

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 659 469**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.06.2013 PCT/IB2013/054888**

87 Fecha y número de publicación internacional: **19.12.2013 WO13186754**

96 Fecha de presentación y número de la solicitud europea: **14.06.2013 E 13747516 (6)**

97 Fecha y número de publicación de la concesión europea: **13.12.2017 EP 2861762**

54 Título: **Método para detectar e identificar Escherichia Coli enterohemorrágica**

30 Prioridad:

14.06.2012 EP 12171941

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

15.03.2018

73 Titular/es:

**AGENCE NATIONALE DE SÉCURITÉ SANITAIRE
DE L'ALIMENTATION, DE L'ENVIRONNEMENT ET
DU TRAVAIL (50.0%)**

**14 rue Pierre et Marie Curie
94701 Maisons-Alfort Cedex, FR y
BUNDESINSTITUT FÜR RISIKOBEWERTUNG
(50.0%)**

72 Inventor/es:

**FACH, PATRICK;
DELANNOY, SABINE y
BEUTIN, LOTHAR**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 659 469 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método para detectar e identificar *Escherichia Coli* enterohemorrágica

La invención se refiere a la identificación de *E. coli* productora de toxina Shiga (ECTS) que constituye un riesgo grave para la salud humana.

La *Escherichia coli* productora de toxina Shiga (ECTS) es un grupo diverso de *E. coli* que pertenece a más de 400 serotipos O:H de *E. coli*, algunos de los cuales causan brotes y casos esporádicos de enfermedades transmitidas por alimentos que van desde diarrea a colitis hemorrágica (CH) y el síndrome urémico hemolítico (SUH). Según su patogenicidad humana, las últimas cepas también se denominaron *E. coli* enterohemorrágica (ECEH) (Levine 1987, Nataro and Kaper 1998). Numerosos casos de CH y SUH se han atribuido a cepas ECEH con serotipo O157:H7, pero ahora se ha reconocido que otros serotipos de ECTS pertenecen al grupo ECEH.

Por lo tanto, la evidencia acumulada de numerosos países indica que hasta el 30-60% de las infecciones humanas por ECTS son causadas por ECTS no O157 y que tan solo de cinco a siete serotipos "prioritarios" de ECTS están implicados en brotes y casos esporádicos de CH y SHU. Estos comprenden los serotipos O26:[H11], O45:[H2], O103:[H2], O111:[H8], O121:[H19], O145:[H28], O157:[H7] y sus derivados no móviles. Además, una cepa inusual de O104:[H4] se ha asociado con el mayor brote de CH y SUH mundial en 2011 (Scheutz et al., 2011; Frank et al., 2011; Struelens et al., 2011; Gault et al., 2011).

En consecuencia, muchas jurisdicciones están considerando la implementación de programas de inspección de alimentos para salvaguardar al público de estas cepas ECTS con alta virulencia para los humanos. Una estrategia racional para la detección de estas cepas de *E. coli* enterohemorrágica (ECEH), como parte de un programa de inspección de alimentos basado en el riesgo, requiere una definición clara de la característica distintiva de la ECTS prioritaria (p. ej., serogrupo, serotipos, virulencia y otros marcadores) y estrategias efectivas para detectar estos ECTS patógenos en los alimentos. La detección de ECEH no O157 es particularmente desafiante por que no tienen características específicas que las distingan de la gran cantidad de *E. coli* comensales inofensivas que comparten los mismos nichos. Karmali et al. (2003) han propuesto una clasificación de seropatótipo como un marco para identificar los O-serogrupos más importantes implicados en brotes transmitidos por alimentos, basado en la gravedad de la enfermedad, frecuencia y asociación con brotes, pero las razones de la diferencia en la virulencia entre las diversas cepas ECTS siguen sin estar claras. Es probable que esta diferencia se deba a diferencias en el patrón de genes de virulencia que poseen las cepas ECTS y se necesitan estudios para fundamentar esto y para identificar marcadores moleculares apropiados.

Existen técnicas para determinar la presencia de una contaminación por ECTS en una muestra detectando, por ejemplo, la presencia de los genes *stx1/stx2* y el gen *eae* localizado en el LEE (locus de borrado del enterocito), un locus que se identificó por primera vez en *E. coli* enteropatógena ECEP. Pero la base genética de la patogenicidad de ECTS es mucho más compleja que la presencia o ausencia de uno o ambos de estos genes. En una muestra compleja (p. ej., muestras de alimentos, fecales y ambientales), que puede comprender una mezcla de cepas (p. ej., una mezcla de cepas ECTS y ECEP), la presencia de los genes *stx1/2* y *eae* no es indicativa de la presencia de una ECEH en esta muestra.

Sin embargo, dado que algunas cepas ECTS pueden causar problemas de salud muy serios en humanos, la detección de una cepa de ECTS en un producto alimenticio conduce a descartar dicho producto, aunque es probable que esta ECTS no represente una amenaza para la salud humana. Esto da como resultado una gran cantidad de desperdicio debido a la falta de discriminación entre cepas ECTS no patógenas y cepas ECEH.

Se han descrito varios métodos: El documento PCT WO 03/062464 y Beutin *et al.* (Journal of Applied Microbiology, vol. 106, n°4, 2009, p. 1122-1132) proponen el uso de genes diana *stx1* y *stx2* en una detección basada en PCR múltiple de serotipos de ECEH; Burgarel *et al.* (Applied and Environmental Microbiology, vol. 77, n°7, 2011, p.2275-2281) utiliza una PCR para detectar y caracterizar el serotipo O26:H11 de ECEH con genes diana *stx1*, *stx2*, *espN*, *espK* y *espM2*; Burgarel *et al.* (BMC Microbiology, Biomed Central, London GB, vol. 11, n°1, 2011 p.142) describe el análisis y la detección de una pluralidad de genes diana, *stx1*, *stx2* y *espK*, en diferentes serotipos diana; el Instituto Superior di Sanita ("detection and identification of Verocytotoxin-producing *Escherichia coli* (ECVT) O104:H4 in food by real time PCR", 2011) describe la detección de O104:H4 mediante la amplificación de un fragmento *wzx* por PCR.

Se ha propuesto utilizar, además de los marcadores *stx1/stx2* y *eae*, otros marcadores genéticos para detectar selectivamente cepas ECEH y diferenciarlas de cepas ECTS no patógenas. Por ejemplo, el documento PCT WO 2011/018762 describe un método que implica la detección combinada de los genes *stx1*, *stx2*, *eae*, *nleB* y *espK* para predecir la presencia de ECEH en una muestra.

Sin embargo, todavía existe la necesidad de pruebas fiables que permitan una selección discriminativa de la presencia de ECEH, que incluya ECEH no O157, y una detección específica de los serotipos ECEH implicados, en particular en el caso de los "siete principales" serotipos O26:[H11], O45:[H2], O103:[H2], O111:[H8], O121:[H19], O145:[H28], O157:[H7].

Los inventores ahora han identificado marcadores genéticos discriminativos asociados con varias cepas ECTS que constituyen un riesgo grave para la salud humana. En particular, han identificado marcadores genéticos localizados dentro de las secuencias CRISPR (Repeticiones palindrómicas cortas agrupadas y regularmente espaciadas, por sus siglas en inglés) de cepas ECEH con alta virulencia para humanos.

Las CRISPR están presentes dentro de los genomas de muchas especies bacterianas, que incluyen *E. coli*. Consisten en secuencias en tándem que contienen repeticiones directas de 21 a 47 pb de longitud y separadas por espaciadores de tamaño similar. Los espaciadores se derivan de ácidos nucleicos extraños, tales como fagos o plásmidos, y se ha formulado la hipótesis de que pueden proteger a bacterias de infecciones posteriores por fagos y plásmidos homólogos.

Ibtissem *et al.* (BMC Bioinformatics, Biomed Central London, vol. 8, n°1, 2007, p.172) sugieren el uso de la base de datos de CRISPR y herramientas para identificar la secuencia CRISPR y Fabre *et al.* (PLOS ONE, vol. 7, n°5, 2012, p. e36995) describen el uso de la secuencia CRISPR para detectar subtipos y cepas de Salmonella; pero ninguna de estas descripciones sugieren la detección de serotipos de ECEH con una selección específica de secuencias diana CRISPR.

Los inventores han secuenciado los loci CRISPR de varias cepas ECEH que están asociadas con los casos clínicos más frecuentes del mundo, y han identificado diferentes espaciadores que se pueden utilizar para una identificación específica de los serotipos O157:[H7], O145:[H28], O103:[H2], O111:[H8], O121:[H19], O45:[H2], O26:[H11], O104:[H4] de ECEH y sus derivados no móviles, que son responsables de la mayoría de Infecciones ECEH en humanos.

Por lo tanto, un objeto de la presente invención es un método para identificar el(los) serotipo(s) de ECEH que se sospecha que están presentes en una muestra, en donde dicho método comprende detectar la presencia o ausencia, en dicha muestra o ADN aislado de la misma, de las siguientes secuencias CRISPR de *E. Coli*:

a) secuencias CRISPR para identificar ECEH O157:[H7] en donde dichas secuencias CRISPR se seleccionan entre:

- las secuencias CRISPR SEQ ID NO: 1, SEQ ID NO: 2, y SEQ ID NO: 3, en donde la presencia de una o más de dichas secuencias SEQ ID NO: 1-3 es indicativa de la presencia de ECEH O157:[H7]; y/o
- la secuencia CRISPR SEQ ID NO: 4, en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O157:[H7]; y

b) una secuencia CRISPR para identificar ECEH O145:[H28], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 5, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O145:[H28]; y

c) una secuencia CRISPR para identificar ECEH O111:[H8], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 6, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O111:[H8]; y

d) una secuencia CRISPR para identificar ECEH O121:[H19], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 7, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O121:[H19]; y

e) una secuencia CRISPR para identificar ECEH O103:[H2] y/o ECEH O45:[H2], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 8, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O103:[H2] y/o de ECEH O45:[H2]; y

f) una secuencia CRISPR para identificar ECEH O104:[H4], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 9, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O104:[H4]; y

g) una secuencia CRISPR para identificar ECEH O26:[H11], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 10, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O26: [H11].

Según una realización preferida de la invención, dicho método comprende realizar un ensayo de PCR sobre dicha muestra o ADN aislado de la misma, con cebadores diseñados para amplificar dichas secuencias CRISPR, y verificar la presencia de los productos de amplificación correspondientes.

Preferiblemente, dicho ensayo de PCR se realiza con una combinación de cebadores que comprende:

- a) cebadores para detectar ECEH O157:[H7], en donde dichos cebadores consisten en:

- un conjunto de cebadores dirigidos tanto a la secuencia CRISPR SEQ ID NO: 1 como a la SEQ ID NO: 2, en donde dichos cebadores están definidos por las siguientes secuencias:

GGGAACACAAACCGAAACACA (SEQ ID NO: 11)

5 CTTAGTGTGTTCCCCGCGC (SEQ ID NO: 12) y

- un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 3 en donde dichos cebadores están definidos por las siguientes secuencias:

GAACACTTTGGTGACAGTTTTTGT (SEQ ID NO: 13); CTTAGTGTGTTCCCCGCGC (SEQ ID NO: 14),

10 en donde la presencia de un producto de amplificación para al menos uno de dichos conjuntos de cebadores es indicativa de la presencia de ECEH O157:[H7]; y/o

- un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 4, en donde dichos cebadores están definidos por las siguientes secuencias:

GAACACAAACCGAAACACACG (SEQ ID NO: 15) ATAAACCGTCACCAAACAGTG (SEQ ID NO: 16),

15 en donde la presencia de un producto de amplificación para dicho conjunto de cebadores es indicativa de la presencia de ECEH O157:[H7];y

b) cebadores para detectar ECEH O145:[H28], en donde dichos cebadores consisten en:

- un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 5, en donde dichos cebadores están definidos por las siguientes secuencias:

GAACCTTGAGCCCTGCCAGAA (SEQ ID NO: 17)

20 ACCGCGATCTTTTCTACCTG (SEQ ID NO: 18),

en donde la presencia de un producto de amplificación para dicho conjunto de conjunto de cebadores es indicativa de la presencia de ECEH O145:[H28]; y

c) cebadores para detectar ECEH O111:[H8], en donde dichos cebadores consisten en:

25 - un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 6, en donde dichos cebadores están definidos por las siguientes secuencias:

GTGACCGCCTGTACACGC (SEQ ID NO: 19)

CGGATATTTGGGCGTAATACC (SEQ ID NO: 20)

CTGCCGCGAGTGTTTCAC (SEQ ID NO: 21),

30 en donde la presencia de un producto de amplificación para al menos uno de los pares de cebadores con SEQ ID NO: 19 y SEQ ID NO: 20 o SEQ ID NO: 19 y SEQ ID NO: 21 es indicativa de la presencia de ECEH O111:[H8]; y

d) cebadores para detectar ECEH O121:[H19], en donde dichos cebadores consisten en:

- un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 7, en donde dichos cebadores están definidos por las siguientes secuencias:

CGGGGAACACTACAGGAAAGAA (SEQ ID NO: 22)

35 GGCGGAATACAGGACGGGTGG (SEQ ID NO: 23),

en donde la presencia de un producto de amplificación para dicho conjunto de cebadores es indicativa de la presencia de ECEH O121:[H19]; y

e) cebadores para detectar ECEH O103:[H2] y/o ECEH O45:[H2], en donde dichos cebadores consisten en:

40 - un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 8, en donde dichos cebadores están definidos por las siguientes secuencias:

GAGTCTATCAGCGACTACC (SEQ ID NO: 24)

AACCGCAGCTCGCAGCGC (SEQ ID NO: 25),

en donde la presencia de un producto de amplificación para dicho conjunto de cebadores es indicativa de la presencia de ECEH O103:[H2] y/o ECEH O45:[H2]; y

f) cebadores para detectar ECEH O104:[H4], en donde dichos cebadores consisten en:

- 5 - un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 9, en donde dichos cebadores están definidos por las siguientes secuencias:

GGAACTCACCGAGCGCCG (SEQ ID NO: 26);

GCCTTTGCAGCGTCTTTCCGATC (SEQ ID NO: 27);

10 en donde la presencia de un producto de amplificación para dicho conjunto de cebadores es indicativa de la presencia de ECEH O104:[H4]; y

g) cebadores para detectar ECEH O26:[H11], en donde dichos cebadores consisten en:

- dos conjuntos de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 10, en donde el primer conjunto de cebadores está definido por las siguientes secuencias:

ACAATCGTGTGTAATTCGCGG (SEQ ID NO: 28)

15 GATAAACCGTGGTACGGAACA (SEQ ID NO: 29) y el segundo dicho conjunto de cebadores está definido por las siguientes secuencias:

TGAAACCACTCGCGGCAGAT (SEQ ID NO: 30);

ATAAACCGATCTCCTCATCCTC (SEQ ID NO: 31);

20 en donde la presencia de un producto de amplificación para al menos uno de los dichos conjuntos de cebadores es indicativa de la presencia de O26:[H11].

Los productos de amplificación se pueden detectar mediante cualquier método apropiado para la detección de productos de PCR. Por ejemplo, se pueden detectar por medio de sondas derivadas de las respectivas secuencias diana.

A continuación se dan ejemplos de sondas preferidas:

25 - una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 1 y SEQ ID NO: 2, definida por la siguiente secuencia: CGATCAATCCGAATATGAGCGGT (SEQ ID NO: 32), y una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 3, definida por la siguiente secuencia: CACTGTTTTGGTGACGGTTTATCC (SEQ ID NO: 33), y/o una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 4, definida por la siguiente secuencia:

30 ACAAAAAAATACTACTACCAAGTGTTT (SEQ ID NO: 34);

- una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 5, definida por la siguiente secuencia:

TGGGCCTCTTTTGTACCCGG (SEQ ID NO: 35);

35 - una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 6, definida por la siguiente secuencia:

TGTAATGGCTCACCGGTTTATCCCC (SEQ ID NO: 36);

- una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 7, definida por la siguiente secuencia:

TCCGCCAACGGCGACAGGGG (SEQ ID NO: 37);

40 - una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 8, definida por la siguiente secuencia:

TCGGAACGTGGCGCTATAGGTG (SEQ ID NO: 38);

- una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 9, definida por la siguiente secuencia:

45 CTGGGAGGCGTATCTCACGTTCCGGT (SEQ ID NO: 39);

- una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 10, definida por la siguiente secuencia:

TGCTGTCTATATTTTCGACCAGTGTTC (SEQ ID NO: 40);

5 - una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 10, definida por la siguiente secuencia:

CCAGCTACCGACAGTAGTGTTC (SEQ ID NO: 41);

10 Según otro aspecto de la presente invención, proporciona un método para predecir si una muestra contiene *Escherichia coli* enterohemorrágica típica (ECEH), (que se definen en la presente memoria como cepas de *Escherichia coli* positivas para *stx* y *eae*), y/o la ECEH O104:H4 atípica que dio positivo para *stx* y negativo para *eae*. Las cepas ECEH típicas incluyen, en particular, los serotipos O157:H7, O145:H28, O103:H2, O111:H8, O121:H19, O26:H11 y O45:H2 de ECEH y sus derivados no móviles.

