

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 398 918**

51 Int. Cl.:

C12Q 1/68 (2006.01)

C12N 15/74 (2006.01)

A23C 9/12 (2006.01)

C12N 1/21 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **25.08.2006 E 10178697 (8)**

97 Fecha y número de publicación de la concesión europea: **31.10.2012 EP 2325332**

54 Título: **Un método y una ordenación para soportar verticalmente elementos de resistencia eléctrica pendientes**

30 Prioridad:

26.08.2005 US 711396 P

19.05.2006 US 747683 P

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

22.03.2013

73 Titular/es:

DUPONT NUTRITION BIOSCIENCES APS

(100.0%)

Langebrogade 1, Postboks 17

1001 Copenhagen K, DK

72 Inventor/es:

HORVATH, PHILIPPE;

BARRANGOU, RODOLPHE;

FREMAUX, CHRISTOPHE;

BOYVAL, PATRICK y

ROMERO, DENNIS

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 398 918 T3

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

DESCRIPCIÓN

Un método y una ordenación para soportar verticalmente elementos de resistencia eléctrica pendientes

Campo de la invención

5 La presente invención se refiere a la modulación entre otras de la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo. En particular, la presente invención se refiere, en un aspecto, al uso de uno o más genes *cas* o proteínas para la modulación de la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.

Antecedentes de la invención

10 Cultivos, tales como los cultivos iniciadores ("starter"), se usan ampliamente en la industria alimenticia en la producción de productos fermentados incluyendo los productos de leche (tales como yogurt, mantequilla y queso), productos cárnicos, productos de panadería, vino y productos vegetales. La preparación de cultivos es labor intensa, que ocupa mucho espacio y equipamiento, y hay un riesgo considerable de contaminación con bacteria de degradación y/o fagos durante la etapa de propagación. El fracaso de los cultivos bacterianos por infección de bacteriófago (fago) y multiplicación es un problema muy importante con el uso industrial de cultivos bacterianos. Hay
15 muchos tipos diferentes de fagos con mecanismos variantes para atacar las bacterias. Además, aparecen nuevas cepas de bacteriófagos.

Las estrategias usadas en la industria para minimizar la infección de bacteriófago, y así el fracaso de un cultivo bacteriano, incluyen el uso de: (i) cultivos iniciadores mezclados; y (ii) el uso alterno de cepas que tienen diferentes perfiles de susceptibilidad a fago (rotación de cepa).

20 (i) Tradicionalmente, los cultivos iniciadores en la industria láctea son mezclas de cepas bacterianas de ácido láctico. La composición compleja de los cultivos iniciadores mezclados asegura que está presente un cierto nivel de resistencia a ataque de fago. Sin embargo, el sub cultivo repetido de cultivos de cepa mezclada conduce a cambios impredecibles en la distribución de cepas individuales y finalmente dominancia de cepa indeseada. Esto por turnos puede conducir a susceptibilidad incrementada a ataque de fago y riesgo de fracasos de fermentación.

25 (ii) La rotación de cepas bacterianas seleccionadas que son sensibles a diferentes fagos es otro enfoque para limitar el desarrollo del fago. Sin embargo, es difícil y engorroso identificar y seleccionar un número suficiente de cepas que tengan perfiles de diferente tipo de fago para proporcionar un programa de rotación eficiente y fiable. Además, el uso continuo de cepas requiere realizar un cuidadoso seguimiento de los nuevos fagos infecciosos y la necesidad de sustituir rápidamente una cepa que está infectada por el nuevo bacteriófago por una cepa resistente. En la producción de plantas donde grandes cantidades de cultivos iniciadores a granel ("bulk starter") se realizan con anticipación, normalmente no es posible dicha respuesta rápida.

Se han realizado varios intentos para mejorar la resistencia de los cultivos para usar en la industria.

35 Pedersen et al. ("7th symposium on lactic acid bacteria: genetics, metabolism and applications", 1-5 de Septiembre, 2002, Egmond aan Zee, Países Bajos) enseña una cepa de *Lactococcus lactis* resistente a fago, la cual tiene no actividad timidilato sintasa y la cual requiere timidina durante la replicación del ADN.

40 El documento WO 01/14520 describe una bacteria de ácido láctico que tiene una susceptibilidad reducida hacia el ataque por al menos un tipo de bacteriófago. Dichas bacterias de ácido láctico comprenden un gen mutado implicado en el metabolismo de pirimidina, principalmente *pyrG* que da como resultado un defecto en la CTP sintetasa.

Kosuge et al. (1998- "Appl. Environ. Microbiol", Volumen: 64, Ejemplar: 11, Página(s): 4.328-4.332) y Kosuge et al (1994-FEMS "Microbiology Letters", 123 (1/2)55-62) enseña una bacteria *Thermus thermophilus* HB27 que está mutada en el gen *proB* y es incapaz de utilizar prolina para el crecimiento.

45 Bolotin et al. "Microbiology" 2005, Vol. 151, p2.551-2.569 enseña que numerosos genomas procariotas contienen estructuras conocidas como CRISPRS.

Jansen et al. "Molecular Microbiology" 2002, Vol. 43, Nº 6, p1.565-1.575 usó análisis *in silico* para estudiar una familia novedosa de secuencias de ADN repetitivas (CRISPR) en procariotas.

Sin embargo, hay una continua necesidad de mejorar los cultivos para usar en la industria.

Compendio de la invención

En la presente memoria se describe el uso de los loci CRISPR o un componente de los mismos para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.

5 CRISPRs (del Inglés “Clustered Regularly Interspaced Short Palindromic Repeats” “Repeticiones Palindrómicas Cortas Regularmente Interespaciadas Agrupadas”) (también denominadas SPIDRs- del Inglés “Spacer Interspersed Direct Repeats”, “Repeticiones directas intercaladas de Espaciador”) constituyen una familia de loci de ADN recientemente descritos. Los loci CRISPR consisten en repeticiones de ADN cortas y altamente conservadas (generalmente de 24 a 40 pb, repetidas entre 1 y 140 veces) que son parcialmente palindrómicas. Las secuencias repetidas (normalmente específicas a una especie) están interespaciadas por secuencias variables de longitud constante (generalmente de 20 a 58 pb dependiendo de CRISPR). Hasta 20 loci CRISPR distintos se han encontrado dentro de un cromosoma sencillo.

Aunque la función biológica de los loci CRISPR es desconocida se han propuesto algunas hipótesis. Por ejemplo, se ha propuesto que pueden estar implicados en el acoplamiento del cromosoma a una estructura celular, o en la replicación del cromosoma y partición del replicón (Jansen et al., 2002; Pourcel et al., 2005). Además, Mojica et al. 2005 plantea la hipótesis de que CRISPR podría estar implicada en conferir inmunidad específica frente a ADN exógeno y Pourcel et al. (2005) plantea la hipótesis de que CRISPRs son estructuras que son capaces de absorber piezas de ADN exógeno como parte de un mecanismo de defensa. Bolotin et al. (2005) sugiere que los elementos espaciador de CRISPR son las trazas de invasiones pasadas por elementos extracromosómicas, y plantea la hipótesis de que proporcionan una célula con inmunidad frente a infección de fago, y más generalmente expresión de ADN exógeno, mediante la codificación de un ARN anti-sentido. Bolotin et al. (2005) también sugiere que los genes *cas* son necesarios para la formación de CRISPR.

Al contrario de las enseñanzas de la técnica anterior que plantean la hipótesis de que CRISPR o los espaciadores CRISPR podrían estar implicados en conferir inmunidad específica, la presente invención se basa, en parte, en el sorprendente descubrimiento de que los genes *cas* o proteínas son requeridos para la inmunidad frente a un ácido nucleico diana o un producto de transcripción del mismo.

Incluso más sorprendentemente, los inventores han descubierto que uno o más genes *cas* o proteínas están asociados con dos o más repeticiones CRISPR dentro de los loci CRISPR. En otras palabras, los genes *cas* o proteínas parecen ser específicos para una repetición CRISPR de ADN dado, que quiere decir que los genes *cas* o proteínas y la secuencia repetida forman un par funcional. Por consiguiente, se pueden usar uno o más espaciadores CRISPR junto con uno o más de estos pares funcionales (es decir, repeticiones CRISPR y genes *cas*) para modular la resistencia de una célula frente a un ácido nucleico o un producto de transcripción del mismo.

En una realización, para que uno o más espaciadores CRISPR confieran inmunidad a la célula, la(s) repetición(es) CRISPR y el(los) gen(es) *cas* o proteínas forma(n) una combinación funcional, es decir, la(s) repetición(es) y el(los) gen(es) *cas* o proteínas son compatible(s).

Por consiguiente, en la presente memoria sugerimos por primera vez que un gen *cas* o proteína influye resistencia, tal como la resistencia de una bacteria a uno o más bacteriófagos. En particular, el conocimiento de dos o más repeticiones CRISPR y/o uno o más genes *cas* o proteínas para una célula dada será una ventaja para predecir, determinar y modificar su resistencia, por ejemplo, su lisotipo, el cual define la resistencia/sensibilidad de una bacteria dada a diversos bacteriófagos. Por lo tanto, la identificación y la detección de loci CRISPR en, por ejemplo, células y bacteriófagos podría ayudar a determinar, predecir y modificar el perfil de resistencia de una célula o interacciones fago-hospedante.

Ventajosamente, la aplicación de uno o más loci CRISPR, dos o más repeticiones CRISPR, uno o más genes *cas* o proteínas y/o uno o más espaciadores CRISPR en ingeniería genética podría conducir a variantes resistentes o sensibles de células para usarse dentro de una amplia diversidad de aplicaciones en la industria de la biotecnología.

En un aspecto se ha proporcionado un método para modular la resistencia de una célula bacteriana frente a un ácido nucleico diana o producto de transcripción de un bacteriófago del mismo que comprende las etapas de:

- (i) identificar uno o más pseudo espaciadores CRISPR en un genoma de bacteriófago que es capaz de proporcionar resistencia al ácido nucleico diana o producto de transcripción del mismo;
- (ii) preparar un ácido nucleico recombinante que comprende un gen *cas1* y al menos dos repeticiones CRISPR junto con dicho identificado uno o más pseudo espaciadores CRISPR identificado; e
- (iii) introducir dicho ácido nucleico recombinante en de dicha célula bacteriana para así volver a la célula bacteriana resistente a dicho ácido nucleico diana o producto de transcripción del mismo.

En otro aspecto se ha proporcionado una célula obtenida u obtenible mediante el(los) método(s) descrito(s) en la presente memoria.

En un aspecto adicional se ha proporcionado el uso de una célula para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.

- 5 En otro aspecto se ha proporcionado un cultivo celular que comprende una célula para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.

En un aspecto adicional se ha proporcionado un producto alimenticio o comida que comprende el cultivo descrito en la presente memoria.

- 10 En otro aspecto se ha proporcionado un proceso para preparar un producto alimenticio o comida que comprende el uso del cultivo descrito en la presente memoria.

En un aspecto adicional se ha proporcionado el uso del cultivo descrito en la presente memoria para preparar un producto alimenticio.

Realización preferida

- 15 En algunas realizaciones, los uno o más genes *cas* o proteínas se usan en combinación con dos o más repeticiones CRISPR.

En algunas realizaciones, los uno o más genes *cas* o proteínas y/o las dos o más repeticiones son o son derivables (preferiblemente, derivadas) a partir de la misma célula.

En algunas realizaciones, los uno o más genes *cas* o proteínas y las dos o más repeticiones CRISPR se dan conjuntamente de manera natural en la misma célula.

- 20 En algunas realizaciones, los uno o más genes *cas* o proteínas se usan en combinación con uno o más espaciadores CRISPR.

En algunas realizaciones, el(los) espaciador(es) CRISPR está(n) o es(son) derivable(s) (preferiblemente, derivado(s)) a partir de un organismo que es diferente a la célula a partir de la cual los uno o más genes *cas* o proteínas y/o las dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivadas).

- 25 En algunas realizaciones, el espaciador se obtiene a partir de una célula que es resistente a un ácido nucleico diana.

En algunas realizaciones, el espaciador CRISPR es una secuencia sintética de ácido nucleico.

En algunas realizaciones, el(los) espaciador(es) tiene(n) homología al ácido nucleico diana.

En algunas realizaciones, el(los) espaciador(es) tiene(n) 100% de identidad al ácido nucleico diana sobre al menos la longitud de la zona central del espaciador CRISPR.

- 30 En algunas realizaciones, los uno o más genes *cas* o proteínas se usan en combinación con al menos uno o más espaciadores CRISPR y al menos dos o más repeticiones CRISPR.

En algunas realizaciones, el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de ADN bacteriófago.

- 35 En algunas realizaciones, el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de ADN plásmido.

En algunas realizaciones, el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de un elemento genético móvil.

En algunas realizaciones, el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de un elemento transponible o una secuencia de inserción.

- 40 En algunas realizaciones, el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de un gen de resistencia a antibiótico.

En algunas realizaciones, el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de un ácido nucleico que codifica un factor de virulencia.

- 45 En algunas realizaciones, el factor de virulencia se selecciona entre el grupo que consiste en un ácido nucleico que codifica una toxina, una intemalina y una hemolisina.

En algunas realizaciones, los uno o más genes *cas* y las dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivados) a partir de la misma célula.

En algunas realizaciones, los uno o más genes *cas* y las dos o más repeticiones CRISPR se dan de conjuntamente de manera natural en la misma célula.

- 5 En algunas realizaciones, los espaciadores CRISPR son o son derivables (preferiblemente, derivados) a partir de un organismo que es diferente a la célula a partir de la cual los uno o más genes *cas* y/o las dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivados).

En algunas realizaciones, la célula es una célula receptora o una célula hospedante.

- 10 En algunas realizaciones, los uno o más genes *cas* o proteínas y/o las dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivados) a partir de la misma célula.

En algunas realizaciones, los espaciadores son o son derivables (preferiblemente, derivados) a partir de un organismo que es diferente a la célula que comprende los uno o más genes *cas* o proteínas y las dos o más repeticiones CRISPR.

- 15 En algunas realizaciones, los uno o más genes *cas* o proteínas y las dos o más repeticiones CRISPR se dan conjuntamente de manera natural en la misma célula.

En algunas realizaciones, dicha modificación comprende la inserción de uno o más espaciadores CRISPR y/o seudo espaciadores CRISPR en la célula.

En algunas realizaciones, el espaciador de la célula tiene 100% de homología al espaciador CRISPR o seudo espaciador CRISPR del organismo.

- 20 En algunas realizaciones, dicha modificación comprende modificar por ingeniería genética el espaciador CRISPR de la célula.

En algunas realizaciones, se modifica todo o parte del espaciador en la célula.

En algunas realizaciones, dicha modificación comprende la modificación de un espaciador recombinante.

En algunas realizaciones, dicha modificación se da a través de mutación o mutagénesis espontánea.

- 25 En algunas realizaciones, se someten a delección los al menos uno o más espaciador(es) CRISPR en la célula.

En algunas realizaciones, se someten a delección los al menos uno o más repetición(es) CRISPR en la célula.

En algunas realizaciones, se someten a delección uno o más genes *cas*.

En algunas realizaciones, se someten a delección CRISPR y/o uno o más genes *cas*.

- 30 En algunas realizaciones, se someten a delección los uno o más genes *cas* o proteínas y/o dos o más repeticiones CRISPR en la célula.

En algunas realizaciones, las secuencias de nucleótidos del gen *cas* y la repetición CRISPR son o son derivables (preferiblemente, derivados) a partir de las mismas o diferentes cepas.

En algunas realizaciones, las secuencias de nucleótidos del gen *cas* y la repetición CRISPR son o son derivables (preferiblemente, derivados) a partir de las mismas o diferentes especies.

- 35 En algunas realizaciones, las secuencias de nucleótidos del gen *cas* y la repetición CRISPR son o son derivables (preferiblemente, derivados) a partir de los mismos o diferentes géneros

En algunas realizaciones, las secuencias de nucleótidos del gen *cas* y la repetición CRISPR son o son derivables (preferiblemente, derivados) a partir de los mismos o diferentes organismos.

- 40 En algunas realizaciones, el ácido nucleico diana en el bacteriófago es una secuencia altamente conservada de ácido nucleico.

En algunas realizaciones, el ácido nucleico diana en el bacteriófago codifica una proteína de especificidad a hospedante.

En algunas realizaciones, el ácido nucleico diana en el bacteriófago codifica una proteína que es esencial para la supervivencia, la replicación o el crecimiento del bacteriófago.

En algunas realizaciones, el ácido nucleico diana en los bacteriófagos codifica una helicasa, una primasa, una proteína estructural de cabeza o cola, una proteína con un dominio conservado (por ejemplo, holin, lisina y otros) o unas secuencias conservadas entre genes importantes de fago.

5 En algunas realizaciones, el método para determinar la resistencia de una célula a un ácido nucleico diana o un producto de transcripción del mismo comprende la etapa adicional de comparar la secuencia de los uno o más espaciadores CRISPR en la célula con la secuencia del ácido nucleico diana.

En algunas realizaciones, el método para determinar la resistencia de una célula a un ácido nucleico diana o un producto de transcripción del mismo comprende la etapa adicional de determinar el perfil de resistencia de la célula.

En algunas realizaciones, dicho cultivo es un cultivo iniciador o un cultivo probiótico.

10 Figuras

Figura 1

Representación esquemática de CRISPR de *S. thermophilus* CNRZ1066 (42 repeticiones).

Figura 2

15 Análisis de gráfico de puntos de secuencias de la proteína Cas (A) y secuencias del locus CRISPR (B). Los nombres del organismo (género, especie, cepa) están indicados en el lado derecho de cada gráfico de puntos (por ejemplo Sth_LMG18311= cepa LMG18311 de *S. thermophilus*).

Figura 3

20 Las secuencias espaciadoras del locus CRISPR de *S. thermophilus* CNRZ1066 se confrontaron (búsquedas de la secuencia corta casi exacta usando BlastN en la página de internet NCBI) frente a la base de datos de virus, y se alinean con las parejas posteriores en bacteriófagos *S. thermophilus*. (A) La tabla indica las secuencias espaciadoras de CNRZ 1066 CRISPR que presenta identidades significantes de secuencia con las secuencias del fago (células oscuras). (B) Alineamiento de la secuencia de secuencia interespaciadora nº29 con ocho secuencias de fago. (Observación: el espaciador nº20 muestra similitud con un número de proteínas de especificidad a hospedante).

25 Figura 4

Estructura secundaria en horquilla ("stem-loop") putativa de una secuencia de repetición CRIS1PR de *S. thermophilus*. Únicamente se muestra una cadena de ADN.

Figura 5

30 La integración de un espaciador CRISPR en el locus CRISPR de *Streptococcus thermophilus* proporciona resistencia frente a un bacteriófago al que el espaciador CRISPER muestra identidad. DGCC7710 parental es sensible a fago, y BIM DCBCC7710RH1 es resistente a fago. BIM DGCC7710RH1 tiene un nuevo espaciador (Sn) en el locus CRISPR, el cual muestra 100% de identidad a la secuencia del fago. En la etapa (b) la cepa se enfrenta al fago 858 y se selecciona un mutante resistente a fago. En la etapa (c) el locus CRISPR I del mutante tiene un espaciador adicional que comparte 100% de identidad con la región 31921-31950 pb del fago.

35 Figura 6

40 La integración de un espaciador CRISPR en el locus CRISPR de *Streptococcus thermophilus* proporciona resistencia frente a un bacteriófago al que el espaciador CRISPR muestra identidad. DGCC7710 parental es sensible a fago, y BIM DGCC7710RH2 es resistente a fago. BIM DGCC7710HR2 tiene un nuevo espaciador (Sn) en el locus CRISPR, el cual muestra 100% de identidad a la secuencia del fago. En la etapa (b) la cepa se enfrenta al fago 858 y se selecciona un mutante resistente a fago. En la etapa (c) el experimento se repitió independientemente y se seleccionó otro mutante. El locus CRISPR I del mutante tiene un espaciador adicional (diferente al del en RH1) que comparte 100% de identidad con la región 17125-17244 pb del fago.

