400), durante 30 minutos a temperatura ambiente, lavadas 3 veces en PBS/saponina y una vez en PBS, y montadas en portaobjetos de vidrio usando mowiol + DABCO. Para la visualización de la actina polimerizada, se incluyó faloidina conjugada con Alexa-488 (Molecular Probes) con los anticuerpos secundarios a una dilución 1:100. Donde así se lo indica, las células se incubaron con 5 μ g/ml de brefeldina A (Boehringer Mannheim) durante 30 minutos antes de la fijación. Las imágenes se detectaron con un microscopio Zeiss Axioskop, se capturaron con una cámara digital Empix DVC1300, y analizadas con el software para tratamiento de

imágenes Northern Eclipse, en un microscopio confocal BioRad Radiance Plus usando el software Lasersharp.

Fraccionamiento de células hospedadoras infectadas

Para cada muestra, dos cápsulas confluentes de 100 mm de células HeLa se infectaron con EHEC-pNleA-HA de tipo salvaje o EHECescN-pNleA-HA usando una MOI inicial de 1:10. A las 6 horas postinfección, las células se lavaron tres veces con PBS helado y se sometieron a fraccionamiento bioquímico como se describiera previamente (30, 49). En breve, las células se resuspendieron en 300 μ L de buffer de homogeneización (imidazola 3 mM, sacarosa 250 mM, EDTA 0,5 mM, pH 7,4) suplementado con el cocktail inhibidor de proteasa COMPLETE (Roche) y desintegradas mecánicamente mediante el pasaje a través de una aguja de tamaño 22. El homogenato fue centrifugado a baja velocidad (3000g) durante 15 minutos a 4°C hasta transformar en pellets las células no desintegradas, las bacterias, los núcleos y los componentes citoesqueléticos (pellet de baja velocidad). El sobrenadante fue sometido a ultracentrifugación de alta velocidad (41.000g) durante 20 minutos a 4°C en un rotor TLS55 en una centrifuga TL100 (Beckman) para separar las membranas de las células hospedadoras (pellet) del citoplasma (sobrenadante). Los pellets fueron resuspendidos en 300 μ L de buffer laemmli 1X, y el sobrenadante se llevó a buffer laemmli 1X usando un stock 5X. Iguaes volúmenes de todas las fracciones se resolvieron por SDS-PAGE (9% de poliacrilamida) y se transfirieron a nitrocelulosa y se analizaron por Western Blot.

Para los estudios de extracción del NleA asociado a la membrana, se fraccionaron dos cápsulas de 100 mm de células HeLa infectadas como se describiera más arriba para cada una de las condiciones de extracción. Los pellets de alta velocidad (la fracción de membrana hospedadora) fueron resuspendidos en 300 μ L de uno de los siguientes buffers de extracción: (i) Tris 10 mM, $MgCl_2$ 5 mM, pH 7,4; (ii) Tris 10 mM, $MgCl_2$ 5 mM, NaCl 1 M, pH 7,4; (iii) $NaHCO_3$ 0.2 M, $MgCl_2$ 5 mM, pH 11,4; (iv) Tris 10 mM, $MgCl_2$ 5 mM, 1% Triton X-100 pH 7,4. La extracción se realizó sobre hielo pipeteando las muestras arriba y abajo cada 5 minutos durante 30 minutos, y las muestras fueron recentrifugadas a 100.000g durante 30 minutos. El pellet (fracción

insoluble) fue resuspendido en 300 μ L de buffer laemmli 1X; el sobrenadante (fracción soluble) fue precipitado en ácido tricloroacético al 10% sobre hielo durante 30 minutos, lavado en acetona al 100% y resuspendido en 300 μ L de buffer laemmli 1X. Igualess volúmenes fueron resueltos por SDS-PAGE (9% de poliacrilamida) y transferidos a nitrocelulosa y analizados por Western Blot.

Análisis de infección de *C. rodentium* en ratones

En el servicio animal de la Universidad de British Columbia se alojaron ratones C3H/HeJ (Jackson Laboratory) y ratones NIH Swiss (Harlan Sprague-Dawley) de 5 semanas de edad, directamente de acuerdo con las normas esbozadas por el Comité de Cuidado Animal de la Universidad de British Columbia y el Consejo Canadiense sobre el Uso de Animales de Laboratorio. *C. rodentium* de tipo salvaje y el mutante de delección *nleA* se cultivaron en caldo LB durante una noche en un agitador a 200 rpm, y 100 μ L del cultivo fueron usados para infectar a los ratones por cebadura oral. El inóculo fue titulado por dilución serial y cultivo en placa, y se halló que era 4×10^8 cfu/ratón para ambos grupos. Para la infección de los ratones C3H/HeJ altamente susceptibles por *C. rodentium*, la supervivencia de los ratones infectados fue evaluada diariamente en el curso de la infección. Cuando un ratón se tomaba moribundo, éste era inmediatamente sacrificado. Para los ensayos sobre virulencia bacteriana usando ratones NIH Swiss, los animales fueron sacrificados al día 10 post-infección. Para cuantificar la hiperplasia colónica, se colectaron y pesaron los primeros 4 cm del colon distal medidos a partir del borde anal, luego de retirar cualquier pellet fecal. Para evaluar la colonización bacteriana, los tejidos colónicos más los pellets fecales fueron homogeneizados en PBS usando un Homogeneizador de Tejido Polytron, y fueron diluidos serialmente antes de cultivarlos en agar MacConkey (Difco Laboratories). El tejido colónico y los pellets fecales se combinaron para determinar la carga bacteriana total en el colon del ratón al momento de sacrificarlo. El agar MacConkey es selectivo para las bacterias Gram negativas, sobre el cual *C. rodentium* forma colonias con una morfología altamente distintiva e identificable, no típica de *E. coli* (26). Para el análisis histológico se fijaron los últimos 0,5 cm del colon de ratones infectados en formalina con buffer neutro al 10%, se procesaron, se cortaron en secciones de 3 μ m y se colorearon con hematoxilina y eosina. El análisis histológico fue hecho por el Laboratorio de

Servicios Morfológicos del Departamento de Patología y Medicina de Laboratorio de la Universidad de British Columbia.

EJEMPLO 2

Análisis de la Regulación de la Expresión Génica del LEE

Para determinar cuáles genes en el LEE son los que regulan la expresión génica del LEE en *C. rodentium*, analizamos mutantes LEE en busca de la expresión de EspB y Tir, las cuales están codificadas por los operones *LEE4* y *LEE5*, respectivamente. Los lisatos celulares totales de bacterias desarrolladas en DMEM se analizaron por Western Blot con sueros anti-Tir y anti-EspB. Nuestros resultados confirmaron el rol esencial del Ler en la expresión del LEE, ya que no se produjeron Tir ni EspB en Δ ler. Tal como se esperaba, Δ tir y Δ espB no produjeron Tir y EspB, respectivamente. No se observó Tir en Δ cesT, lo cual es consistente con el rol de los CesT como acompañante para la estabilidad y secreción de Tir (18, 19). Llamativamente, otra proteína codificada por el LEE, Orf11, también era requerida para la expresión de Tir y EspB. La expresión de Tir y EspB en Δ orf11 fue complementada por un plásmido que portaba sólo el *orf11* de *Citrobacter* (Fig. 8). El gen *orf11* es altamente conservado entre patógenos A/E, y tanto el *orf11* de EHEC como de EPEC complementaban al Δ orf11 de *Citrobacter* (Fig. 8), indicando que Orf11 es funcionalmente equivalente en la regulación positiva de la expresión génica del LEE en patógenos A/E.

El análisis secuencial indicó que Orf11/GrlA muestra 23% de identidad respecto de CaiF, un activador transcripcional de los operones *cai* y *fix* de las *Enterobacteriaceae* (15), y 37% de identidad respecto de la secuencia de aminoácidos deducida de un producto no caracterizado de *Salmonella* codificado por un gen localizado corriente debajo del operón fimbrial *std* (16) (Fig. 7). Este homólogo de *Salmonella* está indicado como SGH (Homólogo GrlA de *Salmonella*) en la figura. Estas tres proteínas contienen una estructura hélice-giro-hélice esperada que es característica de las proteínas de enlace del ADN.

Para determinar la jerarquía de Orf11 y Ler en la regulación de la expresión génica del LEE, nosotros creamos un mutante doble de *ler* y *orf11* en *C. rodentium*. Si

bien la expresión de Tir y EspB en $\Delta ler \Delta orf11$ puede restaurarse mediante la expresión de Ler *in trans*, el Orf11 similarmente expresado no tenía este efecto (Fig. 8), sugiriendo que Orf11 actúa corriente arriba respecto de Ler en la cascada regulatoria. El análisis de extensión de primers confirmó esta jerarquía regulatoria mostrando que el promotor ler de *Citrobacter* es similar al ler de EPEC y su expresión fue reducida en $\Delta orf11$.

El rol de Orf11 en la regulación de la expresión de ler fue ulteriormente demostrada mediante el monitoreo de las actividades de las fusiones transcripcionales entre las regiones regulatorias de los operones *LEE1* (Ler) (pLEE1-CAT) o *LEE5* (Tir) (pLEE5-CAT) y el gen reportador cat en cepas WT, Δler y $\Delta orf11$ de *Citrobacter* desarrolladas en Medio Esencial Modificado de Dulbecco (DMEM) durante 6 horas. La actividad de la fusión *LEE1-cat* fue disminuida en $\Delta orf11$, y la de las fusiones *LEE5-cat* fue drásticamente reducida tanto en Δler como en $\Delta orf11$. Estos resultados indican que Orf11 es un nuevo regulador positivo de la expresión de Ler, que subsecuentemente facilita la expresión de otros operones LEE. Dado que Orf11 actúa corriente arriba de Ler en la cascada regulatoria, fue denominado GrIA (regulador global del activador de LEE).

EJEMPLO 3

Identificación de efectores secretados por TTSS codificado por LEE

Los agentes patógenos A/E secretan varias proteínas hacia los cultivos de tejidos o medios mínimos, pero las proteínas secretadas son, predominantemente, los translocadores EspA, EspB y EspD. Las proteínas secretadas fueron concentradas por precipitación con TCA a partir de sobrenadantes de cultivos celulares desarrollados en DMEM y analizados por SDS-PAGE 12% seguido de coloración con Azul Coomassie. Las *C. rodentium* que portaban un plásmido con orf11/grIA secretaban por lo menos 300% más EspA, EspB y EspD que las cepas WT.

Para definir los efectores codificados por el LEE de *Citrobacter*, marcamos varias proteínas, codificadas por LEE que no están involucradas en TTS y en la adhesión a la célula hospedadora, con un epítipo doble de hemaglutinina (2HA) en el extremo carboxi, y analizamos su secreción en *C. rodentium* WT y mutantes. Solamente Tir,

EspG, EspF, EspH y Map fueron secretados por el TTSS codificado por LEE en *C. rodentium*, sugiriendo que el LEE codifica sólo para 5 efectores. Una proteína de 54 kDa no reconocida previamente (p54) fue fácilmente detectada por coloración de Coomassie en un mutante que también tenía una muy elevada secreción de Tir, sugiriendo que representa un nuevo efector putativo codificado fuera del LEE.

Para identificar a p54 y para determinar si hay otros factores que se codifiquen fuera del LEE en *C. rodentium*, aprovechamos la capacidad de GrIA para aumentar la expresión génica del LEE y/o la secreción de tipo III, e introducimos el plásmido grIA en los mutantes que secretaban efectores pero no translocadores. La sobre-expresión de GrIA aumentó en gran medida (más del 400%) la secreción de Tir en estos mutantes, sin que se secretaran translocadores. Se observaron por lo menos 6 otras proteínas secretadas, indicando que el TTSS codificado por LEE secreta varias otras proteínas no codificadas por LEE.

Para identificar estas proteínas, las proteínas secretadas fueron analizadas por geles 2-D. Dado que algunos de los efectores codificados por LEE (EspF, EspH y Map) tenían valores básicos de pI esperados, con EspF presentando un pI extremo de 11,00, las proteínas secretadas fueron en primer lugar enfocadas en Tiras Secas Immobiline con gradientes tanto ácidos (pH 3-10) como básicos (pH 6-11), y luego resueltas en SDS-PAGE 12% y 14%, respectivamente. Los geles fueron coloreados con Sypro Ruby, y las manchas de proteína seleccionadas fueron extraídas manualmente y analizadas por espectrometría de masa y por secuenciamiento peptídico *de novo*.

Este análisis confirmó que la secreción de Tir, EspF, EspG, EspH y Map codificados por el LEE era de tipo III (Tabla 2).

Tabla 2. Efectores y efectores putativos secretados por el TTSS codificado por el LEE en *C. rodentium*.

Número de serie	Nombre propuesto	MW estimado	pI estimado	Ubicación del gen	Homólogos en EHEC y otros agentes patógenos
5	Tir	68	5,0	LEE	Tir, conservado en todos los patógenos A/E EspG, conservado en todos los patógenos A/E
10	EspG	44	7,3	LEE	A/E
C1&C2	Map	23	9,0	LEE	Map, conservado en todos los patógenos A/E
C3	EspF	31	11,0	LEE	EspF, conservado en todos los patógenos A/E

C5&C6	EspH	21	8,7	LEE	EspH, conservado en todos los patógenos A/E
7	NleA	54	5,8	No en LEE	EHEC Z8024 en isla-O 71 cerca del profago CP-933P
12	NleB	39	5,9	No en LEE	EHEC Z4328 en isla-O 122, RorE de REPEC asociado a LEE, y STMF1 de <i>S. typhimurium</i> . También tiene homología con Z0985 de isla O 36
13	NleC	40	4,6	No en LEE	EHEC Z0986 en isla O 36 cerca del profago CP-933K
14	NleD	28	7,1	No en LEE	EHEC Z0990 en isla O 36, en la misma isla O que Z0985 y Z0986. También tiene similitudes con el efector HopPtoH de <i>P. syringae</i> pv. tomato
17	NleE	27	6,3	No en LEE	EHEC Z4329 en la misma isla O 122 que Z4328, RorE de REPEC asociado a LEE, y ORF122 de <i>S. flexneri</i>
19	NleF	24	4,7	No en LEE	EHEC Z8020, en la misma isla O 71 que Z8024. Algunas similitudes con proteínas hipotéticas en <i>Yersinia pestis</i> y <i>Helicobacter pylori</i>
20	NleG	26	5,8	No en LEE	Secuencia peptídica identificada: QQENAPSS(I/L)QTR. No se hallaron homólogos en base de datos

Además de los cinco efectores codificados por LEE, nosotros identificamos siete proteínas secretadas no codificadas por LEE que probablemente sean efectores. Estas proteínas secretadas codificadas fuera del LEE fueron por lo tanto designadas NleA (p54), NleB, NleC, NleD, NleE, NleF y NleG (QQENAPSS(I/L)QTR; SEQ ID NO: 59) (significando: efectores no codificados por LEE) para distinguirlas de las proteínas secretadas o efectores codificados por LEE (Esp) (Tabla 2). Entre las siete proteínas identificadas, sólo NleG resultó única para *C. rodentium*, y las otras 6 proteínas tienen homólogos altamente conservados en EHEC O157 (Tabla 2). Los homólogos de EHEC están codificados por genes agrupados en tres regiones discretas del genoma, y cada región codifica al menos para dos proteínas que muestran homología con las proteínas secretadas de *C. rodentium* (Tabla 2). Los genes Z6024 y Z8020 que codifican para los homólogos de NleA y NleF están localizados en la isla O 71 de EHEC asociados al profago CP-933P. En forma similar, los genes Z4328 y Z4329 que codifican los homólogos NleB y NleE están localizados en la isla O 122, y aquéllas que codifican para los homólogos NleC y NleD (Z0986 y Z0990) están en la isla O 36 (Tabla 2, Fig. 9). Más aún, Z4328 (isla O 122) tiene una fuerte homología con Z0985, un gen localizado cerca de Z0986 en la isla O 36.

Los homólogos de los seis nuevos genes efectores de EHEC están presentes también, y organizados similarmente, en EPEC, cuyo genoma está siendo secuenciado (<http://www.sanger.ac.uk/Projects/Microbes/>). Salvo Z6024, que mostró un 89% de identidad nucleotídica con un gen de EPEC, los otros 5 genes de EHEC mostraron una identidad superior al 95% respecto de sus homólogos en EPEC. Más aún, algunos de estos efectores también son altamente conservados en otras bacterias patogénicas. NleD/Z0990 tiene similitud con el efector tipo III HopPtoH de *P. syringae* pv. Tomate (41, 42). NleA/Z4329 tiene una homología significativa con el RorfD de EPEC de conejo (REPEC) y el Orf212 de *S. flexneri* (8, 33), mientras que NleB/Z4328 tiene una fuerte similitud con RorfE de REPEC y dos proteínas hipotéticas de *S. typhimurium*. Los genes para Z4328 y Z4329 están localizados en forma adyacente en EHEC, similarmente a la disposición de genes de *rorfD* y *rorfE* en REPEC (8). Sin embargo, *rorfD* y *rorfE* están localizados junto al LEE en REPEC, mientras que sus contrapartes en EHEC residen en una región (isla O 122) distante del LEE. La isla O 122 de EHEC que porta a Z4328 y Z4329 también contiene genes que codifican para dos citotoxinas y para un homólogo de PagC, un importante factor de virulencia regulado por PhoP/PhoQ en *S. enterica* (3, 43). Las tres islas O de EHEC que codifican para los nuevos efectores tienen un sesgo dinucleotídico y bajo contenido de GC%, sellos distintivos de los PAIs (9). Además, están asociadas con un profago o bien flanqueadas por secuencias de inserción móviles, y no están presentes en el genoma de *E. coli* no patogénicas (3), sugiriendo la transferencia horizontal de estos genes. Colectivamente, esto sugiere la importancia de estas islas y de los efectores de EHEC y *Citrobacter* recientemente identificados en cuanto a la virulencia. También indica que, al diferir entre sí, los agentes patógenos relacionados mantienen un conjunto sorprendentemente conservado de PAIs, a pesar de las diversas ubicaciones de los PAIs en el cromosoma bacteriano.

EJEMPLO 4

Identificación de NleA

Si bien la secreción de tipo III es considerada generalmente como contacto-dependiente (46), las condiciones de cultivo definidas *in vitro* pueden inducir a EHEC a secretar efectores de tipo III en el medio extracelular durante el crecimiento en cultivo líquido (28, 51). Los sobrenadantes de cultivo se prepararon a partir de EHEC de tipo

salvaje (wt) y un mutante de secreción tipo III (*escN*-), cultivado en condiciones inductoras de secreción de tipo III. El análisis de las proteínas secretadas por SDS-PAGE reveló una abundante proteína de alto peso molecular, común a las proteínas secretadas de las muestras tanto de tipo salvaje como *escN*-, y varias otras abundantes proteínas exclusivas de la muestra de tipo salvaje (Figura 10A). Las proteínas secretadas fueron separadas por electroforesis en gel bidimensional y las abundantes manchas separadas de proteínas fueron extraídas del gel y analizadas por espectroscopia de masa (Figura 10B, Tabla I).

Tabla I

Mancha No.	ID	Valor de e	Número de péptidos	mw esperado (kDa)	mw experimental (kDa)	pi esperado	pi experimental
1	EspP	5,60E-4	4	105	95	5,9	6,5
2a	Tir	2,50E-39	14	58	68	5	5
2b	Tir	2,10E-29	10	58	65	5	4,8
3	NleA	7,70E-11	6	48	50	5	5
4a	EspB	8,60E-53	19	33	38	5,1	5,2
4b	EspB	2,00E-06	2	33	38	5,1	5,1
5	EspA	1,30E-37	16	21	18	4,8	5

La mancha #1, que estuvo presente en los sobrenadantes tanto de los cultivos del tipo salvaje como en los de *escN*-, fue identificado como EspP, una proteína de EHEC codificada por un plásmido, que es secretada por un mecanismo autotransportador que es independiente de la secreción de tipo III (25). Cuatro manchas principales (#2, 3, 4, 5) fueron exclusivas de los sobrenadantes del tipo salvaje. Tres de estas manchas se identificaron como proteínas secretadas de tipo III ya conocidas, codificadas dentro del LEE: Ti (mancha #2), EspB (mancha #4) y EspA (mancha #5) (Figura 10B, Tabla I). La mancha #3 fue identificada como una proteína con un peso molecular esperado de 48 kDa, codificada por una marco de lectura abierto, dentro del genoma de EHEC pero fuera del LEE (Figura 10C). Hemos denominado a esta proteína NleA (efector A no codificado por LEE).

EJEMPLO 5

Caracterización del locus que contiene a *nleA*

El gen *nleA* está codificado en una isla O: una región del genoma de EHEC que está ausente en el genoma de la cepa no patogénica K-12 (3) de *E. coli*. La región entre el último gen conservado en la espina dorsal de *E. coli* K12 (*YciE*) y los genes que codifican las proteínas estructurales de fago contenía varios fragmentos putativos de transposasa y un fragmento putativo sitio-específico de recombinasa (Figura 11A). El análisis de esta región con Islandpath, un programa diseñado para identificar PAIs (13), reveló que todos los ORFs dentro de esta región tienen un sesgo dinucleotídico y un contenido de GC divergente respecto de la media del genoma de EHEC. 10 ORFs dentro de la región tienen un contenido de GC al menos 1 desviación estándar por debajo de la media del genoma de EHEC, mientras que 2 de los 6 ORFs tienen un contenido de GC al menos 1 desviación estándar más alto que la media del genoma de EHEC. En conjunto, estos resultados sugieren que *nleA* está localizado en un PAI que contiene genes horizontalmente transferidos. Varios otros ORFs dentro de esta región tienen características que sugieren roles en la virulencia, incluyendo un acompañante putativo (Z2565) y dos proteínas con similitud a las proteínas secretadas de tipo III de otros agentes patógenos (Z6021, Z6020),

Para investigar más en profundidad la naturaleza y distribución del gen *nleA*, se preparó una sonda *nleA* y se realizó un análisis Western Blot sobre un panel de ADN genómico de otros patógenos A/E y de otra cepa de *E. coli* no patogénica. Como se muestra en la Figura 11B, el gen *nleA* estuvo presente en todos los patógenos A/E examinados, pero ausente en *E. coli* no patogénicas. El análisis de la secuencia genómica "en progreso" de EPEC reveló que *nleA* está presente en forma muy próxima a un sitio de inserción de fago en el genoma de EPEC. *nleA* también está presente dentro de un profago de una cepa intimina-positiva no-O157 de EHEC, O84:H4, pero ausente en cepas no patogénicas de *E. coli*, *E. coli* uropatogénica, la cual no contiene el LEE. *nleA* también está ausente en otros patógenos que contienen TTSS, tales como especies de *Salmonella* y *Shigella*. Por lo tanto, *nleA* parece haber sido específicamente adquirido o retenido en los agentes patógenos A/E. El alineamiento secuencial múltiple de las secuencias del gen *nleA* de *C. rodentium*, EPEC, EHEC y O84:H4 revela un alto grado de conservación de secuencias en estos cuatro patógenos A/E (Figura 11C).

EJEMPLO 6

NleA es secretada por el sistema de secreción de tipo III codificado por el LEE

Los efectores del TTSS codificado por LEE en EHEC y EPEC descritos hasta hoy están codificados dentro del LEE, en forma muy próxima a los genes que codifican para el propio aparato de secreción. Para determinar si la secreción de NleA era dependiente del TTSS codificado por LEE, una versión de NleA marcada con un epítipo fue expresada a partir de un plásmido en EHEC de tipo salvaje y una cepa *escN*⁻, la cual es deficiente para la secreción de tipo III (47). Como se muestra en la Figura 12A, si bien la NleA marcada con HA fue expresada a niveles similares en EHEC de tipo salvaje y *escN*⁻, la proteína sólo fue secretada hacia el medio extracelular por las bacterias de tipo salvaje transformadas para NleA y HA. DnaK, una proteína bacteriana no secretada, fue usada como control para la ausencia de proteínas no secretadas en las muestras de proteína secretada (Figura 12B). Tir fue secretada en las cepas de tipo salvaje no transformadas y transformadas para NleA-HA, pero no fue secretada por las cepas *escN*⁻ (Figura 12C), verificando los fenotipos TTSS esperados. Similares resultados se obtuvieron para la expresión de la NleA marcada con epítipo en EPEC de tipo salvaje y varios mutantes de EPEC para la secreción de tipo III, indicando que la NleA también puede ser secretada por el TRSS de EPEC.

EJEMPLO 7

NleA es translocada hacia las células hospedadoras

Cuando se cultiva EHEC en condiciones inductoras de secreción de tipo III, son secretados dos tipos de proteínas hacia el medio extracelular. Para determinar si NleA era un efector translocador o translocado, investigamos la translocación y la secreción de tipo III en ausencia de NleA mediante la generación de una cepa mutante por delección. Se analizaron los perfiles de proteínas secretadas de EHEC de tipo salvaje y mutantes $\Delta nleA$. Las muestras del tipo salvaje contenían una abundante proteína de aproximadamente 50 kDa que estaba ausente en las proteínas secretadas por $\Delta nleA$ (Figura 13A). El análisis Western Blot con antisueros dirigidos contra NleA demostró que NleA de 50 kDa estaba presente en las proteínas secretadas por el tipo salvaje, y ausente en la muestra de $\Delta nleA$ (Figura 13B). Sin embargo, además de la presencia o

ausencia de NleA, los perfiles de las proteínas secretadas de las cepas de tipo salvaje y $\Delta nleA$ eran idénticos (Figura 13A). Por lo tanto, NleA no es necesaria para la secreción de otros efectores secretados de tipo III. Para determinar si NleA es necesaria para la translocación de otros efectores de tipo III hacia la célula hospedadora, se infectaron células HeLa con EHEC de tipo salvaje y EHEC $\Delta nleA$, y se monitoreó la translocación y la función de Tir mediante la coloración inmunofluorescente de las células infectadas. Se examinó la formación de pedestal por EHEC de tipo salvaje y mutantes $\Delta nleA$ sometiendo las células infectadas a inmunofluorescencia con anticuerpos anti-EHEC y anti-Tir, y visualizando la actina filamentosa con faloidina. Los resultados indicaron que la EHEC $\Delta nleA$ se adhería a las células HeLa a niveles similares que la EHEC de tipo salvaje. La coloración inmunofluorescente reveló que Tir era translocado hacia las células hospedadoras y enfocado bajo las bacterias infectantes tanto en las cepas EHEC de tipo salvaje como en las $\Delta nleA$. Para confirmar la translocación funcional de Tir, las células infectadas fueron coloreadas con faloidina fluorescente para visualizar la actina polimerizada involucrada en la formación del pedestal debajo de las bacterias adheridas. Los pedestales de actina eran evidentes en células infectadas tanto con EHEC de tipo salvaje como $\Delta nleA$, indicando que la translocación y función de otros efectores de tipo III pueden operar en ausencia de NleA. Estos resultados también indican que NleA no es necesaria para la formación del pedestal.

Dado que NleA no parecía jugar un rol en la secreción o translocación de otros efectores, investigamos si la propia NleA era translocada. Se infectaron células HeLa durante 6 horas con EHEC de tipo salvaje o *escN*- que expresara NleA marcada con HA y se sometieron a fraccionamiento subcelular y análisis Western Blot con un anticuerpo anti-HA. Como se indica en la Figura 14A, NleA es translocada hacia las células hospedadoras donde se asocia con la fracción de membrana de la célula hospedadora. No se observa translocación de NleA durante la infección de células con un mutante de secreción de tipo III que expresa NleA marcado con HA, indicando que la translocación de NleA y la asociación a la membrana del hospedador es TTSS-dependiente. El análisis Western Blot de las fracciones con anticuerpos para proteínas específicas a cada fracción conformó la ausencia de contaminación cruzada de las fracciones. La calnexina, una proteína integral de la membrana de la célula hospedadora, estaba ausente en la fracción citoplasmática del hospedador, y la tubulina, una proteína

citoplasmática de la célula hospedadora, estaba ausente en la fracción de membrana del hospedador. DnaK, una proteína bacteriana no secretada, estaba presente sólo en el pellet de baja velocidad, demostrando la ausencia de contaminación bacteriana de la membrana y fracciones citosólicas del hospedador. NleA y DnaK estaban ausentes en el pellet de baja velocidad de las células infectadas con mutante de tipo III debido a la dependencia de tipo III de la adherencia de EHEC. Para descartar que la ausencia de NleA en las muestras infectadas por el mutante de tipo III pudiera deberse a un artefacto debido a la dependencia de tipo III de la adherencia de EHEC, también llevamos a cabo experimentos similares expresando y liberando NleA-HA por EPEC de tipo salvaje y mutante de tipo III, ya que la adherencia de EPEC a las células HeLa es independiente de la secreción de tipo III. NleA estaba presente en la fracción de membrana de las células infectadas con cepas de EPEC de tipo salvaje, pero no con mutantes de tipo III. Tanto NleA como DnaK estaban presentes en las fracciones de pellets de baja velocidad tanto en las células infectadas con EPEC de tipo salvaje como en aquellas infectadas por EPEC mutantes de tipo III.

Para investigar la naturaleza de la asociación de NleA con las membranas de las células hospedadoras, las fracciones de membranas de células hospedadoras infectadas que contienen NleA marcada con HA fueron extraídas sobre hielo en diversas condiciones y recentrifugadas para obtener fracciones de membrana solubles e insolubles (Figura 14B). Estas fracciones fueron sometidas a un análisis Western Blot con anticuerpo anti-HA para detectar NleA marcado con HA. El tratamiento con alta salinidad (NaCl 1M) o pH alcalino (Na_2CO_3 0,2M, pH 11.4) elimina las proteínas que están asociadas periféricamente con membranas a través de interacciones electrostáticas o hidrofílicas, respectivamente. La asociación de NleA con las membranas de células hospedadoras resistieron la disrupción con estos tratamientos (Figura 14B, panel superior), tal como lo hizo la calnexina, una proteína integral de membrana (Figura 14B, panel medio). En contraste con esto, una proporción significativa de calreticulina, una proteína periférica de membrana, fue extraída de la fracción de membrana tanto durante el tratamiento con alta salinidad como con pH alcalino (Figura 14B, panel inferior). El tratamiento de las fracciones de membrana con el detergente no-iónico Triton X-100, el cual solubiliza las proteínas integrales de membrana tales como la calnexina (Figura 14B, panel medio), NleA casi completamente

solubilizado, dando como resultado un desplazamiento de la proteína NleA marcada con HA desde la fracción insoluble hacia la soluble (Figura 14B, panel superior). Estos resultados indican que NleA es translocada hacia las células hospedadoras, donde se comporta como una proteína integral de membrana. Efectivamente, el análisis secuencial de la proteína NleA mediante varios programas predictivos para dominios de transmembrana predice uno o dos dominios putativos de transmembrana dentro de la secuencia (Figura 11C).

EJEMPLO 8

NleA se localiza en el aparato de Golgi del hospedador

A continuación se determinó la localización subcelular de NleA dentro de la célula hospedadora. Las células HeLa fueron infectadas con EHEC de tipo salvaje o EHEC $\Delta nleA$, y sometidas a inmunofluorescencia con anticuerpos dirigidos contra NleA y manosidasa II. Algunas muestras fueron tratadas con brefeldina A durante 30 minutos antes de la fijación. Se aplicaron dos capas de color a las tinciones de NleA y manosidasa II. Las células HeLa fueron transfectadas con una construcción de expresión que codifica para una proteína de fusión GFP-NleA, y sometidas a inmunofluorescencia con un anticuerpo dirigido contra la manosidasa II.

La coloración inmunofluorescente de las células HeLa infectadas con EHEC de tipo salvaje, usando un anticuerpo anti-NleA, dio como resultado un patrón de coloración perinuclear que estaba ausente en células infectadas con EHEC $\Delta nleA$, o células no infectadas. Este patrón no recordaba a la coloración obtenida con marcadores para endosomas tardíos, lisosomas, RE, mitocondrias o núcleo. Sin embargo, se observó un patrón de coloración muy similar cuando las células fueron co-coloreadas con anti-NleA y anticuerpos a los marcadores del aparato de Golgi, incluyendo a la manosidasa II, donde las dos proteínas se colocalizaban extensivamente. Para confirmar la localización de la NleA en Golgi, las células infectadas se incubaron con brefeldina A, un metabolito fúngico que destruye la estructura del Golgi (27), antes de la fijación y la inmunofluorescencia. El tratamiento con brefeldina A provocó la difusión de la coloración de la manosidasa II y la NleA, como se esperaba para proteínas localizadas en Golgi. La colocalización de NleA se

observó con varios otros marcadores del aparato de Golgi, y también se observó la localización de Golgi en experimentos para examinar NleA marcada con epítipo y coloreada con anticuerpos anti-marca, utilizando tanto marcas epitópicas para HA como para FLAG. Para determinar si la localización de NleA en Golgi requería otros factores bacterianos o si era una propiedad inherente de NleA, las células fueron transfectadas con una construcción de expresión que codificaba una proteína de fusión GFP-NleA. La NleA GFP transfectada también se localizó en el Golgi, donde se superponía con la coloración de la manosidasa II.

Por lo tanto, nuestros resultados indican que NleA se localiza en el Golgi. La observación de que una proteína de fusión NleA-GFP transfectada se localiza en el Golgi sugiere que la proteína NleA contiene información que la direcciona a Golgi, y que no requiere otros factores bacterianos para llegar a este destino. La NleA liberada por bacterias también se localiza en Golgi.

EJEMPLO 9

NleA es necesario para la virulencia

El alto nivel de conservación de secuencia de NleA en agentes patógenos A/E (Figura 11C) sugiere que NleA cumple un papel similar en la infección. *C. Rodentium* es un agente patógeno natural de ratones (53), y ha sido usado como un sistema modelo para estudiar la patogénesis A/E. En cepas susceptibles, la infección por *C. Rodentium* es fatal y típicamente provoca la muerte de los ratones infectados entre los días 6-10 de infección (54). Las cepas de ratón más resistentes no mueren a causa de la infección por *C. Rodentium*, pero son colonizadas y desarrollan inflamación intestinal e hiperplasia colónica (54).

Para probar el rol de NleA en la virulencia, hemos creado una cepa de *C. Rodentium* con delección en *nleA* y verificado la ausencia de NleA mediante análisis Western blot del total de los extractos bacterianos con el antisuero NleA (Figura 16A). Los ratones fueron infectados con igual cantidad de bacterias de tipo salvaje ó $\Delta nleA$ por sonda oral. En ratones C3H-HeJ susceptibles a *C. Rodentium*, NleA fue absolutamente necesario para la virulencia. Todos los ratones C3H-HeJ infectados con

C. Rodentium de tipo salvaje murieron entre los días 6-10 de la infección (n=9), mientras que todos los ratones infectados con $\Delta nleA$ (n=13) demostraron algunos síntomas de enfermedad leve tal como heces blandas, pero igualmente subieron de peso y estuvieron activos durante la infección y sobrevivieron indefinidamente (Figura 16B). Además, los ratones infectados con $\Delta nleA$ fueron resistentes a sometimientos posteriores a *C. Rodentium* salvaje (n=5, figura 16B). Por lo tanto, mientras la cepa $\Delta nleA$ no es patogénica en los ratones susceptibles, interactúa en forma suficiente con el hospedador para estimular la inmunidad protectora.

En oposición a los ratones C3H/HeJ, la infección por *C. Rodentium* no es letal para los ratones NIH Swiss de cruce. En estos ratones, la colonización del intestino grueso por *C. rodentium* produce inflamación intestinal, hiperplasia colónica y leves síntomas de diarrea. Los ratones NIH Swiss fueron infectados con *C. Rodentium* salvaje o la cepa $\Delta nleA$ y sacrificados el día 10 después de la infección. Los ratones infectados con la cepa $\Delta nleA$ tuvieron, en promedio, una titulación de *C. Rodentium* 20 veces menor en el colon en el día 10 (figura 16C).

El análisis histológico del colon de los ratones NIH infectados se realizó infectando ratones con *C. Rodentium* salvaje o la cepa $\Delta nleA$ y sacrificándolos el día 10 después de la infección. Los últimos 0,5 cm del colon de los ratones infectados fue fijado en formalina con buffer neutral al 10%, procesado, cortado en secciones de 3 μ m y teñido con hematoxilina y eosina. Las secciones de tejido fueron observadas y fotografiadas usando objetivos de 5X y 63X. Los resultados indicaron que, en los análisis histológicos de las biopsias tomadas del borde anal de los ratones infectados, se encontraron numerosas bacterias en el tejido infectado con el tipo salvaje, pero las bacterias eran pocas en las muestras infectadas con $\Delta nleA$. Todos los animales infectados con *C. rodentium* de tipo salvaje demostraron signos patológicos de hiperplasia colónica, mientras que todos los ratones infectados con $\Delta nleA$ no tenían signos de hiperplasia. Las muestras infectadas con el tipo salvaje mostraron inflamación severa e hiperplasia hasta el punto en que el lumen intestinal ya no era evidente y la capa de músculo exterior estaba visiblemente distendida para acomodar el gran volumen de epitelio. Opuestamente, las muestras infectadas con $\Delta nleA$ mostraron histología relativamente normal. El grado relativo de inflamación e hiperplasia intestinal

también fue evidente en la diferencia en el peso del colon al momento del sacrificio en los dos grupos de ratones (Figura 16D). Los ratones infectados con el tipo salvaje también tenían bazo más grandes que los ratones infectados con *ΔnleA* tal como se refleja en los pesos esplénicos (Figura 16D).

Por lo tanto, hemos demostrado un efecto llamativo de NleA sobre la virulencia en un modelo de enfermedad en ratones. En el ratón susceptible, la presencia de NleA funcional en *C. Rodentium* produce una infección letal en 10 días. Los ratones infectados con una cepa sin NleA tuvieron pocos síntomas y sobrevivieron a la infección. En una cepa de ratones más resistentes, en los que la infección por *C. Rodentium* no es letal, NleA es necesario para el desarrollo de la hiperplasia colónica, y a los 10 días de infección, se hallan menos bacterias mutantes de NleA presentes en el intestino hospedador. Estos estudios indican un efecto claro de NleA en la virulencia de *C. Rodentium*. Nuestros resultados de la infección por EHEC de células HeLa demuestran que *in vitro*, NleA no afectará la adherencia de las bacterias a las células hospedadoras o translocación de otros efectores, lo que sugiere que NleA puede actuar al nivel de resistir la evacuación por el hospedador en lugar de mejorar la adherencia bacteriana. Además, la resistencia de los ratones infectados con *ΔnleA* a posteriores infecciones por *C. Rodentium* de tipo salvaje provee evidencia de que una cepa mutante de *nle* coloniza e interactúa con el hospedador en forma suficiente para estimular la inmunidad del hospedador. Esto está en oposición a mutantes tipo III de *C. Rodentium* que no colonizan el hospedador y no proveen protección para futuras infecciones. Por lo tanto, las cepas mutantes de NleA pueden ser usadas como una cepa de vacuna atenuada.

REFERENCIAS

Las siguientes publicaciones son incorporadas por referencia.

1. Nataro, J. P. y Kaper J. B. (1998), Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* 11: 142-201.
2. Frankel G., Phillips A. D., Rosenshine I., Dougan G., Kaper J. B., y Knutton S. (1998), Enteropathogenic and enterohaemorrhagic *Escherichia coli*: more subversive elements. *Mol Microbiol* 30: 911-921.