Dicho método comprende la detección del gen *espK* y de uno o más de los siguientes genes diana: *espV*, *ureD*, *Z2098*, *Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156* y *Z6065*.

15 Estas dianas génicas de *E. coli* corresponden a efectores tipo III no codificados por LEE derivados de varias islas O genómicas: OI-43, OI-44, OI-50, OI-57 y OI-71.

20 Las combinaciones de *espK* con uno o más de *espV*, *ureD*, *Z2098*, *Z1151*, *Z1153*, *Z1154*, *Z1155* y *Z1156*, se identificaron por los inventores entre varias combinaciones de marcadores de virulencia putativos, como las más predictivas de ECEH típico (cepas de *E. coli* *stx* y *eae* positivas), y en particular de la presencia de cepas ECEH de los serotipos O157:[H7], O145:[H28], O103:[H2], O111:[H8], O121:[H19], O26:[H11] o O45:[H2] de ECEH. La combinación de *espK* con *Z6065* es predictiva de la presencia de ECEH O104:H4 atípica.

Las combinaciones particularmente preferidas son las siguientes:

- *espK* con uno o más de *espV*, *ureD*, *Z2098*;
- *espK* con *Z6065*;
- 25 - *espK* con uno o más de *espV*, *ureD*, *Z2098* y con *Z6065*.

Según una realización particular, dicho método comprende realizar un ensayo de PCR sobre dicha muestra o ADN aislado de la misma con una combinación de cebadores que comprende un conjunto de cebadores derivados de *espK* y un conjunto de cebadores derivados de al menos uno de *espV*, *ureD*, *Z2098*, *Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156* y *Z6065*;

30 y detectar la presencia o la ausencia de un producto de amplificación para cada conjunto de cebadores de dicha combinación.

35 Según una realización preferida de este método, la combinación de cebadores además comprende un conjunto de cebadores derivados de *stx1* y un conjunto de cebadores derivados de *stx2*. Esto permite seleccionar muestras tanto para los genes *stx*, como marcadores de ECTS, como para los marcadores genéticos adicionales enumerados anteriormente, relacionados con los serotipos de ECTS prioritarios que están asociados con brotes y casos esporádicos de CH y SUH.

En contraste con los métodos de la técnica anterior, el método de la invención no necesita la detección del gen *eae*.

40 Los cebadores derivados de *espK*, *espV*, *ureD*, *Z2098*, *Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156*, *Z6065*, *stx1* o *stx2* y adecuados para uso en el ensayo de PCR de la invención, así como las sondas que permiten la detección de los productos de amplificación obtenidos con estos cebadores, pueden ser fácilmente diseñados por un experto en la técnica, a partir de las secuencias de estos genes disponibles en las bases de datos, por ejemplo dentro de la secuencia anotada de *Escherichia coli* O157:H7 (cepa EDL933) disponible en GenBank bajo el número de acceso AE005174.2.

45 A continuación se dan ejemplos no limitativos de conjuntos de cebadores preferidos para uso en este ensayo de PCR:

- un conjunto de cebadores que se dirigen a *espK*, definidos por las siguientes secuencias:

GCAGRCATCAAAGCGAAATCACACC (SEQ ID NO: 42)

TCGTTTGGTAACTGTGGCAGATACTC (SEQ ID NO: 43)

- un conjunto de cebadores que se dirigen a *espV*, definidos por las siguientes secuencias:

TCAGGTTCCCTCGTCTGATGCCGC (SEQ ID NO: 44)

CTGGTTCAGGCCTGGAGCAGTCC (SEQ ID NO: 45)

- un conjunto de cebadores que se dirigen a *ureD*, definidos por las siguientes secuencias:

5 GCAATAATTGACTCTGATTGCC (SEQ ID NO: 46)

GCTGCTGCGGTAAAATTTACT (SEQ ID NO: 47)

- un conjunto de cebadores que se dirigen a *Z2098*, definidos por las siguientes secuencias:

CTGAAAAGAGCCAGAACGTGC (SEQ ID NO: 48)

TGCCTAAGATCATTACCCGGAC (SEQ ID NO: 49)

10 - un conjunto de cebadores que se dirigen a *Z1153*, definidos por las siguientes secuencias:

CGATCATTGTGGGCATGTTATGCC (SEQ ID NO: 50)

CCTGAATTCACACGGTGATGCG (SEQ ID NO: 51)

- un conjunto de cebadores que se dirigen a *Z1154*, definidos por las siguientes secuencias:

GCCTTTTTATGTTTATTGCGTTG (SEQ ID NO: 52)

15 GTATAGTTTTAGCAATACCTTCTGC (SEQ ID NO: 53)

- un conjunto de cebadores que se dirigen a *Z1155*, definidas por las siguientes secuencias:

GATTGTGGCGATTAATGGGGG (SEQ ID NO: 54)

ACACCGATCTGGTCATTGGCG (SEQ ID NO: 55)

- un conjunto de cebadores que se dirigen a *Z1156*, definidos por las siguientes secuencias:

20 AAACGCCTTTAAAATCTGCGTCT (SEQ ID NO: 56)

TGCCGTGCGCACAGTCATAAG (SEQ ID NO: 57)

- un conjunto de cebadores que se dirigen a *Z1151*, definidos por las siguientes secuencias:

GCCCATGGCTCCACATCCTG (SEQ ID NO: 58)

CCAAAAAAGTTATGATGATTGCACTG (SEQ ID NO: 59)

25 - un conjunto de cebadores que se dirigen a *Z6065*, definidos por las siguientes secuencias:

GCACTGGCCCTTGTTGCTCAGGC (SEQ ID NO: 60)

GCTCTTCCAGTGAGAATGTCTTTCCGG (SEQ ID NO: 61)

- un conjunto de cebadores que se dirigen a *stx1* y *stx2*, definidos por las siguientes secuencias:

TTTGTYACTGTSACAGCWGAAGCYTTACG (SEQ ID NO: 62)

30 CCCCAGTTCARWGTRAGRTCMACRTC (SEQ ID NO: 63)

A continuación se dan ejemplos no limitativos de sondas para detectar los productos de amplificación:

- una sonda que permite la detección de productos de amplificación derivados de *espK*, definida por la siguiente secuencia:

ATTCAGATAGAAGAAGCGCGGGCCAG (SEQ ID NO: 64);

35 - una sonda que permite la detección de productos de amplificación derivados de *espV*, definida por la siguiente secuencia:

CTTGCAACACGTTACGCTGCCGAGTATT (SEQ ID NO: 65);

- una sonda que permite la detección de productos de amplificación derivados de *UreD*, definida por la siguiente secuencia:

TACGCTGATCACCATGCCTGGTGC (SEQ ID NO: 66);

- una sonda que permite la detección de productos de amplificación derivados de Z2098, definida por la siguiente secuencia:

5 TAACTGCTATACCTCCGCGCCG (SEQ ID NO: 67);

- una sonda que permite la detección de productos de amplificación derivada de Z1153, definida por la siguiente secuencia:

TGTAACACCCAGACGGTCAGCAACATG (SEQ ID NO: 68);

10 - una sonda que permite la detección de productos de amplificación derivada de Z1154, definida por la siguiente secuencia:

TCACTTCCAGTTTCTGGTGATGTTTTGAT (SEQ ID NO: 69);

- una sonda que permite la detección de productos de amplificación derivada de Z1155, definida por la siguiente secuencia:

TGGGTGAGTTAAAATATAAAGAACGATTGC (SEQ ID NO: 70);

15 - una sonda que permite la detección de productos de amplificación derivada de Z1156, definida por la siguiente secuencia:

TAAGATATTTTCTGACTTTCCGCATGCGCTT (SEQ ID NO: 71);

- una sonda que permite la detección de productos de amplificación derivada de Z1151, definida por la siguiente secuencia:

20 AAAGAGCCAGCGCAGAGCTGACCAG (SEQ ID NO: 72);

- una sonda que permite la detección de productos de amplificación derivada de Z6065, definida por la siguiente secuencia:

TTCGCTGGAAGCAGAGCCCGTGC (SEQ ID NO: 73);

25 - una sonda que permite la detección de productos de amplificación derivada de *stx1*, definida por la siguiente secuencia:

CTGGATGATCTCAGTGGGCGTTCTTATGTAA (SEQ ID NO: 74);

- una sonda que permite la detección de productos de amplificación derivada de *stx2*, definida por la siguiente secuencia:

TCGTCAGGCACTGTCTGAAACTGCTCC (SEQ ID NO: 75);

30 de forma ventajosa, la invención proporciona un método para predecir si una muestra contiene *Escherichia coli* enterohemorrágica (ECEH) típica de al menos uno de los serotipos O157:[H7], O145:[H28], O103:[H2], O111:[H8], O121:[H19], O26:[H11] y O45:[H2] de ECEH, y que además identifica el(los) serotipo(s) de dicha ECEH, en donde dicho método comprende:

35 - realizar un ensayo de PCR para evaluar si dicha muestra comprende o no ECEH de al menos uno de los serotipos O157:[H7], O145:[H28], O103:[H2], O111:[H8], O121:[H19], O26:[H11], O45:[H2] y O104:H4, como se describió anteriormente, y si los resultados de dicho ensayo de PCR son positivos,

- realizar un ensayo de PCR para identificar el(los) serotipo(s) de dicho ECEH, como se describió anteriormente.

40 Los ensayos de PCR de la invención se pueden utilizar para analizar cualquier muestra de una sustancia que potencialmente contenga ECEH, tal como muestras de alimentos, muestras de agua, muestras de suelo, etc.

45 Los ensayos de PCR de la invención se pueden llevar a cabo utilizando cualquier método adecuado para la amplificación de secuencias diana por PCR, utilizando cualquiera de las diversas enzimas naturales o modificadas disponibles para este fin. También se pueden utilizar métodos alternativos tales como amplificación basada en secuencias de ácidos nucleicos (NASBA, por sus siglas en inglés), ADN ramificado, amplificación por desplazamiento de cadena o el método de amplificación isotérmica mediada por bucle (LAMP, por sus siglas en inglés) (Compton 1991, Chang 1991, Walker et al. 1992, Notomi et al., 2000).

Los métodos particularmente preferidos son aquellos que implican amplificación por PCR en tiempo real como se describe por Ian M. Mackay en "Real-time PCR in Microbiology: from diagnosis to characterization" (2007) Caister Academic Press, Norfolk, UK.

5 La PCR en tiempo real, también llamada reacción en cadena de la polimerasa en tiempo real cuantitativa (qPCR, por sus siglas en inglés) o reacción en cadena de la polimerasa cinética, se utiliza para amplificar y cuantificar simultáneamente una molécula de ADN diana. Permite tanto la detección como la cuantificación (como número absoluto de copias o cantidad relativa cuando se normaliza para la entrada de ADN o genes normalizadores adicionales) de una secuencia específica en una muestra de ADN. El procedimiento sigue el principio general de la
10 reacción en cadena de la polimerasa; su característica clave es que el ADN amplificado se cuantifica a medida que se acumula en la reacción en tiempo real después de cada ciclo de amplificación (Mackay 2007). Dos métodos comunes de cuantificación son el uso de tintes fluorescentes que se intercalan con el ADN bicatenario, y las sondas de oligonucleótidos de ADN modificado que emiten fluorescencia cuando se hibridan con un ADN complementario (Mackay 2007). En la presente invención, los inventores han mostrado el segundo de estos dos métodos, pero el
15 otro método de cuantificación de productos de PCR basado en colorantes fluorescentes intercalados también está dentro del alcance de la presente invención.

Ejemplos no limitantes de marcadores fluorescentes adecuados incluyen 6-carboxifluoresceína (FAM), tetracloro-6-carboxifluoresceína (TET), 6-carboxi-X-rodamina (ROX). Ejemplos no limitativos de desactivadores adecuados para marcar sondas doblemente marcadas incluyen 6-carboxi-tetrametil-rodamina (TAMRA), DABCYL, desactivadores no fluorescentes tales como desactivadores de la familia Black Hole Quencher (BHQ), o que incluyen un grupo de
20 ligandos de unión al surco menor (MGB).

Cada uno de los ensayos de PCR de la invención se puede llevar a cabo realizando una reacción de PCR separada para cada secuencia diana a detectar (PCR simple). Sin embargo, en muchos casos se preferirá llevar a cabo la PCR múltiple, que permite la amplificación de varias secuencias diana en una única reacción. De forma ventajosa,
25 se puede utilizar una macromatriz, es decir, una estructura preformada tal como un sustrato sobre la que se han detectado los cebadores de ADN deseados. Tal macromatriz permite la realización rutinaria de ensayos de PCR múltiple descritos en la presente memoria. A modo de ejemplo, se puede utilizar la macromatriz GeneDisc® (Pall-GeneDisc Technology, Bruz, Francia) descrita, por ejemplo, por Beutin et al. (Beutin et al.2009) que permite la detección simultánea de múltiples dianas en microcámaras de reacción precargadas con los reactivos necesarios
30 para detectar y cuantificar las dianas requeridas.

Para garantizar que los resultados del ensayo son representativos del contenido real de la muestra, también puede comprender un control de amplificación negativo para garantizar que cualquiera de los productos detectados son verdaderos positivos y también un control de inhibición para garantizar que el ADN de la muestra se puede amplificar y, por lo tanto, que no se generan falsos negativos.

35 La invención también incluye los conjuntos de cebadores y las sondas definidos anteriormente, que permiten llevar a cabo los ensayos de PCR de la invención, así como kits que asocian estos conjuntos de cebadores y estas sondas, eventualmente asociados con reactivos para realizar una reacción de PCR. Estos conjuntos también pueden comprender instrucciones para realizar dicha reacción de amplificación. Los productos de amplificación que utilizan los cebadores de la invención también son parte de la invención.

40 Según una primera realización, un kit de la invención comprende una combinación de cebadores que comprende:

- un conjunto de cebadores definidos por las secuencias SEQ ID NO: 11 y SEQ ID NO: 12 y un conjunto de cebadores definidos por las secuencias SEQ ID NO: 13 y SEQ ID NO: 14, y/o un conjunto de cebadores definidos por las secuencias SEQ ID NO: 15 y SEQ ID NO: 16;

- un conjunto de cebadores definidos por las secuencias SEQ ID NO: 17 y SEQ ID NO: 18;

45 - un conjunto de cebadores definidos por las secuencias SEQ ID NO: 19, SEQ ID NO: 20, y SEQ ID NO: 21;

- un conjunto de cebadores definidos por las secuencias SEQ ID NO: 22 y SEQ ID NO: 23;

- un conjunto de cebadores definidos por las secuencias SEQ ID NO: 24 y SEQ ID NO: 25;

- un conjunto de cebadores definidos por las secuencias SEQ ID NO: 26 y SEQ ID NO: 27;

- un conjunto de cebadores definidos por las secuencias SEQ ID NO: 28 y SEQ ID NO: 29;

50 - un conjunto de cebadores definidos por las secuencias SEQ ID NO: 30 y SEQ ID NO: 31;

preferiblemente, dicho kit también comprende:

- 5 - una sonda que permite la detección de productos de amplificación derivados de las SEQ ID NO: 1 y SEQ ID NO: 2, y una sonda que permite la detección de productos de amplificación derivados de SEQ ID NO: 3, y/o una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 4, como se definió anteriormente;
- una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 5, como se definió anteriormente;
- una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 6, como se definió anteriormente;
- 10 - una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 7, como se definió anteriormente;
- una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 8, como se definió anteriormente;
- 15 - una sonda que permite la detección de productos de amplificación derivados de la SEQ ID NO: 9, como se definió anteriormente;
- dos sondas que permiten la detección de productos de amplificación derivados de la SEQ ID NO: 10, como se definió anteriormente;

Según una segunda realización, un kit de la invención comprende:

- un conjunto de cebadores derivados de *espK*, y
- 20 - uno o más conjuntos de cebadores seleccionados entre: un conjunto de cebadores derivados de *espV*, un conjunto de cebadores derivados de *ureD*, un conjunto de cebadores derivados de *Z2098*, un conjunto de cebadores derivados de *Z1151*, un conjunto de cebadores derivados de *Z1153*, un conjunto de cebadores derivados de *Z1154*, un conjunto de cebadores derivados de *Z1155*, un conjunto de cebadores derivados de *Z1156*, un conjunto de cebadores derivados de *Z6065*.
- 25 Preferiblemente, dicho kit también comprende una sonda que permite la detección de productos de amplificación derivados de *espK*, y una o más sondas seleccionadas entre: una sonda que permite la detección de productos de amplificación derivados de *espV*, una sonda que permite la detección de productos de amplificación derivados de *ureD*, o una sonda que permite la detección de productos de amplificación derivados de *Z2098*, una sonda que permite la detección de productos de amplificación derivados de *Z1151*, una sonda que permite la detección de productos de amplificación derivados de *Z1153*, una sonda que permite la detección de productos de amplificación derivados de *Z1154*, una sonda que permite la detección de productos de amplificación derivados de *Z1155*, una sonda que permite la detección de productos de amplificación derivados de *Z1156*, una sonda que permite la detección de productos de amplificación derivados de *Z6065*.
- 30

35 Los kits según la segunda realización descrita anteriormente pueden comprender además un conjunto de cebadores que se dirigen a *stx1* y un conjunto de cebadores que se dirigen a *stx2*, y preferiblemente una sonda que permite la detección de productos de amplificación derivados de *stx1*, y una sonda que permite la detección de productos de amplificación derivados de *stx2*.