Figura 7

45 La ordenación de espaciador de CRISPR I en diversas cepas de *Streptococcus thermophilus*. Los números indican la posición del espaciador. Los nombres de la cepa están enumerados a la izquierda. Las letras indican el tipo de espaciador CRISPR, con espaciadores idénticos descritos con un código de 2 letras similares. Los espaciadores con polimorfismos de nucleótido sencillo se etiquetan con combinación de letra idéntica, complementado con una etiqueta "prime". Los espaciadores únicos no están descritos mediante una combinación de letra, y se dejan en blanco.

Figura 8

Homología de espaciadores CRISPR con secuencias conocidas, incluyendo secuencias cromosómicas bacterianas (sombreado en gris), secuencias de ADN plásmido (cuadrados negros) y secuencias genómicas de fago (sombreado en negro).

5 **Figura 9**

Una representación gráfica del sistema de plásmido usado para producir por ingeniería genética un número de construcciones en *Streptococcus thermophilus* tal como se describe por Russell, M. W., y T.R. Klaenhammer (2001) "Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasserii* chromosomes via homologous recombination. "Applied and Environmental Microbiology" 67:4.361-4.364.

10 **Figura 10**

Una representación gráfica del plásmido usado para productos de PCR de subclón de las diversas construcciones descritas en la presente memoria (casKO, cas4 KO, RT y S1S2). El plásmido está comercialmente disponible por Invitrogen en el kit TOPOTA Cloning®.

Figura 11

15 Una representación gráfica del plásmido usado para la recombinación homóloga en una realización de la presente invención.

Figura 12

Una representación gráfica que ilustra la preparación de la construcción CAS1KO en la cual el gen *cas1* está alterado por recombinación homóloga.

20 **Figura 13**

Una representación gráfica que ilustra la preparación de la construcción CAS4KO en la cual el gen *cas4* está alterado por recombinación homóloga.

Figura 14

25 Una representación gráfica que ilustra la modificación por ingeniería de la construcción S1S2 usando cebadores ("primers") específicos y reacciones de PCR iterativas. El primer panel ilustra todos los cebadores usados y la disposición para las dos primeras reacciones de PCR (reacción nº1 con cebadores P1 y P2 y la reacción nº2 con cebadores P2 y P3). El segundo panel muestra los productos de PCR obtenidos a partir de las dos primeras reacciones de PCR, con el producto de la reacción nº1 a la izquierda y el producto de la reacción nº2 a la derecha. El tercer panel muestra la tercera reacción de PCR, usando una combinación de los productos de las dos primeras
30 PCR como la plantilla para la tercera reacción de PCR, y el primer P1 de la primera reacción junto con cebador P4 de la segunda reacción. El cuarto panel muestra el producto de la PCR nº3, la cual técnicamente genera la construcción S1S2.

Figura 15

35 Una representación gráfica de los detalles para el diseño de cebador para los cebadores 2 y 3, los cuales contienen las secuencias clave para el experimento, derivadas de los espaciadores idénticos a las secuencias de fago (los productos de PCR derivados a partir de estos cebadores de PCR generarán los espaciadores que proporcionarán por último resistencia a los fagos).

Figura 16

Una representación gráfica de la integración de la construcción S1S2.

40 **Figura 17**

Una representación gráfica de la preparación de la construcción RT usando una enzima de restricción para generar la construcción RT a partir de la construcción S1S2. Hay sitios de restricción *BglII* dentro de las repeticiones que permiten que se corte la parte "media" de la construcción. Después de la digestión enzimática, se usa una ligasa para juntar las dos piezas finales, generando así una nueva construcción que tiene RT, pero no espaciadores.

45 **Figura 18**

Una representación gráfica de la integración de la construcción RT.

Figura 19

Una representación gráfica de la construcción RT'.

Figura 20

Una representación gráfica de la construcción RT'.

5 Descripción detallada de la invención**Locus CRISPR**

Los loci CRISPR son una clase distinta de repeticiones de secuencia cortas intercaladas (SSRs, del Inglés "Short Sequence Repeats") que son reconocidas por primera vez en *E. coli* (Ishino et al. (1987) "J. Bacteriol". 169:5.429-5.433; Nakata et al. (1989) "J. Bacteriol". 171:3.553-3.556). Se han identificado SSRs intercalados similares en *Halobacterium salinarum*, *Streptococcus pyogenes*, *Anabaena* y *Mycobacterium tuberculosis* (Groenen et al. (1993) "Mol. Microbiol". 10:1.057-1.065; Hoe et al. (1999) "Emerg. Infect. Dis". 5:254-263; Masepohl et al. (1996) "Biochim. Biophys. Acta" 1.307:26-30; Mojica et al. (1995) "Mol. Microbiol". 17:85-93). Los loci CRISPR difieren de las otras SSRs por la estructura de las repeticiones, las cuales se han calificado repeticiones cortas regularmente espaciadas (SRSRs, del Inglés "Short Regularly Spaced Repeats")(Janssen et al. (2002) OMICS "J. Integ. Biol". 6:23-33; Mojica et al. (2000) "Mol. Microbiol". 36:244-246). Las repeticiones son elementos cortos que se dan en grupos, que están siempre regularmente espaciados por secuencias intermedias únicas con una longitud constante (Mojica et al. (2000) "Mol. Microbiol". 36:244-246). Aunque las secuencias de repetición están altamente conservadas entre las cepas, el número de repeticiones intercaladas y las secuencias de las regiones espaciador difieren entre cepa y cepa (van Embden et al. (2000) "J. Bacteriol". 182:2.393-2.401).

Las características estructurales comunes de los loci CRISPR están descritas en Jansen et al. (2002) como (i) la presencia de múltiples repeticiones cortas y directas, las cuales muestran ninguna o muy pequeña variación de secuencia dentro de un locus dado; (ii) la presencia de secuencias espaciadoras no repetitivas entre las repeticiones de tamaño similar; (iii) la presencia de una secuencia líder ("leader") común de unos pocos cientos de bases de pares en la mayoría de especies que abrigan los loci CRISPR múltiples; (iv) la ausencia de largas pautas de lectura abierta dentro del locus; y (v) la presencia de uno o más genes *cas*.

Las CRISPRs generalmente son secuencias cortas parcialmente palindrómicas de 24-40 pb que contienen repeticiones invertidas internas y terminales de hasta 11 pb. Aunque se han detectado elementos aislados, generalmente están ordenados en grupos (hasta aproximadamente 20 o más por genoma) de unidades repetidas espaciadas por secuencias de 20-58 pb intermedias únicas. Las CRISPRs generalmente son homogéneas dentro de un genoma dado con la mayoría de ellos siendo idénticos. Sin embargo, hay ejemplos de heterogeneidad en, por ejemplo, las Arqueas (Mojica et al. 2000).

A modo de ejemplo, el genoma de *Streptococcus thermophilus* LMG18311 contiene 3 loci CRISPR; las secuencias repetidas de 36 pb son diferentes en CRISPR1 (34 repeticiones), CRISPR2 (5 repeticiones) y CRISPR3 (una secuencia sencilla). No obstante, están perfectamente conservadas dentro de cada locus. Las repeticiones CRISPR1 y CRISPR2 están respectivamente interespaciadas por 33 y 4 secuencias de 30 pb de longitud. Todas estas secuencias interespaciadoras son diferentes una de otras. También son diferentes de las encontradas en la cepa CNRZ1066 (41 secuencias interespaciadoras dentro de CRISPR1) y en la cepa LMD-9 (16 dentro de CRISPR1 y 8 dentro de CRISPR3), las cuales ambas son *S. thermophilus*. La Figura 1 describe una de las CRISPRs identificadas en *S. thermophilus*.

Diversos métodos para identificar los loci CRISPR son ya conocidos en la técnica. A modo de ejemplo, Jensen et al. (2002) describe un enfoque basado en la informática en el cual se buscan los motivos CRISPR en las secuencias de nucleótidos usando el programa PATSCAN en el servidor del "Mathematics and Computer Science Division at the Argonne National Laboratory", Argonne, IL, USA. El algoritmo que se usó para identificar los motivos CRISPR fueron $p1=a...b\ c...d\ p1\ c...d\ p1\ c...d\ p1$, en el que a y b son el límite inferior y superior de tamaño de la repetición y p1 y c y d son el límite inferior y superior de tamaño de las secuencias espaciadora. Los valores de a, b, c y d se pueden variar a partir de aproximadamente 15 a aproximadamente 70 pb en incrementos de aproximadamente 5 pb.

Los loci CRISPR se pueden identificar usando gráficos de puntos (usando, por ejemplo, un programa informático denominado Dotter).

El análisis de similitud de secuencia se puede realizar usando diversos métodos que son bien conocidos en la técnica. A modo de ejemplo, el análisis se pueden realizar usando NCBI BLAST con una base de datos de genomas microbianos (<http://www.ncbi.nlm.nih.gov>) y GenBank.

La amplificación de los loci CRISPR ha sido descrita en, por ejemplo, Mojica et al. (2005) y Pourcel et al. (2005). La amplificación de la región deseada de ADN se puede alcanzar mediante cualquier método conocido en la técnica,

incluyendo la reacción en cadena de la polimerasa (PCR, del Inglés “Polymerase Chain Reaction”). Por “amplificación” queremos decir la producción de copias adicionales de una secuencia de ácido nucleico. Esto generalmente se lleva a cabo usando tecnologías de PCR bien conocidas en la técnica (Dieffenbach y Dveksler (1995) “PCR Primer, a Laboratory Manual” (Cold Spring Harbor Press, Plainview, New York).

5 Por “reacción en cadena de la polimerasa” o “PCR” queremos decir un método tal como el descrito en los documentos US4.683.195 y US4.683.202, los cuales describen un método para incrementar la concentración de un segmento de una secuencia diana en una mezcla de un ADN genómico sin clonación o purificación. La longitud del segmento amplificado de la secuencia diana deseada se determina mediante las posiciones relativas de dos cebadores de oligonucleótido con respecto el uno al otro, y por lo tanto, esta longitud es un parámetro controlable. 10 En virtud del aspecto de repetición del proceso, el método se refiere a “PCR”. Debido a los segmentos amplificados deseados de la secuencia diana llegan a ser las secuencias predominantes (en términos de concentración) en la mezcla, se dice que es “PCR amplificada”.

En el enfoque de PCR, se pueden diseñar los cebadores de oligonucleótido para usarse en las reacciones PCR para amplificar todo o parte de un locus CRISPR. Por “cebador” queremos decir un oligonucleótido, si se da de manera natural como en una digestión de restricción purificada o sintéticamente producido, el cual es capaz de actuar como un punto de inicio de síntesis cuando se coloca bajo condiciones en las cuales se induce la síntesis de un producto de extensión de cebador que es complementario a una cadena de ácido nucleico (es decir, en presencia de nucleótidos y un agente de inducción, tal como ADN polimerasa y a una temperatura y pH adecuados). El cebador puede ser de cadena sencilla para máxima eficacia en la amplificación, pero alternativamente puede ser de doble 15 cadena. Si es de doble cadena, el cebador se trata primero para separar sus cadenas antes de usarse para preparar productos de extensión. El cebador puede ser un oligodeoxiribonucleótido. El cebador debe ser suficientemente largo para cebar la síntesis de los productos de extensión en presencia del agente de inducción. Las longitudes exactas de los cebadores dependerán de muchos factores, incluyendo la temperatura, la fuente de cebador y el uso del método. Los cebadores de PCR generalmente son al menos aproximadamente 10 nucleótidos de longitud, y la mayoría generalmente al menos aproximadamente 20 nucleótidos de longitud. 20 25

Los métodos para diseñar los cebadores de PCR y la clonación de PCR generalmente se conocen en la técnica y están descritos en Sambrook et al. (1989) “Molecular Cloning: A Laboratory Manual” (2ª Edición, Cold Spring Harbor Laboratory Press, Plainview, New York). Véase también Innis et al. eds. (1990) “PCR Protocols: A Guide to Methods and Applications” (Academic Press, New York); Innis y Gelfand, eds. (1995) “PCR Strategies” (Academic Press, New York); y Innis y Gelfand, eds. (1999) “PCR Methods Manual” (Academic Press, New York). Métodos conocidos de PCR incluyen, pero no se limitan a, métodos que usan cebadores emparejados, cebadores anidados, cebadores específicos sencillos, cebadores degenerados, cebadores específicos a gen, cebadores específicos a vector, cebadores parcialmente desemparejados y similares. 30

Los loci CRISPR pueden comprender, consistir o consistir esencialmente en ADN o ARN de origen genómico, sintético o recombinante. 35

Los loci CRISPR pueden ser de doble cadena o de cadena sencilla si representan la cadena sentido o antisentido o combinaciones de las mismas.

Los loci CRISPR se pueden preparar mediante el uso de técnicas de ADN recombinante (por ejemplo, ADN recombinante), tal como se describe en la presente memoria.

40 Las secuencias de nucleótidos descritas en la presente memoria se pueden obtener a partir de bases de datos, tales como GenBank o la página de internet de JGI en http://genome.jgi-psf.org/mic_homc.html.

Orientación del CRISPER

Para evitar la duda, en el contexto de la presente invención el locus CRISPR está orientado tal como sigue.

45 El líder CRISPR es un segmento de ADN conservado de tamaño definido. Por ejemplo, la secuencia líder de CRISPR1 de *S. thermophilus* es el segmento de ADN que comienza inmediatamente después del codon de parada del gen str0660 y que finaliza justo antes de la primera repetición. El líder CRISPR está localizado en el extremo 5' del locus CRISPR. El líder de CRISPER está localizado inmediatamente hacia arriba (“upstream”) de la primera repetición CRISPR del locus CRISPR.

50 El tráiler CRISPR es un segmento de ADN conservado de tamaño definido. Por ejemplo, la secuencia tráiler del CRISPR1 de *S. thermophilus* es el segmento de ADN que comienza inmediatamente después de la repetición terminal, y que finaliza justo antes del codon de parada del gen str0661 (localizado en la cadena opuesta de ADN). El tráiler CRISPR está localizado en el extremo 3' del locus CRISPR. El tráiler CRISPR está localizado inmediatamente hacia abajo (“downstream”) de la repetición terminal.

A modo de ejemplo, las secuencias del líder CRISPR y el tráiler CRISPR en el locus CRISPR1 de la cepa CNRZ1066 de *Streptococcus thermophilus* son:

Líder CRISPR

5'-CAAGGACAGTTATTGATTTTATAATCACTATGTGGGTATAAAAAACGTCAAATTTCAATTTGAG-3'

5 Trailer CRISPR

5'-TTGATTCAACATAAAAAAGCCAGTTCAATTGAACTTGGCTTT-3'

El líder CRISPR corresponde a las posiciones 625038 a 625100 y el tráiler CRISPR corresponde a las posiciones 627845 a 627885 en el genoma completo (CP000024) de *Streptococcus thermophilus*.

Para evitar la duda "upstream" quiere decir en la dirección 5' y "downstream" quiere decir en la dirección 3'.

10 **Cas**

Tal como se usa en la presente memoria, el término "gen cas" tiene el significado convencional como el usado en la técnica y se refiere a uno o más genes cas que generalmente están acoplados, asociados o cerca de o en las inmediaciones de locis CRISPR flanqueantes.

15 Una exhaustiva revisión de la familia de la proteína Cas está presentada en Haft et al. (2005) "Computational Biology 1, 6 e60". Tal como se describe en la misma, se describen 41 familias del gen asociado a CRISPR (cas), además de las cuatro familias de gen previamente conocidas. Se muestra que los sistemas de CRISPR pertenecen a diferentes clases, con diferentes patrones de replicación, conjuntos de genes e intervalos de especies.

El número de genes cas en un locus CRISPR dado puede variar entre especies.

20 En un aspecto, la presente invención se refiere al uso de uno o más genes cas o proteínas para modular la resistencia en una célula (por ejemplo, una célula receptora) frente a un ácido nucleico diana o un producto de transcripción del mismo.

En un aspecto adicional, la presente invención se refiere al uso de uno o más genes cas o proteínas y uno o más espaciadores CRISPR para modular la resistencia en una célula (por ejemplo, una célula receptora) frente a un ácido nucleico diana o un producto de transcripción del mismo.

25 En algunas realizaciones, uno o más de los genes cas y/o proteínas se pueden dar de manera natural en una célula receptora y uno o más espaciadores heterólogos pueden estar integrados o insertados adyacentes a los uno o más genes cas o proteínas.

30 En algunas realizaciones, uno o más de los genes cas y/o proteínas adecuados pueden ser heterólogos a la célula receptora y uno o más de los espaciadores pueden ser homólogos o heterólogos. En este caso, los espaciadores pueden estar integrados o insertados adyacentes al uno o más de los gen cas o proteínas.

En un aspecto, la presente invención se refiere al uso de uno o más genes cas o proteínas y al menos dos repeticiones CRISPR para modular la resistencia en una célula (por ejemplo, una célula receptora) frente a un ácido nucleico diana o un producto de transcripción del mismo.

35 En un aspecto, la presente invención se refiere al uso de uno o más genes cas o proteínas, al menos dos repeticiones CRISPR y al menos un espaciador CRISPR para modular la resistencia en una célula (por ejemplo una célula receptora) frente a un ácido nucleico diana o un producto de transcripción del mismo.

40 Las estructuras CRISPR generalmente se encuentran en las inmediaciones de cuatro genes denominados cas1 a cas4. La ordenación más común de estos genes es cas3-cas4-cas1-cas2. La proteína Cas3 parece ser una helicasa, mientras que Cas4 se parece a la familia RecB de exonucleasas y contiene un motivo rico en cisteína, sugerente de la unión a ADN. Cas1 generalmente es altamente básico y es la única proteína Cas encontrada sistemáticamente en todas las especies que contienen loci CRISPR. Cas2 queda por ser caracterizado. Cas1-4 generalmente se caracterizan por su cercana proximidad a los loci CRISPR y su amplia distribución a través de especies bacterianas y de arqueas. Aunque no todos los genes cas1-4 se asocian con todos los loci CRISPR, todos se encuentran en múltiples subtipos.

45 Bolotin et al. (2005) recientemente ha informado de otro grupo de tres genes asociados con estructuras CRISPR en muchas especies bacterianas, denominados aquí como cas1B, cas5 y cas6.

El gen cas puede ser cas1, cas2, cas3, cas4, cas1B, cas5 y/o cas6. En una realización, el gen cas es cas1.

ES 2 398 918 T3

El gen *cas* puede ser *cas1*, *cas2*, *cas3*, *cas4*, *cas1B*, *cas5* y/o *cas6* o un fragmento, variante, homólogo o derivado del mismo.