3. Perna N. T., Plunkett G., 3^{ra}. Burland V., Mau B., Glasner J. D., Rose D. J., Mayhew, G. F., Evans P. S., Gregor J., Kirkpatrick H. A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grotbeck E. J., Davis N. W., Lim A., Dimalanta E. T., Potamouisis K. D., Apodaca J., Anantharaman T. S., Lin J., Yen G., Schwartz D.C., Welch R. A. y Blattner F. R. (2001), Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 409: 529-533.
4. Elliott S.J., Wainwright L. A., McDaniel T. K., Jarvis K. G., Deng Y. K., Lai L. C., McNamara B. P., Donnenberg M. S. y Kaper J. B. (1998). The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Mol Microbiol* 28:1-4.
5. Perna N. T., Mayhew G. F., Posfai G, Elliot S., Donnenberg M. S., Kaper J. B., y Blattner F. R. (1998). Molecular evolution of a pathogenicity island from enterohaemorrhagic *Escherichia coli* O157:H7. *Infect. Immun* 66: 3810-3817.-
6. Zhu C., Agin T. S., Elliot S., Johnson L. A., Thate T. E., Kaper J. B., y Boedeker E. C. (2001). Complete nucleotide sequence and analysis of the locus of enterocyte Effacement from rabbit diarrheagenic *Escherichia coli* RDEC-1. *Infect Immun* 69: 2107-2115.
7. Deng W., Li Y., Vallance B. A., y Finlay B. B. (2001). Locus enterocyte effacement from *Citrobacter rodentium*: sequence analysis and evidence for horizontal transfer among attaching and effacing pathogens. *Infect Immun* 69: 6323-6335.
8. Tauscheck M., Strugnell R.A., y Robins-Browne R. M. (2002). Characterization and evidence of mobilization of the LEE pathogenicity island of rabbit-specific strains of enteropathogenic *Escherichia coli*. *Mol Microbiol* 44: 1533-1550.
9. Hacker J. y Kaper J. B. (2000). Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* 54: 641-679.
10. G. R. Cornelis, *J. Cell Biol.* 158, 401 (2002).
11. S. Gruenheid, B. B. Finlay, *Nature* 422, 775 (2003).
12. I. Chernushevich, A. Loboda, B. Thomson, *J. Mass Spectrom*, 36, 849 (2001).
13. Hsiao W., Wan I., Jones S. J., y Brinkman F. S. (2003). Island Path: aiding detection of genomic islands in prokaryotes. *Bioinformatics* 19: 418-420.

14. Edwards R. A., Keller L. H. y Schifferli D. M. (1998). Improved allelic exchange vectors and their use to analyze 987P fimbria gene expression. *Gene* 207:149-157.
15. V. Sperandio, J. L. Mellies, W. Nguyen, S. Shin, J. B. Kaper, *Proc. Natl. Acad. Sci. USA* 96, 15196 (1999).
16. J. L. Mellies, S. J. Elliot, V. Sperandio, M. S. Donnenberg, J. B. Kaper, *Mol. Microbiol* 33, 296 (1999).
17. K. A. Datsenko, B. L. Wanner, *Proc. Natl. Acad. Sci. USA* 97, 6640 (2000).
18. A. Abe et. al., *Mol. Microbiol*, 33, 1162 (1999).
19. S. J. Elliot et. al., *Mol. Microbiol* 33, 1176 (1999).
20. Vallance B. A., Deng W., DeGrado M., Chan C., Jacobson K. y Finlay B. B. (2002). Modulation of inducible nitric oxide synthase expression by the attaching and effacing bacterial pathogen *Citrobacter rodentium* in infected mice. *Infect Immun* 70:6424-6435.
21. V. H. Bustamante, F. J. Santana, E. Calva, J. L. Puente, *Mol. Microbiol.*, 39, 664 (2001).
22. J. L. Puente, D. Bieber, S. W. Ramer, W. Murray, G. K. Murray, G. K. Schoolnik, *Mol. Microbiol*, 20, 87 (1996).
23. T. Houyhaeve, H. Gausepohl, M. Mann, K. Ashman, *FEBS Lett.*, 376, 91 (1995).
24. Z. Ziegler, *Anal. Chem.*, 74, 489A (2002).
25. Brunder, W., Schmidt H., and Karch H. (1997), EspP, A novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol. Microbiol* 24: 767-778.
26. Buchet A., Nasser W, Eichler K., y Mandrand-Berthelot M. A. (1999). Positive co-regulation of the *Escherichia coli* carnitine pathway *cai* and *fix* operons by CRP and the CaiF activator. *Mol. Microbiol.* 34: 562-575.
27. Chardin P., y McCormick F. (1999) Brefeldin A: the advantage of being uncompetitive. *Cell* 97: 153-155.
28. DeVinney R., Stein M., Reinscheid D., Abe A., Ruschkowski S. Y Finlay B. B. (1999). Enterohemorrhagic *Escherichia coli* O157:H7 produces phosphorylated. *Infect Immun* 67: 2389-2398.

29. **Donnenberg M. S., y Kaper J. B. (1991). Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infect Immun* 59: 4310-4317.**
30. **Gauthier A., DeGrado M. y Finlay B. B. (2000). Mechanical fractionation reveals structural requirements for enteropathogenic *Escherichia coli* Tir insertion into host membranes. *Infect Immun* 68: 4344-4348.**
31. **Goosney D. L., DeVinney R. y Finlay B. B. (2001). Recruitment of cytoskeletal and signalling proteins to enteropathogenic and enterohemorrhagic *Escherichia coli* pedestals. *Infect Immun* 69: 3315-3322.**
32. **Griffin P. M., Ostroff S. M., Tauxe R. V., Greene K. D., Wells J. G., Lewis J. H. y Blake P. A. (1998). Illnesses associated with *Escherichia coli* O157:H7 infections. A broad clinical spectrum. *Ann Intern Med* 100: 705-712.**
33. **C. Buchrieser et. al , *Mol. Microbiol.* 38, 760 (2000).**
34. **Galan J. E. (2001). Salmonella interactions with host cells: type III secretion at work. *Annu Rev Cell Dev Biol.* 17: 53-88.**
35. **Kenny B., DeVinney R., Stein M., Reinscheid D. J., Frey E. A. y Finlay B. B. (1997). Enteropathogenic *E. Coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell* 91: 511-520.**
36. **Elliott S. J., Krejany E. O., Meliies J. L., Robins-Browne R. M., Sasakawa C. y Kaper J. B. (2001). EspG, a novel type III system-secreted protein from enteropathogenic *Escherichia coli* with similarities to VirA of *Shigella flexneri*. *Infect Immun.* 69: 4027-4033.**
37. **Kenny B. y Jepson M. (2000). Targeting of an enteropathogenic *Echerichia coli* (EPEC) effector protein to host mitochondria. *Cell Microbiol.* 2: 579-590.—**
38. **McNamara B. P., Koutsouris A., O'Connell C. B., Nougayrede J. P., Sonnenberg M. S. y Hecht G. (2001). Translocated EspF protein from enteropathogenic *Escherichia coli* disrupts host intestinal barrier function. *J. Clin Invest* 107: 621-629.**
39. **O. Marches et. al., *Infect Immun.* 68, 2171 (2000).**
40. **Tu X., Nisan I., Yona C., Hanski E., y Rosenshine I. (2003). EspH, a new cytoskeleton-modulating effector of enterohaemorrhagic and enteropathogenic *Escherichia coli*. *Mol. Microbiol.* 47:595-606.**
41. **D. S. Guttman et. al., *Science* 295, 1722 (2002).**
42. **T. Petnicki-Ocwieja et. al., *Proc Natl. Acad. Sci. USA* 99, 7652 (2002).**

43. S. I. Miller, A. M. Kukral, J. J. Mekalanos, *Proc Natl. Acad. Sci. USA* **86**, 5054 (1989).
44. Gruenheid S., DeVinney R., Bladt F., Goosney D., Gelkop S., Gish G. D., Pawson T. y Finlay B. B. (2001). Enteropathogenic *E. Coli* Tir binds Nck to initiate actin pedestal formation in host cells. *Nat Cell Biol.* **3**: 856-859.
45. Harlow E., y Lane D. (1988) *Antibodies: a laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
46. Hueck C. J. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol. Rev.* **62**: 379-433.
47. Jarvis K. G. y Kaper J. B. (1996). Secretion of extracellular proteins by enterohemorrhagic *Escherichia coli* via a putative type III secretion system. *Infect Immun* **64**: 4826-4829.
48. Kenny B., Ellis S., Leard A. D., Warawa J., Mellor H. y Jepson M. A. (2002). Co-ordinate regulation of distinct host cell signalling pathways by multifunctional enteropathogenic *Escherichia coli* effector molecules. *Mol. Microbiol* **44**: 1095-1107.
49. Knodler L. A., Vallance B. A., Hensel M., Jackel D, Finlay B. B., y Steele-Mortimer, O. (2003). Salmonella type III effectors PiB and PipB2 are targeted to detergent-resistant microdomains on internal host cell membranes. *Mol. Microbiol* **49**: 685-704.
50. Levine M. M., Bergquist E. J., Nalin D. R., Watman D. H., Hornick R. B., Young C. R., y Sotman S. (1978). *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *Lancet* **1**: 1119-1122.
51. Li Y., Frey E., Mackenzie A. M., y Finlay B. B. (2000). Human response to *Escherichia coli* O157:H7 infection: antibodies to secreted virulence factors. *Infect Immun* **68**: 5090-5095.
52. Peeters J. E., Geeroms R. y Orskov F. (1998). Biotype, serotype and pathogenicity of attaching and effacing enteropathogenic *Escherichia coli* strains isolated from diarrheic commercial rabbits. *Infect Immun* **56**: 1442-1448.
53. Schauer D. B. y Falkow S. (1993). The *eae* gene of *Citrobacter freundii* biotype 4280 is necessary for colonization in transmissible murine colonic hyperplasia. *Infect Immun* **61**: 4654-4661.

54. Vallance B.A., Deng W., Jacobson K. y Finlay B. B. (2003). Host susceptibility to the attaching and effacing bacterial pathogen *Citrobacter rodentium*. *Infect Immun* 71: 3443-3453.
55. Yoshida S., Katayama E., Kuwae A., Mimuro H., Suzuki T. y Sasakawa C. (2002). Shigella deliver an effector protein to trigger host microtubule destabilization, which promotes Rac1 activity and efficient bacterial internalization. *Embo J.* 21: 2923-2935.
56. Deng W., Vallance B. A., Li Y., Puente J. L. y Finlay B. B. (2003). *Citrobacter rodentium* translocated intimin receptor (Tir) is an essential virulence factor needed for actin condensation, intestinal colonization and colonic hyperplasia in mice. *Mol. Microbiol.* 48:95-115.
57. "Remington's pharmaceutical Sciences" (19th edition), ed. A. Gennaro, 1995, Mack Publishing Company, Easton, Pa.
58. Abe A., U. Heczko, R. Hegele y B. Finlay, 1998. Two enteropathogenic *Escherichia coli* type III secreted proteins, EspA and EspB, are virulence factors. *Journal of Experimental Medicine* 188: 1907-1916.
59. Adv. DrugDeliv. Rev. 5 (3): 163-187 (1990).
60. Altschul S. F., 1991, "Amino acid substitution matrices from an information theoretic perspective", *Journal of Molecular Biology*, 219: 555-665.
61. Ausubel et. al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY 1998.
62. Babiuk et.al.(1986) *Virology* 159:57-66.
63. Beuzon C. R. y D. W. Olden 2001. Use of mixed infections with *Salmonella* strains to study virulence genes and tier interactions *in vivo*. *Microbes Infect* 3: 1345-52.
64. Brown D. et. al. (2002) *TechNotes* 9:3-5.
65. Brummerkamp TR, et. al.(2002) *Science* 296:550-553.
66. Canil C., I. Rosenshine, S. Ruschkowski, M. S. Donnenberg, J. B. Kaper y B. B. Finlay. 1993. Enteropathogenic *Escherichia coli* decreases the transepithelial electrical resistance of polarized epithelial monolayers. *Infect Immun* 61:2755-62.
67. Caplen NJ, et. al. (2001) *Proc. Natl. Acad. Sci. USA* 98:9746-9747.
68. Clin. Exp. Immunol, 78 (2):258-262 (1989).

69. Cloning Vectors: A laboratory Manual (P. H. Pouwels et. al., 1985, Supp. 1987.
70. Dayhoff M. O., Schwartz R. M., Orcutt B. C. 1978. "A model of evolutionary change in proteins". In "Atlas of Protein Sequence and Structure" 5 (3) M. O. Dayhoff (ed.), 345-352, National Biomedical Research Foundation, Washington.
71. Eisenberg et. al. *J. Mol. Bio.* 179:125-142, 184.
72. Elder et. al. *Proc Natl. Acad. Sci. USA* (2000) 97:2999.
73. Epitope Mapping Protocols in Methods in Molecular Biology, Vol. 66 (Glenn E. Morris, Ed. 1996) Humana Press, Totowa, NJ.
74. Fey Et. al., *Emerg. Infect. Dis* (2000) Volume 6.
75. Gali D. (1986) *Immunology* 11:369-386.
76. Geysen et. al. (1984) *Proc. Natl. Acad. Sci. USA* 81: 3998-4002.
77. Geysen et. al. (1986) *Molec. Immunol.* 23:709-715.
78. Hammerling et. al., In *Monoclonal Antibodies and T Cell Hybridomas*, Elsevier NY, 1981.
79. Hammond SM et. al. (2001) *Nature Rev. Gen* 2: 110-119.
80. Hauf N. y T. Chakraborty. 2003. Supression of NF-kappa B activation and proinflammatory cytokine expression by Shiga toxin-producing *Escherichia coli*. *J. Immunol* 170:2074-82.
81. Hopp et. al., *Proc. Natl. Acad. Sci. USA* (1981) 78:3824-3828.
82. *Immunology* 58: (2): 245-250 (1986).
83. *Int. Arch. Allergy Appl. Immunol.* 68 (3): 201-208 (1982).
84. *J. Immunol Methods* 97 (2): 159-164 (1987).
85. *J. Controlled Release* 7:123-132 (1988).
86. Karlin S. y Altschul S. F. 1990. "Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes". *Proc. Natl. Acad. Sci USA* 87:2264-2268.
87. Knodler L. A., J. Celli, W. D. Hardt, B- A. Vallance, C. Yip y B. B. Finlay. 2002. Salmonella effectors within a single pathogenicity island are differentially expressed and translocated by separate type III secretion systems. *Mol Microbiol.* 43:1089-103.
88. Kodak Laboratory Chemicals Bulletin 56 (1): 1-5 (1986).
89. Kohler et. al., *Eur. J. Immunol*, 6:292-1976.

90. Kohler et. al., Eur. J. Immunol, 6:511, 1976.
91. Kohler et. al., Nature 256:495, 1975.
92. Kyte et. al., J. Mol. Biol. (1982) 157:105-132.
93. Lee NS et. al. (2002) Nature Biotechnol 20:500-505.
94. M. S. Johnson y J. P. Overington 1993. "A Structural Basis of Sequence Comparisons: An evaluation of scoring methodologies". Journal of Molecular Biology, 233:719-738.
95. Miyagishi M. y Talra K. (2002). Nature Biotechnol 20:497-500.
96. Myers and Miller, CABIOS, 1989, 4:11-17.
97. Naylor S W., J. C. Low, T. E. Besser, A. Mahajan, G. J. Gunn, M. C. Pearce, I. J. McKendrick, D. G. Smith y D. L. Gally 2003. Lymphoid follicle-dense mucosa at the terminal rectum is the principal site for colonization of enterohemorrhagic *Escherichia coli* O157:H7 in the bovine host. Infect Immun 71: 1505-12.
98. Paddison P.J. Et. al. (2002). Genes & Dev. 16:948-958.
99. Paul CP, et. al. (2002) Nature Biotechnol 20:505-508.
100. Scijs et. al., Curr. Opl. Immunol (2000) 12:456.
101. Sharp P. A. (2001) Genes Dev 15:485-490
102. States D. J., Gish W., Altschul S. F. 1991. "Improved Sensitivity of Nucleic Acid Database Search Using Application-Specific coring Matrices". Methods: A companion to methods in enzymology 3(1):66-77
103. Steven Henikoff y Jorja G. Henikoff 1992. "Amino acid substitution matrices from protein blocks". Proc. Natl. Acad. Sci. USA 89 (biochemistry):10915-10919.
104. Steven Henikoff y Jorja G. Henikoff 1993. "Performance Evaluation of Amino Acid Substitution Matrices". Proteins: structure, function and Genetics 17:49-61.
105. Sui G. at. al. (2002) Proc. Natl. Acad. Sci. USA 99: 5515-5520.
106. Vallance B.A., W. Deng, L. A. Knodler y B. B. Finlay 2002. Mice lacking T and B lymphocytes develop transient colitis and crypt hyperplasia yet suffer impaired bacterial clearance during *Citrobacter rodentium* infection. Infect Immun 70:2070-81.
107. Van Donkersgoed et.al., Can. Vet. J. (1990) 40:332.
108. Van Donkersgoed et. al., Can. Vet. J. (2001) 42:714.
109. Yu J-Y et. al (2002) Proc. Natl. Acad. Sci. USA 99:6047-6052.

110. Sambrook et. al, Molecular Cloning: A laboratory Manual, 2nd. Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 1989.
111. Ausubel et. al., Current Protocols in Molecular Biology, John Wiley & Sons 1994.
112. Vaughan et. al., Nature Biotech 14:309-314, 1996.
113. Larsson A. et. al. Chicken antibodies: taking advantage of evolution-a review. Poult. Sci 1993 Oct; 72 (10):1807-12.
114. Alymova IV, Kodihalli S., Govorkova E. A., Fanget B., Gerdil C., Webster RG. Immunogenicity and protective efficacy in mice of influenza B virus vaccines grown in mammalian cells or embryonated chicken eggs. J. Virol. 1998, May 72(5):4472-7
115. O'Farrelly C., Branton D., Wanka CA. Oral ingestion of egg yolk immunoglobulin from hens immunized with an enterotoxigenic *Escherichia coli* strain prevents diarrhea in rabbits challenged with the same strain. Infect Immun. 1992 Jul 60(7): 2593-7.
116. Romito M., Vlijoen GJ, Du Plessis DH. Eliciting antigen-specific egg-yolk IgY with naked DNA. Biotechniques 2001 Sep. 31 (3) 670, 672, 674-5.
117. Yokoyama H., Umeda K., Peralta RC, Hashi T., Icatlo FC, Kuroki M., Ikemori Y, Kodama Y. Oral passive immunization against experimental salmonellosis in mice using chicken egg yolk antibodies specific for *Salmonella enteritidis* and *S. typhimurium*. Vaccine, 1998 Feb. 16(4):388-93.

OTRAS REALIZACIONES

Si bien en la presente se revelan otras realizaciones de la invención, se pueden realizar muchas adaptaciones y modificaciones dentro del alcance de la invención de acuerdo con el conocimiento general común de los expertos en el arte. Tales modificaciones incluyen la sustitución de los equivalentes conocidos para cualquier aspecto de la invención a fin de lograr el mismo resultado esencialmente de la misma manera. Los números de acceso, tal como se usan aquí, se refieren a los números de acceso de las bases de datos múltiples, incluso GenBank, el European Molecular Biology Laboratory (Laboratorio de biología molecular europeo) (EMBL), la DNA Database of Japan (Base de datos de ADN de Japón) (DDBJ), o la Genome Sequence

Database (Base de datos de secuencias de genomas) (GSDB), para secuencias nucleotídicas, e incluso Protein Information Resource (Fuente de información de proteínas) (PIR), Swissprot, Protein Research Foundation (Fundación de Investigaciones sobre proteínas) (PRF) y Protein Data Bank (PDB) (secuencias de estructuras resueltas), así como las traducciones de las regiones codificadas anotadas de las secuencias nucleotídicas en GenBank, EMBL, DDBJ o RefSeq, para secuencias de polipéptidos. Los rangos numéricos incluyen los números que definen el rango. En la memoria descriptiva, la palabra "que comprende" se usa como un término abierto, sustancialmente equivalente a la frase "que incluye pero no se limita a" y la palabra "que comprende" tiene un significado correspondiente. Las citas de referencias en la presente no deberán entenderse como un reconocimiento de tales referencias como arte previo a la presente invención. Todas las publicaciones son incorporadas a la presente por referencia como si cada publicación individual fuera indicada en forma específica e individual como incorporada a la presente por referencia. La invención incluye todas las realizaciones y variaciones sustancialmente como se describe aquí y con referencia a los ejemplos y figuras.

REIVINDICACIONES

Tras haber descripto e ilustrado la naturaleza y objeto principal de la presente invención, así como también la manera en que la misma se puede llevar a la práctica, se declara reivindicar como de propiedad y de derechos exclusivos:

1. Una composición caracterizada porque comprende un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a la secuencia de las SEQ ID Nos. 22-43, 59 ó 73-84; o un fragmento o variante de ellos, en combinación con un portador fisiológicamente aceptable.

2. Una composición caracterizada porque comprende una molécula de ácido nucleico codificadora de un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 22-43, 59 ó 73-84, o un fragmento o variante de ellos, en combinación con un portador fisiológicamente aceptable.

3. Una composición caracterizada porque comprende una molécula de ácido nucleico que comprende una secuencia nucleotídica sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 1-21 ó 60-72 o un fragmento o variante de ellos, en combinación con un portador fisiológicamente aceptable.

4. Una composición caracterizada porque comprende un sobrenadante de cultivo celular que comprende un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 22-43, 59 ó 73-84, o un fragmento o variante de ellos, en combinación con un portador fisiológicamente aceptable.

5. La composición tal y como se reivindica en la Reivindicación 4, caracterizada porque el polipéptido comprende 20% de la proteína celular presente en la composición.

6. La composición tal y como se reivindica en cualquiera de las reivindicaciones 1 a 5 caracterizada además porque comprende EspA, EspB, EspD, EspP, Tir, toxina Shiga 1, toxina Shiga 2 o polipéptido de intimina.

7. La composición de cualquiera de las reivindicaciones 1 a 6 caracterizada además porque comprende un adyuvante.

8. Una bacteria o una preparación de la misma, caracterizada porque la bacteria comprende una mutación en el genoma bacteriano en una secuencia nucleotídica que es sustancialmente idéntica a las SEQ. ID Nos. 1-21 ó 60-72.

9. Una bacteria o una preparación de la misma, caracterizada porque la bacteria comprende una mutación en un gen sustancialmente idéntico a *nleA*, *nleB*, *nleC*, *nleD*, *nleE*, *nleF*, *nleG*, ó *nleH* o un homólogo de los mismos.

10. La bacteria de las reivindicaciones 8 ó 9 caracterizada porque la bacteria es un agente patógeno A/E.

11. La bacteria de cualquiera de las reivindicaciones 8 a 10 caracterizada porque la bacteria es *E. Coli* enterohemorrágica (EHEC), *E. Coli* enteropatógena (EPEC) ó *Citrobacter rodentium*.

12. La bacteria tal y como se reivindica en la reivindicación 11, caracterizada porque la EHEC es EHEC O157:H7 ó EHEC O157:NM.

13. La bacteria tal y como se reivindica en la reivindicación 11, caracterizada porque la EPEC es EPEC O157:H6.

14. La bacteria de cualquiera de las reivindicaciones 8 a 13 caracterizada porque la mutación atenúa la virulencia.

15. Una composición caracterizada porque comprende la bacteria de cualquiera de las reivindicaciones 8 a 14, en combinación con un portador fisiológicamente aceptable.

16. La composición tal y como se reivindica en la reivindicación 15 caracterizada porque la bacteria está viva.

17. La composición tal y como se reivindica en la reivindicación 16 caracterizada porque la bacteria es administrada por vía oral.

18. La composición tal y como se reivindica en la reivindicación 15, caracterizada porque la bacteria está muerta.

19. La composición tal y como se reivindica en la reivindicación 18 caracterizada porque la bacteria es administrada por vía parenteral.

20. La composición tal y como se reivindica en cualquiera de las reivindicaciones 15 a 19 caracterizada además porque comprende un adyuvante.

21. La composición tal y como se reivindica en las reivindicaciones 7 ó 20 caracterizada porque el adyuvante es una emulsión de aceite en agua o una emulsión de no-aceite en agua.

22. La composición tal y como se reivindica en las reivindicaciones 7 ó 20 caracterizada porque el adyuvante comprende un aceite mineral y bromuro de dimetildioctadecilamonio.

23. La composición tal y como se reivindica en las reivindicaciones 7 ó 20 caracterizada porque el adyuvante comprende uno o más agentes seleccionados del grupo que consiste en: un agente emulsionante, un dipéptido de muramilo, un agente acuoso, un agente a base de quitosano, un saponino, un aceite, un lipopolisacárido, un extracto de pared de célula bacteriana, un ADN bacteriano, un complejo bacteriano, un oligonucleótido sintético y una base nitrogenosa alifática.

24. La composición de la reivindicación 23 caracterizada porque el agente emulsionante es seleccionado del grupo que consiste en uno o más de un agente emulsionante natural, un agente emulsionante sintético, un agente emulsionante aniónico, un agente emulsionante catiónico y un agente no iónico.

25. La composición de la reivindicación 24, caracterizada porque el agente emulsionante natural es seleccionado del grupo que consiste en uno o más de acacia, gelatina, lecitina y colesterol.

26. La composición de la reivindicación 24, caracterizada porque el agente emulsionante aniónico es seleccionado de uno o más del grupo que consiste en una sal de potasio de ácido láurico, una sal de potasio de ácido oleico, una sal de sodio de ácido láurico, una sal de sodio de ácido oleico, una sal de amonio de ácido láurico, una sal de amonio de ácido oleico, una sal de calcio de un ácido graso, una sal de magnesio de un ácido graso, una sal de aluminio de un ácido graso, un jabón metálico y un sulfonato orgánico.

27. La composición de la reivindicación 26, caracterizada porque el sulfonato orgánico es sulfato laurilo de sodio.

28. La composición de la reivindicación 24 caracterizada porque el agente emulsionante catiónico es bromuro de cetiltrimetilamonio.

29. La composición de la reivindicación 24 caracterizada porque el agente emulsionante sintético es seleccionado del grupo que consiste en uno o más de éster de glicerilo, éster de polioxietilenglicol, éter de polioxietilenglicol y éster de ácido graso de sorbitán.

30. La composición de la reivindicación 28 caracterizada porque el éster de glicerilo es monosteato de glicerilo.

31. La composición de la reivindicación 28 caracterizada porque el éster de ácido graso de sorbitán es seleccionado de uno o más del grupo que consiste en monopalmitato de sorbitán y un derivado de polioxietileno del mismo.

32. La composición de la reivindicación 30 caracterizada porque el derivado de polioxietileno es monopalmitato de sorbitán polioxietilenado.

33. La composición de la reivindicación 23 caracterizada porque el agente acuoso es hidróxido de aluminio.

34. La composición de la reivindicación 23 caracterizada porque el aceite es seleccionado del grupo que consiste en uno o más de un aceite mineral, un aceite vegetal o un aceite animal.

35. La composición de la reivindicación 33 caracterizada porque el aceite vegetal es seleccionado del grupo que consiste en uno o más de aceite de canola, aceite de almendra, aceite de semilla de algodón, aceite de maíz, aceite de oliva, aceite de maní, aceite de cártamo, aceite de sésamo y aceite de soja.

36. La composición de la reivindicación 33 caracterizada porque el aceite animal es seleccionado del grupo que consiste en uno o más de aceite de hígado de bacalao, aceite de hipogloso, aceite de lacha, aceite de "reloj anaranjado" y aceite de hígado de tiburón.

37. La composición de las reivindicaciones 7 ó 20 caracterizada porque el adyuvante comprende un componente de aceite.

38. La composición de la reivindicación 36 caracterizada porque el componente de aceite es seleccionado del grupo que consiste en uno o más de un aceite único y una mezcla de aceites.

39. La composición de la reivindicación 21, caracterizada porque la emulsión sin aceite en agua es seleccionada del grupo que consiste en uno o más de una emulsión de aceite, una emulsión de agua en aceite y una emulsión de agua en aceite en agua.

40. La composición de la reivindicación 21 caracterizada porque la emulsión de aceite en agua es EMULSIGEN [™] ó EMULSIGEN PLUS[™].

41. La composición de la reivindicación 24 caracterizada porque el aceite es Amphigen®.

42. La composición de la reivindicación 7 ó 20 caracterizada porque el adyuvante comprende un extracto de pared celular micobacteriana.

43. La composición de la reivindicación 7 ó 20 caracterizada porque el adyuvante comprende un ADN micobacteriano.

44. La composición de la reivindicación 7 ó 20 caracterizada porque el adyuvante comprende un complejo de pared de célula micobacteriana.

45. La composición de la reivindicación 23 caracterizada porque la base nitrogenada alifática se selecciona del grupo que consiste en uno o más de una amina, un compuesto de amonio cuaternario, una guanidina, una benzamidina y un tiouronio.

46. La composición de la reivindicación 23 caracterizada porque la base nitrogenada alifática es N,N-dioctadecil-N,N-bis(2-hidroxetil)propanodiamina.

47. La composición de la reivindicación 7 ó 20 caracterizada porque el adyuvante comprende bromuro de dimetil-dioctadecilamonio.

48. La composición de cualquiera de las reivindicaciones 21 a 46 caracterizada porque el adyuvante está presente en una composición a una concentración de entre 20% y 40% (v/v).

49. Un método para detectar la presencia de un agente patógeno A/E en una muestra, caracterizado porque el método comprende:

- a) proporcionar una muestra; y
- b) detectar la presencia de una molécula de ácido nucleico que comprende una secuencia nucleotídica sustancialmente idéntica a una secuencia seleccionada de uno o más del grupo que consiste en las SEQ. ID Nos. 1-21 ó 60-72 o un fragmento o una variante de las mismas.
- c) detectar la presencia de una molécula de ácido nucleico que codifica para un polipéptido sustancialmente idéntico a una secuencia seleccionada de uno o más del grupo que consiste en las SEQ.ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas;
- d) detectar la presencia de un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a una secuencia seleccionada de uno o más del grupo que consiste en las SEQ. ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas;

en donde la presencia de una molécula de ácido nucleico o el polipéptido indica la presencia de un agente patógeno A/E en la muestra.

50. El método de la reivindicación 48 caracterizado porque la muestra es materia fecal o sangre.

51. El método de la reivindicación 48 ó 49 caracterizado porque el paso de detectar comprende poner en contacto la secuencia nucleotídica con una sonda o iniciador sustancialmente idénticos a:

- a) una secuencia nucleotídica seleccionada del grupo que consiste en uno o más de las SEQ: ID Nos. 1-21 ó 60-72 o un fragmento o una variante de las mismas, o
- b) una secuencia nucleotídica codificadora de un polipéptido sustancialmente idéntico a una o más de las SEQ.ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas.

52. El método de las reivindicaciones 48 ó 49, caracterizado porque la detección comprende poner en contacto la secuencia de aminoácido con un anticuerpo que une

específicamente una secuencia seleccionada del grupo que consiste en una o más de las SEQ ID Nos: 22-43, 59, 73-84 o un fragmento o una variante de las mismas.

53. Un método para producir una respuesta inmune contra un agente patógeno A/E, o un componente del mismo, en un animal, caracterizado porque el método comprende administrar al animal una cantidad efectiva de una composición tal y como se reivindica en cualquiera de las reivindicaciones 1-7 ó 15-47, o que comprende administrar al animal una cantidad efectiva de la bacteria de cualquiera de las reivindicaciones 8-14, produciendo así una respuesta inmune en el animal.

54. Un método para reducir la colonización de un agente patógeno A/E en un animal, caracterizado porque el método comprende administrar al animal una cantidad efectiva de la composición de cualquiera de las reivindicaciones 1-7 ó 15-47 o que comprende administrar al animal una cantidad efectiva de la bacteria de cualquiera de las reivindicaciones 8-14, reduciendo así la colonización del agente patógeno A/E en el animal.

55. Un método para reducir la propagación de un agente patógeno A/E en un animal caracterizado porque comprende administrar al animal una cantidad efectiva de la composición de cualquiera de las reivindicaciones 1-7 ó 15-47, o que comprend administrar al animal una cantidad efectiva de la bacteria de cualquiera de las reivindicaciones 8-14, reduciendo así la propagación del agente patógeno A/E en el animal.

56. El método de cualquiera de las reivindicaciones 52 a 54 caracterizado porque el animal es un rumiante.

57. El método de la reivindicación 55, caracterizado porque el rumiante es un bovino o un ovino.

58. El método de cualquiera de las reivindicaciones 52 a 54 caracterizado porque el animal es un ser humano.

59. Un método para tratar o prevenir la infección causada por un agente patógeno A/E, caracterizado porque dicho método comprende:

a) identificar un animal que presente, o que esté en riesgo de contraer, una infección por agente patógeno A/E; y b) administrar al animal una cantidad efectiva de un compuesto que atenúa la virulencia de un agente patógeno A/E, en donde el compuesto inhibe la expresión, secreción o actividad biológica de un polipéptido que comprende una secuencia de aminoácido sustancialmente idéntica a la secuencia de cualquiera de las SEQ. ID. Nos: 22-43, 59, 73-84.

60. El método de la reivindicación 58, caracterizado porque el compuesto es una molécula de ácido nucleico antisentido que es complementaria de una secuencia nucleotídica sustancialmente idéntica a la secuencia de cualquiera de las SEQ. ID Nos. 1-21, 59, 73-84.

61. El método de la reivindicación 58, caracterizado porque el compuesto es un ARNsi.

62. Un método para atenuar la virulencia de un agente patógeno A/E, caracterizado porque el método comprende mutar uno o más de un gen seleccionado del grupo que consiste en *nleA*, *nleB*, *nleC*, *nleD*, *nleE*, *nleF*, *nleG* y *nleH*, o un homólogo de los mismos en el agente patógeno A/E, o mutar una o más de una secuencia nucleotídica en el genoma del agente patógeno A/E, en donde la secuencia nucleotídica es seleccionada de las SEQ. ID. Nos. 1-21 ó 60-72, atenuando así la virulencia.

63. Un método para seleccionar un compuesto que atenúa la virulencia del agente patógeno A/E, caracterizado porque el método comprende:

a) proporcionar un sistema que comprende:

(i) una molécula de ácido nucleico que comprende una secuencia nucleotídica sustancialmente idéntica a las SEQ. ID Nos 1-21 ó 60-72 o un fragmento o una variante de las mismas; ó

(ii) una molécula de ácido nucleico codificadora de un polipéptido que comprende una secuencia de aminoácido sustancialmente idéntica a las SEQ. ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas; ó

(iii) un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a las SEQ. ID. Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas;

b) proporcionar un compuesto de prueba; y

c) determinar si el compuesto de prueba modula la expresión, secreción o actividad biológica del polipéptido o la molécula de ácido nucleico, en donde un cambio en la expresión, secreción o actividad biológica del polipéptido o la molécula de ácido nucleico indica un compuesto que atenúa la virulencia de un agente patógeno A/E.

64. El método de la reivindicación 62, caracterizado porque el sistema es una célula.

65. El método de la reivindicación 62 ó 63, caracterizado porque la célula es *E. Coli* enterohemorrágica (EHEC), *E. Coli* enteropatogénica (EPEC) o *Citrobacter rodentium*.

66. El método de la reivindicación 62, caracterizado porque el sistema es un sistema *in vitro*.

67. Un método para producir un polipéptido de agente patógeno A/E caracterizado porque comprende:

a) proporcionar una célula recombinante que comprende:

(i) una molécula de ácido nucleico que comprende una secuencia nucleotídica sustancialmente idéntica a las SEQ. ID. Nos. 1-21 ó 60-72; ó

(ii) una molécula de ácido nucleico codificadora de un polipéptido que comprende una secuencia de aminoácido sustancialmente idéntica a las SEQ. ID. Nos. 22-43, 59 ó 73-84; y

b) cultivar la célula recombinante en condiciones que permitan la expresión del polipéptido.

68. El método de la reivindicación 68 caracterizado además porque comprende cultivar la célula recombinante en condiciones que permitan la secreción del polipéptido.

69. El método de la reivindicación 66 ó 67 caracterizado porque el polipéptido se secreta desde la célula.

70. El método de cualquiera de las reivindicaciones 66 a 68 caracterizado además porque comprende aislar el polipéptido.

71. El método de cualquiera de las reivindicaciones 48 a 69 caracterizado porque el agente patógeno A/E es *E. Coli* enterohemorrágica (EHEC), *E. Coli* enteropatogénica (EPEC) ó *Citrobacter rodentium*.

72. El método de la reivindicación 70 caracterizado porque la EHEC es EHEC O157:H7 ó EHEC O157:NM.

73. El método de la reivindicación 70 caracterizado porque la EPEC es EPEC O127:H6.

74. Un polipéptido recombinante caracterizado porque comprende una secuencia de aminoácido sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 22-43, 59 ó 73-84.

75. Una molécula de ácido nucleico aislada caracterizada porque comprende una secuencia nucleotídica sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 1-21 ó 60-72.

76. Un vector caracterizado porque comprende la secuencia nucleotídica de la reivindicación 74.

77. El vector de la reivindicación 75 caracterizado porque el vector es capaz de integrarse al genoma del agente patógeno A/E.

78. El vector de la reivindicación 75 caracterizado porque el vector es capaz de integrarse al genoma del agente patógeno A/E.

79. Una célula hospedadora caracterizada porque comprende el vector de cualquiera de las reivindicaciones 75 a 77.

80. La célula hospedadora de la reivindicación 78, caracterizada porque la célula hospedadora es un agente patógeno A/E.

81. La célula hospedadora de la reivindicación 79 caracterizada porque el agente patógeno A/E es *E. Coli* enterohemorrágica (EHEC), *E. Coli* enteropatógena (EPEC) o *Citrobacter rodentium*.

82. El uso de la composición de cualquiera de las reivindicaciones 1-7 ó 15-47, la bacteria de cualquiera de las reivindicaciones 8-14, el polipéptido de la reivindicación 73 o la molécula de ácido nucleico de la reivindicación 74, para la preparación de un medicamento para producir una respuesta inmune contra un agente patógeno A/E, o un componente del mismo, o para reducir la propagación o la colonización de un agente patógeno A/E en un animal, o para tratar o prevenir la infección causada por un agente patógeno A/E.

83. Un *kit* caracterizado porque comprende un reactivo para detectar un agente patógeno A/E en una muestra y un prospecto con instrucciones para detectar el agente patógeno A/E en la muestra.

84. El *kit* de la reivindicación 83, caracterizado porque el reactivo comprende una sonda o una sonda iniciadora o un iniciador sustancialmente idéntico a:

a) una secuencia nucleotídica seleccionada del grupo que consiste en una o más de las SEQ. ID No. 1-21 ó 60-72 o un fragmento o una variante de las mismas;

b) una secuencia nucleotídica que codifica para un polipéptido sustancialmente idéntico a una o más de las SEQ. ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas.

85. El *kit* de la reivindicación 83, caracterizado porque el reactivo comprende un anticuerpo que une específicamente una secuencia seleccionada del grupo que consiste en una o más de las SEQ. ID No. 22-43, 59, 73-84 o un fragmento o una variante de las mismas.

RESUMEN

La invención se relaciona, en parte, a proteínas de secreción de agentes patógenos bacterianos y a métodos para su utilización. Más específicamente, la invención provee en parte varias proteínas de secreción comunes para agentes patógenos de adherencia y destrucción (agentes patógenos "A/E"). En algunas realizaciones de la invención, estos polipéptidos y las moléculas de ácido nucleico que codifican para estos polipéptidos, o porciones de las mismas, son útiles como vacunas, como diagnósticos o como herramientas de selección de fármacos para infecciones patogénicas A/E, o como reactivos.