40 Para un mejor entendimiento de la invención y para mostrar cómo se puede llevar a cabo la misma, a continuación se mostrarán solo a modo de ejemplo, realizaciones específicas, métodos y procedimientos según la presente invención.

Ejemplo 1: Identificación de secuencias de ADN derivadas de los loci CRISPR de *E. coli* para la identificación específica de *E. coli* enterohaemorrágica (ECEH)

Materiales y métodos

Cepas bacterianas

45 Las cepas de *E. coli* (n = 955) que se investigaron en busca de sus loci CRISPR mediante PCR en tiempo real de alto rendimiento se dan a conocer en la Tabla I a continuación.

Tabla I: Cepas de *E. coli*

ECEH* (n = 331)
O103:[H25] (n=6), O103:H2 (n=38), O111:H8 (n=49), O118:[H16] (n=3), O119:[H25] (n=4), O121:H19 (n=12), O123:H11, O127:H8s, O145, O145:[H28] (n=29), O156:H21, O156:H25 (n=10), O157:[H7] (n=75), O165:H25, O172:[H25], O172:H25, O172:NM, O177 (n=2), O177:[H25], O182:[H25], O26:[H11] (n=76), O3, O45:H2, O49:H16, O5 (n=8), O55, O76:H51, O84:H2, Ont:[H2], Or:H16, OX186:[H2]
ECEP (n = 344)
O100:[H25] (n=2), O102:H19, O103:H21, O103:H8, O108:H9 (n=6), O109:H25, O111, O111:B4, O111:H11, O111:H19 (n=3), O111:H2 (n=13), O111:H25 (n=2), O111:H47, O111:H9 (n=3), O113:H6 (n=2), O114:H2 (n=6), O114:H49 (n=5), O115:H38 (n=3), O117:H25, O117:H40b (n=3), O118:H5, O118:H8a (n=3), O119:[H25], O119:H2 (n=3), O119:H6 (n=4), O119:H8 (n=2), O119:H9, O119s:H2, O123/O4:H45 (n=2), O123:H25, O125:H6, O125ac:H6 (n=3), O126:H27, O126:H6, O127, O127:H19, O127:H40 (n=4), O127:H40b (n=2), O127:H6 (n=2), O128:[H2] (n=12), O128:H8, O128ac:H2, O142:H34, O142:H6 (n=3), O145:H34 (n=5), O15:H11, O15:H2 (n=3), O153:H14, O156, O156:[H8] (n=7), O156:H1 (n=2), O156:H25 (n=3), O157, O157:[H45] (n=2), O157:H16 (n=5), O157:H2, O157:H26 (n=2), O157:H39, O157:H45 (n=3), O177:H26, O186:[H45], O2:[H40] (n=2), O2:H40b, O2:H8, O21:H25, O22:H7, O26:[H11] (n=38), O26:H31, O26:H34, O28:H28 (n=4), O3:H40b, O3:H5, O3:H8a (n=3), O37:H10, O4:H16, O45, O45:H7, O45:H9, O49:[H10] (n=2), O49:H-, O5:H11, O51:H49 (n=3), O55:[H51], O55:[H7] (n=26), O55:H6 (n=5), O62:H9, O63:H6 (n=2), O66:H8/8a, O69:[H2], O69:H16 (n=2), O70/O86:H2, O70:H11 (n=5), O71:H40b, O76:H41, O76:H7 (n=5), O80:[H2] (n=3), O84:[H2], O86:[H34] (n=4), O86:H11 (n=2), O86:H40, O86:H8 (n=4), O86:H8a, O88:H8a, O89:[H2], O9:H10, OK8:H10, Ont:[H10], Ont:[H6], Ont:H11, Ont:H14, Ont:H2 (n=2), Ont:H21, Or:H40b, Or:H8a, Or:H9, OX177:H11 (n=2), OX177:H6
ECTS** (n = 160)
O100:NM (n=2), O101:H- (n=3), O104:H7, O105:H18, O109:H-, O110:H28, O111, O111:H10, O113:H4, O115:H18 (n=2), O116:H28, O117 (n=2), O117:H7 (n=2), O118:H12 (n=3), O125, O126, O126:H8, O128ab:H2, O130:H11, O136 (n=2), O138, O139, O139:H1, O141:[H4], O141ac, O146:H21, O146:H28 (n=2), O146:H8, O147, O149:H19, O15:H16, O153:H25 (n=3), O165:H11, O168:H8, O177/H7:H41, O171:H-, O171:H2, O172:H21, O174:[H21] (n=11), O174:H2, O174:H8 (n=4), O176:H-, O178:H19, O179:H8, O181:H49, O2:H25, O2:H27, O21:H21 (n=3), O22/O83, O22:H16 (n=2), O22:H8 (n=3), O23:H15, O3, O39:H48, O40:21, O40:H8, O46:H38 (n=2), O48:H21, O5, O5:[H19], O53, O6 (n=7), O6:H10 (n=2), O6:H34 (n=2), O68:H12, O73:H18, O74:H42, O75:H8, O76, O76:H19 (n=3), O77 (n=2), O79, O79:H48, O8:H10, O8:H19 (n=6), O8:H8, O85:H11, O86, O88:H25, O91 (n=6), O91:[H21] (n=5), O91:H10 (n=3), O91:H14 (n=2), O92, O107:H-, O92, O107:H48, O96:H19, Ont:H-, Ont:H7, Or:[H16], Or:H12, Or:H29, Or:H33, Or:H4, Or:H48, OX178:H19, OX185:H28, OX187:Hbev, OX3:H-, OX3:H2, OX3:H21, OX7:H16
<i>E. coli</i> no patógena (n = 120)
O103 (n=2), O103:H8, O104:H7, O110, O111:H12, O111:H21, O121:[H45], O126 (n=33), O126:H11, O126:H27 (n=3), O127 (n=8), O127:H10, O127:H21, O142 (n=8), O145:H2 (n=2), O150:H8, O153:H12, O156:H33, O156:H47, O156:H56, O157 (n=5), O157:H27, O180:H-, O26:H? (n=4), O26:H21/32, O26:H32 (n=6), O26:NM, O4:H5, O41:H7, O45:H7, O55 (n=8), O55:H19, O55:H21, O6:H4, O62:H30 (n=2), O8/O104:H10, O8/O104:H45, O86 (n=6), O86/O125ac, O86:H2, O86:H27, O88, O9:K9:H12, OX183:H18

Para cada serotipo, n=1 a menos que se indique lo contrario.

*Incluye los derivados de ECEH como se describe en (Bugarel et al., 2010). ** Incluye ECEH atípicas.

5

Las cepas de *E. coli* se dividieron en *E. coli* productora de toxina Shiga o ECTS (n = 160), *E. coli* enteropatógena o ECEP (n = 344), *E. coli* enterohemorrágica o ECEH (n = 331) y *E. coli* no patógena (n = 120). El tipo ECTS/ECEH se definió con la presencia de genes *stx*- y *eae*-. Las cepas ECEH se definieron como portadoras tanto de un gen *stx* (*stx1* y/o *stx2*) como de un gen *eae*, mientras que las cepas ECTS solo portaban *stx*. Las ECTS incluyeron cepas de *E. coli* *stx*-positivas y *eae*-negativas de los serotipos O91:[H21], O113:[H21], O104:[H21], también denominadas ECEH atípicas, que están implicadas con menos frecuencia en enfermedades hemorrágicas que otras ECEH, pero son una causa frecuente de diarrea. Los derivados *stx*-negativos de cepas ECEH se designaron como similares a ECEH y se definieron en base a su perfil genético *n/e*, subtipo y serotipo *eae* como se describe por Bugarel et al. (2010; 2011) excepto para las cepas similares a ECEH de serotipo O26:H11 que se identificaron en base a la presencia del gen *espK* y su forma alélica 2 del gen *arcA* (Bugarel et al., 2011). Las cepas ECEP se definieron como se describe por Bugarel et al. (2011). Las *E. coli* no patógenas se definieron como cepas *stx*- y *eae*- negativas.

15

Todas las cepas investigadas en este trabajo se identificaron para los antígenos O (LPS) y H (flagelares) de *E. coli* y se han caracterizado por los genes *stx*- y *eae*- como se dio a conocer anteriormente (Bugarel et al. 2010). Para el examen, las bacterias se cultivaron en colonias individuales sobre placas de Caldo Luria y se dejaron crecer durante

la noche a 37°C. Se recogió una colonia y se extrajo el ADN utilizando la matriz InstaGene (Bio-Rad Laboratories, Marnes La Coquette, Francia) antes de la prueba de PCR en tiempo real de alto rendimiento.

Secuenciación de ADN

- 5 Los loci CRISPR de las cepas de *E. coli* se amplificaron mediante PCR con los cebadores enumerados en la Tabla II. La secuenciación de ADN bicatenario de los amplicones CRISPR fue realizada por Eurofins MWG Operon (Courtaboeuf, Francia) utilizando los cebadores de secuenciación enumerados en la Tabla II.

Tabla II

Nombre del cebador	Secuencias (5'-3') del cebador directo y del cebador inverso	SEQ ID NO:	Número de acceso	Ubicación dentro de la secuencia
CRISPR-I-F	GGTGAAGGAGYTGGCGAAGGCGTC	76	AE005174	3665412-3665435
CRISPR-I-R	CCGGTGGATTTGGATGGGTTACG	77	AE005174	3665885-3665863
CRISPR-II-F	TGTGAACCTCTCTGGCATGGAG	78	AP010953	3786919-3786940
CRISPR-II-R	TAAAGTTGGTAGATTGTGACTGGC	79	AP010953	3787672-3787649

- 10 PCR en tiempo real de alto rendimiento

El LightCycler® 1536 (Roche, Meylan, Francia) se utilizó para realizar amplificaciones mediante PCR en tiempo real de alto rendimiento. Para la configuración de la PCR de las placas multipocillo LightCycler® 1536, se utilizó el inyector automático de líquido Bravo (Agilent Technologies, Massy, Francia) equipado con un enfriador y el sellador térmico de microplacas PlateLoc (Agilent Technologies). Las reacciones de PCR contenían 0,5 µl de muestra y 1 µl de mezcla maestra que contenía 1x RealTime ready DNA Probes Master (Roche) (que corresponde a 0,7x final), 300 nM de cada cebador y 300 nM de cada sonda (que corresponde a 200 nM final de cada uno). Las amplificaciones se realizaron utilizando sondas TaqMan® marcadas con FAM o HEX. Los cebadores y las sondas utilizados para amplificaciones mediante PCR se enumeran en la Tabla III. El sistema de PCR en tiempo real LightCycler® 1536 se utilizó con el siguiente perfil térmico: 95°C durante 1 min seguido de 35 ciclos de 95°C durante 0 s (incremento gradual de temperatura: 4,8°C/s) y 60°C durante 30 s (incremento gradual de temperatura: 2,5°C/s) y una etapa de enfriamiento final a 40°C durante 30 s. Las configuraciones del software eran sondas de hidrólisis /sondas UPL de doble color y control maestro.

Tabla III

	Secuencias (5'-3') del cebador directo, del cebador inverso y de la sonda	SEQ ID NO:	Secuencia diana
SP_O157_A	GAACACAAACCGAAACACACG	15	(SEQ ID NO:4)
	ATAAACCGTCACCAAAACAGTG	16	
	[FAM]-ACAAAACTGTCACCAAAAGTGTC-[BHQ1]	34	
SP_O157_B	GGAACACAAACCGAAACACA	11	(SEQ ID NO: 1 y 2)
	CTTAGTGTGTTCCCCGCGC	12	
	[HEX]-CGATCAATCCGAATATGAGCGGT-[BHQ1]	32	
SP_0157_C	GAACACTTTGGTGACAGTTTTTGT	13	(SEQ ID NO:3)
	CTTAGTGTGTTCCCCGCGC	14	
	[HEX]-CACTGTTTTGGTGACGGTTTATCC-[BHQ1]	33	
SP_O121	CGGGGAACACTACAGGAAAGAA	22	(SEQ ID NO: 7)
	GGCGGAATACAGGACGGGTGG	23	
	[HEX]-TCCGCCAACGGCGACAGGGG-[BHQ1]	37	

	Secuencias (5'-3') del cebador directo, del cebador inverso y de la sonda	SEQ ID NO:	Secuencia diana
SP_O45	GAGTCTATCAGCGACTACC	24	(SEQ ID NO: 8)
	AACCGCAGCTCGCAGCGC	25	
	[HEX]-TCGGAACGTGGCGCTATAGGTG-[BHQ1]	38	
SP_O145	GAACTTGAGCCCTGCCAGAA	17	(SEQ ID NO: 5)
	ACCGCATCTTTTCTACCTG	18	
	[HEX]-TGGGGCCTCTTTTGTACCCGG-[BHQ1]	35	
SP_O104	GGAACTCACCGAGCGCCG	26	(SEQ ID NO: 9)
	GCCTTTGCAGCGTCTTTCCGATC	27	
	[HEX]-CTGGGAGGCGTATCTCACGTTCCGGT-[BHQ1]	39	
SP_O26_C	ACAATCGTGTGTAATTCGCGG	28	(SEQ ID NO: 10)
	GATAAACCGTGGTACGGAACA	29	
	[HEX]-TGCTGTCTATATTTCCGACCAGTGTCC-[BHQ1]	40	
SP_O26_D	TGAAACCACTCGCGGCAGAT	30	(SEQ ID NO: 10)
	ATAAACCGATCTCCTCATCCTC	31	
	[HEX]-CCAGCTACCGACAGTAGTGTGTTCC-[BHQ1]	41	
SP_O111	GTGACCGCCTGTACACGC	19	(SEQ ID NO: 6)
	CGGATATTTGGCGTAATACC	20	
	CTGCCGCGAGTGGTTTCAC	21	
	[HEX]-TGTAATGGCTCACCGTTTATCCCC-[BHQ1]	36	

Resultados

Identificación de secuencias de ADN específicas que se dirigen a los loci CRISPR de ECEH O157:H7

- 5 La secuenciación de los loci CRISPR de varias cepas ECEH O157:[H7] ha demostrado el polimorfismo de este locus para este serotipo. Las secuencias características de los loci CRISPR de cepas de ECEH O157:[H7] se dan a conocer en SEQ ID NO: 1, 2, 3 y 4. En base a estas secuencias y al locus CRISPR de la cepa EDL933 (número de acceso AE005174), se diseñaron varios ensayos de PCR en tiempo real (SP_O157_A, SP_O157_B y SP_O157_C) para detectar ECEH O157:[H7]. La especificidad y sensibilidad de los ensayos se analizaron contra un panel de 955 cepas de *E. coli*, que incluía 75 cepas ECEH O157:[H7] (Tabla I). Los ensayos de PCR demostraron ser altamente sensibles y específicas para ECEH O157:[H7]. La sensibilidad de los ensayos oscilaba de 92,0% a 97,3% con solo pocas cepas de O157:[H7] que no fueron detectadas por cada ensayo. La especificidad de los ensayos de PCR fue alta, oscilando de 99,6 a 100%. El análisis de PCR SP_O157_B fue el único ensayo que dio reacción cruzada con muy pocas cepas del serogrupo O55. Combinando los ensayos de PCR SP_O157_B y SP_O157_C, se detectaron correctamente las 75 cepas ECEH O157:[H7] (100% de sensibilidad) y solo 3 cepas aisladas del serogrupo O55 presentaron reacción cruzada (99,6% de especificidad).