5 Los genes *cas* pueden ser *cas1*, *cas2*, *cas3*, *cas4*, *cas1B*, *cas5* y/o *cas6* o una pluralidad de los mismos o una combinación de los mismos, tal como *cas1* y *cas2*; *cas1* y *cas3*; *cas1* y *cas4*; *cas1* y *cas1B*; *cas1* y *cas5*; *cas1* y *cas6*; *cas2* y *cas3*; *cas2* y *cas4*; *cas2* y *cas1B*; *cas2* y *cas5*; *cas2* y *cas6*; *cas3* y *cas4*; *cas3* y *cas1B*; *cas3* y *cas5*; *cas3* y *cas6*; *cas4* y *cas1B*; *cas4* y *cas5*; *cas4* y *cas6*; *cas1B* y *cas5*; *cas1B* y *cas6*; *cas1*, *cas2* y *cas3*; *cas1*, *cas2* y *cas4*; *cas1*, *cas2* y *cas1B*; *cas1*, *cas2*, *cas3* y *cas4*; *cas1*, *cas2*, *cas3* y *cas1B*; *cas1*, *cas2*, *cas3* y *cas5*; *cas1*, *cas2*, *cas3* y *cas6*; *cas1*, *cas2*, *cas3*, *cas4* y *cas1B*; *cas1*, *cas2*, *cas3*, *cas4* y *cas5*; *cas1*, *cas2*, *cas3*, *cas4*, *cas1B* y *cas6*; *cas1*, *cas2*, *cas3*, *cas4*, *cas1B*, *cas5*; *cas1*, *cas2*, *cas3*, *cas4*, *cas1B* y *cas6*; *cas1*, *cas2*, *cas3*, *cas4*, *cas1B*, *cas5* y *cas6*; *cas2*, *cas3* y *cas4*; *cas2*, *cas3* y *cas1B*; *cas2*, *cas3* y *cas5*; *cas2*, *cas3* y *cas6*; *cas2*, *cas3*, *cas4* y *cas1B*; *cas2*, *cas3*, *cas4* y *cas5*; *cas2*, *cas3*, *cas4* y *cas6*; *cas2*, *cas3*, *cas4*, *cas1B* y *cas5*; *cas2*, *cas3*, *cas4*, *cas1B* y *cas6*; *cas2*, *cas3*, *cas4*, *cas1B*, *cas5* y *cas6*; *cas3*, *cas4*, *cas1B* y *cas6*; *cas3*, *cas4*, *cas1B*, *cas5* y *cas6*; *cas3*, *cas4*, *cas1B*, *cas5* y *cas6*; *cas4*, *cas1B* y *cas5*; *cas4*, *cas1B* y *cas6*; *cas4*, *cas1B*, *cas5* y *cas6*; *cas5* y *cas6* o combinaciones de los mismos.

15 Los genes *cas* pueden ser *cas1* y *cas2*; *cas1* y *cas3*; *cas1* y *cas4*; *cas1* y *cas1B*; *cas1* y *cas5*; *cas1* y *cas6*; *cas2* y *cas3*; *cas2* y *cas4*; *cas2* y *cas1B*; *cas2* y *cas5*; *cas2* y *cas6*; *cas3* y *cas4*; *cas3* y *cas1B*; *cas3* y *cas5*; *cas3* y *cas6*; *cas4* y *cas1B*; *cas4* y *cas5*; *cas4* y *cas6*; *cas1B* y *cas5* o *cas1B* y *cas6* o combinaciones de los mismos.

20 Los genes *cas* pueden ser un *cas1*, *cas2* y *cas3*; *cas1*, *cas2* y *cas4*; *cas1*, *cas2* y *cas1B*; *cas1*, *cas2*, *cas3* y *cas4*; *cas1*, *cas2*, *cas3* y *cas1B*; *cas1*, *cas2*, *cas3* y *cas5*; *cas1*, *cas2*, *cas3* y *cas6*; *cas1*, *cas2*, *cas3*, *cas4* y *cas1B*; *cas1*, *cas2*, *cas3*, *cas4* y *cas5*; *cas1*, *cas2*, *cas3*, *cas4*, *cas1B* y *cas6*; *cas1*, *cas2*, *cas3*, *cas4*, *cas1B* y *cas5*; *cas1*, *cas2*, *cas3*, *cas4*, *cas1B* y *cas6*; *cas1*, *cas2*, *cas3*, *cas4*, *cas1B*, *cas5* y *cas6* o combinaciones de los mismos.

Los genes pueden ser *cas2*, *cas3* y *cas4*; *cas2*, *cas3* y *cas1B*; *cas2*, *cas3* y *cas5*; *cas2*, *cas3* y *cas6*; *cas2*, *cas3*, *cas4* y *cas1B*; *cas2*, *cas3*, *cas4* y *cas5*; *cas2*, *cas3*, *cas4* y *cas6*; *cas2*, *cas3*, *cas4*, *cas1B* y *cas5*; *cas2*, *cas3*, *cas4*, *cas1B* y *cas6*; *cas2*, *cas3*, *cas4*, *cas1B*, *cas5* y *cas6* o combinaciones de los mismos.

25 Los genes *cas* pueden ser *cas3*, *cas4* y *cas1B*; *cas3*, *cas4* y *cas5*; *cas3*, *cas4* y *cas6*; *cas3*, *cas4*, *cas1B* y *cas5*; *cas3*, *cas4*, *cas1B* y *cas6*; *cas3*, *cas4*, *cas1B*, *cas5* y *cas6*; *cas4*, *cas1B* y *cas5*; *cas4*, *cas1B* y *cas6*; *cas4*, *cas1B*, *cas5* y *cas6*; *cas5* y *cas6* o combinaciones de los mismos.

30 El gen *cas* puede ser uno o más de *cas1*, *cas2*, *cas3*, *cas1B*, *cas5* y/o *cas6* o una pluralidad de los mismos, tal como una pluralidad de cualquiera de 2 genes *cas*, cualquiera de 3 genes *cas*, cualquiera de 4 genes *cas*, cualquiera de 5 genes *cas*, cualquiera de 6 genes *cas*, o cualquiera de 7 genes *cas*.

La pluralidad de genes *cas* puede comprender, consistir o consistir esencialmente en una pluralidad de los mismos genes *cas*, tal como 2 genes *cas*, 3 genes *cas*, 4 genes *cas*, 5 genes *cas*, 6 genes *cas*, 7 genes *cas*, 8 genes *cas*, 9 genes *cas*, 10 genes *cas*, 15 genes *cas*, 20 genes *cas*, 25 genes *cas*, 30 genes *cas*, 35 genes *cas*, 40 genes *cas* o incluso 50 o más genes *cas*.

35 La pluralidad de genes *cas* puede comprender, consistir o consistir esencialmente en una pluralidad de diferentes genes *cas*, tal como 2 genes *cas* diferentes, 3 genes *cas* diferentes, 4 genes *cas* diferentes, 5 genes *cas* diferentes, 6 genes *cas* diferentes, 7 genes *cas* diferentes, 8 genes *cas* diferentes, 9 genes *cas* diferentes, 10 genes *cas* diferentes, 15 genes *cas* diferentes, 20 genes *cas* diferentes, 25 genes *cas* diferentes, 30 genes *cas* diferentes, 35 genes *cas* diferentes, 40 genes *cas* diferentes o incluso 50 o más genes *cas* diferentes.

40 La pluralidad de genes *cas* puede comprender, consistir o consistir esencialmente en una pluralidad de los mismos y/o diferentes genes *cas*, tal como 2 genes *cas* diferentes, 3 genes *cas* diferentes, 4 genes *cas* diferentes, 5 genes *cas* diferentes, 6 genes *cas* diferentes, 7 genes *cas* diferentes, 8 genes *cas* diferentes, 9 genes *cas* diferentes, 10 genes *cas* diferentes, 15 genes *cas* diferentes, 20 genes *cas* diferentes, 25 genes *cas* diferentes, 30 genes *cas* diferentes, 35 genes *cas* diferentes, 40 genes *cas* diferentes o incluso 50 o más genes *cas* diferentes. El mismo gen *cas* puede duplicarse una pluralidad de veces.

45 Adecuadamente, el término "gen *cas*" se refiere a una pluralidad de genes *cas*, tal como entre 2 y 12 genes *cas*, más preferiblemente, entre 3 y 11 genes *cas*, más preferiblemente, entre 4 y 10 genes *cas*, más preferiblemente, entre 4 y 9 genes *cas*, más preferiblemente, entre 4 y 8 genes *cas*, más preferiblemente, entre 4 y 7 genes *cas*, tales como 4, 5, 6 ó 7 genes *cas*.

50 El(los) gen(es) *cas* puede(n) comprender, consistir o consistir esencialmente en ADN o ARN de origen genómico, sintético o recombinante.

El(los) gen(es) puede(n) ser de doble cadena o cadena sencilla si se representa la cadena sentido o antisentido o combinaciones de las mismas.

El(los) gen(es) se puede(n) preparar mediante el uso de técnicas de ADN recombinante (por ejemplo, ADN recombinante), tal como se describe en la presente memoria.

5 Tal como se describe a continuación en la presente memoria, el gen *cas* puede ser un fragmento de gen *cas*, indicando de ese modo que el gen *cas* comprende una fracción de una secuencia tipo natural. Adecuadamente, la secuencia comprende al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 65%, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98 o al menos 99% de la secuencia tipo natural.

Para algunas realizaciones se prefiere que el gen *cas* sea el gen *cas* que está el más cercano de la secuencia líder o de la primera repetición CRISPR en el extremo 5' del locus CRISPR, tal como *cas4* o *cas6*.

10 La proteína Cas puede ser Cas1, Cas2, Cas3, Cas4, Cas1B, Cas5 y/o Cas6.

La proteína Cas puede ser Cas1, Cas2, Cas3, Cas4, Cas1B, Cas5 y/o Cas6 o un fragmento, variante, homólogo o derivado de las mismas.

15 La proteína Cas puede ser Cas1, Cas2, Cas3, Cas4, Cas1B, Cas5 y/o Cas6 o una combinación de las mismas, tal como Cas1 y Cas2; Cas1 y Cas3; Cas1 y Cas4; Cas1 y Cas1B; Cas1 y Cas5; Cas1 y Cas6; Cas2 y Cas3; Cas2 y Cas4; Cas2 y Cas1B; Cas2 y Cas5; Cas2 y Cas6; Cas3 y Cas4; Cas3 y Cas1B; Cas3 y Cas5; Cas3 y Cas6; Cas4 y Cas1B; Cas4 y Cas5; Cas4 y Cas6; Cas1B y Cas5; Cas1B y Cas6; Cas1, Cas2 y Cas3; Cas1, Cas2 y Cas4; Cas1, Cas2 y Cas1B; Cas1, Cas2, Cas3 y Cas4; Cas1, Cas2, Cas3 y Cas1B; Cas1, Cas2, Cas3 y Cas5; Cas1, Cas2, Cas3 y Cas6; Cas1, Cas2, Cas3, Cas4 y Cas1B; Cas1, Cas2, Cas3, Cas4 y Cas5; Cas1, Cas2, Cas3, Cas4, Cas1B y Cas6; Cas1, Cas2, Cas3, Cas4, Cas1B y Cas5; Cas1, Cas2, Cas3, Cas4, Cas1B y Cas6; Cas1, Cas2, Cas3, Cas4, Cas1B, Cas5 y Cas6; Cas2, Cas3 y Cas4; Cas2, Cas3 y Cas1B; Cas2, Cas3 y Cas5; Cas2, Cas3 y Cas6; Cas2, Cas3, Cas4 y Cas1B; Cas2, Cas3, Cas4 y Cas5; Cas2, Cas3, Cas4 y Cas6; Cas2, Cas3, Cas4, Cas1B y Cas5; Cas2, Cas3, Cas4, Cas1B y Cas6; Cas2, Cas3, Cas4, Cas1B, Cas5 y Cas6; Cas3, Cas4 y Cas1B; Cas3, Cas4 y Cas5; Cas3, Cas4 y Cas6; Cas3, Cas4, Cas1B y Cas5; Cas3, Cas4; Cas1B y Cas6; Cas3, Cas4, Cas1B, Cas5 y Cas6; Cas4, Cas1B y Cas5; Cas4, Cas1B y Cas6 o Cas4, Cas1B, Cas5 y Cas6, Cas5 y Cas6.

25 La proteína Cas puede ser Cas1 y Cas2; Cas1 y Cas3; Cas1 y Cas4; Cas1 y Cas1B; Cas1 y Cas5; Cas1 y Cas6; Cas2 y Cas3; Cas2 y Cas4; Cas2 y Cas1B; Cas2 y Cas5; Cas2 y Cas6; Cas3 y Cas4; Cas3 y Cas1B; Cas3 y Cas5; Cas3 y Cas6; Cas4 y Cas1B; Cas4 y Cas5; Cas4 y Cas6; Cas1B y Cas5 o Cas1B y Cas6 o combinaciones de las mismas.

30 La proteína Cas puede ser Cas1, Cas2 y Cas3; Cas1, Cas2 y Cas4; Cas1, Cas2 y Cas1B; Cas1, Cas2, Cas3 y Cas4; Cas1, Cas2, Cas3 y Cas1B; Cas1, Cas2, Cas3 y Cas5; Cas1, Cas2, Cas3 y Cas6; Cas1, Cas2, Cas3, Cas4 y Cas1B; Cas1, Cas2, Cas3, Cas4 y Cas5; Cas1, Cas2, Cas3, Cas4, Cas1B y Cas6; Cas1, Cas2, Cas3, Cas4, Cas1B y Cas5; Cas1, Cas2, Cas3, Cas4, Cas1B y Cas6; Cas1, Cas2, Cas3, Cas4, Cas1B, Cas5 y Cas6 o combinaciones de los mismos.

35 La proteína Cas puede ser Cas2, Cas3 y Cas4; Cas2, Cas3 y Cas1B; Cas2, Cas3 y Cas5; Cas2, Cas3 y Cas6; Cas2, Cas3, Cas4 y Cas1B; Cas2, Cas3, Cas4 y Cas5; Cas2, Cas3, Cas4 y Cas6; Cas2, Cas3, Cas4, Cas1B y Cas5; Cas2, Cas3, Cas4, Cas1B y Cas6; Cas2, Cas3, Cas4, Cas1B, Cas5 y Cas6 o combinaciones de las mismas.

La proteína puede ser Cas3, Cas4 y Cas1B; Cas3, Cas4 y Cas5; Cas3, Cas4 y Cas6; Cas3, Cas4, Cas1B y Cas5; Cas3, Cas4, Cas1B y Cas6; Cas3, Cas4, Cas1B, Cas5 y Cas6; Cas4, Cas1B y Cas5; Cas4, Cas1B y Cas6; Cas4, Cas1B, Cas5 y Cas6; Cas5 y Cas6 o combinaciones de las mismas.

40 La proteína Cas puede ser una o más de Cas1, Cas2, Cas3, Cas4, Cas1B, Cas5 y/o Cas6 o una pluralidad de las mismas, tal como una pluralidad de cualquiera de 2 genes Cas, cualquiera de 3 genes Cas, cualquiera de 4 genes Cas, cualquiera de 5 genes Cas, cualquiera de 6 genes Cas o cualquiera de 7 genes Cas.

45 La pluralidad de proteínas Cas puede comprender, consistir o consistir esencialmente en una pluralidad de las mismas proteínas Cas, tal como 2 proteínas Cas, 3 proteínas Cas, 4 proteínas Cas, 5 proteínas Cas, 6 proteínas Cas, 7 proteínas Cas, 8 proteínas Cas, 9 proteínas Cas, 10 proteínas Cas, 15 proteínas Cas, 20 proteínas Cas, 25 proteínas Cas, 30 proteínas Cas, 35 proteínas Cas, 40 proteínas Cas o incluso 50 o más proteínas Cas.

50 La pluralidad de proteínas Cas puede comprender, consistir o consistir esencialmente en una pluralidad de diferentes proteínas Cas, tal como 2 proteínas Cas diferentes, 3 proteínas Cas diferentes, 4 proteínas Cas diferentes, 5 proteínas Cas diferentes, 6 proteínas Cas diferentes, 7 proteínas Cas diferentes, 8 proteínas Cas diferentes, 9 proteínas Cas diferentes, 10 proteínas Cas diferentes, 15 proteínas Cas diferentes, 20 proteínas Cas diferentes, 25 proteínas Cas diferentes, 30 proteínas Cas diferentes, 35 proteínas Cas diferentes, 40 proteínas Cas diferentes o incluso 50 o más proteínas Cas diferentes.

La pluralidad de proteínas Cas puede comprender, consistir o consistir esencialmente en una pluralidad de las mismas y/o diferentes proteínas Cas, tal como 2 proteínas Cas diferentes, 3 proteínas Cas diferentes, 4 proteínas Cas diferentes, 5 proteínas Cas diferentes, 6 proteínas Cas diferentes, 7 proteínas Cas diferentes, 8 proteínas Cas diferentes, 9 proteínas Cas diferentes, 10 proteínas Cas diferentes, 15 proteínas Cas diferentes, 20 proteínas Cas diferentes, 25 proteínas Cas diferentes, 30 proteínas Cas diferentes, 35 proteínas Cas diferentes, 40 proteínas Cas diferentes o incluso 50 o más proteínas Cas diferentes. Las mismas proteínas Cas pueden duplicarse una pluralidad de veces.

Adecuadamente, el término "proteína Cas" se refiere a una pluralidad de proteínas Cas, tal como entre 2 y 12 proteínas Cas, más preferiblemente, entre 3 y 11 proteínas Cas, más preferiblemente, entre 4 y 10 proteínas Cas, más preferiblemente, entre 4 y 9 proteínas Cas, más preferiblemente, entre 4 y 8 proteínas Cas, más preferiblemente, entre 4 y 7 genes de proteínas, tales como 4, 5, 6 ó 7 proteínas Cas.

La(s) proteína(s) Cas se puede(n) codificar mediante un gen *cas* que comprende ADN o ARN de origen genómico, sintético o recombinante.

La(s) proteína(s) Cas se puede(n) codificar mediante un gen *cas* que puede ser de doble cadena o de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas.

La(s) proteína(s) Cas se puede(n) preparar mediante el uso de técnicas de ADN recombinante (por ejemplo, ADN recombinante), tal como se describe en la presente memoria.

En un aspecto adicional, se ha proporcionado un método para identificar un gen *cas* para usarse en la modulación de la resistencia de una célula frente a un ácido nucleico diana o producto de transcripción del mismo que comprende las etapas de: (i) preparar una célula que comprenda al menos un espaciador CRISPR y al menos dos repeticiones CRISPR; (ii) modificar por ingeniería la célula de modo que comprenda al menos un gen *cas*; y (iii) determinar si la célula modula la resistencia frente el ácido nucleico o producto de transcripción del mismo, donde la modulación de la resistencia de la célula frente al ácido nucleico diana o producto de transcripción del mismo es indicativo de que el gen *cas* se puede usar para modular la resistencia de la célula.

Uno o más de los genes *cas* se pueden usar para modificar por ingeniería una célula, tal como una célula receptora. En particular, uno o más genes *cas* se pueden usar para modificar por ingeniería una célula, tal como una célula receptora, que en combinación con una o más, preferiblemente, dos o más repeticiones CRISPR y uno o más espaciadores CRISPR se pueden usar para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo. A modo de ejemplo, el(los) gen(es) se puede(n) insertar en el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico de una célula, usando diversos métodos que son bien conocidos en la técnica. A modo de ejemplo adicional, los genes *cas* se pueden usar como una plantilla sobre la cual modificar (por ejemplo, mutar) el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, de modo que los genes *cas* se crean o forman en el ADN de la célula. A modo de ejemplo adicional, los genes *cas* se pueden clonar en una construcción, un plásmido o un vector y similares que a continuación se transforma en la célula, usando métodos tales como los descritos en la presente memoria.

Los genes *cas* pueden comprender o consistir en un grupo *cas* seleccionado entre el grupo que consiste en cualquiera de una o más de SEC. ID. N° 461, SEC. ID. N° 466, SEC. ID. N° 473, SEC. ID. N° 478, SEC. ID. N° 488, SEC. ID. N° 493, SEC. ID. N° 498, SEC. ID. N° 504, SEC. ID. N° 509, SEC. ID. N° 517.

Los genes *cas* pueden comprender o consistir en cualquiera de una o más de SEC. ID. N°s 462-465, 467-472, 474-477, 479-487, 489-492, 494-497, 499-503, 505-508, 510-516 y/o 517-521.

