

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 382 408**

51 Int. Cl.:
C12N 15/82 (2006.01)
C12N 15/61 (2006.01)
C12N 9/90 (2006.01)
A01H 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **08708634 .4**
96 Fecha de presentación: **04.02.2008**
97 Número de publicación de la solicitud: **2121933**
97 Fecha de publicación de la solicitud: **25.11.2009**

54 Título: **Uso de genes de la alanina racemasa para conferir a las plantas resistencia a los nemátodos**

30 Prioridad:
06.02.2007 US 899746 P

45 Fecha de publicación de la mención BOPI:
08.06.2012

45 Fecha de la publicación del folleto de la patente:
08.06.2012

73 Titular/es:
BASF Plant Science GmbH
67056 Ludwigshafen, DE

72 Inventor/es:
HILL, Steve

74 Agente/Representante:
Carvajal y Urquijo, Isabel

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).

ES 2 382 408 T3

DESCRIPCIÓN

Uso de genes de la alanina racemasa para conferir a las plantas resistencia a los nemátodos

Campo de la invención

- 5 La invención se refiere al control de los nematodos, en particular, el control de los nematodos del quiste de la soja. Se divulgan aquí métodos de producción de plantas transgénicas con mayor resistencia a los nematodos, vectores de expresión que contienen polinucleótidos que codifican para proteínas funcionales, y plantas transgénicas y semillas generadas a partir de los mismos.

Antecedentes de la invención

- 10 Los nematodos son gusanos microscópicos que se alimentan de las raíces, hojas y tallos de más de 2.000 vegetales, frutas y plantas ornamentales, causando una pérdida estimada de cultivos por valor de \$ 100 mil millones en todo el mundo. Un tipo común de nematodo es el nematodo agallador (RKN), cuya alimentación provoca las características agallas sobre las raíces en una amplia variedad de especies de plantas. Otros nemátodos que se alimentan de raíces son los que forman quistes y lesiones, que son más específico del huésped.

- 15 Los nematodos están presentes en todos los Estados Unidos, pero son sobre todo un problema en áreas cálidas y húmedas del sur y el oeste, y en suelos arenosos. El nematodo del quiste de la soja (SCN), *Heterodera glycines*, fue descubierto por primera vez en los Estados Unidos en Carolina del Norte en 1954. Es la plaga más grave de las plantas de soja. Algunas áreas están tan fuertemente infestadas por el SCN que la producción de soja ya no es económicamente posible sin medidas de control. Aunque la soja es el principal cultivo de importancia económica atacado por el SCN, el SCN parasita a unos cincuenta huéspedes en total, incluyendo cultivos de campo, hortalizas, plantas ornamentales y malezas.

- 20 Los signos de daño causado por nematodos incluyen retraso del crecimiento y amarillamiento de las hojas y el marchitamiento de las plantas durante periodos de calor. Sin embargo, los nematodos, incluyendo al SCN, pueden causar pérdidas significativas de rendimiento sin síntomas obvios por encima del suelo. Además, las raíces infectadas con el SCN se quedan enanas o atrofiadas. La infestación de nematodos pueden reducir el número de nódulos fijadores de nitrógeno sobre las raíces, y puede hacer que estas sean más susceptibles a los ataques de otros patógenos de las plantas transmitidas por el suelo.

- 25 El ciclo de vida del nematodo tiene tres grandes etapas: de huevo, juvenil y adulta. El ciclo de vida varía entre las especies de nematodos. Por ejemplo, el ciclo de vida del SCN por lo general se puede completar en 24 a 30 días en condiciones óptimas mientras que a otras especies les puede tomar hasta un año o más, para completar el ciclo de vida. Cuando los niveles de temperatura y de humedad se tornan adecuados en la primavera, eclosionan en forma de gusanos jóvenes en el suelo. Estos nemátodos en etapa juvenil son los únicos que puede infectar las raíces de la soja.

- 30 El ciclo de vida del SCN ha sido objeto de muchos estudios y por lo tanto puede ser utilizado como un ejemplo para la comprensión del ciclo de vida de un nemátodo. Después de penetrar en las raíces de la soja, los SCN juveniles se mueven a través de la raíz hasta que se ponen en contacto con el tejido vascular, donde se detienen y comienzan a alimentarse. El nematodo inyecta secreciones que modifican ciertas células de la raíz y las transforman en sitios de alimentación especializados. Las células de la raíz se transforma morfológicamente en grandes sincitios multinucleados (o células gigantes en el caso del RKN), que son utilizados como fuente de nutrientes por los nematodos. Los nematodos se alimentan activamente robando así nutrientes esenciales de la planta lo que resulta en la pérdida de rendimiento. A medida que se alimentan los nematodos, se hinchan y eventualmente los nematodos hembra llegan a ser tan grandes que se rompen a través del tejido de la raíz y quedan expuestos sobre la superficie de la raíz.

- 35 Los nematodos SCN macho, que no se hinchan en etapa adulta, migrar fuera de la raíz hacia el suelo y fertilizan a las hembras adultas en forma de limón. Los machos mueren luego, mientras que las hembras permanecen adheridas a las raíces y continúan alimentándose. Los huevos de las hembras hinchadas comienzan a desarrollarse, inicialmente en una masa o saco de huevos fuera del cuerpo, y más tarde dentro de la cavidad corporal. Eventualmente, toda la cavidad corporal de la hembra adulta está llena de huevos, y el nematodo hembra muere. Es el cuerpo lleno de huevo de la hembra muerta lo que se conoce como el quiste. Los quistes eventualmente desalojan la raíz y se encuentran libres en el suelo. Las paredes del quiste se hacen muy duras, proporcionando una excelente protección para los aproximadamente 200 a 400 huevos que contiene. Los huevos del SCN sobreviven dentro del quiste hasta que se presentan condiciones adecuadas para la eclosión. Aunque muchos de los huevos pueden eclosionar durante el primer año, muchos de ellos también pueden sobrevivir dentro de los quistes durante
- 40
- 45
- 50
- 55

- Los nematodos pueden moverse a través del suelo únicamente unas pocas pulgadas por año por sus propios medios. Sin embargo, infestación por nematodos se puede expandir distancias significativas en una variedad de formas. Cualquier cosa que pueda mover el suelo infestado es capaz de propagar la infestación, incluida la maquinaria agrícola, vehículos y herramientas, viento, agua, animales, y los trabajadores agrícolas. Partículas del suelo del tamaño de las semillas a menudo contaminan a la semilla cosechada. En consecuencia, la infestación de nematodos se puede propagar cuando la semilla contaminada de los campos infestados se siembra en campos no infestados. Incluso, existe evidencia de que ciertas especies de nematodos pueden ser esparcidas por las aves. Sólo algunas de estas causas pueden ser prevenidas.
- 10 Las prácticas tradicionales de manejo de infestación por nematodos incluyen: el mantenimiento apropiado de los nutrientes del suelo y los niveles de pH del suelo en tierras infestadas con nematodos; el control de otras enfermedades de las plantas, así como de las plagas de insectos y de malas hierbas; la utilización de prácticas de saneamiento tales como el arado, la siembra y el cultivo de los campos infestados con nematodos sólo después del laboreo de los campos no infestados; limpieza completa de los equipos con agua a alta presión o vapor de agua
- 15 después de trabajar en los campos infestados; no utilizar semillas cultivadas en tierras infestadas para la siembra de campos no infestados, a menos que la semilla haya sido correctamente limpiada; rotación de los campos infestados y alternación de los cultivos huéspedes con los cultivos no huéspedes; uso de nematicidas; y la siembra de variedades de plantas resistentes.
- 20 Se han propuesto métodos para la transformación genética de las plantas con el fin de conferirles una mayor resistencia a los nematodos que parasitan las plantas. Las Patentes de Los Estados Unidos Nos. 5.589.622 y 5.824.876 están dirigidas a la identificación de genes de las plantas que se expresan específicamente en o cerca al sitio de alimentación de la planta después de la fijación por el nematodo.
- 25 La alanina racemasa (EC 5.1.1.1) cataliza la interconversión de los enantiómeros L y D de alanina y representa el primer paso implicado en la biosíntesis de la pared celular bacteriana. La D-alanina es un componente esencial del peptidoglicano de la pared celular (mureína) en todas las bacterias y es producido por la racemización de L-alanina. La actividad de la alanina racemasa en *E. coli* es debida a dos productos génicos distintos. Una alanina racemasa (Alr) es constitutiva / poco abundante y es codificada por Alr (Neidhardt et al., *J. Biol. Chem.* 1989, 15: 264 (5): 2393 - 6). La otra alanina racemasa, DadX, es inducida por D-alanina o L-alanina y reprimida por glucosa y es codificada
- 30 por el gen para DadX (Hennig et al., *Mol Gen Genet* 1985, 198 (2): 315 - 22).
- No obstante lo anterior, existe la necesidad de identificar composiciones seguras y efectivas y métodos para el control de los nematodos que parasitan las plantas, y para la producción de plantas que tengan mayor resistencia a los nematodos que parasitan las plantas.

Resumen de la invención

- 35 Los presentes inventores encontraron que la expresión de un transgén que comprende un gen para racemasa alanina en una planta puede conferirle a la planta resistencia a los nematodos. La presente invención proporciona plantas transgénicas y semillas, y métodos para superar, o al menos aliviar, la infestación por nematodos en las plantas de cultivo.
- 40 Por lo tanto, en una realización, la invención se refiere a una planta transgénica que contiene un vector de expresión que comprende un polinucleótido aislado que codifica una racemasa alanina. Preferiblemente, el polinucleótido que codifica alanina racemasa es sobreexpresado en los sincitios inducidos por nemátodos.
- Otra realización de la invención proporciona una semilla transgénica que es una línea genéticamente pura para un transgén de alanina racemasa.
- 45 Otra realización de la invención se relaciona con el uso de un vector de expresión que comprende un elemento regulador de la transcripción operativamente enlazado a un polinucleótido que codifica una alanina racemasa, en donde la expresión del polinucleótido confiere resistencia a los nematodos a una planta transgénica. Preferiblemente, el vector de expresión comprende también un promotor operativamente enlazado al polinucleótido que codifica alanina racemasa, siendo capaz el promotor de dirigir la expresión del polinucleótido que codifica alanina racemasa en las raíces o en los sincitios de las plantas infectadas con nematodos.
- 50 En otra realización, la invención proporciona un método para producir una planta transgénica que tiene mayor resistencia a los nematodos, en donde el método comprende las etapas de introducir en la planta un vector de expresión que contiene un promotor operativamente enlazado a un polinucleótido que codifica alanina racemasa, en donde la expresión del polinucleótido confiere a la planta una mayor resistencia a los nemátodos.

Breve descripción de los dibujos

La Figura 1 muestra el porcentaje de identidad de aminoácidos de los homólogos de DadX con la proteína DadX (SEQ ID NO: 6).

5 La Figura 2 muestra el porcentaje de identidad de aminoácidos de los homólogos de Alr con la proteína Alr (SEQ ID NO: 8).

La Figura 3 muestra las secuencias de ADN y de proteína del gen para DadX.

La Figura 4 muestra las secuencias de ADN y de proteína del gen para Alr.

Las Figuras 5a y 5b muestran las secuencias de ADN para los promotores MTN3, POX y TPP.

Descripción detallada de las realizaciones preferidas

10 La presente invención se puede entender más fácilmente por medio de la siguiente descripción detallada de las realizaciones de la invención y de los ejemplos incluidos en ella. A menos que se especifique lo contrario, los términos utilizados aquí deben entenderse también de acuerdo con el uso convencional por parte de aquellos ordinariamente capacitados en la técnica pertinente.

15 A todo lo largo de esta solicitud, se hace referencia a diferentes patentes y publicaciones científicas. Las divulgaciones de todas estas publicaciones y de aquellas referencias citadas en esas publicaciones en su totalidad se incorporan aquí como referencia en esta solicitud con el fin de describir más completamente el estado de la técnica a la que pertenece esta invención. Las abreviaturas y nomenclatura, donde se emplean, se consideran estándar en el campo y se utilizan normalmente en las revistas profesionales, tales como las citadas en este documento. Tal como se utiliza aquí, la palabra "o" significa cualquier miembro de un listado en particular, y también incluye cualquier combinación de miembros de esa lista.

25 El término "aproximadamente" se utiliza aquí para referirse a, aproximadamente, más o menos, alrededor de, o en las regiones de. Cuando el término "aproximadamente" se utiliza junción con un rango numérico, modifica ese rango extendiendo los límites por encima y por debajo de los valores numéricos expuestos. En general, el término "aproximadamente" se utiliza aquí para modificar un valor numérico por encima y por debajo del valor establecido por medio de una diferencia del 10 por ciento, hacia arriba o hacia abajo (superior o inferior). Como se usa aquí, el término "ácido nucleico", "nucleótido", o "polinucleótido" está destinado a incluir las moléculas de ADN (por ejemplo, ADNc o ADN genómico), las moléculas de ARN (por ejemplo, ARNm), de origen natural, mutadas, moléculas sintéticas de ADN o de ARN, y análogos del ADN o del ARN generados utilizando análogos de nucleótidos. Puede ser monocatenario o bicatenario. Tales ácidos nucleicos o polinucleótidos incluyen, pero no se limitan a, las secuencias de codificación de los genes estructurales, secuencias antisentido, y secuencias reguladoras no codificadoras que no codifican los ARNm o productos proteicos. Un polinucleótido puede codificar para un rasgo fenotípico o agronómicamente valioso.

35 Como se usa aquí, un polinucleótido "aislado" está sustancialmente libre de otros materiales celulares o medio de cultivo cuando es producido mediante técnicas recombinantes, o sustancialmente libre de precursores químicos cuando es sintetizado químicamente.

40 El término "gen" se utiliza ampliamente para referirse a cualquier segmento de ácido nucleico asociado con una función biológica. De este modo, el término genes incluye intrones y exones como en una secuencia genómica, o sólo las secuencias de codificación como en los ADNc y / o en las secuencias reguladoras necesarias para su expresión. Por ejemplo, gen se refiere a un fragmento de ácido nucleico que expresa ARNm o ARN funcional, o que codifica una proteína específica, y que incluye secuencias reguladoras.

45 Los términos "polipéptido" y "proteína" se utilizan aquí indistintamente para referirse a un polímero de residuos de aminoácidos consecutivos.

50 El término "operativamente enlazado" o "funcionalmente enlazado" como se utilizan aquí se refieren a la asociación de secuencias de ácido nucleico sobre un solo fragmento de ácido nucleico de tal manera que la función de una se ve afectada por la otra. Por ejemplo, se dice que un ADN regulador está "operativamente enlazado a" un ADN que expresa un ARN o codifica a un polipéptido si los dos ADN están situados de tal manera que el ADN regulador afecta la expresión del ADN de codificación.

55 El término "expresión específica" como se usa aquí se refiere a la expresión de productos génicos que se limita a uno o a unos pocos tejidos vegetales (limitación especial) y / o a una o a unas pocas etapas de desarrollo de la planta (limitación temporal). Se reconoce que difícilmente existe una verdadera especificidad: los promotores

parecen ser preferiblemente activados en algunos tejidos, mientras que en otros tejidos no puede haber ninguna actividad o sólo poca actividad. Este fenómeno se conoce como expresión defectuosa. Sin embargo, con el término expresión específica como se define aquí, se entiende la expresión en uno o en unos pocos tejidos vegetales o sitios específicos en una planta.

El término "promotor" como se usa aquí se refiere a una secuencia de ADN que, cuando se liga a una secuencia de nucleótidos de interés, es capaz de controlar la transcripción de la secuencia de nucleótidos de interés en ARNm. Un promotor está típicamente, aunque no necesariamente, localizado 5' (por ejemplo, secuencia arriba) de un nucleótido de interés (por ejemplo, próximo al sitio de inicio de la transcripción de un gen estructural) cuya transcripción en ARNm controla, y proporciona un sitio para enlazamiento específico por medio de la ARN polimerasa y otros factores de transcripción para la iniciación de la transcripción.

El término "elemento regulador de la transcripción" como se usa aquí se refiere a un polinucleótido que es capaz de regular la transcripción de un polinucleótido operativamente enlazado. Este incluye, pero no se limita a, promotores, reforzadores, intrones, los UTR 5', y los UTR 3'.

Como se usa aquí, el término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico con el que se ha enlazado. Un tipo de vector es un "plásmido", que se refiere a un bucle circular de ADN bicatenario dentro del cual se pueden ligar segmentos adicionales de ADN. En la presente memoria descriptiva, "plásmido" y "vector" se puede utilizar indistintamente ya que el plásmido es la forma más comúnmente utilizada de vector. Un vector puede ser un vector binario o un ADN-T que incluye el borde izquierdo y el borde derecho y puede incluir un gen de interés en el medio. El término "vector de expresión" como se usa aquí, significa un vector capaz de dirigir la expresión de un nucleótido particular en una célula huésped apropiada. Un vector de expresión comprende un elemento regulador de ácido nucleico operativamente enlazado a un ácido nucleico de interés, que está - opcionalmente - operativamente enlazado a una señal de terminación y / o a otros elementos reguladores.

El término "homólogos" tal como aquí se utiliza se refiere a un gen relacionado con un segundo gen que descende de una secuencia ancestral común de ADN. El término "homólogos" puede aplicarse a la relación entre genes separados por el evento de especiación (por ejemplo, ortólogos) o a la relación entre genes separados por el evento de la duplicación genética (por ejemplo, parálogos).

Como se usa aquí, el término "ortólogos" se refiere a genes de diferentes especies, pero que han evolucionado a partir de un gen ancestral común por especiación. Los ortólogos retienen la misma función en el transcurso de la evolución. Los ortólogos codifican proteínas que tienen las mismas funciones o funciones similares. Como se usa aquí, el término "parálogos" se refiere a los genes que están relacionados por duplicación dentro de un genoma. Los parálogos usualmente tienen funciones diferentes o funciones nuevas, pero estas funciones pueden estar relacionadas.

Como se usa aquí, el término "hibrida bajo condiciones rigurosas" pretende describir las condiciones para hibridación y lavado bajo las cuales las secuencias de nucleótidos que son al menos 60% similares o idénticas entre sí típicamente permanecen hibridadas entre sí. En otra realización, las condiciones son tales que las secuencias que son al menos aproximadamente 65%, o al menos aproximadamente 70%, o al menos aproximadamente 75% o más, similares o idénticas entre sí, típicamente permanecen hibridadas entre sí. Tales condiciones rigurosas son conocidas por aquellos capacitados en el arte y se describen más adelante. Un ejemplo preferido, no limitante de condiciones estrictas de hibridación es la hibridación en cloruro de sodio / citrato de sodio 6X aproximadamente a 45° C, seguido por uno o más lavados en 0,2 X SSC, SDS al 0,1% a 50° - 65°C.