Identificación de secuencias de ADN específicas que se dirigen al locus CRISPR de ECEH O145:H28

- 20 El locus CRISPR de ECEH O145:[H28] se caracterizó (SEQ ID NO: 5) secuenciando uno de los dos loci CRISPR identificados en *E. coli*. Se diseñó un ensayo de PCR (SP_O145) a partir de esta secuencia CRISPR para dirigirse a ECEH O145:[H28]. Entre las 955 cepas de *E. coli* que se investigaron con este ensayo de PCR, solo las 29 cepas ECEH O145:[H28] y las 4 cepas ECEP O28:H28 dieron positivo. La sensibilidad y especificidad del ensayo de PCR SP_O145 fueron, respectivamente, de 100% y 99,5%.

Identificación de secuencias de ADN específicas que se dirigen al locus CRISPR de ECEH O111:H8

En base a la secuencia del locus CRISPR de ECEH O111:H8, (SEQ ID NO: 6), se diseñó un ensayo de PCR en tiempo real (SP_O111) para detectar ECEH O111:[H8]. La investigación de 980 cepas de *E. coli* mediante el ensayo de PCR SP_O111 dio resultados positivos para 47 cepas ECEH O111:[H8] de las 49 cepas O111:[H8] analizadas. Solo una cepa ECEP del serotipo O45:H7 dio positivo. La sensibilidad y especificidad de este ensayo de PCR fueron altas, 95,9% y 99,9% respectivamente.

Identificación de secuencias de ADN específicas dirigidas al locus CRISPR de ECEH O121:H19

En este estudio se secuenció el locus CRISPR de ECEH O121:[H19] (SEQ ID NO: 7). Se diseñó un ensayo de PCR (SP_O121) a partir de esta secuencia para dirigirse a ECEH O121:[H19]. Entre las 955 cepas de *E. coli* analizadas por el ensayo de PCR SP_O121, solo una cepa O104:H7 y las 12 cepas ECEH O121:[H19] dieron positivo, lo que demuestra que este ensayo de PCR fue altamente sensible (100%) y específico (99,9 %).

Identificación de secuencias de ADN específicas dirigidas a los loci CRISPR de ECEH O103:H2 y O45:H2

En base a la determinación de la secuencia del locus CRISPR de ECEH O45:[H2] (SEQ ID NO: 8) y de la secuencia del locus CRISPR de ECEH O103:H2, emitida por la cepa 12009 (número de acceso APO10958), se diseñó un ensayo de PCR (SP_O45) y dio positivo una cepa ECEH O45:H2 y las 38 cepas ECEH O103:H2 investigadas en este estudio. Por lo tanto, el ensayo de PCR SP_O45 ha demostrado una alta sensibilidad (100%) para ECEH O103:[H2] y O45:[H2]. Este ensayo tiene una especificidad de 98,6% cuando se ensayo sobre un panel grande de *E. coli*, que da solo reacciones cruzadas menores con pocas cepas de los siguientes serotipos: O118:H8, O128:[H2], O128:H8, O128:H2, O89:[H2], O46:H38, O8:H8, O142, O145:H2 y una cepa O103 que dio negativo para el flagelo H2.

Identificación de secuencias de ADN específicas que se dirigen al locus CRISPR de ECEH O104:H4

En este estudio se secuenció el locus CRISPR de ECEH O104:[H4] (SEQ ID NO: 9). Se diseñó un ensayo de PCR (SP_O104) a partir de esta secuencia para dirigirse a ECEH O104: [H4]. El ensayo de PCR dirigido al locus CRISPR de *E. coli* O104:H4 se evaluó sobre un panel de 1303 cepas de *E. coli* que incluía los 186 serogrupos O conocidos y 56 tipos de H. Este ensayo de PCR dio resultados positivos para las 48 cepas aisladas O104:H4 (que incluye una cepa aislada Or:H4), relacionadas con el brote que se produjo en mayo de 2011, y para una cepa aislada clínica O104:H4 notificada en 2001. Las 39 cepas de *E. coli* O104 que tienen tipos de H distintos de H4 dieron negativo. Las cepas de *E. coli* que portan un antígeno capsular K9 (O8:K9:H10, O8:K9:H45, O9:K9:H1, O9:K9:H12 y O9:K9:H51) que presentan reactividad cruzada mediante aglutinación con el suero anti-O104 dieron negativo. En definitiva, entre las demás cepas de *E. coli* que incluían los 186 serogrupos O conocidos y 56 tipos de H, solo 5 cepas aisladas pertenecientes a los serotipos Ont:H2, O43:H2, O141:H2 y O174:H2 presentaron reacción cruzada con los cebadores y sondas diseñados en el locus CRISPR de ECEH O104:H4. Por lo tanto, se ensayaron cepas adicionales O174:H2, O141:H2 y O43:H2 para CRISPR-O104. Tres de doce O174:H2 dieron positivo, así como 3/4 O43:H2 y 1/8 de O141:H2. En conjunto, los datos mostraron que este ensayo de PCR fue altamente sensible (100%) y específico (99.6%).

Identificación de secuencias de ADN específicas dirigidas al locus CRISPR de ECEH O26:H11

La secuenciación de los loci CRISPR de varias cepas ECEH O26:[H11] ha mostrado el polimorfismo de este locus para este serotipo. Se dio a conocer una secuencia característica de los loci CRISPR de ECEH O26:[H11] en la SEQ ID NO: 10. En base a estas secuencias y al locus CRISPR de la cepa ECEH O26:H11 11368 (números de acceso AP010953, NC_013361), se diseñaron dos ensayos de PCR en tiempo real (SP_O26_C y SP_O26_D) para detectar ECEH O26:[H11]. La especificidad y sensibilidad de los ensayos se analizaron contra un panel de 980 cepas de *E. coli*, que incluían 77 cepas de ECEH O26:[H11] y similares a ECEH O26:[H11]. Los dos ensayos de PCR demostraron ser sensibles y específicos para ECEH O26:[H11]. La sensibilidad del ensayo de PCR SP_O26_C fue 87%, mientras que la sensibilidad del ensayo de PCR SP_O26_D fue 90,9%. Sólo pocas cepas O26:[H11] no fueron detectadas por cada ensayo. La especificidad del ensayo de PCR SP_O26_C fue 98,7% (12 cepas presentaron reacción cruzada) mientras que la especificidad del ensayo de PCR SP_O26_D fue 98,1% (17 cepas presentaron reacción cruzada). Al combinar los ensayos de PCR SP_O26_C y SP_O26_D, solo 4 cepas similares a ECEH O26:H11 de las 77 cepas similares a ECEH y ECEH O26:[H11] no se detectaron (sensibilidad de 94,8%) y solo 26 *E. coli* presentaron reacción cruzada (especificidad de 97,1%).

Conclusión

Los resultados de este estudio se resumen en la Tabla IV a continuación.

Tabla IV: Sensibilidad y especificidad

Serotipo	Número	PCR	Sensibilidad	Especificidad	Reacción cruzada
O157:[H7] ^a	75	SP_O157_A	92,0%	100%	-
		SP_O157_B	97,3%	99,6%	O55:[H7] ^a , O55:[H7] (n=2) ^b
		SP_O157_C	94,7%	100%	-
		SP_O157_B+C	100%	99,6%	O55:[H7] ^a , O55:[H7] (n=2) ^b
O103:H2 ^a , O45:H2 ^a	38 1	SP_O45	100%	98,6%	O118:H8a (n=3) ^b , O128:[H2] ^b , O128:H8 ^b , O128ac:H2 ^b , O89:[H2] ^b , O46:H38 ^c , O8:H8 ^c , O103 ^d , O142 ^d , O145:H2 ^d
O111:H8 ^a	49	SP_O111	95,9%	99,9%	O45:H7 (n=1) ^b
O121:H19 ^a	12	SP_O121	100%	99,9%	O104:H7 ^d
O145:[H28] ^a	29	SP_O145	100%	99,5%	O28:H28 (n=4) ^b
O104:[H4] ^a	49	SP_O104	100%	99,6%	Ont:H2, O43:H2 (n=4), O141:H2 (n=2), y O174:H2 (n=4)
O26:[H11] ^a	77	SP_O26_C	87%	98,7%	O111:H11 ^b , O111:H47 ^b , O118:H16 (n=2) ^a , O118:H8a (n=3) ^b , O128:H8 ^b , O26:H11 ^b , O118:H2 ^a , O103:H11 ^a , O111 ^b
		SP_O26_D	90,9%	98,1%	O118:H16 (n=3) ^a , O123:H11 ^a , O26:H11 (n=9) ^b , O118:H2 (n=2) ^a , O86:H11 (n=2) ^b , O103:H11 ^a
		SP_O26_C+D	94,8%	97,1%	O111:H11 ^b , O111:H47 ^b , O118:H16 (n=4) ^a , O118:H8a (n=3) ^b , O123:H11 ^a , O128:H8 ^b , O26:H11 (n=10) ^b , O86:H11 (n=2) ^b , O118:H2 (n=2) ^a , O103:H11 ^a , O111 ^b

Para cada serotipo, n=1 a menos que se indique lo contrario.

^a)ECEH y similar a ECEH; ^b)ECEP; ^c)ECTS y ECEH atípica; ^d)*E. coli* no patógena

5

La secuenciación de los loci CRISPR de varias cepas ECEH ha mostrado la diversidad genética de las secuencias CRISPR emitidas a partir de ECEH asociadas con los casos clínicos más frecuentes en el mundo. El análisis de las secuencias espaciadoras localizadas entre las secuencias de repetición palindrómicas cortas de los loci CRISPR, permitió identificar marcadores genéticos útiles para detectar con alta sensibilidad y especificidad cepas de ECEH. En base a una estrategia de PCR en tiempo real de alto rendimiento, se investigó un panel muy grande de cepas de *E. coli*, que comprendía ECEH, ECEP, ECTS y *E. coli* no patógena con respecto a su contenido de loci CRISPR. En definitiva, se detectaron ECEH O145:H28 (n = 29), O103:H2 (n = 38), O121:H19 (n = 12), O104:H4 (n = 49) y O45:H2 (n = 1) con sensibilidad de 100% con cada ensayo de PCR que se dirigen a varias secuencias de CRISPR derivadas de estos serotipos de ECEH. Se detectó ECEH O157:[H7] (n = 75) con 100% de sensibilidad cuando se combinaron los ensayos de PCR SP_O157_B y SP_O157_C que se dirigen a dos secuencias diferentes de los loci CRISPR de ECEH O157. Se detectó ECEH O111:[H8] (n = 49) con sensibilidad de 95,9% (se detectaron 47/49 O111:[H8], solo dos no se detectaron). Al combinar los ensayos de PCR SP_O26_C y SP_O26_D que se dirigen a dos secuencias diferentes de los loci CRISPR de O26, se detectó ECEH O26:[H11] (n = 77) con sensibilidad de 94,8% (se detectaron 73/77 O26:[H11]; las únicas 4 cepas que no se detectaron fueron cepas similares a ECEH O26:H11).

20

Los ensayos de PCR desarrollados en este estudio para dirigirse a los loci CRISPR de ECEH asociadas con los casos clínicos más frecuentes del mundo también fueron muy específicos. Estos ensayos tuvieron una especificidad de 97,1% a 100% cuando se analizaron en un panel muy grande de cepas de *E. coli*, dando solo reacciones cruzadas muy pequeñas (Tabla IV).

Ejemplo 2: Identificación de marcadores genético para identificar *Escherichia coli* productora de toxina shiga (ECTS), asociada con alta virulencia para humanos

El amplio repertorio de efectores tipo III no codificados por LEE (Tobe et al., 2006; Creuzburg et al., 2011) y adhesinas (Spears et al., 2006; Cergole-Novella et al., 2007;) representa la fuente más probable de determinantes de virulencia de ECTS. Sin embargo, las dianas genéticas que respaldan mejor una estrategia de evaluación del riesgo molecular se tienen que definir todavía. El control de ECEH en los alimentos requiere, en particular, la selección de marcadores genéticos capaces de discriminar claramente cepas ECEH de ECEP.

En un intento de identificar tales factores, exploramos la idoneidad de ciertos genes *nle* derivados de las islas O genómicas OI-43, OI-44, OI-50, OI-57 y OI-71 como candidatos para distinguir las cepas ECTS que constituyen un riesgo severo para la salud humana de las cepas ECEP y ECTS que no están asociadas con enfermedades severas y epidémicas. Los dianas del gen de *E. coli* utilizadas para la amplificación por PCR en tiempo real se dan a conocer en la Tabla V a continuación.

Tabla V

Gen (nombre de ORF si es cromosómico) ^a	Proteína codificada o efector de familia	Soporte genético (elementos móviles) ^a
<i>ureD</i> (Z1142, Z1581)	Proteína UreD asociada a ureasa	OI-43 & OI-48
Z1151	Proteína hipotética	OI-43
Z1153	Proteína hipotética	OI-43
Z1154	Proteína de inmunidad a la colicina	OI-43
Z1155	Supuestas proteínas de membrana	OI-43
Z1156	Proteína hipotética	OI-43
<i>espV</i> (Z1387)	Efector de la familia AvrA	OI-44
<i>espK</i> (Z1829)	Repeticiones ricas en leucina	OI-50
Z2098	Proteína hipotética	OI-57
Z6065	Proteína hipotética	OI-71

a) La nomenclatura de los ORF y elementos móviles se refiere a la secuencia de *E. coli* O157:H7 EDL933 (GenBank AE005174)

1) Marcadores genéticos *espK*, Z1151, Z1153, Z1154, Z1155, Z1156 y Z6065.

La distribución de marcadores genéticos derivados de la OI-43 (Z1151, Z1153, Z1154, Z1155, Z1156), OI-50 (*espK*) y OI-71 (Z6065) se examinó entre diversos grupos patógenos de *E. coli* para evaluar su asociación con cepas ECTS con alta virulencia para humanos.

Materiales y métodos

Las 1252 cepas de *E. coli* investigadas en este estudio se dividieron en *E. coli* enterohemorrágica o ECEH (n = 466), *E. coli* enteropatógena o ECEP (n = 468), *E. coli* productora de toxina Shiga o ECTS (n = 179) y *E. coli* no patógena (n = 139), en base a la presencia de genes *stx*- y *eae*-. Las cepas ECTS solo portaban *stx*. Las cepas ECEP solo portaban *eae*. Las cepas de *E. coli* no patógenas (n = 139) se definieron como *stx* y *eae* negativas.

Se realizó una prueba de PCR en tiempo real de alto rendimiento como se describe en el Ejemplo 1 anterior.

Los cebadores y las sondas utilizados para las amplificaciones por PCR de los marcadores genéticos *espK*, Z1151, Z1153, Z1154, Z1155, Z1156 y Z6065 se enumeran en la Tabla VI. Los cebadores y las sondas para la detección de *stx1*, *stx2* y *eae* se describieron anteriormente (Bugarel et al. 2010). Las amplificaciones de los genes *stx1*, *stx2* y *eae* se utilizaron como controles internos y para fines de asignación grupal.

Tabla VI

	Secuencias (5'-3') del cebador directo, del cebador inverso y de la sonda	SEQ ID NO:
<i>espK</i> (1829)	GCAGRCATCAAAAGCGAAATCACACC	42
	TCGTTTGGTAACTGTGGCAGATACTC	43
	[6FAM]-ATTCAGATAGAAGAAGCGCGGGCCAG-[BHQ1]	64
Z1153	CGATCATTGTGGGCATGTTATGCC	50
	CCTGAATTCACACGGTGATGCG	51
	[6FAM]-TGTAACACCCAGACGGTCAGCAACATG-[BHQ1]	68
Z1154	GCCTTTTTATGTTTCATTATTGCGGTTG	52
	GTATAGTTTTAGCAATACCTTCCTGC	53
	[6FAM]-TCACTCCAGTTTCTGGTGATGTTTTGAT-[BHQ1]	69
Z1155	GATTGTGGCGATTAATGGGGG	54
	ACACCGATCTGGTCATTGGCG	55
	[6FAM]-TGGGTGAGGTTAAAATATAAAGAACGATTGC-[BHQ1]	70
Z1156	AAACGCCTTTAAATCTGCGTCT	56
	TGCCGTGCGCACAGTCATAAG	57
	[6FAM]-TAAGATATTTTCTGACTTTCCGCATGCGCTT-[BHQ1]	71
Z1151	GCCCATGGCTCCACATCCTG	58
	CCAAAAAAGTTATGATGATTGCACTG	59
	[6FAM]-AAAGAGCCAGCGCAGAGCTGACCAG-[BHQ1]	72
Z6065	GCACTGGCCCTTGTTGCTCAGGC	60
	GCTCTTCCAGTGAGAATGTCTTTCCGG	61
	[6FAM]-TTCGCTGGAAGCAGAGCCCGTGC-[BHQ1]	73

Resultados

- 5 Distribución de *espK*, Z1151, Z1153, Z1154, Z1155, Z1156 y Z6065 y combinación de los mismos entre los grupos patógenos de *E. coli*

10 La distribución de los diferentes marcadores genéticos *espK*, Z1151, Z1153, Z1154, Z1155, Z1156 y Z6065 entre los diferentes grupos patógenos de *E. coli* se muestra en la Tabla VII a continuación. En general, los marcadores genéticos investigados se detectaron principalmente en cepas ECEH con frecuencias que oscilan entre 51,9% (Z6065) y 90,8% (*espK*). Estos marcadores estaban menos asociados con cepas ECEP con frecuencias que oscilan entre 17,7% (Z1154) y 53,8% (Z1155) y raramente detectadas en ECTS (3,4 a 20,7%) y *E. coli* no patógena (3,6 a 9,4%).