Adecuadamente, los uno o más genes *cas* o proteínas se usan junto con o en combinación con una o más, preferiblemente, dos o más repeticiones CRISPR y opcionalmente uno o más espaciadores CRISPR.

Repetición CRISPR

Tal como se usa en la presente memoria, el término "repetición CRISPR" tiene el significado convencional como el usado en la técnica, es decir, múltiples repeticiones directas y cortas, las cuales muestran ninguna o muy poca variación de secuencia dentro de un locus CRISPR dado.

Tal como se usa en la presente memoria, el término "CRISPR" es sinónimo del término "repetición CRISPR"

El número de nucleótidos en una repetición generalmente es aproximadamente 20 a aproximadamente 40 pares de bases, pero puede ser aproximadamente 20 a aproximadamente 39 pares de bases, aproximadamente 20 a aproximadamente 37 pares de bases, aproximadamente 20 a aproximadamente 35 pares de bases, aproximadamente 20 a aproximadamente 33 pares de bases, aproximadamente 20 a aproximadamente 30 pares de bases, aproximadamente 21 a aproximadamente 40 pares de bases, aproximadamente 21 a aproximadamente 39 pares de base, aproximadamente 21 a aproximadamente 37 pares de bases, aproximadamente 23 a

5 aproximadamente 40 pares de base, aproximadamente 23 a aproximadamente 39 pares de bases, aproximadamente 23 a aproximadamente 37 pares de bases, aproximadamente 25 a aproximadamente 40 pares de bases, aproximadamente 25 a aproximadamente 39 pares de bases, aproximadamente 25 a aproximadamente 37 pares de bases, aproximadamente 25 a aproximadamente 35 pares de bases, o aproximadamente 28 ó 29 pares de bases. El número de repeticiones puede oscilar entre aproximadamente 1 y aproximadamente 140, entre aproximadamente 1 y aproximadamente 100, entre aproximadamente 2 y aproximadamente 100, entre aproximadamente 5 y aproximadamente 100, entre aproximadamente 10 y aproximadamente 100, entre aproximadamente 15 y aproximadamente 100, entre aproximadamente 20 y aproximadamente 100, entre aproximadamente 25 y aproximadamente 100, entre aproximadamente 30 y aproximadamente 100, entre

10 aproximadamente 35 y aproximadamente 100, entre aproximadamente 40 y aproximadamente 100, entre aproximadamente 45 y aproximadamente 100, entre aproximadamente 50 y aproximadamente 100, entre aproximadamente 1 y aproximadamente 135, entre aproximadamente 1 y aproximadamente 130, entre aproximadamente 1 y aproximadamente 125, entre aproximadamente 1 y aproximadamente 120, entre aproximadamente 1 y aproximadamente 115, entre aproximadamente 1 y aproximadamente 110, entre

15 aproximadamente 1 y aproximadamente 105, entre aproximadamente 1 y aproximadamente 100, entre aproximadamente 1 y aproximadamente 95, entre aproximadamente 1 y aproximadamente 90, entre aproximadamente 1 y aproximadamente 80, entre aproximadamente 1 y aproximadamente 70, entre aproximadamente 1 y aproximadamente 60, entre aproximadamente 1 y aproximadamente 50, entre

20 aproximadamente 10 y aproximadamente 140, entre aproximadamente 10 y aproximadamente 130, entre aproximadamente 10 y aproximadamente 120, entre aproximadamente 10 y aproximadamente 110, entre aproximadamente 10 y aproximadamente 95, entre aproximadamente 10 y aproximadamente 90, entre aproximadamente 20 y aproximadamente 80, entre aproximadamente 30 y aproximadamente 70, entre aproximadamente 30 y aproximadamente 60, entre aproximadamente 30 y aproximadamente 50, entre aproximadamente 30 y aproximadamente 40, o aproximadamente 32.

25 Adecuadamente, el número de nucleótidos en una repetición generalmente es aproximadamente 20 a aproximadamente 40 pares de bases, pero puede ser aproximadamente 20 a aproximadamente 39 pares de bases, aproximadamente 20 a aproximadamente 37 pares de bases, aproximadamente 20 a aproximadamente 35 pares de bases, aproximadamente 20 a aproximadamente 33 pares de bases, aproximadamente 20 a aproximadamente 30 pares de bases, aproximadamente 21 a aproximadamente 40 pares de bases, aproximadamente 21 a

30 aproximadamente 39 pares de bases, aproximadamente 21 a aproximadamente 37 pares de bases, aproximadamente 23 a aproximadamente 40 pares de bases, aproximadamente 23 a aproximadamente 39 pares de bases, aproximadamente 23 a aproximadamente 37 pares de bases, aproximadamente 25 a aproximadamente 40 pares de bases, aproximadamente 25 a aproximadamente 39 pares de bases, aproximadamente 25 a aproximadamente 37 pares de bases, aproximadamente 25 a aproximadamente 35 pares de bases o

35 aproximadamente 28 o 29 pares de bases.

Adecuadamente, el número de repeticiones puede oscilar entre aproximadamente 2 y aproximadamente 140, entre aproximadamente 2 y aproximadamente 100, entre aproximadamente 2 y aproximadamente 100, entre aproximadamente 5 y aproximadamente 100, entre aproximadamente 10 y aproximadamente 100, entre aproximadamente 15 y aproximadamente 100, entre aproximadamente 20 y aproximadamente 100, entre

40 aproximadamente 25 y aproximadamente 100, entre aproximadamente 30 y aproximadamente 100, entre aproximadamente 35 y aproximadamente 100, entre aproximadamente 40 y aproximadamente 100, entre aproximadamente 45 y aproximadamente 100, entre aproximadamente 50 y aproximadamente 100.

Adecuadamente, el número de repeticiones puede oscilar entre aproximadamente 2 y aproximadamente 135, entre aproximadamente 2 y aproximadamente 130, entre aproximadamente 2 y aproximadamente 125, entre

45 aproximadamente 2 y aproximadamente 120, entre aproximadamente 2 y aproximadamente 115, entre aproximadamente 2 y aproximadamente 110, entre aproximadamente 2 y aproximadamente 105, entre aproximadamente 2 y aproximadamente 100, entre aproximadamente 2 y aproximadamente 95, entre aproximadamente 2 y aproximadamente 90, entre aproximadamente 2 y aproximadamente 80, entre

50 aproximadamente 2 y aproximadamente 70, entre aproximadamente 2 y aproximadamente 60, entre aproximadamente 2 y aproximadamente 50, entre aproximadamente 2 y aproximadamente 40, entre aproximadamente 2 y aproximadamente 30, entre aproximadamente 2 y aproximadamente 20, entre aproximadamente 2 y aproximadamente 10, entre aproximadamente 2 y aproximadamente 9, entre

55 aproximadamente 2 y aproximadamente 8, entre aproximadamente 2 y aproximadamente 7, entre aproximadamente 2 y aproximadamente 6, entre aproximadamente 2 y aproximadamente 5, entre aproximadamente 2 y aproximadamente 4, o entre aproximadamente 2 y aproximadamente 3.

La(s) repetición(es) CRISPR puede(n) comprender, consistir o consistir esencialmente en ADN o ARN de origen genómico, sintético o recombinante.

La(s) repetición(es) CRISPR puede(n) ser de doble cadena o de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de los mismas.

La(s) repetición(es) CRISPR se puede(n) preparar mediante el uso de técnicas de ADN recombinante (por ejemplo, ADN recombinante), tal como se describe en la presente memoria.

5 Una o más de las repeticiones CRISPR se pueden usar para modificar por ingeniería una célula, tal como una célula receptora. En particular, una o más, preferiblemente, dos o más repeticiones CRISPR se pueden usar para modificar por ingeniería una célula, tal como una célula receptora, que en combinación con uno o más genes *cas* o proteínas y uno o más espaciadores CRISPR se pueden usar para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo. A modo de ejemplo, las(s) repetición(es) se pueden insertar en el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico de una célula, usando diversos métodos que son bien conocidos en la técnica. A modo de ejemplo adicional, la(s) repetición(es) CRISPR se puede(n) usar como una plantilla sobre la cual modificar (por ejemplo, mutar) el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, de modo que la(s) repetición(es) CRISPR se crea(n) o modifica(n) por ingeniería en el ADN de la célula. A modo de ejemplo adicional, la(s) repetición(es) CRISPR se puede(n) clonar en una construcción, un plásmido o un vector y similares que a continuación se transforma en la célula, usando métodos tales como los descritos en la presente memoria.

15 En un aspecto adicional de la presente invención, se ha proporcionado también un método para identificar una repetición CRISPR para usarse en la modulación de la resistencia de una célula frente a un ácido nucleico diana o producto de transcripción del mismo que comprende las etapas de: (i) preparar una célula que comprenda al menos un espaciador CRISPR y al menos un gen *cas*; (ii) modificar por ingeniería la célula de modo que contenga una repetición CRISPR; y (iii) determinar si la célula modula la resistencia frente el ácido nucleico diana o producto de transcripción del mismo, donde la modulación de la resistencia de la célula frente al ácido nucleico diana o producto de transcripción del mismo es indicativo de que la repetición CRISPR se puede usar para modular la resistencia.

20 Adecuadamente, uno o más genes *cas* o proteínas se usan junto con o en combinación con una o más, preferiblemente, dos o más repeticiones CRISPR y opcionalmente uno o más espaciadores CRISPR. Adecuadamente, el(los) gen(es) *cas* o proteína(s) y repetición(es) CRISPR forma(n) una combinación funcional tal como se describe a continuación.

Un espaciador está flanqueado por dos repeticiones CRISPR. En otras palabras, un espaciador CRISPR tiene al menos una repetición CRISPR a cada lado.

Las repeticiones CRISPR pueden comprender o consistir en la secuencia de nucleótidos descrita en cualquiera de una o más de SEC. ID. N°s 1-22.

30 **Combinación funcional**

Tal como se ha mencionado anteriormente, sorprendentemente, los inventores han descubierto que un conjunto dado de genes *cas* o proteínas está siempre asociado con una secuencia repetida dada dentro de un locus CRISPR particular. En otras palabras, los genes *cas* o proteínas parecen ser específicos para una repetición de ADN dada, queriendo decir que los genes *cas* o proteínas y la secuencia repetida forman un par funcional.

35 Por consiguiente, las combinaciones particulares de uno o más genes *cas* o proteínas y una o más, preferiblemente, dos o más repeticiones CRISPR se usan para que un espaciador CRISPR confiera resistencia frente a un ácido nucleico diana o producto de transcripción del mismo en una célula (por ejemplo, una célula receptora). Por consiguiente, sorprendentemente se ha encontrado que no es posible usar simplemente cualquiera de los genes *cas* o proteínas o cualquier repetición CRISPR. En su lugar es una característica de la presente invención que la combinación sea funcional.

40 En el contexto de la combinación repetición CRISPR-gen *cas* o proteína descrita en la presente memoria, el término "funcional" quiere decir que la combinación es capaz de conferir resistencia a un ácido nucleico diana o un producto de transcripción del mismo cuando se usa junto con un espaciador CRISPR que se alinea con o es homólogo a un ácido nucleico diana o producto de transcripción del mismo.

45 Tal como se usa en la presente memoria el término "combinación funcional repetición CRISPR-*cas*" y "combinación funcional repetición CRISPR-gen *cas*" incluye una combinación funcional en la cual *cas* es un gen *cas* o una proteína Cas.

50 Adecuadamente, los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivados) a partir de la misma célula (por ejemplo, la misma célula receptora).

En una realización, el término "derivable" es sinónimo del término "obtenible".

En una realización, el término "derivado" es sinónimo del término "obtenido".

- Adecuadamente, los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR son derivables (preferiblemente, derivados) a partir del mismo locus CRISPR dentro de un genoma o plásmido, preferiblemente un genoma o plásmido de la misma cepa, especie o género.
- 5 Adecuadamente, los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR son derivables (preferiblemente, derivados) a partir del mismo locus CRISPR dentro de un genoma o plásmido sencillo, preferiblemente un genoma o plásmido sencillo de la misma cepa, especie o género.
- Adecuadamente, los uno o más genes *cas* o proteínas y las una o más, preferiblemente, dos o más repeticiones CRISPR se dan conjuntamente de manera natural.
- 10 Adecuadamente, los uno o más genes *cas* o proteínas y las una o más, preferiblemente, dos o más repeticiones CRISPR se dan conjuntamente de manera natural en la misma célula (por ejemplo, célula receptora).
- Adecuadamente, los uno o más genes *cas* o proteínas y las una o más, preferiblemente, dos o más repeticiones CRISPR se dan conjuntamente de manera natural en el mismo genoma de una célula (por ejemplo, célula receptora).
- 15 Adecuadamente, los uno o más genes *cas* o proteínas y las una o más, preferiblemente, dos o más repeticiones CRISPR se dan conjuntamente de manera natural en el mismo genoma de una cepa, especie o género.
- Por consiguiente, en un aspecto adicional, se ha proporcionado una combinación o ácido nucleico que consiste esencialmente en al menos dos repeticiones CRISPR y al menos un gen *cas* o proteína.
- 20 En una realización, el término “consiste esencialmente en” se refiere a una combinación de al menos dos repeticiones CRISPR y al menos un gen *cas* o proteína y excluyendo al menos un componente adicional de un locus CRISPR, tal como la ausencia de uno o más espaciador(es) CRISPR y/o la ausencia de una o más secuencia(s) líder común(es) o un locus CRISPR.
- En una realización, el término “consiste esencialmente en” se refiere a una combinación de al menos dos repeticiones CRISPR y al menos un gen *cas* o proteína sólo y excluyendo todos los otros componentes de un locus CRISPR, tal como un locus CRISPR que se da de manera natural.
- 25 En una realización adicional, el término “consiste esencialmente en” se refiere a una combinación de al menos dos repeticiones CRISPR y al menos un gen *cas* o proteína solo y excluyendo al menos un componente adicional de un locus CRISPR, excluyendo preferiblemente al menos un componente adicional de un locus CRISPR que se da de manera natural.
- 30 En una realización adicional, el término “consiste esencialmente en” se refiere a una combinación de al menos dos repeticiones CRISPR y al menos un gen *cas* o proteína a condición de que al menos un componente adicional del locus CRISPR natural esté ausente (por ejemplo, sustancialmente ausente).
- Adecuadamente, se ha proporcionado una combinación de al menos dos repeticiones CRISPR y al menos un gen *cas* o proteína a condición de que todos los otros componentes del locus CRISPR estén ausentes (por ejemplo, sustancialmente ausentes), preferiblemente que todos los otros componentes del locus CRISPR de la combinación natural de repetición(es) CRISPR y gen(es) *cas* estén ausentes.
- 35 Adecuadamente, los uno o más genes *cas* o proteínas se usan en combinación o junto con uno o más espaciadores CRISPR.
- Adecuadamente, los uno o más genes *cas* o proteínas se usan en combinación o junto con al menos uno o más espaciadores CRISPR y al menos una o más, preferiblemente, dos o más repeticiones CRISPR.
- 40 En una realización, el(los) espaciador(es) CRISPR está(n) o es(son) derivable(s) (preferiblemente, derivados) a partir de un organismo (por ejemplo, un organismo donante) que es diferente a la célula (por ejemplo, la célula receptora) a partir de la cual los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivados).
- 45 Se contemplan diversas ordenaciones de repetición(es) CRISPR y gen(es) *cas* o proteína(s), tales como combinaciones funcionales repetición CRISPR-*cas*.
- La combinación puede comprender, consistir o consistir esencialmente en al menos cualquiera de 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ó 20 repetición(es) CRISPR en combinación con cualquiera de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ó 20 genes *cas* o proteínas, tal como 16 repeticiones CRISPR y 12 genes *cas* o proteínas o 18 repeticiones CRISPR y 20 genes *cas* o proteínas o cualquiera de otras combinaciones de los mismos.

5 repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas1*, *cas2*, *cas3*, *cas4*, *cas1B* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas1*, *cas2*, *cas3*, *cas4*, *cas1B* y *cas5*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas1*, *cas2*, *cas3*, *cas4*, *cas1B* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas1*, *cas2*, *cas3*, *cas4*, *cas1B*, *cas5* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones) o combinaciones de los mismos.

10 Los genes *cas* pueden ser *cas2*, *cas3* y *cas4*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas2*, *cas3* y *cas1B*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas2*, *cas3* y *cas5*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas2*, *cas3* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas2*, *cas3*, *cas4* y *cas1B*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas2*, *cas3*, *cas4* y *cas5*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas2*, *cas3*, *cas4* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas2*, *cas3*, *cas4*, *cas1B* y *cas5*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas2*, *cas3*, *cas4*, *cas1B* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas2*, *cas3*, *cas4*, *cas1B*, *cas5* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones) o combinaciones de los mismos.

25 Los genes *cas* pueden ser *cas3*, *cas4* y *cas1B*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas3*, *cas4* y *cas5*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas3*, *cas4* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas3*, *cas4*, *cas1B* y *cas5*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas3*, *cas4*, *cas1B* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas3*, *cas4*, *cas1B*, *cas5* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas4*, *cas1B* y *cas5*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas4*, *cas1B* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas4*, *cas1B*, *cas5* y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones); *cas5*, y *cas6*-repetición (en los que la repetición es al menos dos repeticiones, preferiblemente, con al menos un espaciador entre las repeticiones) o combinaciones de los mismos.

40 Cuando la combinación de un gen *cas* y una repetición CRISPR comprende más de un gen *cas*, se entenderá que la repetición CRISPR puede estar insertada en el extremo 3' de los genes *cas*, el extremo 5' de los genes *cas*, o incluso entre los genes *cas*, siempre que al menos uno de los genes *cas* se mantenga funcional.

45 En una realización, una primera combinación repetición CRISPR-gen *cas* o proteína (que comprende al menos un gen *cas* o proteína y al menos dos repeticiones CRISPR, donde ambas son derivables (preferiblemente, derivadas) a partir del mismo locus CRISPR dentro de un genoma) se puede usar en combinación con una segunda combinación repetición CRISPR-gen *cas* o proteína (que comprende al menos un gen *cas* o proteína y al menos dos repeticiones CRISPR, donde ambas son derivables (preferiblemente, derivadas) a partir del mismo o un locus CRISPR diferente dentro de un genoma). Por consiguiente, en esta realización de la invención, la primera y segunda combinación son derivables (preferiblemente, derivadas) a partir de los mismos o diferentes loci CRISPR dentro de un genoma.

50 Por tanto, las primera y segunda combinación repetición CRISPR-gen *cas* puede incluso ser a partir de genomas diferentes, tales como los genomas diferentes dentro del mismo grupo, tal como se describe en más detalle en la presente memoria.

55 En una realización aún adicional de la presente invención, una primera y/o segunda combinación repetición CRISPR-gen *cas* o proteína (que comprende al menos un gen *cas* y al menos dos repeticiones CRISPR derivables (preferiblemente, derivadas) a partir del mismo locus CRISPR dentro de un genoma) se puede usar en combinación con 3, 4, 5, 6, 7, 8, 9 ó 10 o más combinaciones repetición CRISPR-gen *cas* o proteína (comprendiendo cada una al menos un gen *cas* o proteína y al menos dos repeticiones CRISPR derivables (preferiblemente, derivadas) a partir de los mismos o un loci CRISPR diferente dentro de un genoma). Por consiguiente, en esta realización de la

invención, las combinaciones son derivables (preferiblemente, derivadas) a partir de los mismo o diferentes loci CRISPR dentro de un genoma.

5 En una realización adicional de la invención, las combinaciones pueden incluso ser a partir de diferentes genomas, tales como los genomas diferentes dentro del mismo grupo, tal como se describe en más detalle en la presente memoria.