El término "identidad de secuencia" o "identidad" en el contexto de dos secuencias de ácido nucleico o de polipéptido hace referencia a aquellas posiciones en las dos secuencias en donde pares idénticos de símbolos caen juntos cuando las secuencias se alinean para correspondencia máxima sobre una ventana de comparación especificada, por ejemplo, ya sea la secuencia completa como en una alineación global o la región de similitud en una alineación local. Cuando se utiliza el porcentaje de identidad de secuencia en referencia a las proteínas, se reconoce que las posiciones de los residuos que no son idénticas a menudo difieren por sustituciones conservadoras de aminoácidos, donde los residuos aminoácidos están sustituidos por otros residuos de aminoácidos con propiedades químicas similares (por ejemplo, carga o hidrofobicidad) y por lo tanto no cambian las propiedades funcionales de la molécula. Cuando las secuencias difieren en sustituciones conservadoras, se puede ajustar el porcentaje de identidad de la secuencia hacia arriba para corregir la naturaleza conservadora de la sustitución. Las secuencias que se diferencian en tales sustituciones conservadoras se dice que tienen "similitud de secuencia" o "similitud". Los medios para hacer este ajuste son bien conocidos por aquellos capacitados en el arte. Generalmente, esto implica dar un puntaje a una sustitución conservadora como una correspondencia parcial en lugar de una falta de correspondencia, incrementando así el porcentaje de similitud de la secuencia.

Como se usa aquí, "porcentaje de identidad de secuencia" se refiere a un valor determinado señalando primero en dos secuencias óptimamente alineadas sobre una ventana de comparación, ya sea en forma global o localizada, en cada posición constituyente en cuanto a si se presentan en ambas secuencias la misma base de ácidos nucleicos o

- de residuos de aminoácidos, lo que denota una correspondencia, o no, lo que denotaría una falta de correspondencia. Como dicha alineación se construye mediante la optimización del número de bases coincidentes, mientras que al mismo tiempo se permiten desajustes en cualquier posición y la introducción de espacios de tamaño arbitrario, o regiones nulas o vacías ya que al hacerlo así aumenta la importancia o calidad de la alineación, el cálculo determina el número total de posiciones para las cuales existe una condición de correspondencia, y luego se divide este número por el número total de posiciones en la ventana de comparación, y por último se multiplica el resultado por 100 para obtener el porcentaje de identidad de secuencia. El "porcentaje de similitud de secuencia" para las secuencias de proteína se puede calcular utilizando el mismo principio, en el que la sustitución conservadora se calcula como una falta parcial de correspondencia en lugar de una falta total de correspondencia. Así, por ejemplo, en donde a un aminoácido idéntico se le da una puntuación de 1 y a una sustitución no conservadora se le da una puntuación de cero, a una sustitución conservadora se le da una puntuación entre cero y uno. La puntuación de las sustituciones conservadoras se puede obtener a partir de matrices de aminoácidos conocidas en el arte, por ejemplo, matrices Blosum o PAM.
- Los métodos de alineación de secuencias para comparación son bien conocidos en la técnica. La determinación del porcentaje de identidad o el porcentaje de similitud (para proteínas) entre dos secuencias se puede lograr utilizando un algoritmo matemático. Los ejemplos no limitantes preferidos de tales algoritmos matemáticos son, el algoritmo de Myers y Miller (Bioinformatics, 4 (1): 11 - 17, 1988), la alineación global de Needleman - Wunsch (J. Mol. Biol. 48 (3): 443 - 53, 1970), la alineación local de Smith - Waterman (J. Mol. Biol., 147: 195 - 197, 1981), el método de búsqueda de similitud de Pearson y Lipman (PNAS, 85 (8.): 2444 - 2448, 1988), el algoritmo de Karlin y Altschul (J. Mol. Biol., 215 (3.): 403 - 410, 1990; PNAS, 90: 5873 - 5877, 1993). Se pueden utilizar las implementaciones por ordenador de estos algoritmos matemáticos para la comparación de las secuencias para determinar la identidad de secuencia o para identificar homólogos. Tales implementaciones incluyen, pero no se limitan a, los programas descritos a continuación.
- El término "región conservada" o "dominio conservado" como se usan aquí se refieren a una región en las secuencias heterólogas de polinucleótidos o de polipéptidos donde existe un grado relativamente alto de identidad de secuencia entre las distintas secuencias. La "región conservada" puede ser identificada, por ejemplo, a partir de la alineación de secuencias múltiples utilizando el algoritmo de ClustalW.
- El término "célula" o "célula de una planta" tal como se utilizan aquí se refieren a una sola célula, y también incluyen una población de células. La población puede ser una población pura que contiene un tipo de célula. Asimismo, la población puede incluir más de un tipo de célula. Una célula de una planta en el sentido de la invención puede estar aislada (por ejemplo, en cultivo en suspensión) o incluida en un tejido de una planta, órgano de una planta o planta en cualquier etapa de desarrollo.
- El término "tejido" con respecto a una planta (o "tejido vegetal") significa la disposición de múltiples células de plantas, incluyendo tejidos de plantas diferenciados y no diferenciados. Los tejidos vegetales puede constituir parte de un órgano de la planta (por ejemplo, la epidermis de una hoja de la planta), pero también pueden constituir tejidos tumorales (por ejemplo, tejido de calloso) y diferentes tipos de células en cultivo (por ejemplo, células individuales, protoplastos, embriones, callos, cuerpos tipo protocormo, etc.). Los tejidos de la planta puede estar en una planta, en un cultivo de órganos, en un cultivo de tejidos o en cultivos celulares.
- El término "órgano" con respecto a una planta (u "órgano de una planta") significa partes de una planta y pueden incluir, pero sin limitarse a, por ejemplo, raíces, frutos, brotes, tallos, hojas, hipocotiledones, cotiledones, anteras, sépalos, pétalos, polen, semillas, etc.
- El término "planta" como se usa aquí, dependiendo del contexto, puede entenderse que se refiere a plantas enteras, células de plantas, órganos de plantas, semillas de plantas, y la progenie de las mismas. La palabra "planta" se refiere también a cualquier planta, en particular, a plantas de semilla, y pueden incluir, pero sin limitarse a, plantas de cultivo. Partes de la planta incluyen, pero no se limitan a, tallos, raíces, brotes, frutos, óvulos, estambres, hojas, embriones, regiones meristemáticas, tejido calloso, gametófitos, esporofitos, polen, microesporas, hipocotiledones, cotiledones, anteras, sépalos, pétalos, polen, semillas y similares. La clase de plantas que pueden ser utilizadas en el método de la invención es generalmente tan amplia como la clase de plantas superiores e inferiores susceptibles a técnicas de transformación, incluyendo angiospermas (plantas monocotiledóneas y dicotiledóneas), gimnospermas, helechos, colas de caballo, psilofitas, briofitas, y algas multicelulares.
- El término "transgénico" como se usa aquí, pretende referirse a las células y / o plantas que contienen un transgén, o cuyo genoma ha sido alterado por medio de la introducción de un transgén, o que tienen incorporado genes exógenos o polinucleótidos. Las células, tejidos, órganos y plantas transgénicas pueden ser producidos por diferentes métodos, incluyendo la introducción de un "transgén" que comprende un polinucleótido (usualmente ADN) en una célula objetivo o la integración del transgén en un cromosoma de una célula objetivo por medio de la intervención humana, tal como mediante los métodos descritos aquí.

El término "línea genéticamente pura" como se usa aquí se refiere a una variedad de planta para un rasgo particular si esta es genéticamente homocigota para ese rasgo hasta un grado en que, cuando la variedad de la línea genéticamente pura es autógama, no se observa una cantidad significativa de segregación independiente del rasgo entre la progenie.

- 5 El término "de tipo silvestre" como se usa aquí se refiere a una célula de la planta, semilla, componente de la planta, tejido de la planta, órgano de la planta, o planta entera que no ha sido modificada genéticamente o tratada en sentido experimental.

- 10 El término "planta de control" o "planta de tipo silvestre" como se usan aquí se refieren a una célula de una planta, un explante, semilla, componente de una planta, tejido de una planta, órgano de una planta, o una planta entera utilizados para comparación contra una planta transgénica o genéticamente modificada para el propósito de identificar un fenotipo mejorado o un rasgo deseable en la planta transgénica o genéticamente modificada. Una "planta de control" en algunos casos puede ser una línea de una planta transgénica que comprende un vector vacío o gen marcador, pero que no contiene el polinucleótido recombinante de interés que está presente en la planta transgénica o genéticamente modificada que está siendo evaluada. Una planta de control puede ser una planta de la misma línea o variedad que la planta transgénica o genéticamente modificada que está siendo analizada, o puede ser otra línea o variedad, tal como una planta que se sabe que tiene un fenotipo específico, característico, o genotipo conocido. Una planta de control adecuada incluiría una planta genéticamente no alterado o no transgénica de la línea progenitora utilizada para generar una planta transgénica aquí.

- 20 El término "resistente a la infección por nematodos" o "una planta que tiene resistencia a los nematodos" como se usa aquí se refiere a la capacidad de una planta para evitar la infección por nematodos, para matar los nematodos o para dificultar, reducir o detener el desarrollo, crecimiento o multiplicación de los nematodos. Esto podría lograrse por medio de un proceso activo, por ejemplo, mediante la producción de una sustancia perjudicial para el nematodo, o por medio de un proceso pasivo, como tener un valor nutricional reducido para el nematodo o estructuras que no se desarrollan inducidas por el sitio alimentación del nematodo como células sincitiales o gigantes. El nivel de resistencia a los nematodos de una planta se puede determinar de diferentes maneras, por ejemplo, por medio del recuento de los nematodos que son capaces de establecer parasitismo en esa planta, o midiendo los tiempos de desarrollo de los nematodos, la proporción de nematodos macho y hembra o el número de quistes o de huevos de nematodos producidos. Una planta con una mayor resistencia a la infección por nematodos es una planta, que es más resistente a la infección por nematodos en comparación con otra planta que tiene un genotipo similar o preferiblemente un genotipo idéntico que carece del gen o de los genes que confieren mayor resistencia a los nematodos, por ejemplo, una planta de control o de tipo silvestre.

- 30 El término "sitio de alimentación" o "sitio de sincitios" se usan indistintamente y se refieren como se usa aquí al sitio de alimentación formado en las raíces de la planta después de la infestación por nematodos. El sitio se utiliza como una fuente de nutrientes para los nematodos. Los sincitios son el sitio de alimentación para los nematodos de quiste y las células gigantes son los sitios de alimentación de los nematodos agalladores.

- 40 En una realización, la invención proporciona una planta transgénica transformada con un vector de expresión que comprende un polinucleótido que codifica alanina racemasa aislada, en donde la expresión del polinucleótido le confiere a la planta una mayor resistencia a los nematodos. Preferiblemente, el polinucleótido que codifica alanina racemasa se selecciona del grupo que consiste de un polinucleótido que tiene una secuencia tal como se define en las SEQ ID NOs: 5 ó 7; un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en las SEQ ID NOs: 6 u 8; un polinucleótido que tiene al menos 70% de identidad de secuencia con un polinucleótido que tiene una secuencia tal como se define en las SEQ ID NOs: 5 ó 7; un polinucleótido que codifica un polipéptido que tiene al menos 70% de identidad de secuencia con un polipéptido que tiene una secuencia tal como se define en las SEQ ID NOs: 6 u 8; un polinucleótido que hibrida bajo condiciones rigurosas con un polinucleótido que tiene una secuencia tal como se define en las SEQ ID NOs: 5 ó 7; un polinucleótido que se hibrida bajo condiciones rigurosas con un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en las SEQ ID NOs: 6 u 8; un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en cualquiera de las SEQ ID NOs: 12 a 44, y un polinucleótido que codifica un polipéptido que tiene al menos 90% de identidad de secuencia con cualquiera de las secuencias como se define en cualquiera de las SEQ ID NOs: 12 a 44.

- 50 Un polinucleótido que codifica alanina racemasa como se define aquí abarca también homólogos, ortólogos, parálogos y variantes alélicas del polinucleótido que codifica alanina racemasa de las SEQ ID NOs: 5 ó 7, o un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en las SEQ ID NOs: 6 u 8 o como se define en cualquiera de las SEQ ID NOs: 12 a 44. Tal como se usa aquí, el término "variante alélica" se refiere a un polinucleótido que contiene polimorfismos que conducen a cambios en las secuencias de aminoácidos de una proteína codificada por el nucleótido y que existen dentro de una población natural (por ejemplo, una especie o variedad de una planta). Tales variaciones alélicas naturales típicamente puede resultar en una varianza de 1 - 5% en un polinucleótido que codifica una proteína, o una varianza de 1 - 5% en la proteína codificada. Las variantes alélicas pueden ser identificadas por medio de la secuenciación del ácido nucleico de interés en una cantidad de

plantas diferentes, que puede ser fácilmente llevada a cabo utilizando, por ejemplo, sondas de hibridación para identificar el mismo locus genético del gen en esas plantas. Todas y cada una de tales variaciones del ácido nucleico en un polinucleótido y los polimorfismos resultantes del aminoácido o las variaciones de una proteína que son el resultado de una variante alélica natural y que no alteran la actividad funcional de la proteína codificada, se pretende que estén dentro del alcance de la invención. Para clonar variantes alélicas u homólogos de los polinucleótidos de la invención, se puede utilizar la información de la secuencia dada en este documento. Por ejemplo, los iniciadores descritos por las SEQ ID NOs: 1, 2, 3 y 4 puede ser utilizados para clonar variantes alélicas u homólogos.

En aún otra realización, la planta puede ser una planta seleccionada del grupo que consiste de plantas monocotiledóneas y plantas dicotiledóneas. La planta puede ser de un género seleccionado del grupo que consiste en maíz, trigo, arroz, cebada, avena, centeno, sorgo, plátano, y ballico. La planta puede ser de un género seleccionado del grupo que consiste de arveja, alfalfa, soja, zanahoria, apio, tomate, patata, algodón, tabaco, pimienta, colza, remolacha, col, coliflor, brócoli, lechuga y *Arabidopsis thaliana*.

La presente invención también proporciona semilla transgénica que es una línea genéticamente pura para un polinucleótido que codifica alanina racemasa, y partes de las plantas transgénicas que contienen al polinucleótido que codifica alanina racemasa, y plantas de la progenie de dicha planta, incluidas híbridas y endogámicas. La invención también proporciona un método de fitomejoramiento de plantas, por ejemplo, para preparar una planta transgénica fértil cruzada. El método comprende el cruzamiento de una planta transgénica fértil que contiene un vector de expresión particular de la invención consigo misma o con una segunda planta, por ejemplo, una que carece del vector de expresión particular, para preparar la semilla de una planta transgénica fértil cruzada que contiene al vector de expresión particular. La semilla se planta luego para obtener una planta transgénica fértil cruzada. La planta puede ser una monocotiledónea o una dicotiledónea. La planta transgénica fértil cruzada puede tener al vector de expresión particular, heredado a través de un progenitor hembra o a través de un progenitor macho. La segunda planta puede ser una planta endogámica. La planta transgénica fértil cruzada puede ser un híbrido. También se incluyen dentro de la presente invención las semillas de cualquiera de esas plantas transgénicas fértiles cruzadas.

Otra realización de la invención se relaciona con el uso de un vector de expresión que comprende un elemento regulador de la transcripción operativamente enlazado a un polinucleótido que codifica alanina racemasa, en donde la expresión del polinucleótido confiere una mayor resistencia a los nematodos a una planta transgénica. En una realización, el elemento regulador de la transcripción es un promotor capaz de regular la expresión constitutiva de un polinucleótido operativamente enlazado. Un "promotor constitutivo" se refiere a un promotor que es capaz de expresar el marco de lectura abierto o el elemento regulador que éste controla en todos o casi todos los tejidos de la planta durante todas o casi todas las etapas de desarrollo de la planta. Los promotores constitutivos incluyen, pero no se limitan a, el promotor 35S del CaMV de virus de las plantas (Franck et al., Cell 21: 285 - 294, 1980.), el promotor Nos (An G. et al, The Plant Cell 3: 225 - 233, 1990), el promotor de ubiquitina (Christensen et al., Plant Mol Biol. 12: 619 - 632, 1992 y 18: 581 - 8, 1991), el promotor MAS (Velten et al., EMBO J. 3: 2723 - 30, 1984), el promotor de la histona H3 del maíz (Lepetit et al., Mol. Gen. Genet 231: 276 - 85, 1992), el promotor ALS (WO96/30530), el promotor 19S del CaMV (Patente de los Estados Unidos No. 5.352.605), el súper promotor (Patente de los Estados Unidos No. 5.955.646), el promotor del virus del mosaico de escrofularia (Patente de los Estados Unidos No. 6.051,753), el promotor de actina del arroz (Patente de los Estados Unidos No. 5.641, 876), y el promotor de la subunidad pequeña de Rubisco (Patente de los Estados Unidos No. 4.962.028).

En otra realización, el elemento regulador de la transcripción es un promotor regulado. Un "promotor regulado" se refiere a un promotor que dirige la expresión génica en forma no constitutiva, pero en una forma temporal y / o espacial, e incluye tanto promotores inducibles como específicos del tejido. Promotores diferentes pueden dirigir la expresión de un gen o elemento regulador en diferentes tejidos o tipos de células, o en diferentes fases de desarrollo, o en respuesta a diferentes condiciones ambientales.

Un "promotor específico del tejido" se refiere a un promotor regulado que no se expresa en todas las células de una planta, sino únicamente en uno o más tipos de células en órganos específicos (por ejemplo, hojas o semillas), tejidos específicos (por ejemplo, embrión o cotiledón) o tipos específicos de células (por ejemplo el parénquima de la hoja o células de almacenamiento de las semillas). También se incluyen los promotores que están temporalmente regulados, tal como en la embriogénesis temprana o tardía, durante la maduración del fruto en semillas o fruto en desarrollo, en la hoja completamente diferenciada, o al comienzo de la secuencia. Los promotores adecuados incluyen el promotor del gen de la napina de colza (Patente de los Estados Unidos No. 5.608.152), la promotor de USP de Vicia faba (Baeumlein et al., Mol Gen Genet. 225 (3): 459 - 67, 1991), el promotor de oleosina de *Arabidopsis* (WO 98/45461), la promotor de faseolina de *Phaseolus vulgaris* (Patente de los Estados Unidos No. 5.504.200), al promotor de Bce4 de *Brassica* (WO 91/13980) o al promotor B4 de legúmina (LeB4; Baeumlein et al., Plant Journal, 2 (2): 233 - 9, 1992), así como promotores que confieren expresión específica de la semilla en plantas monocotiledóneas como el maíz, la cebada, el trigo, el centeno, el arroz, etc. Los promotores adecuados para tener en cuenta son el promotor del gen *lpt2* o *lpt1* de cebada (WO 95/15389 y WO 95/23230) o los que se describen en el documento WO 99/16890 (promotores del gen de la hordelina de la cebada, el gen de la glutelina del arroz, el gen

de la orizina del arroz, el gen de la prolamina del arroz, el gen de la gliadina del trigo, el gen de la glutelina del trigo, el gen de la zeína del maíz, el gen de la glutelina de la avena, el gen de la kasirina del sorgo y el gen de la secalina del centeno). Los promotores adecuados para expresión preferencial en tejidos de la raíz de plantas incluyen, por ejemplo, al promotor derivado del gen de la nicotinamina sintasa del maíz (solicitud de patente de los Estados Unidos No. 2003/0131377) y el promotor RCC3 del arroz (solicitud de patente de los Estados Unidos No. 11/075.113). Los promotores adecuados para la expresión preferencial en los tejidos vegetales verdes incluyen a los promotores de genes tales como el gen FDA de la aldolasa del maíz (solicitud de patente de los Estados Unidos No. 2004/0216189), aldolasa y piruvato ortofosfato diquinasa (PPDK) (Taniguchi et al., Plant Cell Physiol. 41 (1): 42 - 48, 2000).