15 Ninguno de los marcadores genéticos *espK*, Z1151, Z1153, Z1154, Z1155, Z1156 y Z6065 es, por sí solo, capaz de identificar de manera fiable todas las cepas ECEH. Sin embargo, cuando se combinó *espK* con marcadores genéticos de la OI-43 (Z1151, Z1153, Z1154, Z1155 y Z1156) u OI-71 (Z6065), la mayoría de las cepas ECEH se detectaron con frecuencias que oscilaban entre 95,5% (*espK*/Z6065) y 98,3% (*espK*/Z1155). Las mismas combinaciones detectaron cepas ECEP con frecuencias que oscilan entre 31,2% (*espK*/Z1156) y 61,8% (*espK*/Z1155), cepas ECTS con frecuencias de 6,7% a 23,5% y cepas de *E. coli* no patógenas con frecuencias entre 7,9% y 13,7%.

20

Tabla VII

Marcadores genéticos	ECEH (n=466)	ECEP (n=468)	ECTS (n=179)	EC (n=139)
Z1151	79,8%	20,3%	20,7%	7,9%
Z1153	89,3%	23,9%	12,3%	9,4%
Z1154	83,3%	17,7%	3,4%	3,6%
Z1155	79,4%	53,8%	16,8%	8,6%
Z1156	88,8%	18,8%	12,8%	6,5%
Z6065	51,9%	20,1%	5,0%	8,6%
<i>espK</i>	90,8%	28,0%	3,4%	5,0%
<i>espK/Z1151</i>	97,2%	34,0%	23,5%	12,2%
<i>espK/Z1153</i>	97,4%	35,7%	15,1%	13,7%
<i>espK/Z1154</i>	97,0%	31,8%	6,7%	7,9%
<i>espK/Z1155</i>	98,3%	61,8%	19,6%	12,9%
<i>espK/Z1156</i>	97,4%	31,2%	15,6%	10,1%
<i>espK/Z6065</i>	95,5%	36,8%	8,4%	13,7%

espK/Z1151 representan cepas que dan un resultado positivo para *espK* y/o Z1151; *espK/Z1153* representan cepas que dan un resultado positivo para *espK* y/o Z1153; *espK/Z1154* representan cepas que dan un resultado positivo para *espK* y/o Z1154; *espK/Z1155* representan cepas que dan un resultado positivo para *espK* y/o Z1155; *espK/Z1156* representan cepas que dan un resultado positivo para *espK* y/o Z1156; *espK/Z6065* representan cepas que dan un resultado positivo para *espK* y/o Z6065

Distribución de los marcadores genéticos en *E. coli* enterohemorrágica

La distribución de cada marcador genético *espK*, Z1151, Z1153, Z1154, Z1155, Z1156 y Z6065 fue significativamente diferente según los serotipos de ECEH (Tabla VIII). Curiosamente, el marcador genético Z6065 es el único marcador genético capaz de detectar ECEH O104:H4 (*stx* positivo, *eae* negativo, *aggR* positivo) que ha estado involucrado en el gran brote alemán en 2011.

Excepto Z1151 que no se detectó en ninguna ECEH O45:[H2] y Z6065 que estaba ausente en 18 de los 19 O121:[H19] probados (5,3%), todos los demás marcadores genéticos investigados se encontraron en cepas ECEH de los 7 serotipos principales, con frecuencias que oscilan entre 15,4% (prevalencia de Z6065 en O26:[H11]) y 100%.

Al combinar *espK* con uno de los siguientes marcadores genéticos de la OI-43: Z1151, Z1153, Z1154, Z1155 y Z1156, se detectaron la mayoría de las cepas ECEH de los 7 principales serotipos de ECEH. Por lo tanto, cualquiera que sea la combinación de marcadores genéticos utilizados, todas las cepas de ECEH de los 7 serotipos principales dieron positivo, con la excepción de 1 a 2 cepas ECEH O121:[H19] que dieron negativo con *espK/Z1154* y *espK/Z6065* respectivamente; una cepa de O103:[H2] que no se pudo detectar con *espK/Z1154* y 7 a 8 cepas de ECEH O26:[H11] que se detectaron negativas con todas las asociaciones probadas de marcadores genéticos. Por lo tanto, solo unas pocas cepas ECEH no reaccionaron con los marcadores genéticos probados aquí. Estas podrían ser cepas anómalas, no representativas para los tipos clásicos de ECEH. Considerar otros genes en estas cepas anecdóticas o secuenciar su genoma podría revelar más diferencias que aclaren las cosas con respecto a su estado. Deberíamos asumir, en principio, que no es necesariamente el caso que todos los miembros de un serotipo particular sean EHEC.

Curiosamente, otras cepas ECEH, con otros serotipos que aquellos de los 7 serotipos principales, fueron altamente detectadas con frecuencias que oscilan entre 87,5% y 95,5%. Este hallazgo indicó que las combinaciones probadas de los marcadores genéticos podrían detectar ECEH típica (cepas de *E. coli* tanto *stx* como *eae* positivas) con alta sensibilidad. La introducción del marcador genético Z6065 permite detectar además ECEH O104:H4 (*stx* positivo, *eae* negativo, *aggR* positivo) que ha estado involucrado en el gran brote alemán en 2011.

Tabla VIII

ES 2 659 469 T3

Marcadores genéticos	O26:H11 (n=117)	O45:H2 (n=19)	O103:H2 (n=61)	O111:H8 (n=33)	O121:H19 (n=19)	O145:H28 (n=31)	O157:H7 (n=98)	Otras ECEH (n=88)
Z1151	105/117 (89,7%)	0/19 (0%)	44/61 (72,1%)	33/33 (100%)	5/19 (26,3%)	30/31 (96,8%)	91/98 (92,9%)	64/88 (72,7%)
Z1153	107/117 (91,5%)	19/19 (100%)	48/61 (78,7%)	33/33 (100%)	18/19 (94,7%)	31/31 (100%)	91/98 (92,9%)	69/88 (78,4%)
Z1154	87/117 (74,4%)	19/19 (100%)	48/61 (78,7%)	31/33 (93,9%)	15/19 (78,9%)	29/31 (93,5%)	91/98 (92,9%)	68/88 (77,3%)
Z1155	75/117 64,1%	16/19 (84,2%)	41/61 (67,2%)	33/33 (100%)	14/19 (73,7%)	25/31 (80,6%)	97/98 (99,0%)	69/88 (78,4%)
Z1156	106/117 (90,6%)	19/19 (100%)	48/61 (78,7%)	33/33 (100%)	18/19 (94,7%)	31/31 (100%)	91/98 (92,9%)	68/88 (77,3%)
Z6065	18/117 (15,4%)	19/19 (100%)	59/61 (96,7%)	7/33 (21,2%)	1/19 (5,3%)	6/31 (19,4%)	85/98 (86,7%)	47/88 (53,4%)
espK	108/117 (92,3%)	19/19 (100%)	60/61 (98,4%)	33/33 (100%)	17/19 (89,5%)	31/31 (100%)	92/98 (93,9%)	63/88 (71,6%)
espK/Z1151	110/117 (94,0%)	19/19 (100%)	61/61 (100%)	33/33 (100%)	19/19 (100%)	31/31 (100%)	98/98 (100%)	82/88 (93,2%)
espK/Z1153	110/117 (94,0%)	19/19 (100%)	61/61 (100%)	33/33 (100%)	19/19 (100%)	31/31 (100%)	98/98 (100%)	83/88 (94,3%)
espK/Z1154	110/117 (94,0%)	19/19 (100%)	60/61 (98,4%)	33/33 (100%)	18/19 (94,7%)	31/31 (100%)	98/98 (100%)	83/88 (94,3%)
espK/Z1155	113/117 (96,6%)	19/19 (100%)	61/61 (100%)	33/33 (100%)	19/19 (100%)	31/31 (100%)	98/98 (100%)	84/88 (95,5%)
espK/Z1156	110/117 (94,0%)	19/19 (100%)	61/61 (100%)	33/33 (100%)	19/19 (100%)	31/31 (100%)	98/98 (100%)	83/88 (94,3%)
espK/Z6065	109/117 (93,2%)	19/19 (100%)	61/61 (100%)	33/33 (100%)	17/19 (89,5%)	31/31 (100%)	98/98 (100%)	77/88 (87,5%)

5 *espK/Z1151* representan cepas que dan un resultado positivo para *espK* y/o *Z1151*; *espK/Z1153* representan cepas que dan un resultado positivo para *espK* y/o *Z1153*; *espK/Z1154* representan cepas que dan un resultado positivo para *espK* y/o *Z1154*; *espK/Z1155* representan cepas que dan un resultado positivo para *espK* y/o *Z1155*; *espK/Z1156* representan cepas que dan un resultado positivo para *espK* y/o *Z1156*; *espK/Z6065* representan cepas que dan un resultado positivo para *espK* y/o *Z6065*

2) Marcadores genéticos *espK*, *espV*, *Z2098* y *UreD*

10 La producción de toxina Shiga (Stx) por *E. coli* enterohemorrágica (ECEH) es el rasgo de virulencia primario responsable de colitis hemorrágica (CH) y síndrome urémico hemolítico (SUH), pero muchas cepas de *E. coli* que producen Stx (ECTS) no causan CH ni SUH. Además de la capacidad de producir uno o más tipos de toxinas Shiga, las cepas ECTS asociadas con infecciones humanas portan otros factores que podrían utilizarse para distinguir cepas ECTS que constituyen un riesgo grave para la salud humana de cepas ECTS que no están asociadas con enfermedades graves y epidémicas. En un intento de identificar tales factores, exploramos la idoneidad de ciertos

genes *nle* derivados de las islas O genómicas OI-43, OI-44, OI-50 y OI-57 como candidatos para distinguir cepas ECTS que constituyen un riesgo grave para la salud humana de cepas ECEP y ECTS que no están asociadas con enfermedades graves ni epidémicas. Nos centramos en *ureD* (actividad ureasa) codificada por OI-43 y/o OI-48, *espK* (EspK) portado por OI-50, un locus implicado en la persistencia de ECEH O157:H7 en los intestinos de terneros inoculados vía oral (Vlisidou et al. 2006). Además, nos centramos en Z2098, una secuencia derivada de OI-57, una isla genómica que puede estar asociada con una mayor virulencia de cepas ECTS en humanos (Coombes et al., 2008; Imamovic et al, 2010; Bugarel et al., 2011). La secuenciación del genoma de cepas ECEH (ECEH O157:H7, O111, O103 y O26) también ha mostrado otros marcadores genéticos, tales como *espV*, cuyo papel en la enfermedad no se ha evaluado. Este gen está localizado en OI-44 de ECEH O157:H7, pero su prevalencia en otros grupos patógenos de *E. coli* aún no se ha documentado. En este estudio, evaluamos la distribución de *ureD*, *espV*, *espK* y Z2098 en diversos grupos patógenos de *E. coli* para evaluar su asociación con cepas ECTS con alta virulencia para humanos y para probar su idoneidad para distinguir claramente ECEH de otros grupos patógenos de *E. coli*.

15 Materiales y métodos

Las cepas de *E. coli* (n=1100) utilizadas en este estudio fueron principalmente las descritas en los estudios anteriores. Las cepas tipo ECEH (n=340) y se definieron en presencia de genes *stx*- y *eae*-. Las cepas ECTS (n=193) portaron solo *stx*. Las cepas ECEP (n=392) portaron solo *eae*. La *E. coli* no patógena (n=175) se definió como cepas *stx*- y *eae*- negativas. El cultivo de bacterias y la preparación de ADN se realizó como se describió anteriormente.

También se realizaron ampliaciones por PCR en tiempo real de alto rendimiento como se describió anteriormente.

Los cebadores y las sondas TaqMan® marcadas con FAM utilizadas para ampliaciones por PCR de *stx1*, *stx2* y *eae* se describieron previamente (Bugarel et al. 2010). Los cebadores y las sondas utilizados para dirigirse a *ureD*, *espK*, Z2098 y *espV* se enumeran en la Tabla IX a continuación.

25

Tabla IX

Gen diana ^a	Secuencias (5'-3') del cebador directo, del cebador inverso y de la sonda	SEQ ID NO:	Localización dentro de la secuencia AE005174
<i>espK</i> (Z1829)	GCAGRCATCAAAGCGAAATCACACC	42	1673422-1673397
	TCGTTTGGTAACTGTGGCAGATACTC	43	1673312-1673338
	[6FAM]-ATTCAGATAGAAGAAGCGCGGGCCAG-[BHQ]	64	1673395 -16673370
<i>espV</i> (Z1387)	TCAGGTTCTCGTCTGATGCCGC	44	1295446-1295424
	CTGGTTCAGGCCTGGAGCAGTCC	45	1295360-1295382
	[6FAM]-CTTGCAACACGTTACGCTGCCGAGTATT-[BHQ]	65	1295422-1295395
<i>ureD</i> (Z1142)	GCAATAATTGACTCTGATTGCC	46	1078824-1078845
	GCTGCTGCGGTAAAATTTACT	47	1078892-1078872
	TACGCTGATCACCATGCCTGGTGC-[BHQ]	66	1078847-1078870
Z2098	CTGAAAAGAGCCAGAACGTGC	48	1888173-1888193
	TGCCTAAGATCATTACCCGGAC	49	1888308-1888287
	[6FAM]TAACTGCTATACCTCCGCGCCG-[BHQ]	67	1888286-1888265

a) Numeración como en EDL933

Resultados

Distribución de *ureD*, *espV*, *espK* y Z2098 y combinación de los mismos entre grupos patógenos de *E. coli*

La distribución de los marcadores genéticos *ureD*, *espV*, *espK* y Z2098 entre los diferentes grupos patógenos de *E. coli* se muestra en la Tabla X. En general, los marcadores genéticos investigados se detectaron principalmente en cepas ECEH con frecuencias que oscilan entre 84,4% (*espV*) y 92,4% (*espK*). Estos marcadores estaban menos asociados con cepas ECEP con frecuencias que oscilan entre 18,1% (*ureD*) y 45,2% (*espV*) y raramente detectadas en ECTS (0,5 a 3,6%) y *E. coli* no patógenas (0,6 a 2,9%). En general, observamos que el 26,5% de las cepas ECEP que dieron positivo para al menos uno de los marcadores genéticos investigados pertenecían a los 7 principales serotipos de ECEH. Por lo tanto, es digno de mención que 57/113 cepas ECEP que son positivas para

35

5 *espK* pertenecían a los 7 principales serotipos de ECEH. Asimismo, 59/177 cepas ECEP positivas para *espV* pertenecían a los 7 principales serotipos de ECEH. También es notable que 68/91 cepas ECEP positivas para Z2098 y 58/71 cepas ECEP positivas para *ureD* también pertenecían a los 7 principales serotipos de ECEH. Curiosamente, otras cepas ECEP que tienen un serotipo de ECEH conocido como O55:H7, O103:H25 y O156:H25 también se detectaron positivas para al menos uno de estos marcadores genéticos (datos no mostrados). Estos hallazgos indicarían que tales cepas aisladas podrían ser *Stx* negativas derivadas de ECEH que también se designan como cepas similares a ECEH (Bugarel et al., 2011). Supusimos que estas cepas aisladas eran derivadas de ECEH según sus serotipos y el contenido de genes *nle*, pero también podrían ser cepas ECEP que aún no podemos discriminar de las derivadas de ECEH. Investigaciones adicionales que utilizan la secuenciación completa del genoma podrían aclarar la designación exacta de estas cepas en el futuro.