10 En otras palabras, para que la combinación repetición CRISPR-gen *cas* o proteína confiera resistencia, en algunas realizaciones, la(s) repetición(es) CRISPR y el(los) gen(es) o proteína(s) se da(n) conjuntamente de manera natural dentro de un locus CRISPR dado de un genoma. En algunas realizaciones, la(s) repetición(es) CRISPR y el(los) gen(es) *cas* o proteína(s) que se dan conjuntamente de manera natural dentro del mismo locus CRISPR de un genoma. Estas combinaciones funcionales juntas pueden conferir resistencia frente a un ácido nucleico o un producto de transcripción del mismo.

15 En un aspecto adicional, se ha proporcionado un método para identificar una combinación funcional de un gen *cas* o proteína y una repetición CRISPR que comprende las etapas de: (i) analizar las secuencias (por ejemplo, secuencias de ácido nucleico o proteína) del gen *cas* o proteína y la repetición CRISPR; (ii) identificar uno o más grupos de genes *cas* o proteínas; (iii) identificar uno o más grupos de repeticiones CRISPR; y (iv) combinar aquellas secuencias de genes *cas* o proteína y la repetición CRISPR que caen dentro del mismo grupo.

20 En un aspecto adicional, se ha proporcionado un método para identificar una combinación funcional de un gen *cas* o proteína y una repetición CRISPR para usarse en la modulación de la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo que comprende las etapas de: (i) preparar una célula que comprende una combinación de uno o más genes *cas* o proteínas y una o más, preferiblemente, dos o más repeticiones CRISPR; (ii) modificar por ingeniería la célula de modo que contenga uno o más espaciadores CRISPR; y (iii) determinar si la célula modula la resistencia frente a un ácido nucleico diana, donde la modulación de la resistencia de la célula frente al ácido nucleico diana o un producto de transcripción del mismo es indicativo de que la combinación se puede usar para modular la resistencia de la célula frente al ácido nucleico diana.

25 Adecuadamente, las secuencias del gen *cas* o proteína y la repetición CRISPR son o son derivables (preferiblemente, derivadas) a partir de las mismas o diferentes cepas.

Adecuadamente, las secuencias del gen *cas* o proteína y la repetición CRISPR son o son derivables (preferiblemente, derivadas) a partir de las mismas o diferentes especies.

30 Adecuadamente, las secuencias del gen *cas* o proteína y la repetición CRISPR son o son derivables (preferiblemente, derivadas) a partir de los mismos o diferentes géneros.

Adecuadamente, las secuencias del gen *cas* o proteína y la repetición CRISPR son o son derivables (preferiblemente, derivadas) a partir de los mismos o diferentes organismos.

Adecuadamente, el análisis se realiza usando análisis de grafico de punto.

35 La combinación puede comprender, consistir o consistir esencialmente en ADN o ARN de origen genómico, sintético o recombinante.

La combinación puede comprender, consistir o consistir esencialmente en ADN y ARN de origen genómico, sintético o recombinante.

La combinación puede comprender, consistir o consistir esencialmente en una repetición CRISPR de ADN de origen genómico, sintético o recombinante.

40 La combinación puede comprender, consistir o consistir esencialmente en una repetición CRISPR de ARN de origen genómico, sintético o recombinante.

La combinación puede comprender, consistir o consistir esencialmente en una repetición de gen *cas* de ADN de origen genómico, sintético o recombinante.

45 La combinación puede comprender, consistir o consistir esencialmente en un gen *cas* de ARN de origen genómico, sintético o recombinante.

La combinación puede comprender, consistir o consistir esencialmente en una repetición CRISPR de ADN y gen *cas* de ADN de origen genómico, sintético o recombinante.

La combinación puede comprender, consistir o consistir esencialmente en una repetición CRISPR de ADN y gen *cas* de ARN de origen genómico, sintético o recombinante.

La combinación puede comprender, consistir o consistir esencialmente en una repetición CRISPER de ARN y gen *cas* de ADN de origen genómico, sintético o recombinante.

La combinación puede comprender, consistir o consistir esencialmente en una repetición CRISPR de ARN y gen *cas* de ARN de origen genómico, sintético o recombinante.

5 La repetición CRISPER puede ser de doble cadena o cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas.

El gen *cas* puede ser de doble cadena o de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas.

10 La repetición CRISPR puede ser de doble cadena o de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas y el gen *cas* puede ser de doble cadena o de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas.

La repetición CRISPR puede ser de doble cadena si representa la cadena sentido o antisentido o combinaciones de las mismas y el gen *cas* puede ser de doble cadena si representa la cadena sentido o antisentido o combinaciones de las mismas.

15 La repetición CRISPR puede ser de doble cadena si representa la cadena sentido o antisentido o combinaciones de las mismas y el gen *cas* puede ser de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas.

20 La repetición CRISPR puede ser de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas y el gen *cas* puede ser de doble cadena si representa la cadena sentido o antisentido o combinaciones de las mismas.

La repetición CRISPR puede ser de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas y el gen *cas* puede ser de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas.

25 Una o más de las combinaciones funcionales como las descritas anteriormente se pueden usar para modificar por ingeniería una célula, tal como una célula receptora. En particular, una o más combinaciones funcionales se pueden usar para modificar por ingeniería una célula, tal como una célula receptora, que en combinación con uno o más espaciadores CRISPR se pueden usar para modular la resistencia de una célula frente a un ácido nucleico diana o producto de transcripción del mismo. A modo de ejemplo, las combinaciones funcionales se pueden insertar en el ADN de una célula (por ejemplo, una célula receptora), tal como ADN de plásmido o ADN genómico de una célula, usando diversos métodos que son bien conocidos en la técnica. A modo de ejemplo adicional, las combinaciones funcionales se pueden usar como plantilla sobre la cual modificar (por ejemplo, mutar) el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, se crean de modo que las combinaciones funcionales en el ADN de la célula. A modo de ejemplo adicional, las combinaciones funcionales se pueden clonar en una construcción, un plásmido o un vector y similares el cual a continuación se transforman en la célula, usando métodos tales como los descritos en la presente memoria.

35 En una realización, la combinación funcional se obtiene o es obtenible mediante un método que comprende las etapas de: (a) analizar las secuencias de un gen *cas* y una repetición CRISPR; (b) identificar uno o más grupos de genes *cas*; (c) identificar uno o más grupos de repeticiones CRISPR; y (d) combinar aquellos genes *cas* y secuencias de repetición CRISPR que caen dentro del mismo grupo, donde la combinación del gen *cas* y las secuencias de repetición CRISPR dentro del mismo grupo es indicativo de que la combinación es una combinación funcional.

Los grupos se describen a continuación en más detalle.

Espaciador CRISPR

45 Tal como se usa en la presente memoria, el término "espaciador CRISPR" tiene el significado convencional como el usado en la técnica y se refiere a las secuencias de espaciador no repetitivas que se encuentran entre las múltiples repeticiones directas y cortas (es decir, repeticiones CRISPR) de loci CRISPR. En otras palabras, un espaciador CRISPR se encuentra entre dos repeticiones CRISPR.

50 Se ha encontrado que las secuencias de espaciador CRISPR en procariontes con frecuencia tienen similitudes significantes a una diversidad de moléculas de ADN, tales como los elementos genéticos (que incluyen, pero no se limitan a, cromosomas, bacteriófagos y plásmidos conjugativos). Curiosamente, las células que llevan estos espaciadores CRISPR son incapaces de ser infectadas por moléculas de ADN que contienen secuencias homólogas a los espaciadores (Mojica et al. 2005).

Generalmente, el espaciador CRISPR de manera natural está presente entre dos múltiples repeticiones directas, cortas, e idénticas que son palindrómicas.

5 Adecuadamente, el espaciador CRISPR es homólogo al ácido nucleico diana o un producto de transcripción del mismo o una secuencia identificada. Aunque la homología también se puede considerar en términos de similitud, en el contexto de la presente invención se prefiere expresar la homología en términos de identidad de secuencia. Una secuencia homóloga se supone que incluye un espaciador CRISPR, el cual puede ser al menos idéntico al 70, 75, 80, 85 ó 90%, o al menos idéntico al 91, 92, 93, 94, 95, 96, 97, 98 ó 99% a la secuencia de ácido nucleico diana o un producto de transcripción de la misma o una secuencia identificada.

En algunas realizaciones, el espaciador CRISPR es idéntico al 100% a la secuencia del ácido nucleico diana.

10 El número de espaciadores CRISPR en un loci o locus CRISPR dado puede variar entre especies.

Adecuadamente, el número de espaciadores puede oscilar entre aproximadamente 1 y aproximadamente 140, entre aproximadamente 1 y aproximadamente 100, entre aproximadamente 2 y aproximadamente 100, entre aproximadamente 5 y aproximadamente 100, entre aproximadamente 10 y aproximadamente 100, entre aproximadamente 15 y aproximadamente 100, entre aproximadamente 20 y aproximadamente 100, entre aproximadamente 25 y aproximadamente 100, entre aproximadamente 30 y aproximadamente 100, entre aproximadamente 35 y aproximadamente 100, entre aproximadamente 40 y aproximadamente 100, entre aproximadamente 45 y aproximadamente 100, entre aproximadamente 50 y aproximadamente 100.

20 Adecuadamente, el número de espaciadores puede oscilar entre aproximadamente 1 y aproximadamente 135, entre aproximadamente 1 y aproximadamente 130, entre aproximadamente 1 y aproximadamente 125, entre aproximadamente 1 y aproximadamente 120, entre aproximadamente 1 y aproximadamente 115, entre aproximadamente 1 y aproximadamente 110, entre aproximadamente 1 y aproximadamente 105, entre aproximadamente 1 y aproximadamente 100, entre aproximadamente 1 y aproximadamente 95, entre aproximadamente 1 y aproximadamente 90, entre aproximadamente 1 y aproximadamente 80, entre aproximadamente 1 y aproximadamente 70, entre aproximadamente 1 y aproximadamente 60, entre aproximadamente 1 y aproximadamente 50, entre aproximadamente 1 y aproximadamente 40, entre aproximadamente 1 y aproximadamente 30, entre aproximadamente 1 y aproximadamente 20, entre aproximadamente 1 y aproximadamente 10, entre aproximadamente 1 y aproximadamente 9, entre aproximadamente 1 y aproximadamente 8, entre aproximadamente 1 y aproximadamente 7, entre aproximadamente 1 y aproximadamente 6, entre aproximadamente 1 y aproximadamente 5, entre aproximadamente 1 y aproximadamente 4, entre aproximadamente 1 y aproximadamente 3, entre aproximadamente 1 y aproximadamente 2.

Generalmente, los espaciadores CRISPR se identifican mediante análisis de secuencia como los tramos de ADN localizados entre dos repeticiones.

35 Tal como se describe en la presente memoria, los inventores sorprendentemente han descubierto que el uso de uno o más genes *cas* o proteínas en combinación con una o más, preferiblemente, dos o más repeticiones CRISPR (preferiblemente, combinación(es) funcional(es) de las mismas) mantienen la especificidad de la inmunidad a conferir por uno o más espaciadores CRISPR en una célula, tal como una célula receptora.

40 Tal como se usa en la presente memoria, el término "especificidad de inmunidad" quiere decir que se puede conferir inmunidad frente a una secuencia específica de ácido nucleico o producto de transcripción de la misma usando una secuencia de espaciador CRISPR específico (o seudo espaciador CRISPR). Por consiguiente, un espaciador CRISPR dado no confiere resistencia frente a cualquier secuencia de ácido nucleico o producto de transcripción de la misma sino solamente a aquellas secuencias frente a las cuales el espaciador CRISPR (o seudo espaciador CRISPR) es homólogo, tal como idéntico al 100%.

45 El(los) espaciador(es) CRISPR puede(n) ser derivables (preferiblemente derivados) a partir de un organismo, tal como un organismo donante, que es diferente a la célula, tal como la célula receptora o incluso un organismo donante adicional, a partir del cual los uno o más genes o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivadas). Los espaciadores CRISPR pueden estar o pueden ser derivables (preferiblemente, derivados) a partir de un organismo, tal como un organismo donante, que es heterólogo a la célula, tal como la célula receptora o incluso un organismo donante adicional, a partir del cual los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivadas). Los uno o más genes *cas* y/o las una o más, preferiblemente, dos o más repeticiones CRISPR pueden estar o pueden ser derivables (preferiblemente, derivados) a partir de una célula homóloga (es decir, la misma), tal como una célula receptora homóloga.

55 Para evitar la duda, los espaciador(es) CRISPR se puede(n) diseñar y producir sintéticamente (por ejemplo, usando técnicas de ADN recombinante).

- 5 En una realización, los espaciadores CRISPR son heterólogos (es decir, diferentes) a la célula, tal como la célula receptora, a partir de la cual los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR son o son derivables (preferiblemente derivadas) y los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivadas) a partir de una célula homóloga, tal como una célula receptora homóloga.
- En otra realización, los espaciadores CRISPR son heterólogos (es decir, diferentes) a la célula, tal como la célula receptora, a partir de la cual los uno o más genes *cas* o proteínas son o son derivables (preferiblemente, derivados).
- 10 En otra realización, los espaciadores CRISPR son heterólogos a la célula, tal como la célula receptora y los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivados) a partir de una célula homóloga, tal como una célula receptora homóloga.
- En otra realización, los espaciadores CRISPR son heterólogos a la célula, tal como la célula receptora, mientras que los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR es/son homólogas a la célula, tal como la célula receptora.
- 15 En otra realización, los espaciadores CRISPR son heterólogos a la célula receptora, mientras que la célula receptora es homóloga para los uno o más genes *cas* o proteínas y/o las una o más, preferiblemente, dos o más repeticiones CRISPR.
- En otra realización, el espaciador CRISPR usado de acuerdo con la presente invención es uno que no está asociado de manera natural con la repetición CRISPR y/o genes *cas* y/o la combinación funcional repetición CRISPR-gen *cas*. En otras palabras, el espaciador CRISPR en el locus CRISPR recombinante de acuerdo con la presente invención es heterólogo a la repetición CRISPR y/o los genes *cas* del locus CRISPR.
- 20 Uno o más de los espaciadores CRISPR se pueden usar para modificar por ingeniería una célula, tal como una célula receptora. En particular, uno o más espaciadores CRISPR se pueden usar para modificar por ingeniería una célula, tal como una célula receptora, que en combinación con uno o más genes *cas* o proteínas y/o una o más, preferiblemente, dos o más repeticiones CRISPR (preferiblemente, una o más combinaciones funcionales de las mismas) se pueden usar para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.
- 25 Adecuadamente, uno o más espaciadores CRISPR se pueden usar para modificar por ingeniería una célula, tal como una célula receptora. En particular, uno o más espaciadores CRISPR se pueden usar para modificar por ingeniería una célula, tal como una célula receptora, que en combinación con uno o más genes *cas* o proteínas se pueden usar para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.
- 30 A modo de ejemplo, los espaciadores CRISPR se pueden insertar en el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico de una célula, usando diversos métodos que son bien conocidos en la técnica.
- 35 A modo de ejemplo adicional, los espaciadores CRISPR se pueden usar como plantilla sobre la cual modificar (por ejemplo, mutar) el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, de modo que los espaciadores CRISPR se crean en el ADN de la célula.
- A modo de ejemplo adicional, los espaciadores CRISPR se pueden clonar en una construcción, un plásmido o un vector y similares que a continuación se transforman en la célula, usando métodos como los descritos en la presente invención.
- 40 En un aspecto adicional, también se ha proporcionado un método para identificar un espaciador CRISPR para usarse en la modulación de la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo que comprende las etapas de: (i) preparar una célula que comprenda al menos dos repeticiones CRISPR y al menos un gen *cas* o proteína; (ii) identificar al menos un espaciador CRISPR en un organismo, tal como un organismo donante; (iii) modificar la secuencia del espaciador CRISPR de la célula de modo que tenga homología al espaciador CRISPR del organismo donante que comprende el ácido nucleico diana; y (iv) determinar si la célula modula la resistencia frente al ácido nucleico diana, donde la modulación de la resistencia de la célula frente al ácido nucleico diana o producto de transcripción del mismo es indicativo de que el espaciador CRISPR modula la resistencia de la célula frente al ácido nucleico diana.
- 45 Los espaciadores CRISPR pueden comprender o consistir en la secuencia de nucleótidos descrita en cualquiera o más de cualquiera de SEC. ID. N°s 23-460 y/o SEC. ID. N°s 522-665.
- 50 Un espaciador CRISPR está flanqueado por dos repeticiones CRISPR. En otras palabras, un espaciador CRISPR tiene al menos una repetición a cada lado.

Sin estar obligado por ninguna teoría particular, cuanto más lejos esté un espaciador CRISPR dado del extremo 5' del locus CRISPR (que comprende el(los) gen(es) o la secuencia líder), puede ser menor la resistencia conferida por ese espaciador CRISPR. Por consiguiente, en una realización de la presente invención se prefiere que uno o más de los primeros 100 espaciadores CRISPR a partir del extremo 5' del locus CRISPR (que comprende los genes *cas* y/o la secuencia líder) estén modificados, más preferiblemente, que uno o más de los primeros 50 espaciadores CRISPR a partir del extremo 5' del locus CRISPR (que comprende los genes *cas* y/o la secuencia líder) estén modificados, más preferiblemente, que uno o más de los primeros 40 espaciadores CRISPR a partir del extremo 5' del locus CRISPR (que comprende los genes *cas* y/o la secuencia líder) estén modificados, más preferiblemente, que uno o más de los primeros 30 espaciadores CRISPR a partir del extremo 5' del locus CRISPR (que comprende los genes *cas* y/o la secuencia líder) estén modificados, más preferiblemente, que uno o más de los primeros 20 espaciadores CRISPR a partir del extremo 5' del locus CRISPR (que comprende los genes *cas* y/o la secuencia líder) estén modificados, más preferiblemente, que uno o más de los primeros 15 espaciadores CRISPR a partir del extremo 5' del locus CRISPR (que comprende los genes *cas* y/o la secuencia líder) estén modificados, más preferiblemente, que uno o más de los primeros 10 espaciadores CRISPR a partir del extremo 5' del locus CRISPR (que comprende los genes *cas* y/o la secuencia líder) estén modificados.

Tal como se apreciará por el experto en la técnica, diferentes bacterias tienen diferentes números de espaciadores CRISPR.

Zona central del Espaciador CRISPR

Para un tipo de CRISPR específico dentro de una especie microbiana, el espaciador CRISPR generalmente se representa mediante una longitud predominante definida, pero el tamaño puede variar. Los tipos de CRISPR descritos hasta la fecha se han encontrado que contienen una longitud de espaciador predominante de entre aproximadamente 20 pb y aproximadamente 58 pb.

Tal como se usa en la presente memoria, el término "zona central de espaciador CRISPR" quiere decir la longitud del espaciador observado más corto dentro de un tipo de CRISPR. Por tanto, a modo de ejemplo, dentro de CRISPR Tipo 1 de *Streptococcus thermophilus*, la longitud del espaciador dominante es 30 pb con una minoría de espaciadores entre 28 pb y 32 pb en tamaño. Así en este ejemplo particular (CRISPR Tipo 1 de *S. thermophilus*), la zona central de espaciador CRISPR se define como un tramo continuo de 28 pb.

Adecuadamente, la zona central del espaciador CRISPR es homóloga al ácido nucleico diana o un producto de transcripción del mismo o una secuencia identificada sobre la longitud de la secuencia de la zona central. Aunque la homología también se puede considerar en términos de similitud, en el contexto de la presente invención se prefiere expresar homología en términos de identidad de secuencia. Una secuencia homóloga se supone que incluye una zona central de espaciador CRISPR, la cual puede ser al menos idéntico al 90% o al menos idéntico al 91, 92, 93, 94, 95, 96, 97, 98 o 99 % a la secuencia de ácido nucleico diana o un producto de transcripción del mismo o una secuencia identificada sobre la longitud de la secuencia de la zona central.