Citrobacter rodentium NleA DNA Sequence

ATGAACATTCAACCGAACATACATTCCGGAATCACCACACAGAATAATCAACAACATCATCAACGAGAA
 CAAGTGCCTGTCTCTAGCTCAATACCGGATCAGATTACCTCCAAATGCGAAGCTGGATTGTGTG
 CATATTCCAGAGGATATACAGCAACATGTACCGGAATGTGGTGAACAACGGCTCTATTAAGCTTGATA
 AAAGATGAAGGCCTGCTCTCAGGACTAGATAAATATCTTGCTCCTCACCTTGAAGAAGGCTCCCTTGGG
 AAAAAAGCATTGGATACGTTTGGTTTATTCAATGTTACTCAAATGGCATTAGAGATACCCAGTTCCGTT
 CCAGGCATATCTGGTAAATATGGTGTTCAGATGAACATTGTAAAACCAGACATACATCCAACAACCGGG
 AACTATTTTTTACAGCTATTTCTCTGTCATGACGAAATAGGTTTTAACTTCAAAGATCTTCTGGCCCA
 TTAATAATGCATTAAACCAACAGCAGTATATCGGCTACTGCATCGACTGTAGCCCCCACACCAACGAC
 CCAATGCCATGGTTTGGATTAACTGCTCAAGTGGTTCGTAATCATGGTGTAGAACTTCTTATAGTCAA
 ACCGAAATGGATGGAAGCTTGTAGGGGAAACACCTCTTACTCCAGATGGCCCCAAAGCCAATTATACG
 GAAGATGGGTATCAGGCCGGGAGAAGCAGATTTTAAATATGGAACATCGCCATTACAGGCAACTCTT
 GGACTGGAGTTTGGTGCGCATTTTAAGTGGGATTAGATAATCCTAATACCAATATGCCATCCTTACC
 AATGCTGCCGCAAATGCTATTGGTGTCTGCTGGAGGTTTGGCGGTATCCAAAGTCCCCGGCATAGATCCA
 ATGCTGTCCCTCATGCTCGGTGCAATGCTTGGGCAAGCAGCGGGGCATGCCGTACAAATGTAATACCCCC
 GGATTAAAGCCAGACACTATTTTATGGTGGGCAGGCGCGACATTTGGAGCTGCTGATTAAATAAAGCC
 GAATTTGATAAAGTCCGTTCACTGACTACCTCTGTATATGGTTCCATGCACGGGAAGGAGCTTTATTC
 CCAATAAGCAAGACATTGCCCGTGTAAACAGGCGCAGACATAAAAGCTATGGAAGAAGGCGTACCCGTT
 GGACATCAACATCCAAACCGGAGGATGTGGTCATCGATATCGAAGGTGGCAATTCACCACATCATAAT
 CCATCAAATTATGTTGACACCTTTGAAATAATCCAAGAAACAAGGGTCTAA (SEQ ID NO:1)

FIG. 1ACitrobacter rodentium NleA Protein Sequence

MNIQPNIHSGITTONNQHHHAHQVPVSSSI PRSDLPFNCEAGFVWHIPEDIQQHVPECGETTALLSLI
 KDEGLLSGLDKYLAPHLEEGSLGKKALDTFGLFNVTQMALEIPSSVPGISGKYGVQMNIVKPD IHPTTG
 NYFLQLFPLHDEIGFNFKDLPGPLKNALTNSSISATASTVAPTNPDPMPWFGLTAQVVRNHGVELPIVK
 TENGWKL VGETPLTPDGPKANYTEEWVIRPGEADFKYGTSPLOATLGLFGAHFKNDLDNPNTKYAILT
 NAAANAIGAAGGFAVSKVPGIDFMLS PHVGAMLGQAAGHAVQCNT PGLKPDITLWAGATPGAADLNKA
 EFDKVRFTDYPRIFWFAREGALFPNKQDIARVTGADIKAMEEGVPVGHQHPKPEDVVIDIEGGNSPHHN
 PSNYVDTFEIIQETRV (SEQ ID NO:22)

FIG. 1B

Enteropathogenic *E. coli* NleA DNA Sequence

atgaacattcaaccgatcgtaacatccggaatcaccacacaaaacaatcgacatcatcacgcagaacaa
acgtcccctacacaaataccgcaatccgaattacctaattggatgcgaaacgggattttgttgcataatc
ccagaggatatgcagcgacatgcaacggaatgcggtgaaacaacagctctactgagcttgetaaaagat
gaaggctctgctctctgggctggataaatatcttgcacctcatcttgaagaaggctctgcaggaaaaaaa
gcattggatatgtttggtttattcaatgtctctcagatggcattagaaatacccagcacggttccgggt
atctctggtaaatatgggtgtccagctaaacattgtaaaaccagatattcatcctacatcaggttaattat
tttttacagatattccctttgcatgatgaaataggtattaatttttaagaccttcctgggtccattaaaa
aatgcattaagcaacagcaatataccaaccactgtatcgactgctgcatccactattgcatcagccact
acttcgacggtaaccacccgctcaaaagacccaataacatgggtttggattaacagctcaagtagttcgt
aatcatgggtgtggaacttcctatagtcaaaaactgaaaatggatggaagcttgttggagaaactcctctt
actcctgatggcccaagcaatatactgaagagtgggtgatcagaccgggagaagcagatttttaa
tatgggtgcatctccactacaggcaactctagggctggagtttggcgacatttcaagtgggatttagat
aacccctaataactaatatgccgttcttaccatgtctgccgcaaatgcgcttgggtgctgtagggggattt
gcagtatccagatttactgggtacagatccaatgttaagtcctcatatcgggtgcaatgggttgggcaagca
gcggggcatgccatacagtataataccccgggattaaagccagacactattttatgggtgggcaggtact
actcttggactggctgatttaacaagccgagtttggagaggccagattcactgactatcctcgtata
tgggtggcatgcaagagaaggtgccattttcccaataaaagcagatattgaacatgccacaggggtgat
atacgcgcaatggaagaaggtgtatctgttggacaacggcatccaaatccagaggatgtggtcatcaat
atcgaagcaataactcaccacatcataacccatcaaattatgttgataccggtgatataatccaagaa
acaagagtctaa (SEQ ID NO:2)

FIG. 1C

Enteropathogenic *E. coli* NleA Protein Sequence

MNIQPIVTSGITTONNRHHAEQTSPTQIPQSELPNGCETGFVVHIPEDMQRHAPECGETTALLSLIKD
EGLLSGLDKYLAPHLEEGSAGKKALDMFGLFNVSQMALEIPSTVPGISGKYGVQLNIVKPDINHPTSGNY
FLQIFPLHDEIGINFKDLPGPLKNALSNSNIPTTVSTAASIASATTSTVTTASKDPIFWFGLTAQVVR
NHGVELPIVKTENGWKLVGETPLTPDGPKANYTEEWVIRPGEADFKYGASPLQATLGLLEFGAHFKWDL
NPNTKYAVLTNAAANALGAVGGFAVSRFTGTDPMLSPHIGAMVGQAAGHAIQYNTPGLKPDITLWAGT
TLGLADLNKAEFGARFTDYPRIWWHAREGAIFPNKADIEHATGADIRAMEEGVSVGQRHPNPNEDVVIN
IESNNSPHNPSNYVDTVDIQETRV (SEQ ID NO:23)

FIG. 1D

Enterohemorrhagic E. coli NleA DNA Sequence

ATGAACATTCAACCGACCATACAATCTGGAATCACCTCACAAAACAATCAACATCATCAAACAGAACAA
 ATACCCCTCTACACAAATACCGCAATCCGAATTACCTCTAGGATGCCAAGCTGGATTTGTTGTTAATATT
 CCAGATGATATACAGCAACATGCACCGGAATGCGGTGAAACAACAGCTCTACTGAGCTTGATAAAAGAT
 AAAGGTCTGCTCTCAGGGCTAGACGAATATATAGCTCCTCACCTTGAAGAAGGATCCATAGGAAAAAAA
 ACATTGGATATGTTTGGTTTATTCAATGTTACCCAAATGGCATTAGAGATACCTAGTTCCGTTTCAGGC
 ATCTCTGGTAAATATGGTGTCCAGCTAAACATTGTAAAACCAGATATTCATCCTACATCAGGTAATTAT
 TTTTACAGATATTCCCTCTGCATGATGAAATAGGTTTTAATTTTAAAGACCTTCCTGGCCCGTTAAAA
 AATGCATTAAGCAACAGTAATATATCAACCACTGCAGTGTGACTATTGCATCGACTGGAACATCAGCC
 ACTACTTCGACGGTAACCACCGAGCCAAAAGACCCAATACCATGGTTTGGATTAACAGCTCAAGTGGTT
 CGTAATCATGGTGTAGAACTTCCTATAGTCAAACTGAAAATGGATGGAAGCTTGTTGGAGAAACACCA
 CTTACTCCTGATGGGCCGAAAGCAAATTACACGGAGGAGTGGGTTATCAGACCGGGAGAAGCAGATTTT
 AAATATGGTGCATCTCCATTACAGGCAACTCTAGGGCTGGAGTTTGGCGCACATTTCAAGTGGGATTTA
 GATAACCCCTAATACTAAATATGCCGTTCTTACCAATGTGCCGCAAATGCGCTTGGTGCTTTAGGGGGA
 TTTGCAGTATCCAGATTTGCTAGTACAGATCCAATGTTAAGTCCTCATATCGGTGCAATGGTTGGGCAA
 GCAGCAGGGCATGCCATACAGTATAATACCCCTGGATTAAAGCCAGACACTATTTTATGGTGGGCTGGT
 GCGACACTGGGGGCTGCCGATTTAAACAAGGCCGAGTTTGAAGTAGCTAGATTCACTGACTATCCTCGT
 ATATGGTGGCAGCAAGAGAAGGAGCTATTTTCCCAATAAAGCAGATATTGAACATGCCACAGGTGCT
 GATATACGCGCAATGGAAGAAGGTATCCCTGTTGGACAGCGGCATCCAAATCCAGAGGATGTGGTAATC
 GATATCGAAAGCAATGGCTTACCACATCATAATCCATCAAATCATGTTGATATCTTTGATATAATCCAA
 GAAACAAGAGTCTAA (SEQ ID NO:3)

FIG. 1EEnterohemorrhagic E. coli NleA Protein Sequence

mnigptiqsgitsqnnqhhqteqipstqipqselpigcqagfvvniipddiqqhapedcgetallslidk
 kgllsgldeyiaphleegsigkktldmfglfnvtqmaleipssvsgisgkygvqlnivkpdihptsgny
 flqifplhdeigfnfkdlpgplknalsnennisttavstiasgttsattstvttepkdpipwfgltaqv
 rnhgvelpivktengwklvgetpltpdgpkanyteewvirpgeadfkylgasplqatlglefghfkwdl
 dnpntkyavltnaaanalgalgffavsrfastrpmlsphigamvgqaaghaigqyntpglkpdtllwwag
 atlgadlnkaefevartdyprwharegaifpnkadiehatgadirameegipvgqrhpnpedvvi
 diessnglphhnpshvdiidliqetrv (SEQ ID NO:24)

FIG. 1F

Citrobacter rodentium NleB DNA Sequences

tattcttttttcagtgggtttgaagcAAGGCCAGAGCGATACGGAAAAGGTGAAGTACCGATATTGAATAC
 CAAAGAGCATCCGTATTTGAGCAATATTATAAATGCTGCAAAAATAGAAAATGAGCGcgTAATaGGAGT
 ACTGGTAGACGGAGACTTTACTTATGAGCAAAGAAAAGAATTTCTCAGTCTTGAAGATGAACATCAAAA
 TATAAAGATAATATATCGGGAAAATGTTGATTTCACTATGTATGATAAAAACTGTCTGATATTTATCT
 TGAATAATATTCATGAACAAGAATCATATCCAGCGAGTGAGAGAGATAATTATCTGTTaGGcTTaTTAAG
 AGAAGAGTTAAAAAATATTCCATACGGAAAGGACTCTTTGATTGAATCATATGCAGAAAAAAGAGGTCA
 TACTTGGTTTGATTTTTTTTAGAAACTTGGCGGTATTGAAGGGGGGGGGTGTGTTACAGAGACGGGTAA
 AACTGGATGCCATAACATATCTCCATGTGGGGGATGTATATATCTTGATGCAGATATGATTATTACTGA
 TAAATTAGGTGTCTGTATGCTCCTGATGGTATcgctgtgcatgtagattgtaatgatgaga (SEQ
 ID NO:4)

FIG. 2A

AAGGCCAGAGCGATACGGAAAAGGTGAAGTACCGATATTGAATACCAAAGAGCATCCGTATTTGAGCAA
 TATTATAAATGCTGCAAAAATAGAAAATGAGCGTAATGGAGTACTGGTAGACGGAGACTTTACTTATGA
 GCAAAGAAAAGAATTTCTCAGTCTTGAAGATGAACATCAAAAATATAAAGATAATATATCGGGAAAATGT
 TGATTTCACTATGTATGATAAAAACTGTCTGATATTTATCTTGAAAATATTCATGAACAAGAATCATA
 TCCAGCGAGTGAGAGAGATAATTATCTGTTGGTTTTAAGAGAAGAGTTAAAAAATATTCCATACGGAAA
 GGACTCTTTGATTGAATCATATGCAGAAAAAAGAGGTCATACTTGGTTTGATTTTTTTAGAACTTGGC
 GGTATTGAAGGGGGGGGGTGTGTTACAGAGACGGGTAAAACCTGGATGCCATAACATATCTCCATGTGG
 GGGATGTATATATCTTGATGCAGATATGATTATTACTGATAAATTAGGTGTCTGTATGCTCCTGATGG
 TAT (SEQ ID NO:5)

FIG. 2BCitrobacter rodentium NleB Protein Sequences

ILFQWFEPARYGKGEVPILNTKEHPYLSNIIINAAKIENERVIGVLVDGDFTYEQRKEFLSLEDEHQN
 IKIIYRENVDFSMYDKKLSDIYLENIEHESYPASERDNYLLGLLREELKNIPYGKDSLIESYAEKRGH
 TWFDFFRNLAVLKGGGLFTETGKTGCHNISPCGGCIYLDADMIITDKLGVLYAPDG IAVHVDCNDE
 (SEQ ID NO:25)

FIG. 2C

RPERYGKGEVPILNTKEHPYLSNIIINAAKIENERVIGVLVDGDFTYEQRKEFLSLEDEHQNIKIYREN
 VDFSMYDKKLSDIYLENIEHESYPASERDNYLLGLLREELKNIPYGKDSLIESYAEKRGH
 LAVLKGGGLFTETGKTGCHNISPCGGCIYLDADMIITDKLGVLYAPDG (SEQ ID NO:26)

FIG. 2D

Enteropathogenic *E. coli* NleB DNA Sequence

atgttatcttcattaaatgtccttcaatccagcttcagaggaaagacagctttatcaaatagtacactt
ctccagaaagtttcttttgctggaaaagaatattctctggaaacctattgatgaaagaacccctattctt
tttcagtggtttgaagcaaggccagagcgatacgaaaaaggagagtgaccaatattgaataccaaagaa
catccgtatttgagcaatattataaatgctgcataaataagaaaatgagcgtataatcgggtgtgctggtg
gatggaaattttactttatgaacaaaaaagggaatttctcaatcttgaaaatgaacatcaaaatataaaa
ataatctaccgagcagatgtggatttcagcatgtatgataaaaaactatctgatattttaccttgaaaat
atccataaacaagaatcataccctgccagtggagaggataattatctgttaggcttattaagagaagag
ttaaaaaatatccagaaggtaaggactctttgattgagtcatatgcagaaaaagagaaacatacttgg
tttgattttttcaggaattttggccatattgaaggctggaagtttgtttacagagacgggaaaaactgga
tgccataacatatcgccctgtagcggatgtatatatcttgatgccgacatgattattaccgataaatta
ggagtcctgtatgctcctgatggtatcgctgtgcatgtagattgtaatgatgagataaaaagtcttgaa
aatgggtgcgatagtttgtcaatcgtagtaatcatccagcattacttgaggcctcgatattatgaagagt
aaagttgacgctcatccatattatgatggtctaggaagggtatcaagcggcatttttaactattcatcg
ttacacaaattataatgctttttgtgattttattgaatttaagcatgaaaatattataccgaataccagt
atgtataccagcagttcatggtaa (SEQ ID NO:6)

FIG. 2E

Enteropathogenic *E. coli* NleB Protein Sequence

MLSSLNLVQSSFRGKTALSNSTLLQKVSFAGKEYSLEPIDERTPILFOWFEARPERYKGEVPILNTKE
HPYLSNIINAAKIENERIIGVLVDGNFTYEQKKEFLNLENEHQNIKIYRADVDFSMYDKKLSDIYLEN
IHQESYPASERDNYLLGLLREELKNIPEGKOSLIESYAEKREHTWFDFFRNLAAILKAGSLFTETGKTG
CHNISPCSGCIYLDADMIITDKLGVLYAPDGLAVHVDNCDEIKSLENGAIVVNRSNHPALLAGLDIMKS
KVDAPHYYDGLGKGIRHFNYSLLHNYNAFCDFIEFKHENIIPNTSMYTSSSW (SEQ ID NO:27)

FIG. 2F

Enterohemorrhagic *E. coli* NleB DNA Sequence

ATGTTATCTTCATTAAATGTCCTTCAATCCAGCTTCAGAGGAAAGACAGCTTTATCAAATAGTACACTT
CTCCAGAAAGTTTCTTTTGCTGGAAAAGAATATCCTCTGGAACCTATTGATGAAAAAACCCCTATTCTT
TTTCAGTGTTTGAAGCAAGGCCAGAGCGATACGAAAAAGGAGAAGTACCAATATTGAATACCAAAGAA
CATCCGTATTTGAGCAATATTATAAATGCTGCAAAAAATAGAAAATGAGCGTATAATCGGTGTGCTGGTA
GATGGAAATTTTACTTATGAACAAAAAAGGAATTTCTCAGTCTTGAAAATGAATATCAAAATATAAAA
ATAATCTACCGAGCAGATGTGGATTTCAGCATGTATGATAAAAAACTATCTGATATTTACCTTGAAAAT
ATCCATAACAAGAATCATACCCTGCCAGTGAGAGGGATAATTATCTGTTAGGCTTATTAGAGAAGAG
TTAAAAAATATCCAGAGGTAAGGACTCTTTGATTGAGTCATATGCAGAAAAAGAGAACATACTTGG
TTTGATTTTTTCAGGAATTTGGCCATGTTGAAGGCTGGAAGTTTGTTTACAGAGACGGGAAAAACTGGA
TGCCATAACATATCGCCCTGTAGCGGATGTATATATCTTGATGCCGACATGATTATTACCGATAAATTA
GGAGTCCTGTATGCTCCTGATGGTATCGCTGTGCATGTAGATTGTAATGATGAGATAAAAAAGTCTTGAA
AATGGTGCGATAGTTGTCAATCGTAGTAATCATCCAGCATTACTTGACGGCCTCGATATTATGAAGAGT
AAAGTTGACGCTCATCCATATTATGATGGTCTAGGAAAGGGTATCAAGCGGCATTTTAACCTATTCATCG
TTACACGATTATAATGCTTTTTTGATTTTATTGAATTTAAGCATGAAAATATTATACCGAATACCACT
ATGTATACCTGCAGTTCATGGTAA (SEQ ID NO:7)

FIG. 2G

Enterohemorrhagic *E. coli* NleB Protein Sequences

mlsslrvlqssfrgktalsnstllqkvsfaqkeypleidektpilfqwfearperyekegevpilntke
 hpylsniinaakieneriigvlvdgnftyeqkkeflaleneyqnikiiyradvdfsmysdkklsdiylen
 ihkqesypaserdnyllgllreelknipegkdsliesyaehtwdfdfnrlamkagslfttetgktg
 chnispcagciyldadmiitdklgvlyapdqiavhvdcndeikslengaiivnrshpallagldimks
 kvdahpyydglgkgikrhfnyslhdynafcdfiefkheniipntamytcsaw (SEQ ID NO:28)

FIG. 2H

MLSPIRTTFHNSVNIQSSPCQTVSFAGKEYELKVIDEKTPIILQWFEPNPERYKKEDEVPIVNTKQHPY
 LDWVTNAARIESDRMIGIFVDGDFSVNQKTAFSKLERDFENVMIIYREDVDFSMYDRKLSDIYHDIICE
 QRLRTEDKRDEYLLNLLLEKELREISKQDSLISMYAKGRNHAWDFDFRNLALLKAGEIFRCTYNTKNHG
 ISFGEGCIYLDMDMILTGKLGTIYAPDGISMHVDRRNDVNIENSAIIVNRSNHPALLEGLSFMHRSKVD
 AHPIYDGLGKGVKKYFNFTPLHNYNHFCDFIEFNHPNIIMNTSQYTCSSW (SEQ ID NO:29)

FIG. 2IEnterohemorrhagic *E. coli* NleB DNA Sequence

atgctttcac cgataaggac aactttccat aactcagtaa atatagtgc gagttcaccc
 tgtcaaacgg tttcttttgc aggaaggaa tatgagtaa aggtcattga tgaaaaaag
 cctattcttt ttcagtgggt tgaacctaat cctgaacgat ataagaaaga tgagggtcca
 atagttaata ctaagcagca tccctattta gataatgtca caaatgcggc aaggatagag
 agtgatcgta tgataggtat ttttggtgat ggcgattttt cagtcaacca aaagactgct
 ttttcaaaat tggaaacgaga ttttgaaaat gtaatgataa tctatcggga agatgttgac
 ttcagtatgt atgacagaaa actatcagat atttatcatg atatttatatg tgaacaaagg
 ttacgaactg aagacaaaag agatgaatac ttgttgaatc tgtagagaa agagctgagg
 gaaatttcaa aggcgcagga ttctttgatt tctatgtatg caaagaaaag aatcatgca
 tggtttgatt tcttcagaaa tttagcctta ttaaagcag gagagatatt caggtgcaca
 tataatacaa agaatacagg tatttcattc ggggaggggt gtatctatct tgataggtat
 atgatactta caggtaagct tggtaacata tatgctcctg atggaatttc aatgcatgtg
 gatcgtcgta atgatatgtt aaatattgaa aatagtgcaa taattgttaa ccgtagtaat
 catcctgctc tacttgaggg actttctttt atgcatagta aagtagatgc tcatccatat
 tatgatgggt tggggaaagg agttaagaaa tattttaatt ttacaccatt acataattat
 aatcattttt gtgactttat tgagtttaac caccctaata taatcatgaa cacaagtcag
 tatacatgca gttcatggt a (SEQ ID NO:60)

FIG. 2J

Citrobacter rodentium NleC DNA Sequence

ATGAAAATTCCCTCACTCCAGCCAGCTTCAACTTTTTTCGCCCCAGCAGGATACTCTGCTGCGGTTGCT
 CCCAATCGTTCGGACAATGCCTATGCTGATTACGTATTGGATATAGGCAAGCGAATACCACTTTCCGCG
 GAAGATTTAGGCAACCTATATGAAAATGTCATTGCGGCCGTTGCTGACAGCCGTAGCAAGCTCATAGAT
 CAGCATACGGTCGATATGATTGGTAACACTATACTTGATGCTTTGAGCCGATCACAAACCTTTGCTGAT
 GCCGTAAGCTATGGCATTCAATAAAGGAGGTACACATTGGTTGCATTAAATACAGAAACGAATACGAG
 CTC AACGGAGAATCCCCCGTCAAAGTTGATGATATTCATCACTAACCTGTACCGAATTATATGAATAC
 GATGTCGGGCAAGAACCAATTTTACCCATTTGCGAGGCAGGAGAAAACGATAACGAAGAGOCCTTATGTC
 AGTTTTAGTGTTGCGCCAGATACTGACTCTTATGAGATGCCATCGTGGCAGGAAGGGCTGATTCAOGAG
 ATTATTCATCATGTGACTGGAGCTAGCGATCCGTCTGGAGATAGTAATATAGAGCTAGGACCCACGGAG
 ATTCTCGCAGTCGTGTGCTCAAGAGCTGGGATGGACTGTCCCGACTTCATAGGATATGCAGAGCCA
 GATCGTGAAGCTCATCTTAGGGGACGTAACTGAATGCCCTTCGACAGGCGGCCATGCCGACATGAAGAT
 AATGAGAGGACTTTCTTCGAAAGGCTGGGTATGATCAGTGATCGATATGAGGCGAGTCTGATTTTACA
 GAGTATTCGCTGTGTCTAACATAGAATATGGATTTATCCAGCAACATGATTTTCCCGGGTTGGCTATC
 GACGATAATTTACAGGATGCAATCAGATCCAACTCTATCATGGAGCACCTTATATCTTTACATTCCGG
 GATGTGGACAAACACAATCAGCGCTGA (SEQ ID NO:8)

FIG. 3ACitrobacter rodentium NleC Protein Sequence

MKIFSLQPSFNFFAPAGYSAAVAPNRS DNAYADYVLDIGKRIPLSAEDLGNLYENVIRAVRDSRSKLID
 QHTVDMIGNTILDALSRSTFRDAVS YGIHNKEVHIGCIKYRNEYELNGESFVKVDDIQSLTCTELYEY
 DVGQEPILPICEAGENDNEEPYVSFSVAPDTSYEMPSWQEGLIHEIIHHVTGASDPSGDSNIELGPTE
 ILARRVAQELGNTVPDFIGYAE PDREAHLRGRNLNALRQAMRHEDNERTFFERLGMISDRYEASPDFT
 EYSAVSNIEYGFIIQHDFFPGLAIDNLDANQIQLYHGAPYIFTFGDVOKHNQR (SEQ ID NO:30)

FIG. 3B

Enteropathogenic *E. coli* NleC DNA Sequence

atgaaaattccctcattacagtcctaacttcaacttttccgccccggcaggatactctgctccatttgc
 cctaatacgtgctgaaaatgcctatgcggattacgttttggatataggttaagcgaataccactttccgca
 gcagatttaagcaacgtatacgaagtgtataacgcgccgtccatgacagccgtagcaggcttatcgat
 cagcatacagtcgatatgatcggcaacactgtacttgatgctttgagccgatcacagacatttcgtgat
 gccgtaagctatggcattcataatgagaaggtagacattgggttcattaaatacagaaacgaatacgag
 cttaacgaagaatcttctgtcaaaattgatgatattcaatcactaacctgtaacgaattatatgaatat
 gatgtcgggcaagagccaattttcccccatttgcgaagcaggagaaaacgataacgaagagccttatgtc
 agtttttagtggttgcgcagatactgactcttatgagatgccatcgtggcaggaaggactgattcacgag
 attattcatcatgttactggatctagcgatccatctggagatagtaatatagagttaggacccacgag
 attctcgcaagtcgtgtcgtcgaagaactgggatggagtgttcccgacttcaaaggatatgcagagcca
 gaacgtgaagctcatcttaggttaacgtaacctgaatgccottcgacaggctgccatgaggcatgaagag
 aatgagagggctttcttcgaaggctgggtacgatcagtgaccgatatgaggcgagtcctgatttcaca
 gagtattccgctgtgtctaacataggatacggatttatccagcaacatgattttcctggattggctatc
 aacgataatttacaggatgcaaatcagatccaactgtatcatggcgccccttatatttttacatttggg
 gatgtggacaaacacaatcagcgatga (SEQ ID NO:9)

FIG. 3CEnteropathogenic *E. coli* NleC Protein Sequence

MKIPSLQSNFNFSAAGYSAPIAPNRAENAYADYVLDIGKRIPLSAADLSNVYESVIRAVHDSRSLID
 QHTVDMIGNTVLDALSRSTFRDAVSYGIEHNEKVHIGCIKYRNEYELNEESSVKIDDIQSLTCNELY
 DVGQEPFIPICEAGENDNEEPYVSFSVAPDTSYEMPSWQGLIHEIINHVTGSSDPSGDSNIELGPTE
 ILARRVAQELGNSVPDFKGYAEPEREHLRLRLNLALRQAAMRHEENERAFFERLGTISDRYEASPDF
 EYSAVSNIGYGFYQQRDFPGLAINDNLQDANQIQLYHGAPYIFTFGUVDKHNQR (SEQ ID NO:31)

FIG. 3D

Citrobacter rodentium NleF DNA Sequences

atgttacccaacaagtgggttccttcAGCAAATCTTTACTCATGGATGTATATCTCAGGAAAAGAGAATCCT
TCGACTCCGGAATCAGTAAGTGAACCTTAATCATAATCATTTTTCTTTCTCCTGAATTACAGGAGAACTG
GATGTTATGTTCCGCATATATTCATGTGCCAGAAACAATGATGAGCGTGAGAATATTTACCCGGAGCTA
AGGGATTTTGTAAGTAGCCTAATGGATAAGAGAAACAATGTGTTTGAGGTGATAAATGAAGATACTGAT
GAGGTGACCGGAGCTCTGAGAGCGGGAATGACGATAGAGGACAGGGATAGTTATATCAGGGATCTTTTT
TTTCTGCATTCAATTGAAAGTAAAAATTGAGGAAAGCAGACAAGATAAAGAGGATTGGAAATGTAAAGTT
TATGATCTGCTATGTCCGCATCATTCTTCAGAGCTATATGGGGATCTACGGGCAATCAAATGCCTCGTT
GAAGGATGCAGTGATGATTTTAGTCCTTTTGATACTATTAAGGTGCCGGATCTTACTTAcaacaaagga
tcctttacaatgtggatga (SEQ ID NO:18)

FIG. 6A

AGCAAATCTTTACTCATGGATGTATATCTCAGGAAAAGAGAATCCTTCGACTCCGGAATCAGTAAGTGA
ACTTAATCATAATCATTTTTCTTTCTCCTGAATTACAGGAGAACTGGATGTTATGTTCCGCATATATTC
ATGTGCCAGAAACAATGATGAGCGTGAGAATATTTACCCGGAGCTAAGGGATTTTGTAAGTAGCCTAAT
GGATAAGAGAAACAATGTGTTTGAGGTGATAAATGAAGATACTGATGAGGTGACCGGAGCTCTGAGAGC
GGGAATGACGATAGAGGACAGGGATAGTTATATCAGGGATCTTTTTTTTCTGCATTCAATTGAAAGTAAA
AATTGAGGAAAGCAGACAAGATAAAGAGGATTGGAAATGTAAAGTTTATGATCTGCTATGTCCGCATCA
TTCTTCAGAGCTATATGGGGATCTACGGGCAATCAAATGCCTCGTTGAAGGATGCAGTGATGATTTTAG
TCCTTTTGATACTATTAAGGTGCCGGATCTTACTTA (SEQ ID NO:19)

FIG. 6BCitrobacter rodentium NleF Protein Sequences

MLPTSGSSANLYSWMYISGKENPSTPESVSELNHNHFLSPELQEKLDVMFAIYSCARNNDERENIYPEL
RDFVSSLMDKRNNVFEVINEDTDEVTGALRAGMTIEDRDSYIRDLFFLHSLKVXIEESRQDKEDWKCKV
YDLLCPHHSSELYGDLRAIKCLVEGCSDDFSPFDTIKVPDLTYNKGSLQC (SEQ ID NO:40)

FIG. 6C

ANLYSWMYISGKENPSTPESVSELNHNHFLSPELQEKLDVMFAIYSCARNNDERENIYPELRDFVSSLMD
DKRNNVFEVINEDTDEVTGALRAGMTIEDRDSYIRDLFFLHSLKVXIEESRQDKEDWKCKVYDLLCPH
SSELYGDLRAIKCLVEGCSDDFSPFDTIKVPDL (SEQ ID NO:41)

FIG. 6D

Enterohemorrhagic *E. coli* NleE DNA Sequence

ATGATTAACTCCTGTTACTAATACTCAGGGCGTGTCCTTATAAATACTAAATATGCTGAACATGTGGTG
 AAAAATATTTACCGGAAATTAAACATGATTACTTTAATGAATCACCAATATATATGATAAGAAGTAT
 ATATCCGGTATAACCGAGGAGTAGCTGAACTAAAACAGGAAGAATTTGTTAACGAGAAGCCAGACGG
 TTTTCTTATATGAAGACTATGTATTCTGTATGTCCAGAAGCGTTTGAACCTATTTCCAGAAATGAAGCC
 AGTACACCGGAAGGAAGCTGGCTAACAGTTATATCCGGAAAACGCCCAATGGGGCAGTTTTCTGTAGAT
 AGTTTATACAATCCTGATTACATGCATTATGTGAGCTTCCGGACATTTGTTGTAAGATCTTCCCTAAA
 GAAATAATGATTTTTTATACATAGTTGTTGTGTACAGAAATGACAGCCCTCTAGGAGAACACGGGCA
 AATAGATTTATAGAATTATATAATATAAAAAGAGATATCATGCAGGAATTAAATTATGAGTTACCAGAG
 TTAAGGCAGTAAATCTGAAATGATTATCGCACGTGAAATGGGAGAAATCTTTAGCTACATGCCTGGG
 GAAATAGACAGTTATATGAAATACATAAATAATAAACTTTCTAAAATTGAGTAG (SEQ ID NO:17)

FIG. 5GEnterohemorrhagic *E. coli* NleE Protein Sequence

minpvtntqgvspintkyahvvnkniypeikhdyfnespnlydkkyisgitrgvaelkqeefvnexarr
 fsymktmysvcpeafepisrneastpegswltvisgkrpmgqfsvdslynpdlhalcelpdicckifpk
 enndflyivvvyrndsplgeqranrfielynikrdimgelnyelpelkavksemiaremgeifsympg
 eidsymkyinnkleie (SEQ ID NO:39)

FIG. 5H

9/37

Enterohemorrhagic *E. coli* NleC DNA Sequence

ATGAAAATTCCCTCATTACAGTCCAACCTTCAACTTTTCCGCCCCGGCAGGATACTCTGCTCCCATTGCT
 CCTAATCGTGCTGAAAATGCCTATGCGGATTAGGTTTTGGATATAGGTAAGCGAATACCACCTTTCCGCA
 GCAGATTTAAGCAACGTATACGAAAGTGTAATACGCCCGTCCATGACAGCCGTAGCAGGCTTATCGAT
 CAGCATAACAGTCGATATGATCGGCAACACTGTACTTGATGCTTTGAGCCGATCACAGACATTTCTGTGAT
 GCCGTAAGCTATGGCATTGATAATGAGAAGGTACACATTGGTTGCATTAAATACAGAAACGAATACGAG
 CTTAACGAAGAATCTTCTGTCAAAATTGATGATATTCAATCACTAACCTGTAACGAATTATATGAATAT
 GATGTCGGGCAAGAGCCAATTTTCCCCATTTGCGAAGCAGGAGRAAACGATAACGAAGAGCCTTATGTC
 AGTTTTAGTGTTGCGCCAGATACTGACTCTTATGAGATGCCATCGTGGCAGGAAGGACTGATTCACGAG
 ATTATTCATCATGTTACTGGATCTAGCGATCCATCTGGAGATAGTAATATAGAGTTAGGACCCACCGAG
 ATTCTCGCACGTCGTGTCGCTCAAGAACTGGGATGGAGTGTTCCTGACTTCAAAGGATATGCAGAGCCA
 GAACGTGAAGCTCATCTTAGGCTACGTAACCTGAATGCCCTTCGACAGGCTGCCATGAGGCATGAAGAG
 AATGAGAGGGCTTTCTTCGAAAGGCTGGGTACGATCAGTGACCGATATGAGGCGAGTCTTGATTTCA
 GAGTATTCGCTGTGTCTAACATAGGATACGGATTTATCCAGCAACATGATTTTCCCTGGATTGGCTATC
 AACGATAATTTACAGGATGCAATCAGATCCAACCTGTATCATGGCGCCCTTATATTTTTACATTTGGG
 GATGTGGACAAACACAATCAGCAATGA (SEQ ID NO:10)

FIG. 3EEnterohemorrhagic *E. coli* NleC Protein Sequence

mkipslqsnfnfsapagysapiapnraenayadyvldigkriplsadlsvyesviravhdersrlid
 qhtvdmigntvldalsrsqtfrdavsygihnekvhigcikyrneyelneessvkiddiqsltcnelyy
 dvqqepifpiceagendneepyvafsvapdttsyempswqegliheihhvtgssdpqsdnielqpte
 ilarrvaqelgwsvpdfkgyaepereahrlrlnlalrqaaamrheenerafferlgtisdryeaspdft
 eysavsnigygfliqqhdfpglainsnlqdanqilqlyhgapyiftfgdvdkhnqq (SEQ ID NO:32)

FIG. 3F

Citrobacter rodentium NleD DNA Sequence

ATGCGCCCTACATCCCTTAACCTGACATTACCTTCGTTACCTCTACCCCTCATCTTCAAATTC AATTTCA
 GCCACAGACATTCAATCTCTTGTAAAAATGTGCGGGTGTGCGCTGGGTGAAAAACAACCAACRACTCTGT
 TTCCACGGGACTGACCTTAAAAATCTACCAGCATCTTGAAGCTGCCCTCGATAAGATCGAATCCACAGAC
 ACTGGACGTA CTCTTTTGAAC TGTATTGAATTAACATCCCGACTCAAATCAGAAAACTGGCAATACAT
 CTCGATTCTGCTGAGTTAGGGGTGATAGCACACTGCAATGCGGATGCTGAAAACTCCCGAGGAAGTGGC
 TOCGACTTTCACTGTAATCTGAATGCAGTTGAATATCCCTGCGGGCAAGGAATTAGCCTGGTAGACTTT
 CATGCATGCATTGTTTTCCATGAACTTCTCCAGTTTTCCACAATTTAAATGGAGAGCGCCTGAAAGTT
 GAGAGTTCTCAACCAGAATTACAAACACACTCCCCACTTTTACTCGAAGAAGCCAGGACTGTTGGGTTT
 GGTGCTTTTTCTGAAGAAGTTCTTTCAGAAAATAAATTTCTGTAAGAGATTGGGATGCCCGCAGAAC
 TTCTACCCGCACGATTCATCTCTCATTTCATGATGACAATACAGTGACTCAGAGATTCCAGCGGAAAAA
 CTGCATCCGTTACTTTAG (SEQ ID NO:11)

FIG. 4ACitrobacter rodentium NleD Protein Sequence

MRPTSLNLTLPSPFLPSSSNSISATDIQSLVFMMSGVRVWVKNQQLCFHGTDLKIYQHLEAALDKIESTD
 TGRILLNCIELTSRLKSEKLAHLDSAE LGVIAHCNADAENSRGTGSD FHCNLNAVEYPCGGGISLVDF
 HACIVFHELLHVFHNLNGERLKVSSQPELQTHSPLLLLEEARIVGLGAFSEEVLSENKFREEIGMPRT
 FYPHDSSLIHDDNTVTQRFORKNLHPLL (SEQ ID NO:33)

FIG. 4BEnteropathogenic E. coli NleD DNA Sequence

atgcgccctacgtccctcaacttgggtattacatcagtcacacgctcgagctcaatgtcagatacagat
 atcgagctctcttgtaaaagcatcgagcggttcaatggataaaaaataatccgcaacttcgtttcagggg
 actgatcataatatatatcagcagattgaagcagcactcgataagattggctctacagagacagggggt
 gtactcctgaatgctattgaatcaatatcccgacttaaatcagaaacagtggttaatacacctcaactct
 tccagactaggagttatggcacatagagatatagatgctgagaacccatcgggggactggttccgatttt
 cactgtaatctgaatgcagttgaatatccctgtggggaggggatttagcgtggtggactttcatgacgact
 attgtttttcatgagttgctccatgttttccacaatttaaatggggagcggttgaaagttgagagttcc
 cgaccagaatcacaaaaatactctccacttttactcgaagaagccaggactgttgggttgggggctttt
 tcagaggagggtgctttcagaaaaataaattccgcgaagagattgggatgcoccgtagaacctcctacccg
 cagcactcagctcttattcatgatgacaatacagtgagtcctgggattccaacaggttaagactgcatcca
 ttgcttttag (SEQ ID NO:12)

FIG. 4C

Enteropathogenic *E. coli* NleD Protein Sequence

MRPTSLNLVLHQSSSTSSMSDIEDISLVKASSVQWIKNNPQLRFQGTDHNIYQQIEAALDKIGSTETGR
VLLNAIESISRLKSETVVIHLNSSRLGVMARDIDAENHRGTGSDFHCNLNAVEYPCGEGISVVDPHAT
IVFHLLHVFHNLNGERLKVESRRPESQKYSPLLLLEEARTVGLGAFSEEVLSNKFREEIGMPRRTSYP
HDSALIHDDNTVSLGFQQVRLHPLL (SEQ ID NO:34)

FIG. 4DEnterohemorrhagic *E. coli* NleD DNA Sequence

ATGCGCCCTACGTCCCTCAACTTGGTATTACATCAGTCATCAAGGTCGAGCTCAATGTCAGATACAGAT
ATCGAGTCTCTTGTAAAAGCATCGAGCGTTCAATGGATAAAAAATAATCCGCAACTTCGTTTCCAGGGG
ACTGATCATAATATATATCAGCAGATTGAAGCAGCACTCGATAAGATTGGCTCTACAGAGACAGGGCGT
GTACTCCTGAATGCTATTGAATCAATATCCCGACTTAAATCAGAAACAGTGGTAATACACCTCAACTCT
TCCAGACTAGGAGTTATGGCACATAGAGATATAGATGCTGAGAAOCATCGGGGGACTGGTTCGATTTT
CACTGTAATCTGAATGCAGTTGAATATCCCTGTGGGGAGGGGATTAGCGTGGTGGACTTTCATGCGACT
ATTGTTTTTTCATGAGTTGCTCCATGTTTTCCACAATTTAAATGGGGAGCGTTTGAAAGTTGAGAGTTCC
CGAGCAGAATCACAAAATACTCTCCACTTTTACTCGAAGAAGCCAGGACTGTTGGGTTGGGGGCTTTT
TCAGAGGAGGTGCTTTCAGAAAATAAATTCCACGAAGAGATTGGGATGCCCGTAGAACCTCCTACOCG
CRCGACTCAGCTCTTATTCATGATGACAATACAGTGAGTCTGGGATTCCAACAGGTAAGACTGCATCCA
TTGCTTTAG (SEQ ID NO:13)