15 Ninguno de los marcadores genéticos *ureD*, *espV*, *espK* y Z2098 es, por sí solo, capaz de identificar de manera fiable todas las cepas ECEH. Se exploraron las combinaciones de los marcadores genéticos para identificar aquellos que detectan ECEH con la mejor especificidad. Los resultados se presentan en la Tabla X. En combinación, esos marcadores genéticos estaban altamente asociados con ECEH con frecuencias que oscilaban entre 97,9% (*espK/Z2098*) y 98,8% (*espK/ureD*). Las mismas combinaciones detectadas de cepas ECEP con frecuencias que oscilaban entre 33,4% (*espK/ureD*) y 54,1% (*espK/espV*), cepas ECTS con frecuencias de 1,6% a 3,6% y cepas de *E. coli* no patógenas con frecuencia entre 1,1% y 3,4%.

Tabla X

Marcadores genéticos	ECEH (n=340)	ECEP (n=392)	ECTS (n=193)	EC (n=175)
<i>espK</i>	92,4%	28,8%	0,5%	1,1%
<i>ureD</i>	89,4%	18,1%	3,1%	2,9%
Z2098	87,4%	23,2%	3,6%	1,1%
<i>espV</i>	84,4%	45,2%	1,6%	0,6%
<i>espK/espV</i>	98,5%	54,1%	1,6%	1,1%
<i>espK/ureD</i>	98,8%	33,4%	3,6%	3,4%
<i>espK/Z2098</i>	97,9%	36,7%	3,6%	2,3%

20 *espK/espV* representan cepas que dan un resultado positivo para *espK* y/o *espV*; *espK/ureD* representan cepas que dan un resultado positivo para *espK* y/o *ureD*; *espK/Z2098* representan cepas que dan un resultado positivo para Z2098 y/o *espK*

Distribución de *ureD*, *espV*, *espK*, *espN*, Z2098 y *espM1* y combinación de los mismos entre los serotipos de ECEH

25 La distribución de cada marcador genético *ureD*, *espV*, *espK* y Z2098 fue significativamente diferente según los serotipos de ECEH. La distribución de cada marcador genético en varios serogrupos de ECEH se notifica en la Tabla XI. Excepto *espV* que no se detectó en ningún ECEH O45:[H2], todos los demás marcadores genéticos investigados se encontraron altamente prevalentes en cepas ECEH de los 7 serotipos principales, con frecuencias que oscilan entre 71,4% (prevalencia de *ureD* en O103:[H2]) y 100%.

Tabla XI

Marca dores genétic os	7 seroti pos de ECEH principa les	O103: H2	O111: H8	O121: H19	O145: H28	O157: H7	O26: H11	O45: H2	Otra ECEH (nueva ECEH emer gente) ^a	ECEH total
Z2098	250/277 (90,3%)	49/49 (100%)	47/51 (92,2%)	17/20 (85,0%)	30/30 (100%)	49/66 (74,2%)	44/44 (100%)	14/17 (82,4%)	47/63 (74,6%)	297/340 (87,4%)
<i>espK</i>	269/277 (97,1%)	48/49 (98,0%)	51/51 (100%)	19/20 (95,0%)	29/30 (96,7%)	61/66 (92,4%)	43/44 (97,7%)	17/17 (100%)	45/63 (71,4%)	314/340 (92,4%)

<i>espV</i>	248/277 (89,5%)	48/49 (98,0%)	51/51 (100%)	20/20 (100%)	30/30 (100%)	65/66 (98,5%)	34/44 (77,3%)	0/17 (0%)	39/63 (61,9%)	287/340 (84,4%)
<i>ureD</i>	257/277 (92,8%)	35/49 (71,4%)	51/51 (100%)	16/20 (80,0%)	30/30 (100%)	64/66 (97,0%)	44/44 (100%)	17/17 (100%)	47/63 (74,6%)	304/340 (89,4%)

a) O103:[H25] (n=2), O118:[H16] (n=4), O118:H2, O119:[H25] (n=5), O123:H11, O127:H8s, O145, O145:[H25] (n=5), O156:H21, O156:H25 (n=11), O165:H25 (n=2), O172:[H25] (n=2), O172:NM, O177 (n=2), O177:[H25], O182:[H25], O3, O49:H16, O5 (n=11), O55:[H7] (n=2), O76:H51, O84:H2, Ont:[H2], Ont:H25 (n=2), Or:H16, OX186:[H2].

5 La detección de los 7 principales serotipos de ECEH basada en diferentes combinaciones de estos marcadores genéticos se dan a conocer en la Tabla XII. La detección de *espK* y/o *Z2098* permitió detectar la mayoría de los serotipos de ECEH asociados con infecciones humanas. Por lo tanto, todas las cepas ECEH O111:[H8], O26:[H11], O45:[H2], O103:[H2] y O145:[H28] dieron un resultado positivo para *espK* y/o *Z2098*, mientras que el 97,0% de O157:[H7] y el 95% de O121:[H19] dieron positivo. La asociación de *espK* con *espV* o *ureD* también permitió

10 detectar la mayoría de las cepas de los 7 principales serotipos de ECEH. Por lo tanto, todas las cepas de los serotipos O157:[H7], O145:[H28], O111:[H8], O103:[H2], O45:[H2] y O121:[H19] dieron resultados positivos para *espK* y/o *espV*, y el 97,7% de O26:[H11] dio un resultado positivo para *espK* y/o *espV*. Los datos fueron muy similares al probar *espK* en asociación con *ureD*. En ese caso, todas las cepas de los 7 principales serotipos de ECEH dieron un resultado positivo para *espK* y/o *ureD*.

15 Tabla XII

Asociación de genes	7 serotipos de ECEH principales	O103:H2	O111:H8	O121:H19	O145:H28	O157:H7	O26:H11	O45:H2	Otra ECEH (nueva ECEH emergente) ^a	ECEH total
<i>espK/espV</i>	276/277 (99,6%)	49/49 (100%)	51/51 (100%)	20/20 (100%)	30/30 (100%)	66/66 (100%)	43/44 (97,7%)	17/17 (100%)	59/63 (93,7%)	335/340 (98,5%)
<i>espK/ureD</i>	277/277 (100%)	49/49 (100%)	51/51 (100%)	20/20 (100%)	30/30 (100%)	66/66 (100%)	44/44 (100%)	17/17 (100%)	59/63 (93,7%)	336/340 (98,8%)
<i>espK/Z2098</i>	275/277 (99,3%)	49/49 (100%)	51/51 (100%)	19/20 (95,0%)	30/30 (100%)	65/66 (98,5%)	44/44 (100%)	17/17 (100%)	59/63 (93,7%)	334/340 (98,2%)

a) O103:[H25] (n=2), O118:[H16] (n=4), O118:H2, O119:[H25] (n=5), O123:H11, O127:H8s, O145, O145:[H25] (n=5), O156:H21, O156:H25 (n=11), O165:H25 (n=2), O172:[H25] (n=2), O172:NM, O177 (n=2), O177:[H25], O182:[H25], O3, O49:H16, O5 (n=11), O55:[H7] (n=2), O76:H51, O84:H2, Ont:[H2], Ont:H25 (n=2), Or:H16, OX186:[H2].

20 *espK/espV* representan cepas que dan un resultado positivo para *espK* y/o *espV*; *espK/ureD* representan cepas que dan un resultado positivo para *espK* y/o *ureD*; *espK/Z2098* representan cepas que dan un resultado positivo para *Z2098* y/o *espK*

3) Compendio

25 Los estudios anteriores permitieron seleccionar los marcadores genéticos *Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156*, *Z6065*, *ureD*, *espV*, *espK* y *Z2098* útiles para detectar cepas ECEH típicas y, en particular, las pertenecientes a los siete serotipos de ECEH principales notificados en todo el mundo en infecciones humanas. La distribución de estos marcadores genéticos diferentes se ha investigado entre los diferentes grupos patógenos de *E. coli*, que permite diseñar subcombinaciones óptimas de estos marcadores. Los resultados de estos estudios se resumen a continuación.

30 Los marcadores genéticos *ureD*, *espV*, *espK*, *Z2098*, *Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156* y *Z6065* se detectaron a diferentes frecuencias entre los serotipos de ECEH. Exploramos las diversas asociaciones de estos marcadores genéticos para buscar las mejores combinaciones de marcadores que proporcionan la mayor especificidad y

5 sensibilidad para detectar ECEH. La asociación del marcador genético *espK* con uno de los otros nueve marcadores genéticos permite detectar la mayoría de las cepas ECEH típicas y, en particular, las pertenecientes a los 7 principales serotipos de ECEH. Los marcadores genéticos *espV*, *ureD* y *Z2098* se mostraron los mejores candidatos para ser combinados con *espK* para detectar ECEH. Tomados individualmente no fueron capaces de detectar todas las cepas de los 7 principales serotipos de ECEH, mientras que en asociación detectaron de 99,3% a 100% de las 7 principales cepas ECEH. La asociación de *espK* con *espV*, *ureD* o *Z2098* demostró ser la mejor combinación para una detección más específica y sensible de cepas ECEH. Por lo tanto, se observó un resultado positivo para *espK* y/o *espV* en 99,6% de cepas ECEH pertenecientes a los siete principales serotipos de ECEH notificados en todo el mundo en infecciones humanas (solo una cepa aislada ECEH O26:H11 dio negativo). Además, el 93,7% de cepas ECEH con serotipos distintos a los de los 7 principales serotipos dieron positivo para *espK* y/o *espV*. En definitiva, solo un subconjunto (54,1%) de cepas ECEP dieron positivo para *espK* y/o *espV*. La mayoría de las cepas ECTS y *E. coli* no virulentas se detectaron negativas con *espK* y *espV*. Otra estrategia interesante fue asociar *espK* con *Z2098*. Esta combinación de marcadores genéticos dio como resultado la detección de 99,3% de cepas ECEH pertenecientes a los siete principales serotipos de ECEH y de 93,7% de cepas ECEH con serotipos diferentes de los 7 principales serotipos. La detección de *espK* y/o *Z2098* se notificó solo para 36,7% de ECEP, 3,6% de ECTS y 2,3% de cepas no patógenas de *E. coli*. La mejor estrategia para detectar ECEH con la mayor especificidad y sensibilidad fue combinar *espK* con *ureD*. Esta asociación permitió detectar 100% de ECEH de los 7 principales serotipos y 93,7% de cepas ECEH con otros serotipos. La detección de *espK* y/o *ureD* también se notificó para solo 33,4% de ECEP, 3,6% de ECTS y 3,4% de cepas no patógenas de *E. coli*.

Estos hallazgos mostraron que al combinar la detección de *espK* con *espV*, *ureD* o *Z2098* es una estrategia altamente sensible y específica para identificar con $\geq 99\%$ de confianza los serotipos de ECEH relacionados con los casos clínicos más frecuentes en el mundo. La detección de estos marcadores genéticos en combinación con *stx* en muestras complejas (alimentos o muestras fecales) proporcionaría un diagnóstico más dirigido a ECEH que al combinar solo *stx* y *eae*. Curiosamente, la introducción de *Z6065* en el esquema de detección permite detectar el EHEC O104:H4 atípico que estuvo involucrado en el brote grave y más grande de STEC que ocurrió en Europa. Dada la rapidez de estos ensayos de PCR, esta estrategia debería tener un impacto importante en las investigaciones de vigilancia y brotes de los 7 principales ECEH y es probable que sea beneficioso para la salud pública. Además, la detección de estos conjuntos de marcadores genéticos en 93,7% de cepas ECEH que tienen serotipos diferentes de los 7 principales serotipos de ECEH puede ser útil para identificar nuevas cepas ECEH emergentes.

Conclusión

Utilizamos una estrategia de PCR de alto rendimiento para explorar el viruloma de diferentes grupos patógenos de *E. coli* en un intento de identificar rasgos genéticos que caracterizarían las cepas ECTS patógenas. Se investigó la distribución de diez marcadores genéticos (*Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156*, *Z6065*, *ureD*, *espV*, *espK* y *Z2098*) en un gran panel de *E. coli* que comprende cepas ECEH, ECEP, ECTS y de *E. coli* no patógenas. La distribución de estos marcadores genéticos varió entre los grupos patógenos de *E. coli* y según los serotipos.

En general, las asociaciones de *espK* con los otros nueve genes (*Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156*, *Z6065*, *ureD*, *espV* y *Z2098*) mostraron las mejores combinaciones para detectar cepas de ECEH pertenecientes a los siete principales serotipos de ECEH notificados en todo el mundo en infecciones humanas. Estos hallazgos mostraron que al utilizar estas combinaciones relevantes de genes, la mayoría de las cepas ECEH dieron positivo, mientras que solo un subconjunto de las cepas ECEP presentaba reacción cruzada. Además, solo las cepas ECTS muy pequeñas y de *E. coli* no virulentas presentaron reacción cruzada al utilizar tal estrategia. Además de la detección de cepas ECEH típicas, la combinación *espK/Z6065* permite detectar la ECEH O104:H4 atípica (*stx* positivo, *eae* negativo, *aggR* positivo) que estuvo involucrada en la epidemia más grande de CH y SUH que se produjo en Europa en 2011.

Bibliografía

- 5 Beutin, L., S. Jahn, and P. Fach. 2009. Evaluation of the 'GeneDisc' real-time PCR system for detection of enterohaemorrhagic *Escherichia coli* (ECEH) O26, O103, O111, O145 and O157 strains according to their virulence markers and their O- and H-antigen-associated genes. *J. Appl. Microbiol.* 106(4): 1122-1132.
- Bugarel M, Martin A, Fach P, Beutin L. 2011. Virulence gene profiling of enterohemorrhagic (ECEH) and enteropathogenic (ECEP) *Escherichia coli* strains: a basis for molecular risk assessment of typical and atypical ECEP strains. *BMC Microbiol.* Jun 21 ;11:142.
- 10 Bugarel M, Beutin L, Scheutz F, Loukiadis E, Fach P. 2011. Identification of genetic markers for differentiation of Shiga toxin-producing, enteropathogenic, and avirulent strains of *Escherichia coli* O26. *Appl Environ Microbiol.* Apr;77(7):2275-81.
- Bugarel M, Beutin L, Martin A, Gill A, Fach P. 2010. Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (ECTS) seropathotypes associated with Hemorrhagic Colitis and Hemolytic Uremic Syndrome in humans. *Int J Food Microbiol.* Sep 1 ;142(3):318-29.
- 15 Bugarel M, Beutin L, Fach P. 2010. Low-density microarray targeting non-locus of enterocyte effacement effectors (nle genes) and major virulence factors of Shiga toxin-producing *Escherichia coli* (ECTS): a new approach for molecular risk assessment of ECTS isolates. *Appl Environ Microbiol.* Jan;76(1):203-11.
- Cergole-Novella MC, Nishimura LS, Dos Santos LF, Irino K, Vaz TM, Bergamini AM, Guth BE. 2007. Distribution of virulence profiles related to new toxins and putative adhesins in Shiga toxin-producing *Escherichia coli* isolated from diverse sources in Brazil. *FEMS Microbiol Lett.* Sep;274(2):329-34. Epub 2007 Jul 25.
- 20 Chang, C. 1991 "Branched DNA Amplification Multimers for the Sensitive, Direct Detection of Human Hepatitis Viruses," *Nucleic Acids Symposium Series*, no. 24: 197-200.
- Compton, J. 1991 "Nucleic Acid Sequence-Based Amplification," *Nature* 350, no. 6313: 91- 92.
- 25 Coombes BK, Wickham ME, Mascarenhas M, Gruenheid S, Finlay BB, Karmali MA. 2008. Molecular analysis as an aid to assess the public health risk of non-O157 Shiga toxin-producing *Escherichia coli* strains. *Appl Environ Microbiol.* Apr;74(7):2153-60.
- Creuzburg K, Middendorf B, Mellmann A, Martaler T, Holz C, Fruth A, Karch H, Schmidt H. 2011. Evolutionary analysis and distribution of type III effector genes in pathogenic *Escherichia coli* from human, animal and food sources. *Environ Microbiol.* Feb;13(2):439-52.
- 30 Frank C, Werber D, Cramer JP, Askar M, Faber M, an der Heiden M, Bernard H, Fruth A, Prager R, Spode A, Wadi M, Zoufaly A, Jordan S, Kemper MJ, Follin P, Muller L, King LA, Rosner B, Buchholz U, Stark K, Krause G; SUH Investigation Team. Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. *N Engl J Med.* 2011 Nov 10;365(19): 1771-80
- 35 Gault G, Weill FX, Mariani-Kurkdjian P, Jourdan-da Silva N, King L, Aldabe B, Charron M, Ong N, Castor C, Mace M, Bingen E, Noel H, Vaillant V, Bone A, Vendrely B, Delmas Y, Combe C, Bercion R, d'Andigne E, Desjardin M, de Valk H, Rolland P. Outbreak of haemolytic uraemic syndrome and bloody diarrhoea due to *Escherichia coli* O104:H4, southwest France, June 2011. *Euro Surveill.* 2011 Jun 30;16(26).
- 40 Imamovic L, Tozzoli R, Michelacci V, Minelli F, Marziano ML, Caprioli A, Morabito S. 2010. OI-57, a genomic island of *Escherichia coli* O157, is present in other seropathotypes of Shiga toxin-producing *E. coli* associated with severe human disease. *Infect Immun.* Nov;78(11):4697-704.
- Karmali, M. A., M. Mascarenhas, S. Shen, K. Ziebell, S. Johnson, R. Reid-Smith, J. Isaac-Renton, C. Clark, K. Rahn, and J. B. Kaper. 2003. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. *J. Clin. Microbiol.* 41: 4930-4940.
- 45 Levine, M. M. 1987. *Escherichia coli* That Cause Diarrhea - Enterotoxigenic, Enteropathogenic, Enteroinvasive, Enterohemorrhagic, and Enteroadherent. *J. Infect. Dis.* 155: 377-389.
- Mackay, I. 2007. *Real-time PCR in Microbiology, from diagnosis to characterization.* Caister Academic Press, Norfolk, UK.
- Nataro, J. P. and J. B. Kaper. 1998. Diarrheagenic *Escherichia coli*. *Clinical Microbiol. Rev.* 11: 142-201.
- 50 Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N., and Hase, T. 2000 Loop-Mediated Isothermal Amplification of DNA. *Nucleic Acids Research* 28, no. 12: E63.