"Promotores inducibles" se refiere a aquellos promotores regulados que pueden ser activados en uno o más tipos de células por un estímulo externo, por ejemplo, un compuesto químico, luz, hormonas, estrés, o un agente patógeno tales como nematodos. Los promotores químicamente inducibles son especialmente adecuados si se desea que la expresión del gen se produzca de una forma específica en el tiempo. Los ejemplos de tales promotores son un promotor inducible por ácido salicílico (WO 95/19443), un promotor inducible por tetraciclina (Gatz et al., Plant J. 2: 397 - 404, 1992), al promotor inducible por la luz de la subunidad pequeña de la Ribulosa-1,5-bis-fosfato carboxilasa (ss-RUBISCO), y un promotor inducible por etanol (WO 93/21334). Además, los promotores adecuados que respondan a condiciones de estrés biótico o abiótico son aquellos tales como el promotor del gen PRP1 inducible por el patógeno (Ward et al., Plant. Mol. Biol. 22: 361 - 366, 1993), el promotor del tomate hsp80 inducible por calor (Patente de los Estados Unidos No. 5.187.267), promotor de alfa-amilasa inducible por frío de la patata (WO 96/12814), el promotor inducible por sequía del maíz (Busk et al., Plant J. 11: 1285 - 1295, 1997), el promotor inducible por frío, sequía, y alta salinidad de la patata (Kirch, Plant Mol. Biol. 33: 897 - 909, 1997) o el promotor RD29A de Arabidopsis (Yamaguchi-Shinozaki et al., Mol. Gen. Genet. 236: 331 - 340, 1993), muchos promotores inducibles por frío tales como el promotor de cor15a de Arabidopsis (número de acceso del GenBank U01377), but101 y blt4.8 de la cebada (números de acceso del GenBank AJ310994 y U63993), wcs120 de trigo (número de acceso del GenBank AF031235), mlip15 del maíz (número de acceso del GenBank D26563), bn115 de Brassica (número de acceso del GenBank U01377), y el promotor de pinII inducible por lesiones (Patente Europea No. 375091).

Los promotores preferidos son los promotores específicos de la raíz, específicos del sitio de alimentación, inducibles por patógenos o inducibles por nemátodos.

Se divulga un método para producir una planta transgénica que comprende un polinucleótido que codifica alanina racemasa, en donde el método comprende las etapas de introducir en la planta el vector de expresión que contiene al polinucleótido que codifica alanina racemasa; y la selección de plantas transgénicas para mayor resistencia a los nematodos.

Se conocen una variedad de métodos para la introducción de polinucleótidos en el genoma de las plantas y para la regeneración de plantas a partir de tejidos de la planta o de células de la planta, por ejemplo, en Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), capítulo 6/7, páginas 71 - 119 (1993); White FF (1993) Vectors for Gene Transfer in Higher Plants; Transgenic Plants, vol. 1, Engineering and Utilization, Ed.: Kung and Wu R, Academic Press, 15 - 38; Jené B et al. (1993) Techniques for Gene Transfer; Transgenic Plants, vol. 1, Engineering and Utilization, Ed.: Kung and R. Wu, Academic Press, páginas 128 - 143; Potrykus (1991) Annu Rev Plant Physiol Plant Molec Biol 42: 205 - 225; Halford NG, Shewry PR (2000) Br Med Bull 56(1): 62 - 73.

Los métodos de transformación puede incluir métodos directos e indirectos de transformación. Los métodos directos adecuados incluyen la incorporación de ADN inducida por polietilén glicol, transformación mediada por liposomas (Patente de los Estados Unidos No. 4.536.475), métodos biolísticos utilizando la pistola de genes (Fromm M. E. et al, Bio / Technology 8 (9): 833 - 9, 1990; Gordon-Kamm et al., Plant Cell 2: 603, 1990), electroporación, incubación de embriones secos en solución que contiene ADN, y microinyección. En el caso de estos métodos de transformación directa, el plásmido utilizado no tienen que cumplir con ningún requerimiento particular. Los plásmidos simples, tales como aquellos de la serie pUC, pBR322, de la serie M13mp, pACYC184 y similares, pueden ser utilizados. Si se van a regenerar plantas intactas a partir de las células transformadas, se localiza preferiblemente un gen marcador seleccionable adicional en el plásmido. Las técnicas de transformación directa son igualmente adecuadas para plantas dicotiledóneas y monocotiledóneas.

La transformación también puede llevarse a cabo por infección bacteriana por medio de Agrobacterium (por ejemplo el documento EP 0 116 718), infección viral por medio de vectores virales (EP 0 067 553 ; Patente de los Estados Unidos No. 4.407.956, WO 95/34668, WO 93/03161) o por medio de polen (EP 0 270 356, WO 85/01856; Patente de los Estados Unidos No. 4.684.611). Técnicas de transformación basadas en Agrobacterium (especialmente para plantas dicotiledóneas) son bien conocidas en el arte. La cepa de Agrobacterium (por ejemplo, Agrobacterium tumefaciens o Agrobacterium rhizogenes) comprende un plásmido (plásmido Ti o Ri) y un elemento de ADN-T que es transferido a la planta después de la infección con Agrobacterium. El ADN-T (ADN transferido) se integra dentro del genoma de la célula de la planta. El ADN-T puede ser localizado sobre el plásmido Ti o Ri o incluido

separadamente en el así llamado vector binario. Los métodos para transformación mediada por *Agrobacterium* se describen, por ejemplo, en Horsch R. B. et al. (1985) *Science* 225: 1229. La transformación mediada por *Agrobacterium* se adapta mejor a las plantas dicotiledóneas pero también ha sido adoptada para las plantas monocotiledóneas. La transformación de plantas por *Agrobacterias* se describe, por ejemplo, en White FF, *Vectors for Gene Transfer in Higher Plants, Transgenic Plants*, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, 1993, páginas 15 - 38; Jené B et al. *Techniques for Gene Transfer, Transgenic Plants*, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, 1993, páginas 128 - 143; Potrykus (1991) *Annu Rev Plant Physiol Plant Molec Biol* 42: 205 - 225.

La transformación puede resultar en transformación transitoria o estable y expresión. Aunque un polinucleótido que codifica alanina racemasa se puede insertar en cualquier planta o célula de una planta que cae dentro de estas clases amplias, es particularmente útil en células de plantas de cultivo.

Los polinucleótidos que codifican alanina racemasa pueden ser transformados directamente en el genoma del plástido. La expresión del plástido, en cuyos genes se insertan por recombinación homóloga en las diferentes miles de copias del genoma circular del plástido presente en cada célula de la planta, toma ventaja de la enorme ventaja del número de copias sobre los genes expresados en el núcleo para permitir altos niveles de expresión. En una realización, se insertan los nucleótidos en un vector plástido objetivo y se transforma en el genoma del plástido de una planta huésped deseada. Se obtienen plantas homoplásmicas para genomas de plástidos que contienen las secuencias de nucleótidos, y son preferentemente capaces de alta expresión de los nucleótidos.

La tecnología de transformación plástidos es, por ejemplo ampliamente descrita en las Patentes de los Estados Unidos Nos. 5.451.513, 5.545.817, 5.545.818, y 5.877.462, en los documentos WO 95/16783 y WO 97/32977, y en McBride et al. (1994) *PNAS* 91, 7301 - 7305, todas incorporadas aquí como referencia en su totalidad. La técnica básica para la transformación de plástidos implica la introducción de regiones de ADN del plástido clonado que flanquean un marcador seleccionable junto con la secuencia de nucleótidos en un tejido objetivo adecuado, por ejemplo, utilizando transformación biolística o de protoplastos (por ejemplo, transformación mediada por cloruro de calcio o por PEG). Las regiones de flaqueo de 1 a 1,5 kb, denominada secuencias objetivo, facilitan la recombinación homóloga con el genoma del plástido y permitiendo así la sustitución o modificación de regiones específicas del plástido. Inicialmente, se utilizan mutaciones puntuales en el ARNr 16S del cloroplasto y genes *rps12* que confieren resistencia a la espectinomicina y / o a la estreptomycinina como marcadores seleccionables para transformación (Svab et al., *PNAS* 87, 8526 - 8530, 1990; Staub et al., *Plant Cell* 4, 39 - 45, 1992). La presencia de sitios de clonación entre estos marcadores permite la creación de un vector objetivo plástido para la introducción de genes foráneos (Staub et al., *EMBO J.* 12, 601 - 606, 1993). Aumentos sustanciales en la frecuencia de transformación se obtienen mediante la sustitución del ARNr recesivo o los genes de resistencia a los antibióticos para la proteína *r* con un marcador seleccionable dominante, el gen bacteriano *aadA* que codifica la enzima desintoxicante de espectinomicina, aminoglicósido-3'-adeniltransferasa (Svab et al., *PNAS* 90, 913 - 917, 1993). Se conocen en el arte otros marcadores seleccionables útiles para la transformación de plástidos y están abarcados dentro del alcance de la invención.

La planta o la planta transgénica pueden ser cualquier planta, tales como, pero sin limitarse a árboles, flores cortadas, plantas ornamentales, vegetales o plantas de cultivo. La planta puede ser de un género seleccionado del grupo que consiste de *Medicago*, *Lycopersicon*, *Brassica*, *Cucumis*, *Solanum*, *Juglans*, *Gossypium*, *Malus*, *Vitis*, *Antirrhinum*, *Populus*, *Fragaria*, *Arabidopsis*, *Picea*, *Capsicum*, *Chenopodium*, *Dendranthema*, *Pharbitis*, *Pinus*, *Pisum*, *Oryza*, *Zea*, *Triticum*, *Triticale*, *Secale*, *Lolium*, *Hordeum*, *glicina*, *Pseudotsuga*, *Kalanchoe*, *Beta*, *Helianthus*, *Nicotiana*, *Cucurbita*, *Rosa*, *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranio*, *Manihot*, *Daucus*, *Raphanus*, *Sinapis*, *Atropa*, *Datura*, *Hyoscyamus*, *Nicotiana*, *Petunia*, *Digitalis*, *Majorana*, *Ciahorium*, *Lactuca*, *Bromus*, espárragos, *Antirrhinum*, *Heterocallis*, *Némesis*, *Pelargonium*, *Panieum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Browaalia*, *Phaseolus*, *Avena*, y *Allium*, o se puede seleccionar la planta del grupo que consiste de cereales como trigo, cebada, sorgo, centeno, triticale, maíz, arroz, caña de azúcar, y de árboles como manzana, pera, membrillo, ciruela, cereza, melocotón, nectarina, albaricoque, papaya, mango, álamo, pino, secoya, cedro y roble. El término "planta" como se usa aquí, pueden ser para plantas de cultivo dicotiledóneas, tales como guisantes, alfalfa, soja, zanahoria, apio, tomate, patata, algodón, tabaco, pimienta, colza, remolacha, col, coliflor, brócoli, lechuga y *Arabidopsis thaliana*. En una realización, la planta es una planta monocotiledónea o una planta dicotiledónea.

Preferiblemente, la planta es una planta de cultivo. Las plantas de cultivo son todas plantas utilizadas en la agricultura. Por consiguiente, en una realización la planta es una planta monocotiledónea, preferiblemente una planta de la familia Poaceae, Musaceae, Liliaceae o Bromeliaceae, preferentemente de la familia Poaceae. Por consiguiente, en aún otra realización la planta es una planta Poaceae del género *Zea*, *Triticum*, *Oryza*, *Hordeum*, *Secale*, *Avena*, *Saccharum*, *Sorghum*, *Pennisetum*, *Setaria*, *Panieum*, *Eleusine*, *Miscanthus*, *Brachypodium*, *Festuca* o *Lolium*. Cuando la planta es del género *Zea*, la especie preferida es *Z. mays*. Cuando la planta es del género *Triticum*, la especie preferida es *T. aestivum*, *T. speltae* o *T. durum*. Cuando la planta es del género *Oryza*, la especie preferida es *O. sativa*. Cuando la planta es del género *Hordeum*, la especie preferida es *H. vulgare*.

Cuando la planta es del género *Secale*, la especie preferida es *S. cereale*. Cuando la planta es del género *Avena*, la especie preferida es *A. sativa*. Cuando la planta es del género *Saccarum*, la especie preferida es *S. officinarum*. Cuando la planta es del género *Sorghum*, la especie preferida es la de *S. vulgare*, *S. bicolor* o *S. sudanense*. Cuando la planta es del género *Pennisetum*, la especie preferida es *P. glaucum*. Cuando la planta es del género *Setaria*, la especie preferida es *S. italica*. Cuando la planta es del género *Panicum*, la especie preferida es *P. miliaceum* o *P. virgatum*. Cuando la planta es del género *Eleusine*, la especie preferida es *E. coracana*. Cuando la planta es del género *Miscanthus*, la especie preferida es *M. sinensis*. Cuando la planta es una planta del género *Festuca*, la especie preferida es *F. arundinaria*, *F. rubra* o *F. pratensis*. Cuando la planta es del género *Lolium*, la especie preferida es *L. perenne* o *L. multiflorum*. Alternativamente, la planta puede ser *Triticosecale*.

Alternativamente, en una realización la planta es una planta dicotiledónea, preferentemente una planta de la familia *Fabaceae*, *Solanaceae*, *Brassicaceae*, *Chenopodiaceae*, *Asteraceae*, *Malvaceae*, *Linaceae*, *Euphorbiaceae*, *Rosaceae*, *Convolvulaceae*, *Cucurbitaceae*, *Theaceae*, *Rubiaceae*, *Sterculiaceae* o *Citrus*. En una realización, la planta es una planta de la familia *Fabaceae*, *Solanaceae* o *Brassicaceae*. Por consiguiente, en una realización la planta es de la familia *Fabaceae*, preferiblemente del género *Glycine*, *Pisum*, *Arachis*, *Cicer*, *Vicia*, *Phaseolus*, *Lupinus*, *Medicago* o *Lens*. Las especies preferidas de la familia *Fabaceae* son *M. truncatula*, *H. sativa*, *G. max*, *P. sativum*, *A. hypogea*, *C. arietinum*, *V. faba*, *P. vulgaris*, *Lupinus albus*, *Lupinus luteus*, *Lupinus angustifolius* o *Lens culinaris*. Más preferidas son la especies *G. max*, *A. hypogea* y *M. sativa*. La más preferida es la especie *G. max*. Cuando la planta es de la familia *Solanaceae*, el género preferido es *Solanum*, *Lycopersicon*, *Nicotiana* o *Capsicum*. Las especies preferidas de la familia *Solanaceae* son *S. tuberosum*, *L. esculentum*, *N. tabaccum* o *C. chinense*. Más preferida es *S. tuberosum*. Por consiguiente, en una realización la planta es de la familia *Solanaceae*, preferiblemente del género *Brassica* o *Raphanus*. Las especies preferidas de la familia *Brassicaceae* son las especies *B. napus*, *B. oleracea*, *B. juncea* y *B. rapa*. Más preferida es la especie *B. napus*. Cuando la planta es de la familia *Chenopodiaceae*, el género preferido es *Beta* y la especie preferida es la *B. vulgaris*. Cuando la planta es de la familia *Asteraceae*, el género preferido es *Helianthus* y la especie preferida es *H. annuus*. Cuando la planta es de la familia *Malvaceae*, el género preferido es *Gossypium* o *Abelmoschus*. Cuando el género es *Gossypium*, la especie preferida es *G. hirsutum* o *G. barbadense* y la especie más preferidas es *G. hirsutum*. Una especie preferida del género *Abelmoschus* es la especie *A. esculentus*. Cuando la planta es de la familia *Linaceae*, el género preferido es *Linum* y la especie preferida es *L. usitatissimum*. Cuando la planta es de la familia *Euforbiaceae*, el género preferido es *Manihot*, *Jatropha* o *Rhizinus* y las especies preferidas son *M. esculenta*, *J. curcas* o *R. communis*. Cuando la planta es de la familia *Convolvulaceae*, el género preferido es *Ipomea* y la especie preferida es *I. batatas*. Cuando la planta es de la familia *Rosaceae*, el género preferido es *Rosa*, *Malus*, *Pyrus*, *Prunus*, *Rubus*, *Ribes*, *Vaccinium* o *Fragaria* y la especie preferida es el híbrido *Fragaria xananassa*. Cuando la planta es de la familia *Cucurbitaceae*, el género preferido es *Cucumis*, *Citrullus* o *Cucurbita* y la especie preferida es *Cucumis sativus*, *Citrullus lanatus* o *Cucurbita pepo*. Cuando la planta es de la familia *Theaceae*, el género preferido es *Camellia* y la especie preferida es *C. sinensis*. Cuando la planta es de la familia *Rubiaceae*, el género preferido es *Coffea* y la especie preferida es *C. arábica* o *C. canephora*. Cuando la planta es de la familia *Sterculiaceae*, el género preferido es *Theobroma* y la especie preferida es *T. cacao*. Cuando la planta es del género *Citrus*, la especie preferida es *C. sinensis*, *C. limon*, *C. reticulata*, *C. maxima* y los híbridos de especies *Citrus*, o similares. En una realización preferida de la invención, la planta es una planta de soja, de patata o de maíz.

Las plantas transgénicas de la invención pueden ser utilizadas en un método para control de la infestación de un cultivo por un nematodo parásito de las plantas, que comprende la etapa de cultivar dicho cultivo de semillas que contienen un casete de expresión que comprende un elemento regulador de la transcripción enlazado operativamente a un polinucleótido que codifica una racemasa alanina, en donde el casete de expresión está integrado de forma estable en los genomas de las semillas y la planta tiene mayor resistencia a los nematodos.

La invención también proporciona un método para conferir resistencia a los nematodos a una planta, que comprende las etapas de a) transformar una célula de una planta con un casete de expresión de la invención, b) regenerar una planta a partir de esa célula y c) seleccionar dicha planta para resistencia a los nematodos. Más específicamente, el método para aumentar la resistencia a los nematodos en una planta comprende las etapas de introducir en la planta de un vector de expresión que contiene un elemento regulador de la transcripción operativamente enlazado a un polinucleótido de la invención, en donde la expresión del polinucleótido confiere una mayor resistencia a la planta contra los nematodos, y en donde el polinucleótido que codifica alanina racemasa se selecciona del grupo que consiste de un polinucleótido que tiene una secuencia tal como se define en las SEQ ID NOs: 5 ó 7; un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en las SEQ ID NO: 6 u 8; un polinucleótido que tiene al menos 70% de identidad de secuencia con un polinucleótido que tiene una secuencia tal como se define en las SEQ ID NOs: 5 ó 7; un polinucleótido que codifica un polipéptido que tiene al menos 70% de identidad de secuencia con un polipéptido que tiene una secuencia tal como se define en las SEQ ID NOs: 6 u 8; un polinucleótido que hibrida bajo condiciones rigurosas con un polinucleótido que tiene una secuencia tal como se define en las SEQ ID NOs: 5 ó 7; un polinucleótido que hibrida bajo condiciones rigurosas con un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en las SEQ ID NOs: 6 u 8; un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en cualquiera de las SEQ ID NOs: 12 a 44, y un polinucleótido que codifica un polipéptido que tiene al menos 90% de identidad de secuencia con cualquiera de las secuencias como se define en cualquiera de las SEQ ID NOs: 12 a 44.