FIG. 4EEnterohemorrhagic *E. coli* NleD Protein Sequence

mrptslnlvlhqsrrssmsdtdieslvkassvqwiknnpqlrfqgtdhniyqqieaaldkigstetgr
vllnaiesisrlksetvvihlnessrlgvmahrdidaenhrgtgdsdfhcnlnaveypcgegisvvdphat
ivfhellhvfhnlngerlkvessrsesqkyspllleartvglgafseevlsnkfheeiqmprrtsyp
xdsalihddntvslgfqqvrlhpll (SEQ ID NO:35)

FIG. 4F

Citrobacter rodentium NleE DNA Sequences

tacttttaaatgaatcaccCAATGTATATGATAAGAAGTATATATCTGGCGTAACTAGAGGAGTAGCTGAA
CTAAAACAGGAAGGATTTTATTAACGAGAAAGCCAGGCGACTTGCTTATATGCAAGCAATGTATTCTGTA
TGTCCGGAAGAGTTTAAACCTATTTCCAGAAACGAAGCTAGTACACCGGAAGGCAGCTGGCTAACAGTT
ATATCCGGAACGCCCCAATGGGACAGTTTTCTGTAGATAGCTTATATCATCCTGACTTACATGCATTG
TGTGAGCTTCCGGATATTTGTTGCAAGATCTTCCCTAAAGAAAACAATGATTTTTTGTATATAGTGATT
GTGTACAGAAATGACAGCCCTCTGGGAGAACAACGAGCAAATCGATTTATAGAATTATATAATATAAAA
AGAGACATCATGCAGGAATTAAATTATGAATCTCCAGAGTTAAAGGCTGTGAAATCTGAAATGATTATT
gcacgtgaatgggagaaatctt (SEQ ID NO:14)

FIG. 5A

CAATGTATATGATAAGAAGTATATATCTGGCGTAACTAGAGGAGTAGCTGAACTAAAACAGGAAGGATT
TATTAACGAGAAAGCCAGGCGACTTGCTTATATGCAAGCAATGTATTCTGTATGTCCGGAAGAGTTTAA
ACCTATTTCCAGAAACGAAGCTAGTACACCGGAAGGCAGCTGGCTAACAGTTATATCCGGAACGCCC
AATGGGACAGTTTTCTGTAGATAGCTTATATCATCCTGACTTACATGCATTGTGTGAGCTTCCGGATAT
TTGTTGCAAGATCTTCCCTAAAGAAAACAATGATTTTTTGTATATAGTGATTGTGTACAGAAATGACAG
CCCTCTGGGAGAACAACGAGCAAATCGATTTATAGAATTATATAATATAAAAAGAGACATCATGCAGGA
ATTAAATTATGAATCTCCAGAGTTAAAGGCTGTGAAATCTGAAATGATTATT (SEQ ID NO:15)

FIG. 5B

Citrobacter rodentium NleE Protein Sequences

YFNESPNVYDKKYISGVTRGVaelKQEGFINEKARRLAYMQAMYSVCPEEFKPISRNEASTPEGSWLTV
ISGKRPMGQFSVDSLYHPDLHALCELDPICCKIFPKENNDFLYIVIVYRNDSPLGEQRANRFIELYNIK
RDIMQELNYESPELKAVKSEMI IAREMGEI (SEQ ID NO:36)

FIG. 5C

NVYDKKYISGVTRGVaelKQEGFINEKARRLAYMQAMYSVCPEEFKPISRNEASTPEGSWLTVISGKRPM
GQFSVDSLYHPDLHALCELDPICCKIFPKENNDFLYIVIVYRNDSPLGEQRANRFIELYNIKRDIMQE
LWYESPELKAVKSEMI (SEQ ID NO:37)

FIG. 5D

Enteropathogenic *E. coli* N1eE DNA Sequence

atgattaatcctgttactaatactcagggcgtgtcccctataaataactaaatatgctgaacatgtggtg
 aaaaatatttaccgaaaattaaacatgattactttaatgaatcaccgaatatatatgataagaagtat
 atatccggtataaccagaggagtagctgaactaaaacaggaagaatttgtaacgagaaagccagacgg
 ttttcttatatgaagactatgtattctgtatgtccagaagcgtttgaacctatttccagaatgaagcc
 agtacaccggaaggaagctggctaacagttatatccggaaaacgcccgaatggggcagttttctgtagat
 agttttatacaatcctgattttacatgcattatgtgagcttcoggacatttggttgtaagatcttccctaaa
 gaaaataatgattttttatacatagttgttgtgtacagaaatgacagccctctaggagaaacaacgggca
 aatagattttatagaattatataataaaaaagagatatcatgcaggaattaaattatgagttaccagag
 ttaaaggcagtaaaatctgaatgattatcgacgctgaatgggagaaatctttagctacatgcctggg
 gaaatagacagtttatatgaatacataaataataaacttttctaasattgagtag (SEQ ID NO:16)

FIG. 5E

Enteropathogenic *E. coli* N1eE Protein Sequence

MINPVTNTQGVSPINTKYAEHVVKNIYPKIKHDYFNESPNIDKKYISGITAGVAELKQEEFVNEKARR
 FSYMKTMYSCPEAFEPISRNEASTPEGSWLTVISGKRPMGQFSVDSLYNPDLHALCELPDICCKIFPK
 ENNDFLYIVVVYRNDSPLGEQRANRFIELYNIKRDIMQELNYELPELKAVKSEMIAREMGEIFSYPG
 EIDSYMKYINNKLKIE (SEQ ID NO:38)

FIG. 5F

Enteropathogenic *E. coli* NleF DNA Sequence

atgttaccaacaagtgggtcttcagcaaatctttattcatggatgtatgtatcaggaagaggtaaccct
 tcgactccggaatcagtaagtgaagcttaataataatcactttctttctcctgaattacaagataaactt
 gatgttatgggtctctatatattcatgtgccagaaataataatgagcttgaggaaatctttcaagagcta
 agtgcttttgtaagtgggctgatggataagagaaatagtgtatttgaggtgagaaatgaaatactgat
 gaggttgtcggagcgctgagggcggaatgacgatagaggatagggatagttatatcagggatctttt
 tttctgcattcattgaaagtaaaaattgaggaaagtagacaaggcaagaagattcgaaatgtaaaagt
 tataatctgctatgtccgcacactcttcagagctatatgggtgatctacgagcaatgaaatgcctcgtg
 gaaggatgcagtgatgattttaatccttttgatattattagggtagcagatcttacttacaacaaggga
 tctttacaatgtggatga (SEQ ID NO:20)

FIG. 6EEnteropathogenic *E. coli* NleF Protein Sequence

MLPTSGSSANLYSWMYVSGRGNPSTPESVSELNHNHFLSPELQDKLDVMVSIYSCARNNNELEEIFQEL
 SAFVSGLMDKRNSVFEVRNENTDEVV GALRAGMTIEDRDSYIRD LFFLHSLKVKIEESRQKEDSKCKV
 YNLLCPHSSSELYGDLRAMKCLVEGCSDDFNPFDIIRVPDLTYNKGSLQCG (SEQ ID NO:42)

FIG. 6FEnterohemorrhagic *E. coli* NleF DNA Sequence

ATGTTACCAACAAGTGGTTCCTTCAGCAAACTCTTTATTTCATGGATGTATGTATCAGGAAGAGGTAACCT
 TCGACTCCGGAATCAGTAAGTGAGCTTAATCATAATCACTTTCTTTCTCCTGAATTACAAGATAAACTT
 GATGTTATGGTCTCTATATATTTCATGTGCCAGAAATAATAATGAGCTTGAGGAAATTTTCAAGAGCTA
 AGTGCTTTTGTAAAGTGGGCTGATGGATAAGAGAAATAGTGTATTTGAGGTGAGAAATGAAATACTGAT
 GAGGTTGTCGGAGCGCTGAGGGCGGGAATGACGATAGAGGACAGGGATAGTTATATCAGGGATCTTTT
 TTTCTGCATTCAATTGAAAGTAAAAATTGAGGAAAGTAGACAAGGCAAGAAGATTGAAATGTAAAGTT
 TATAATCTGCTATGTCCGCATCACTCTTCAGAGCTATATGGTGATCTACGAGCAATGAAATGOCTOGTG
 GAAGGATGCAGTGATGATTTTAATCCTTTTGATATTATTAGGGTACCAGATCTTACTTACAACAAGGA
 TCTTTACAATGTGGATGA (SEQ ID NO:21)

FIG. 6GEnterohemorrhagic *E. coli* NleF Protein Sequence

mlptsgssaanlyswmyvsgrgnpstpesvselnhnhflspelqdkldvmvsiyscarnnneleeifqel
 safvsglmdkrnsvfevrnentdevvgalragmtiedrdsyirdlfflhlslkvkieesrqgkedskckv
 ynllcphssselygdlramkclvegcsddfnpfdiirvpdltynkgsllqcg (SEQ ID NO:43)

FIG. 6H

*

SGN	MCPDNTHAKKQYLTPGNDIHYPGQTNHDACFIPVSVRQYA
Caif	MCEG-----YV
GrlA	M-----ESKNSDYVIPDSVKNYN

	***++++ * +++*+ +++ *** ++ ++
SGN	GEPLYIIVAHWCLLOQNWVORNOIAEAFHITARRASYLIA
Caif	EKPLYLLIAEWMMAENRWVIAREISIHFDIEHSAVNTLT
GrlA	GEPLYILVSLWCKLQEKWISRNDIAEAFGINLRRASFITT

Helix-turn-helix motif

	**+ + + + + + + **
SGN	YLRSKTSRVVSICRHQTLPN-KARRYEIIY-VIRVLDSPTP
Caif	YILSEVTEISCEVKMIP-NKLEGRGCQCQRLVKVVDIDEQ
GrlA	YISRRKEKISFRVRYVSYGNLHYKRL EIF-IYNVNLEAAP

	+ + ** ** +
SGN	STRREKAGPP-----LVSKRRVGN GDRSM--ANELNWRLC
Caif	IYARLRNNSREKLVGVKTPRI PAVPLTELNREQKNQMM-
GrlA	TESHVSTGPK-----RKT LRVGN GIVG---QSSIWNEM-

	+++++ + +
SGN	SNRNAGKILKKKEDEDDGT (170 aa)
Caif	----LSKSMRR (131 aa)
GrlA	-----INRRKKE----S (135 aa)

FIG. 7

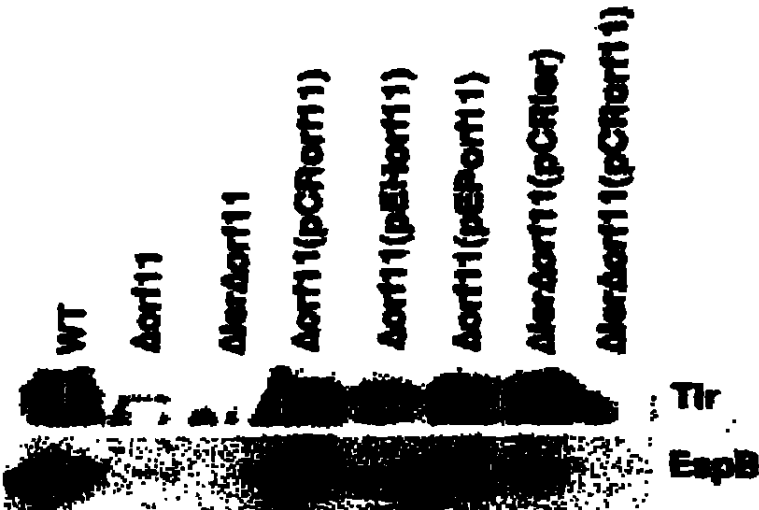


FIG. 8

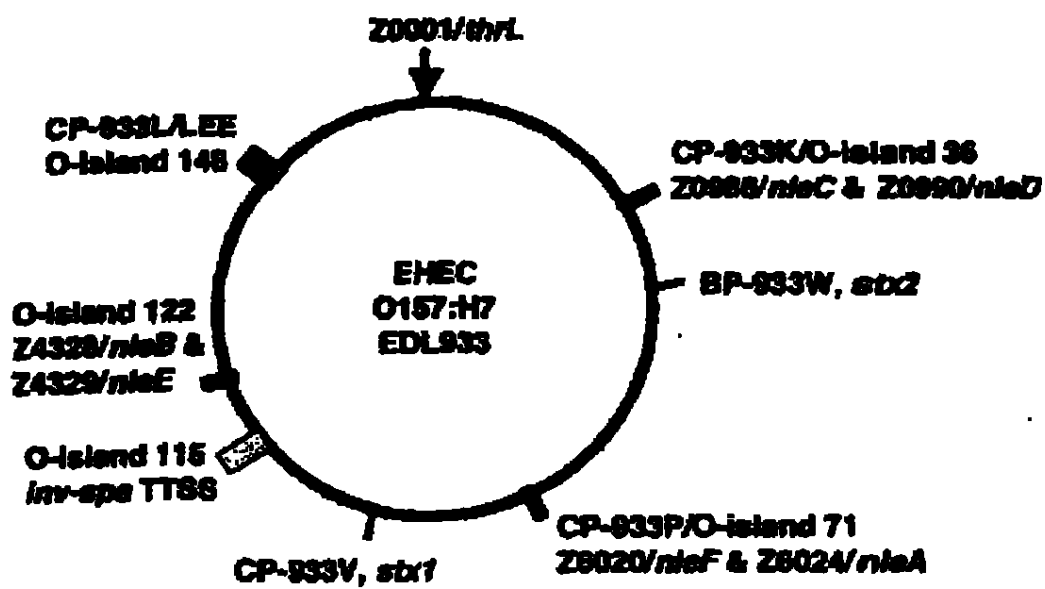


FIG. 9

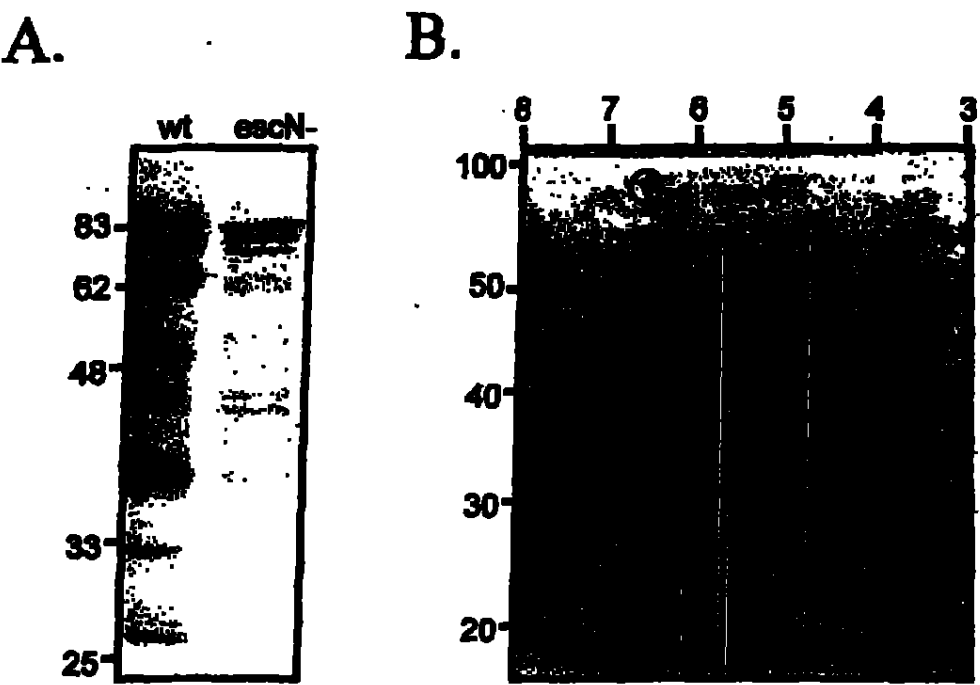
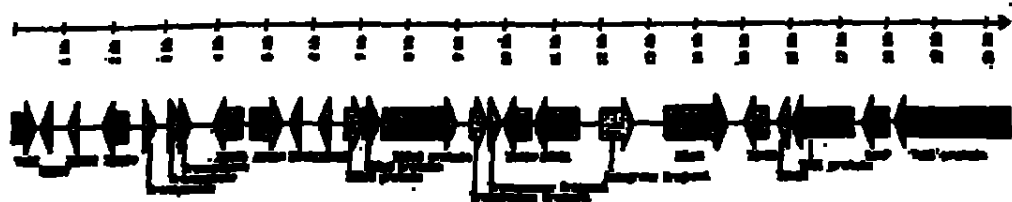


FIG. 10

A.



B.



C.

EPEC	MTQVYQSTTQSGNQ-SAGTQSLPS-TQYDQMLPLQSGNPPVQVHIFQDQDQANQ	88
OS4_M1	IV.....T.....S.....P.....W..K.....	90
EPEC	IVT.....T..S...NR...TSP.....W..ST.....E..E..NR.....	94
Citrobacter	E..E.....T.....Q..NR..V.VSGS..E.D..SR.E.....N..E.....V....	98
EPEC	STVLLSLTQVLLSGLSTLAPLSSGSRQCKTLPQPLPSTQVALLSTPS-VGSLIS	117
OS4_M1		118
EPEC	L.....E.....A..A.....S.....T..P....	117
Citrobacter	E.....K.L.....L..A..T.....P....	118
EPEC	GRVQGLSTVQDGLPSTGNTPLQPLPLSGRQSPVQSLPSTLALSLSSGSRQV--	178
OS4_M1	N.....S.....	178
EPEC		178
Citrobacter	N.....T.....L.....T..S.....	178
EPEC	-----TATGSRGATSTVSTGSRVLPQPLPSTQVALLSTPS-VGSLIS	218
OS4_M1	ASSTGSRGATSTV	218
EPEC	-----A...ZA.....AS.....	217
Citrobacter	-----A...AST..S..N.....	207
EPEC	TQVLLSLTQVLLSGLSTLAPLSSGSRQCKTLPQPLPSTQVALLSTPS-VGSLIS	278
OS4_M1		280
EPEC		277
Citrobacter	T.....	267
EPEC	PTSTGSRGATSTVSTGSRVLPQPLPSTQVALLSTPS-VGSLIS	318
OS4_M1	T..V..A.....LSPV.....A...S.A...L...V.CY.....	318
EPEC		317
Citrobacter	I.....I..A.....EVSIS.....V..L.....V.C.....	327
EPEC	TLVQVLLSLTQVLLSGLSTLAPLSSGSRQCKTLPQPLPSTQVALLSTPS-VGSLIS	398
OS4_M1	T..L.....G.....E..AR.....	418
EPEC	T..L.....GR.....	397
Citrobacter	P.....SRV.....F.....L...Q..SRV...R....	387
EPEC	QLPQVLLSLTQVLLSGLSTLAPLSSGSRQCKTLPQPLPSTQVALLSTPS-VGSLIS	441
OS4_M1	.V..SR...N.....S...N...Y..TV...R....	439
EPEC	.VS.....N.....SR.....Y..TV.....	448
Citrobacter	.V...SQ..X.....GRS.....Y..T.R.....	438

FIG. 11

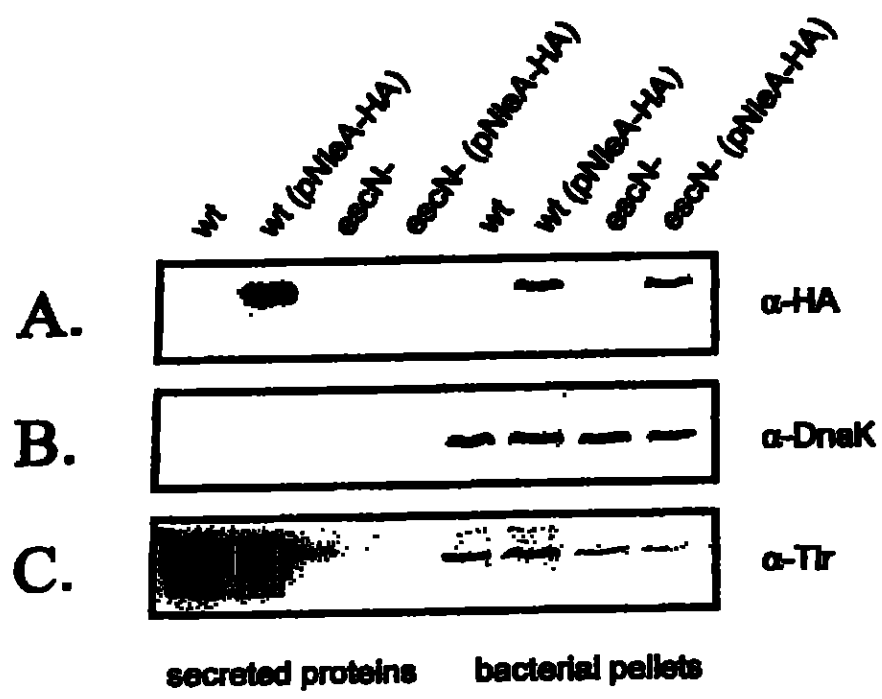


FIG. 12

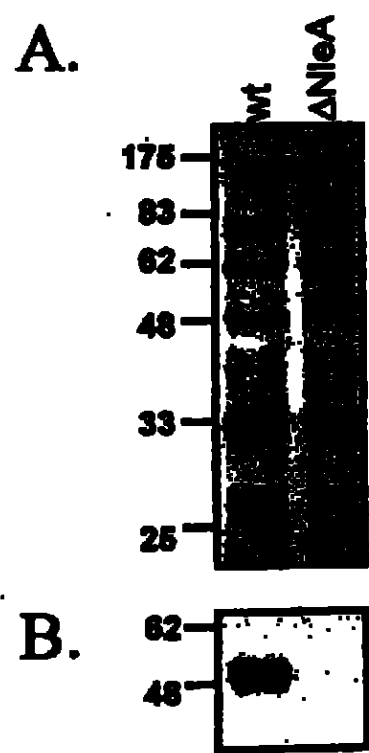
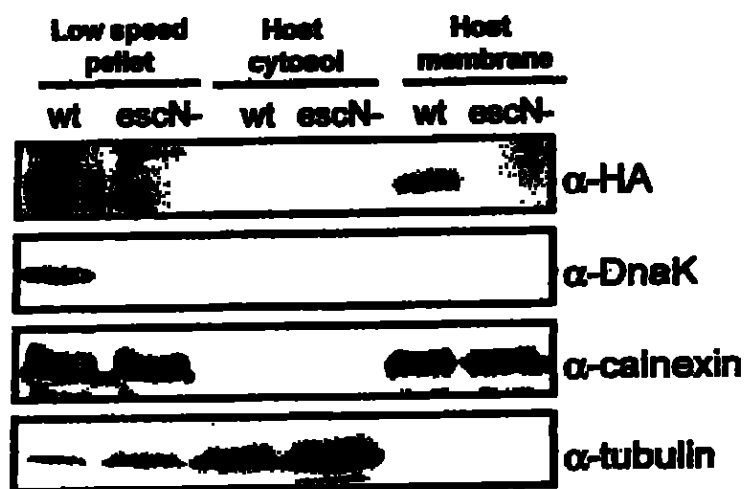


FIG. 13

24/37

A.



B.

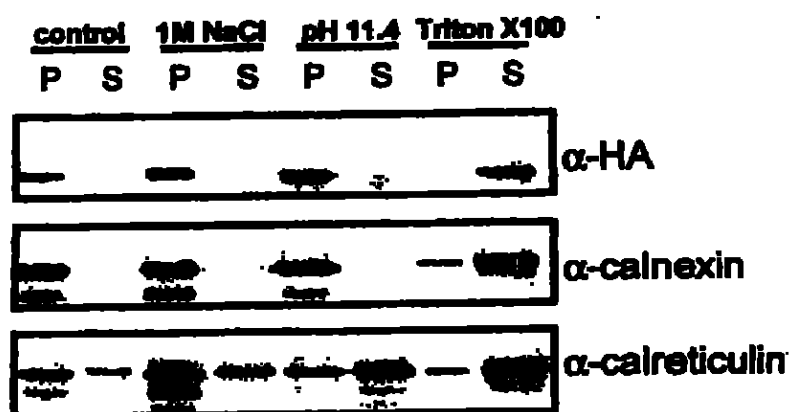
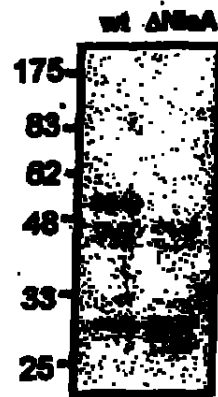
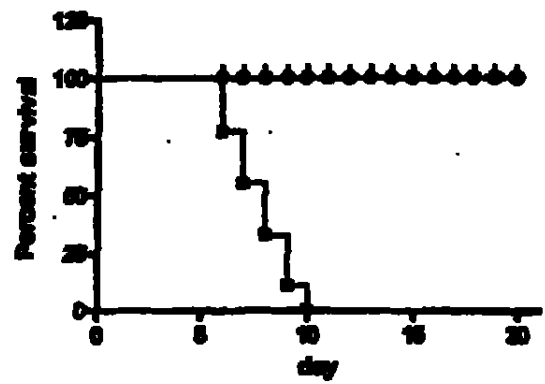


FIG. 14

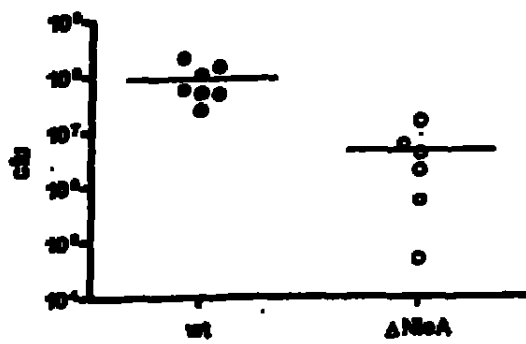
A.



B.



C.



D.

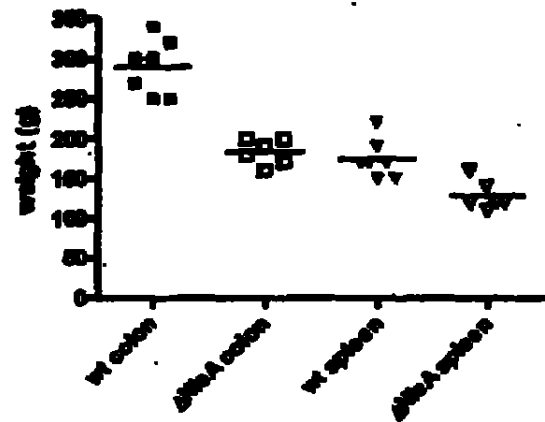


FIG. 15

Enterohemorrhagic *E. coli* NleG homolog Nucleotide Sequence

ATGAATGTCCTTCGAGCTCAAGTAGCATCTAGCGGTCGAGGGGAGTTTACATTAGGTAA
GAGACTGTCAGCATTGTATTTAATGAAACCGATGGGCGTTTTCTATCCAGCGGCAGTAGT
GGGGGATTGCTTACTGAGTTATTCTTTATGGGTTTAATAACGGCCCTGAAGCTCTTCGC
GATAGGATGCTCAGTATGCTTTCGGACTCAGGTGAAGCACAATCGCAAGAGAGTATTCAG
GACAAAATATCTCAATGTAAGTTTCCTGTTAGTTCAGGAAATTTCCAGTGCCCGCCAGAG
TCTATTCAGTGTCCAATTACACTAGAGAGACCCGAAGAAGGAGTGTTTGTCAAAAATTCA
GATAGTTCGGCAGTATGCTGCTTATTTGATTTTGATGCATTTTCTCGTTTAGCTAGTGAA
GGCTCATATCATCCACTGACCGAGAAOCCAATAACGGCATCAATGATTATAAGTCCTGAT
AAATGTGTTTATGATCCTATCAAGGGAACTTCATTATAAAAGATAGTTAA
(SEQ ID NO: 61)

FIG. 16AEnterohemorrhagic *E. coli* NleG homolog Protein Sequence

MNVLRAQVASSGRGEFTLGNETVSIVFNETDGRFLSSGSSGGLLTELFLYGFNNGPEALRDRMLSMLSD
SGEAQSQESIQDKISQCKFPVSSGNFQCPPESIQCPTILERPEEGVFVKNSDSSAVCCCLFDFDAFSRLA
SEGSYHPLTREPITASMIISPDKCVYDPKGNFIKDS (SEQ ID NO: 73)

FIG. 16B

Enterohemorrhagic E. coli NleH1 Nucleotide Sequence

ATGTTATCGCCCTCTTCTATAAATTTGGGATGTTTCATGGAATTCCTTAACCAGAAACCTG
 ACTTCGCCTGATAATCGTGTTTTATCCTCTGTAAGGGATGCTGCTGTTCACTCTGATAGC
 GGGACGCAAGTAACGGTTGGCAACAGAACATATCGTGTTGTGGTCACTGATAATAAGTTT
 TGOGTTACAAGAGAAAGTCATAGTGGTTGTTTTACTAATCTGTTGCACAGGTTGGGATGG
 CCTAAGGGAGAGATTAGCAGAAAAATTGAGGCTATGCTGAATACATCGCCAGTGAGCACG
 ACTATAGAAAGAGGCTCTGTTCAATCGAACAGACCTGATTTACCTCCAGTGGATTATGCG
 CAGCCGGAGTTACCTCCAGCGGATTATACTCAATCAGAGTTGCCGAGGGTTAGCAACAAT
 AAATCACCCGTCGCCAGGTAATGTTATTGGTAAAGGTGGTAATGCTGTCGTGTATGAAGAT
 ATGGAAGATACAACAAAAGTGTTGAAGATGTTTACTATATCTCAAAGOCATGAAGAGGTG
 ACAAGCGAAGTTCGTTGTTTCAATCAGTATTATGGTTCGGGAGTGCGAGAGAAATATAT
 AATGATAATGGAAATGTTATTGGTATTAGAATGAATAAAATAAATGGGGAATCTCTTTTG
 GATATTCATCATTACCAGCACAAAGCTGAACAGGCTATTTACGATATGTTTGACAGACTG
 GAGAAAAAAGGAATTCCTTTTGTGATACAACAGAAACAAATGTTTTATATGATCGTATG
 AGAAATGAATTTAATCCAATAGATATATCATCTTATAATGTTTCTGATATTTTCATGGAGT
 GAACATCAAGTCATGCAATCTTATCACGGAGGAAAGCTGGATCTTATTAGTGTAGTATTA
 AGTAAGATATAA

(SEQ ID NO: 62)

FIG. 17AEnterohemorrhagic E. coli NleH1 Protein Sequence

MLSPSSINLGCSWNSLTRNLSPDNRLSSVRDAAVHSDSGTQVTVGNRTYRVVVTDNKF
 CVTRESHSGCFTNLLHRLGNPKGEISRKIEAMLNTSPVSTTIERGSVHSNRPDLPPVDYA
 QPELPPADYTQSELPRVSNNKSPVPGNVIGKGGNAVVEDMEDTTKVLKMFISQSHEEV
 TSEVRCFNQYYGSGSAEKIYNDNGNVIGIRMNKINGESLLDIPSLPAQAEQAIYDMFDRL
 EKKGILFVDTTETNVLYDMRNEFNPIDISSYNVSDISWSEHQVMQSYHGGKLDLISVVL
 SKI (SEQ ID NO: 74)

FIG. 17B

Enterohemorrhagic E. coli NleH2 Nucleotide Sequence

ATGTTATCGCCATATTCTGTAAATTTGGGATGTTTCATGGAATTCCTTAACCAGAAACCTG
 ACTTCGCCTGATAATCGTGTTTTATCCTCTGTAAAGGGATGCTGCCGTTCACTCTGATAAT
 GGGGCGCAAGTAAAGGTTGGCAACAGAACATATCGTGTTGTTGCCACCGATAATAAGTTT
 TGGGTTACAAGAGAAAGTCATAGTGGTTGTTTTACTAATCTGTTGCACAGGCTGGGATGG
 CCTAAGGGGGAGATTAGCAGGAAAATTGAGGTCATGCTGAATGCATCACCAGTGAGCGCT
 GCTATGQAAAGAGGCATTGTTCAATCGAACAGACCTGATTTACCTCCTGTTGATTATGCA
 CCGCCAGAGTTACCGAGTGTGGACTATAACAGGTTGTCAGTACCTGGTAATGTTATTGGC
 AAAGGGGGGAACGCTGTAGTATATGAAGATGCTGAGGATGCAACAAAAGTCCTGAAGATG
 TTTACTACATCTCAAAGCAATGAAGAGGTGACAAGCGAAGTTCGTTGCTTCAACCAATAT
 TATGGTGCCGGGAGTGCAGAAAAATATATGGCAATAATGGTGATATTATTGGTATTAGA
 ATGGATAAAATAAATGGAGAAATCGCTTTTAAATATTTTCGTCCTTGCCAGCACAGGCTGAG
 CATGCTATTTACGATATGTTTGATAGACTGGAGCAAAAAGGAATTCCTTTTGTGATACA
 ACAGAGACAAATGTCTTATATGACCGCGCGAAGAATGAGTTTAAATCCAATAGATATATCA
 TCTTATAATGTTTCCGACCGTTCATGGAGTGAAGTCAAATAATGCAATCTTATCATGGC
 GGAAAGCAAGATCTTATTAGTGTGGTATTAAGTAAAATTTAG (SEQ ID NO: 63)

FIG. 18AEnterohemorrhagic E. coli NleH2 Protein Sequence

MLSPYSVNLGCSWNSLTRNLTS PDNRVLSSVRDAAVHSDNGAQVKVGNRTYRVVATDNKF
 CVTRESHSGCFTNLLHRLGWPKGEISRKIEVMLNASPVSAAMERGIVHSNRPOLPPVDYA
 PPFLPSVDYNRLSVPGNVIGKGGNAVVEYDAEDATKVLKMFSTTSQSNEEVTSEVRCFNQY
 YGAGSAEKIYGNGDIIGIRMDKINGESLLNISSLPAQAEHAIYDMFDRLEQKGILFVDI
 TETNVLYDRAKNEFNPIDISSYNVSDRSWSESQIMQSYHGGKQDLISVVLSKI (SEQ ID NO: 75)

FIG. 18B

29/37

Enterohemorrhagic E. coli Z2076 Nucleotide Sequence

ATGGTAATGCCTGGATTAGTATCATATATATCATCGACTTCATTGCGAATGAGATGGCG
GAGATGCGTCAGCAGGTAATGGAAGGGCAGATTGGTGGATTCTCCTGGGAGGGGAGAGA
GTTAGAGTTTCTTATTTATTTCAATTGCATTAA (SEQ ID NO: 64)

FIG. 19AEnterohemorrhagic E. coli Z2076 Protein Sequence

MVMPGLVSYISSTSFANEMAEMRQQVMEGQIGGFLLGGERVRVSYLFQLR
(SEQ ID NO: 76)

FIG. 19B

Enterohemorrhagic *E. coli* Z2149 Nucleotide Sequence

ATGCCATTAACTCAGATATTAGATCACATTCATTTAATCTTGGGGTGGAGGTTGTTTCGT
GCCCGAATTGTAGCCAATGGGCGCGGAGATATTACAGTCGGTGGTGAAACTGTCAGTATT
GTGTATGATTCTACTAATGGCGCTTTTCATCCAGTGGCGGTAATGGCGGATTGCTTTCT
GAGTTATTGCTTTTGGGATTTAATAGTGGTCCTCGAGCCCTTGGTGAGAGAATGCTAAGT
ATGCTTTCGGACTCAGGTGAAGCACAAATCGCAAGAGAGTATTCAGAACAAAATATCTCAA
TGTAAGTTTCTGTTTGTCAGAGAGACTTCAGTGCCCGCTTGAGGCTATTCAGTGTCCA
ATTACACTGGAGCAGCCTGAAAAAGGTATTTTGTGAAGAATTCAGATGGTTCAGATGTA
TGTACTTTATTGATGCCGCTGCATTTTCTCGTTTGTTGGTGAAGGCTTACCCACCCA
CTGACCCGGGAACCAATAACGGCATCAATAATTGTAAAACATGAAGAATGCATTTATGAC
GATACCAGAGGAACTTCATTATAAAGGGTAATTGA (SEQ ID NO: 65)

FIG. 20A

Enterohemorrhagic *E. coli* Z2149 Protein Sequence

MPLTSDIRSHSFNLGVEVVRARIVANGRGDITVGGETVSIVYDSTNGRFSSSGNGGLLS
ELLLLGFNSGPRALGERMLSMLSDSGEAQSQESIQNKISQCKFSVCPERLQCPLEAIQCP
ITLEQPENGIFVKNSDGSVDVCTLFDAAAF SRLVGEGLPHFLTREPITASIIVKHEECIYD
DTRGNFIKGN (SEQ ID NO: 77)

FIG. 20B

SEQUENCE LISTING

<110> THE UNIVERSITY OF BRITISH COLUMBIA, et al.