Scheutz F, Nielsen EM, Frimodt-Møller J, Boisen N, Morabito S, Tozzoli R, Nataro JP, Caprioli A. Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing *Escherichia coli* O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. *Euro Surveill.* 2011 Jun 16; 16(24).

- 5 Spears KJ, Roe AJ, Gaily DL. 2006. A comparison of enteropathogenic and enterohaemorrhagic *Escherichia coli* pathogenesis. *FEMS Microbiol Lett.* Feb;255(2):187-202.

Struelens MJ, Palm D, Takkinen J. Enteroaggregative, Shiga toxin-producing *Escherichia coli* O104:H4 outbreak: new microbiological findings boost coordinated investigations by European public health laboratories. *Euro Surveill.* 2011 Jun 16; 16(24).

- 10 Tobe T, Beatson SA, Taniguchi H, Abe H, Bailey CM, Fivian A, Younis R, Matthews S, Marches O, Frankel G, Hayashi T, Pallen MJ. 2006. An extensive repertoire of type III secretion effectors in *Escherichia coli* O157 and the role of lambdoid phages in their dissemination. *Proc Natl Acad Sci USA.* Oct 3; 103 (40): 14941-6.

- 15 Vlisidou I, Marches O, Dziva F, Mundy R, Frankel G, Stevens MP. 2006. Identification and characterization of EspK, a type III secreted effector protein of enterohaemorrhagic *Escherichia coli* O157:H7. *FEMS Microbiol Lett.* Oct;263(1):32-40.

Walker, G., Fraiser, M., Schram, J., Little, M., Nadeau, J., and Douglas P. Malinowski, D. 1992 Strand Displacement Amplification—An Isothermal, In Vitro DNA Amplification Technique, *Nucleic Acids Research* 20, no. 7: 1691-1696.

LISTA DE SECUENCIAS

	<110> AGENCE NATIONALE DE SECURITE SANITAIRE DE L'ALIMENTATION, DE L'ENVIRONNEMENT ET DU TRAVAIL FACH, Patrick DELANNOY, Sabine BEUTIN, Lothar	
5	<120> Método para detectar e identificar Escherichia Coli enterohemorrágica	
	<130> MJP-11-F1159-8PCT	
10	<150> EP 12171941.3	
	<151> 14-06-2012	
	<160> 79	
15	<170> PatentIn versión 3.5	
	<210> 1	
	<211> 334	
	<212> DNA	
20	<213> Escherichia coli	
	<400> 1	
	gggtgttcta cgcaggggc ggggaattct aagtgatatc catcatcgca tccagtgcgc	60
	ccggtttatc cccgctgatg cggggaacac agcggcacgc tggattgaac aaatccctgg	120
	gccggtttat ccccgctggc gcggggaaca ctttatacac ggatcctgtg tgccgtggac	180
	cgccggttta tccccgctgg cgcggggaac acgagctggc gatagccctg aatgaactgg	240
	cagacggtta atccccgctg gcgcggggaa cacaaaccga aacacacgat caatccgaat	300
	atgagcggtt tatccccgct ggcgcgggga acac	334
25	<210> 2	
	<211> 212	
	<212> DNA	
	<213> Escherichia coli	
30	<400> 2	
	gggtgttcta cgcaggggc ggggaattct aagtgatatc catcatcgca tccagtgcgc	60
	ccggtttatc cccgctgatg cggggaacac agcggcacgc tggattgaac aaatccctgg	120
	gccggtttat ccccgctggc gcggggaaca caaaccgaaa cacacgatca atccgaatat	180
	gagcggttta tccccgctgg cgcggggaac ac	212
	<210> 3	
	<211> 212	
35	<212> DNA	
	<213> Escherichia coli	
	<400> 3	
	gggtgttcta cgcaggggc ggggaattct aagtgatatc catcatcgca tccagtgcgc	60
	ccggtttatc cccgctgatg cggggaacac agcggcacgc tggattgaac aaatccctgg	120
40	gccggtttat ccccgctggc gcggggaaca ctttggtgac agttttgtc actgttttgg	180
	tgacggttta tccccgctgg cgcggggaac ac	212
	<210> 4	
	<211> 273	
45	<212> DNA	
	<213> Escherichia coli	

ES 2 659 469 T3

<400> 4
 ggggtgttcta ccgcaggggc ggggaattct aagtgatatc catcatcgca tccagtgcgc 60
 ccggtttatc cccgctgatg cggggaacac agcggcacgc tggattgaac aaatccctgg 120
 gccggtttat ccccgctggc gcggggaaca caaacggaaa cacacgatca atccgaatat 180
 gagcggttta tccccgctgg cgcggggaac actttggtga cagtttttgt cactgttttg 240
 gtgacggttt atccccgctg gcgcggggaa cac 273
 5
 <210> 5
 <211> 272
 <212> DNA
 <213> Escherichia coli
 10
 <400> 5
 tggtttatcc ccaactggcgc ggggaacttg agccctgccca gaatggggcc tcttttgtac 60
 ccggtttatc cccgctgccg cggggaacac atttactgat tgtaccagggt aggaaaagat 120
 cgcggtttat ccccgctggc gcggggaact catgattgtc gattttttcc gacaggattt 180
 tgcggtttat ccccgctggc gcggggaact cgtcacactg ccgtgcatta ttacagccag 240
 gcgcggttta tccccgctgg cgcggggaac tc 272
 15
 <210> 6
 <211> 395
 <212> DNA
 <213> Escherichia coli
 <400> 6
 tggtttatcc ccgctggcgc ggggaactcg acagaacggc ctcagtagtc tcgtcaggct 60
 ccggtttatc cccgctggcg cggggagctc tacgtgaaga atatttgcaa caccgcgaag 120
 aacggtttat ccccgctggc gcggggaaca ctccaacctt ccatgagata cgcgcattag 180
 cgcggtttta tccccgctgg cgcggggaac accgtgaccg cctgtacacg ctgtaatggc 240
 tcaccggttt atccccgctg gcgcggggaa cacacactat ccgggcggtta ttacgcccaa 300
 atatccggtt tatccccgct ggcgcgggga acacctgccg ggtgaaacca ctgcgcgcag 360
 20
 atcttgcggt ttatccccgc tggcgcgggg aacac 395
 25
 <210> 7
 <211> 334
 <212> DNA
 <213> Escherichia coli
 <400> 7
 gggagttcta ccgcaggggc ggggaattct aagtgatatc catcatcgca tccagtgcgc 60
 ccggtttatc cccgctggcg cggggaacac tcacgcgctg ttaccagtc cctttttttc 120
 agcgggtttat ccccgctggc gcggggaaca ctacaggaaa gaatccgcca acggcgacag 180
 gggcggttta tccccgctgg cgcggggaac acgaatatcc acccgtcctg tattccgcca 240
 attgcggttt atccccgctg gcgcggggaa cactaccaaa ccggaatctt tccatataac 300
 ggcgcgcggtt tatccccgct ggcgcgggga acac 334
 30
 <210> 8
 <211> 273
 <212> DNA

ES 2 659 469 T3

<213> Escherichia coli

<400> 8

```

gggagttcta cgcaggggc ggggaattct aagtgatatc catcatcgca tccagtgcgc      60
ccggtttatc cccgctgacg cggggaacac aaaaccaaac ttctccataa attccatagc      120
cgcggtttat ccccgctggc gcggggaaca cgagtctatc agcgacacta cgggcaatag      180
cgacggttta tccccgctgg cgcggggaac acctatagcg ccacgttccg agcgctgcga      240
5  gctgcggttt atccccgctg gcgcggggaa cac                                273

```

<210> 9

<211> 822

<212> DNA

10 <213> Escherichia coli

<400> 9

```

tggtttatcc ccgctggcac ggggaactcg cctgacattg cagacgttac aaattgagag      60
gcggtttatc cccgctggcg cggggaactc taatcacggt ttagcgcgcc ctcgccgggt      120
ttcggtttat ccccgctggc gcggggaact catcttcttg ttcaggaacg tcagtagagg      180
tgtcggttta tccccgctgg cgcggggaac tcagcattaa cccccaccag ctcgacgtgt      240
gtggcggttt atccccgctg gcgcggggaa ctcgcgggcg ttaacgcggt gatactgttt      300
gacgcggtt tatccccgct ggcgcgggga actcacgag cgccgctggg aggcgtatct      360
cacgttcggt ttatccccgc tggcgcgggg aactctgaat gatcggaag acgctgcaaa      420
ggcaatgogg tttatccccg ctggcgcggg gaactcatta aaggattatt ttgatgagtc      480
tgaaaaatcg gtttatcccc gctggcgcgg ggaactctga ttttgggtgc gggggtgact      540
gggggaatac ggtttatccc cgctggcgcg gggaaactca ccggtgtgaa ttacagcacc      600
accgccacc cggtttatcc ccgctggcgc ggggaactcc aacaggtcg acatgtttgg      660
ctaacagcta acggtttatc cccgctggcg cggggaactc cgataacaca gggccgtcaa      720
tgcggccctt ttcggtttat ccccgctggc gcggggaact cataccgggt tttgctccag      780
15 cagcgccagt tcctggttta tccccgctgg cgcggggaac tc                                822

```

<210> 10

<211> 578

<212> DNA

20 <213> Escherichia coli

<400> 10

ES 2 659 469 T3

	tggtttatcc ccgctggcgc ggggagctca acatcgaaa cggcttcgcg gcggcggcgt	60
	ccggtttatc cccgctggcg cggggaactc gacagaacgg cctcagtagt ctcgtcaggg	120
	tccggtttat ccccgctggc gcggggagct ctacgtgaag aatatttgca acaccgcaa	180
	gaacggttta tccccgctgg cgcggggagc tcggcatagc caggctgatc cggcgacggc	240
	cttacggttt atccccgctg gcgcggggaa cacaatcgtg tgtaaattcg cggcggctcc	300
	actggcgggt tatccccgct ggcgcgggga aactcggctg aaatatagac agcatgttcc	360
	gtaccacggt ttatccccgc tggcgcgggg aacacacact atccgggcgg tattacgccc	420
	aaatatccgg tttatccccg ctggcgcggg gaacacctgc cgggtgaaac cactcgcggc	480
	agatcttgcg gtttatcccc gctggcgcgg ggaacacact actgtcggta gctgggagga	540
	tgaggagatc ggtttatccc cgctggcgcg gggaacac	578
5	<210> 11 <211> 21 <212> DNA <213> Escherichia coli	
10	<400> 11 gggaacacaa accgaaacac a 21	
15	<210> 12 <211> 19 <212> DNA <213> Escherichia coli	
	<400> 12 cttagtgtgt tccccgcg 19	
20	<210> 13 <211> 24 <212> DNA <213> Escherichia coli	
25	<400> 13 gaacacttg gtgacagttt ttgt 24	
30	<210> 14 <211> 19 <212> DNA <213> Escherichia coli	
35	<400> 14 cttagtgtgt tccccgcg 19	
40	<210> 15 <211> 21 <212> DNA <213> Escherichia coli	
	<400> 15 gaacacaaac cgaaacacac g 21	
45	<210> 16 <211> 22 <212> DNA <213> Escherichia coli	
50	<400> 16 ataaacgctc accaaaacag tg 22	

5 <210> 17
 <211> 20
 <212> DNA
 <213> Escherichia coli

 10 <400> 17
 gaacttgagc cctgccagaa 20

 15 <210> 18
 <211> 21
 <212> DNA
 <213> Escherichia coli

 20 <400> 18
 accgcatct tttctacct g 21

 25 <210> 19
 <211> 18
 <212> DNA
 <213> Escherichia coli

 30 <400> 19
 gtgaccgct gtacacgc 18

 35 <210> 20
 <211> 21
 <212> DNA
 <213> Escherichia coli

 40 <400> 20
 cggatattg ggcgtaatac c 21

 45 <210> 21
 <211> 19
 <212> DNA
 <213> Escherichia coli

 50 <400> 21
 ctgccgagcagg tggttcac 19

 55 <210> 22
 <211> 22
 <212> DNA
 <213> Escherichia coli

 60 <400> 22
 cggggaacac tacaggaaag aa 22

 65 <210> 23
 <211> 21
 <212> DNA
 <213> Escherichia coli

 70 <400> 23
 ggcggaatac aggacgggtg g 21

 75 <210> 24
 <211> 21
 <212> DNA
 <213> Escherichia coli

 80 <400> 24
 gagtctatca gcgacactac c 21

 85 <210> 25

<211> 18
 <212> DNA
 <213> Escherichia coli
 5
 <400> 25
 aaccgcagct cgcagcgc 18
 <210> 26
 <211> 18
 <212> DNA
 <213> Escherichia coli
 10
 <400> 26
 ggaactcacc gagcgccg 18
 <210> 27
 <211> 23
 <212> DNA
 <213> Escherichia coli
 15
 <400> 27
 gcctttgcag cgtctttccg atc 23
 <210> 28
 <211> 22
 <212> DNA
 <213> Escherichia coli
 20
 <400> 28
 acaatcgtgt gtaaattcgc gg 22
 <210> 29
 <211> 21
 <212> DNA
 <213> Escherichia coli
 25
 <400> 29
 gataaacctt ggtacggaac a 21
 <210> 30
 <211> 20
 <212> DNA
 <213> Escherichia coli
 30
 <400> 30
 tgaaaccact cgcggcagat 20
 <210> 31
 <211> 22
 <212> DNA
 <213> Escherichia coli
 35
 <400> 31
 ataaaccgat ctcctcatcc tc 22
 <210> 32
 <211> 23
 <212> DNA
 <213> Escherichia coli
 40
 <400> 32
 cgatcaatcc gaatatgagc ggt 23
 <210> 33
 <211> 24
 <212> DNA

<213> Escherichia coli
 <400> 33
 5 cactgttttg gtagcggtt atcc 24
 <210> 34
 <211> 24
 <212> DNA
 10 <213> Escherichia coli
 <400> 34
 acaaaaactg tcaccaaagt gttc 24
 15 <210> 35
 <211> 21
 <212> DNA
 <213> Escherichia coli
 20 <400> 35
 tggggcctct tttgtaccg g 21
 <210> 36
 <211> 25
 25 <212> DNA
 <213> Escherichia coli
 <400> 36
 tgtaatggct caccggtta tcccc 25
 30 <210> 37
 <211> 20
 <212> DNA
 <213> Escherichia coli
 35 <400> 37
 tccgccaacg gcgacagggg 20
 <210> 38
 <211> 22
 <212> DNA
 <213> Escherichia coli
 40 <400> 38
 tcggaacgtg gcgctatagg tg 22
 <210> 39
 <211> 25
 <212> DNA
 50 <213> Escherichia coli
 <400> 39
 ctgggaggcg tatctcacgt tcggt 25
 55 <210> 40
 <211> 27
 <212> DNA
 <213> Escherichia coli
 60 <400> 40
 tgctgtctat atttcgacca gtttcc 27
 <210> 41
 <211> 25
 65 <212> DNA
 <213> Escherichia coli