Adecuadamente, la zona central del espaciador CRISPR es idéntico al 100% a la secuencia de ácido nucleico diana o un producto de transcripción del mismo o una secuencia identificada sobre la longitud de la secuencia de la zona central.

Seudo espaciador CRISPR

Tal como se usa en la presente memoria, el término "seudo espaciador CRISPR" se refiere a una secuencia de ácido nucleico presente en un organismo (por ejemplo, un organismo donante), tal como bacteriófago, la cual es preferiblemente esencial para la función y/o supervivencia y/o replicación y/o infectividad y similares, y la cual forma una secuencia espaciador CRISPR; y/o se puede usar para formar o preparar una secuencia de espaciador CRISPR la cual es complementaria a o homóloga al seudo espaciador CRISPR; y/o se puede usar para modular la resistencia.

Uno o más de los seudo espaciadores CRISPR o espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) al uno o más seudo espaciador(es) CRISPR se puede(n) usar para modificar por ingeniería una célula, tal como una célula receptora. En particular, uno o más seudo espaciadores CRISPR o espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPR se puede(n) usar para modificar por ingeniería una célula, tal como una célula receptora, que en combinación con uno o más genes *cas* o proteínas y/o una o más repeticiones CRISPR (por ejemplo, una o más combinaciones funcionales de las mismas) se pueden usar para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.

Uno o más seudo espaciadores CRISPR o espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPR se puede(n) usar para modificar por ingeniería una célula, tal como

una célula receptora, que en combinación con uno o más genes *cas* o proteínas se pueden usar para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.

5 A modo de ejemplo, los seudo espaciadores CRISPR o espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPR se puede(n) insertar en el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico de una célula, usando diversos métodos que son bien conocidos en la técnica.

10 A modo de ejemplo adicional, los seudo espaciadores CRISPR se pueden usar como plantilla sobre la cual modificar (por ejemplo, mutar) el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, de modo que los espaciadores CRISPR se crean en el ADN de la célula. A modo de ejemplo adicional, los seudo espaciadores CRISPR o espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPR se puede(n) clonar en una construcción, un plásmido o un vector y similares que a continuación se transforman en la célula, usando métodos como los descritos en la presente memoria.

Secuencia de ácido nucleico

15 En un aspecto adicional, se ha proporcionado una secuencia de ácido nucleico (por ejemplo, una secuencia recombinante o una de ácido nucleico aislada) que consiste esencialmente en al menos un gen *cas* o proteína.

20 La secuencia de ácido nucleico puede ser ADN o ARN de origen genómico, sintético o recombinante, por ejemplo, ADNc. La secuencia de nucleótidos puede ser de doble cadena o de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas. Las secuencias de ácido nucleico recombinante se pueden preparar mediante el uso de técnicas de ADN recombinante, tal como se ha descrito en la presente memoria. La secuencia de ácido nucleico diana puede ser o puede ser derivada a partir de un gen.

Tal como se usa en la presente memoria, el término “que esencialmente consiste en” en el contexto de la secuencia de ácido nucleico se refiere a una secuencia de ácido nucleico que comprende uno o más genes *cas* y que excluye al menos un componente adicional de un locus CRISPR, tal como las repeticiones CRISPR, los espaciadores CRISPR y/o la secuencia líder común.

25 Por consiguiente, en una realización, se ha proporcionado una secuencia de ácido nucleico que consiste esencialmente en al menos un gen *cas* y al menos dos repeticiones CRISPR.

En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que consiste esencialmente en al menos un gen *cas* y al menos un espaciador CRISPR.

30 En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que consiste esencialmente en al menos un gen *cas*, al menos un espaciador CRISPR y al menos dos repeticiones CRISPR.

En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que comprende al menos un gen *cas* a condición de que al menos un componente adicional de un locus CRISPR esté ausente, adecuadamente, a condición de que al menos un componente adicional de un locus CRISPR natural esté ausente (por ejemplo, sustancialmente ausente).

35 En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que comprende al menos un gen *cas* a condición de que los espaciadores CRISPR del locus CRISPR estén ausentes, adecuadamente, a condición de que los espaciadores CRISPR de un locus CRISPR natural estén ausentes (por ejemplo, sustancialmente ausente).

40 En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que comprende al menos un gen *cas* a condición de que las repeticiones CRISPR del locus CRISPR estén ausentes, adecuadamente, a condición de que las repeticiones CRISPR de un locus CRISPR natural estén ausentes.

En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que comprende al menos un gen *cas* a condición de que las secuencias líder comunes del locus CRISPR estén ausentes, adecuadamente, a condición de que las secuencias líder comunes del locus CRISPR natural estén ausentes.

45 En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que comprende al menos un gen *cas* a condición de que los espaciadores CRISPR y las repeticiones CRISPR del locus CRISPR estén ausentes, adecuadamente, a condición de que los espaciadores CRISPR y las repeticiones CRISPR del locus CRISPR natural estén ausentes.

50 En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que comprende al menos un gen *cas* a condición de que los espaciadores CRISPR y las repeticiones CRISPR del locus CRISPR estén ausentes,

adecuadamente, a condición de que los espaciadores CRISPR y las repeticiones CRISPR del locus CRISPR natural estén ausentes.

5 En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que comprende al menos un gen cas a condición de que los espaciadores CRISPR y las secuencias líder comunes del locus CRISPR estén ausentes, adecuadamente, a condición de que los espaciadores CRISPR y las secuencias líder comunes del locus CRISPR natural estén ausentes.

10 En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que comprende al menos un gen cas a condición de que las repeticiones CRISPR y las secuencias líder comunes del locus CRISPR estén ausentes, adecuadamente, a condición de que las repeticiones CRISPR y las secuencias líder comunes del locus CRISPR natural estén ausentes.

En una realización adicional, se ha proporcionado una secuencia de ácido nucleico que comprende al menos un gen cas a condición de que las repeticiones CRISPR, los espaciadores CRISPR y las secuencias líder comunes del locus CRISPR estén ausentes, adecuadamente, a condición de que las repeticiones CRISPR, los espaciadores CRISPR y las secuencias líder comunes del locus CRISPR natural estén ausentes.

15 La secuencia de ácido nucleico y los ácidos nucleicos se pueden aislar o sustancialmente purificar. Por "aislar" o "sustancialmente purificar" se hace pensar que las moléculas de ácido nucleico, o fragmentos biológicamente activos o variantes, homólogos, o derivados de los mismos están sustancialmente o esencialmente libres de componentes normalmente encontrados en asociación con el ácido nucleico en su estado natural. Dichos
20 componentes incluyen otro material, medios de cultivo a partir de la producción recombinante y diversos componentes químicos usados en sintetizar químicamente los ácidos nucleicos.

Una secuencia de ácido nucleico o ácido nucleico "aislado" generalmente está libre de secuencias de ácido nucleico que flanquean el ácido nucleico de interés en el ADN genómico del organismo a partir del cual se derivó el ácido nucleico (tal como secuencias codificadoras presentes en los extremos 5' o 3'). Sin embargo, la molécula puede incluir algunas bases o restos adicionales que no afectan de manera perjudicial las características básicas de la
25 composición.

La(s) secuencia(s) de ácido nucleico se puede(n) usar en la modificación por ingeniería de una célula, tal como una célula receptora. A modo de ejemplo, la secuencia de ácido nucleico puede estar insertada en el ADN, tal como ADN plásmido o ADN genómico, de una célula receptora, usando métodos, tal como recombinación homóloga. A modo
30 de ejemplo adicional, la(s) secuencia(s) de ácido nucleico se puede(n) usar como plantilla sobre la cual modificar (por ejemplo, mutar) el ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, de modo que la(s) secuencia(s) de ácido nucleico se crea(n) en el ADN de la célula. A modo de ejemplo adicional, la(s) secuencia(s) de ácido nucleico se puede(n) clonar en una construcción, un plásmido o un vector y similares que a continuación se transforman en la célula, usando métodos tales como los descritos en la presente memoria.

35 Un espaciador CRISPR está flanqueado por dos repeticiones CRISPR. En otras palabras, un espaciador CRISPR tiene al menos una repetición CRISPR a cada lado.

Secuencia de Ácido Nucleico Diana

40 Tal como se usa en la presente memoria, el término "secuencia de ácido nucleico diana" se refiere a cualquier secuencia de ácido nucleico o producto de transcripción de la misma, frente a la cual se modula la resistencia en una célula, tal como una célula receptora,.

La resistencia puede ser frente a la secuencia de ácido nucleico diana *per se*. Ventajosamente, esto confiere resistencia a una célula frente a un organismo donante a partir del cual el(los) ácido(s) nucleico(s) diana es(son) derivable(s) (preferiblemente, derivado(s)). Por tanto, a modo de ejemplo, la inserción de un seudo espaciador CRISPR derivable (preferiblemente, derivado) a partir de un bacteriófago o un(os) espaciador(es) CRISPR que
45 es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPR en una célula, tal como una célula receptora, puede conferir resistencia al bacteriófago. Por tanto, a modo de ejemplo adicional, la inserción entre dos repeticiones CRISPR de un seudo espaciador CRISPR derivable (preferiblemente, derivado) a partir de un bacteriófago o espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPER en una célula, tal como una célula receptora, puede conferir resistencia al bacteriófago.

50 La resistencia puede ser frente a un producto de transcripción de la secuencia de ácido nucleico diana, tal como un transcrito de la secuencia de ácido nucleico diana (por ejemplo, un transcrito de ARN (por ejemplo, ARNm) (por ejemplo, un transcrito sentido o antisentido) o incluso un producto de transcripción de polipéptido. Ventajosamente, esto confiere resistencia a una célula frente a un organismo donante a partir del cual el producto de transcripción es derivable (preferiblemente, derivado).

La secuencia de nucleótidos diana puede ser ADN o ARN de origen genómico, sintético o recombinante.

La secuencia de nucleótidos puede ser de doble cadena o de cadena sencilla si representa la cadena sentido o antisentido o combinaciones de las mismas.

5 La secuencia de nucleótidos se puede preparar mediante el uso de técnicas de ADN recombinante (por ejemplo, ADN recombinante).

La secuencia de nucleótidos puede ser la misma que una forma que se da de manera natural, o puede ser derivable (preferiblemente, derivada) a partir de la misma.

La secuencia de ácido nucleico diana puede ser o puede ser derivable (preferiblemente, derivada) a partir de un gen.

10 La secuencia de ácido nucleico diana puede ser o puede ser derivable (preferiblemente, derivada) a partir de una variante, homólogo, fragmento o derivado de un gen.

En una realización, la secuencia nucleica diana es o es derivable (preferiblemente, derivada) a partir de bacteriófago.

En una realización, la secuencia nucleica diana es o es derivable (preferiblemente, derivada) a partir de ADN plásmido.

15 En una realización, la secuencia nucleica diana es o es derivable (preferiblemente, derivada) a partir de un elemento genético móvil.

En una realización, la secuencia nucleica diana es o es derivable (preferiblemente, derivada) a partir de un elemento transponible o una secuencia de inserción.

20 En una realización, la secuencia nucleica diana es o es derivable (preferiblemente, derivada) a partir de un gen que confiere resistencia.

En una realización, la secuencia nucleica diana es o es derivable (preferiblemente, derivada) a partir de un gen que confiere resistencia a un antibiótico.

En una realización, la secuencia nucleica diana es o es derivable (preferiblemente, derivada) a partir de un factor de virulencia.

25 En una realización, la secuencia nucleica diana es o es derivable (preferiblemente, derivada) a partir de una toxina, una intermalina o una hemolisina.

Modular resistencia

En un aspecto adicional, se ha proporcionado un método para modular la resistencia de una célula, tal como una célula receptora, frente a un ácido nucleico diana o un producto de transcripción del mismo.

30 Tal como se usa en la presente memoria, el término “modular la resistencia” puede referirse a suprimir, reducir, disminuir, inducir, conferir, restaurar, elevar, incrementar o de lo contrario afectar la resistencia de una célula a un ácido nucleico diana.

35 Tal como se usa en la presente memoria, el término “resistencia” no quiere decir que implique que una célula es 100% resistente a un ácido nucleico diana o un producto de transcripción del mismo, pero incluye células que son tolerantes al ácido nucleico diana o un producto de transcripción del mismo.

Tal como se usa en la presente memoria el término “resistencia a ácido nucleico diana o producto de transcripción del mismo” quiere decir que se confiere resistencia frente a una célula o un organismo, tal como un fago, que comprende o produce el ácido nucleico diana o producto de transcripción del mismo.

40 Sin estar obligado por cualquier teoría particular, creemos que la resistencia o inmunidad no está enlazada a la “entrada” de ADN exógeno en una célula (es decir, penetración a través de la membrana celular). La inmunidad o resistencia se prefiere que corresponda a una obstrucción, obstáculo, impedimento, barrera o anulación de la persistencia, mantenimiento o supervivencia del ácido nucleico entrante (o bien, por ejemplo, en una forma lineal libre, o integrado dentro del cromosoma bacteriano, fuera de un locus CRISPR o dentro de una molécula circular, tal como un plásmido), o a una obstrucción, obstáculo, impedimento, barrera o anulación de su replicación y/o transcripción y/o expresión.

45

En una realización, los componentes mínimos que confieren inmunidad o resistencia frente a un ácido nucleico diana o producto de expresión del mismo son al menos un gen *cas* (o una proteína Cas) y al menos dos repeticiones CRISPR que flanquean un espaciador.

5 En una realización, se prefiere que “modular resistencia” quiera decir inducir, conferir, elevar o incrementar la resistencia de una célula a un ácido nucleico diana.

10 En un aspecto, se ha proporcionado un método para modular (por ejemplo, conferir o incrementar) la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo que comprende las etapas de: (i) identificar una secuencia (por ejemplo, una secuencia conservada) en un organismo (preferiblemente, una secuencia esencial a la función o supervivencia del organismo); (ii) preparar un espaciador CRISPR que sea un homólogo de secuencia, (adecuadamente idéntico al 100%), a la secuencia identificada; (iii) preparar un ácido nucleico que comprenda al menos un gen *cas* y al menos dos repeticiones CRISPR junto con el espaciador CRISPR; y (iv) transformar una célula con dicho ácido nucleico para así volver a la resistencia celular a dicho ácido nucleico diana o producto de transcripción del mismo.

15 Tal como se usa en la presente memoria, el término “secuencia conservada” en el contexto de identificar una secuencia conservada en un organismo necesariamente no tiene que ser conservada en su sentido más estricto puesto que el conocimiento de una secuencia a partir de un organismo dado sería suficiente. Además, la secuencia no necesita ser parte de una entidad esencial, puesto que creemos que un espaciador inspirado a partir de un gen esencial sería más eficiente en conferir inmunidad o resistencia.

20 En una realización, la secuencia conservada es una secuencia que es esencial para la función y/o la supervivencia y/o la replicación y/o la infectividad y similares de un organismo o una célula. A modo de ejemplo, la secuencia conservada puede ser una helicasa, una primasa, una proteína estructural de cabeza o cola, una proteína con un dominio conservado (por ejemplo, holin, lisina y otros) o unas secuencia conservadas entre genes importantes de fago.

25 En un aspecto adicional, se ha proporcionado un método para modular (por ejemplo, conferir o incrementar) la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo que comprende las etapas de: (i) identificar uno o más espaciadores CRISPR en un organismo resistente al ácido nucleico diana o producto de transcripción del mismo; (ii) preparar un ácido nucleico recombinante que comprenda al menos un gen *cas* o proteína y al menos dos repeticiones CRISPR junto con dichos uno o más espaciadores identificados; y (iii) transformar una célula con dicho ácido nucleico recombinante para así volver a la célula receptora resistente a dicho ácido nucleico diana o producto de transcripción del mismo.

35 En un aspecto adicional, se ha proporcionado un método para modular (por ejemplo, conferir o incrementar) la resistencia de una célula que comprende al menos uno o más genes *cas* o proteínas y una o más, preferiblemente, dos o más repeticiones CRISPR frente a ácido nucleico diana o un producto de transcripción del mismo que comprende las etapas de: (i) identificar uno o más espaciadores CRISPR en un organismo resistente al ácido nucleico diana o producto de transcripción del mismo, y (ii) modificar la secuencia de uno o más espaciador(es) CRISPR en la célula de modo que el(los) espaciador(es) tiene(n) homología a los espaciador(es) CRISPR en el organismo.

40 En una realización, uno o más espaciadores CRISPR en una célula, tal como una célula receptora, se modifican (por ejemplo, por ingeniería genética) de modo que el(los) espaciador(es) CRISPR tiene(n) homología a uno o más espaciador(es) CRISPR en un organismo, tal como un organismo donante, que es sustancialmente resistente a un ácido nucleico diana o un producto de transcripción del mismo para volver a la célula resistente al ácido nucleico diana.

Adecuadamente, los uno o más genes *cas* o proteínas y uno o más, preferiblemente, dos o más repeticiones CRISPR en la célula son una combinación funcional como las descritas en la presente memoria.

45 La ingeniería genética puede incluir, pero no se limita a, añadir (por ejemplo, insertar), someter a delección (por ejemplo, separar) o modificar (por ejemplo, mutar) la secuencia de los uno o más espaciadores CRISPR o en una célula de modo que el espaciador CRISPR tenga homología (por ejemplo, homología incrementada después de la ingeniería genética) a uno o más espaciadores CRISPR de un organismo donante. Esta etapa de ingeniería dará como resultado una célula que será sustancialmente sensible a un ácido nucleico diana o un producto de transcripción del mismo al ser sustancialmente resistente al ácido nucleico diana o un producto de transcripción del mismo.

50 La ingeniería genética puede incluso incluir, pero no se limita a, añadir (por ejemplo, insertar) o someter a delección (por ejemplo, separar) la secuencia de los uno o más pseudo espaciadores CRISPR en una célula. Esta etapa de ingeniería dará como resultado una célula que será sustancialmente sensible a un ácido nucleico diana o un

producto de transcripción del mismo al ser sustancialmente resistente al ácido nucleico diana o un producto de transcripción del mismo.

En otra realización, “modular la resistencia” quiere decir suprimir, reducir o disminuir la resistencia de una célula a un ácido nucleico diana.

5 Por tanto, en un aspecto adicional, se ha proporcionado un método para disminuir o reducir la resistencia de una célula, tal como una célula receptora, que comprende al menos uno o más genes *cas* o proteínas y una o más, preferiblemente, dos o más repeticiones CRISPR frente a un ácido nucleico diana o un producto de transcripción del mismo.

10 De acuerdo con esta realización, el método comprende las etapas de: (i) identificar uno o más espaciadores CRISPR en un organismo que es sustancialmente resistente al ácido nucleico diana o un producto de transcripción del mismo; y (ii) modificar la secuencia de uno o más espaciador(es) CRISPR en la célula de modo que el(los) espaciador(es) CRISPR tenga(n) un grado reducido de homología a el(los) espaciador(es) CRISPR en el organismo.

15 En otra realización, se ha proporcionado un método para modular (por ejemplo, disminuir) la resistencia de una célula que comprende uno o más genes *cas* o proteínas y una o más, preferiblemente, dos o más repeticiones CRISPR frente a un ácido nucleico diana o producto de transcripción del mismo que comprende las etapas de: (i) identificar un espaciador CRISPR o un seudo espaciador CRISPR en un organismo que comprenda un ácido nucleico diana o producto de transcripción del mismo frente al cual se modula la resistencia; y (ii) identificar el espaciador CRISPR en el organismo en el cual se modula la resistencia; y (iii) adaptar la secuencia del espaciador CRISPR en el organismo en el cual se modula la resistencia de modo que el espaciador CRISPR tenga un grado inferior de homología al espaciador CRISPR o seudo espaciador CRISPR del organismo que comprende el ácido nucleico diana o producto de transcripción del mismo frente al cual se modula la resistencia.