<120> BACTERIAL VIRULENCE FACTORS AND USES THEREOF

<130> 80021-735

<140> NOT YET ASSIGNED

<141> 2004-10-29

<150> US 60/515,703

<151> 2003-10-31

<160> 84

<170> PatentIn version 3.3

<210> 1

<211> 1293

<212> DNA

<213> Citrobacter rodentium

<400> 1

atgaacattc	aaccgaacat	acattccgga	atcaccacac	agaataatca	acaacatcat	60
cacgcagaac	aagtgcctgt	ctctagctca	ataccgcgat	cagatttacc	tccaaattgc	120
gaagctggat	ttgtttgtga	tattccagag	gatatacagc	aacatgtacc	ggaatgtggt	180
gaaacaacgg	ctctattaag	cttgataaaa	gatgaaggcc	tgctctcagg	actagataaa	240
tatcttgctc	ctcaccttga	agaaggctcc	cttgggaaaa	aagcattgga	tacgtttggt	300
ttattcaatg	ttactcaaat	ggcattagag	ataccagtt	ccgttccagg	catatctggt	360
aaatatggtg	ttcagatgaa	cattgtaaaa	ccagacatac	atccaacaac	cgggaactat	420
tttttacagc	tatttcctct	gcatgacgaa	ataggtttta	acttcaaaga	tcttcctggc	480
ccattaaaaa	atgcattaac	caacagcagt	atatcggcta	ctgcatcgac	tgtagccccc	540
acaccaaacy	acccaatgcc	atggtttgga	ttaactgctc	aagtggttcg	taatcatggt	600
gtagaacttc	ctatagtcaa	aaccgaaaat	ggatggaagc	ttgtagggga	aacacctctt	660
actccagatg	gccccaaagc	caattatacg	gaagaatggg	ttatcaggcc	gggagaagca	720
gattttaaat	atggaacatc	gccattacag	gcaactcttg	gactggagtt	tgggtgcgcat	780
tttaagtggg	atttagataa	tcctaatacc	aaatatgcc	tccttaccaa	tgctgccgca	840
aatgctattg	gtgctgctgg	agggtttgcg	gtatccaaag	tccccggcat	agatccaatg	900
ctgtcccctc	atgtcgggtg	aatgcttggt	caagcagcgg	ggcatgccgt	acaatgtaat	960
acccccggat	taaagccaga	cactatttta	tgggtggcag	gcgcgacatt	tggagctgct	1020
gatttaataa	aagccgaatt	tgataaagtg	cggttcactg	actaccctcg	tatatggttc	1080
catgacgggg	aaggagcttt	attcccaaat	aagcaagaca	ttgcccggtg	aacaggcgca	1140
gacataaaaag	ctatggaaga	agggcgtaacc	gttgacatc	aacatccaaa	accggaggat	1200
gtggtcatcg	atatcgaagg	tggcaattca	ccacatcata	atccatcaaa	ttatgttgac	1260
acctttgaaa	taatccaaga	aacaagggtc	taa			1293

<210> 2

<211> 1323

<212> DNA

<213> Enteropathogenic E. coli

<400> 2

atgaacattc	aaccgatcgt	aacatccgga	atcaccacac	aaaacaatcg	acatcatcac	60
gcagaacaaa	cgtcccctac	acaaataccg	caatccgaat	tacctaattg	atgcgaaacy	120
ggatttggtg	ttcatatccc	agaggatatg	cagcgacatg	caccggaatg	cggtgaaaca	180
acagctctac	tgagcttgat	aaaagatgaa	ggtctgctct	ctgggctgga	taaatactct	240
gcacctcatc	ttgaagaagg	ctctgcagga	aaaaaagcat	tggatatggt	tggtttattc	300
aatgtctctc	agatggcatt	agaaataccc	agcaccgttc	cgggtatctc	tggtaaatac	360
ggtgtccagc	taaacattgt	aaaaccagat	attcatccta	catcaggtaa	ttatTTTTTA	420
cagatatattc	ctttgcatga	tgaaataggt	attaatttta	aagaccttcc	tgggtccatta	480
aaaaatgcat	taagcaacag	caatatacca	accactgtat	cgactgctgc	atccactatt	540
gcatcagcca	ctacttcgac	ggtaaccacc	gcgtcaaaag	acccaatacc	atggtttgga	600
ttacagctc	aagtagttcg	taatcatggt	gtggaacttc	ctatagtcaa	aactgaaaat	660
ggatggaagc	ttgttgagga	aactcctctt	actcctgatg	gccccaaagc	aaattatact	720
gaagagtggg	tgatcagacc	gggagaagca	gattttaaat	atgggtgcatc	tccactacag	780
gcaactctag	ggctggagtt	tggcgcacat	ttcaagtggg	atttagataa	ccctaatact	840
aaatatgccg	ttcttaccaa	tgctgccgca	aatgcgcttg	gtgctgtagg	gggatttgca	900
gtatccagat	ttactggtac	agatccaatg	ttaagtcctc	atatcgggtg	aatggttggg	960
caagcagcgg	ggcatgccat	acagtataat	acccccggat	taaagccaga	cactatttta	1020

Enterohemorrhagic *E. coli* Z2150 Nucleotide Sequence

```

atgcctgtta ccaccttaag tatcccaagt atatctcaat tatctcctgc aagagtacag
tctttgcagg atgcagccag acttgaaaagt ggaataagaa tatccattgg tagtggccaa
tattctgttc actatgtcca actactggat ggattttcag ttgaaccggg gagaggaggc
ttactggata ggctattggg gcgtgagcat cgaatggata gaagggctgt ggctctggaa
aggcaattaa atggaggtgt cgatttttta agtagtgta ataactattt tcagagtgtc
atggcagaac acagagaaaa taaaacaggc aataaaatat taatggaaaa aataaattct
tgtgtatttg gaacggattc taatcacttt tcttgcccgg agtcattttt gacatgcccg
ataacgctgg acacacctna gactggagtg ttcattgagaa actcacgagg tgctgagata
tgctctctat atgataagga tgcgttagtg caacttggtg aaactgggtg aactcatcct
ctgagtcgag aacctataac agaatcaatg attatgagaa aagacgaatg tcactttgat
gcmetaagag aagctttttg ttgtaagtga
(SEQ ID NO: 66)

```

FIG. 21AEnterohemorrhagic *E. coli* Z2150 Protein Sequence

```

MPVTTLSPISISQLSPARVQSLQDAARLESGIRISIGSGQYSVHYVQLLDGFSVEPVRGGLLDRLIGRE
HRMDRRAVALERQLNGGVDFLSSVINNYFQSVMAEHRENKTGNKILMEKINSCVFGTDSNHFSCPESFLT
CPITLDTPXTGVFMRNSRGAEICSLYDKDALVQLVETGGTHPLSREPITESMIMRKDECHFDAKREAF
CK (SEQ ID NO:78)

```

FIG. 21B

Enterohemorrhagic *E. coli* 22151 Nucleotide Sequence

ATGCCTGTAGATTTAACGCCCTTATATTTTACCTGGGGTTAGTTTTTTTGTCTGACATTCCT
CAAGAAACCTTGTCTGAGATACGTAATCAGACTATTCTGGAGAAGCTCAAGTAAGACTG
GGTGAGTTGATGGTGTCAATACGACCTATGCAGGTAAATGGATATTTTATGGGAAGTCTT
AACCAGGATGGTTTATCGAATGATAACATCCAGATTGGCCTTCAATATATAGAACATATT
GAACGTACACTTAATCATGGTAGTTTGACAAGCCGTGAAGTTACAGTACTGCGTGAAATT
GAGATGCTCGAAAAATATGGAATTGCTTTCTAACTACCAGTTAGAGGAGTTGTTAGATAAA
ATTGAAGTATGTGCATTTAATGTGGAGCATGCACAATTGCAAGTGCCAGAGAGCTTACGA
ACATGCCCTGTTACATTATGTGAACCAGAAGATGGGGTATTTATGAGGAATTC AATGAAT
TCAAATGTTTGTATGTTGTATGATAAAATGTCATTAAATATATCTTGTTAAAACAAGGGCG
GCTCATCCTTTGAGCAGGGAAATCAATCGCAGTTTCAATGATTGTAGGAAGAGATAATTGT
GCTTTTGACTCTGACAGAGGTAACTTCGTTTTAAAAAATTAA (SEQ ID NO: 67)

FIG. 22A**Enterohemorrhagic *E. coli* 22151 Protein Sequence**

MPVDLTPTYILPGVSFLSDIPQETLSEIRMQTIERGEAQVRLGELMVSIRPMQVNGYFMGSL
NQDGLSNDNIQIGLQYIEHIERTLNHGSLTSREVTVLREIEMLENMELLSNYQLEELLDK
IEVCAFNV EHAQLQVPESLRTCPVTLCEPEDGVFMHNSMNSNVCMLYDKMSLIYLVKTRA
AHPLSRESIAVSMIVGRONCAFDSDRGNEVLKN (SEQ ID NO: 79)

FIG. 22B

Enterohemorrhagic E. coli Z2337 Nucleotide Sequence

ATGCCTGTAGATTTAACGCCTTATATTTTACCTGGGGTTAGTTTTTTGTCTGACATTCCT
CAAGAAACCTTGCTCGAGTAACGTAATCAGACTATTCGTGGAGAAGCTCAAATAAGACTG
GGTGAGTTGATGGTGTCAATACGACCTATGCAGGTAAATGGATATTTTATGGGAAGTCTT
AACCAGGATGGTTTATCGAATGATAATATCCAGATTGGCCTTCAATATATAGAACATATT
GAACGTACACTTAATCATGGTAGTTTGACAAGCCGTGAAGTTACAGTACTGCGTGAAATT
GAGATGCTCGAAAATATGGATTGCTTTCTAACCTACCAGTTAGAGGAGTTGTTAGATAAA
ATTGAAGTATGTGCATTTAATGTGGAGCATGCACAAATGCAAGTGCCAGAGAGCTTACGA
ACATGCCCTGTTACATTATGTGAACCAGAGATGGGGTATTTATGAGGAATTCAATGAAT
TCAAATGTTTGTATGTTGTATGATAAAATGGCATTAAATACATCTTGTAAAACAAGGGCG
GCTCATCCTTTGAGCAGGGAATCAATCGCAGTTTCAATGATTGTAGGAAGAGATAATTGT
GCTTTTGACCCTGACAGAGGTAACCTTCGTTTTAAAAAATTAA (SEQ ID NO: 68)

FIG. 23AEnterohemorrhagic E. coli Z2337 Protein Sequence

MFVDLTPYILPGVSFLSDIPQETLSEIRNQTIRGEAQIRLGELMVSIRPMQVNGYFMGSL
NQDGLSNDNIQIGLOYIEHIERTLNHGSLTSREVTVLREIEMLENMDLLSNYQLEELLDK
IEVCAFVVEHAQLQVPESLRTCPVTLCPEPEDGVFMRNSMNSNVCHLYDKMALIHLVKTRA
AHPLSRESIAVSMIVGRDNCAFDPRGNFVLKN (SEQ ID NO: 80)

FIG. 23B

Enterohemorrhagic E. coli Z2338 Nucleotide Sequence

ATGCCTGTTACCACCTTAAGTATCCCAAGTATATCTCAATTATCTCCTGCAGGAGTACAG
TCTTTGCAGGATGCTGCCAGACTTGAAAGTGGAAATAAGAATATCCATTGGTAGTGGCCAA
TATTCTGTTCACTATGTCCAGCTACTGGATGGATTTTCAGTTGAACCGGTGAGAGGAGGC
TTACTGGATAGGCTATTGGGGCGTGAGCATCGAATGGAGAGAAGGGCTGTGGCTCTGGAA
AGGCAATTAAATGGAGGTGTCGATTTTTTAAGTAGTGTTAATAACTATTTTCAGAGTGTC
ATGGCAGAACACAGAGAAAATAAAACAAGTAATAAATATTAATGGAAAAATAAATTCT
TGTTTATTAGACCTGATTCTAATCACTTTTCTTGCCCGGAGTCATTTTGGACATGCCCG
ATAACGCTGGACACACCTGAGACTGGGGTGTTTCATGAGAACTCACGAGGTGCTGAGATA
TGCTCTCTATATGATAAGGACGCGTTAGTGCAACTTGTTGAACTGGTGGAGCTCATCCT
CTGAGTCGAGAACCTATAACAGAATCAATGATTATGAGAAAAGATGAATGTCACTTTGAT
ACAAAAGAGAGAAGCTTTTTGTTGTAAGTGA (SEQ ID NO: 69)

FIG. 24AEnterohemorrhagic E. coli Z2338 Protein Sequence

MPVTTLSPISISQLSPAGVQSLQDAARLES GIRISIGSGQYSVHYVQLLDGFSVEPVRGG
LLDRLLGREHRMERRAVALERQLNGGVDFLSSVNNYFQSVMAEHRENKTSNKILMEKINS
CLFRPDSNHFSCFESFLTCPITLDTPETGVFMRNSRGAEICSLYDKDALVQLVETGGAHP
LSREPITESMIMRKDECHFDTKREAFCK (SEQ ID NO: 81)

FIG. 24B

Enterohemorrhagic E. coli 22339 Nucleotide Sequence

ATGCCATTAACTCAGATATTAGATCACATTCATTTAATCTTGGGGTGGAGGTTGTTTCGT
GCCCGAATTGTAGCCAATGGGCGCGGAGATATTACAGTCGGTGGTGAACTGTCAGTATT
GTGTATGATTCTACTAATGGGCGCTTTTCATCCAGTGGCGGTAATGGCGGATTGCTTCT
GAGTTATTGCTTTTGGGATTTAATAGTGGTCCTCGAGCCCTTGGTGAGAGAATGCTAAGT
ATGCTTTCGGACTCAGGTGAAGCACAAATCGCAAGAGAGTATTCAGAACAATAATCTCAA
TGTAAGTTTCTGTTTGTCCAGAGAGACTTCAGTGCCCGCTTGAGGCTATTCARTGTCCA
ATTACACTGGAGCAGCCTGAAAAAGGTATTTTGTGAAGAATTCAGATGGTTCAGATGTA
TGTACTTTATTTGATGCCGCTGCATTTTCTCGTTTGGTTGGTGAAGGCTTACCCACCCA
CTGACCCGGGAACCAATAACGGCATCAATAATTGTAAACATGAAGAATGCATTTATGAC
GATACCAGAGGAACTTCGTTATAAAGGGTAATTGA (SEQ ID NO: 70)

FIG. 25AEnterohemorrhagic E. coli 22339 Protein Sequence

MPLTSDIRSHSFNLGVEVVRARIVANGRGDITVGGETVSVIVYDSTNGRFSSSGNGGLLS
ELLLLGFNSGPRALGERMLSLSDSGEAQSQESIQNKISQCKFSVCPERLQCPLEAIQCP
ITLQPEKGIFVKNSDGSVDVCTLFDAAAFSLVGEGLPHPLTREPI TASIIVKHEECIYD
DTRGNFVIKGN (SEQ ID NO: 82)

FIG. 25B

Enterohemorrhagic E. coli Z2560 Nucleotide Sequence

atggacgctt ttattgtaga tectgttcaa ggggaactat attcgggttt aagccataca
gaactagccg atatcattag attggtgat tctgttgaaa atcaattgaa tggaggcaat
tcatttcttg atgtattcag tacatatatg gggcaggtta tttctgaatt tatgcatagt
aatgataaca gaattgaatt gttacagcgg cgattacatt catgttcatt tttagttaat
attgaagaaa tgtcttacat agatgaagca ttacagtgcc cgattacgct ggcaattcct
caacgagggtg tttttttaag aaatgctgaa gggtccagag tatgtagttt atatgatgaa
atggctcttt ctogtataat taatgatggg atgcatcacc cactaagcag agagccaata
acattatcaa tgcttggtgc cagagagcag tgtgagtttg attgcagtat cggtcacttt
acggtgagga gtgattgtta ttcagtgtag (SEQ ID NO: 71)

FIG. 26AEnterohemorrhagic E. coli Z2560 Protein Sequence

MDAFIVDPVQGLYSGLSHTELADIIRLADSVENQLNGGNSFLDVSTYMGQVISEFMHSNDNRIELQ
RRLHSCSFLVNIEMSYIDEALQCPITLAIPQGVFLRNAEGSRVCSLYDEMALSRIINDGMHHPLSRE
PITLSMLVAREQCEFDCSIGHFTVRSDCYSV (SEQ ID NO: 83)

FIG. 26B

Enterohemorrhagic E. coli Z2976 Nucleotide Sequence

ATGGCAGACCGCAAACAGCACCGCGCTATCGCGGAGCGTCGTCACATCCAGACTGAAATC
AACCGCAGACTTTCCCGCGCATCACGCGTCGCGCAAATCATGCACATCAATATGCTGCAT
GAGCGCAGCCACGCACTATCAAACATTTATTCGCGCTCTGTTTTCAGCTATCTGGCGGAT
GATCTGCACGAGTTTCAACAGCTCATCCAGCAGCAAAACAAACTCCATTAA (SEQ ID NO:72)

FIG. 27A

Enterohemorrhagic E. coli Z2976 Protein Sequence

MADRKQHRAIAERRHIQTEINRRLSRASRVAQIMHINMLHERSHALSNIYSASVFSYLAD
DLHEFQQLIQQONKLH (SEQ ID NO:84)

FIG. 27B

tggtgggcag	gtactactct	tggaactggct	gatttaaaca	aggccgagtt	tggaagagcc	1080
agattcactg	actatcctcg	tatatgggtg	catgcaagag	aaggtgccat	tttcccaaat	1140
aaagcagata	ttgaacatgc	cacaggggct	gatatacgcg	caatggaaga	aggtgtatct	1200
gttggacaac	ggcatccaaa	tccagaggat	gtggatcatca	atatcgaaag	caataactca	1260
ccacatcata	acccatcaaa	ttatgttgat	accgttgata	taatccaaga	aacaagagtc	1320
taa						1323

<210> 3
 <211> 1326
 <212> DNA
 <213> Enterohemorrhagic E. coli

<400> 3						
atgaacattc	aacgcaccat	acaatctgga	atcacctcac	aaaacaatca	acatcatcaa	60
acagaacaaa	taccctctac	acaaataccg	caatccgaat	tacctctagg	atgccaaagt	120
ggattttgtt	ttaatattcc	agatgatata	cagcaacatg	caccggaatg	cgggtgaaaca	180
acagctctac	tgagcttgat	aaaagataaa	ggctctgctct	cagggctaga	cgaatatata	240
gctcctcacc	ttgaagaagg	atccatagga	aaaaaaacat	tgatatgtt	tggtttattc	300
aatgttacct	aaatggcatt	agagatacct	agttccgttt	caggcatctc	tggtaaatat	360
ggtgtccagc	taaacattgt	aaaaccagat	attcatccta	catcaggtaa	ttatTTTTTA	420
cagatattcc	ctctgcatga	tgaaataggt	tttaatttta	aagaccttcc	tggtccgtta	480
aaaaatgcat	taagcaacag	taatatatca	accactgcag	tgtcgactat	tgcatcgact	540
ggaacatcag	ccactacttc	gacggtaacc	accgagccaa	aagacccaat	accatggttt	600
ggattaacag	ctcaagtggg	tcgtaatcat	ggtgtagaac	ttcctatagt	caaaactgaa	660
aatggatgga	agcttggttg	agaaacacca	cttactcctg	atgggccgaa	agcaaattac	720
acggaggagt	gggttatcag	accgggagaa	gcagatttta	aatatgggtg	atctccatta	780
caggcaactc	tagggctgga	gtttggcgca	catttcaagt	gggatttaga	taaccctaata	840
actaaatatg	ccgttcttac	caatgctgcc	gcaaatgcgc	ttggtgcttt	agggggattt	900
gcagtatcca	gatttgctag	tacagatcca	atgttaagtc	ctcatatcgg	tgcaatgggt	960
gggcaagcag	cagggcatgc	catacagtat	aatacccctg	gattaaagcc	agacactatt	1020
ttatggtggg	ctggtgcgac	actgggggct	gccgatttaa	acaaggccga	gtttgaagta	1080
gctagattca	ctgactatcc	tcgtatatgg	tgccacgcaa	gagaaggagc	tattttcccc	1140
aataaagcag	atattgaaca	tgccacaggt	gctgatatac	gcgcaatgga	agaagggtatc	1200
cctgttggac	agcggcatcc	aaatccagag	gatgtggtta	tcgatatcga	aagcaatggc	1260
ttaccacatc	ataatccatc	aaatcatgtt	gatatctttg	atataatcca	agaaacaaga	1320
gtctaa						1326

<210> 4
 <211> 614
 <212> DNA
 <213> Citrobacter rodentium

<400> 4						
tattcttttt	cagtgggttg	aagcaaggcc	agagcgatac	ggaaaagggtg	aagtaccgat	60
attgaatacc	aaagagcatc	cgtattttgag	caatattata	aatgctgcaa	aaatagaaaa	120
tgagcgcgta	ataggagtac	tggtagacgg	agactttact	tatgagcaaa	gaaaagaatt	180
tctcagtctt	gaagatgaac	atcaaaatat	aaagataata	tatcgggaaa	atgttgattt	240
cagtatgtat	gataaaaaac	tgtctgatat	ttatcttgaa	aataattcatg	aacaagaatc	300
atatccagcg	agtgaagagag	ataattatct	gttaggctta	ttaagagaag	agttaaaaaa	360
tattccatac	ggaaaggact	ctttgattga	atcatatgca	gaaaaagag	gtcatacttg	420
gttttgatttt	tttagaaact	tggcggtatt	gaaggggggg	gggttgttta	cagagacggg	480
taaaactgga	tgccataaca	tatctccatg	tgggggatgt	atatatcttg	atgcagatat	540
gattattact	gataaattag	gtgtcctgta	tgctcctgat	ggtatcgctg	tgcatgtaga	600
ttgtaatgat	gaga					614

<210> 5
 <211> 555
 <212> DNA
 <213> Citrobacter rodentium

<400> 5						
aaggccagag	cgatacggaa	aagggtgaagt	accgatattg	aataccaaag	agcatccgta	60
tttgagcaat	attataaatg	ctgcaaaaaat	agaaaatgag	cgtaatggag	tactggtaga	120
cggagacttt	acttatgagc	aaagaaaaga	atctctcagt	cttgaagatg	aacatcaaaa	180
tataaagata	atatatcggg	aaaatgttga	tttcagtatg	tatgataaaa	aactgtctga	240
tattttatctt	gaaaatatct	atgaacaaga	atcatatcca	gcgagtgaga	gagataatta	300
tctgttggtt	ttaagagaag	agttaaaaaa	tattccatac	ggaaaggact	ctttgattga	360
atcatatgca	gaaaaagag	gtcatacttg	gtttgatttt	tttagaaact	tggcggtatt	420
gaaggggggg	gggttgttta	cagagacggg	taaaactgga	tgccataaca	tatctccatg	480
tgggggatgt	atatatcttg	atgcagatat	gattattact	gataaattag	gtgtcctgta	540

tgctcctgat ggtat

555

<210> 6
 <211> 990
 <212> DNA
 <213> Enteropathogenic *E. coli*

<400> 6
 atgttatctt cattaaatgt ccttcaatcc agcttcagag gaaagacagc tttatcaaat 60
 agtacacttc tccagaaagt ttcttttgct ggaaaagaat attctctgga acctattgat 120
 gaaagaaccc ctattctttt tcagtgggtt gaagcaaggc cagagcgata cgaaaaagga 180
 gaagtaccaa tattgaatac caaagaacat ccgtatttga gcaatattat aaatgctgca 240
 aaaatagaaa atgagcgtat aatcgggtgt ctggtagatg gaaattttac ttatgaacaa 300
 aaaaaggaat ttctcaatct tgaaaatgaa catcaaaata taaaaataat ctaccgagca 360
 gatgtggatt tcagcatgta tgataaaaaa ctatctgata tttaccttga aaatatccat 420
 aaacaagaat cataccctgc cagtgaagag gataattatc tgttaggctt attaagagaa 480
 gagttaaaaa atatcccaga aggtaaggac tctttgattg agtcatatgc agaaaaaaga 540
 gaacatactt ggtttgattt tttcaggaat ttggccatat tgaaggctgg aagtttgttt 600
 acagagacgg gaaaaactgg atgccataac atatcgccct gtagcggatg tatatatctt 660
 gatgccgaca tgattattac cgataaaatta ggagtcctgt atgctcctga tggtatcgct 720
 gtgcatgtag attgtaatga tgagataaaa agtcttgaaa atggtgcgat agttgtcaat 780
 cgtagtaatc atccagcatt acttgcaggc ctcgatatta tgaagagtaa agttgacgct 840
 catccatatt atgatggtct aggaaagggg atcaagcggc attttaacta ttcacgttta 900
 cacaattata atgctttttg tgattttatt gaatttaagc atgaaaatat tataccgaat 960
 accagtatgt ataccagcag ttcattggtaa 990

<210> 7
 <211> 990
 <212> DNA
 <213> Enterohemorrhagic *E. coli*

<400> 7
 atgttatctt cattaaatgt ccttcaatcc agcttcagag gaaagacagc tttatcaaat 60
 agtacacttc tccagaaagt ttcttttgct ggaaaagaat atcctctgga acctattgat 120
 gaaaaaaccc ctattctttt tcagtgggtt gaagcaaggc cagagcgata cgaaaaagga 180
 gaagtaccaa tattgaatac caaagaacat ccgtatttga gcaatattat aaatgctgca 240
 aaaatagaaa atgagcgtat aatcgggtgt ctggtagatg gaaattttac ttatgaacaa 300
 aaaaaggaat ttctcagtct tgaaaatgaa tatcaaaata taaaaataat ctaccgagca 360
 gatgtggatt tcagcatgta tgataaaaaa ctatctgata tttaccttga aaatatccat 420
 aaacaagaat cataccctgc cagtgaagag gataattatc tgttaggctt attaagagaa 480
 gagttaaaaa atatcccaga aggtaaggac tctttgattg agtcatatgc agaaaaaaga 540
 gaacatactt ggtttgattt tttcaggaat ttggccatat tgaaggctgg aagtttgttt 600
 acagagacgg gaaaaactgg atgccataac atatcgccct gtagcggatg tatatatctt 660
 gatgccgaca tgattattac cgataaaatta ggagtcctgt atgctcctga tggtatcgct 720
 gtgcatgtag attgtaatga tgagataaaa agtcttgaaa atggtgcgat agttgtcaat 780
 cgtagtaatc atccagcatt acttgcaggc ctcgatatta tgaagagtaa agttgacgct 840
 catccatatt atgatggtct aggaaagggg atcaagcggc attttaacta ttcacgttta 900
 cacgattata atgctttttg tgattttatt gaatttaagc atgaaaatat tataccgaat 960
 accagtatgt atacctgcag ttcattggtaa 990

<210> 8
 <211> 993
 <212> DNA
 <213> *Citrobacter rodentium*

<400> 8
 atgaaaattc cctcactcca gccagcttc aactttttcg ccccagcagg atactctgct 60
 gccgttgctc ccaatcgttc ggacaatgcc tatgctgatt acgtattgga tataggcaag 120
 cgaataccac tttccgcgga agatttaggc aacctatatg aaaatgtcat tcgcgccgtt 180
 cgtgacagcc gtagcaagct catagatcag catacggtcg atatgattgg taacactata 240
 cttgatgctt tgagccgatc acaaaccttt cgtgatgccg taagctatgg cattcataat 300
 aaggaggtag acattggttg cattaaatac agaaacgaat acgagctcaa cggagaatcc 360
 cccgtcaaag ttgatgatat tcaatcacta acctgtaccg aattatatga atacgatgtc 420
 gggcaagaac caattttacc catttgcgag gcaggagaaa acgataacga agagccttat 480
 gtcagtttta gtgttgccgc agatactgac tcttatgaga tgccatcgtg gcaggaaggg 540
 ctgattcacg agattattca tcatgtgact ggagctagcg atccgtctgg agatagtaat 600
 atagagctag gacccacgga gattctcgca cgtcgtgtcg ctcaagagct gggatggact 660
 gtccccgact tcataggata tgcagagcca gatcgtgaag ctcatcttag gggacgtaac 720
 ctgaatgcc ttcgacaggg gccatgcga catgaagata atgagaggac tttcttcgaa 780
 aggctgggta tgatcagtga tcgatatgag gcgagtcctg atttcacaga gtattccgct 840

gtgtctaaca	tagaatatgg	atttatccag	caacatgatt	ttccccgggt	ggctatcgac	900
gataatttac	aggatgcaaa	tcagatccaa	ctctatcatg	gagcacctta	tatctttaca	960
ttcgggggatg	tggaacaaca	caatcagcgc	tga			993

<210> 9
 <211> 993
 <212> DNA
 <213> Enteropathogenic *E. coli*

<400> 9						
atgaaaattc	cctcattaca	gtccaacttc	aactttttccg	ccccggcagg	atactctgct	60
cccattgctc	ctaactcgtgc	tgaaaatgcc	tatgcggtatt	acgtttttgga	tataggtaag	120
cgaataccac	tttccgcagc	agattttaagc	aacgtatacg	aaagtgtaat	acgcgccgtc	180
catgacagcc	gtagcaggct	tatcgatcag	catacagtcg	atatgatcgg	caacactgta	240
cttgatgctt	tgagccgac	acagacattt	cgtgatgccg	taagctatgg	cattcataat	300
gagaaggtag	acattgggtt	cattaaatac	agaaacgaat	acgagcttaa	cgaagaatct	360
tctgtcaaaa	ttgatgat	tcaatcacta	acctgtaacg	aattatatga	atatgatgtc	420
gggcaagagc	caattttccc	catttgcgaa	gcaggagaaa	acgataacga	agagccttat	480
gtcagtttta	gtgttgccgc	agatactgac	tcttatgaga	tgccatcgtg	gcaggaagga	540
ctgattcacg	agattattca	tcatgttact	ggatctagcg	atccatctgg	agatagtaat	600
atagagttag	gacccaccga	gattctcgc	cgtcgtgtcg	ctcaagaact	gggatggagt	660
gttccccgact	tcaaaggata	tgagagcca	gaacgtgaag	ctcatcttag	gttacgtaac	720
ctgaatgccc	ttcgacaggg	tgccatgagg	catgaagaga	atgagagggc	tttcttcgaa	780
aggctgggta	cgatcagtga	ccgatatgag	gcgagtcctg	atttcacaga	gtattccgct	840
gtgtctaaca	taggatacgg	atttatccag	caacatgatt	ttcctggatt	ggctatcaac	900
gataatttac	aggatgcaaa	tcagatccaa	ctgtatcatg	gcgcccctta	tatttttaca	960
tttgggggatg	tggaacaaca	caatcagcga	tga			993

<210> 10
 <211> 993
 <212> DNA
 <213> Enterohemorrhagic *E. coli*

<400> 10						
atgaaaattc	cctcattaca	gtccaacttc	aactttttccg	ccccggcagg	atactctgct	60
cccattgctc	ctaactcgtgc	tgaaaatgcc	tatgcggtatt	acgtttttgga	tataggtaag	120
cgaataccac	tttccgcagc	agattttaagc	aacgtatacg	aaagtgtaat	acgcgccgtc	180
catgacagcc	gtagcaggct	tatcgatcag	catacagtcg	atatgatcgg	caacactgta	240
cttgatgctt	tgagccgac	acagacattt	cgtgatgccg	taagctatgg	cattcataat	300
gagaaggtag	acattgggtt	cattaaatac	agaaacgaat	acgagcttaa	cgaagaatct	360
tctgtcaaaa	ttgatgat	tcaatcacta	acctgtaacg	aattatatga	atatgatgtc	420
gggcaagagc	caattttccc	catttgcgaa	gcaggagaaa	acgataacga	agagccttat	480
gtcagtttta	gtgttgccgc	agatactgac	tcttatgaga	tgccatcgtg	gcaggaagga	540
ctgattcacg	agattattca	tcatgttact	ggatctagcg	atccatctgg	agatagtaat	600
atagagttag	gacccaccga	gattctcgc	cgtcgtgtcg	ctcaagaact	gggatggagt	660
gttccccgact	tcaaaggata	tgagagcca	gaacgtgaag	ctcatcttag	gttacgtaac	720
ctgaatgccc	ttcgacaggg	tgccatgagg	catgaagaga	atgagagggc	tttcttcgaa	780
aggctgggta	cgatcagtga	ccgatatgag	gcgagtcctg	atttcacaga	gtattccgct	840
gtgtctaaca	taggatacgg	atttatccag	caacatgatt	ttcctggatt	ggctatcaac	900
gataatttac	aggatgcaaa	tcagatccaa	ctgtatcatg	gcgcccctta	tatttttaca	960
tttgggggatg	tggaacaaca	caatcagcaa	tga			993

<210> 11
 <211> 708
 <212> DNA
 <213> *Citrobacter rodentium*

<400> 11						
atgcgcccta	catcccttaa	cctgacatta	ccttcgttac	ctctaccctc	atcttcaaat	60
tcaatttcag	ccacagacat	tcaatctctt	gtaaaaatgt	cgggtgtgcg	ctgggtgaaa	120
aacaaccaac	aactctgttt	ccacgggact	gaccttaaaa	tctaccagca	tcttgaagct	180
gccctcgata	agatcgaatc	cacagacact	ggacgtactc	ttttgaactg	tattgaatta	240
acatcccga	tcaaatcaga	aaaactggca	atacatctcg	attctgctga	gttaggggtg	300
atagcacact	gcaatgcgga	tgctgaaaac	tcccagaggaa	ctggctccga	ctttcactgt	360
aatctgaatg	cagttgaata	tccctgcggg	caaggaaata	gcctggtaga	ctttcatgca	420
tgcattgttt	tccatgaact	tctccacggt	ttccacaatt	taaatggaga	gcgcctgaaa	480
gttgagagtt	ctcaaccaga	attacaaaca	cactccccac	ttttactcga	agaagccagg	540
actgttgggt	tggtgtcttt	ttctgaagaa	gttctttcag	aaaataaatt	tcgtgaagag	600
attgggatgc	cccgcagaac	attctacccg	cacgattcat	ctctcattca	tgatgacaat	660
acagtgactc	agagattcca	gcggaaaaaa	ctgcatccgt	tacttttag		708

<210> 12
 <211> 699
 <212> DNA
 <213> Enteropathogenic *E. coli*

<400> 12
 atgcgcccta cgtccctcaa cttggtatta catcagtcac caacgtcgag ctcaatgtca 60
 gatacagata tcgagtctct tgtaaaagca tcgagcgttc aatggataaa aaataatccg 120
 caacttcggt tccaggggac tgatcataat atatatcagc agattgaagc agcactcgat 180
 aagattggct ctacagagac agggcggtga ctcttgaatg ctattgaatc aatatcccg 240
 cttaaatcag aaacagtggc aatacacctc aactcttcca gactaggagt tatggcacat 300
 agagatatag atgctgagaa ccatcggggg actggttccg attttactg taatctgaat 360
 gcagttgaat atccctgtgg ggaggggatt agcgtggtgg actttcatgc gactattgtt 420
 tttcatgagt tgctccatgt tttccacaat ttaaatgggg agcgtttgaa agttgagagt 480
 tcccgaccag aatcacaaaa atactctcca cttttactcg aagaagccag gactgttggg 540
 ttgggggctt tttcagagga ggtgctttca gaaaataaat tccgcgaaga gattgggatg 600
 ccccgtagaa cctcctaccc gcacgactca gctcttattc atgatgacaa tacagtgagt 660
 ctgggattcc aacaggtaag actgcatcca ttgctttag 699

<210> 13
 <211> 699
 <212> DNA
 <213> Enterohemorrhagic *E. coli*

<400> 13
 atgcgcccta cgtccctcaa cttggtatta catcagtcac caaggtcgag ctcaatgtca 60
 gatacagata tcgagtctct tgtaaaagca tcgagcgttc aatggataaa aaataatccg 120
 caacttcggt tccaggggac tgatcataat atatatcagc agattgaagc agcactcgat 180
 aagattggct ctacagagac agggcggtga ctcttgaatg ctattgaatc aatatcccg 240
 cttaaatcag aaacagtggc aatacacctc aactcttcca gactaggagt tatggcacat 300
 agagatatag atgctgagaa ccatcggggg actggttccg attttactg taatctgaat 360
 gcagttgaat atccctgtgg ggaggggatt agcgtggtgg actttcatgc gactattgtt 420
 tttcatgagt tgctccatgt tttccacaat ttaaatgggg agcgtttgaa agttgagagt 480
 tcccgagcag aatcacaaaa atactctcca cttttactcg aagaagccag gactgttggg 540
 ttgggggctt tttcagagga ggtgctttca gaaaataaat tccacgaaga gattgggatg 600
 ccccgtagaa cctcctaccc gcrcgactca gctcttattc atgatgacaa tacagtgagt 660
 ctgggattcc aacaggtaag actgcatcca ttgctttag 699

<210> 14
 <211> 506
 <212> DNA
 <213> *Citrobacter rodentium*

<400> 14
 tacttttaatg aatcacccaa tgtatatgat aagaagtata tatctggcgt aactagagga 60
 gtagctgaac taaaacagga aggatattatt aacgagaaag ccaggcgact tgcttatatg 120
 caagcaatgt attctgtatg tccggaagag tttaaacctt tttccagaaa cgaagctagt 180
 acaccggaag gcagctggct aacagttata tccggaaaac gcccaatggg acagttttct 240
 gtagatagct tatatcatcc tgacttacat gcattgtgtg agcttccgga tatttgttgc 300
 aagatcttcc ctaaagaaaa caatgatttt ttgtatatag tgattgtgta cagaaatgac 360
 agccctctgg gagaacaacg agcaaatcga tttatagaat tatataatat aaaaagagac 420
 atcatgcagg aattaaatta tgaatctcca gagttaaagg ctgtgaaatc tgaaatgatt 480
 attgcacgtg aaatgggaga aatctt 506

<210> 15
 <211> 466
 <212> DNA
 <213> *Citrobacter rodentium*

<400> 15
 caatgtatat gataagaagt atatctctgg cgtaactaga ggagtagctg aactaaaaca 60
 ggaaggattt attaacgaga aagccaggcg acttgcttat atgcaagcaa tgtattctgt 120
 atgtccggaa gagtttaaac ctatttccag aaacgaagct agtacaccgg aaggcagctg 180
 gctaacagtt atatccggaa aacgccaat gggacagttt tctgtagata gcttatatca 240
 tcttgactta catgcattgt gtgagcttcc ggatatttgt tgcaagatct tccctaaaga 300
 aaacaatgat tttttgtata tagtgattgt gtacagaaat gacagccctc tgggagaaca 360
 acgagcaaat cgatttatag aattatataa tataaaaaga gacatcatgc aggaattaaa 420
 ttatgaatct ccagagttaa aggctgtgaa atctgaaatg attatt 466

<210> 16
 <211> 675
 <212> DNA
 <213> Enteropathogenic *E. coli*

<400> 16
 atgattaatc ctgttactaa tactcagggc gtgtccccta taaatactaa atatgctgaa 60
 catgtggtga aaaatattta cccgaaaatt aaacatgatt actttaatga atcacccaat 120
 atatatgata agaagtatat atccggtata accagaggag tagctgaact aaaacaggaa 180
 gaatttggtta acgagaaagc cagacgggtt tcttatatga agactatgta ttctgtatgt 240
 ccagaagcgt ttgaacctat ttccagaaat gaagccagta caccggaagg aagctggcta 300
 acagttatat ccggaaaacg cccaatgggg cagttttctg tagatagttt atacaatcct 360
 gattttacatg cattatgtga gcttccggac atttgttgta agatcttccc taaagaaaat 420
 aatgattttt tatacatagt tgttgtgtac agaaatgaca gccctctagg agaacaacgg 480
 gcaaatagat ttatagaatt atataatata aaaagagata tcatgcagga attaaattat 540
 gagttaccag agttaaaggc agtaaaatct gaaatgatta tcgcacgtga aatgggagaa 600
 atcttttagct acatgcctgg ggaaatagac agttatatga aatacataaa taataaactt 660
 tctaaaattg agtag 675

<210> 17
 <211> 675
 <212> DNA
 <213> Enterohemorrhagic *E. coli*

<400> 17
 atgattaatc ctgttactaa tactcagggc gtgtccccta taaatactaa atatgctgaa 60
 catgtggtga aaaatattta cccggaaaatt aaacatgatt actttaatga atcacccaat 120
 atatatgata agaagtatat atccggtata accagaggag tagctgaact aaaacaggaa 180
 gaatttggtta acgagaaagc cagacgggtt tcttatatga agactatgta ttctgtatgt 240
 ccagaagcgt ttgaacctat ttccagaaat gaagccagta caccggaagg aagctggcta 300
 acagttatat ccggaaaacg cccaatgggg cagttttctg tagatagttt atacaatcct 360
 gattttacatg cattatgtga gcttccggac atttgttgta agatcttccc taaagaaaat 420
 aatgattttt tatacatagt tgttgtgtac agaaatgaca gccctctagg agaacaacgg 480
 gcaaatagat ttatagaatt atataatata aaaagagata tcatgcagga attaaattat 540
 gagttaccag agttaaaggc agtaaaatct gaaatgatta tcgcacgtga aatgggagaa 600
 atcttttagct acatgcctgg ggaaatagac agttatatga aatacataaa taataaactt 660
 tctaaaattg agtag 675

<210> 18
 <211> 570
 <212> DNA
 <213> *Citrobacter rodentium*

<400> 18
 atgttaccaa caagtgggtc ttcagcaaatt ctttactcat ggatgtatat ctcaggaaaa 60
 gagaatcctt cgactccgga atcagtaagt gaacttaatc ataatcattt tctttctcct 120
 gaattacagg agaaactgga tgttatgttc gccatatatt catgtgccag aaacaatgat 180
 gagcgtgaga atattttacc ggagctaagg gattttgtaa gtatgcctaatt ggataagaga 240
 aacaatgtgt ttgaggtgat aaatgaagat actgatgagg tgaccggagc tctgagagcg 300
 ggaatgacga tagaggacag ggatagttat atcagggatc ttttttttct gcattcattg 360
 aaagtaaaaa ttgaggaaag cagacaagat aaagaggatt ggaaatgtaa agtttatgat 420
 ctgctatgtc cgcattcattc ttcagagcta tatggggatc tacgggcaat caaatgcctc 480
 gttgaaggat gcagtgatga ttttagtcct tttgatacta ttaaggtgcc ggatcttact 540
 tacaacaaag gatctttaca atgtggatga 570

<210> 19
 <211> 519
 <212> DNA
 <213> *Citrobacter rodentium*

<400> 19
 agcaaatctt tactcatgga tgtatatctc aggaaaagag aatccttcga ctccggaatc 60
 agtaagtga cttaatcata atcattttct tctcctgaa ttacaggaga aactggatgt 120
 tatgttcgcc atatattcat gtgccagaaa caatgatgag cgtgagaata tttaccggga 180
 gctaagggat tttgtaagta gcctaattgga taagagaaac aatgtgtttg aggtgataaa 240
 tgaagatact gatgaggtga ccggagctct gagagcggga atgacgatag aggacaggga 300
 tagttatata agggatcttt tttttctgca ttcatgaaa gtaaaaattg aggaaagcag 360
 acaagataaa gaggattgga aatgtaaagt ttatgatctg ctatgtccgc atcattcttc 420
 agagctatat ggggatctac gggcaatcaa atgcctcgtt gaaggatgca gtgatgattt 480
 tagtccctttt gatactatta aggtgccgga tcttactta 519