La presente invención puede ser utilizada para reducir la destrucción de cultivos por parte de los nematodos

parásitos de las plantas o para conferir resistencia contra los nematodos a una planta. El nematodo puede ser cualquier nematodo parásito de las plantas, como los nematodos de las familias Longidoridae, Trichodoridae, Aphelenchoididae, Anguinidae, Belonolaimidae, Criconematidae, Heterodidae, Hoplolaimidae, Meloidogynidae, Paratylenchidae, Pratylenchidae, Tylenchulidae, Tylenchidae, o similares. Preferiblemente, los nematodos parásitos pertenecen a las familias de los nematodos que inducen células gigantes o sincitiales. Los nematodos que inducen células gigantes o sincitiales se encuentran en las familias Longidoridae, Trichodoridae, Heterodidae, Meloidogynidae, Pratylenchidae o Tylenchulidae. En particular, en las familias Heterodidae y Meloidogynidae.

En consecuencia, los nematodos parásitos objetivo de la presente invención pertenecen a uno o más géneros seleccionados entre el grupo de Nacobus, Cactodera, Dolichodera, Globodera, Heterodera, Punctodera, Longidorus o Meloidogyne. En una realización preferida, los nematodos parásitos pertenecen a uno o más géneros seleccionados entre el grupo de Nacobus, Cactodera, Dolichodera, Globodera, Heterodera, Meloidogyne o Punctodera. En una realización más preferida, los nematodos parásitos pertenecen a uno o más géneros seleccionados entre el grupo de Globodera, Heterodera o Meloidogyne. En una realización aún más preferida, los nematodos parásitos pertenecen a uno o a ambos géneros seleccionados entre el grupo de Globodera o Heterodera. En otra realización, los nematodos parásitos pertenecen al género Meloidogyne.

Cuando los nematodos parásitos son del género Globodera, las especies son preferentemente del grupo que consiste de *G. achilleae*, *G. artemisiae*, *G. hypolysi*, *G. mexicana*, *G. millefolii*, *G. mali*, *G. pallida*, *G. rostochiensis*, *G. tabacum*, y *G. virginiae*. En otra realización preferida, los nematodos parásitos del género Globodera incluyen al menos una de las especies *G. pallida*, *G. tabacum*, o *G. rostochiensis*. Cuando los nematodos parásitos son del género Heterodera, las especies pueden ser preferiblemente del grupo que consiste de *H. avenae*, *H. carotae*, *H. ciceri*, *H. cruciferae*, *H. delvii*, *H. elachista*, *H. filipjevi*, *H. gambiensis*, *H. glycines*, *H. goettingiana*, *H. graduni*, *H. humuli*, *H. hordecalis*, *H. latipons*, *H. major*, *H. medicaginis*, *H. oryzicola*, *H. pakistanensis*, *H. rosii*, *H. sacchari*, *H. schachtii*, *H. sorghi*, *H. trifolii*, *H. urticae*, *H. vigni* y *H. zae*. En otra realización preferida, los nematodos parásitos de Heterodera incluyen al menos una de las especies *H. glycines*, *H. avenae*, *H. cajani*, *H. Gottingiana*, *H. trifolii*, *H. Zeae* o *H. schachtii*. En una realización más preferida, los nematodos parásitos incluyen al menos una de las especies *H. glycines* o *H. schachtii*. En una realización más preferida, el nematodo parásito es la especie *H. glycines*.

Cuando los nematodos parásitos son del género Meloidogyne, el nematodo parásito puede ser seleccionado del grupo que consiste de *M. acronea*, *M. arabica*, *M. arenaria*, *M. artiellia*, *M. brevicauda*, *M. camelliae*, *M. chitwoodi*, *M. coeicola*, *M. esigua*, *M. graminicola*, *M. hapla*, *M. incognita*, *M. indica*, *M. inornata*, *M. javanica*, *M. lini*, *M. mali*, *M. microcephala*, *M. microtyla*, *M. naasi*, *M. salasi* y *M. thamesi*. En una realización preferida, los nematodos parásitos incluyen al menos una de las especies *M. javanica*, *M. incognita*, *M. hapla*, *M. arenaria* o *M. chitwoodi*.

En consecuencia, la invención comprende un procedimiento para conferir a una planta resistencia a los nematodos, comprendiendo dicho método las etapas de: a) preparar un casete de expresión que contiene un polinucleótido de la invención operativamente enlazado a un promotor, b) transformar una planta receptora con dicho casete de expresión, c) producir una o más descendencias transgénicas de dicha planta receptora; y d) seleccionar la progenie para resistencia a los nematodos. Preferiblemente el promotor es un promotor inducible por nemátodos o preferido de la raíz o un promotor que media la expresión en los sitios de alimentación de los nematodos, por ejemplo sincitios o células gigantes.

EJEMPLOS

Ejemplo 1: Clonación de los genes que codifican alanina racemasa

Las dos formas de alanina racemasa, Alr (SEQ ID NO: 7) y DadX (SEQ ID NO: 5), fueron clonados a partir del ADN genómico de *E. coli* utilizando los iniciadores de la PCR mostrados en la Tabla 1.

Tabla 1: Iniciadores utilizados para clonar los genes que codifican ARLNCP

Nombre del iniciador	Secuencia	Propósito	SEQ ID NO:
Iniciador 1 - DadX	GCGGCGCGCCACCATGACCCGTCCGATACAGGC	Iniciador 5' DadX	1
Iniciador 2 - DadX	GCCTCGAGTTACACCGTCACAACCGGGACGC	Iniciador 3' DadX	2
Iniciador 1 - Alr	GCGGCGCGCCACCATGCAAGCGGCAACTGTTGTG	Iniciador 5' Alr	3
Iniciador 2 - Alr	GCCTCGAGTTAATCCACGTATTTTCATCGCGAC	Iniciador 3' Alr	4

Ejemplo 2: Construcción de vectores para la transformación y la generación de raíces transgénicas

5 Los productos de PCR generados en el Ejemplo 1 fueron secuenciados y clonados en una cantidad de vectores de expresión que contienen los promotores preferidos de los sincitios (inducidos por nemátodos). Los promotores preferidos de los sincitios incluyen MTN3 de soja SEQ ID NO: 9 (p-47116125) (USSN 60/899.714), POX de Arabidopsis peroxidasa SEQ ID NO: 10 (p-At5g05340) (USSN 60/876.416) y TPP trehalosa-6- fosfato fosfatasa de Arabidopsis SEQ ID NO: 11 (p-AT1 g35910) (USSN 60/874.375). Se utilizó también el promotor súper constitutivo. El

10 marcador de selección para la transformación fue un gen mutado para la ácido acetohidroxi sintasa (AHAS) de Arabidopsis thaliana que confiere resistencia al herbicida ARSENAL (Imazepyr, BASF Corporation, Mount Olive, Nueva Jersey). La expresión de AHAS mutada fue conducida por el promotor de actina 2 de Arabidopsis.

Tabla 2. vector de expresión que comprende las SEQ ID NOs: 5 ó 7

vector	Composición del vector de expresión (promotor::gen codificador ARLNCP)
RSH118	MTN3::DadX
RSH120	POX::DadX
RSH122	TPP::DadX
RSH117	Súper Promotor::DadX
RSH125	MTN3::Alr
RSH127	POX::Alr
RSH129	TPP::Alr
RSH124	Súper Promotor::Alr

Ejemplo 3: Generación de bioensayo soja transgénica peluda de la raíz y nematodos

15 Los vectores RSH118, RSH120, RSH122, RSH117, RSH125, RSH127, RSH129 y RSH124 fueron transformados en la cepa K599 de A. rhizogenes por medio de electroporación. Las cepas transformadas de Agrobacterium fueron utilizadas para inducir la formación de raíces vellosas en la soja utilizando métodos conocidos. También se

20 generaron raíces vellosas no transgénicas de la variedad de cultivo Williams 82 de soja (susceptible al SCN) y Jack (resistente al SCN) mediante el uso de A. Rhizogenes no transformado, para que sirvieran como controles para el crecimiento de nematodos en el ensayo.

Se llevó a cabo un bioensayo para evaluar la resistencia a los nematodos sobre la raíz velluda transgénica transformada con los vectores y sobre raíces velludas no transgénicas de Williams 82 y Jack como controles. Los

25 cultivos de raíces velludas de cada línea que ocupaba al menos la mitad del pozo fueron inoculados con la raza 3 descontaminada en la superficie del nematodo juvenil del quiste de la soja (SCN) de segunda fase (J2). Las placas fueron luego selladas, y llevadas nuevamente a la incubadora a 25° C en la oscuridad. Se generaron varias líneas de raíces peludas independientes a partir de cada transformación del vector binario y se utilizaron las líneas para el

bioensayo. Cuatro semanas después de la inoculación del nemátodo, se hizo el recuento del número de quistes en cada pozo.

5 Los resultados del bioensayo para los constructos RSM118, RSH120, RSH122, y RSH125 muestran una reducción estadísticamente significativa (valor de $p < 0,05$) en el recuento de quistes sobre múltiples líneas transgénicas y una tendencia general de recuento reducido de quistes en la mayoría de las líneas transgénicas analizadas.

Aquellos capacitados en el arte reconocerán, o serán capaces de determinar utilizando únicamente experimentación rutinaria, muchos equivalentes a las realizaciones específicas de la invención aquí descrita.

LISTADO DE SECUENCIAS

<110> BASF Plant Science GmbH

10 Hill, Steve

<120> GENES PARA CONTROL DE NEMÁTODOS TIPO ALANINA RACEMASA Y METODOS DE USO EN PLANTAS

<130> PF 58854

<160> 44

15 <170> PatentIn versión 3.4

<210> 1

<211> 33

<212> ADN

<213> Escherichia coli

20 <400> 1

gcggcgcgcc accatgaccc gtccgataca ggc 33

<210> 2

<211> 31

<212> ADN

25 <213> Escherichia coli

<400> 2

gcctcgagtt acaccgtcac aaccgggacg c 31

<210> 3

<211> 34

30 <212> ADN

<213> Escherichia coli

<400> 3

gcggcgcgcc accatgcaag cggcaactgt tgtg 34

<210> 4

<211> 32

<212> ADN

<213> Escherichia coli

5 <400> 4

gcctcgagtt aatccacgta ttcatcgcg ac 32

<210> 5

<211> 1071

<212> ADN

10 <213> Escherichia coli

<400> 5

atgacccgtc cgatacaggc cagcctcgat ctgcaggcat taaaacagaa tctgtccatt	60
gtccgccagg ccgcgacgca cgcgcgcgtc tggtcggtgg taaaagcgaa cgcttacggg	120
catggtattg agcgtatctg gagcgcgacg ggggccaccg atggctttgc attgcttaac	180
ctggaagagg caataacggt acgtgagcgc ggctggaaag gaccgatcct gatgctggaa	240
ggatTTTTCC atgctcagga tctggagatt tatgaccagc accgcctgac cacctgcgta	300
cacagcaact ggcagctcaa agcactgcaa aatgcgcggc taaaagcacc gttggatatt	360
tatcttaaag tgaacagtgg gatgaatcgg ttgggcttcc agcccgatcg cgtgcttacc	420

ES 2 382 408 T3

```

gtctggcagc agttgcgggc aatggcgaat gttggcgaaa tgaccctgat gtcgcatttt 480
gccgaagcgg aacatcctga tggaaatttcc ggcgcgatgg cgcgtattga gcaggcggcg 540
gaggggctgg agtgtcggcg ttcgttgtcc aattcggcgg cgactctgtg gcacccggaa 600
gcgcattttg actgggttcg gcctggcatt attttgtatg gcgcttcgcc gtccgggtcag 660
tggcgtgata tcgccaatac cggattacgt ccggtgatga cgctaagcag tgagattatt 720
ggtgtccaga cgctaaaagc gggcgagcgt gtgggctacg gcggtcgcta tactgcgcgc 780
gatgaacagc gaatcggcat tgtcgccgca gggtagcccg acggttatcc gcgccacgcg 840
cctaccggta cccctgtttt agtggacggc gtgcgcacca tgacggtggg gaccgtctcg 900
atggatatgc tagcggtcga tttaacgcct tgcccgcagg cgggtattgg tacgccgggtt 960
gagctgtggg gcaaggagat caaaattgat gatgtcgccg ccgctgccgg aacggtgggc 1020
tatgagttga tgtgcgcgct ggcgctacgc gtcccgggtg tgacggtgta a 1071

```

<210> 6

<211> 356

5 <212> PRT

<213> Escherichia coli

<400> 6

ES 2 382 408 T3

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

Ala Glu Ala Glu His Pro Asp Gly Ile Ser Gly Ala Met Ala Arg Ile
165 170 175

Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
180 185 190

Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
195 200 205

Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
210 215 220

Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
225 230 235 240

Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
245 250 255

Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
260 265 270

Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
275 280 285

Asp Gly Val Arg Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
290 295 300

Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
305 310 315 320

Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
325 330 335

Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
340 345 350

Val Val Thr Val
355

<210> 7

<211> 1080

<212> ADN

5 <213> Escherichia coli

<400> 7

```

atgcaagcgg caactgttgt gattaaccgc cgcgctctgc gacacaacct gcaacgtctt    60
cgtgaactgg cccctgccag taaaatgggt gcggtggtga aagcgaacgc ttatgggtcac    120
ggtcttcttg agaccgcgcg aacgctcccc gatgctgacg cctttggcgt agcccgcttc    180
gaagaagctc tgcgactgcg tgcgggggga atcaccaaac ctgtactgtt actcgaaggc    240
ttttttgatg ccagagatct gccgacgatt tctgcgcaac attttcatac cgccgtgcat    300

aacgaagaac agctggctgc gctggaagag gctagcctgg acgagccggt taccgtcttg    360
atgaaactcg ataccggtat gcaccgtctg ggcgtaaggc cggaacaggc tgaggcgttt    420
tatcatcgcc tgaccagtg caaaaacgtt cgtcagccgg tgaatatcgt cagccatttt    480
gcgcgcgcgg atgaaccaa atgtggcgca accgagaaac aactcgctat ctttaatacc    540
ttttgcgaag gcaaacttg tcaacgttcc attgccgcgt cgggtggcat tctgctgtgg    600
ccacagtcgc attttgactg ggtgcgccc ggcattcattc tttatggcgt ctgcgcgctg    660
gaagatcgct ccaccggtgc cgattttggc tgtcagccag tgatgtcact aacctccagc    720
ctgattgccg tgcgtgagca taaagccgga gagcctgttg gttatggtgg aacctgggta    780
agcgaacgtg ataccgtct tggcgtagtc gcgatgggct atggcgatgg ttatccgcgc    840
gccgcgccgt ccggtacgcc agtgctgggt aacggtcgcg aagtaccgat tgtcgggcgc    900
gtggcgatgg atatgatctg cgtagactta ggtccacagg cgcaggacaa agccggggat    960
ccggtcattt tatggggcga aggtttgccc gtagaacgta tcgctgaaat gacgaaagta   1020
agcgcttacg aacttattac gcgcctgact tcaagggtcg cgatgaaata cgtggattaa   1080

```

- 5 <210> 8
- <211> 359
- <212> PRT
- <213> Escherichia coli
- <400> 8

Met Gln Ala Ala Thr Val Val Ile Asn Arg Arg Ala Leu Arg His Asn
1 5 10 15

Leu Gln Arg Leu Arg Glu Leu Ala Pro Ala Ser Lys Met Val Ala Val
20 25 30

Val Lys Ala Asn Ala Tyr Gly His Gly Leu Leu Glu Thr Ala Arg Thr
35 40 45

Leu Pro Asp Ala Asp Ala Phe Gly Val Ala Arg Leu Glu Glu Ala Leu
50 55 60

Arg Leu Arg Ala Gly Gly Ile Thr Lys Pro Val Leu Leu Leu Glu Gly
65 70 75 80

Phe Phe Asp Ala Arg Asp Leu Pro Thr Ile Ser Ala Gln His Phe His
85 90 95

Thr Ala Val His Asn Glu Glu Gln Leu Ala Ala Leu Glu Glu Ala Ser
100 105 110

Leu Asp Glu Pro Val Thr Val Trp Met Lys Leu Asp Thr Gly Met His
115 120 125

Arg Leu Gly Val Arg Pro Glu Gln Ala Glu Ala Phe Tyr His Arg Leu
130 135 140

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

<210> 9

<211> 609

5 <212> ADN

ES 2 382 408 T3

<213> Glycine max

<400> 9

gaagccacgt catgaagagt atatcatttc agtaatgttt tgagacgcct ctataatgct	60
ttaccaacaa aacaaaacaa aaaaaagaac atttgaaacc atttgattta aaaaaaaaaa	120
ggtatattag gccataatat tataggtaac atgaaatatc aaatgacacg caagagtttt	180

5

gtcaaaaatg aaaccatcac acatcagaga ttatggcaaa taatgttttg tgtgtctctt	240
gcttcaccca taacataagc ctctataact ggagagaaga aaaaaaaaaag tggaggggct	300
agggtgggaa tttggaagaa tacagttata ttgagcattg agcaagttga tagaaagctt	360
ctcaatttgt acaaaatttg catccacatg attattaaag acgtagacag cacttcttcc	420
ttcttttttt ctataagttt cttatatatt gttcttcatg ttttaattatt attactttat	480
gtacgcgtct aacagtagtc ctcccaaact gctataaata gagcctcttc aacgcacctc	540
ttggcagtac aaaaattatt catctcttct aagttctaatt tttctaagca ttcagtaaaa	600
gaactaacc	609