<400> 41
 ccagctaccg acagtagtgt gttcc 25

5 <210> 42
 <211> 26
 <212> DNA
 <213> Escherichia coli

10 <400> 42
 gcagrcatca aaagcgaat cacacc 26

<210> 43
 <211> 26
 15 <212> DNA
 <213> Escherichia coli

<400> 43
 tcgtttgta actgtggcag atactc 26

20 <210> 44
 <211> 23
 <212> DNA
 <213> Escherichia coli

25 <400> 44
 tcaggtcct cgtctgatgc cgc 23

30 <210> 45
 <211> 23
 <212> DNA
 <213> Escherichia coli

35 <400> 45
 ctggttcagg cctggagcag tcc 23

<210> 46
 <211> 22
 <212> DNA
 40 <213> Escherichia coli

<400> 46
 gcaataattg actctgattg cc 22

45 <210> 47
 <211> 21
 <212> DNA
 <213> Escherichia coli

50 <400> 47
 gctgctgcgg taaaattac t 21

<210> 48
 <211> 21
 55 <212> DNA
 <213> Escherichia coli

<400> 48
 ctgaaaagag ccagaacgtg c 21

60 <210> 49
 <211> 22
 <212> DNA
 <213> Escherichia coli

65 <400> 49
 tgcctaagat cattaccgg ac 22

5 <210> 50
 <211> 24
 <212> DNA
 <213> Escherichia coli

 10 <400> 50
 cgatcattgt gggcatgta tgcc 24

 15 <210> 51
 <211> 22
 <212> DNA
 <213> Escherichia coli

 20 <400> 51
 cctgaattca cacggtgatg cg 22

 25 <210> 52
 <211> 27
 <212> DNA
 <213> Escherichia coli

 30 <400> 52
 gccttttat gttcattatt gcggttg 27

 35 <210> 53
 <211> 26
 <212> DNA
 <213> Escherichia coli

 40 <400> 53
 gtatagtttt agcaatacct tctgc 26

 45 <210> 54
 <211> 21
 <212> DNA
 <213> Escherichia coli

 50 <400> 54
 gattgtggcg attaatgggg g 21

 55 <210> 55
 <211> 21
 <212> DNA
 <213> Escherichia coli

 60 <400> 55
 acaccgatct ggtcattggc g 21

 65 <210> 56
 <211> 23
 <212> DNA
 <213> Escherichia coli

 70 <400> 56
 aaacgccttt aaaatctgcg tct 23

 75 <210> 57
 <211> 21
 <212> DNA
 <213> Escherichia coli

 80 <400> 57
 tgccgtgcgc acagcataa g 21

 85 <210> 58

<211> 20
 <212> DNA
 <213> Escherichia coli
 5
 <400> 58
 gcccatggct ccacatcctg 20
 10
 <210> 59
 <211> 26
 <212> DNA
 <213> Escherichia coli
 15
 <400> 59
 ccaaaaaagt tatgatgatt gcactg 26
 20
 <210> 60
 <211> 23
 <212> DNA
 <213> Escherichia coli
 <400> 60
 gcactggccc ttgtgctca ggc 23
 25
 <210> 61
 <211> 27
 <212> DNA
 <213> Escherichia coli
 30
 <400> 61
 gctcttcag tgagaatgc ttcccg 27
 35
 <210> 62
 <211> 29
 <212> DNA
 <213> Escherichia coli
 <400> 62
 tttgyactg tsacagcwga agcyttacg 29
 40
 <210> 63
 <211> 26
 <212> DNA
 <213> Escherichia coli
 45
 <400> 63
 ccccagtca rwgtragrtc macrtc 26
 50
 <210> 64
 <211> 26
 <212> DNA
 <213> Escherichia coli
 55
 <400> 64
 attcagatag aagaagcgcg ggccag 26
 60
 <210> 65
 <211> 28
 <212> DNA
 <213> Escherichia coli
 <400> 65
 cttgcaacac gttacgctgc cgagtatt 28
 65
 <210> 66
 <211> 24
 <212> DNA

<213> Escherichia coli
 <400> 66
 5 tacgctgac accatgctg gtgc 24
 <210> 67
 <211> 22
 <212> DNA
 10 <213> Escherichia coli
 <400> 67
 taactgctat acctccgcg cg 22
 15 <210> 68
 <211> 27
 <212> DNA
 <213> Escherichia coli
 20 <400> 68
 tgtaacaccc agacggtcag caacatg 27
 <210> 69
 <211> 29
 25 <212> DNA
 <213> Escherichia coli
 <400> 69
 tcacttccag tttctggtga tgtttgat 29
 30 <210> 70
 <211> 31
 <212> DNA
 <213> Escherichia coli
 35 <400> 70
 tgggtgaggt taaaataaa agaacgattg c 31
 <210> 71
 <211> 31
 <212> DNA
 <213> Escherichia coli
 40 <400> 71
 45 taagatatt tctgactttc cgcatgcgct t 31
 <210> 72
 <211> 25
 <212> DNA
 50 <213> Escherichia coli
 <400> 72
 aaagagccag cgcagagctg accag 25
 55 <210> 73
 <211> 23
 <212> DNA
 <213> Escherichia coli
 60 <400> 73
 ttcgctggaa gcagagcccg tgc 23
 <210> 74
 <211> 31
 65 <212> DNA
 <213> Escherichia coli

<400> 74
ctggatgac tcagtggcg ttctatgta a 31

5 <210> 75
<211> 27
<212> DNA
<213> Escherichia coli

10 <400> 75
tcgtcaggca ctgtctgaaa ctgctcc 27

<210> 76
<211> 24
15 <212> DNA
<213> Escherichia coli

<400> 76
20 ggtgaaggag ytgccaagg cgtc 24

<210> 77
<211> 23
<212> DNA
<213> Escherichia coli

25 <400> 77
ccggtggatt tggatgggtt acg 23

<210> 78
<211> 22
30 <212> DNA
<213> Escherichia coli

<400> 78
35 tgtgaacctc tctggcatgg ag 22

<210> 79
<211> 24
<212> DNA
40 <213> Escherichia coli

<400> 79
taaagttggt agattgtgac tggc 24

45

REIVINDICACIONES

- 5 1. Un método para identificar el(los) serotipo(s) de *Escherichia coli* enterohemorrágica (ECEH) que se sospecha que están presentes en una muestra, donde dicho método comprende detectar la presencia o la ausencia, en dicha muestra o ADN aislado de la misma, de las siguientes secuencias CRISPR de *E. coli*:
- a) Secuencias CRISPR para identificar ECEH O157:[H7] en donde dichas secuencias CRISPR se seleccionan entre
- 10 - las secuencias CRISPR SEQ ID NO: 1, SEQ ID NO: 2, y SEQ ID NO: 3, en donde la presencia de una o más de dichas CRISPR SEQ ID NO: 1-3 es indicativa de la presencia de ECEH O157:[H7]; y/o;
- la secuencia CRISPR SEQ ID NO: 4, en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O157:[H7];
- b) una secuencia CRISPR para identificar ECEH O145:[H28], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 5, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O145:[H28]; y
- 15 c) una secuencia CRISPR para identificar ECEH O111:[H8], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 6, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O111:[H8]; y
- d) una secuencia CRISPR para identificar ECEH O121:[H19], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 7, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O121:[H19]; y
- 20 e) una secuencia CRISPR para identificar ECEH O103:[H2] y/o ECEH O45:[H2], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 8, y en donde la presencia de la secuencia CRISPR es indicativa de la presencia de O103:[H2] y/o de ECEH O45:[H2]; y
- 25 f) una secuencia CRISPR para identificar ECEH O104:[H4], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 9, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O104:[H4]; y
- g) una secuencia CRISPR para identificar ECEH O26:[H11], en donde dicha secuencia CRISPR es la secuencia SEQ ID NO: 10, y en donde la presencia de dicha secuencia CRISPR es indicativa de la presencia de ECEH O26:[H11].
- 30 2. Un método de la reivindicación 1, en donde dicho método comprende realizar un ensayo de PCR sobre dicha muestra o ADN aislado de la misma con una combinación de cebadores que se dirigen a dicha secuencia CRISPR.
3. Un método de la reivindicación 2, en donde dicha combinación de cebadores comprende:
- 35 a) cebadores para detectar ECEH O157:[H7], en donde dichos cebadores consisten en:
- un conjunto de cebadores que se dirigen a las secuencias CRISPR SEQ ID NO: 1 y SEQ ID NO: 2, en donde dichos cebadores se definen por las siguientes secuencias:
- GGAACACAAACCGAAACACA (SEQ ID NO: 11)
- CTTAGTGTGTTCCCGCGC (SEQ ID NO: 12) y
- 40 - un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 3 en donde dichos cebadores se definen por las siguientes secuencias:
- GAACACTTTGGTGACAGTTTTTGT (SEQ ID NO: 13);
- CTTAGTGTGTTCCCGCGC (SEQ ID NO: 14),
- 45 en donde la presencia de un producto de amplificación para al menos uno de dichos conjuntos de cebadores es indicativa de la presencia de ECEH O157:[H7]; y/o:
- un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 4, en donde dichos cebadores se definen por las siguientes secuencias:
- GAACACAAACCGAAACACACG (SEQ ID NO: 15)

ATAAACCGTCACCAAACAGTG (SEQ ID NO: 16),

en donde la presencia de un producto de amplificación para dicho conjunto de cebadores es indicativa de la presencia de ECEH O157:[H7]; y

5 b) cebadores para detectar ECEH O145:[H28], en donde dichos cebadores consisten en:

- un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 5, en donde dichos cebadores se definen por las siguientes secuencias:

GAACTTGAGCCCTGCCAGAA (SEQ ID NO: 17)

ACCGCGATCTTTTCCTACCTG (SEQ ID NO: 18),

10 en donde la presencia de un producto de amplificación para dicho conjunto de cebadores es indicativa de la presencia de la presencia de ECEH O145:[H28]; y

c) cebadores para detectar ECEH O111:[H8], en donde dichos cebadores consisten en:

- un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 6, en donde dichos cebadores se definen por las siguientes secuencias:

15 GTGACCGCCTGTACACGC (SEQ ID NO: 19)

CGGATATTTGGGCGTAATACC (SEQ ID NO: 20)

CTGCCGCGAGTGGTTTCAC (SEQ ID NO: 21),

en donde la presencia de un producto de amplificación para al menos uno de los pares de cebadores SEQ ID NO: 19 y SEQ ID NO: 20 o SEQ ID NO: 19 y SEQ ID NO: 21 es indicativa de la presencia de ECEH O111:[H8]; y

20 d) cebadores para detectar ECEH O121:[H19], en donde dichos cebadores consisten en:

- un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 7, en donde dichos cebadores se definen por las siguientes secuencias:

CGGGGAACACTACAGGAAAGAA (SEQ ID NO: 22)

GGCGGAATACAGGACGGGTGG (SEQ ID NO: 23),

25 en donde la presencia de un producto de amplificación para dicho conjunto de cebadores es indicativa de la presencia de ECEH O121:[H19]; y

e) cebadores para detectar ECEH O103:[H2] y/o ECEH O45:[H2], en donde dichos cebadores consisten en:

- un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 8, en donde dichos cebadores se definen por las siguientes secuencias:

30 GAGTCTATCAGCGACTACC (SEQ ID NO: 24)

AACCGCAGCTCGCAGCGC (SEQ ID NO: 25),

en donde la presencia de un producto de amplificación para dicho conjunto de cebadores es indicativa de la presencia de ECEH O103:[H2] y/o de ECEH O45:[H2]; y

f) cebadores para detectar ECEH O104:[H4], en donde dichos cebadores consisten en:

35 - un conjunto de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 9, en donde dichos cebadores se definen por las siguientes secuencias:

GGAACTCACCGAGCGCCG (SEQ ID NO: 26);

GCCTTTGCAGCGTCTTTCCGATC (SEQ ID NO: 27);

40 en donde la presencia de un producto de amplificación para dicho conjunto de cebadores es indicativa de la presencia de ECEH O104:[H4]; y

g) cebadores para detectar ECEH O26:[H11], en donde dichos cebadores consisten en:

- dos conjuntos de cebadores que se dirigen a la secuencia CRISPR SEQ ID NO: 10, en donde el primer conjunto de cebadores está definido por las siguientes secuencias:

ACAATCGTGTGTAATTCGCGG (SEQ ID NO: 28)

GATAAACCGTGGTACGGAACA (SEQ ID NO: 29) y el segundo dicho conjunto de cebadores está definido por las siguientes secuencias:

5 TGAAACCACTCGCGGCAGAT (SEQ ID NO: 30);

ATAAACCGATCTCCTCATCCTC (SEQ ID NO: 31);

en donde la presencia de un producto de amplificación para al menos uno de los dichos conjuntos de cebadores es indicativa de la presencia de ECEH O26:[H11].

10 4. Un método de cualquiera de las reivindicaciones 1 a 3, que comprende una etapa previa para predecir si dicha muestra contiene *Escherichia coli* enterohemorrágica (ECEH) de al menos uno de los serotipos O157:[H7], O145:[H28], O103:[H2], O111:[H8], O121:[H19], O26:[H11], O45:[H2] y O104:[H4] de ECEH, en donde dicho método además comprende la detección del gen *espK* y de uno o más de los siguientes genes diana: *espV*, *ureD*, *Z2098*, *Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156*, y *Z6065*.

15 5. Un método de la reivindicación 4, en donde dicha etapa previa para predecir si dicha muestra contiene *Escherichia coli* enterohemorrágica (ECEH) de al menos uno de los serotipos O157:[H7], O145:[H28], O103:[H2], O111:[H8], O121:[H19], O26:[H11], O45:[H2] y O104:[H4] de ECEH, comprende la detección del gen *espK*, de al menos un gen seleccionado entre *espV*, *ureD*, *Z2098*, *Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156*, y del gen *Z6065*.

20 6. Un método de la reivindicación 4 o la reivindicación 5, en donde dicha etapa previa para predecir si dicha muestra contiene *Escherichia coli* enterohemorrágica (ECEH) de al menos uno de los serotipos O157:[H7], O145:[H28], O103:[H2], O111:[H8], O121:[H19], O26:[H11], O45:[H2] y O104:[H4] de ECEH comprende realizar un ensayo de PCR sobre dicha muestra o ADN aislado de la misma con una combinación de cebadores que comprende un conjunto de cebadores derivados de *espK* y un conjunto de cebadores derivados de al menos uno de *espV*, *ureD*, *Z2098*, *Z1151*, *Z1153*, *Z1154*, *Z1155*, *Z1156*, y *Z6065* y detectar la presencia o la ausencia de un producto de amplificación para cada conjunto de cebadores de dicha combinación.

25 7. Un método de cualquiera de las reivindicaciones 4 a 6, en donde dicha etapa previa para predecir si dicha muestra contiene *Escherichia coli* enterohemorrágica (ECEH) de al menos uno de los serotipos O157:[H7], O145:[H28], O103:[H2], O111:[H8], O121:[H19], O26:[H11], O45:[H2] y O104:[H4] de ECEH comprende además realizar un ensayo de PCR sobre dicha muestra o ADN aislado de la misma con una combinación de cebadores que comprende un conjunto de cebadores derivados de *stx1* y un conjunto de cebadores derivados de *stx2* y detectar la presencia o la ausencia de un producto de amplificación para cada conjunto de cebadores de dicha combinación.

30 8. Un kit para la identificación del(de los) serotipo(s) de *Escherichia coli* enterohemorrágica (ECEH), que comprende los conjuntos de cebadores definidos en la reivindicación 3, y opcionalmente las sondas para detectar los productos de amplificación para cada dicho conjunto de cebadores.

35 9. Un kit de la reivindicación 8, que comprende además un conjunto de cebadores derivados de *espK*, y al menos un conjunto de cebadores seleccionados entre: un conjunto de cebadores derivados de *espV*, un conjunto de cebadores derivados de *ureD*, un conjunto de cebadores derivados de *Z2098*, un conjunto de cebadores derivados de *Z1151*, un conjunto de cebadores derivados de *Z1153*, un conjunto de cebadores derivados de *Z1154*, un conjunto de cebadores derivados de *Z1155*, un conjunto de cebadores derivados de *Z1156*, un conjunto de cebadores derivados de *Z6065*, y opcionalmente las sondas para detectar los productos de amplificación para cada dicho conjunto de cebadores.

40 10. Un kit de una de las reivindicaciones 8 o 9, que comprende además un conjunto de cebadores que se dirigen a *stx1* y *stx2*, y opcionalmente una sonda para detectar el producto de amplificación a partir de *stx1* y una sonda para detectar el producto de amplificación a partir de *stx2*.