<210> 20
<211> 570
<212> DNA
<213> Enteropathogenic E. coli

<400> 20
atgttaccaa caagtgggtc ttcagcaa atttattcat ggatgtatgt atcaggaaga 60
ggtaaccctt cgactccgga atcagtaagt gagcttaatc ataatcactt tctttctcct 120
gaattacaag ataaacttga tgttatggtc tctatatatt catgtgccag aaataataat 180
gagcttgagg aaatttttca agagctaagt gcttttgtaa gtgggctgat ggataagaga 240
aatagtgtat ttgaggtgag aaatgaaaat actgatgagg ttgtcggagc gctgagggcg 300
ggaatgacga tagaggatag ggatagttat atcagggatc ttttttttct gcattcattg 360
aaagtaaaaa ttgaggaaaag tagacaaggc aaagaagatt cgaaatgtaa agtttataat 420
ctgctatgtc cgcactcactc ttcagagcta tatggtgatc tacgagcaat gaaatgcctc 480
gtggaaggat gcagtgatga ttttaatcct tttgatatta ttaggggtacc agatcttact 540
tacaacaaag gatctttaca atgtggatga 570

<210> 21
<211> 570
<212> DNA
<213> Enterohemorrhagic E. coli

<400> 21
atgttaccaa caagtgggtc ttcagcaa atttattcat ggatgtatgt atcaggaaga 60
ggtaaccctt cgactccgga atcagtaagt gagcttaatc ataatcactt tctttctcct 120
gaattacaag ataaacttga tgttatggtc tctatatatt catgtgccag aaataataat 180
gagcttgagg aaatttttca agagctaagt gcttttgtaa gtgggctgat ggataagaga 240
aatagtgtat ttgaggtgag aaatgaaaat actgatgagg ttgtcggagc gctgagggcg 300
ggaatgacga tagaggacag ggatagttat atcagggatc ttttttttct gcattcattg 360
aaagtaaaaa ttgaggaaaag tagacaaggc aaagaagatt cgaaatgtaa agtttataat 420
ctgctatgtc cgcactcactc ttcagagcta tatggtgatc tacgagcaat gaaatgcctc 480
gtggaaggat gcagtgatga ttttaatcct tttgatatta ttaggggtacc agatcttact 540
tacaacaaag gatctttaca atgtggatga 570

<210> 22
<211> 430
<212> PRT
<213> Citrobacter rodentium

<400> 22

Met Asn Ile Gln Pro Asn Ile His Ser Gly Ile Thr Thr Gln Asn Asn
1 5 10 15
Gln Gln His His His Ala Glu Gln Val Pro Val Ser Ser Ser Ile Pro
20 25 30
Arg Ser Asp Leu Pro Pro Asn Cys Glu Ala Gly Phe Val Val His Ile
35 40 45
Pro Glu Asp Ile Gln Gln His Val Pro Glu Cys Gly Glu Thr Thr Ala
50 55 60
Leu Leu Ser Leu Ile Lys Asp Glu Gly Leu Leu Ser Gly Leu Asp Lys
65 70 75 80
Tyr Leu Ala Pro His Leu Glu Glu Gly Ser Leu Gly Lys Lys Ala Leu
85 90 95
Asp Thr Phe Gly Leu Phe Asn Val Thr Gln Met Ala Leu Glu Ile Pro
100 105 110
Ser Ser Val Pro Gly Ile Ser Gly Lys Tyr Gly Val Gln Met Asn Ile
115 120 125
Val Lys Pro Asp Ile His Pro Thr Thr Gly Asn Tyr Phe Leu Gln Leu
130 135 140

Phe Pro Leu His Asp Glu Ile Gly Phe Asn Phe Lys Asp Leu Pro Gly
 145 150 155 160
 Pro Leu Lys Asn Ala Leu Thr Asn Ser Ser Ile Ser Ala Thr Ala Ser
 165 170 175
 Thr Val Ala Pro Thr Pro Asn Asp Pro Met Pro Trp Phe Gly Leu Thr
 180 185 190
 Ala Gln Val Val Arg Asn His Gly Val Glu Leu Pro Ile Val Lys Thr
 195 200 205
 Glu Asn Gly Trp Lys Leu Val Gly Glu Thr Pro Leu Thr Pro Asp Gly
 210 215 220
 Pro Lys Ala Asn Tyr Thr Glu Glu Trp Val Ile Arg Pro Gly Glu Ala
 225 230 235 240
 Asp Phe Lys Tyr Gly Thr Ser Pro Leu Gln Ala Thr Leu Gly Leu Glu
 245 250 255
 Phe Gly Ala His Phe Lys Trp Asp Leu Asp Asn Pro Asn Thr Lys Tyr
 260 265 270
 Ala Ile Leu Thr Asn Ala Ala Ala Asn Ala Ile Gly Ala Ala Gly Gly
 275 280 285
 Phe Ala Val Ser Lys Val Pro Gly Ile Asp Pro Met Leu Ser Pro His
 290 295 300
 Val Gly Ala Met Leu Gly Gln Ala Ala Gly His Ala Val Gln Cys Asn
 305 310 315 320
 Thr Pro Gly Leu Lys Pro Asp Thr Ile Leu Trp Trp Ala Gly Ala Thr
 325 330 335
 Phe Gly Ala Ala Asp Leu Asn Lys Ala Glu Phe Asp Lys Val Arg Phe
 340 345 350
 Thr Asp Tyr Pro Arg Ile Trp Phe His Ala Arg Glu Gly Ala Leu Phe
 355 360 365
 Pro Asn Lys Gln Asp Ile Ala Arg Val Thr Gly Ala Asp Ile Lys Ala
 370 375 380
 Met Glu Glu Gly Val Pro Val Gly His Gln His Pro Lys Pro Glu Asp
 385 390 395 400
 Val Val Ile Asp Ile Glu Gly Gly Asn Ser Pro His His Asn Pro Ser
 405 410 415
 Asn Tyr Val Asp Thr Phe Glu Ile Ile Gln Glu Thr Arg Val
 420 425 430

 <210> 23
 <211> 440
 <212> PRT
 <213> Enteropathogenic E. coli

 <400> 23

 Met Asn Ile Gln Pro Ile Val Thr Ser Gly Ile Thr Thr Gln Asn Asn
 1 5 10 15
 Arg His His His Ala Glu Gln Thr Ser Pro Thr Gln Ile Pro Gln Ser
 20 25 30
 Glu Leu Pro Asn Gly Cys Glu Thr Gly Phe Val Val His Ile Pro Glu
 35 40 45

Asp	Met	Gln	Arg	His	Ala	Pro	Glu	Cys	Gly	Glu	Thr	Thr	Ala	Leu	Leu	50	55	60	
Ser	Leu	Ile	Lys	Asp	Glu	Gly	Leu	Leu	Ser	Gly	Leu	Asp	Lys	Tyr	Leu	65	70	75	80
Ala	Pro	His	Leu	Glu	Glu	Gly	Ser	Ala	Gly	Lys	Lys	Ala	Leu	Asp	Met	85	90	95	
Phe	Gly	Leu	Phe	Asn	Val	Ser	Gln	Met	Ala	Leu	Glu	Ile	Pro	Ser	Thr	100	105	110	
Val	Pro	Gly	Ile	Ser	Gly	Lys	Tyr	Gly	Val	Gln	Leu	Asn	Ile	Val	Lys	115	120	125	
Pro	Asp	Ile	His	Pro	Thr	Ser	Gly	Asn	Tyr	Phe	Leu	Gln	Ile	Phe	Pro	130	135	140	
Leu	His	Asp	Glu	Ile	Gly	Ile	Asn	Phe	Lys	Asp	Leu	Pro	Gly	Pro	Leu	145	150	155	160
Lys	Asn	Ala	Leu	Ser	Asn	Ser	Asn	Ile	Pro	Thr	Thr	Val	Ser	Thr	Ala	165	170	175	
Ala	Ser	Thr	Ile	Ala	Ser	Ala	Thr	Thr	Ser	Thr	Val	Thr	Thr	Ala	Ser	180	185	190	
Lys	Asp	Pro	Ile	Pro	Trp	Phe	Gly	Leu	Thr	Ala	Gln	Val	Val	Arg	Asn	195	200	205	
His	Gly	Val	Glu	Leu	Pro	Ile	Val	Lys	Thr	Glu	Asn	Gly	Trp	Lys	Leu	210	215	220	
Val	Gly	Glu	Thr	Pro	Leu	Thr	Pro	Asp	Gly	Pro	Lys	Ala	Asn	Tyr	Thr	225	230	235	240
Glu	Glu	Trp	Val	Ile	Arg	Pro	Gly	Glu	Ala	Asp	Phe	Lys	Tyr	Gly	Ala	245	250	255	
Ser	Pro	Leu	Gln	Ala	Thr	Leu	Gly	Leu	Glu	Phe	Gly	Ala	His	Phe	Lys	260	265	270	
Trp	Asp	Leu	Asp	Asn	Pro	Asn	Thr	Lys	Tyr	Ala	Val	Leu	Thr	Asn	Ala	275	280	285	
Ala	Ala	Asn	Ala	Leu	Gly	Ala	Val	Gly	Gly	Phe	Ala	Val	Ser	Arg	Phe	290	295	300	
Thr	Gly	Thr	Asp	Pro	Met	Leu	Ser	Pro	His	Ile	Gly	Ala	Met	Val	Gly	305	310	315	320
Gln	Ala	Ala	Gly	His	Ala	Ile	Gln	Tyr	Asn	Thr	Pro	Gly	Leu	Lys	Pro	325	330	335	
Asp	Thr	Ile	Leu	Trp	Trp	Ala	Gly	Thr	Thr	Leu	Gly	Leu	Ala	Asp	Leu	340	345	350	
Asn	Lys	Ala	Glu	Phe	Gly	Glu	Ala	Arg	Phe	Thr	Asp	Tyr	Pro	Arg	Ile	355	360	365	
Trp	Trp	His	Ala	Arg	Glu	Gly	Ala	Ile	Phe	Pro	Asn	Lys	Ala	Asp	Ile	370	375	380	
Glu	His	Ala	Thr	Gly	Ala	Asp	Ile	Arg	Ala	Met	Glu	Glu	Gly	Val	Ser	385	390	395	400
Val	Gly	Gln	Arg	His	Pro	Asn	Pro	Glu	Asp	Val	Val	Ile	Asn	Ile	Glu	405	410	415	

Ser Asn Asn Ser Pro His His Asn Pro Ser Asn Tyr Val Asp Thr Val
420 425 430

Asp Ile Ile Gln Glu Thr Arg Val
435 440

<210> 24
<211> 441
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 24

Met Asn Ile Gln Pro Thr Ile Gln Ser Gly Ile Thr Ser Gln Asn Asn
1 5 10 15

Gln His His Gln Thr Glu Gln Ile Pro Ser Thr Gln Ile Pro Gln Ser
20 25 30

Glu Leu Pro Leu Gly Cys Gln Ala Gly Phe Val Val Asn Ile Pro Asp
35 40 45

Asp Ile Gln Gln His Ala Pro Glu Cys Gly Glu Thr Thr Ala Leu Leu
50 55 60

Ser Leu Ile Lys Asp Lys Gly Leu Leu Ser Gly Leu Asp Glu Tyr Ile
65 70 75 80

Ala Pro His Leu Glu Glu Gly Ser Ile Gly Lys Lys Thr Leu Asp Met
85 90 95

Phe Gly Leu Phe Asn Val Thr Gln Met Ala Leu Glu Ile Pro Ser Ser
100 105 110

Val Ser Gly Ile Ser Gly Lys Tyr Gly Val Gln Leu Asn Ile Val Lys
115 120 125

Pro Asp Ile His Pro Thr Ser Gly Asn Tyr Phe Leu Gln Ile Phe Pro
130 135 140

Leu His Asp Glu Ile Gly Phe Asn Phe Lys Asp Leu Pro Gly Pro Leu
145 150 155 160

Lys Asn Ala Leu Ser Asn Ser Asn Ile Ser Thr Thr Ala Val Ser Thr
165 170 175

Ile Ala Ser Thr Gly Thr Ser Ala Thr Thr Ser Thr Val Thr Thr Glu
180 185 190

Pro Lys Asp Pro Ile Pro Trp Phe Gly Leu Thr Ala Gln Val Val Arg
195 200 205

Asn His Gly Val Glu Leu Pro Ile Val Lys Thr Glu Asn Gly Trp Lys
210 215 220

Leu Val Gly Glu Thr Pro Leu Thr Pro Asp Gly Pro Lys Ala Asn Tyr
225 230 235 240

Thr Glu Glu Trp Val Ile Arg Pro Gly Glu Ala Asp Phe Lys Tyr Gly
245 250 255

Ala Ser Pro Leu Gln Ala Thr Leu Gly Leu Glu Phe Gly Ala His Phe
260 265 270

Lys Trp Asp Leu Asp Asn Pro Asn Thr Lys Tyr Ala Val Leu Thr Asn
275 280 285

Ala Ala Ala Asn Ala Leu Gly Ala Leu Gly Gly Phe Ala Val Ser Arg
290 295 300

108
157

Phe Ala Ser Thr Asp Pro Met Leu Ser Pro His Ile Gly Ala Met Val
305 310 315 320

Gly Gln Ala Ala Gly His Ala Ile Gln Tyr Asn Thr Pro Gly Leu Lys
325 330 335

Pro Asp Thr Ile Leu Trp Trp Ala Gly Ala Thr Leu Gly Ala Ala Asp
340 345 350

Leu Asn Lys Ala Glu Phe Glu Val Ala Arg Phe Thr Asp Tyr Pro Arg
355 360 365

Ile Trp Trp His Ala Arg Glu Gly Ala Ile Phe Pro Asn Lys Ala Asp
370 375 380

Ile Glu His Ala Thr Gly Ala Asp Ile Arg Ala Met Glu Glu Gly Ile
385 390 395 400

Pro Val Gly Gln Arg His Pro Asn Pro Glu Asp Val Val Ile Asp Ile
405 410 415

Glu Ser Asn Gly Leu Pro His His Asn Pro Ser Asn His Val Asp Ile
420 425 430

Phe Asp Ile Ile Gln Glu Thr Arg Val
435 440

<210> 25
<211> 204
<212> PRT
<213> Citrobacter rodentium

<400> 25

Ile Leu Phe Gln Trp Phe Glu Ala Arg Pro Glu Arg Tyr Gly Lys Gly
1 5 10 15

Glu Val Pro Ile Leu Asn Thr Lys Glu His Pro Tyr Leu Ser Asn Ile
20 25 30

Ile Asn Ala Ala Lys Ile Glu Asn Glu Arg Val Ile Gly Val Leu Val
35 40 45

Asp Gly Asp Phe Thr Tyr Glu Gln Arg Lys Glu Phe Leu Ser Leu Glu
50 55 60

Asp Glu His Gln Asn Ile Lys Ile Ile Tyr Arg Glu Asn Val Asp Phe
65 70 75 80

Ser Met Tyr Asp Lys Lys Leu Ser Asp Ile Tyr Leu Glu Asn Ile His
85 90 95

Glu Gln Glu Ser Tyr Pro Ala Ser Glu Arg Asp Asn Tyr Leu Leu Gly
100 105 110

Leu Leu Arg Glu Glu Leu Lys Asn Ile Pro Tyr Gly Lys Asp Ser Leu
115 120 125

Ile Glu Ser Tyr Ala Glu Lys Arg Gly His Thr Trp Phe Asp Phe Phe
130 135 140

Arg Asn Leu Ala Val Leu Lys Gly Gly Gly Leu Phe Thr Glu Thr Gly
145 150 155 160

Lys Thr Gly Cys His Asn Ile Ser Pro Cys Gly Gly Cys Ile Tyr Leu
165 170 175

Asp Ala Asp Met Ile Ile Thr Asp Lys Leu Gly Val Leu Tyr Ala Pro
180 185 190

Asp Gly Ile Ala Val His Val Asp Cys Asn Asp Glu
195 200

<210> 26
<211> 186
<212> PRT
<213> Citrobacter rodentium

<400> 26

Arg Pro Glu Arg Tyr Gly Lys Gly Glu Val Pro Ile Leu Asn Thr Lys
1 5 10 15
Glu His Pro Tyr Leu Ser Asn Ile Ile Asn Ala Ala Lys Ile Glu Asn
20 25 30
Glu Arg Val Ile Gly Val Leu Val Asp Gly Asp Phe Thr Tyr Glu Gln
35 40 45
Arg Lys Glu Phe Leu Ser Leu Glu Asp Glu His Gln Asn Ile Lys Ile
50 55 60
Ile Tyr Arg Glu Asn Val Asp Phe Ser Met Tyr Asp Lys Lys Leu Ser
65 70 75 80
Asp Ile Tyr Leu Glu Asn Ile His Glu Gln Glu Ser Tyr Pro Ala Ser
85 90 95
Glu Arg Asp Asn Tyr Leu Leu Gly Leu Leu Arg Glu Glu Leu Lys Asn
100 105 110
Ile Pro Tyr Gly Lys Asp Ser Leu Ile Glu Ser Tyr Ala Glu Lys Arg
115 120 125
Gly His Thr Trp Phe Asp Phe Phe Arg Asn Leu Ala Val Leu Lys Gly
130 135 140
Gly Gly Leu Phe Thr Glu Thr Gly Lys Thr Gly Cys His Asn Ile Ser
145 150 155 160
Pro Cys Gly Gly Cys Ile Tyr Leu Asp Ala Asp Met Ile Ile Thr Asp
165 170 175
Lys Leu Gly Val Leu Tyr Ala Pro Asp Gly
180 185

<210> 27
<211> 329
<212> PRT
<213> Enteropathogenic E. coli

<400> 27

Met Leu Ser Ser Leu Asn Val Leu Gln Ser Ser Phe Arg Gly Lys Thr
1 5 10 15
Ala Leu Ser Asn Ser Thr Leu Leu Gln Lys Val Ser Phe Ala Gly Lys
20 25 30
Glu Tyr Ser Leu Glu Pro Ile Asp Glu Arg Thr Pro Ile Leu Phe Gln
35 40 45
Trp Phe Glu Ala Arg Pro Glu Arg Tyr Glu Lys Gly Glu Val Pro Ile
50 55 60
Leu Asn Thr Lys Glu His Pro Tyr Leu Ser Asn Ile Ile Asn Ala Ala
65 70 75 80
Lys Ile Glu Asn Glu Arg Ile Ile Gly Val Leu Val Asp Gly Asn Phe
85 90 95

Thr Tyr Glu Gln Lys Lys Glu Phe Leu Asn Leu Glu Asn Glu His Gln
100 105 110

Asn Ile Lys Ile Ile Tyr Arg Ala Asp Val Asp Phe Ser Met Tyr Asp
115 120 125

Lys Lys Leu Ser Asp Ile Tyr Leu Glu Asn Ile His Lys Gln Glu Ser
130 135 140

Tyr Pro Ala Ser Glu Arg Asp Asn Tyr Leu Leu Gly Leu Leu Arg Glu
145 150 155 160

Glu Leu Lys Asn Ile Pro Glu Gly Lys Asp Ser Leu Ile Glu Ser Tyr
165 170 175

Ala Glu Lys Arg Glu His Thr Trp Phe Asp Phe Phe Arg Asn Leu Ala
180 185 190

Ile Leu Lys Ala Gly Ser Leu Phe Thr Glu Thr Gly Lys Thr Gly Cys
195 200 205

His Asn Ile Ser Pro Cys Ser Gly Cys Ile Tyr Leu Asp Ala Asp Met
210 215 220

Ile Ile Thr Asp Lys Leu Gly Val Leu Tyr Ala Pro Asp Gly Ile Ala
225 230 235 240

Val His Val Asp Cys Asn Asp Glu Ile Lys Ser Leu Glu Asn Gly Ala
245 250 255

Ile Val Val Asn Arg Ser Asn His Pro Ala Leu Leu Ala Gly Leu Asp
260 265 270

Ile Met Lys Ser Lys Val Asp Ala His Pro Tyr Tyr Asp Gly Leu Gly
275 280 285

Lys Gly Ile Lys Arg His Phe Asn Tyr Ser Ser Leu His Asn Tyr Asn
290 295 300

Ala Phe Cys Asp Phe Ile Glu Phe Lys His Glu Asn Ile Ile Pro Asn
305 310 315 320

Thr Ser Met Tyr Thr Ser Ser Ser Trp
325

<210> 28
<211> 329
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 28

Met Leu Ser Ser Leu Asn Val Leu Gln Ser Ser Phe Arg Gly Lys Thr
1 5 10 15

Ala Leu Ser Asn Ser Thr Leu Leu Gln Lys Val Ser Phe Ala Gly Lys
20 25 30

Glu Tyr Pro Leu Glu Pro Ile Asp Glu Lys Thr Pro Ile Leu Phe Gln
35 40 45

Trp Phe Glu Ala Arg Pro Glu Arg Tyr Glu Lys Gly Glu Val Pro Ile
50 55 60

Leu Asn Thr Lys Glu His Pro Tyr Leu Ser Asn Ile Ile Asn Ala Ala
65 70 75 80

Lys Ile Glu Asn Glu Arg Ile Ile Gly Val Leu Val Asp Gly Asn Phe
85 90 95

Thr Tyr Glu Gln Lys Lys Glu Phe Leu Ser Leu Glu Asn Glu Tyr Gln
100 105 110
Asn Ile Lys Ile Ile Tyr Arg Ala Asp Val Asp Phe Ser Met Tyr Asp
115 120 125
Lys Lys Leu Ser Asp Ile Tyr Leu Glu Asn Ile His Lys Gln Glu Ser
130 135 140
Tyr Pro Ala Ser Glu Arg Asp Asn Tyr Leu Leu Gly Leu Leu Arg Glu
145 150 155 160
Glu Leu Lys Asn Ile Pro Glu Gly Lys Asp Ser Leu Ile Glu Ser Tyr
165 170 175
Ala Glu Lys Arg Glu His Thr Trp Phe Asp Phe Phe Arg Asn Leu Ala
180 185 190
Met Leu Lys Ala Gly Ser Leu Phe Thr Glu Thr Gly Lys Thr Gly Cys
195 200 205
His Asn Ile Ser Pro Cys Ser Gly Cys Ile Tyr Leu Asp Ala Asp Met
210 215 220
Ile Ile Thr Asp Lys Leu Gly Val Leu Tyr Ala Pro Asp Gly Ile Ala
225 230 235 240
Val His Val Asp Cys Asn Asp Glu Ile Lys Ser Leu Glu Asn Gly Ala
245 250 255
Ile Val Val Asn Arg Ser Asn His Pro Ala Leu Leu Ala Gly Leu Asp
260 265 270
Ile Met Lys Ser Lys Val Asp Ala His Pro Tyr Tyr Asp Gly Leu Gly
275 280 285
Lys Gly Ile Lys Arg His Phe Asn Tyr Ser Ser Leu His Asp Tyr Asn
290 295 300
Ala Phe Cys Asp Phe Ile Glu Phe Lys His Glu Asn Ile Ile Pro Asn
305 310 315 320
Thr Ser Met Tyr Thr Cys Ser Ser Trp
325

<210> 29
<211> 326
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 29

Met Leu Ser Pro Ile Arg Thr Thr Phe His Asn Ser Val Asn Ile Val
1 5 10 15
Gln Ser Ser Pro Cys Gln Thr Val Ser Phe Ala Gly Lys Glu Tyr Glu
20 25 30
Leu Lys Val Ile Asp Glu Lys Thr Pro Ile Leu Phe Gln Trp Phe Glu
35 40 45
Pro Asn Pro Glu Arg Tyr Lys Lys Asp Glu Val Pro Ile Val Asn Thr
50 55 60
Lys Gln His Pro Tyr Leu Asp Asn Val Thr Asn Ala Ala Arg Ile Glu
65 70 75 80
Ser Asp Arg Met Ile Gly Ile Phe Val Asp Gly Asp Phe Ser Val Asn
85 90 95

Gln Lys Thr Ala Phe Ser Lys Leu Glu Arg Asp Phe Glu Asn Val Met
100 105 110

Ile Ile Tyr Arg Glu Asp Val Asp Phe Ser Met Tyr Asp Arg Lys Leu
115 120 125

Ser Asp Ile Tyr His Asp Ile Ile Cys Glu Gln Arg Leu Arg Thr Glu
130 135 140

Asp Lys Arg Asp Glu Tyr Leu Leu Asn Leu Leu Glu Lys Glu Leu Arg
145 150 155 160

Glu Ile Ser Lys Ala Gln Asp Ser Leu Ile Ser Met Tyr Ala Lys Lys
165 170 175

Arg Asn His Ala Trp Phe Asp Phe Phe Arg Asn Leu Ala Leu Leu Lys
180 185 190

Ala Gly Glu Ile Phe Arg Cys Thr Tyr Asn Thr Lys Asn His Gly Ile
195 200 205

Ser Phe Gly Glu Gly Cys Ile Tyr Leu Asp Met Asp Met Ile Leu Thr
210 215 220

Gly Lys Leu Gly Thr Ile Tyr Ala Pro Asp Gly Ile Ser Met His Val
225 230 235 240

Asp Arg Arg Asn Asp Ser Val Asn Ile Glu Asn Ser Ala Ile Ile Val
245 250 255

Asn Arg Ser Asn His Pro Ala Leu Leu Glu Gly Leu Ser Phe Met His
260 265 270

Ser Lys Val Asp Ala His Pro Tyr Tyr Asp Gly Leu Gly Lys Gly Val
275 280 285

Lys Lys Tyr Phe Asn Phe Thr Pro Leu His Asn Tyr Asn His Phe Cys
290 295 300

Asp Phe Ile Glu Phe Asn His Pro Asn Ile Ile Met Asn Thr Ser Gln
305 310 315 320

Tyr Thr Cys Ser Ser Trp
325

<210> 30
<211> 330
<212> PRT
<213> Citrobacter rodentium

<400> 30
Met Lys Ile Pro Ser Leu Gln Pro Ser Phe Asn Phe Phe Ala Pro Ala
1 5 10 15

Gly Tyr Ser Ala Ala Val Ala Pro Asn Arg Ser Asp Asn Ala Tyr Ala
20 25 30

Asp Tyr Val Leu Asp Ile Gly Lys Arg Ile Pro Leu Ser Ala Glu Asp
35 40 45

Leu Gly Asn Leu Tyr Glu Asn Val Ile Arg Ala Val Arg Asp Ser Arg
50 55 60

Ser Lys Leu Ile Asp Gln His Thr Val Asp Met Ile Gly Asn Thr Ile
65 70 75 80

Leu Asp Ala Leu Ser Arg Ser Gln Thr Phe Arg Asp Ala Val Ser Tyr
85 90 95

167
162

Gly Ile His Asn Lys Glu Val His Ile Gly Cys Ile Lys Tyr Arg Asn
100 105 110
Glu Tyr Glu Leu Asn Gly Glu Ser Pro Val Lys Val Asp Asp Ile Gln
115 120 125
Ser Leu Thr Cys Thr Glu Leu Tyr Glu Tyr Asp Val Gly Gln Glu Pro
130 135 140
Ile Leu Pro Ile Cys Glu Ala Gly Glu Asn Asp Asn Glu Glu Pro Tyr
145 150 155 160
Val Ser Phe Ser Val Ala Pro Asp Thr Asp Ser Tyr Glu Met Pro Ser
165 170 175
Trp Gln Glu Gly Leu Ile His Glu Ile Ile His His Val Thr Gly Ala
180 185 190
Ser Asp Pro Ser Gly Asp Ser Asn Ile Glu Leu Gly Pro Thr Glu Ile
195 200 205
Leu Ala Arg Arg Val Ala Gln Glu Leu Gly Trp Thr Val Pro Asp Phe
210 215 220
Ile Gly Tyr Ala Glu Pro Asp Arg Glu Ala His Leu Arg Gly Arg Asn
225 230 235 240
Leu Asn Ala Leu Arg Gln Ala Ala Met Arg His Glu Asp Asn Glu Arg
245 250 255
Thr Phe Phe Glu Arg Leu Gly Met Ile Ser Asp Arg Tyr Glu Ala Ser
260 265 270
Pro Asp Phe Thr Glu Tyr Ser Ala Val Ser Asn Ile Glu Tyr Gly Phe
275 280 285
Ile Gln Gln His Asp Phe Pro Gly Leu Ala Ile Asp Asp Asn Leu Gln
290 295 300
Asp Ala Asn Gln Ile Gln Leu Tyr His Gly Ala Pro Tyr Ile Phe Thr
305 310 315 320
Phe Gly Asp Val Asp Lys His Asn Gln Arg
325 330

<210> 31
<211> 330
<212> PRT
<213> Enteropathogenic E. coli

<400> 31

Met Lys Ile Pro Ser Leu Gln Ser Asn Phe Asn Phe Ser Ala Pro Ala
1 5 10 15
Gly Tyr Ser Ala Pro Ile Ala Pro Asn Arg Ala Glu Asn Ala Tyr Ala
20 25 30
Asp Tyr Val Leu Asp Ile Gly Lys Arg Ile Pro Leu Ser Ala Ala Asp
35 40 45
Leu Ser Asn Val Tyr Glu Ser Val Ile Arg Ala Val His Asp Ser Arg
50 55 60
Ser Arg Leu Ile Asp Gln His Thr Val Asp Met Ile Gly Asn Thr Val
65 70 75 80
Leu Asp Ala Leu Ser Arg Ser Gln Thr Phe Arg Asp Ala Val Ser Tyr
85 90 95

Gly Ile His Asn Glu Lys Val His Ile Gly Cys Ile Lys Tyr Arg Asn
100 105 110
Glu Tyr Glu Leu Asn Glu Glu Ser Ser Val Lys Ile Asp Asp Ile Gln
115 120 125
Ser Leu Thr Cys Asn Glu Leu Tyr Glu Tyr Asp Val Gly Gln Glu Pro
130 135 140
Ile Phe Pro Ile Cys Glu Ala Gly Glu Asn Asp Asn Glu Glu Pro Tyr
145 150 155 160
Val Ser Phe Ser Val Ala Pro Asp Thr Asp Ser Tyr Glu Met Pro Ser
165 170 175
Trp Gln Glu Gly Leu Ile His Glu Ile Ile His His Val Thr Gly Ser
180 185 190
Ser Asp Pro Ser Gly Asp Ser Asn Ile Glu Leu Gly Pro Thr Glu Ile
195 200 205
Leu Ala Arg Arg Val Ala Gln Glu Leu Gly Trp Ser Val Pro Asp Phe
210 215 220
Lys Gly Tyr Ala Glu Pro Glu Arg Glu Ala His Leu Arg Leu Arg Asn
225 230 235 240
Leu Asn Ala Leu Arg Gln Ala Ala Met Arg His Glu Glu Asn Glu Arg
245 250 255
Ala Phe Phe Glu Arg Leu Gly Thr Ile Ser Asp Arg Tyr Glu Ala Ser
260 265 270
Pro Asp Phe Thr Glu Tyr Ser Ala Val Ser Asn Ile Gly Tyr Gly Phe
275 280 285
Ile Gln Gln His Asp Phe Pro Gly Leu Ala Ile Asn Asp Asn Leu Gln
290 295 300
Asp Ala Asn Gln Ile Gln Leu Tyr His Gly Ala Pro Tyr Ile Phe Thr
305 310 315 320
Phe Gly Asp Val Asp Lys His Asn Gln Arg
325 330

<210> 32
<211> 330
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 32

Met Lys Ile Pro Ser Leu Gln Ser Asn Phe Asn Phe Ser Ala Pro Ala
1 5 10 15
Gly Tyr Ser Ala Pro Ile Ala Pro Asn Arg Ala Glu Asn Ala Tyr Ala
20 25 30
Asp Tyr Val Leu Asp Ile Gly Lys Arg Ile Pro Leu Ser Ala Ala Asp
35 40 45
Leu Ser Asn Val Tyr Glu Ser Val Ile Arg Ala Val His Asp Ser Arg
50 55 60
Ser Arg Leu Ile Asp Gln His Thr Val Asp Met Ile Gly Asn Thr Val
65 70 75 80
Leu Asp Ala Leu Ser Arg Ser Gln Thr Phe Arg Asp Ala Val Ser Tyr
85 90 95

Gly Ile His Asn Glu Lys Val His Ile Gly Cys Ile Lys Tyr Arg Asn
100 105 110

Glu Tyr Glu Leu Asn Glu Glu Ser Ser Val Lys Ile Asp Asp Ile Gln
115 120 125

Ser Leu Thr Cys Asn Glu Leu Tyr Glu Tyr Asp Val Gly Gln Glu Pro
130 135 140

Ile Phe Pro Ile Cys Glu Ala Gly Glu Asn Asp Asn Glu Glu Pro Tyr
145 150 155 160

Val Ser Phe Ser Val Ala Pro Asp Thr Asp Ser Tyr Glu Met Pro Ser
165 170 175

Trp Gln Glu Gly Leu Ile His Glu Ile Ile His His Val Thr Gly Ser
180 185 190

Ser Asp Pro Ser Gly Asp Ser Asn Ile Glu Leu Gly Pro Thr Glu Ile
195 200 205

Leu Ala Arg Arg Val Ala Gln Glu Leu Gly Trp Ser Val Pro Asp Phe
210 215 220

Lys Gly Tyr Ala Glu Pro Glu Arg Glu Ala His Leu Arg Leu Arg Asn
225 230 235 240

Leu Asn Ala Leu Arg Gln Ala Ala Met Arg His Glu Glu Asn Glu Arg
245 250 255

Ala Phe Phe Glu Arg Leu Gly Thr Ile Ser Asp Arg Tyr Glu Ala Ser
260 265 270

Pro Asp Phe Thr Glu Tyr Ser Ala Val Ser Asn Ile Gly Tyr Gly Phe
275 280 285

Ile Gln Gln His Asp Phe Pro Gly Leu Ala Ile Asn Asp Asn Leu Gln
290 295 300

Asp Ala Asn Gln Ile Gln Leu Tyr His Gly Ala Pro Tyr Ile Phe Thr
305 310 315 320

Phe Gly Asp Val Asp Lys His Asn Gln Gln
325 330

<210> 33
<211> 235
<212> PRT
<213> Citrobacter rodentium

<400> 33

Met Arg Pro Thr Ser Leu Asn Leu Thr Leu Pro Ser Leu Pro Leu Pro
1 5 10 15

Ser Ser Ser Asn Ser Ile Ser Ala Thr Asp Ile Gln Ser Leu Val Lys
20 25 30

Met Ser Gly Val Arg Trp Val Lys Asn Asn Gln Gln Leu Cys Phe His
35 40 45

Gly Thr Asp Leu Lys Ile Tyr Gln His Leu Glu Ala Ala Leu Asp Lys
50 55 60

Ile Glu Ser Thr Asp Thr Gly Arg Thr Leu Leu Asn Cys Ile Glu Leu
65 70 75 80

Thr Ser Arg Leu Lys Ser Glu Lys Leu Ala Ile His Leu Asp Ser Ala
85 90 95

165

Glu Leu Gly Val Ile Ala His Cys Asn Ala Asp Ala Glu Asn Ser Arg
100 105 110

Gly Thr Gly Ser Asp Phe His Cys Asn Leu Asn Ala Val Glu Tyr Pro
115 120 125

Cys Gly Gln Gly Ile Ser Leu Val Asp Phe His Ala Cys Ile Val Phe
130 135 140

His Glu Leu Leu His Val Phe His Asn Leu Asn Gly Glu Arg Leu Lys
145 150 155 160

Val Glu Ser Ser Gln Pro Glu Leu Gln Thr His Ser Pro Leu Leu Leu
165 170 175

Glu Glu Ala Arg Thr Val Gly Leu Gly Ala Phe Ser Glu Glu Val Leu
180 185 190

Ser Glu Asn Lys Phe Arg Glu Glu Ile Gly Met Pro Arg Arg Thr Phe
195 200 205

Tyr Pro His Asp Ser Ser Leu Ile His Asp Asp Asn Thr Val Thr Gln
210 215 220

Arg Phe Gln Arg Lys Lys Leu His Pro Leu Leu
225 230 235

<210> 34
<211> 232
<212> PRT
<213> Enteropathogenic E. coli

<400> 34

Met Arg Pro Thr Ser Leu Asn Leu Val Leu His Gln Ser Ser Thr Ser
1 5 10 15

Ser Ser Met Ser Asp Thr Asp Ile Glu Ser Leu Val Lys Ala Ser Ser
20 25 30

Val Gln Trp Ile Lys Asn Asn Pro Gln Leu Arg Phe Gln Gly Thr Asp
35 40 45

His Asn Ile Tyr Gln Gln Ile Glu Ala Ala Leu Asp Lys Ile Gly Ser
50 55 60

Thr Glu Thr Gly Arg Val Leu Leu Asn Ala Ile Glu Ser Ile Ser Arg
65 70 75 80

Leu Lys Ser Glu Thr Val Val Ile His Leu Asn Ser Ser Arg Leu Gly
85 90 95

Val Met Ala His Arg Asp Ile Asp Ala Glu Asn His Arg Gly Thr Gly
100 105 110

Ser Asp Phe His Cys Asn Leu Asn Ala Val Glu Tyr Pro Cys Gly Glu
115 120 125

Gly Ile Ser Val Val Asp Phe His Ala Thr Ile Val Phe His Glu Leu
130 135 140

Leu His Val Phe His Asn Leu Asn Gly Glu Arg Leu Lys Val Glu Ser
145 150 155 160

Ser Arg Pro Glu Ser Gln Lys Tyr Ser Pro Leu Leu Leu Glu Glu Ala
165 170 175

Arg Thr Val Gly Leu Gly Ala Phe Ser Glu Glu Val Leu Ser Glu Asn
180 185 190

Lys Phe Arg Glu Glu Ile Gly Met Pro Arg Arg Thr Ser Tyr Pro His
195 200 205

Asp Ser Ala Leu Ile His Asp Asp Asn Thr Val Ser Leu Gly Phe Gln
210 215 220

Gln Val Arg Leu His Pro Leu Leu
225 230

<210> 35
<211> 232
<212> PRT
<213> Enterohemorrhagic E. coli

<220>
<221> MISC_FEATURE
<222> (208)..(208)
<223> Xaa = Arg or His

<400> 35

Met Arg Pro Thr Ser Leu Asn Leu Val Leu His Gln Ser Ser Arg Ser
1 5 10 15

Ser Ser Met Ser Asp Thr Asp Ile Glu Ser Leu Val Lys Ala Ser Ser
20 25 30

Val Gln Trp Ile Lys Asn Asn Pro Gln Leu Arg Phe Gln Gly Thr Asp
35 40 45

His Asn Ile Tyr Gln Gln Ile Glu Ala Ala Leu Asp Lys Ile Gly Ser
50 55 60

Thr Glu Thr Gly Arg Val Leu Leu Asn Ala Ile Glu Ser Ile Ser Arg
65 70 75 80

Leu Lys Ser Glu Thr Val Val Ile His Leu Asn Ser Ser Arg Leu Gly
85 90 95

Val Met Ala His Arg Asp Ile Asp Ala Glu Asn His Arg Gly Thr Gly
100 105 110

Ser Asp Phe His Cys Asn Leu Asn Ala Val Glu Tyr Pro Cys Gly Glu
115 120 125

Gly Ile Ser Val Val Asp Phe His Ala Thr Ile Val Phe His Glu Leu
130 135 140

Leu His Val Phe His Asn Leu Asn Gly Glu Arg Leu Lys Val Glu Ser
145 150 155 160

Ser Arg Ala Glu Ser Gln Lys Tyr Ser Pro Leu Leu Leu Glu Glu Ala
165 170 175

Arg Thr Val Gly Leu Gly Ala Phe Ser Glu Glu Val Leu Ser Glu Asn
180 185 190

Lys Phe His Glu Glu Ile Gly Met Pro Arg Arg Thr Ser Tyr Pro Xaa
195 200 205

Asp Ser Ala Leu Ile His Asp Asp Asn Thr Val Ser Leu Gly Phe Gln
210 215 220

Gln Val Arg Leu His Pro Leu Leu
225 230

<210> 36
<211> 168
<212> PRT
<213> Citrobacter rodentium

165
167

<400> 36

Tyr Phe Asn Glu Ser Pro Asn Val Tyr Asp Lys Lys Tyr Ile Ser Gly
1 5 10 15
Val Thr Arg Gly Val Ala Glu Leu Lys Gln Glu Gly Phe Ile Asn Glu
20 25 30
Lys Ala Arg Arg Leu Ala Tyr Met Gln Ala Met Tyr Ser Val Cys Pro
35 40 45
Glu Glu Phe Lys Pro Ile Ser Arg Asn Glu Ala Ser Thr Pro Glu Gly
50 55 60
Ser Trp Leu Thr Val Ile Ser Gly Lys Arg Pro Met Gly Gln Phe Ser
65 70 75 80
Val Asp Ser Leu Tyr His Pro Asp Leu His Ala Leu Cys Glu Leu Pro
85 90 95
Asp Ile Cys Cys Lys Ile Phe Pro Lys Glu Asn Asn Asp Phe Leu Tyr
100 105 110
Ile Val Ile Val Tyr Arg Asn Asp Ser Pro Leu Gly Glu Gln Arg Ala
115 120 125
Asn Arg Phe Ile Glu Leu Tyr Asn Ile Lys Arg Asp Ile Met Gln Glu
130 135 140
Leu Asn Tyr Glu Ser Pro Glu Leu Lys Ala Val Lys Ser Glu Met Ile
145 150 155 160
Ile Ala Arg Glu Met Gly Glu Ile
165