<210> 10

<211> 2085

10 <212> ADN

<213> Arabidopsis thaliana

<400> 10

cgaagagcat aagttttgtt caaatggccc aataacaaat taaaaacatg taaagtagtc	60
agtttaaaca agcatttgca taaagtgtgg ttaatatatt attaaacttc acatccaatg	120
agcattcatg taattttaaag taactgaagt taagtatcta gaagcctttt tcttctattg	180
gttattaatt tgcttaattt tctttataag ttaatttctg gttggtgtga aaatgtgacc	240
ggagaaggta tctaactttt ttttttcttt aatgaattcc actaaaattt aattctgtat	300
gtaacgcata tagtaaaatc tagaaagcga ccggcgtgcc tcctttggaa agtaatcctg	360
taaaagtaaa agccgcgtag tgtaaaagta tatgacttct tcttcccata attattttat	420
aattagtctt taatctaaat atttaaacat ataattcgtt ttacgagaaa gatcttcaca	480
ctcgattagt atacattaca tttaattccc tagttcataa aatggataac aaaaggctgt	540
gcgagattac aactgtactt gataattttg tataaatata tcctttatga atatatatta	600
gcattgatga ccgtacatgg ttaatccagt ctgcagcata acggagtatg atattaaatg	660
aacactttct gttcgtatca aatggtatcg aatattatta gagtgatcat tcagaagaaa	720
aaaagagaga gaagaaaacc tacagtgtaa acattttttt ttttgctaaa tacctacagt	780
gtaaacatga agtgctataa tttctgcaaa tagaaatcaa gaacagaaag agttgcttgg	840
agggaaagaa atagaaaatt aagaaatcta gtgatgtaat aaatctttcc ataaaaatcaa	900
atgtttggtc caaagtatta gttaaataat tagggcacta ttcttgacaa ctctttttaa	960
caaactcttc tatattttct cgtggtacat atgctgaaaa agatgtatgt ctaatccata	1020
atatatctgt ataatgcgac ttctattatc tattagtacg acttctaacc tagaagataa	1080
caagcattag ctagggcacg aaaatcaacg tggaaaaacc tacgaaaagc acgaagtgat	1140
taatctgtgt aggggtggcg taagggtaaa gactaaagac tgagaatcta ggggtcaagg	1200
cgtaaaacttg ttctgctttt tgggtttcat ttatttggcg aacaacattg atgtgtgtgg	1260
accatttggt gttcagggat tgagacaaga taatatgttt gctctcacct tctaggatta	1320
ctcgggtgct aagactcact tagtactatt gctatatcga tatactagtt cattaccaa	1380
aaatggagtc ttcaaatttc gagttccaat atctgaaagc attgtttaaa gagatttggt	1440
ttctccctgc acaattagtt tataacttca tatatacaca atcttatcaa tttaacaaca	1500
ggtgtgtgtg aaccttcaca taatctctct tattcattca tgtatatatc caataaaagt	1560
tcgatatgtg aaattatata tctccatcta atggttagact attcccgggt cttgactata	1620
aatttaaagt attagacgag ctaattatat ttagcacaaa caatttcttc tgtaacagtg	1680
tcacgcttat cactaccaa gaataaacac tgatctgttt taatctctta tttctcacc	1740
catattcaaa gtcaactatt gcaagacttc gagataatta atttgatggc tatactattt	1800
acttgacatt tgggaaaata tattttcgct gataaatttg gtttttactt ctctctccga	1860
cggatataga aacaattcaa ttacatgcga aaatgataat tcaaccctat aaaccaaacc	1920
aaataacaga atgcacattt ttttcaacgc gttagggtcac ctatctttca ctttagaaca	1980
tcccttcacg tctctatata aacctcgact ctgttatcct ttgttcttca agtacaacaa	2040
tcaactctaa gtctattata ttcaagtctt tgttttaacc taaca	2085

<210> 11

<211> 1999

<212> ADN

<213> Arabidopsis thaliana

5 <400> 11

```

gtagtgcctt tcattgatac caaaagagaa aatttgattt agtgcataca tataacaata    60
taacgccgca taataatact gtataaaaca gtcattgtaac gatattgacag cagtaataca    120
gttccaagag acgtttataat cgtattgcaat catattgcttg cgtattgattt ccaacagttt    180
tgtttcgctg ataggaggaa ctcaaacactc tagggtagtg attggtagac actatttagca    240
caaaaaatat taatttttact ctgattgttta ccaaaaaagt taccaatcaa atattttaaga    300
gatcgtactc ttccacggcg actctaaaaa ccaaagatat aggttagact cataactact    360
ttataaagaa aatgttttaac gataactacc gagattctaataaataaacct tcattttcaa    420
gtatattata tttgcttctt ttgtttatat atcaaaccac gttctggttt ataaaaatat    480
tagataaaac tcgtctaaat aggttagtgtt aaaataaaat tttaaatttt tatcgataat    540
atttaaaatt tgaaaagtta ataattgatcc acacattttt tctaattttt aatttagtaa    600
tttttgattt aaataaaatt tcaatcatat acattcgatt tttctataca ttttaactat    660
ctattttctgc ataataaact gtattttcat tttatacgct tcattcttatg gatgatattt    720
aaattttaaa tagtaattca tacacttttt aatatttaatt ttagtatttt cttaaattcca    780
aatttttaac ttacaattta aatatctact ttaacataat acaaatatac ttttaatttca    840
ttgtattaaa ttcaaataa atttgattat aataaaatac aatttaattc taaaaagtcc    900
atcttagatt ttaattttcc tttttagttt tgaaaattaa aaatttaaat ttattagata    960
tatattgtac tttttcagtt ttctatttta ttttaagaaa aaatattttt taacacatgt   1020
caacttgtaa acaatagact gaacacgtca ttttatatta tgtttagttt tgaaaattaa   1080
agttaattaa atattttatat ttcttttttt tagcttttct aattattttt aaaatagtaa   1140

```


ES 2 382 408 T3

```

atatttttaa tacaaatcaa tatctgaaca atagatttga tacataacat aatcctataa 1200
attattaact tggaaaacga tagtttatat aataaaatta ttttcttaag ttctctaacc 1260
ataacaatta aactatattt tagcgaagaa aagaagagaa taccgagaga acgcaacttg 1320
cactaaaagc taccactttg gcaaatcact catttatatt attatatact atcacctcaa 1380
ttcaatcgaa acctcaaaat aacactaata tatacacaaa gaaacaacag aataacaccg 1440
aagaatatag gtttaggaaa atccagaatt tgttgagact aaagagatca aattttcgat 1500
acaagggttt gctcaatttg tattttcata ataaaattct ttatttcacc atagacttac 1560
atgattagtt tttcttttaa taaaaaaaaa cgcgcgacat gaaaattata ttatctcagt 1620
gttgctgaat ttgaatttga attttgagtt aaatactaca catttgttga caacttatta 1680
aactttacaa gtctgctaca aatattgtca aatatttact aattaatgga ccaaaatcct 1740
ctaacttgca aatttgatc tacatcaact taaaaattag gaatatgcga ccaaaaaaaaa 1800
aaaaaactag gaataataat aaaaaaatgg aatgatgtgg aggaagctct ttactctttg 1860
agaggaagtt tataaattga ccacacattt agtctattat catcacatgt attaagactt 1920
gacaacttgt ctttctcaca ccaaaccctt ctcctctggt tcataacatc tgctctttct 1980
tttttttcct aagccccta 1999

```

<210> 12

<211> 356

5 <212> PRT

<213> HS de Escherichia coli

<400> 12

ES 2 382 408 T3

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

ES 2 382 408 T3

Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln
 130 135 140
 Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
 145 150 155 160
 Ala Glu Ala Glu His Pro Asp Gly Ile Ser Gly Ala Met Ala Arg Ile
 165 170 175
 Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
 180 185 190
 Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
 195 200 205
 Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
 210 215 220
 Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
 225 230 235 240
 Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
 245 250 255
 Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
 260 265 270
 Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
 275 280 285
 Asp Gly Val Arg Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
 290 295 300
 Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
 305 310 315 320
 Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
 325 330 335
 Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
 340 345 350
 Val Val Thr Val
 355

<210> 13

<211> 356

<212> PRT

<213> EDL933 0157:H7 de Escherichia coli

<400> 13

ES 2 382 408 T3

Met Thr Arg Pro Ile Gln Ala Ser Leu Asp Leu Gln Ala Leu Lys Gln
 1 5 10 15
 Asn Leu Ser Ile Val Arg Gln Ala Ala Pro His Ala Arg Val Trp Ser
 20 25 30
 Val Val Lys Ala Asn Ala Tyr Gly His Gly Ile Glu Arg Ile Trp Ser
 35 40 45
 Ala Leu Gly Ala Thr Asp Gly Phe Ala Leu Leu Asn Leu Glu Glu Ala
 50 55 60
 Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu
 65 70 75 80
 Gly Phe Phe His Ala Gln Asp Leu Glu Met Tyr Asp Gln His Arg Leu
 85 90 95
 Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
 100 105 110
 Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
 115 120 125
 Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln
 130 135 140
 Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
 145 150 155 160
 Ala Glu Ala Glu His Pro Asp Gly Ile Ser Gly Ala Met Ala Arg Ile
 165 170 175
 Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
 180 185 190
 Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
 195 200 205
 Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
 210 215 220
 Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
 225 230 235 240
 Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
 245 250 255
 Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
 260 265 270
 Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val

ES 2 382 408 T3

275 280 285
 Asp Gly Val Arg Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
 290 295 300
 Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
 305 310 315 320
 Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
 325 330 335
 Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
 340 345 350
 Val Val Thr Val
 355

<210> 14

<211> 356

5 <212> PRT

<213> UTI89 de Escherichia coli

<400> 14

Met Thr Arg Pro Ile Gln Ala Ser Leu Asp Leu Gln Ala Leu Lys Gln
1 5 10 15 .

Asn Leu Ser Ile Val Arg Gln Ala Ala Pro Arg Ala Arg Val Trp Ser
20 25 30

Val Val Lys Ala Asn Ala Tyr Gly His Gly Ile Glu Arg Ile Trp Ser
35 40 45

Ala Ile Gly Ala Thr Asp Gly Phe Ala Leu Leu Asn Leu Glu Glu Ala
50 55 60

Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu
65 70 75 80

Gly Phe Phe His Ala Gln Asp Leu Glu Ile Tyr Asp Gln His Arg Leu
85 90 95

Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
100 105 110

Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
115 120 125

Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln
130 135 140

Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
145 150 155 160

Ala Glu Ala Glu His Pro Asp Gly Ile Ser Gly Ala Met Ala Arg Ile
165 170 175

Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
180 185 190

Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
195 200 205

Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
210 215 220

Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
225 230 235 240

Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
245 250 255

Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
260 265 270

Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
275 280 285

Asp Gly Val Arg Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
290 295 300

Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
305 310 315 320

Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
325 330 335

Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
340 345 350

Val Val Thr Val
355

<210> 15

<211> 356

5 <212> PRT

<213> 53683 de Escherichia coli

<400> 15

ES 2 382 408 T3

Met Thr Arg Pro Ile Gln Ala Ser Leu Asp Leu Gln Ala Leu Lys Gln
 1 5 10 15

Asn Leu Ser Ile Val Arg Gln Ala Ala Thr His Ala Arg Val Trp Ser
 20 25 30

Val Val Lys Ala Asn Ala Tyr Gly His Gly Ile Glu Arg Ile Trp Ser

ES 2 382 408 T3

35	40	45	
Ala Leu Gly Ala Thr Asp Gly Phe Ala Leu Leu Asn Leu Glu Glu Ala	50	55	60
Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu	65	70	75
Gly Phe Phe His Ala Gln Asp Leu Glu Ile Tyr Asp Gln His Arg Leu	85	90	95
Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala	100	105	110
Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met	115	120	125
Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln	130	135	140
Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe	145	150	155
Ala Glu Ala Glu His Pro Asp Gly Ile Ser Gly Ala Met Ala Arg Ile	165	170	175
Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser	180	185	190
Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro	195	200	205
Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile	210	215	220
Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile	225	230	235
Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg	245	250	255
Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr	260	265	270
Ala Asp Gly Tyr Pro Pro His Ala Pro Thr Gly Thr Pro Val Leu Val	275	280	285
Asp Gly Val Asn Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu	290	295	300
Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val	305	310	315
		320	

ES 2 382 408 T3

Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
 325 330 335

Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
 340 345 350

Val Val Thr Val
 355

<210> 16

<211> 356

5 <212> PRT

<213> E24377A de Escherichia coli

<400> 16

ES 2 382 408 T3

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

Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
195 200 205

Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
210 215 220

Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
225 230 235 240

Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
245 250 255

Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
260 265 270

Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
275 280 285

Asp Gly Val Arg Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
290 295 300

Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
305 310 315 320

Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
325 330 335

Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
340 345 350

Val Val Thr Val
355

<210> 17

<211> 356

5 <212> PRT

<213> Cepa 8401 de Shigella flexneri 5

<400> 17

ES 2 382 408 T3

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

Gly Phe Phe His Ala Gln Asp Leu Glu Ile Tyr Asp Gln His Arg Leu
85 90 95

Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
100 105 110

Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
115 120 125

Asn Arg Leu Gly Phe Gln Ser Asp Arg Val Leu Thr Val Trp Gln Gln
130 135 140

Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
145 150 155 160

Ala Glu Ala Glu His Pro Asp Gly Ile Ser Gly Ala Met Ala Arg Ile
165 170 175

Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
180 185 190

Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
195 200 205

Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
210 215 220

Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
225 230 235 240

Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
245 250 255

Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
260 265 270

Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
275 280 285

Asp Gly Val Arg Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
290 295 300

Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
305 310 315 320

Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
325 330 335

Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
340 345 350

ES 2 382 408 T3

Val Val Thr Val
355

<210> 18

<211> 356

<212> PRT

5 <213> Ss046 de Shigella sonnei

<400> 18

Met Thr Arg Pro Ile Gln Ala Ser Leu Asp Leu Gln Ala Leu Lys Gln
1 5 10 15

Asn Leu Ser Ile Val Arg Gln Ala Ala Pro His Ala Arg Val Trp Ser
20 25 30

Val Val Lys Ala Asn Ala Tyr Gly His Gly Ile Glu Arg Ile Trp Ser
35 40 45

Ala Leu Gly Ala Thr Asp Gly Phe Ala Leu Leu Asn Leu Glu Glu Ala
50 55 60

Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu
65 70 75 80

Gly Phe Phe His Ala Gln Asp Leu Glu Ile Tyr Asp Gln His Arg Leu
85 90 95

Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
100 105 110

Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
115 120 125

Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln
130 135 140

Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
145 150 155 160

Ala Glu Ala Glu His Pro Asp Gly Ile Ser Ser Ala Met Ala Arg Ile
165 170 175

Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
180 185 190

Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
195 200 205

Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
210 215 220

Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile

ES 2 382 408 T3

[illegible]

<210> 19

<211> 356

5 <212> PRT

<213> B7A de Escherichia coli

<400> 19

ES 2 382 408 T3

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

ES 2 382 408 T3

Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
 115 120 125
 Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln
 130 135 140
 Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
 145 150 155 160
 Ala Glu Ala Glu His Pro Asp Gly Ile Ser Ser Ala Met Ala Arg Ile
 165 170 175
 Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
 180 185 190
 Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
 195 200 205
 Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
 210 215 220
 Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
 225 230 235 240
 Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
 245 250 255
 Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
 260 265 270
 Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
 275 280 285
 Asp Gly Val Arg Thr Met Thr Val Gly Ala Val Ser Met Asp Met Leu
 290 295 300
 Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
 305 310 315 320
 Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
 325 330 335
 Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
 340 345 350
 Val Val Thr Val
 355

<210> 20

<211> 356

<212> PRT

<213> Escherichia coli

<400> 20

Met Thr Arg Pro Ile Gln Ala Ser Leu Asp Leu Gln Ala Leu Lys Gln
1 5 10 15

Asn Leu Ser Ile Val Arg Gln Ala Ala Thr His Ala Arg Val Trp Ser
20 25 30

Val Val Lys Ala Asn Ala Tyr Gly His Gly Ile Glu Arg Ile Trp Ser
35 40 45

Ala Ile Gly Ala Thr Asp Gly Phe Ala Leu Leu Asn Leu Glu Glu Ala
50 55 60

Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu
65 70 75 80

Gly Phe Phe His Ala Gln Asp Leu Glu Ile Tyr Asp Gln His Arg Leu
85 90 95

Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
100 105 110

Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
115 120 125

Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln
130 135 140

Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
145 150 155 160

Ala Glu Ala Glu His Pro Asp Gly Ile Ser Gly Arg Met Ala Arg Ile
165 170 175

Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
180 185 190

Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
195 200 205

Gly Ile Ile Leu Tyr Gly Arg Ser Pro Ser Gly Gln Trp Arg Asp Ile
210 215 220

Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
225 230 235 240

Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
245 250 255

Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
260 265 270

Ala Asp Gly Tyr Pro Arg His Ala Leu Thr Gly Thr Pro Val Leu Val
275 280 285

Asp Gly Val Arg Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
290 295 300

Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
305 310 315 320

Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
325 330 335

Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
340 345 350

Val Val Thr Val
355

<210> 21

<211> 356

5 <212> PRT

<213> Cepa 301 de Shigella flexneri 2a

<400> 21

ES 2 382 408 T3

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

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

<210> 22

<211> 356

ES 2 382 408 T3

<212> PRT

<213> F11 de Escherichia coli

<400> 22

Met Thr Arg Pro Ile Gln Ala Ser Leu Asp Leu Gln Ala Leu Lys Gln
1 5 10 15

Asn Leu Ser Ile Val Arg Gln Ala Ala Pro Tyr Ala Arg Val Trp Ser
20 25 30

5

ES 2 382 408 T3

Val Val Lys Ala Asn Ala Tyr Gly His Gly Ile Glu Arg Ile Trp Ser
 35 40 45
 Ala Leu Gly Ala Thr Asp Gly Phe Ala Leu Leu Asn Leu Glu Glu Ala
 50 55 60
 Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu
 65 70 75 80
 Gly Phe Phe His Ala Gln Asp Leu Glu Ile Tyr Asp Gln His Arg Leu
 85 90 95
 Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
 100 105 110
 Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
 115 120 125
 Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln
 130 135 140
 Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
 145 150 155 160
 Ala Glu Ala Glu His Pro Asp Gly Ile Ser Gly Ala Met Ala Arg Ile
 165 170 175
 Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
 180 185 190
 Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
 195 200 205
 Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
 210 215 220
 Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
 225 230 235 240
 Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
 245 250 255
 Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
 260 265 270
 Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
 275 280 285
 Asp Gly Val Leu Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
 290 295 300

ES 2 382 408 T3

Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
305 310 315 320

Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
325 330 335

Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
340 345 350

Val Val Thr Val
355

<210> 23

<211> 356

5 <212> PRT

<213> Escherichia coli

<400> 23

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

ES 2 382 408 T3

[illegible]

<210> 24

<211> 356

5 <212> PRT

<213> CFT073 de Escherichia coli

<400> 24

ES 2 382 408 T3

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

Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu
 65 70 75 80
 Gly Phe Phe His Ala Gln Asp Leu Glu Ile Tyr Asp Gln His Arg Leu
 85 90 95
 Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
 100 105 110
 Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
 115 120 125
 Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln
 130 135 140
 Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
 145 150 155 160
 Ala Glu Ala Glu His Pro Asp Gly Ile Ser Ser Ala Met Ala Arg Ile
 165 170 175
 Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ala Asn Ser
 180 185 190
 Ala Ala Thr Leu Trp His Gln Glu Ala His Phe Asp Trp Val Arg Pro
 195 200 205
 Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
 210 215 220
 Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
 225 230 235 240
 Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
 245 250 255
 Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
 260 265 270
 Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
 275 280 285
 Asp Gly Val Leu Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
 290 295 300
 Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
 305 310 315 320
 Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
 325 330 335