20 Se modifican uno o más espaciadores CRISPR en una célula sustancialmente resistente por ingeniería para volver a la célula sensible a un ácido nucleico diana. La ingeniería genética puede incluir, pero no se limita a, la adición (por ejemplo, inserción), deleción (por ejemplo, separación) o modificación de una o más combinaciones funcionales repetición CRISPR-*cas* o porciones o fragmentos de los mismos en la célula sustancialmente resistente y/o la adición (por ejemplo, inserción), deleción (por ejemplo, separación) o modificación de uno o más espaciadores CRISPR o porciones o fragmentos de los mismos en la célula sustancialmente resistente.

25 A continuación, la etapa de ingeniería dará como resultado una célula que será sustancialmente resistente a un ácido nucleico diana o un producto de transcripción del mismo llegando a ser sustancialmente sensible a un ácido nucleico diana o un producto de transcripción del mismo.

Generalmente, para conferir sensibilidad a una célula, se espera que uno o más espaciadores CRISPR, uno o más genes *cas* o proteínas, una o más, preferiblemente, dos o más repeticiones CRISPR o una o más combinaciones funcionales repetición CRISPR-*cas* a partir de una célula sustancialmente resistente se separarán, someterán a deleción o modificarán de modo que no se confiere resistencia durante más tiempo.

35 Ventajosamente, las células que son sensibles a un ácido nucleico diana o un producto de transcripción del mismo se pueden preparar de modo que sus niveles dentro de un cultivo dado, tal como un cultivo iniciador, se puede modular (por ejemplo, disminuir) tal como se desee. Por tanto, a modo de ejemplo, se puede desarrollar un cultivo iniciador que comprenda dos o cepas bacterianas de modo que todos los miembros del cultivo son sensibles al mismo agente (por ejemplo, bacteriófago). Por consiguiente, una vez que no se desee que el cultivo esté vivo durante más tiempo, el cultivo se puede poner en contacto con el mismo agente sencillo para matar todos los miembros del cultivo.

Además, incluso puede ser posible modular la sensibilidad de una célula a uno o más agentes (por ejemplo, bacteriófago) de modo que el agente solamente mate una cierta proporción de las células en un cultivo dado, tal como el 10, 20, 30, 40, 50, 60, 70, 80, 90 ó 95% de las células en un cultivo dado.

45 En un aspecto, una célula, tal como una célula receptora, puede ser modificada por ingeniería de modo que comprenda un espaciador CRISPR o una secuencia que corresponde con un seudo espaciador CRISPR volviendo de ese modo a la célula resistente a un ácido nucleico diana o producto de transcripción del mismo. Adecuadamente, la célula se modifica por ingeniería de modo que el espaciador CRISPR o secuencia que corresponde al seudo espaciador CRISPR se usa junto con una combinación funcional gen *cas*-repetición CRISPR, tal como las descritas en la presente memoria.

50 En un aspecto, una célula que es resistente a un ácido nucleico diana o producto de transcripción del mismo se modifica por ingeniería de modo que el espaciador CRISPR que confiere la inmunidad frente al ácido nucleico diana o producto de transcripción del mismo se inserta en una célula que comprende una combinación funcional gen *cas*-

repetición CRISPR, volviendo de ese modo a la célula resistente al ácido nucleico diana o producto de transcripción del mismo.

5 En un aspecto, se determina la secuencia de uno o más espaciadores CRISPR o seudo espaciadores CRISPR de una célula que es resistente a un ácido nucleico diana o producto de transcripción del mismo. A continuación, se modifica por ingeniería una célula, tal como una célula receptora, de modo que comprenda la secuencia del espaciador CRISPR y una combinación funcional gen *cas*-repetición CRISPR, volviendo así a la célula resistente al ácido nucleico diana o producto de transcripción del mismo.

10 En un aspecto, se preparan un espaciador CRISPR de una célula, tal como una célula receptora, y una combinación funcional gen *cas*-repetición CRISPR de la misma o diferente célula, tal como la misma o diferente célula receptora. A continuación, se modifica por ingeniería una célula adicional, tal como una célula receptora, de modo que comprenda la secuencia de espaciador CRISPR y combinación funcional de gen *cas*-repetición CRISPR volviendo de ese modo a la célula resistente al ácido nucleico diana o producto de transcripción del mismo.

Un espaciador CRISPR está flanqueado por dos repeticiones CRISPR. En otras palabras, un espaciador CRISPR tiene al menos una repetición CRISPR a cada lado.

15 **Bacteriófago**

En un aspecto particularmente preferido de la presente invención, se modula la resistencia de una célula frente un bacteriófago.

El bacteriófago es virulento a la célula.

El bacteriófago puede ser un bacteriófago virulento o uno templado.

20 Tal como se usa en la presente memoria, el término “bacteriófago” tiene su significado convencional como se entiende en la técnica, es decir, un virus que infecta selectivamente procariontes, tales como bacterias. Muchos bacteriófagos son específicos a un género o especie particular o cepa de célula.

El bacteriófago puede ser un bacteriófago lítico o un bacteriófago lisogénico.

25 Un bacteriófago lítico es uno que sigue la ruta lítica hasta la finalización del ciclo lítico, más que entrar en la ruta lisogénica. Un bacteriófago lítico se somete a replicación viral que conduce a lisis de la membrana celular, destrucción de la célula, y liberación de las partículas del bacteriófago progenie capaces de infectar otras células.

Un bacteriófago lisogénico es uno capaz de entrar en la ruta lisogénica, en la cual el bacteriófago llega a ser una parte latente y pasiva del genoma de la célula hasta antes de la finalización de su ciclo lítico.

El término “bacteriófago” es sinónimo del término “fago”.

30 Aunque se puede emplear la resistencia frente a cualquier bacteriófago (incluyendo el bacteriófago de tipo natural, que se da de manera natural, aislado o recombinante), se prefiere el bacteriófago activo frente a bacterias. Más adecuadamente, es de particular interés el bacteriófago activo frente a bacterias que son patógenas a plantas y/o animales (incluyendo humanos).

35 A modo de ejemplo, el bacteriófago incluye, pero no se limita a, aquellos bacteriófagos capaces de infectar una bacteria que comprende de manera natural uno o más loci CRISPR. Los loci CRISPR se han identificado en más de 40 procariontes (Jansen et al. 2002b; Mojica et al. 2005; Haft et al., 2005) que incluyen, pero no se limitan a, *Aeropyrum*, *Pyrobaculum*, *Sulfolobus*, *Archaeoglobus*, *Halocarcula*, *Methanobacterium*, *Methanococcus*, *Methanosarcina*, *Methanopyrus*, *Pyrococcus*, *Picrophilus*, *Thermoplasma*, *Corynebacterium*, *Mycobacterium*, *Streptomyces*, *Aquifex*, *Porphyromonas*, *Chlorobium*, *Thermus*, *Bacillus*, *Listeria*, *Staphylococcus*, *Clostridium*, *Thermoanaerobacter*, *Mycoplasma*, *Fusobacterium*, *Azarcus*, *Chromobacterium*, *Neisseria*, *Nitrosomonas*, *Desulfovibrio*, *Geobacter*, *Myxococcus*, *Campylobacter*, *Wolinella*, *Acinetobacter*, *Erwinia*, *Escherichia*, *Legionella*, *Metliylococcus*, *Pasteurella*, *Photobacterium*, *Salmonella*, *Xanthomonas*, *Yersinia*, *Treponema* y *Thermotoga*.

40 A modo de ejemplo, el bacteriófago incluye, pero no se limita a, aquellos bacteriófagos capaces de infectar bacterias pertenecientes a los siguientes géneros: *Escherichia*, *Shigella*, *Salmonella*, *Erwinia*, *Yersinia*, *Bacillus*, *Vibrio*, *Legionella*, *Pseudomonas*, *Neisseria*, *Bordetella*, *Flelicobacter*, *Listeria*, *Agrobacterium*, *Staphylococcus*, *Streptococcus*, *Enterococcus*, *Clostridium*, *Corynebacterium*, *Mycobacterium*, *Treponema*, *Borrelia*, *Francisella*, *Brucella* y *Xanthomonas*.

50 A modo de ejemplo adicional, el bacteriófago incluye, pero no se limita a, aquellos bacteriófagos capaces de infectar (o someter a transducción) especies de bacterias de ácido láctico, una especie de *Bifidobacterium*, una especie de *Brevibacterium*, una especie de *Propionibacterium*, una especie de *Lactococcus*, una especie de *Streptococcus*, una

especie de *Lactobacillus* incluyendo el *Lactobacillus acidophilus*, especies de *Enterococcus*, especies de *Pediococcus*, una especie de *Leuconostoc* y especie de *Oenococcus*.

5 A modo de ejemplo adicional, el bacteriófago incluye, pero no se limita a, aquellos bacteriófagos capaces de infectar *Lactococcus lactis*, incluyendo *Lactococcus lactis* subsp. *lactis* y *Lactococcus lactis* subsp. *cremoris*, *Lactococcus lactis* subsp. *lactis* biovar *diacetylactis*, *Streptococcus thermophilus*, *Lactobacillus delbrueckii* subsp. *bulgaricus*, *Lactobacillus helveticus*, *Bifidobacterium lactis*, *Lactobacillus acidophilus*, *Lactobacillus casei*, *Bifidobacterium infantis*, *Lactobacillus paracasei*, *Lactobacillus salivarius*, *Lactobacillus plantarum*, *Lactobacillus reuteri*, *Lactobacillus gasserii*, *Lactobacillus johnsonii* o *Bifidobacterium longum*.

10 A modo de ejemplo adicional, los bacteriófagos incluyen, pero no se limitan a, aquellos bacteriófagos capaces de infectar cualquier bacteria fermentativa susceptible de alteración mediante infección de bacteriófago, incluyendo pero no limitado a procesos para la producción de antibióticos, aminoácidos y disolventes. Los productos producidos por fermentación que se sabe que han encontrado infección de bacteriófago, y las correspondientes bacterias de fermentación infectadas, incluyen queso Cheddar y cottage (*Lactococcus lactis* subsp. *lactis*, *Lactococcus lactis* subsp. *cremoris*), Yogurt (*Lactobacillus delbrueckii* subsp. *bulgaricus*, *Streptococcus thermophilus*), queso Suizo, (*S. thermophilus*, *Lactobacillus lactis*, *Lactobacillus helveticus*), queso Azul (*Leuconostoc cremoris*), queso Italiano (*L. bulgaricus*, *S. thermophilus*), Viili (*Lactococcus lactis* subsp. *cremoris*, *Lactococcus lactis* subsp. *lactis* biovar *diacetylactis*, *Leuconostoc cremoris*), Yakult (*Lactobacillus casei*), caseína (*Lactococcus lactis* subsp. *cremoris*), Natto (*Bacillus subtilis* var. *natto*), Vino (*Leuconostoc oenos*), Sake (*Leuconostoc mesenteroides*), Polimixina (*Bacillus polymyxa*), Colistina (*Bacillus colistrium*), Bacitracina (*Bacillus licheniformis*), ácido L-Glutámico (*Brevibacterium lactofermentum*, *Microbacterium ammoniaphilum*) y acetona y butanol (*Colstridium acetobutylicum*, *Clostridium saccharoperbutylacetonicum*).

15

20

Las bacterias preferidas son *S. thermophilus*, *L. delbrueckii* subsp. *bulgaricus* y/o *L. acidophilus*.

25 A modo de ejemplo adicional, los bacteriófagos incluyen, pero no se limitan a, aquellos bacteriófagos capaces de infectar una bacteria que comprende uno o más loci CRISPR heterólogos. La bacteria puede comprender uno o más loci CRISPR homólogos, uno o más genes *cas* heterólogos, una o más repeticiones CRISPR heterólogas y/o uno o más espaciadores CRISPR heterólogos.

Los bacteriófagos pueden incluir, pero no se limitan a, bacteriófagos que pertenecen a cualquiera de las siguientes familias de virus: Corticoviridae, Cystoviridae, Inoviridae, Leviviridae, Microvindae, Myoviridae, Podoviridae, Siphoviridae o Tectiviridae.

30 Para causar infección por bacteriófago de las células, se "infecta" una célula cuando se inyecta o transfiere su ácido nucleico en la célula, con el ácido nucleico del fago que existe independientemente del genoma de la célula. La infección puede conducir a la expresión (transcripción y traducción) del ácido nucleico del bacteriófago dentro de la célula y la continuación del ciclo de vida del bacteriófago. En el caso de bacteriófago recombinante, también se pueden expresar secuencias recombinantes dentro del genoma del fago, tal como los ácido nucleico informadores.

35 Se ha encontrado que las secuencias espaciador CRISPR en procariotas tienen con frecuencia similitudes significantes a una diversidad de moléculas ADN, tales como elementos genéticos (incluyendo, pero no limitado a, cromosomas, bacteriófagos, plásmidos conjugativos). Curiosamente, las células que llevan estos espaciadores CRISPR son incapaces de ser infectados por moléculas de ADN que contienen secuencias homólogas a los espaciadores (Mojica et al. 2005).

40 En el contexto de la presente invención, uno o más pseudo espaciadores particulares derivables o derivados a partir del ADN del bacteriófago o espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más pseudo espaciador(es) CRISPR se puede(n) añadir dentro de un locus CRISPR de una célula, tal como una célula receptora, para modular (por ejemplo, proporcionar) resistencia frente a un bacteriófago particular, previniendo así sustancialmente el ataque de fago.

45 Generalmente, las regiones particulares dentro del genoma de fago se pueden dirigir para preparar los pseudo espaciadores, tales como genes que codifican proteínas de especificidad a hospedante, que proporcionan reconocimiento particular fago-hospedante, tal como helicasas, primasa, proteínas estructurales de cabeza o cola, proteínas con un dominio conservado (por ejemplo, holin, lisina y otras) o secuencias conservadas entre genes importantes de fago.

50 Cualquier ácido nucleico que se origina a partir del genoma de fago puede conferir inmunidad frente al fago cuando se inserta, por ejemplo, entre dos repeticiones en un locus CRISPR activo. La inmunidad puede ser más "eficiente" si el espaciador CRISPR corresponde a una secuencia interna de un gen fago, e incluso más "eficiente" cuando este gen codifica proteínas "esenciales" (por ejemplo, el antireceptor).

5 Por consiguiente, en un aspecto adicional, se ha proporcionado un método para conferir resistencia a una célula (adecuadamente, una célula bacteriana) frente a un bacteriófago que comprende las etapas de: (a) proporcionar uno o más seudo espaciadores CRISPR a partir de al menos un bacteriófago; (b) identificar una o más combinaciones funcionales repetición CRISPR-cas en al menos una célula que es sustancialmente sensible al bacteriófago; y (c) modificar por ingeniería los uno o más loci CRISPR en la célula sustancialmente sensible de modo que comprenda uno o más seudo espaciadores CRISPR a partir de un bacteriófago o uno o más espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPR para volver a la célula resistente.

10 En un aspecto adicional, se ha proporcionado un método para conferir resistencia a una célula (adecuadamente, una célula bacteriana) frente a un bacteriófago que comprende las etapas de: (a) proporcionar uno o más seudo espaciadores CRISPR a partir de al menos un bacteriófago; (b) identificar una o más combinaciones funcionales repetición CRISPR-cas en al menos una célula que es sustancialmente sensible al bacteriófago; y (c) insertar uno o más seudo espaciadores CRISPR a partir del bacteriófago o uno o más espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPR en la célula sustancialmente sensible de modo que la célula se vuelva sustancialmente resistente al bacteriófago.

15 En un aspecto adicional, se ha proporcionado un método para modular el lisotipo de una célula que comprende las etapas de: (a) proporcionar uno o más seudo espaciadores CRISPR a partir de al menos un bacteriófago; (b) identificar una o más combinaciones funcionales repetición CRISPR-cas en al menos una célula que es sustancialmente sensible al bacteriófago; y (c) modificar por ingeniería a los uno o más loci CRISPR en la célula sustancialmente sensible de modo que comprenda uno o más seudo espaciadores de un bacteriófago o uno o más espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPR.

20 En un aspecto adicional, se ha proporcionado un método para modular el lisotipo de una célula que comprende las etapas de: (a) proporcionar uno o más seudo espaciadores CRISPR a partir de al menos un bacteriófago; (b) identificar una o más combinaciones funcionales repetición CRISPR-cas en al menos una célula que es sustancialmente sensible al bacteriófago; e (c) insertar uno o más seudo espaciadores de un bacteriófago o uno o más espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) al uno o más seudo espaciador(es) CRISPR en la célula sustancialmente sensible.

25 En un aspecto adicional, se ha proporcionado un método para conferir resistencia a una célula (adecuadamente, una célula bacteriana) frente a un bacteriófago que comprende las etapas de: (i) identificar un seudo espaciador CRISPR en un bacteriófago que comprenda un ácido nucleico diana o un producto de transcripción del mismo frente al cual se modula la resistencia; y (ii) modificar la secuencia del espaciador CRISPR de la célula de modo que el espaciador CRISPR de la célula tenga homología al seudo espaciador CRISPR del bacteriófago que comprende el ácido nucleico diana.

30 En un aspecto adicional, se ha proporcionado un método para conferir resistencia a una célula (adecuadamente, una célula bacteriana) frente a un bacteriófago que comprende las etapas de: (i) identificar un seudo espaciador CRISPR en un bacteriófago que comprenda un ácido nucleico diana o un producto de transcripción del mismo frente al cual se modula la resistencia; y (ii) modificar la secuencia del espaciador CRISPR de la célula de modo que el espaciador CRISPR de la célula tenga 100% de homología o identidad al seudo espaciador CRISPR del bacteriófago que comprende el ácido nucleico diana.

35 En un aspecto adicional, se ha proporcionado un método para modular el lisotipo de una célula bacteriana que comprende las etapas de: comprende las etapas de: (i) identificar un seudo espaciador CRISPR en un bacteriófago que comprenda un ácido nucleico diana o un producto de transcripción del mismo frente al cual se modula la resistencia; y (ii) modificar la secuencia del espaciador CRISPR de la célula de modo que el espaciador CRISPR de la célula tenga homología al seudo espaciador CRISPR del bacteriófago que comprende el ácido nucleico diana.

40 En un aspecto adicional, se ha proporcionado un método para modular el lisotipo de una célula bacteriana que comprende las etapas: (i) identificar un seudo espaciador CRISPR en un bacteriófago que comprenda un ácido nucleico diana o un producto de transcripción del mismo frente al cual se modula la resistencia; y (ii) modificar la secuencia del espaciador CRISPR de la célula de modo que el espaciador CRISPR de la célula tenga 100% de homología o identidad al seudo espaciador CRISPR del bacteriófago que comprende el ácido nucleico diana.

45 Adecuadamente, el espaciador CRISPR de la célula bacteriana tendrá 100% de homología o identidad a una secuencia, tal como un seudo espaciador CRISPR, en el bacteriófago que comprende el ácido nucleico diana.

Adecuadamente, el espaciador CRISPR de la célula bacteriana formará una parte componente de un locus CRISPR que comprende una combinación funcional repetición CRISPR-cas tal como se describe en la presente memoria.

50 Adecuadamente, el ácido nucleico diana o un producto de transcripción del mismo en el bacteriófago es una secuencia de ácido nucleico altamente conservada.

Adecuadamente, el ácido nucleico diana o producto de transcripción del mismo en el bacteriófago es un gen que codifica una proteína de especificidad a hospedante.

Adecuadamente, el ácido nucleico diana o producto de transcripción del mismo en el bacteriófago codifica una enzima que es esencial para la supervivencia, replicación o crecimiento del bacteriófago.