<210> 37

<211> 154

<212> PRT

<213> Citrobacter rodentium

<400> 37

Asn Val Tyr Asp Lys Lys Tyr Ile Ser Gly Val Thr Arg Gly Val Ala
1 5 10 15
Glu Leu Lys Gln Glu Gly Phe Ile Asn Glu Lys Ala Arg Arg Leu Ala
20 25 30
Tyr Met Gln Ala Met Tyr Ser Val Cys Pro Glu Glu Phe Lys Pro Ile
35 40 45
Ser Arg Asn Glu Ala Ser Thr Pro Glu Gly Ser Trp Leu Thr Val Ile
50 55 60
Ser Gly Lys Arg Pro Met Gly Gln Phe Ser Val Asp Ser Leu Tyr His
65 70 75 80
Pro Asp Leu His Ala Leu Cys Glu Leu Pro Asp Ile Cys Cys Lys Ile
85 90 95
Phe Pro Lys Glu Asn Asn Asp Phe Leu Tyr Ile Val Ile Val Tyr Arg
100 105 110
Asn Asp Ser Pro Leu Gly Glu Gln Arg Ala Asn Arg Phe Ile Glu Leu
115 120 125
Tyr Asn Ile Lys Arg Asp Ile Met Gln Glu Leu Asn Tyr Glu Ser Pro
130 135 140

167
168

Glu Leu Lys Ala Val Lys Ser Glu Met Ile
145 150

<210> 38
<211> 224
<212> PRT
<213> Enteropathogenic E. coli

<400> 38

Met Ile Asn Pro Val Thr Asn Thr Gln Gly Val Ser Pro Ile Asn Thr
1 5 10 15
Lys Tyr Ala Glu His Val Val Lys Asn Ile Tyr Pro Lys Ile Lys His
20 25 30
Asp Tyr Phe Asn Glu Ser Pro Asn Ile Tyr Asp Lys Lys Tyr Ile Ser
35 40 45
Gly Ile Thr Arg Gly Val Ala Glu Leu Lys Gln Glu Glu Phe Val Asn
50 55 60
Glu Lys Ala Arg Arg Phe Ser Tyr Met Lys Thr Met Tyr Ser Val Cys
65 70 75 80
Pro Glu Ala Phe Glu Pro Ile Ser Arg Asn Glu Ala Ser Thr Pro Glu
85 90 95
Gly Ser Trp Leu Thr Val Ile Ser Gly Lys Arg Pro Met Gly Gln Phe
100 105 110
Ser Val Asp Ser Leu Tyr Asn Pro Asp Leu His Ala Leu Cys Glu Leu
115 120 125
Pro Asp Ile Cys Cys Lys Ile Phe Pro Lys Glu Asn Asn Asp Phe Leu
130 135 140
Tyr Ile Val Val Val Tyr Arg Asn Asp Ser Pro Leu Gly Glu Gln Arg
145 150 155 160
Ala Asn Arg Phe Ile Glu Leu Tyr Asn Ile Lys Arg Asp Ile Met Gln
165 170 175
Glu Leu Asn Tyr Glu Leu Pro Glu Leu Lys Ala Val Lys Ser Glu Met
180 185 190
Ile Ile Ala Arg Glu Met Gly Glu Ile Phe Ser Tyr Met Pro Gly Glu
195 200 205
Ile Asp Ser Tyr Met Lys Tyr Ile Asn Asn Lys Leu Ser Lys Ile Glu
210 215 220

<210> 39
<211> 224
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 39

Met Ile Asn Pro Val Thr Asn Thr Gln Gly Val Ser Pro Ile Asn Thr
1 5 10 15
Lys Tyr Ala Glu His Val Val Lys Asn Ile Tyr Pro Glu Ile Lys His
20 25 30
Asp Tyr Phe Asn Glu Ser Pro Asn Ile Tyr Asp Lys Lys Tyr Ile Ser
35 40 45
Gly Ile Thr Arg Gly Val Ala Glu Leu Lys Gln Glu Glu Phe Val Asn
50 55 60

Glu Lys Ala Arg Arg Phe Ser Tyr Met Lys Thr Met Tyr Ser Val Cys
65 70 75 80

Pro Glu Ala Phe Glu Pro Ile Ser Arg Asn Glu Ala Ser Thr Pro Glu
85 90 95

Gly Ser Trp Leu Thr Val Ile Ser Gly Lys Arg Pro Met Gly Gln Phe
100 105 110

Ser Val Asp Ser Leu Tyr Asn Pro Asp Leu His Ala Leu Cys Glu Leu
115 120 125

Pro Asp Ile Cys Cys Lys Ile Phe Pro Lys Glu Asn Asn Asp Phe Leu
130 135 140

Tyr Ile Val Val Val Tyr Arg Asn Asp Ser Pro Leu Gly Glu Gln Arg
145 150 155 160

Ala Asn Arg Phe Ile Glu Leu Tyr Asn Ile Lys Arg Asp Ile Met Gln
165 170 175

Glu Leu Asn Tyr Glu Leu Pro Glu Leu Lys Ala Val Lys Ser Glu Met
180 185 190

Ile Ile Ala Arg Glu Met Gly Glu Ile Phe Ser Tyr Met Pro Gly Glu
195 200 205

Ile Asp Ser Tyr Met Lys Tyr Ile Asn Asn Lys Leu Ser Lys Ile Glu
210 215 220

<210> 40
<211> 188
<212> PRT
<213> Citrobacter rodentium

<400> 40

Met Leu Pro Thr Ser Gly Ser Ser Ala Asn Leu Tyr Ser Trp Met Tyr
1 5 10 15

Ile Ser Gly Lys Glu Asn Pro Ser Thr Pro Glu Ser Val Ser Glu Leu
20 25 30

Asn His Asn His Phe Leu Ser Pro Glu Leu Gln Glu Lys Leu Asp Val
35 40 45

Met Phe Ala Ile Tyr Ser Cys Ala Arg Asn Asn Asp Glu Arg Glu Asn
50 55 60

Ile Tyr Pro Glu Leu Arg Asp Phe Val Ser Ser Leu Met Asp Lys Arg
65 70 75 80

Asn Asn Val Phe Glu Val Ile Asn Glu Asp Thr Asp Glu Val Thr Gly
85 90 95

Ala Leu Arg Ala Gly Met Thr Ile Glu Asp Arg Asp Ser Tyr Ile Arg
100 105 110

Asp Leu Phe Phe Leu His Ser Leu Lys Val Lys Ile Glu Glu Ser Arg
115 120 125

Gln Asp Lys Glu Asp Trp Lys Cys Lys Val Tyr Asp Leu Leu Cys Pro
130 135 140

His His Ser Ser Glu Leu Tyr Gly Asp Leu Arg Ala Ile Lys Cys Leu
145 150 155 160

Val Glu Gly Cys Ser Asp Asp Phe Ser Pro Phe Asp Thr Ile Lys Val
165 170 175

169
170

Pro Asp Leu Thr Tyr Asn Lys Gly Ser Leu Gln Cys
180 185

<210> 41
<211> 171
<212> PRT
<213> Citrobacter rodentium

<400> 41

Ala Asn Leu Tyr Ser Trp Met Tyr Ile Ser Gly Lys Glu Asn Pro Ser
1 5 10 15
Thr Pro Glu Ser Val Ser Glu Leu Asn His Asn His Phe Leu Ser Pro
20 25 30
Glu Leu Gln Glu Lys Leu Asp Val Met Phe Ala Ile Tyr Ser Cys Ala
35 40 45
Arg Asn Asn Asp Glu Arg Glu Asn Ile Tyr Pro Glu Leu Arg Asp Phe
50 55 60
Val Ser Ser Leu Met Asp Lys Arg Asn Asn Val Phe Glu Val Ile Asn
65 70 75 80
Glu Asp Thr Asp Glu Val Thr Gly Ala Leu Arg Ala Gly Met Thr Ile
85 90 95
Glu Asp Arg Asp Ser Tyr Ile Arg Asp Leu Phe Phe Leu His Ser Leu
100 105 110
Lys Val Lys Ile Glu Glu Ser Arg Gln Asp Lys Glu Asp Trp Lys Cys
115 120 125
Lys Val Tyr Asp Leu Leu Cys Pro His His Ser Ser Glu Leu Tyr Gly
130 135 140
Asp Leu Arg Ala Ile Lys Cys Leu Val Glu Gly Cys Ser Asp Asp Phe
145 150 155 160
Ser Pro Phe Asp Thr Ile Lys Val Pro Asp Leu
165 170

<210> 42
<211> 189
<212> PRT
<213> Enteropathogenic E. coli

<400> 42

Met Leu Pro Thr Ser Gly Ser Ser Ala Asn Leu Tyr Ser Trp Met Tyr
1 5 10 15
Val Ser Gly Arg Gly Asn Pro Ser Thr Pro Glu Ser Val Ser Glu Leu
20 25 30
Asn His Asn His Phe Leu Ser Pro Glu Leu Gln Asp Lys Leu Asp Val
35 40 45
Met Val Ser Ile Tyr Ser Cys Ala Arg Asn Asn Asn Glu Leu Glu Glu
50 55 60
Ile Phe Gln Glu Leu Ser Ala Phe Val Ser Gly Leu Met Asp Lys Arg
65 70 75 80
Asn Ser Val Phe Glu Val Arg Asn Glu Asn Thr Asp Glu Val Val Gly
85 90 95
Ala Leu Arg Ala Gly Met Thr Ile Glu Asp Arg Asp Ser Tyr Ile Arg
100 105 110

130
171

Asp Leu Phe Phe Leu His Ser Leu Lys Val Lys Ile Glu Glu Ser Arg
115 120 125

Gln Gly Lys Glu Asp Ser Lys Cys Lys Val Tyr Asn Leu Leu Cys Pro
130 135 140

His His Ser Ser Glu Leu Tyr Gly Asp Leu Arg Ala Met Lys Cys Leu
145 150 155 160

Val Glu Gly Cys Ser Asp Asp Phe Asn Pro Phe Asp Ile Ile Arg Val
165 170 175

Pro Asp Leu Thr Tyr Asn Lys Gly Ser Leu Gln Cys Gly
180 185

<210> 43
<211> 189
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 43

Met Leu Pro Thr Ser Gly Ser Ser Ala Asn Leu Tyr Ser Trp Met Tyr
1 5 10 15

Val Ser Gly Arg Gly Asn Pro Ser Thr Pro Glu Ser Val Ser Glu Leu
20 25 30

Asn His Asn His Phe Leu Ser Pro Glu Leu Gln Asp Lys Leu Asp Val
35 40 45

Met Val Ser Ile Tyr Ser Cys Ala Arg Asn Asn Asn Glu Leu Glu Glu
50 55 60

Ile Phe Gln Glu Leu Ser Ala Phe Val Ser Gly Leu Met Asp Lys Arg
65 70 75 80

Asn Ser Val Phe Glu Val Arg Asn Glu Asn Thr Asp Glu Val Val Gly
85 90 95

Ala Leu Arg Ala Gly Met Thr Ile Glu Asp Arg Asp Ser Tyr Ile Arg
100 105 110

Asp Leu Phe Phe Leu His Ser Leu Lys Val Lys Ile Glu Glu Ser Arg
115 120 125

Gln Gly Lys Glu Asp Ser Lys Cys Lys Val Tyr Asn Leu Leu Cys Pro
130 135 140

His His Ser Ser Glu Leu Tyr Gly Asp Leu Arg Ala Met Lys Cys Leu
145 150 155 160

Val Glu Gly Cys Ser Asp Asp Phe Asn Pro Phe Asp Ile Ile Arg Val
165 170 175

Pro Asp Leu Thr Tyr Asn Lys Gly Ser Leu Gln Cys Gly
180 185

<210> 44
<211> 40
<212> DNA
<213> Artificial

<220>
<223> primer Z6024F

<400> 44
agatctgaag gagatattat gaacattcaa ccgaccatac

171
772

<210> 45
<211> 34
<212> DNA
<213> Artificial

<220>
<223> primer Z6024R

<400> 45
ctcgaggact cttgtttcctt cgattatatc aaag 34

<210> 46
<211> 28
<212> DNA
<213> Artificial

<220>
<223> primer NT10

<400> 46
ccggtacctc taaccattga cgcaactcg 28

<210> 47
<211> 29
<212> DNA
<213> Artificial

<220>
<223> primer NT11

<400> 47
aacctgcaga actaggtatc tctaattgcc 29

<210> 48
<211> 29
<212> DNA
<213> Artificial

<220>
<223> primer NT12

<400> 48
aacctgcagc tgactatcct cgtatatgg 29

<210> 49
<211> 27
<212> DNA
<213> Artificial

<220>
<223> primer NT13

<400> 49
ccgagctcag gtaatgagac tgtcagc 27

<210> 50
<211> 22
<212> DNA
<213> Artificial

<220>
<223> primer del1F

<400> 50
ggtaccacca cacagaataa tc 22

<210> 51
<211> 26
<212> DNA
<213> Artificial

<220>
<223> primer del1R

<400> 51
cgctagccta tatactgctg ttgggt 26

<210> 52
<211> 28
<212> DNA
<213> Artificial

<220>
<223> primer del2F

<400> 52
gctagctgac aggcaactct tggactgg 28

<210> 53
<211> 29
<212> DNA
<213> Artificial

<220>
<223> primer del2R

<400> 53
gagctcaaca taatttgatg gattatgat 29

<210> 54
<211> 24
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 54
ttccatatga acattcaacc gacc 24

<210> 55
<211> 24
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 55
ggaattcaat aatagctgcc atcc 24

<210> 56
<211> 135
<212> PRT
<213> Salmonella

<400> 56

Met Glu Ser Lys Asn Ser Asp Tyr Val Ile Pro Asp Ser Val Lys Asn
1 5 10 15

Tyr Asn Gly Glu Pro Leu Tyr Ile Leu Val Ser Leu Trp Cys Lys Leu
20 25 30

Gln Glu Lys Trp Ile Ser Arg Asn Asp Ile Ala Glu Ala Phe Gly Ile
35 40 45

Asn Leu Arg Arg Ala Ser Phe Ile Ile Thr Tyr Ile Ser Arg Arg Lys
50 55 60

Glu Lys Ile Ser Phe Arg Val Arg Tyr Val Ser Tyr Gly Asn Leu His
65 70 75 80

Tyr Lys Arg Leu Glu Ile Phe Ile Tyr Asn Val Asn Leu Glu Ala Ala
85 90 95

Pro Thr Glu Ser His Val Ser Thr Gly Pro Lys Arg Lys Thr Leu Arg
100 105 110

Val Gly Asn Gly Ile Val Gly Gln Ser Ser Ile Trp Asn Glu Met Ile
115 120 125

Met Arg Arg Lys Lys Glu Ser
130 135

<210> 57

<211> 131

<212> PRT

<213> Enterobacteriaceae

<400> 57

Met Cys Glu Gly Tyr Val Glu Lys Pro Leu Tyr Leu Leu Ile Ala Glu
1 5 10 15

Trp Met Met Ala Glu Asn Arg Trp Val Ile Ala Arg Glu Ile Ser Ile
20 25 30

His Phe Asp Ile Glu His Ser Lys Ala Val Asn Thr Leu Thr Tyr Ile
35 40 45

Leu Ser Glu Val Thr Glu Ile Ser Cys Glu Val Lys Met Ile Pro Asn
50 55 60

Lys Leu Glu Gly Arg Gly Cys Gln Cys Gln Arg Leu Val Lys Val Val
65 70 75 80

Asp Ile Asp Glu Gln Ile Tyr Ala Arg Leu Arg Asn Asn Ser Arg Glu
85 90 95

Lys Leu Val Gly Val Arg Lys Thr Pro Arg Ile Pro Ala Val Pro Leu
100 105 110

Thr Glu Leu Asn Arg Glu Gln Lys Trp Gln Met Met Leu Ser Lys Ser
115 120 125

Met Arg Arg
130

<210> 58

<211> 170

<212> PRT

<213> Citrobacter rodentium

<400> 58

Met Cys Pro Asp Asn Thr His Ala Lys Lys Gln Tyr Leu Thr Pro Gly
1 5 10 15

Asn Asp Ile His Tyr Pro Gly Gln Thr Asn His Asp Ala Cys Phe Ile
20 25 30

Pro Val Ser Val Arg Gln Tyr Ala Gly Glu Pro Leu Tyr Ile Ile Val
35 40 45

Ala His Trp Cys Leu Leu Gln Gln Asn Trp Val Gln Arg Asn Gln Ile
50 55 60

Ala Glu Ala Phe His Ile Thr Ala Arg Arg Ala Ser Tyr Leu Ile Ala
65 70 75 80

134
175

Tyr Leu Arg Ser Lys Thr Ser Arg Val Val Ser Ile Cys Arg His Gln
85 90 95

Thr Leu Pro Asn Lys Ala Arg Arg Tyr Glu Ile Tyr Val Ile Arg Val
100 105 110

Leu Asp Ser Pro Thr Pro Ser Thr Arg Arg Glu Lys Ala Gly Pro Pro
115 120 125

Leu Val Ser Lys Arg Arg Val Gly Asn Gly Asp Arg Ser Met Ala Asn
130 135 140

Glu Leu Trp Asn Arg Leu Cys Ser Asn Arg Asn Ala Gly Lys Ile Leu
145 150 155 160

Lys Lys Lys Glu Asp Glu Asp Asp Gly Thr
165 170

<210> 59
<211> 12
<212> PRT
<213> Citrobacter rodentium

<220>
<221> MISC_FEATURE
<222> (9)..(9)
<223> Xaa = Ile or Leu

<400> 59

Gln Gln Glu Asn Ala Pro Ser Ser Xaa Gln Thr Arg
1 5 10

<210> 60
<211> 981
<212> DNA
<213> Enterohemorrhagic E. coli

<400> 60
atgctttcac cgataaggac aactttccat aactcagtaa atatagtgc gagttcacc 60
tgtcaaacgg tttcttttgc aggaaaggaa tatgagttaa aggtcattga tgaaaaaacg 120
cctattcttt ttcagtgggt tgaacctaat cctgaacgat ataagaaaga tgaggttcca 180
atagttaata ctaagcagca tccctattta gataatgtca caaatgcggc aaggatagag 240
agtgatcgta tgataggtat ttttgttgat ggcgattttt cagtcaacca aaagactgct 300
ttttcaaaat tggaaacgaga ttttgaaaat gtaatgataa tctatcggga agatgttgac 360
ttcagtatgt atgacagaaa actatcagat atttatcatg atattatatg tgaacaaagg 420
ttacgaactg aagacaaaag agatgaatac ttggtgaatc tgtagagaa agagctgagg 480
gaaatttcaa aggcgcagga ttctttgatt tctatgtatg caaagaaaag aaatcatgca 540
tggtttgatt tcttcagaaa tttagcctta ttaaaagcag gagagatatt cagggtgcaca 600
tataatacaa agaatcacgg tatttcattc ggggaggggt gtatctatct tgatatggat 660
atgatactta caggtaaagt tggtaacata tatgctcctg atggaatttc aatgcatgtg 720
gatcgctgta atgatatgtt aaatattgaa aatagtgc aaattgttaa ccgtagtaat 780
catcctgctc tacttgaggg actttctttt atgcatagta aagtagatgc tcatccatat 840
tatgatgggt tggggaaagg agttaagaaa tattttaatt ttacaccatt acataattat 900
aatcattttt gtgactttat tgagtttaac caccctaata taatcatgaa cacaagtcag 960
tatacatgca gttcatggta a 981

<210> 61
<211> 531
<212> DNA
<213> Enterohemorrhagic E. coli

<400> 61
atgaatgtcc ttcgagctca agtagcatct agcggctcgag gggagtttac attaggtaat 60
gagactgtca gcattgtatt taatgaaacc gatgggcggt ttctatccag cggcagtagt 120
gggggattgc ttactgagtt attcctttat ggggttaata acggccctga agctcttcgc 180
gataggatgc tcagtatgct ttcggactca ggtgaagcac aatcgcaaga gagtattcag 240
gacaaaatat ctcaatgtaa gtttcctggt agttcaggaa atttccagtg cccgccagag 300
tctatttcagt gtccaattac actagagaga cccgaagaag gagtgtttgt caaaaattca 360

170
176

gatagttcgg	cagtatgctg	cttatttgat	tttcatgcat	tttctcggtt	agctagttaa	420
ggctcatatc	atccactgac	ccgagaacca	ataacggcat	caatgattat	aagtcctgat	480
aatgtgtgtt	atgatcctat	caagggaaac	ttcattataa	aagatagtta	a	531

<210> 62
 <211> 912
 <212> DNA
 <213> Enterohemorrhagic E. coli

<400> 62						
atgttatcgc	cctcttctat	aaatttggga	tgttcattga	attctttaac	cagaaacctg	60
acttcgcctg	ataatcgtgt	tttatcctct	gtaagggatg	ctgctgttca	ctctgatagc	120
gggacgcaag	taacggttgg	caacagaaca	tatcgtgttg	tggctactga	taataagttt	180
tgcgttacaa	gagaaagtca	tagtggttgt	tttactaatc	tgttgccacg	gttgggatgg	240
cctaagggag	agattagcag	aaaaattgag	gctatgctga	atacatcgcc	agtgagcacg	300
actatagaaa	gaggctctgt	tcattcgaac	agacctgatt	tacctccagt	ggattatgcy	360
cagccggagt	tacctccagc	ggattatact	caatcagagt	tgccgagggg	tagcaacaat	420
aaatcacccg	tgccaggtaa	tgttattggg	aaaggtggta	atgctgtcgt	gtatgaagat	480
atggaagata	caacaaaagt	gttgaagatg	tttactatat	ctcaaagcca	tgaagaggtg	540
acaagcgaa	ttcgttgttt	caatcagtat	tatgggtccg	ggagtgcaga	gaaaatatat	600
aatgataatg	gaaatgttat	tgggtattaga	atgaataaaa	taaatgggga	atctcttttg	660
gatattccat	cattaccagc	acaagctgaa	caggctatit	acgatatgtt	tgacagactg	720
gagaaaaaag	gaattctttt	tgttgataca	acagaaacaa	atgtttttata	tgatcgtatg	780
agaaatgaat	ttaatccaat	agatatatca	tcttataatg	tttctgatat	ttcatggagt	840
gaacatcaag	tcattgcaatc	ttatcacgga	ggaaagctgg	atctttattag	tgtagtatta	900
agtaagatat	aa					912

<210> 63
 <211> 882
 <212> DNA
 <213> Enterohemorrhagic E. coli

<400> 63						
atgttatcgc	catattctgt	aaatttggga	tgttcattga	attctttaac	cagaaacctg	60
acttcgcctg	ataatcgtgt	tttatcctct	gtaagggatg	ctgccgttca	ttctgataat	120
ggggcgcaag	taaagggttg	caacagaaca	tatcgtgttg	ttgccaccga	taataagttt	180
tgcgttacaa	gagaaagtca	tagtggttgt	tttactaatc	tgttgccacg	gctgggatgg	240
cctaaggggag	agattagcag	gaaaattgag	gtcatgctga	atgcatcacc	agtgagcgct	300
gctatggaaa	gaggcattgt	tcattcgaac	agacctgatt	tacctcctgt	tgattatgca	360
ccgccagagt	taccgagtgt	ggactataac	agggttgtcag	tacctggtaa	tgttattggc	420
aaagggggga	acgctgtagt	atatgaagat	gctgaggatg	caacaaaagt	cctgaagatg	480
tttactacat	ctcaaagcaa	tgaagaggtg	acaagcgaa	ttcgtttgctt	caaccaatat	540
tatgggtgccg	ggagtgcaga	aaaaatatat	ggcaataatg	gtgatattat	tgggtattaga	600
atggataaaa	taaatggaga	atcgctttta	aatatttctgt	ccttgccagc	acaggctgag	660
catgctatit	acgatatgtt	tgatagactg	gagcaaaaag	gaattctttt	tgtcgataca	720
acagagacaa	atgtctttata	tgaccgcgcg	aagaatgagt	ttaatccaat	agatatatca	780
tcttataatg	tttccgaccg	ttcatggagt	gaaagtcaaa	taatgcaatc	ttatcatggc	840
ggaaagcaag	atctttattag	tgtgggtatta	agtaaaattt	ag		882

<210> 64
 <211> 153
 <212> DNA
 <213> Enterohemorrhagic E. coli

<400> 64						
atggtaatgc	ctggattagt	atcatatata	tcattcgactt	cattcgcgaa	tgagatggcg	60
gagatgcgtc	agcaggtaat	ggaagggcag	attggtggat	ttctcctggg	aggggagaga	120
gttagagttt	cttattttatt	tcaattgcat	taa			153

<210> 65
 <211> 576
 <212> DNA
 <213> Enterohemorrhagic E. coli

<400> 65						
atgccattaa	cctcagatat	tagatcacat	tcattttaac	ttgggggtga	ggttggtcgt	60
gcccgaaattg	tagccaatgg	gcgcggagat	attacagtcg	gtgggtgaaac	tgtcagttat	120
gtgtatgatt	ctactaatgg	gcgccttttca	tccagtggcg	gtaatggcgg	attgctttct	180
gagttattgc	ttttgggatt	taatagtggg	cctcgagccc	ttgggtgagag	aatgctaagt	240
atgctttcgg	actcagggtga	agcacaatcg	caagagagta	ttcagaacaa	aatatctcaa	300

176
177

```
tgtaagtttt ctgtttgtcc agagagactt cagtgcccgcc ttgaggctat tcagtgtcca 360
attacactgg agcagcctga aaaagggtatt tttgtgaaga attcagatgg ttcagatgta 420
tgtactttat ttgatgccgc tgcattttct cgttttggtg gtgaaggctt accccaccca 480
ctgacccggg aaccaataac ggcatcaata attgtaaaac atgaagaatg cattttatgac 540
gataccagag gaaacttcat tataaagggt aattga 576
```

<210> 66
<211> 630
<212> DNA
<213> Enterohemorrhagic E. coli

<220>
<221> misc feature
<222> (439)..(439)
<223> n = any nucleotide

```
<400> 66
atgcctgtta ccaccttaag tatcccaagt atatctcaat tatctcctgc aagagtacag 60
tctttgcagg atgcagccag acttgaaagt ggaataagaa tatccattgg tagtggccaa 120
tattctgttc actatgtcca actactggat ggattttcag ttgaaccggg gagaggaggc 180
ttactggata ggctattggg gcgtgagcat cgaatggata gaagggctgt ggctctggaa 240
aggcaattaa atggagggtgt cgatttttta agtagtggtta ataactattt tcagagtgtc 300
atggcagaac acagagaaaa taaaacagggt aataaaatat taatggaaaa aataaattct 360
tgtgtatttg gaacggattc taatcacttt tcttgcccg agtcattttt gacatgcccg 420
ataacgctgg acacacctna gactggagtg ttcattgagaa actcacgagg tgctgagata 480
tgctctctat atgataagga tgcgttagtg caacttggtg aaactgggtg aactcatcct 540
ctgagtcgag aacctataac agaattcaatg attatgagaa aagacgaatg tcactttgat 600
gcaaaaagag aagctttttg ttgtaagtga 630
```

<210> 67
<211> 642
<212> DNA
<213> Enterohemorrhagic E. coli

```
<400> 67
atgcctgtag atttaacgcc ttatatttta cctgggggta gttttttgtc tgacattcct 60
caagaaacct tgtctgagat acgtaatcag actattcgtg gagaagctca agtaagactg 120
ggtgagttga tgggtgtcaat acgacctatg caggtaaagt gatattttat ggaagtcctt 180
aaccaggatg gtttatcgaa tgataacatc cagattggcc ttcaatatat agaacatatt 240
gaacgtacac ttaatcatgg tagtttgaca agccgtgaag ttacagtact gcgtgaaatt 300
gagatgctcg aaaatatgga attgctttct aactaccagt tagaggagt gttagataaa 360
attgaagtat gtgcatttaa tgtggagcat gcacaattgc aagtgccaga gagcttacga 420
acatgccctg ttacattatg tgaaccagaa gatggggtat ttatgaggaa ttcaatgaat 480
tcaaagtgtt gtatgttgta tgataaaatg tcattaatat atcttggtta aacaagggcg 540
gctcatcctt tgagcagggg atcaatcgca gtttcaatga ttgtaggaag agataattgt 600
gcttttgact ctgacagagg taacttcgtt ttaaaaaatt aa 642
```

<210> 68
<211> 642
<212> DNA
<213> Enterohemorrhagic E. coli

```
<400> 68
atgcctgtag atttaacgcc ttatatttta cctgggggta gttttttgtc tgacattcct 60
caagaaacct tgtctgagat acgtaatcag actattcgtg gagaagctca aataagactg 120
ggtgagttga tgggtgtcaat acgacctatg caggtaaagt gatattttat ggaagtcctt 180
aaccaggatg gtttatcgaa tgataaatc cagattggcc ttcaatatat agaacatatt 240
gaacgtacac ttaatcatgg tagtttgaca agccgtgaag ttacagtact gcgtgaaatt 300
gagatgctcg aaaatatgga tttgctttct aactaccagt tagaggagt gttagataaa 360
attgaagtat gtgcatttaa tgtggagcat gcacaattgc aagtgccaga gagcttacga 420
acatgccctg ttacattatg tgaaccagaa gatggggtat ttatgaggaa ttcaatgaat 480
tcaaagtgtt gtatgttgta tgataaaatg gcattaatac atcttggtta aacaagggcg 540
gctcatcctt tgagcagggg atcaatcgca gtttcaatga ttgtaggaag agataattgt 600
gcttttgacc ctgacagagg taacttcgtt ttaaaaaatt aa 642
```

<210> 69
<211> 630
<212> DNA
<213> Enterohemorrhagic E. coli

<400> 69
atgcctgtta ccaccttaag tatcccaagt atatctcaat tatctcctgc aggagtacag 60
tctttgcagg atgctgccag acttgaaagt ggaataagaa tatccattgg tagtggccaa 120
tattctgttc actatgtcca gctactggat ggattttcag ttgaaccggg gagaggaggc 180
ttactggata ggctattggg gcgtagcat cgaatggaga gaagggtgt ggctctggaa 240
aggcaattaa atggaggtgt cgatttttta agtagtgta ataactattt tcagagtgtc 300
atggcagaac acagagaaaa taaaacaagt aataaaatat taatggaaaa aataaattct 360
tgtttattta gacctgattc taatcacttt tcttgcccg agtcattttt gacatgcccg 420
ataacgctgg acacacctga gactggggtg ttcagtgaag actcacgagg tgctgagata 480
tgctctctat atgataagga cgcgttagtg caacttgttg aaactgggtg agctcatcct 540
ctgagtcgag aacctataac agaataatg attatgagaa aagatgaatg tcactttgat 600
acaaaaagag aagctttttg ttgtaagtga 630

<210> 70
<211> 576
<212> DNA
<213> Enterohemorrhagic E. coli

<400> 70
atgccattaa cctcagatat tagatcacat tcatttaatc ttgggggtgga ggttgttcgt 60
gcccgaattg tagccaatgg gcgcgagat attacagtcg gtgggtgaaac tgcagattt 120
gtgtatgatt ctactaatgg gcgcttttca tccagtggcg gtaatggcgg attgctttct 180
gagttattgc ttttgggatt taatagtggg cctcagagccc ttggtgagag aatgctaagt 240
atgctttcgg actcaggtga agcacaatcg caagagagta ttcagaacaa aatatctcaa 300
tgtaagtttt ctgtttgtcc agagagactt cagtgcccg ttgaggctat tcartgtcca 360
attacactgg agcagcctga aaaagggtatt ttgtgaaga attcagatgg ttcagatgta 420
tgtactttat ttgatgccgc tgcattttct cgtttggttg gtgaaggctt accccaccca 480
ctgacccggg aaccaataac ggcatcaata attgtaaaac atgaagaatg catttatgac 540
gataccagag gaaacttcgt tataaagggt aattga 576

<210> 71
<211> 510
<212> DNA
<213> Enterohemorrhagic E. coli

<400> 71
atggacgctt ttattgtaga tcctgttcaa ggggaactat attcgggttt aagccataca 60
gaactagccg atatcattag attggctgat tctgttgaaa atcaattgaa tggaggcaat 120
tcatttcttg atgtattcag tacatatatg gggcagggtta tttctgaatt tatgcatagt 180
aatgataaca gaattgaatt gttacagcgg cgattacatt catgttcatt tttagttaat 240
attgaagaaa tgtcttacat agatgaagca ttacagtgcc cgattacgct ggcaattcct 300
caacgaggtg tttttttaag aaatgctgaa ggttccagag tatgtagttt atatgatgaa 360
atggctcttt ctcgtataat taatgatggg atgcatcacc cactaagcag agagccaata 420
acattatcaa tgcttggtgc cagagagcag tgtgagtttg attgcagtat cggtcacttt 480
acggtgagga gtgattgtta ttcagtgtag 510

<210> 72
<211> 231
<212> DNA
<213> Enterohemorrhagic E. coli

<400> 72
atggcagacc gcaaacagca ccgcgctatc gcggagcgtc gtcacatcca gactgaaatc 60
aaccgcagac tttcccgcgc atcacgcgtc gcgcaaatca tgcacatcaa tatgctgcat 120
gagcgcagcc acgcactatc aaacatttat tccgcctctg ttttcagcta tctggcggat 180
gatctgcacg agtttcaaca gtcacatccag cagcaaaaca aactccatta a 231

<210> 73
<211> 176
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 73

Met Asn Val Leu Arg Ala Gln Val Ala Ser Ser Gly Arg Gly Glu Phe
1 5 10 15

Thr Leu Gly Asn Glu Thr Val Ser Ile Val Phe Asn Glu Thr Asp Gly
20 25 30

170
179

Arg	Phe	Leu	Ser	Ser	Gly	Ser	Ser	Gly	Gly	Leu	Leu	Thr	Glu	Leu	Phe
	35						40					45			
Leu	Tyr	Gly	Phe	Asn	Asn	Gly	Pro	Glu	Ala	Leu	Arg	Asp	Arg	Met	Leu
	50					55					60				
Ser	Met	Leu	Ser	Asp	Ser	Gly	Glu	Ala	Gln	Ser	Gln	Glu	Ser	Ile	Gln
	65				70					75				80	
Asp	Lys	Ile	Ser	Gln	Cys	Lys	Phe	Pro	Val	Ser	Ser	Gly	Asn	Phe	Gln
				85					90					95	
Cys	Pro	Pro	Glu	Ser	Ile	Gln	Cys	Pro	Ile	Thr	Leu	Glu	Arg	Pro	Glu
			100					105					110		
Glu	Gly	Val	Phe	Val	Lys	Asn	Ser	Asp	Ser	Ser	Ala	Val	Cys	Cys	Leu
		115					120					125			
Phe	Asp	Phe	Asp	Ala	Phe	Ser	Arg	Leu	Ala	Ser	Glu	Gly	Ser	Tyr	His
	130					135					140				
Pro	Leu	Thr	Arg	Glu	Pro	Ile	Thr	Ala	Ser	Met	Ile	Ile	Ser	Pro	Asp
	145				150					155					160
Lys	Cys	Val	Tyr	Asp	Pro	Ile	Lys	Gly	Asn	Phe	Ile	Ile	Lys	Asp	Ser
				165					170					175	

<210> 74
 <211> 303
 <212> PRT
 <213> Enterohemorrhagic E. coli

<400> 74

Met	Leu	Ser	Pro	Ser	Ser	Ile	Asn	Leu	Gly	Cys	Ser	Trp	Asn	Ser	Leu
1				5					10					15	
Thr	Arg	Asn	Leu	Thr	Ser	Pro	Asp	Asn	Arg	Val	Leu	Ser	Ser	Val	Arg
			20					25					30		
Asp	Ala	Ala	Val	His	Ser	Asp	Ser	Gly	Thr	Gln	Val	Thr	Val	Gly	Asn
	35						40					45			
Arg	Thr	Tyr	Arg	Val	Val	Val	Thr	Asp	Asn	Lys	Phe	Cys	Val	Thr	Arg
	50					55					60				
Glu	Ser	His	Ser	Gly	Cys	Phe	Thr	Asn	Leu	Leu	His	Arg	Leu	Gly	Trp
	65				70					75				80	
Pro	Lys	Gly	Glu	Ile	Ser	Arg	Lys	Ile	Glu	Ala	Met	Leu	Asn	Thr	Ser
				85					90					95	
Pro	Val	Ser	Thr	Thr	Ile	Glu	Arg	Gly	Ser	Val	His	Ser	Asn	Arg	Pro
			100					105					110		
Asp	Leu	Pro	Pro	Val	Asp	Tyr	Ala	Gln	Pro	Glu	Leu	Pro	Pro	Ala	Asp
	115						120					125			
Tyr	Thr	Gln	Ser	Glu	Leu	Pro	Arg	Val	Ser	Asn	Asn	Lys	Ser	Pro	Val
	130					135						140			
Pro	Gly	Asn	Val	Ile	Gly	Lys	Gly	Gly	Asn	Ala	Val	Val	Tyr	Glu	Asp
	145				150					155				160	
Met	Glu	Asp	Thr	Thr	Lys	Val	Leu	Lys	Met	Phe	Thr	Ile	Ser	Gln	Ser
				165					170					175	
His	Glu	Glu	Val	Thr	Ser	Glu	Val	Arg	Cys	Phe	Asn	Gln	Tyr	Tyr	Gly
			180					185					190		