ES 2 382 408 T3

Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
340 345 350

Val Val Thr Val
355

<210> 25

<211> 324

5 <212> PRT

<213> BS512 de Shigella boydii

<400> 25

Met Val Lys Ala Asn Ala Tyr Gly His Gly Ile Glu Arg Ile Trp Ser
1 5 10 15

Ala Leu Gly Ala Thr Asp Gly Phe Ala Leu Leu Asn Leu Glu Glu Ala
20 25 30

Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu
35 40 45

Gly Phe Phe His Ala Gln Asp Leu Glu Ile Tyr Asp Gln His Arg Leu
50 55 60

Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
65 70 75 80

Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
85 90 95

Asn Arg Leu Gly Phe Gln Pro Asp Arg Val Leu Thr Val Trp Gln Gln
100 105 110

Leu Arg Ala Met Ala Asn Val Gly Glu Met Thr Leu Met Ser His Phe
115 120 125

Ala Glu Ala Glu His Pro Asp Gly Ile Ser Ser Ala Met Ala Arg Ile
130 135 140

Glu Gln Ala Ala Glu Gly Leu Glu Cys Arg Arg Ser Leu Ser Asn Ser
145 150 155 160

Ala Ala Thr Leu Trp His Pro Glu Ala His Phe Asp Trp Val Arg Pro
165 170 175

Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
180 185 190

Ala Asn Thr Gly Leu Arg Pro Val Met Thr Leu Ser Ser Glu Ile Ile
195 200 205

Gly Val Gln Thr Leu Lys Ala Gly Glu Arg Val Gly Tyr Gly Gly Arg
210 215 220

ES 2 382 408 T3

Tyr Thr Ala Arg Asp Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
225 230 235 240

Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
245 250 255

Asp Gly Val Arg Thr Met Thr Val Gly Thr Val Ser Met Asp Met Leu
260 265 270

Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
275 280 285

Glu Leu Trp Gly Lys Glu Ile Lys Ile Asp Asp Val Ala Ala Ala Ala
290 295 300

Gly Thr Val Gly Tyr Glu Leu Met Cys Ala Leu Ala Leu Arg Val Pro
305 310 315 320

Val Val Thr Val

<210> 26

<211> 329

5 <212> PRT

<213> Cepa Sakai 0157:H7 de Escherichia coli

<400> 26

ES 2 382 408 T3

Met Thr Arg Pro Ile Gln Ala Ser Leu Asp Leu Gln Ala Leu Lys Gln
1 5 10 15

Asn Leu Ser Ile Val Arg Gln Ala Ala Pro His Ala Arg Val Trp Ser
20 25 30

Val Val Lys Ala Asn Ala Tyr Gly His Gly Ile Glu Arg Ile Trp Ser
35 40 45

Ala Leu Gly Ala Thr Asp Gly Phe Ala Leu Leu Asn Leu Glu Glu Ala
50 55 60

Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu
65 70 75 80

Gly Phe Phe His Ala Gln Asp Leu Glu Met Tyr Asp Gln His Arg Leu
85 90 95

Thr Thr Cys Val His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
100 105 110

Arg Leu Lys Ala Pro Leu Asp Ile Tyr Leu Lys Val Asn Ser Gly Met
115 120 125

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

<210> 27

<211> 356

5 <212> PRT

<213> Salmonella entérica subespecie entérica serovar

<400> 27

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

ES 2 382 408 T3

Ala Leu Gly Ala Thr Asp Gly Phe Ala Met Leu Asn Leu Glu Glu Ala
 50 55 60
 Ile Thr Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Met Leu Glu
 65 70 75 80
 Gly Phe Phe His Ala Gln Asp Leu Glu Ala Tyr Asp Thr Tyr Arg Leu
 85 90 95
 Thr Thr Cys Ile His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
 100 105 110
 Arg Leu Asn Ala Pro Leu Asp Ile Tyr Val Lys Val Asn Ser Gly Met
 115 120 125
 Asn Arg Leu Gly Phe Gln Pro Glu Arg Ala Gln Thr Val Trp Gln Gln
 130 135 140
 Leu Arg Ala Met Arg Asn Val Gly Glu Met Thr Leu Met Ser His Phe
 145 150 155 160
 Ala Gln Ala Asp His Pro Glu Gly Ile Gly Glu Ala Met Arg Arg Ile
 165 170 175
 Ala Leu Ala Thr Glu Gly Leu Gln Cys Ala Tyr Ser Leu Ser Asn Ser
 180 185 190
 Ala Ala Thr Leu Trp His Pro Gln Ala His Tyr Asp Trp Val Arg Pro
 195 200 205
 Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
 210 215 220
 Ala Asp Thr Gly Leu Lys Pro Val Met Thr Leu Ser Ser Glu Ile Ile
 225 230 235 240
 Gly Val Gln Thr Leu Ser Val Gly Glu Arg Val Gly Tyr Gly Gly Arg
 245 250 255
 Tyr Ser Val Thr Gln Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
 260 265 270
 Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
 275 280 285
 Asp Gly Ile Arg Thr Arg Thr Val Gly Thr Val Ser Met Asp Met Leu
 290 295 300
 Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
 305 310 315 320

ES 2 382 408 T3

Glu Leu Trp Gly Lys Glu Ile Lys Val Asp Asp Val Ala Ser Ala Ala
 325 330 335

Gly Thr Leu Gly Tyr Glu Leu Leu Cys Ala Val Ala Pro Arg Val Pro
 340 345 350

Phe Val Thr Thr
 355

<210> 28

<211> 356

5 <212> PRT

<213> LT2 de Salmonella typhimurium

<400> 28

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

195	200	205
Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile 210 215 220		
Ala Asp Thr Gly Leu Lys Pro Val Met Thr Leu Ser Ser Glu Ile Ile 225 230 235 240		
Gly Val Gln Thr Leu Ser Ala Gly Glu Arg Val Gly Tyr Gly Gly Gly 245 250 255		
Tyr Ser Val Thr Gln Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr 260 265 270		
Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val 275 280 285		
Asp Gly Ile Arg Thr Arg Thr Val Gly Thr Val Ser Met Asp Met Leu 290 295 300		
Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val 305 310 315 320		
Glu Leu Trp Gly Lys Glu Ile Lys Val Asp Asp Val Ala Ser Ala Ala 325 330 335		
Gly Thr Leu Gly Tyr Glu Leu Leu Cys Ala Val Ala Pro Arg Val Pro 340 345 350		
Phe Val Thr Thr 355		

<210> 29

<211> 356

5 <212> PRT

<213> Salmonella entérica subespecie entérica serovar Typhi

<400> 29

ES 2 382 408 T3

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

Gly Phe Phe His Ala Gln Asp Leu Glu Ala Tyr Asp Thr Tyr Arg Leu
85 90 95

Thr Thr Cys Ile His Ser Asn Trp Gln Leu Lys Ala Leu Gln Asn Ala
100 105 110

Arg Leu Asn Ala Pro Leu Asp Ile Tyr Val Lys Val Asn Ser Gly Met
115 120 125

Asn Arg Leu Gly Phe Gln Pro Glu Arg Ala Gln Thr Val Trp Gln Gln
130 135 140

Leu Arg Ala Met Arg Asn Val Gly Glu Met Thr Leu Met Ser His Phe
145 150 155 160

Ala Gln Ala Asp His Pro Glu Gly Ile Gly Glu Ala Met Arg Arg Ile
165 170 175

Ala Leu Ala Thr Glu Gly Leu Gln Cys Ala Tyr Ser Leu Ser Asn Ser
180 185 190

Ala Ala Thr Leu Trp His Pro Gln Ala His Tyr Asp Trp Val Arg Pro
195 200 205

Gly Ile Ile Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Arg Asp Ile
210 215 220

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

Gly Val Gln Thr Leu Ser Ala Gly Glu Arg Val Gly Tyr Gly Gly Gly
245 250 255

Tyr Ser Val Thr Gln Glu Gln Arg Ile Gly Ile Val Ala Ala Gly Tyr
260 265 270

Ala Asp Gly Tyr Pro Arg His Ala Pro Thr Gly Thr Pro Val Leu Val
275 280 285

Asp Gly Ile Arg Thr Arg Thr Val Gly Thr Val Ser Met Asp Met Leu
290 295 300

Ala Val Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Thr Pro Val
305 310 315 320

Glu Leu Trp Gly Lys Glu Ile Lys Val Asp Asp Val Ala Ser Ala Ala
325 330 335

Gly Thr Leu Gly Tyr Gly Leu Leu Cys Ala Val Ala Pro Arg Val Pro
340 345 350

ES 2 382 408 T3

Phe Val Thr Thr
355

<210> 30

<211> 356

<212> PRT

5 <213> Enterobacter sp. 638

<400> 30

ES 2 382 408 T3

Met Ser Arg Pro Ile Leu Ala Gln Leu Asp Leu Gln Ala Leu Lys Asn
1 5 10 15

Asn Leu Gln Ile Val Arg Arg Ala Ala Pro Gly Ser Arg Val Trp Ser
20 25 30

Val Val Lys Ala Asn Gly Tyr Gly His Gly Ile Glu Ser Ile Trp Ser
35 40 45

Ala Leu Ser Ala Thr Asp Gly Phe Ala Leu Leu Asn Leu Glu Glu Ala
50 55 60

Ile Leu Leu Arg Glu Arg Gly Trp Lys Gly Pro Ile Leu Leu Leu Glu
65 70 75 80

Gly Phe Phe His Ala Asp Glu Leu Pro Leu Leu Asp Lys Tyr Arg Leu
85 90 95

Thr Thr Ser Val His Ser Asn Trp Gln Ile Lys Ala Leu Gln Asn Ala
100 105 110

Lys Leu His Ser Pro Leu Asp Ile Tyr Leu Lys Met Asn Ser Gly Met
115 120 125

Asn Arg Leu Gly Phe Gln Pro Glu Arg Val Asn Thr Val Trp Gln Gln
130 135 140

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

Ala Asp Ala Glu Lys Pro Asp Gly Ile Thr Asn Ala Met Ala Arg Ile
165 170 175

Glu Gln Ala Ala Glu Gly Leu Asp Cys Ala Arg Ser Leu Ser Asn Ser
180 185 190

Ala Ala Thr Leu Trp His Pro Glu Ala His Tyr Gln Trp Val Arg Pro
195 200 205

Gly Ile Val Leu Tyr Gly Ala Ser Pro Ser Gly Gln Trp Gln Asp Ile
210 215 220

Ala Asn Ser Gly Leu Lys Pro Val Met Thr Leu Arg Ser Glu Ile Ile
225 230 235 240

Gly Ile Gln Thr Leu Lys Ala Gly Asp Thr Val Gly Tyr Gly Ser Arg
245 250 255

Tyr Arg Ser Thr Gly Glu Gln Arg Ile Gly Ile Val Ala Gly Gly Tyr
260 265 270

Ala Asp Gly Tyr Pro Arg Leu Ala Pro Ser Gly Thr Pro Val Trp Ile
275 280 285

Asp Gly Val Arg Thr Gly Thr Val Gly Thr Ile Ser Met Asp Met Met
290 295 300

Ala Ile Asp Leu Thr Pro Cys Pro Gln Ala Gly Ile Gly Ser Pro Val
305 310 315 320

Glu Leu Trp Gly Asn Glu Val Lys Val Asp Asp Val Ala Ser Ala Ala
325 330 335

Gly Thr Val Gly Tyr Glu Leu Leu Thr Ala Leu Ala Pro Arg Val Pro
340 345 350

Val Val Thr Val
355

<210> 31

<211> 359

5 <212> PRT

<213> Artificial

<220>

<223> gen sintético

<400> 31

ES 2 382 408 T3

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

ES 2 382 408 T3

[illegible]

<210> 32

<211> 359

<212> PRT

<213> EDL933 0157:H7 de Escherichia coli

<400> 32

Met Gln Ala Ala Thr Val Val Ile Asn Arg Arg Ala Leu Arg His Asn
 1 5 10 15
 Leu Gln Arg Leu Arg Glu Leu Ala Pro Ala Ser Lys Met Val Ala Val
 20 25 30
 Val Lys Ala Asn Ala Tyr Gly His Gly Leu Leu Glu Thr Ala Arg Thr
 35 40 45
 Leu Pro Asp Ala Asp Ala Phe Gly Val Ala Arg Leu Glu Glu Ala Leu
 50 55 60
 Arg Leu Arg Ala Gly Gly Ile Thr Lys Pro Val Leu Leu Leu Glu Gly
 65 70 75 80
 Phe Phe Asp Ala Arg Asp Leu Pro Thr Ile Ser Ala Gln His Phe His
 85 90 95
 Thr Ala Val His Asn Glu Glu Gln Leu Ala Ala Leu Glu Glu Ala Ser
 100 105 110
 Leu Asp Glu Pro Val Thr Val Trp Met Lys Leu Asp Thr Gly Met His
 115 120 125
 Arg Leu Gly Val Arg Pro Glu Gln Ala Glu Ala Phe Tyr His Arg Leu
 130 135 140
 Thr Gln Cys Lys Asn Val Arg Gln Pro Val Asn Ile Val Ser His Phe
 145 150 155 160
 Ala Arg Ala Asp Glu Pro Lys Cys Gly Ala Thr Glu Lys Gln Leu Ala
 165 170 175
 Ile Phe Asn Thr Phe Cys Glu Gly Lys Pro Gly Gln Arg Ser Ile Ala
 180 185 190
 Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Val
 195 200 205
 Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asp Arg Ser
 210 215 220
 Thr Gly Ala Asp Phe Gly Cys Gln Pro Val Met Ser Leu Thr Ser Ser
 225 230 235 240
 Leu Ile Ala Val Arg Glu His Lys Val Gly Glu Pro Val Gly Tyr Gly
 245 250 255

Gly Thr Trp Ile Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
260 265 270

Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
275 280 285

Leu Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp
290 295 300

Met Ile Cys Val Asp Leu Gly Pro Gln Ala Gln Asp Lys Ala Gly Asp
305 310 315 320

Pro Val Ile Leu Trp Gly Glu Gly Leu Pro Val Glu Arg Ile Ala Glu
325 330 335

Met Thr Lys Val Ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg
340 345 350

Val Ala Met Lys Tyr Val Asp
355

<210> 33

<211> 359

5 <212> PRT

<213> Ss046 de Shigella sonnei

<400> 33

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

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

<210> 34

<211> 359

ES 2 382 408 T3

<212> PRT

<213> E110019 de Escherichia coli

<400> 34

5	Met	Gln	Ala	Ala	Thr	Val	Val	Ile	Asn	Arg	Arg	Ala	Leu	Arg	His	Asn
	1				5					10					15	

ES 2 382 408 T3

Leu Gln Arg Leu Arg Glu Leu Ala Pro Ala Ser Lys Met Val Ala Val
 20 25 30
 Val Lys Ala Asn Ala Tyr Gly His Gly Leu Leu Glu Thr Ala Arg Thr
 35 40 45
 Leu Pro Asp Ala Asp Ala Phe Gly Val Ala Arg Leu Glu Glu Ala Leu
 50 55 60
 Arg Leu Arg Ala Gly Gly Ile Thr Lys Pro Val Leu Leu Leu Glu Gly
 65 70 75 80
 Phe Phe Asp Ala Arg Asp Leu Pro Met Ile Ser Ala Gln His Phe His
 85 90 95
 Thr Ala Val His Asn Glu Glu Gln Leu Ala Ala Leu Glu Glu Ala Ser
 100 105 110
 Leu Asp Glu Pro Val Thr Val Trp Met Lys Leu Asp Thr Gly Met His
 115 120 125
 Arg Leu Gly Val Arg Pro Glu Gln Ala Glu Ala Phe Tyr His Arg Leu
 130 135 140
 Thr Gln Cys Lys Asn Val Arg Gln Pro Val Asn Ile Val Ser His Phe
 145 150 155 160
 Ala Arg Ala Asp Glu Pro Lys Cys Gly Ala Thr Glu Lys Gln Leu Ala
 165 170 175
 Ile Phe Asn Thr Phe Cys Glu Gly Lys Pro Gly Gln Arg Ser Ile Ala
 180 185 190
 Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Val
 195 200 205
 Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asp Arg Ser
 210 215 220
 Thr Gly Ala Asp Phe Gly Cys Gln Pro Val Met Ser Leu Thr Ser Ser
 225 230 235 240
 Leu Ile Ala Val Arg Glu His Lys Ala Gly Glu Pro Val Gly Tyr Gly
 245 250 255
 Gly Thr Trp Val Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
 260 265 270
 Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
 275 280 285
 Leu Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp

ES 2 382 408 T3

290		295		300
Met	Ile	Cys	Val	Asp
305				Leu
				310
				Gly
				Pro
				Gln
				Ala
				Gln
				315
				Asp
				Lys
				Ala
				Gly
				Asp
				320
				Pro
				Val
				Ile
				Ala
				335
				Glu
				Arg
				Glu
				Leu
				Thr
				350
				Ser
				Arg
				Met
				355
				Val
				Ala
				Lys
				Tyr
				Val
				Asp

<210> 35

<211> 359

5 <212> PRT

<213> B171 de Escherichia coli

<400> 35

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

Ile Phe Asn Thr Phe Cys Glu Gly Lys Pro Gly Gln Arg Ser Ile Ala
180 185 190

Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Val
195 200 205

Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asp Arg Ser
210 215 220

Thr Gly Ala Asp Phe Gly Cys Gln Pro Val Met Ser Leu Thr Ser Ser
225 230 235 240

Leu Ile Ala Val Arg Glu His Lys Ala Gly Glu Pro Val Gly Tyr Gly
245 250 255

Gly Thr Trp Val Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
260 265 270

Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
275 280 285

Leu Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp
290 295 300

Met Ile Cys Val Asp Leu Gly Pro Gln Ala Gln Asp Lys Ala Gly Asp
305 310 315 320

Pro Val Ile Leu Trp Gly Glu Gly Leu Pro Val Glu Arg Ile Ala Glu
325 330 335

Met Thr Lys Val ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg
340 345 350

Val Ala Met Lys Tyr Val Asp
355

<210> 36

<211> 359

5 <212> PRT

<213> BS512 de shigella boydii

<400> 36

ES 2 382 408 T3

Met Gln Ala Ala Thr Val Val Ile Asn Arg Arg Ala Leu Arg His Asn
1 5 10 15

Leu Gln Arg Leu Arg Glu Leu Ala Pro Ala Ser Lys Met Val Ala Val
20 25 30

Val Lys Ala Asn Ala Tyr Gly His Gly Leu Leu Glu Thr Ala Arg Thr
35 40 45

Leu Pro Asp Ala Asp Ala Phe Gly Val Ala Arg Leu Glu Glu Ala Leu

ES 2 382 408 T3

50	55	60
Arg 65	Leu Arg Ala Gly 70	Gly Ile Thr Lys Pro Val 75
Leu Leu Leu Glu 80		
Phe Phe Asp Ala 85	Arg Asp Leu Pro Thr 90	Ile Ser Ala Gln His 95
Thr Ala Val His 100	Asn Glu Glu Gln 105	Leu Ala Ala Leu Glu 110
Leu Asp Glu 115	Pro Val Thr Val Trp 120	Met Lys Leu Asp Thr 125
Arg Leu Gly Val Arg Pro Glu 130	Gln Ala Gly Ala Phe 140	Tyr His Arg Leu
Thr Gln Cys Lys Asn Val 145	Arg Gln Pro Val Asn 155	Ile Val Ser His Phe 160
Ala Arg Ala Asp Glu 165	Pro Lys Cys Gly Ala 170	Thr Glu Lys Gln Leu Ala 175
Ile Phe Asn Thr 180	Phe Cys Glu Gly Lys 185	Pro Gly Gln Arg Ser 190
Ala Ser Gly 195	Gly Ile Leu Leu Trp 200	Pro Gln Ser His Phe 205
Arg Pro Gly 210	Ile Ile Leu Tyr 215	Gly Val Ser Pro Leu Glu Asp Arg Ser 220
Thr Gly Ala Asp Phe 225	Gly Cys Gln Pro Val Met 235	Ser Leu Thr Ser Ser 240
Leu Ile Ala Val Arg 245	Glu His Lys Ala Gly 250	Glu Pro Val Gly Tyr Gly 255
Gly Thr Trp Val 260	Ser Glu Arg Asp Thr 265	Arg Leu Gly Val Val Ala Met 270
Gly Tyr Gly 275	Asp Gly Tyr Pro Arg Ala Ala 280	Pro Ser Gly Thr Pro Val 285
Leu Val Asn Gly Arg Glu Val 290	Pro Ile Val Gly Arg Val Ala Met Asp 300	
Met Ile Cys Val Asp 305	Leu Gly Pro Gln Ala Gln 310	Asp Lys Ala Gly Asp 320
Pro Val Ile Leu Trp 325	Gly Glu Gly Leu Pro Val 330	Glu Arg Ile Ala Glu 335

ES 2 382 408 T3

Met Thr Lys Val Ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg
340 345 350

Val Ala Met Lys Tyr Val Asp
355

<210> 37

<211> 359

5 <212> PRT

<213> E24377A de Escherichia coli

<400> 37

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

Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asp Arg Ser
 210 215 220
 Thr Gly Ala Asp Phe Gly Cys Gln Pro Val Met Ser Leu Thr Ser Ser
 225 230 235 240
 Leu Ile Ala Val Arg Glu His Lys Ala Gly Glu Pro Val Gly Tyr Gly
 245 250 255
 Gly Thr Trp Val Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
 260 265 270
 Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
 275 280 285
 Leu Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp
 290 295 300
 Met Ile Cys Val Asp Leu Gly Pro Gln Ala Gln Asp Lys Ala Gly Asp
 305 310 315 320
 Pro Val Ile Leu Trp Gly Glu Gly Leu Pro Val Glu Arg Ile Ala Glu
 325 330 335
 Met Thr Lys Val Ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg
 340 345 350
 Val Ala Met Lys Tyr Val Asp
 355

<210> 38

<211> 359

5 <212> PRT

<213> sd197 de shigella dysenteriae

<400> 38

ES 2 382 408 T3

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

ES 2 382 408 T3

Thr Ala Val His Asn Glu Glu Gln Leu Ala Ala Leu Glu Glu Ala Ser
 100 105 110
 Leu Asp Glu Pro Val Thr Val Trp Met Lys Leu Asp Thr Gly Met His
 115 120 125
 Arg Leu Gly Val Arg Pro Glu Gln Ala Glu Ala Phe Tyr His Arg Leu
 130 135 140
 Thr Gln Cys Lys Asn Val Arg Gln Pro Val Asn Ile Val Ser His Phe
 145 150 155 160
 Ala Arg Ala Asp Glu Pro Lys Cys Gly Ala Thr Glu Lys Gln Leu Ala
 165 170 175
 Ile Phe Asn Thr Phe Cys Glu Gly Lys Pro Gly Gln Arg Ser Ile Ala
 180 185 190
 Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Val
 195 200 205
 Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asp Arg Ser
 210 215 220
 Thr Gly Ala Asp Phe Gly Cys Gln Pro Val Met Ser Leu Thr Ser Ser
 225 230 235 240
 Leu Ile Ala Val Arg Glu His Lys Val Gly Glu Pro Val Gly Tyr Gly
 245 250 255
 Gly Thr Trp Ile Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
 260 265 270
 Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
 275 280 285
 Leu Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp
 290 295 300
 Met Ile Cys Val Asp Leu Gly Pro Gln Ala Gln Asp Lys Ala Gly Asp
 305 310 315 320
 Pro Val Ile Leu Trp Gly Glu Gly Leu Pro Val Glu Arg Ile Ala Glu
 325 330 335
 Met Thr Lys Val Ser Ala Tyr Glu Leu Ile Ala Arg Leu Thr Ser Arg
 340 345 350
 Val Ala Met Lys Tyr Val Asp
 355

<210> 39

<211> 359

<212> PRT

<213> Artificial

<220>

<223> gen sintético

5 <400> 39

Met Gln Val Ala Thr Val Val Ile Asn Arg Arg Ala Leu Arg His Asn
 1 5 10 15
 Leu Gln Arg Leu Arg Glu Leu Ala Pro Ala Ser Lys Leu Val Ala Val
 20 25 30
 Val Lys Ala Asn Ala Tyr Gly His Gly Leu Leu Glu Thr Ala Arg Thr
 35 40 45
 Leu Pro Asp Ala Asp Ala Phe Gly Val Ala Arg Leu Glu Glu Ala Leu
 50 55 60
 Arg Leu Arg Ala Gly Gly Ile Thr Lys Pro Val Leu Leu Leu Glu Gly
 65 70 75 80
 Phe Phe Asp Ala Arg Asp Leu Pro Thr Ile Ser Ala Gln His Phe His
 85 90 95
 Thr Ala Val His Asn Glu Gly Gln Leu Ala Ala Leu Glu Glu Ala Ser
 100 105 110
 Leu Asp Glu Pro Val Thr Val Trp Met Lys Leu Asp Thr Gly Met His
 115 120 125
 Arg Leu Gly Val Arg Pro Glu Gln Ala Glu Ala Phe Tyr His Arg Leu
 130 135 140
 Thr Gln Cys Lys Asn Val Arg Gln Pro Val Asn Ile Val Ser His Phe
 145 150 155 160
 Ala Arg Ala Asp Glu Pro Lys Cys Gly Ala Thr Glu Lys Gln Leu Ala
 165 170 175
 Ile Phe Asn Thr Phe Cys Glu Gly Lys Pro Gly Gln Arg Ser Ile Ala
 180 185 190
 Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Val
 195 200 205
 Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asp Arg Ser
 210 215 220
 Thr Gly Ala Asp Phe Gly Cys Gln Pro Val Met Ser Leu Thr Ser Ser
 225 230 235 240

Leu Ile Ala Val Arg Glu His Lys Ala Gly Glu Pro Val Gly Tyr Gly
245 250 255

Gly Thr Trp Val Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
260 265 270

Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
275 280 285

Leu Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp
290 295 300

Met Ile Cys Val Asp Leu Gly Pro Asn Ala Gln Asp Asn Ala Gly Asp
305 310 315 320

Pro Val Val Leu Trp Gly Glu Gly Leu Pro Val Glu Arg Ile Ala Glu
325 330 335

Met Thr Lys Val Ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg
340 345 350

Val Ala Met Lys Tyr Ile Asp
355

<210> 40

<211> 358

5 <212> PRT

<213> Escherichia coli

<220>

<221> característica nueva

<222> (23)..(23)

10 <223> Xaa puede ser cualquier aminoácido de origen natural

<220>

<221> característica nueva

<222> (56)..(56)

<223> Xaa puede ser cualquier aminoácido de origen natural

15 <220>

<221> característica nueva

<222> (279)..(279)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 40

Met Gln Ala Ala Thr Val Val Ile Asn Arg Arg Ala Leu Arg His Asn
1 5 10 15

Leu Gln Arg Leu Arg Glu Xaa Ala Pro Ala Ser Lys Met Val Ala Val
20 25 30

Val Lys Ala Asn Ala Tyr Gly His Gly Leu Leu Glu Thr Ala Arg Thr
35 40 45

Leu Pro Asp Ala Asp Ala Phe Xaa Val Ala Arg Leu Glu Glu Ala Leu
 50 55 60
 Arg Leu Arg Ala Gly Gly Ile Thr Lys Pro Val Leu Leu Leu Glu Gly
 65 70 75 80
 Phe Phe Asp Ala Arg Asp Leu Pro Thr Ile Ser Ala Gln His Phe His
 85 90 95
 Thr Ala Val His Asn Glu Glu Gln Leu Ala Ala Leu Glu Glu Ala Ser
 100 105 110
 Leu Asp Glu Pro Val Thr Val Trp Met Lys Leu Asp Thr Gly Met His
 115 120 125
 Arg Leu Gly Val Arg Pro Glu Gln Ala Glu Ala Phe Tyr His Arg Leu
 130 135 140
 Thr Gln Cys Lys Asn Val Arg Gln Pro Val Asn Ile Val Ser His Phe
 145 150 155 160
 Ala Arg Ala Asp Glu Pro Lys Cys Gly Ala Thr Glu Lys Gln Leu Ala
 165 170 175
 Ile Phe Asn Thr Phe Cys Glu Gly Lys Pro Gly Gln Arg Ser Ile Ala
 180 185 190
 Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Val
 195 200 205
 Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asp Arg Ser
 210 215 220
 Thr Gly Ala Asp Phe Gly Cys Gln Pro Val Met Ser Leu Thr Ser Ser
 225 230 235 240
 Leu Ile Ala Val Arg Glu His Lys Ala Gly Glu Pro Val Gly Tyr Gly
 245 250 255
 Gly Thr Trp Val Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
 260 265 270
 Gly Tyr Gly Asp Gly Tyr Xaa Ala Ala Pro Ser Gly Thr Pro Val Leu
 275 280 285
 Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp Met
 290 295 300
 Ile Cys Val Asp Leu Gly Pro Gln Ala Gln Asp Lys Ala Gly Asp Pro
 305 310 315 320
 Val Ile Leu Trp Gly Glu Gly Leu Pro Val Glu Arg Ile Ala Glu Met

ES 2 382 408 T3

325 330 335
 Thr Lys Val Ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg Val
 340 345 350
 Ala Met Lys Tyr Val Asp
 355

<210> 41

<211> 359

5 <212> PRT

<213> Artificial

<220>

<223> gen sintético

<400> 41

10

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

Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Val
195 200 205

Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu His Lys Pro
210 215 220

Trp Gly Pro Asp Phe Gly Phe Gln Pro Val Met Ser Leu Thr Ser Ser
225 230 235 240

Leu Ile Ala Val Arg Asp His Lys Ala Gly Glu Pro Val Gly Tyr Gly
245 250 255

Gly Thr Trp Val Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
260 265 270

Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
275 280 285

Leu Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp
290 295 300

Met Ile Cys Val Asp Leu Gly Pro Gln Ala Gln Asp Lys Ala Gly Asp
305 310 315 320

Pro Val Ile Leu Trp Gly Glu Gly Phe Pro Val Glu Arg Ile Ala Glu
325 330 335

Met Thr Lys Val Ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg
340 345 350

Val Ala Met Lys Tyr Val Asp
355

<210> 42

<211> 359

5 <212> PRT

<213> CFT073 de Escherichia coli

<400> 42

ES 2 382 408 T3

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

Phe Phe Glu Ala Asp Asp Leu Pro Thr Ile Ser Ala Glu His Leu His
 85 90 95
 Thr Ala Val His Asn Glu Glu Gln Leu Val Ala Leu Glu Asn Ala Glu
 100 105 110
 Leu Lys Glu Pro Val Thr Val Trp Met Lys Leu Asp Thr Gly Met His
 115 120 125
 Arg Leu Gly Val Leu Pro Glu Gln Ala Glu Ala Phe Tyr Gln Arg Leu
 130 135 140
 Ser Gln Cys Lys Asn Val Arg Gln Pro Val Asn Ile Val Ser His Phe
 145 150 155 160
 Ala Arg Ala Asp Glu Pro Gln Ser Gly Ala Thr Glu Lys Gln Leu Asp
 165 170 175
 Ile Phe Asn Thr Phe Cys Glu Gly Lys Pro Gly Gln Arg Ser Ile Ala
 180 185 190
 Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Ala
 195 200 205
 Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asp Gly Thr
 210 215 220
 Thr Gly Ala Asp Phe Gly Cys Gln Pro Val Met Ser Leu Thr Ser Ser
 225 230 235 240
 Leu Ile Ala Val Arg Glu His Lys Ala Gly Glu Pro Val Gly Tyr Gly
 245 250 255
 Gly Thr Trp Val Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
 260 265 270
 Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
 275 280 285
 Leu Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp
 290 295 300
 Met Ile Cys Val Asp Leu Gly Pro Gln Ala Gln Asp Lys Ala Gly Asp
 305 310 315 320
 Pro Val Ile Leu Trp Gly Glu Gly Leu Pro Val Glu Arg Ile Ala Glu
 325 330 335
 Met Thr Lys Val Ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg
 340 345 350

Val Ala Met Lys Tyr Val Asp
355

<210> 43

<211> 359

5 <212> PRT

<213> Salmonella entérica subespecie entérica serovar

<400> 43

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

[illegible]

<210> 44

<211> 359

5 <212> PRT

<213> Enterobacter sp. 638

<400> 44

ES 2 382 408 T3

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

Leu Ser Glu Pro Val Thr Val Trp Met Lys Leu Asp Thr Gly Met His
 115 120 125
 Arg Leu Gly Val Arg Pro Glu Gln Ala Glu Ala Phe Tyr Gln Arg Leu
 130 135 140
 Ser Gln Cys Lys Asn Val His Gln Pro Val Asn Ile Val Ser His Phe
 145 150 155 160
 Ala Arg Ala Asp Glu Pro Glu Cys Gly Ala Thr Glu Gln Gln Leu Asp
 165 170 175
 Ile Phe Asn Thr Phe Cys Glu Gly Lys Pro Gly Met Arg Ser Ile Ala
 180 185 190
 Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Ala
 195 200 205
 Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asn Lys Pro
 210 215 220
 Trp Gly Pro Asp Phe Gly Leu Gln Pro Val Met Ser Leu Val Ser Asn
 225 230 235 240
 Leu Ile Ala Val Arg Glu His Lys Ala Gly Glu Pro Val Gly Tyr Gly
 245 250 255
 Gly Thr Trp Val Ser Glu Ser Asp Thr Arg Leu Gly Val Val Ala Met
 260 265 270
 Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
 275 280 285
 Leu Val Asn Gly Arg Glu Val Lys Ile Val Gly Arg Val Ala Met Asp
 290 295 300
 Met Ile Cys Val Asp Leu Gly Pro Asp Ala Gln Asp Lys Ala Gly Asp
 305 310 315 320
 Ala Val Val Leu Trp Gly Glu Gly Leu Pro Val Glu Arg Ile Ala Glu
 325 330 335
 Ile Thr Lys Val Ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg
 340 345 350
 Val Ala Met Lys Tyr Ile Asp
 355

REIVINDICACIONES

1. Una planta transgénica transformada con un vector de expresión que comprende una racemasa aislado alanina polinucleótido que codifica, en donde la planta transformada demuestra una mayor resistencia a los nematodos en comparación con una variedad de tipo salvaje de la planta.
- 5 2. La planta transgénica de la reivindicación 1, en donde el polinucleótido aislado se selecciona del grupo constituido por:
 - a) un polinucleótido que tiene una secuencia tal como se define en SEQ ID NO: 5 o 7;
 - b) un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en SEQ ID NO: 6 u 8;
 - 10 c) un polinucleótido que tiene al menos 70% de identidad de secuencia a un polinucleótido que tiene la secuencia de aminoácidos tal como se define en SEQ ID NO: 5 o 7;
 - d) un polinucleótido que codifica un polipéptido que tiene al menos 70% secuencia identidad de un polipéptido que tiene una secuencia tal como se define en SEQ ID NO: 6 u 8;
 - e) un polinucleótido que se hibrida en condiciones rigurosas a un polinucleótido que tiene una secuencia tal como se define en SEQ ID NO: 5 o 7;
 - 15 f) un polinucleótido que se hibrida bajo condiciones estrictas a un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en SEQ ID NO: 6 u 8;
 - g) un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en cualquiera de SEQ ID NOS: 12 a 44, y
 - 20 h) un polinucleótido que codifica un polipéptido que tiene al menos 90% de identidad de secuencia con una secuencia tal como se define en cualquiera de SEQ ID NO: 12 a 44.
3. La planta de la reivindicación 1, define además como una monocotiledónea o una dicotiledónea.
4. La planta de la reivindicación 1, caracterizado porque la planta se selecciona del grupo que consiste en maíz, trigo, arroz, cebada, avena, centeno, sorgo, plátano, ballico, guisante, alfalfa, soja, zanahoria, apio, tomate, patata, algodón, tabaco, pimiento, colza, remolacha, col, coliflor, brócoli, lechuga y Arabidopsis thaliana.
- 25 5. Una semilla que es cierto para la cría un transgén que comprende un polinucleótido que codifica alanina racemasa, en donde la expresión del polinucleótido confiere una mayor resistencia a los nematodos de la planta producida a partir de la semilla.
6. La semilla de la reivindicación 5, en el que el polinucleótido se selecciona del grupo constituido por:
 - a) un polinucleótido que tiene una secuencia tal como se define en SEQ ID NO: 5 o 7;
 - 30 b) un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en SEQ ID NO: 6 u 8;
 - c) un polinucleótido que tiene al menos 70% de identidad de secuencia a un polinucleótido que tiene la secuencia de aminoácidos tal como se define en SEQ ID NO: 5 o 7;
 - d) un polinucleótido que codifica un polipéptido que tiene al menos 70% de identidad de secuencia a un polipéptido que tiene una secuencia tal como se define en SEQ ID NO: 6 u 8;
 - 35 e) un polinucleótido que se hibrida en condiciones rigurosas a un polinucleótido que tiene una secuencia tal como se define en SEQ ID NO: 5 o 7;
 - f) un polinucleótido que se hibrida en estricta de condiciones a un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en SEQ ID NO: 6 u 8;
 - 40 g) un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en cualquiera de SEQ ID NO: 12 a 44, y

- h) polinucleótido una que codifica un polipéptido que tiene al menos 90% de identidad de secuencia con una secuencia tal como se define en cualquiera de SEQ ID NO: 12 a 44.
- 5 7. Uso de un vector de expresión que comprende un elemento regulador de la transcripción unido operativamente a un polinucleótido que codifica alanina racemasa, en donde la expresión del polinucleótido confiere una mayor resistencia a los nematodos de una planta.
8. El uso de la reivindicación 7, en el que el polinucleótido se selecciona del grupo constituido por:
- a) un polinucleótido que tiene una secuencia tal como se define en SEQ ID NO: 5 o 7;
- b) un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en SEQ ID NO: 6 u 8;
- 10 c) un polinucleótido que tiene al menos 70% de identidad de secuencia a un polinucleótido que tiene una secuencia tal como se define en SEQ ID NO: 5 o 7;
- d) un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en la SEQ ID NO: 6 u 8;
- e) un polinucleótido que se hibrida en condiciones rigurosas a un polinucleótido que tiene una secuencia tal como se define en SEQ ID NO: 5 o 7, y
- 15 f) un polinucleótido que se hibrida bajo condiciones estrictas a un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en SEQ ID NO: 6 u 8;
- g) un polinucleótido que codifica un polipéptido que tiene una secuencia tal como se define en cualquiera de SEQ ID NO: 12 a 44, y un
- h) polinucleótido que codifica un polipéptido que tiene al menos 90% de identidad de secuencia con una secuencia tal como se define en cualquiera de SEQ ID NO: 12 a 44.
- 20 9. El vector de expresión de la reivindicación 8, en el que el elemento regulador de la transcripción es un promotor de regulación de la expresión específica de raíz o sincitios específica del polinucleótido.
10. Un método para producir una planta transgénica de haber aumentado la resistencia a nematodos, en donde el método comprende las etapas de:
- a) introducir en la planta del vector de expresión que comprende un polinucleótido que codifica alanina racemasa, y
- 25 b) la selección de plantas transgénicas con mayor resistencia a los nematodos.
11. El método de la reivindicación 10, en el que la planta es una monocotiledónea.
12. El método de la reivindicación 11, en el que la planta se selecciona del grupo que consiste en maíz, trigo, arroz, cebada, avena, centeno, sorgo, plátano, y ballico.
13. El método de la reivindicación 10, en el que la planta es una dicotiledónea.
- 30 14. El método de la reivindicación 13, en el que la planta se selecciona del grupo que consiste de arveja, alfalfa, soja, zanahoria, apio, tomate, papa, algodón, tabaco, pimienta, semillas de colza, remolacha, col, coliflor, brócoli, lechuga y Arabidopsis thaliana.
15. El método de la reivindicación 10, en el que el polinucleótido tiene una secuencia tal como se define en SEQ ID NO: 5 o 7.
- 35 16. El método de la reivindicación 10 en el que el polinucleótido codifica un polipéptido que tiene una secuencia tal como se define en SEQ ID NO: 6 u 8.