Ser Gly Ser Ala Glu Lys Ile Tyr Asn Asp Asn Gly Asn Val Ile Gly
195 200 205

Ile Arg Met Asn Lys Ile Asn Gly Glu Ser Leu Leu Asp Ile Pro Ser
210 215 220

Leu Pro Ala Gln Ala Glu Gln Ala Ile Tyr Asp Met Phe Asp Arg Leu
225 230 235 240

Glu Lys Lys Gly Ile Leu Phe Val Asp Thr Thr Glu Thr Asn Val Leu
245 250 255

Tyr Asp Arg Met Arg Asn Glu Phe Asn Pro Ile Asp Ile Ser Ser Tyr
260 265 270

Asn Val Ser Asp Ile Ser Trp Ser Glu His Gln Val Met Gln Ser Tyr
275 280 285

His Gly Gly Lys Leu Asp Leu Ile Ser Val Val Leu Ser Lys Ile
290 295 300

<210> 75
<211> 293
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 75

Met Leu Ser Pro Tyr Ser Val Asn Leu Gly Cys Ser Trp Asn Ser Leu
1 5 10 15

Thr Arg Asn Leu Thr Ser Pro Asp Asn Arg Val Leu Ser Ser Val Arg
20 25 30

Asp Ala Ala Val His Ser Asp Asn Gly Ala Gln Val Lys Val Gly Asn
35 40 45

Arg Thr Tyr Arg Val Val Ala Thr Asp Asn Lys Phe Cys Val Thr Arg
50 55 60

Glu Ser His Ser Gly Cys Phe Thr Asn Leu Leu His Arg Leu Gly Trp
65 70 75 80

Pro Lys Gly Glu Ile Ser Arg Lys Ile Glu Val Met Leu Asn Ala Ser
85 90 95

Pro Val Ser Ala Ala Met Glu Arg Gly Ile Val His Ser Asn Arg Pro
100 105 110

Asp Leu Pro Pro Val Asp Tyr Ala Pro Pro Glu Leu Pro Ser Val Asp
115 120 125

Tyr Asn Arg Leu Ser Val Pro Gly Asn Val Ile Gly Lys Gly Gly Asn
130 135 140

Ala Val Val Tyr Glu Asp Ala Glu Asp Ala Thr Lys Val Leu Lys Met
145 150 155 160

Phe Thr Thr Ser Gln Ser Asn Glu Glu Val Thr Ser Glu Val Arg Cys
165 170 175

Phe Asn Gln Tyr Tyr Gly Ala Gly Ser Ala Glu Lys Ile Tyr Gly Asn
180 185 190

Asn Gly Asp Ile Ile Gly Ile Arg Met Asp Lys Ile Asn Gly Glu Ser
195 200 205

Leu Leu Asn Ile Ser Ser Leu Pro Ala Gln Ala Glu His Ala Ile Tyr
210 215 220

Asp Met Phe Asp Arg Leu Glu Gln Lys Gly Ile Leu Phe Val Asp Thr
225 230 235 240

Thr Glu Thr Asn Val Leu Tyr Asp Arg Ala Lys Asn Glu Phe Asn Pro
245 250 255

Ile Asp Ile Ser Ser Tyr Asn Val Ser Asp Arg Ser Trp Ser Glu Ser
260 265 270

Gln Ile Met Gln Ser Tyr His Gly Gly Lys Gln Asp Leu Ile Ser Val
275 280 285

Val Leu Ser Lys Ile
290

<210> 76

<211> 50

<212> PRT

<213> Enterohemorrhagic E. coli

<400> 76

Met Val Met Pro Gly Leu Val Ser Tyr Ile Ser Ser Thr Ser Phe Ala
1 5 10 15

Asn Glu Met Ala Glu Met Arg Gln Gln Val Met Glu Gly Gln Ile Gly
20 25 30

Gly Phe Leu Leu Gly Gly Glu Arg Val Arg Val Ser Tyr Leu Phe Gln
35 40 45

Leu His
50

<210> 77

<211> 191

<212> PRT

<213> Enterohemorrhagic E. coli

<400> 77

Met Pro Leu Thr Ser Asp Ile Arg Ser His Ser Phe Asn Leu Gly Val
1 5 10 15

Glu Val Val Arg Ala Arg Ile Val Ala Asn Gly Arg Gly Asp Ile Thr
20 25 30

Val Gly Gly Glu Thr Val Ser Ile Val Tyr Asp Ser Thr Asn Gly Arg
35 40 45

Phe Ser Ser Ser Gly Gly Asn Gly Gly Leu Leu Ser Glu Leu Leu Leu
50 55 60

Leu Gly Phe Asn Ser Gly Pro Arg Ala Leu Gly Glu Arg Met Leu Ser
65 70 75 80

Met Leu Ser Asp Ser Gly Glu Ala Gln Ser Gln Glu Ser Ile Gln Asn
85 90 95

Lys Ile Ser Gln Cys Lys Phe Ser Val Cys Pro Glu Arg Leu Gln Cys
100 105 110

Pro Leu Glu Ala Ile Gln Cys Pro Ile Thr Leu Glu Gln Pro Glu Lys
115 120 125

Gly Ile Phe Val Lys Asn Ser Asp Gly Ser Asp Val Cys Thr Leu Phe
130 135 140

Asp Ala Ala Ala Phe Ser Arg Leu Val Gly Glu Gly Leu Pro His Pro
145 150 155 160

Leu Thr Arg Glu Pro Ile Thr Ala Ser Ile Ile Val Lys His Glu Glu
165 170 175

Cys Ile Tyr Asp Asp Thr Arg Gly Asn Phe Ile Ile Lys Gly Asn
180 185 190

<210> 78
<211> 209
<212> PRT
<213> Enterohemorrhagic E. coli

<220>
<221> MISC FEATURE
<222> (147)..(147)
<223> Xaa = any amino acid

<400> 78

Met Pro Val Thr Thr Leu Ser Ile Pro Ser Ile Ser Gln Leu Ser Pro
1 5 10 15

Ala Arg Val Gln Ser Leu Gln Asp Ala Ala Arg Leu Glu Ser Gly Ile
20 25 30

Arg Ile Ser Ile Gly Ser Gly Gln Tyr Ser Val His Tyr Val Gln Leu
35 40 45

Leu Asp Gly Phe Ser Val Glu Pro Val Arg Gly Gly Leu Leu Asp Arg
50 55 60

Leu Leu Gly Arg Glu His Arg Met Asp Arg Arg Ala Val Ala Leu Glu
65 70 75 80

Arg Gln Leu Asn Gly Gly Val Asp Phe Leu Ser Ser Val Asn Asn Tyr
85 90 95

Phe Gln Ser Val Met Ala Glu His Arg Glu Asn Lys Thr Gly Asn Lys
100 105 110

Ile Leu Met Glu Lys Ile Asn Ser Cys Val Phe Gly Thr Asp Ser Asn
115 120 125

His Phe Ser Cys Pro Glu Ser Phe Leu Thr Cys Pro Ile Thr Leu Asp
130 135 140

Thr Pro Xaa Thr Gly Val Phe Met Arg Asn Ser Arg Gly Ala Glu Ile
145 150 155 160

Cys Ser Leu Tyr Asp Lys Asp Ala Leu Val Gln Leu Val Glu Thr Gly
165 170 175

Gly Thr His Pro Leu Ser Arg Glu Pro Ile Thr Glu Ser Met Ile Met
180 185 190

Arg Lys Asp Glu Cys His Phe Asp Ala Lys Arg Glu Ala Phe Cys Cys
195 200 205

Lys

<210> 79
<211> 213
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 79

Met Pro Val Asp Leu Thr Pro Tyr Ile Leu Pro Gly Val Ser Phe Leu
1 5 10 15

Ser Asp Ile Pro Gln Glu Thr Leu Ser Glu Ile Arg Asn Gln Thr Ile
20 25 30

Arg Gly Glu Ala Gln Val Arg Leu Gly Glu Leu Met Val Ser Ile Arg
35 40 45

Pro Met Gln Val Asn Gly Tyr Phe Met Gly Ser Leu Asn Gln Asp Gly
50 55 60

Leu Ser Asn Asp Asn Ile Gln Ile Gly Leu Gln Tyr Ile Glu His Ile
65 70 75 80

Glu Arg Thr Leu Asn His Gly Ser Leu Thr Ser Arg Glu Val Thr Val
85 90 95

Leu Arg Glu Ile Glu Met Leu Glu Asn Met Glu Leu Leu Ser Asn Tyr
100 105 110

Gln Leu Glu Glu Leu Leu Asp Lys Ile Glu Val Cys Ala Phe Asn Val
115 120 125

Glu His Ala Gln Leu Gln Val Pro Glu Ser Leu Arg Thr Cys Pro Val
130 135 140

Thr Leu Cys Glu Pro Glu Asp Gly Val Phe Met Arg Asn Ser Met Asn
145 150 155 160

Ser Asn Val Cys Met Leu Tyr Asp Lys Met Ser Leu Ile Tyr Leu Val
165 170 175

Lys Thr Arg Ala Ala His Pro Leu Ser Arg Glu Ser Ile Ala Val Ser
180 185 190

Met Ile Val Gly Arg Asp Asn Cys Ala Phe Asp Ser Asp Arg Gly Asn
195 200 205

Phe Val Leu Lys Asn
210

<210> 80
<211> 213
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 80
Met Pro Val Asp Leu Thr Pro Tyr Ile Leu Pro Gly Val Ser Phe Leu
1 5 10 15

Ser Asp Ile Pro Gln Glu Thr Leu Ser Glu Ile Arg Asn Gln Thr Ile
20 25 30

Arg Gly Glu Ala Gln Ile Arg Leu Gly Glu Leu Met Val Ser Ile Arg
35 40 45

Pro Met Gln Val Asn Gly Tyr Phe Met Gly Ser Leu Asn Gln Asp Gly
50 55 60

Leu Ser Asn Asp Asn Ile Gln Ile Gly Leu Gln Tyr Ile Glu His Ile
65 70 75 80

Glu Arg Thr Leu Asn His Gly Ser Leu Thr Ser Arg Glu Val Thr Val
85 90 95

Leu Arg Glu Ile Glu Met Leu Glu Asn Met Asp Leu Leu Ser Asn Tyr
100 105 110

Gln Leu Glu Glu Leu Leu Asp Lys Ile Glu Val Cys Ala Phe Asn Val
115 120 125

Glu His Ala Gln Leu Gln Val Pro Glu Ser Leu Arg Thr Cys Pro Val
130 135 140
Thr Leu Cys Glu Pro Glu Asp Gly Val Phe Met Arg Asn Ser Met Asn
145 150 155 160
Ser Asn Val Cys Met Leu Tyr Asp Lys Met Ala Leu Ile His Leu Val
165 170 175
Lys Thr Arg Ala Ala His Pro Leu Ser Arg Glu Ser Ile Ala Val Ser
180 185 190
Met Ile Val Gly Arg Asp Asn Cys Ala Phe Asp Pro Asp Arg Gly Asn
195 200 205
Phe Val Leu Lys Asn
210

<210> 81
<211> 209
<212> PRT
<213> Enterohemorrhagic E. coli

<400> 81

Met Pro Val Thr Thr Leu Ser Ile Pro Ser Ile Ser Gln Leu Ser Pro
1 5 10 15
Ala Gly Val Gln Ser Leu Gln Asp Ala Ala Arg Leu Glu Ser Gly Ile
20 25 30
Arg Ile Ser Ile Gly Ser Gly Gln Tyr Ser Val His Tyr Val Gln Leu
35 40 45
Leu Asp Gly Phe Ser Val Glu Pro Val Arg Gly Gly Leu Leu Asp Arg
50 55 60
Leu Leu Gly Arg Glu His Arg Met Glu Arg Arg Ala Val Ala Leu Glu
65 70 75 80
Arg Gln Leu Asn Gly Gly Val Asp Phe Leu Ser Ser Val Asn Asn Tyr
85 90 95
Phe Gln Ser Val Met Ala Glu His Arg Glu Asn Lys Thr Ser Asn Lys
100 105 110
Ile Leu Met Glu Lys Ile Asn Ser Cys Leu Phe Arg Pro Asp Ser Asn
115 120 125
His Phe Ser Cys Pro Glu Ser Phe Leu Thr Cys Pro Ile Thr Leu Asp
130 135 140
Thr Pro Glu Thr Gly Val Phe Met Arg Asn Ser Arg Gly Ala Glu Ile
145 150 155 160
Cys Ser Leu Tyr Asp Lys Asp Ala Leu Val Gln Leu Val Glu Thr Gly
165 170 175
Gly Ala His Pro Leu Ser Arg Glu Pro Ile Thr Glu Ser Met Ile Met
180 185 190
Arg Lys Asp Glu Cys His Phe Asp Thr Lys Arg Glu Ala Phe Cys Cys
195 200 205

Lys

<210> 82
<211> 191
<212> PRT

<213> Enterohemorrhagic E. coli

<400> 82

Met	Pro	Leu	Thr	Ser	Asp	Ile	Arg	Ser	His	Ser	Phe	Asn	Leu	Gly	Val
1				5					10					15	
Glu	Val	Val	Arg	Ala	Arg	Ile	Val	Ala	Asn	Gly	Arg	Gly	Asp	Ile	Thr
			20					25					30		
Val	Gly	Gly	Glu	Thr	Val	Ser	Ile	Val	Tyr	Asp	Ser	Thr	Asn	Gly	Arg
			35				40					45			
Phe	Ser	Ser	Ser	Gly	Gly	Asn	Gly	Gly	Leu	Leu	Ser	Glu	Leu	Leu	Leu
	50					55					60				
Leu	Gly	Phe	Asn	Ser	Gly	Pro	Arg	Ala	Leu	Gly	Glu	Arg	Met	Leu	Ser
65					70					75				80	
Met	Leu	Ser	Asp	Ser	Gly	Glu	Ala	Gln	Ser	Gln	Glu	Ser	Ile	Gln	Asn
				85					90					95	
Lys	Ile	Ser	Gln	Cys	Lys	Phe	Ser	Val	Cys	Pro	Glu	Arg	Leu	Gln	Cys
			100					105					110		
Pro	Leu	Glu	Ala	Ile	Gln	Cys	Pro	Ile	Thr	Leu	Glu	Gln	Pro	Glu	Lys
		115					120					125			
Gly	Ile	Phe	Val	Lys	Asn	Ser	Asp	Gly	Ser	Asp	Val	Cys	Thr	Leu	Phe
	130					135					140				
Asp	Ala	Ala	Ala	Phe	Ser	Arg	Leu	Val	Gly	Glu	Gly	Leu	Pro	His	Pro
145					150					155					160
Leu	Thr	Arg	Glu	Pro	Ile	Thr	Ala	Ser	Ile	Ile	Val	Lys	His	Glu	Glu
				165					170					175	
Cys	Ile	Tyr	Asp	Asp	Thr	Arg	Gly	Asn	Phe	Val	Ile	Lys	Gly	Asn	
			180					185					190		

<210> 83

<211> 169

<212> PRT

<213> Enterohemorrhagic E. coli

<400> 83

Met	Asp	Ala	Phe	Ile	Val	Asp	Pro	Val	Gln	Gly	Glu	Leu	Tyr	Ser	Gly
1				5					10					15	
Leu	Ser	His	Thr	Glu	Leu	Ala	Asp	Ile	Ile	Arg	Leu	Ala	Asp	Ser	Val
			20					25					30		
Glu	Asn	Gln	Leu	Asn	Gly	Gly	Asn	Ser	Phe	Leu	Asp	Val	Phe	Ser	Thr
		35					40					45			
Tyr	Met	Gly	Gln	Val	Ile	Ser	Glu	Phe	Met	His	Ser	Asn	Asp	Asn	Arg
	50					55					60				
Ile	Glu	Leu	Leu	Gln	Arg	Arg	Leu	His	Ser	Cys	Ser	Phe	Leu	Val	Asn
65					70					75				80	
Ile	Glu	Glu	Met	Ser	Tyr	Ile	Asp	Glu	Ala	Leu	Gln	Cys	Pro	Ile	Thr
				85					90					95	
Leu	Ala	Ile	Pro	Gln	Arg	Gly	Val	Phe	Leu	Arg	Asn	Ala	Glu	Gly	Ser
			100					105					110		
Arg	Val	Cys	Ser	Leu	Tyr	Asp	Glu	Met	Ala	Leu	Ser	Arg	Ile	Ile	Asn
		115					120					125			

Asp Gly Met His His Pro Leu Ser Arg Glu Pro Ile Thr Leu Ser Met
130 135 140

Leu Val Ala Arg Glu Gln Cys Glu Phe Asp Cys Ser Ile Gly His Phe
145 150 155 160

Thr Val Arg Ser Asp Cys Tyr Ser Val
165

<210> 84

<211> 76

<212> PRT

<213> Enterohemorrhagic E. coli

<400> 84

Met Ala Asp Arg Lys Gln His Arg Ala Ile Ala Glu Arg Arg His Ile
1 5 10 15

Gln Thr Glu Ile Asn Arg Arg Leu Ser Arg Ala Ser Arg Val Ala Gln
20 25 30

Ile Met His Ile Asn Met Leu His Glu Arg Ser His Ala Leu Ser Asn
35 40 45

Ile Tyr Ser Ala Ser Val Phe Ser Tyr Leu Ala Asp Asp Leu His Glu
50 55 60

Phe Gln Gln Leu Ile Gln Gln Gln Asn Lys Leu His
65 70 75

Extracto de p. folio 186

283
787



SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
DIVISION DE NUEVAS CREACIONES
EXTRACTO PARA PUBLICACIÓN
SOLICITUD DE PATENTE DE INVENCIÓN

(21) N° de solicitud: 06 - 52194				(51) Int Cl: C12N 15/31 ; A61K 38/16; A61P 31/00	
22) Fecha de solicitud: 2006 - 05 - 31				(71) Solicitante(s) The University of British Columbia Universidad Autónoma de México	
(30) Prioridad	(31) No. Prioridad	32) Fecha	(33) País	(72) Inventor(es) FINLAY, Brett; GRUENHEID, Samantha DENG, Wanyin VALLANCE, Bruce; PUENTE, José L.	
	80/515,703	Octubre 31 de 2003	US	(74) Apoderado: MONICA WOLF WASSERMAN	
(54) Título: FACTORES DE VIRULENCIA BACTERIANA Y SUS USOS					

Reivindicación(es):

1. Una composición caracterizada porque comprende un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a la secuencia de las SEQ ID Nos. 22-43, 59 ó 73-84; o un fragmento o variante de ellos, en combinación con un portador fisiológicamente aceptable.
2. Una composición caracterizada porque comprende una molécula de ácido nucleico codificadora de un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 22-43, 59 ó 73-84, o un fragmento o variante de ellos, en combinación con un portador fisiológicamente aceptable.
3. Una composición caracterizada porque comprende una molécula de ácido nucleico que comprende una secuencia nucleotídica sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 1-21 ó 60-72 o un fragmento o variante de ellos, en combinación con un portador fisiológicamente aceptable.
4. Una composición caracterizada porque comprende un sobrenadante de cultivo celular que comprende un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 22-43, 59 ó 73-84, o un fragmento o variante de ellos, en combinación con un portador fisiológicamente aceptable.
5. La composición tal y como se reivindica en la Reivindicación 4, caracterizada porque el polipéptido comprende 20% de la proteína celular presente en la composición.

6. La composición tal y como se reivindica en cualquiera de las reivindicaciones 1 a 5 caracterizada además porque comprende EspA, EspB, EspD, EspP, Tir, toxina Shiga 1, toxina Shiga 2 o polipéptido de intimina.

7. La composición de cualquiera de las reivindicaciones 1 a 6 caracterizada además porque comprende un adyuvante.

8. Una bacteria o una preparación de la misma, caracterizada porque la bacteria comprende una mutación en el genoma bacteriano en una secuencia nucleotídica que es sustancialmente idéntica a las SEQ. ID Nos. 1-21 ó 60-72.

9. Una bacteria o una preparación de la misma, caracterizada porque la bacteria comprende una mutación en un gen sustancialmente idéntico a *nleA*, *nleB*, *nleC*, *nleD*, *nleE*, *nleF*, *nleG*, ó *nleH* o un homólogo de los mismos.

10. La bacteria de las reivindicaciones 8 ó 9 caracterizada porque la bacteria es un agente patógeno A/E.

11. La bacteria de cualquiera de las reivindicaciones 8 a 10 caracterizada porque la bacteria es *E. Coli* enterohemorrágica (EHEC), *E. Coli* enteropatógena (EPEC) ó *Citrobacter rodentium*.

12. La bacteria tal y como se reivindica en la reivindicación 11, caracterizada porque la EHEC es EHEC O157:H7 ó EHEC O157:NM.

13. La bacteria tal y como se reivindica en la reivindicación 11, caracterizada porque la EPEC es EPEC O157:H6.

14. La bacteria de cualquiera de las reivindicaciones 8 a 13 caracterizada porque la mutación atenúa la virulencia.

15. Una composición caracterizada porque comprende la bacteria de cualquiera de las reivindicaciones 8 a 14, en combinación con un portador fisiológicamente aceptable.

16. La composición tal y como se reivindica en la reivindicación 15 caracterizada porque la bacteria está viva.

17. La composición tal y como se reivindica en la reivindicación 16 caracterizada porque la bacteria es administrada por vía oral.

18. La composición tal y como se reivindica en la reivindicación 15, caracterizada porque la bacteria está muerta.

19. La composición tal y como se reivindica en la reivindicación 18 caracterizada porque la bacteria es administrada por vía parenteral.

20. La composición tal y como se reivindica en cualquiera de las reivindicaciones 15 a 19 caracterizada además porque comprende un adyuvante.

21. La composición tal y como se reivindica en las reivindicaciones 7 ó 20 caracterizada porque el adyuvante es una emulsión de aceite en agua o una emulsión de no-aceite en agua.

22. La composición tal y como se reivindica en las reivindicaciones 7 ó 20 caracterizada porque el adyuvante comprende un aceite mineral y bromuro de dimetildioctadecilamonio.

23. La composición tal y como se reivindica en las reivindicaciones 7 ó 20 caracterizada porque el adyuvante comprende uno o más agentes seleccionados del grupo que consiste en: un agente emulsionante, un dipéptido de muramilo, un agente acuoso, un agente a base de quitosano, un saponino, un aceite, un lipopolisacárido, un extracto de pared de célula bacteriana, un ADN bacteriano, un complejo bacteriano, un oligonucleótido sintético y una base nitrogenosa alifática.

24. La composición de la reivindicación 23 caracterizada porque el agente emulsionante es seleccionado del grupo que consiste en uno o más de un agente emulsionante natural, un agente emulsionante sintético, un agente emulsionante aniónico, un agente emulsionante catiónico y un agente no iónico.

25. La composición de la reivindicación 24, caracterizada porque el agente emulsionante natural es seleccionado del grupo que consiste en uno o más de acacia, gelatina, lecitina y colesterol.

26. La composición de la reivindicación 24, caracterizada porque el agente emulsionante aniónico es seleccionado de uno o más del grupo que consiste en una sal de potasio de ácido láurico, una sal de potasio de ácido oleico, una sal de sodio de ácido láurico, una sal de sodio de ácido oleico, una sal de amonio de ácido láurico, una sal de amonio de ácido oleico, una sal de calcio de un ácido graso, una sal de magnesio de un ácido graso, una sal de aluminio de un ácido graso, un jabón metálico y un sulfonato orgánico.

27. La composición de la reivindicación 26, caracterizada porque el sulfonato orgánico es sulfato laurilo de sodio.

28. La composición de la reivindicación 24 caracterizada porque el agente emulsionante catiónico es bromuro de cetiltrimetilamonio.

29. La composición de la reivindicación 24 caracterizada porque el agente emulsionante sintético es seleccionado del grupo que consiste en uno o más de éster de glicerilo, éster de polioxietilenglicol, éter de polioxietilenglicol y éster de ácido graso de sorbitán.

30. La composición de la reivindicación 28 caracterizada porque el éster de glicerilo es monoestearato de glicerilo.

31. La composición de la reivindicación 28 caracterizada porque el éster de ácido graso de sorbitán es seleccionado de uno o más del grupo que consiste en monopalmitato de sorbitán y un derivado de polioxietileno del mismo.

32. La composición de la reivindicación 30 caracterizada porque el derivado de polioxietileno es monopalmitato de sorbitán polioxietilenado.

33. La composición de la reivindicación 23 caracterizada porque el agente acuoso es hidróxido de aluminio.

34. La composición de la reivindicación 23 caracterizada porque el aceite es seleccionado del grupo que consiste en uno o más de un aceite mineral, un aceite vegetal o un aceite animal.

35. La composición de la reivindicación 33 caracterizada porque el aceite vegetal es seleccionado del grupo que consiste en uno o más de aceite de canola, aceite de almendra, aceite de semilla de algodón, aceite de maíz, aceite de oliva, aceite de maní, aceite de cártamo, aceite de sésamo y aceite de soja.

36. La composición de la reivindicación 33 caracterizada porque el aceite animal es seleccionado del grupo que consiste en uno o más de aceite de hígado de bacalao, aceite de hipogloso, aceite de lacha, aceite de "reloj anaranjado" y aceite de hígado de tiburón.

37. La composición de las reivindicaciones 7 ó 20 caracterizada porque el adyuvante comprende un componente de aceite.

38. La composición de la reivindicación 36 caracterizada porque el componente de aceite es seleccionado del grupo que consiste en uno o más de un aceite único y una mezcla de aceites.

39. La composición de la reivindicación 21, caracterizada porque la emulsión sin aceite en agua es seleccionada del grupo que consiste en uno o más de una emulsión de aceite, una emulsión de agua en aceite y una emulsión de agua en aceite en agua.

40. La composición de la reivindicación 21 caracterizada porque la emulsión de aceite en agua es EMULSIGEN™ ó EMULSIGEN PLUS™.

41. La composición de la reivindicación 24 caracterizada porque el aceite es Amphigen®.

42. La composición de la reivindicación 7 ó 20 caracterizada porque el adyuvante comprende un extracto de pared celular micobacteriana.

43. La composición de la reivindicación 7 ó 20 caracterizada porque el adyuvante comprende un ADN micobacteriano.

44. La composición de la reivindicación 7 ó 20 caracterizada porque el adyuvante comprende un complejo de pared de célula micobacteriana.

45. La composición de la reivindicación 23 caracterizada porque la base nitrogenada alifática se selecciona del grupo que consiste en uno o más de una amina, un compuesto de amonio cuaternario, una guanidina, una benzamidina y un tiouronio.

46. La composición de la reivindicación 23 caracterizada porque la base nitrogenada alifática es N,N-dioctadecil-N,N-bis(2-hidroxietil)propanodiamina.

47. La composición de la reivindicación 7 ó 20 caracterizada porque el adyuvante comprende bromuro de dimetil-dioctadecilamonio.

48. La composición de cualquiera de las reivindicaciones 21 a 46 caracterizada porque el adyuvante está presente en una composición a una concentración de entre 20% y 40% (v/v).

49. Un método para detectar la presencia de un agente patógeno A/E en una muestra, caracterizado porque el método comprende:

a) proporcionar una muestra; y

b) detectar la presencia de una molécula de ácido nucleico que comprende una secuencia nucleotídica sustancialmente idéntica a una secuencia seleccionada de uno o más del grupo que consiste en las SEQ. ID Nos. 1-21 ó 60-72 o un fragmento o una variante de las mismas.

c) detectar la presencia de una molécula de ácido nucleico que codifica para un polipéptido sustancialmente idéntico a una secuencia seleccionada de uno o más del grupo que consiste en las SEQ.ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas;

d) detectar la presencia de un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a una secuencia seleccionada de uno o más del grupo que consiste en las SEQ. ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas;

en donde la presencia de una molécula de ácido nucleico o el polipéptido indica la presencia de un agente patógeno A/E en la muestra.

50. El método de la reivindicación 48 caracterizado porque la muestra es materia fecal o sangre.

51. El método de la reivindicación 48 ó 49 caracterizado porque el paso de detectar comprende poner en contacto la secuencia nucleotídica con una sonda o iniciador sustancialmente idénticos a:

a) una secuencia nucleotídica seleccionada del grupo que consiste en uno o más de las SEQ: ID Nos. 1-21 ó 60-72 o un fragmento o una variante de las mismas, o

b) una secuencia nucleotídica codificadora de un polipéptido sustancialmente idéntico a una o más de las SEQ: ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas.

52. El método de las reivindicaciones 48 ó 49, caracterizado porque la detección comprende poner en contacto la secuencia de aminoácido con un anticuerpo que une específicamente una secuencia seleccionada del grupo que consiste en una o más de las SEQ ID Nos: 22-43, 59, 73-84 o un fragmento o una variante de las mismas.

53. Un método para producir una respuesta inmune contra un agente patógeno A/E, o un componente del mismo, en un animal, caracterizado porque el método comprende administrar al animal una cantidad efectiva de una composición tal y como se reivindica en cualquiera de las reivindicaciones 1-7 ó 15-47, o que comprende administrar al animal una cantidad efectiva de la bacteria de cualquiera de las reivindicaciones 8-14, produciendo así una respuesta inmune en el animal.

54. Un método para reducir la colonización de un agente patógeno A/E en un animal, caracterizado porque el método comprende administrar al animal una cantidad efectiva de la composición de cualquiera de las reivindicaciones 1-7 ó 15-47 o que comprende administrar al animal una cantidad efectiva de la bacteria de cualquiera de las reivindicaciones 8-14, reduciendo así la colonización del agente patógeno A/E en el animal.

55. Un método para reducir la propagación de un agente patógeno A/E en un animal caracterizado porque comprende administrar al animal una cantidad efectiva de la composición de cualquiera de las reivindicaciones 1-7 ó 15-47, o que comprende administrar al animal una cantidad efectiva de la bacteria de cualquiera de las reivindicaciones 8-14, reduciendo así la propagación del agente patógeno A/E en el animal.

56. El método de cualquiera de las reivindicaciones 52 a 54 caracterizado porque el animal es un rumiante.

57. El método de la reivindicación 55, caracterizado porque el rumiante es un bovino o un ovino.

58. El método de cualquiera de las reivindicaciones 52 a 54 caracterizado porque el animal es un ser humano.

59. Un método para tratar o prevenir la infección causada por un agente patógeno A/E, caracterizado porque dicho método comprende:

a) identificar un animal que presente, o que esté en riesgo de contraer, una infección por agente patógeno A/E; y b) administrar al animal una cantidad efectiva de un compuesto que atenúa la virulencia de un agente patógeno A/E, en donde el compuesto inhibe la expresión, secreción o actividad biológica de un polipéptido que comprende una secuencia de aminoácido sustancialmente idéntica a la secuencia de cualquiera de las SEQ. ID. Nos: 22-43, 59, 73-84.

60. El método de la reivindicación 58, caracterizado porque el compuesto es una molécula de ácido nucleico antisentido que es complementaria de una secuencia nucleotídica sustancialmente idéntica a la secuencia de cualquiera de las SEQ. ID Nos. 1-21, 59, 73-84.

61. El método de la reivindicación 58, caracterizado porque el compuesto es un ARNsi.

62. Un método para atenuar la virulencia de un agente patógeno A/E, caracterizado porque el método comprende mutar uno o más de un gen seleccionado del grupo que consiste en *nleA*, *nleB*, *nleC*, *nleD*, *nleE*, *nleF*, *nleG* y *nleH*, o un homólogo de los mismos en el agente patógeno A/E, o mutar una o más de una secuencia nucleotídica en el genoma del agente patógeno A/E, en donde la secuencia nucleotídica es seleccionada de las SEQ. ID. Nos. 1-21 ó 60-72, atenuando así la virulencia.

63. Un método para seleccionar un compuesto que atenúa la virulencia del agente patógeno A/E, caracterizado porque el método comprende:

a) proporcionar un sistema que comprende:

(i) una molécula de ácido nucleico que comprende una secuencia nucleotídica sustancialmente idéntica a las SEQ. ID Nos 1-21 ó 60-72 o un fragmento o una variante de las mismas; ó

(ii) una molécula de ácido nucleico codificadora de un polipéptido que comprende una secuencia de aminoácido sustancialmente idéntica a las SEQ. ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas; ó

(iii) un polipéptido que comprende una secuencia de aminoácidos sustancialmente idéntica a las SEQ. ID. Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas;

b) proporcionar un compuesto de prueba; y

c) determinar si el compuesto de prueba modula la expresión, secreción o actividad biológica del polipéptido o la molécula de ácido nucleico, en donde un cambio en la expresión, secreción o actividad

biológica del polipéptido o la molécula de ácido nucleico indica un compuesto que atenúa la virulencia de un agente patógeno A/E.

64. El método de la reivindicación 62, caracterizado porque el sistema es una célula.

65. El método de la reivindicación 62 ó 63, caracterizado porque la célula es *E. Coli* enterohemorrágica (EHEC), *E. Coli* enteropatogénica (EPEC) o *Citrobacter rodentium*.

66. El método de la reivindicación 62, caracterizado porque el sistema es un sistema *in vitro*.

67. Un método para producir un polipéptido de agente patógeno A/E caracterizado porque comprende:

a) proporcionar una célula recombinante que comprende:

(i) una molécula de ácido nucleico que comprende una secuencia nucleotídica sustancialmente idéntica a las SEQ. ID. Nos. 1-21 ó 60-72; ó

(ii) una molécula de ácido nucleico codificadora de un polipéptido que comprende una secuencia de aminoácido sustancialmente idéntica a las SEQ. ID. Nos. 22-43, 59 ó 73-84; y

b) cultivar la célula recombinante en condiciones que permitan la expresión del polipéptido.

68. El método de la reivindicación 66 caracterizado además porque comprende cultivar la célula recombinante en condiciones que permitan la secreción del polipéptido.

69. El método de la reivindicación 66 ó 67 caracterizado porque el polipéptido se secreta desde la célula.

70. El método de cualquiera de las reivindicaciones 66 a 68 caracterizado además porque comprende aislar el polipéptido.

71. El método de cualquiera de las reivindicaciones 48 a 69 caracterizado porque el agente patógeno A/E es *E. Coli* enterohemorrágica (EHEC), *E. Coli* enteropatogénica (EPEC) ó *Citrobacter rodentium*.

72. El método de la reivindicación 70 caracterizado porque la EHEC es EHEC O157:H7 ó EHEC O157:NM.

73. El método de la reivindicación 70 caracterizado porque la EPEC es EPEC O127:H6.

74. Un polipéptido recombinante caracterizado porque comprende una secuencia de aminoácido sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 22-43, 59 ó 73-84.

75. Una molécula de ácido nucleico aislada caracterizada porque comprende una secuencia nucleotídica sustancialmente idéntica a la secuencia de las SEQ. ID Nos. 1-21 ó 60-72.

76. Un vector caracterizado porque comprende la secuencia nucleotídica de la reivindicación 74.

77. El vector de la reivindicación 75 caracterizado porque el vector es capaz de integrarse al genoma del agente patógeno A/E.

78. El vector de la reivindicación 75 caracterizado porque el vector es capaz de integrarse al genoma del agente patógeno A/E.

79. Una célula hospedadora caracterizada porque comprende el vector de cualquiera de las reivindicaciones 75 a 77.

80. La célula hospedadora de la reivindicación 78, caracterizada porque la célula hospedadora es un agente patógeno A/E.

81. La célula hospedadora de la reivindicación 79 caracterizada porque el agente patógeno A/E es *E. Coli* enterohemorrágica (EHEC), *E. Coli* enteropatogénica (EPEC) o *Citrobacter rodentium*.

82. El uso de la composición de cualquiera de las reivindicaciones 1-7 ó 15-47, la bacteria de cualquiera de las reivindicaciones 8-14, el polipéptido de la reivindicación 73 o la molécula de ácido nucleico de la reivindicación 74, para la preparación de un medicamento para producir una respuesta inmune contra un agente patógeno A/E, o un componente del mismo, o para reducir la propagación o la colonización de un agente patógeno A/E en un animal, o para tratar o prevenir la infección causada por un agente patógeno A/E.

83. Un *kit* caracterizado porque comprende un reactivo para detectar un agente patógeno A/E en una muestra y un prospecto con instrucciones para detectar el agente patógeno A/E en la muestra.

84. El *kit* de la reivindicación 83, caracterizado porque el reactivo comprende una sonda o una sonda iniciadora o un iniciador sustancialmente idéntico a:

a) una secuencia nucleotídica seleccionada del grupo que consiste en una o más de las SEQ. ID No. 1-21 ó 60-72 o un fragmento o una variante de las mismas;

b) una secuencia nucleotídica que codifica para un polipéptido sustancialmente idéntico a una o más de las SEQ. ID Nos. 22-43, 59, 73-84 o un fragmento o una variante de las mismas.

85. El *kit* de la reivindicación 83, caracterizado porque el reactivo comprende un anticuerpo que une específicamente una secuencia seleccionada del grupo que consiste en una o más de las SEQ. ID No. 22-43, 59, 73-84 o un fragmento o una variante de las mismas.

199
798

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800176089-2

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

Rad: 06-052104-00000-0000

Fec: 2006-05-31 18:18:44 Dep. 2020 NUECREACIONES

Tm. 11 PATENTES/II

Ed: 378 FASE NACIONAL I

Act 411 PRESENTACION Fec: 199

RECIBO OFICIAL DE CAJA : 06 - 40,937
FECHA : MAYO 31 DE 2006

***** CONSIGNACION *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	Vr. PAGO
RONICA WOLF	CONSIGNACION	BANCO POPULAR	050-00110-6	55274044	31/05/2006	463,000.00

***** CONCEPTO *****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE IN	463,000.00
		TOTAL :	\$ 463,000.00

SON: CUATROCIENTOS SESENTA Y TRES MIL PESOS

RESPONSABLE : _____

RECIBO DE CAJA APLICADO AL EXPEDIENTE No. _____

199



Industria y Comercio
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
Rad: 06-052184- 00000-0000
Fec: 2006-08-31 16:18:44 Dep. 2020 NUECREACIONES
Trz. 11 PATENTE/VI
EVE: 378 FASENACIONALI
Act 411 PRESENTACION Folio #: 199

199
28

**Delegatura de Propiedad Industrial
División de Nuevas Creaciones**

**REQUISITOS NECESARIOS PARA PRESENTACIÓN Y AGILIZACIÓN DEL
TRAMITE**

	USUARIO	RADICADOR
PATENTE DE INVENCION (X) MODELO DE UTILIDAD ()	()	()
-Indicación de que se solicita la concesión de una patente.✓	()	()
-Datos de identificación del solicitante o de la persona que se presenta la solicitud.✓	()	()
-Descripción de la invención.✓	()	()
-Dibujos de ser estos pertinentes.✓	()	()
-Comprobante de Pago de las tasas establecidas.✓	()	()
COMPLETA () INCOMPLETA ()	()	()
DISEÑO INDUSTRIAL ()		
-Indicación de que se solicita el registro de Diseño Industrial.	()	()
-Datos de identificación del solicitante o de la persona que se presenta la solicitud.	()	()
-Representación gráfica y fotográfica del Diseño Industrial o muestra del Material que incorpora el diseño.	()	()
-Comprobante de Pago de las tasas establecidas.	()	()
COMPLETA () INCOMPLETA ()	()	()
ESQUEMA DE TRAZADO ()		
-Indicación de que se solicita el registro de Diseño Industrial.	()	()
-Datos de identificación del solicitante o de la persona que se presenta la solicitud.	()	()
-Representación gráfica del Esquema Trazado.	()	()
-Comprobante de Pago de las tasas establecidas.	()	()
COMPLETA () INCOMPLETA ()	()	()

Firma Responsable:


Mayo 31 de 2006.

Fecha:

200
10/1

REPUBLICA DE COLOMBIA
SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

Bogotá,
2020

Doctor(a)
WOLF WASSERMAN MONICA

REFERENCIA	Radicación No.	6 52194
	Solicitud internacional	CA2004/001891
	Fecha inicio fase nacional	31/05/2006
	Trámite	11
	Evento	378
	Actuación	430
	Oficio No.	8804

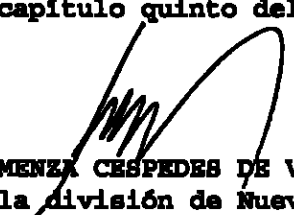
Como resultado del examen de forma, realizado con fundamento en el artículo 38 de la Decisión 486 de la Comisión de la Comunidad Andina, artículo 27 del Tratado de Cooperación en Materia de Patentes (PCT), regla 51 bis del Reglamento y Circular Única de la Superintendencia de Industria y Comercio, se le comunica que:

1. La solicitud no contiene la copia del documento en que consta la cesión del derecho a la patente del inventor al solicitante o a su causante. (Art. 26-k) Decisión 486.

2. La solicitud no contiene el poder debidamente otorgado, con el lleno de los requisitos exigidos por los artículos 63 y subsiguientes del Código de Procedimiento Civil. Adicionalmente, deberá presentar el documento que acredita la existencia y representación legal de la sociedad solicitante, cuando ésta fuere persona jurídica.

En cumplimiento de lo dispuesto por el artículo 39 de la Decisión 486, se le requiere para que dentro del término de dos meses siguientes a la fecha de notificación del presente oficio, complete los requisitos anteriormente enunciados. En caso de no dar cumplimiento al presente requerimiento, la solicitud se considerará abandonada y perderá su prelación.

Notifíquese el presente oficio conforme a lo dispuesto en el numeral 5.2, literal e), del capítulo quinto del título primero de la Circular Unica No.10 de 2001.


ALIX CARMENZA CESPEDES DE VERGEL
Jefe de la división de Nuevas Creaciones

El anterior, requerimiento fue notificado por fijación en lista de fecha 11 AGO. 2006
adpa11