Figura 1

Proteína DadX	Porcentaje de identidad de aminoácidos	No. de acceso del homólogo	Nombre del gen
DadX	100.00%	NP 415708	alanina racemasa [K12 de Escherichia coli]
DadX	99,70%	ZP_00705413	alanina racemasa [HS de Escherichia coli]
DadX	99,20%	NP 287429	alanina racemasa [Escherichia coli]
DadX	99,40%	YP_540389	alanina racemasa 2, catabólica [Escherichia coli]
DadX	99,20%	ZP_00734403	alanina racemasa [53638 de Escherichia coli]
DadX	99,20%	ZP_00703512	alanina racemasa [E24377A de Escherichia coli]
DadX	99,20%	YP_688709	alanina racemasa 2, catabólica [cepa 8401 de Shigella flexneri 5]
DadX	98,90%	YP_310139	alanina racemasa 2, catabólica [Ss046 de Shigella sonnei]
DadX	98,90%	ZP 00714931	alanina racemasa [B7A de Escherichia coli]
DadX	99,20%	AAC61705	alanina racemasa [Escherichia coli]
DadX	98,90%	NP 707087	alanina racemasa [cepa 301 de Shigella flexneri 2a]
DadX	98,90%	ZP 00725784	alanina racemasa [F11 de Escherichia coli]
DadX	98,90%	AAC36881	alanina racemasa catabólica
DadX	98,00%	NP_753544	alanina racemasa [CFT073 de Escherichia coli]
DadX	90,40%	ZP 00699231	alanina racemasa [BS512 de Shigella boydii]
DadX	89,30%	NP 309712	alanina racemasa [cepa Sakai 0157:H7 de Escherichia coli]
DadX	85,70%	YP_216782	alanina racemasa [subespecie de Salmonella entérica]
DadX	85,70%	NP_460758	alanina racemasa [LT2 de Salmonella typhimurium]
DadX	73,60%	NP 456309	alanina racemasa [subespecie de Salmonella entérica]
DadX	80,10%	ZP 01590546	alanina racemasa [Enterobacter sp. 638]

Figura 2

Proteína Alr	Porcentaje de identidad de aminoácidos	No. de acceso del homólogo	Nombre del gen
Alr	99,40%	BAD88778	alanina racemasa [constructo sintético]
Alr	99,40%	NP_290686	alanina racemasa [EDL933 0157:H7 de Escherichia coli]
Alr	99,70%	YP_312966	alanina racemasa 1 [Ss046 de Shigella sonnei]
Alr	99,70%	ZP_00722578	COG0787: Alanina racemasa [E110019 de Escherichia coli]
Alr	99,70%	ZP_00712036	COG0787: Alanina racemasa [B171 de Escherichia coli]
Alr	99,70%	ZP_00699606	COG0787: Alanina racemasa [BS512 de Shigella boydii]
Alr	99,70%	ZP_00702044	COG0787: Alanina racemasa [E24377A de Escherichia coli]
Alr	99,20%	YP_405886	alanina racemasa 1 [Sd197 de Shigella dysenteriae]
Alr	98,10%	BAD88779	alanina racemasa [constructo sintético]
Alr	98,90%	AAC43147	alanina racemasa [Escherichia coli]
Alr	97,50%	BAD88777	alanina racemasa [constructo sintético]
Alr	93,90%	NP_756880	alanina racemasa [CFT073 de Escherichia coli]
Alr	91,10%	YP_219113	alanina racemasa [subespecie de Salmonella entérica]
Alr	89,70%	ZP_01589275	alanina racemasa [Enterobacter sp. 638]

Figura 3

ADN para DadX (SEQ ID NO: 5)

ATGACCCGTCCGATACAGGCCAGCCTCGATCTGCAGGCATTAAACAGAA
TCTGTCCATTGTCCGCCAGGCCGCGACGCACGCGCGCGTCTGGTCGGTGG
TAAAAGCGAACGCTTACGGGCATGGTATTGAGCGTATCTGGAGCGCGATC
GGGGCCACCGATGGCTTTGCATTGCTTAACCTGGAAGAGGCAATAACGTT
ACGTGAGCGCGGCTGGAAAGGACCGATCCTGATGCTGGAAGGATTTTTCC
ATGCTCAGGATCTGGAGATTTATGACCAGCACCGCCTGACCACCTGCGTA
CACAGCAACTGGCAGCTCAAAGCACTGCAAAATGCGCGGCTAAAAGCACC
GTTGGATATTTATCTTAAAGTGAACAGTGGGATGAATCGGTTGGGCTTCC
AGCCCGATCGCGTGCTTACCGTCTGGCAGCAGTTGCGGGCAATGGCGAAT
GTTGGCGAAATGACCCTGATGTTCGATTTTGCCGAAGCGGAACATCCTGA
TGGAATTTCCGGCGCGATGGCGCGTATTGAGCAGGCGGCGGAGGGGCTGG
AGTGTCGGCGTTTCGTTGTCCAATTCGGCGGCGACTCTGTGGCACCCGGAA
GCGCATTTTGACTGGGTTCGGCCTGGCATTATTTTGTATGGCGCTTCGCC
GTCCGGTCAGTGGCGTGATATCGCCAATACCGGATTACGTCCGGTGATGA
CGCTAAGCAGTGAGATTATTGGTGTCCAGACGCTAAAAGCGGGCGAGCGT
GTGGGCTACGGCGGTTCGCTATACTGCGCGCGATGAACAGCGAATCGGCAT
TGTCGCCGCGAGGGTACGCCGACGGTTATCCGCGCCACGCGCCTACCGGTA
CCCCTGTTTTAGTGGACGGCGTGCGCACCATGACGGTGGGGACCGTCTCG
ATGGATATGCTAGCGGTCGATTTAACGCCTTGCCCGCAGGCGGGTATTGG
TACGCCGGTTGAGCTGTGGGGCAAGGAGATCAAAATTGATGATGTGCCCG
CCGCTGCCGGAACGGTGGGCTATGAGTTGATGTGCGCGCTGGCGCTACGC
GTCCCGGTTGTGACGGTGTA

Proteína DadX (SEQ ID NO: 6)

mtrpiqasldlqalkqnlsivrqaatharvwsvkanayghgieriwsaigatdgfallnlecaitrergwkpgpilmlegffh
aqdleiydqhrltccvhsnwqlkalqnarlkapldiylkvnsgmnrllgfpdrvtvwqqlramanvgemtlmshfaeae
hpdgisgamarieqaaeglecrslsnsaatlwhpeahfdwvrpgiilygaspsgqwrldiantglrpvmtlsseiigvqtlka
gervgyggrytardeqrigivaagyadgyprhaptgtpvlvdgvrmtvgtvsmdmlavdltpcpqagigtppvlgkei
kiddvaaaagtvgyelmcalalrvpvtv*

Figura 4

ADN para Alr (SEQ ID NO: 7)

ATGCAAGCGGCAACTGTTGTGATTAACCGCCGCGCTCTGCGACACAACCT
GCAACGTCTTCGTGAACTGGCCCCCTGCCAGTAAAATGGTTGCGGTGGTGA
AAGCGAACGCTTATGGTCACGGTCTTCTTGAGACCGCGCGAACGCTCCCC
GATGCTGACGCCTTTGGCGTAGCCCGTCTCGAAGAAGCTCTGCGACTGCG
TGCGGGGGGAATCACCAAACCTGTACTGTTACTCGAAGGCTTTTTTGATG
CCAGAGATCTGCCGACGATTTCTGCGCAACATTTTCATACCGCCGTGCAT
AACGAAGAACAGCTGGCTGCGCTGGAAGAGGCTAGCCTGGACGAGCCGGT
TACCGTCTGGATGAACTCGATACCGGTATGCACCGTCTGGGCGTAAGGC
CGGAACAGGCTGAGGCGTTTTATCATCGCCTGACCCAGTGCAAAAACGTT
CGTCAGCCGGTGAATATCGTCAGCCATTTTGCGCGCGCGGATGAACCAA
ATGTGGCGCAACCGAGAAACAACCTCGCTATCTTTAATACCTTTTGCGAAG
GCAAACCTGGTCAACGTTCCATTGCCGCGTCGGGTGGCATTCTGCTGTGG
CCACAGTCGCATTTTGACTGGGTGCGCCCGGGCATCATTCTTTATGGCGT
CTCGCCGCTGGAAGATCGCTCCACCGGTGCCGATTTTGGCTGTCAGCCAG
TGATGTCACTAACCTCCAGCCTGATTGCCGTGCGTGAGCATAAAGCCGGA
GAGCCTGTTGGTTATGGTGGAACTGGGTAAGCGAACGTGATACCCGTCT
TGGCGTAGTCGCGATGGGCTATGGCGATGGTTATCCGCGCGCCGCGCCGT
CCGGTACGCCAGTGCTGGTGAACGGTCGCGAAGTACCGATTGTCGGGCGC
GTGGCGATGGATATGATCTGCGTAGACTTAGGTCCACAGGCGCAGGACAA
AGCCGGGGATCCGGTCATTTTATGGGGCGAAGGTTTGCCCGTAGAACGTA
TCGCTGAAATGACGAAAGTAAGCGCTTACGAACTTATTACGCGCCTGACT
TCAAGGGTCGCGATGAAATACGTGGATTAA

Proteína Alr (SEQ ID NO: 8)

mqaatvvinnrralrhnlqrlrelapaskmvavvkanayghlletartlpdadafgvarleearlraggitkpvlllegffdard
lptisaqhftavhneeqlaalecasldepvvwmkldtgmhrlgvrpeqacafyhrltqcknvrqpvnivshfaradepkc
gatekqlaifntfcegkpgqrsiaasggillwpqshfdwvrpgiilygvspledrsgadfgcqpvmstssliavrehkagep
vgyggtwvserdtrlgvvamgygdgypraapsgtpvlvngrevpivgrvamdmicvdlgpqaqdkagdpvilwgegl
pveriaemtkvsayelitrtsrvamkyvd*

Figura 5a**Promotor MtN3 (SEQ ID NO: 9)**

Gaagccacgtcatgaagagtatatcatcttcagtaatgtttgagacgcctctataatgctttaccaacaaaacaaaacaaaaaa
 gaacatttgaaccatttgtattaaaaaaaaaagggtatattagggccataatattataggtaacatgaaatatcaaatgacacgcaag
 agtttgtcaaaaatgaaccatcacacatcagagattatggcaataatgtttgtgtctcttgcctcaccataacataagcctct
 ataactggagagaagaaaaaaaaaagtggaggggctaggggtgggaatttggagaatacagttatattgagcattgagcaagtt
 gatagaaagcttctcaatttgcataaaatttgcacccacatgattattaaagacgtagacagcacttctctctctttttctataagttc
 ttatatattgttctcatgttttaattattactttatgtacgcgtctaacagtagtctctccaaactgctataaataagagcctcttcaacg
 cacctcttggcagtacaaaaattattcatctcttctaagttctaattttctaagcattcagtaaaagaactaacc

Promotor POX (SEQ ID NO: 10)

cgaagagcataagtttgttcaaatggcccaataacaaattaaaaacatgtaaagtagtcagtttaacaagcatttgcataaagtgt
 ggtaatatattatataaacttcacatccaatgagcattcatgtaatttaaagtaactgaagttaagtatctagaagccttttcttattg
 gttattaatttgcctaattttcttataagtttaatttctgttgggtgaaaatgtgaccggagaaggtatctaactttttttctttaaataa
 ttccactaaaaatttattctgtatgtaacgcataatagtaaaatctagaagcgaccggcgctgcctccttggaaagtaactctgtaa
 agtaaaagccgcgtagtgtaaaagtatatgacttcttcccataattattttataattagctttaaataatttaacataataatc
 gttttacgagaaagatcttcacactcgattagtatacattacatttaattccctagttcataaaatggataacaaaaggctgtgcgaga
 ttacaactgtacttgataatttgtataaataatcctttatgaatatatttagcattgatgaccgtacatggtaatccagctctgcagcat
 aacggagtagtatattaaatgaacactttctgttcgtatcaaatggatcgaatattattagagtgatcattcagaagaaaaaagag
 agagaagaaaacctacagtgtaaacatttttttgcataaacctacagtgtaaacatgaagtgtataatttctgcaaatagaaat
 caagaacagaaagagttgcttgaggaaaagaaatagaaaattaagaaatctagtgtatgtaataaatcttccataaaatcaaatg
 tttgtccaaagtattagttaaataattagggccactattcttgacaactcttttaacaaactcttctatatttctcgtgtacatatgctg
 aaaaagatgtatgtctaatccataatataatctgtataatgcgacttccattatctattagtacgacttctaacctagaagataacaagca
 tttagctaggcatcaaaatcaacgtggaaaaacctacgaaaagcacgaagtgttaaatctgtgtaggggtggcgtaaagggtaaa
 gactaaagactgagaatctagggtcaaggcgtaaacctgttctgcttttgggtttcattttattggcgaacaacattgatgtgtgtg
 accatttgggttcagggttgagacaagataatattgttctctcaccctctaggattactcgggtgctaagactcacttagtactatt
 gctatatcgatatactagttcattacaaaaatggagcttcaaatctcaggttccaatatctgaaagcattgtttaaagagatttgttt
 ctccctgcacaattagtttataacttcatatatacacaatcttcaatttacaaccagggtgtgtgtgaaccttcacataatctctcttatt
 cattcatgtatatatccaataaaagtgcgatgtgaaattatatatctccatctaattgttagactattcccggtcttgactataaattta
 aagtattagacgagctaatattatttagcacaacaatttctgtaacagtgctacgcttatcactaccaaagaataaacactgatc
 tgttttaactcttattttctcaccatattcaagtcactattgcaagacttcgagataaataatttgatggctatactatttacttgacat
 ttgggaaaatatatttctgctgataaatttgggttttacttctctccgacggatagaaacaattcaattacatgcgaaaatgataatt
 caacctataaaacaaaacaaataacagaatgcacatttttcaacgcgttaggtcacctatcttcaacttagaacatcccttcacg
 tctctatataaacctcgactctgttatccttgttcttcaagtacaacaatcaactctaagtctattataattcaagctttgtttaacctaac
 a

Figura 5b**Promotor TPP (SEQ ID NO: 11)**

gtagtgccttcatggataccaaaagagaaaatttgatttagtgcatacatataacaatataacgccgcataataactgtataaaa
 cagtcattgtaacgatatgacagcagtaatacagttccaagagacgttataatcgatgcaatcatatgcttgcgtagatttccaaca
 gttttgttcgttgataggaggaactcaacactctagggtagtgattggtagacactattagcacaaaaatattaatttactctgatg
 ttacaaaaaagttaaccaatcaaatatttaagagatcgctactcttcacggcgactctaaaaaccaaagatataggttagactcata
 actactttataaagaaaatgttaacgataactaccgagatctaataaataaaccttcatttcaagtatattatattgcttctttgtttat
 atatcaaaccaagtcttggtttataaaaatattagataaaactcgtctaaataggtagggtgaaaaataaaattttaattttatcgataat
 atttaaaatttgaaaagttaataatgatccacacatttttctaattttaatttagtaattttgtattaaataaaatttcaatcatatacattc
 gattttctatacattttaactatctatttctgcataataaactgtatttcatttatagccttcattcttgatgatatttaaattttaaatag
 aattcatacactttttaataatttaatttagtattttcttaaatccaaattttaattcttaacatttaaatatctactttaacataatacaaatacaat
 ttaatttcattgtattaaattcaaatataatttgattataataaaatacaatttaattctaaaaagtcacatttagattttaatttcttttagtt
 ttgaaaattaaaaatttaatttattagatatatatgttacttttcagtttcttatttatttaagaaaaaatatttttaacacatgtcaactt
 gtaacaatagactgaacacgtcattttatattatgttttagttttgaaaattaaagttaattaaatatttatatttcttttttagcttttcta
 tatttttaaaatagtaaatattttaatacaaatcaatatctgaacaatagatttgatacataacataatcctataaattattaacttgga
 acgatagtttatataataaaattattttcttaagttcttaaccataacaattaaactatatttagcgaagaaaagagaataaccga
 gagaacgcaacttgcaactaaaagctaccactttggcaaatcactcatttatattatatactatcacctcaattcaatcgaaacctc
 aaaataacactaatatatacacaaaagaacaacagaataacaccgaagaatataggttaggaaaatccagaatttggtgagacta
 aagagatcaaattttcgatacaagggtttgctcaatttgattttcataataaaattctttatttcacatagacttacatgattagttttctt
 ttaataaaaaaaaacacgcgacatgaaaattatattatctcagtggttgctgaatttgaatttgaattttgagttaaatactacacattgtt
 gacaacttattaactttacaagtctgctacaaatattgtcaaatatttactaattaatggacccaaaatccttaacttgcaaattgtatc
 tacatcaacttaaaaattaggaatatgcgacccaaaaaaaaaaaaactaggaataataaaaaaaatggaatgatgtggaggaa
 gctctttactctttgagaggaagttataaattgaccacacatttagtctattatcatcacatgtattaagacttgacaacttgctctttctca
 caccaaacctctctctgtttcataacatctgctctttcttttttcttaagccctta