- 5 Adecuadamente, el ácido nucleico diana o producto de transcripción del mismo en el bacteriófago codifica una helicasa, una primasa, una proteína estructural de cabeza o cola, o una proteína con un dominio conservado (por ejemplo, holin, lisina y otras).

10 Ventajosamente, las células bacterianas se pueden preparar de acuerdo con la presente invención que tiene una "susceptibilidad reducida a la multiplicación o infección del bacteriófago". Tal como se usa en la presente memoria, este término se refiere a la bacteria que tiene una baja o no susceptibilidad a la multiplicación o infección del bacteriófago si se compara con la bacteria de tipo natural si se cultiva, en por ejemplo, un medio lácteo.

15 En una realización, el término "baja susceptibilidad a multiplicación de bacteriófago" se refiere al nivel de multiplicación del bacteriófago en una bacteria que está por debajo de un nivel, lo cual causará un efecto perjudicial a un cultivo en un periodo de tiempo dado. Dichos efectos perjudiciales sobre un cultivo incluyen, pero no se limitan a, no coagulación de la leche durante la producción de productos de leche fermentada (tales como yogurt o queso), bajada inadecuada o lenta del pH durante la producción de productos de leche fermentada (tales como yogurt o queso), maduración lenta de queso y deterioro de una textura de alimento al punto de ser poco apetitoso o inadecuado para el consumo humano.

20 Para un conjunto equivalente de condiciones de cultivo la susceptibilidad hacia un bacteriófago de una bacteria de la presente invención es, en comparación con la bacteria tipo natural, 100 veces inferior (eficiencia de formación de placa [EOP, del Inglés "Efficiency of plaquing"]= 10^{-2}), preferiblemente 1.000 veces inferior (EOP= 10^{-3}), preferiblemente 10.000 veces inferior (EOP= 10^{-4}), más preferiblemente 100.000 veces inferior (EOP= 10^{-5}). Preferiblemente, el nivel de multiplicación de bacteriófago en un cultivo se mide después de aproximadamente 14 horas de incubación del cultivo, más preferiblemente después de aproximadamente 12 horas, más preferiblemente después de aproximadamente 7 horas, más preferiblemente después de aproximadamente 6 horas, más preferiblemente después de aproximadamente 5 horas y más preferiblemente después de aproximadamente 4 horas.

30 En un aspecto adicional, se ha proporcionado un método para conferir sensibilidad a una célula (preferiblemente, una célula bacteriana) frente a un bacteriófago que comprende las etapas de: (a) proporcionar un seudo espaciador CRISPR a partir de al menos un bacteriófago; (b) identificar una o más combinaciones funcionales repetición CRISPR-cas en una célula que es sustancialmente resistente al bacteriófago; y (c) modificar por ingeniería los uno o más loci CRISPR en la célula sustancialmente sensible de modo que comprendan uno o más seudo espaciadores CRISPR o uno o más espaciador(es) CRISPR que es/son complementario(s) u homólogo(s) a los uno o más seudo espaciador(es) CRISPR que tienen un grado reducido de homología en comparación con los uno o más loci CRISPR en la célula sustancialmente resistente.

40 En un aspecto adicional, se ha proporcionado un método para modular (por ejemplo, reducir) el lisotipo de una célula (preferiblemente una célula bacteriana), que comprende uno o más genes cas o proteínas y una o más, preferiblemente, dos o más repeticiones CRISPR que comprenden las etapas de: (i) identificar un seudo espaciador CRISPR en un bacteriófago frente al cual se modula la resistencia; y (ii) modificar la secuencia del espaciador CRISPR de la célula de modo que el espaciador CRISPR de la célula tenga un grado reducido de homología al seudo espaciador CRISPR del bacteriófago que comprende el ácido nucleico diana.

45 En incluso un aspecto adicional, se ha proporcionado un método para modular (por ejemplo, reducir o disminuir) la resistencia de una célula bacteriana que comprende uno o más genes cas o proteínas y una o más, preferiblemente, dos o más repeticiones CRISPR frente a un bacteriófago que comprende las etapas de: (i) identificar uno o más seudo espaciadores CRISPR en un bacteriófago frente al cual se modula la resistencia; (ii) identificar un espaciador CRISPR en la célula bacteriana en la cual se modula la resistencia que es homólogo a el(los) seudo espaciador(es) CRISPR; y (iii) modificar la secuencia del espaciador CRISPR en la célula bacteriana en la cual la resistencia se modula de modo que el espaciador CRISPR tenga un grado inferior de homología a seudo espaciador(es) CRISPR del bacteriófago frente al cual se modula la resistencia.

50 Adecuadamente, el espaciador CRISPR de la célula tendrá un grado reducido de homología, tal como una reducción del 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90 ó 95% en homología en comparación con el(los) seudo espaciador(es) CRISPR de bacteriófago frente a los cuales se modula la resistencia.

Por lo tanto las células bacterianas se pueden preparar de acuerdo con la presente invención que tiene una "susceptibilidad incrementada a la multiplicación de bacteriófago". Tal como se usa en la presente memoria, este

término se refiere a la bacteria que tiene una susceptibilidad incrementada o alta a la multiplicación de bacteriófago cuando se compara con la bacteria de tipo natural cuando se cultiva, en por ejemplo, un medio lácteo.

En una realización, el término “alta susceptibilidad a la multiplicación de bacteriófago” se refiere al nivel de multiplicación de bacteriófago en una bacteria que está por encima de un nivel, el cual causaría un efecto perjudicial a un cultivo en un periodo de tiempo dado. Tales efectos perjudiciales sobre un cultivo incluyen, pero no se limitan a, no coagulación de la leche durante la producción de productos de leche fermentada (tales como yogurt o queso), disminución inadecuada o lenta del pH durante la producción de los productos de leche fermentada (tal como yogurt o queso), maduración lenta del queso y deterioro de una textura de alimento al punto de ser poco apetitoso o inadecuado para el consumo humano. Para un conjunto equivalente de condiciones de cultivo la susceptibilidad hacia un bacteriófago de una bacteria de la presente invención es, en comparación con la bacteria tipo natural, 100 veces superior, 1.000 veces superior, 10.000 veces superior, o 100.000 veces superior ($EOP=10^{-5}$). El nivel de multiplicación de bacteriófago en un cultivo se mide después de aproximadamente 14 horas de incubación del cultivo, más preferiblemente después de aproximadamente 12 horas, más preferiblemente después de aproximadamente 7 horas, más preferiblemente después de aproximadamente 6 horas, más preferiblemente después de aproximadamente 5 horas y en una realización altamente preferidas después de aproximadamente 4 horas.

Un espaciador CRISPR está flanqueado por dos repeticiones CRISPR. En otras palabras, un espaciador CRISPR tiene al menos una repetición CRISPR a cada lado.

Bacterias

En una realización adicional, la secuencia de ácido nucleico diana o un producto de transcripción del mismo puede ser o puede ser derivable (preferiblemente, derivada) a partir de una o más bacterias. Por consiguiente, se puede modular la resistencia de una célula, por ejemplo una célula bacteriana, frente a bacterias o un componente de las mismas.

La secuencia de nucleótidos diana puede ser o puede ser derivada a partir de un gen que es o está asociado con la resistencia a la transferencia de plásmido en bacterias. De acuerdo con esta realización de la presente invención, se modifican uno o más espaciadores CRISPR en la célula de modo que el espaciador CRISPR de la célula tiene homología al espaciador CRISPR y/o seudo espaciador CRISPR contenido en el ADN plásmido de la célula bacteriana para así proporcionar resistencia frente a el(los) plásmido(s) particular(es), previniendo así la transferencia de ADN exógeno en la célula. Específicamente, se pueden dirigir regiones particulares dentro del ADN plásmido para proporcionar inmunidad frente a ADN plásmido, tales como las secuencias dentro de origen de replicación de los plásmidos o secuencias dentro de los genes que codifican las proteínas de replicación.

Por tanto, de acuerdo con este aspecto, el método comprende las etapas de: (i) identificar un espaciador CRISPR y/o seudo espaciador CRISPR derivable (preferiblemente, derivado) a partir del ADN plásmido de una célula bacteriana frente al cual se modula la resistencia; (ii) modificar la secuencia de un espaciador CRISPR en la célula en la cual se modula la resistencia de modo que el espaciador CRISPR de la célula tenga homología al espaciador CRISPR y/o seudo espaciador CRISPR contenido en el ADN plásmido de la célula bacteriana.

En aún un aspecto adicional, se ha proporcionado un método adicional para conferir resistencia a una célula frente a la transferencia de plásmido que comprende las etapas de: (a) identificar un espaciador CRISPR y/o seudo espaciador CRISPR derivable (preferiblemente, derivado) a partir del ADN plásmido; (b) identificar una o más combinaciones funcionales repetición CRISPR-gen *cas* en una célula que es sustancialmente sensible al plásmido; y (c) modificar por ingeniería los uno o más loci CRISPR en la célula sustancialmente sensible de modo que comprendan uno o más espaciadores CRISPR y/o seudo espaciadores CRISPR a partir del plásmido para volver a la célula resistente.

La secuencia de nucleótidos diana puede ser o puede ser derivada a partir de un gen que es o está asociado con la resistencia a uno o más elementos genéticos móviles. Los espaciadores CRISPR particulares y/o seudo espaciadores CRISPR derivables (preferiblemente, derivados) a partir de uno o más elementos genéticos móviles se pueden añadir dentro un locus CRISPR de una célula para proporcionar resistencia frente a elementos genéticos móviles, tales como elementos transponibles y secuencias de inserción, previniendo así la transferencia de ADN exógeno y la deriva genética. Específicamente, las regiones particulares dentro de los transposones y las secuencias de inserción se pueden dirigir para proporcionar inmunidad frente a elementos genéticos móviles. Por ejemplo, las dianas pueden incluir transposones conjugativos (Tn916, transposones clase II (Tn501), secuencias de inserciones (IS26) o genes transposasa.

Por tanto, de acuerdo con este aspecto, el método comprende las etapas de: (i) identificar un espaciador CRISPR y/o seudo espaciador CRISPR derivable (preferiblemente, derivado) a partir de uno o más elementos genéticos móviles de una célula frente al cual se modula la resistencia; y (ii) modificar la secuencia de un espaciador CRISPR en una célula en la cual se modula la resistencia de modo que el espaciador CRISPR y/o seudo espaciador CRISPR

de la célula tenga homología al espaciador CRISPR contenido en el(los) elemento(s) genético(s) móviles de la célula.

En aún un aspecto adicional, se ha proporcionado un método adicional para conferir resistencia a una célula frente a uno o más elementos genéticos móviles que comprenden las etapas de: (a) identificar un espaciador CRISPR y/o seudo espaciador CRISPR derivable (preferiblemente, derivado) a partir de uno o más elementos genéticos móviles; (b) identificar una o más combinaciones funcionales repetición CRISPR-cas en una célula que es sustancialmente sensible a los uno o más elementos genéticos móviles; y (c) modificar por ingeniería los uno o más loci CRISPR en la célula sustancialmente sensible de modo que comprendan o tengan homología a uno o más espaciadores CRISPR y/o seudo espaciadores CRISPR a partir de los uno o más elementos genéticos móviles para volver a la célula resistente.

La secuencia de nucleótidos diana puede ser o puede ser derivada a partir de un gen que es o está asociado con la resistencia a antibióticos. Por "antibiótico" se entiende una composición química o resto que disminuye la viabilidad o que inhibe el crecimiento o reproducción de microbios. Los genes de resistencia a antibiótico incluyen, pero no se limitan a *bla_{TEM}*, *bla_{ROB}*, *bla_{SHV}*, *aadB*, *aacC1*, *aacC2*, *aacC3*, *aacA4*, *mecA*, *vanA*, *vanH*, *vanX*, *satA*, *aacA-aphH*, *vat*, *vga*, *msrA* *sul* y/o int. Los genes de resistencia a antibiótico incluyen aquellos que son o son derivables (preferiblemente, derivados) a partir de especies bacterianas que incluyen pero no se limitan a los géneros *Escherichia*, *Klebsiella*, *Pseudomonas*, *Proteus*, *Streptococcus*, *Staphylococcus*, *Enterococcus*, *Haemophilus* y *Moraxella*. Los genes de resistencia a antibiótico también incluyen aquellos que son o son derivables (preferiblemente, derivados) a partir de especies bacterianas que incluyen pero no se limitan a *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Proteus mirabilis*, *Streptococcus pneumoniae*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Enterococcus faecalis*, *Staphylococcus saprophyticus*, *Streptococcus pyogenes*, *Haemophilus influenzae* y *Moraxella catarrhalis*.

Se pueden añadir espaciadores CRISPR particulares y/o seudo espaciadores CRISPR derivables (preferiblemente, derivados) a partir de genes que codifican resistencia a antibiótico dentro de un locus CRISPR de una célula, tal como una célula receptora, para prevenir la transferencia de genes que confieren resistencia a antibióticos en la célula, reduciendo así el riesgo de adquirir marcadores de resistencia a antibiótico. A modo de ejemplo, las dianas también pueden incluir *vanR*, (un gen que confiere resistencia a vancomicina), o *tetR*, un gen que confiere resistencia a tetraciclina, o dirige inhibidores beta lactamasa.

Por tanto, de acuerdo con este aspecto, el método comprende las etapas de: (i) identificar uno o más espaciadores CRISPR y/o seudo espaciadores CRISPR derivables (preferiblemente derivados) a partir de una célula que comprenda uno o más genes o marcadores de resistencia a antibiótico; y (ii) modificar la secuencia del espaciador CRISPR en una célula que no comprende o no expresa los genes o marcadores de resistencia a antibiótico de modo que el espaciador CRISPR de la célula tenga homología a los uno o más espaciadores CRISPR y/o seudo espaciadores CRISPR contenidos en la célula que comprende uno o más genes o marcadores de resistencia a antibióticos.

En aún un aspecto adicional, se ha proporcionado un método para modular la adquisición de marcadores de resistencia a antibiótico en una célula que comprende las etapas de: (a) identificar uno o más espaciadores CRISPR y/o seudo espaciadores CRISPR derivables (preferiblemente, derivados) a partir de una célula que comprende uno o más genes o marcadores de resistencia a antibiótico; (b) identificar uno o más loci CRISPR en una célula que no comprenden o no expresan los genes o marcadores de resistencia a antibiótico; y (c) modificar la secuencia del espaciador CRISPR en la célula que no comprende o no expresa los genes o marcadores de resistencia a antibiótico de modo que el espaciador CRISPR y/o seudo espaciadores CRISPR tenga homología al espaciador CRISPR contenido en la célula resistente a la transferencia de genes que confieren resistencia a uno o más antibióticos.

La secuencia de nucleótidos diana puede ser o puede ser derivada de un gen que es o está asociado con los genes que codifican los factores de virulencia. Los espaciadores CRISPR particulares y/o seudo espaciadores CRISPR derivables (preferiblemente, derivados) a partir de genes que codifican factores de virulencia se pueden añadir dentro de un locus CRISPR de bacteria para proporcionar resistencia frente a la transferencia de genes que confieren virulencia en la bacteria. Por ejemplo, se pueden dirigir los factores que comúnmente contribuyen a la virulencia en los patógenos microbianos, tales como toxinas, internalinas y hemolisinas.

Por tanto, de acuerdo con este aspecto, el método comprende las etapas de: (i) identificar uno o más espaciadores CRISPR y/o seudo espaciadores CRISPR derivables (preferiblemente, derivados) a partir de una célula que comprende uno o más factores de virulencia; y (ii) modificar la secuencia del espaciador CRISPR en una célula que no comprende o no expresa el(los) factor(es) o marcador(es) de virulencia de modo que el espaciador CRISPR de la célula tenga homología a los uno o más espaciadores CRISPR y/o seudo espaciadores CRISPR contenidos en la célula que comprende uno o más factores de virulencia.

5 En aún un aspecto adicional, se ha proporcionado un método adicional para conferir resistencia a una célula frente a uno o más factor(es) o marcador(es) de virulencia que comprende las etapas de: (a) identificar un espaciador CRISPR y/o seudo espaciador CRISPR derivable (preferiblemente, derivado) a partir de uno o más factor(es) o marcador(es) de virulencia; (b) identificar una o más combinaciones funcionales repetición CRISPR-cas en una célula que es sustancialmente sensible a los uno o más factor(es) o marcador(es) de virulencia; y (c) modificar por ingeniería los uno o más loci CRISPR en la célula sustancialmente sensible de modo que comprendan uno o más espaciadores CRISPR y/o seudo espaciadores CRISPR a partir de los uno o más factor(es) o marcador(es) de virulencia; para volver a la célula resistente.

10 Un espaciador CRISPR está flanqueado por dos repeticiones CRISPR. En otras palabras, un espaciador CRISPR tiene al menos una repetición CRISPR a cada lado.

Modificación

Las secuencias de ácido nucleico se pueden modificar por modificación por ingeniería genética secuencias de ácido nucleico.

Se puede modificar toda o parte de una secuencia de ácido nucleico.

15 Se puede modificar toda o parte de uno o más espaciadores CRISPR, genes *cas* o proteínas, repeticiones CRISPR o loci CRISPR.

Se pueden modificar espaciadores CRISPR, genes *cas* o proteínas, repeticiones CRISPR o loci CRISPR recombinantes.

20 Se pueden modificar espaciadores CRISPR, genes *cas* o proteínas, repeticiones CRISPR o loci CRISPR que se dan de manera natural.

Se pueden modificar genes *cas* o proteínas y repeticiones CRISPR que se dan conjuntamente de manera natural.

25 La ingeniería genética se puede mediar usando diversos métodos que se conocen en la técnica y generalmente incluirán métodos bien conocidos, tales como amplificación por PCR, clonación y mutagénesis dirigida a sitio. Las mutaciones se pueden introducir usando oligonucleótidos sintéticos. Estos oligonucleótidos contienen secuencias de nucleótidos que flanquean los sitios de mutación deseados. Un método adecuado está descrito en Morinaga et al., (Biotechnology (1984) 2, p646-649). Otro método de introducción de mutaciones en secuencias de nucleótidos codificadores de enzima está descrito en Nelson y Long ("Analytical Biochemistry" (1989), 180, p147-151). Un método adicional está descrito en Sarkar y Sommer (Biotechniques (1990), 8, p404-407-"The megaprimer method of site directed mutagenesis"). Los kits comercialmente disponibles ahora también están ampliamente disponibles para realizar mutagénesis dirigida a sitio. Los métodos de ingeniería genética están descritos en detalle en J. Sambrook, E. F. Fritsch, y T. Maniatis, 1989, "Molecular Cloning: A Laboratory Manual", Segunda Edición, Libros 1-3, Cold Spring Harbor Laboratory Press; Ausubel, F. M. et al. (1995 y suplementos periódicos; "Current Protocols in Molecular Biology", ch. 9, 13 y 16, John Wiley & Sons, New York, N.Y.); B. Roe, J. Crabtree, y A. Kahn, 1996, "ADN Isolation and Sequencing: Essential Techniques", John Wiley & Sons; M. J. Gait (Editor), 1984, "Oligonucleotide Synthesis: A Practical Approach". Irl Press; y D. M. J. Lilley y J. E. Dahlberg, 1992, "Methods of Enzymology DNA Structure Part A: Synthesis and Physical Analysis of DNA Methods in Enzymology", Academic Press.

30 La etapa de ingeniería genética incluso puede incluir métodos tales como recombinación homóloga que puede ser particularmente útil cuando, por ejemplo, los espaciadores CRISPR están siendo insertados o sometidos a delección.

40 La etapa de ingeniería genética incluso puede incluir la activación de una o más secuencias de ácido nucleico, tales como un loci CRISPR, repeticiones CRISPR, espaciadores CRISPR, genes *cas* o proteínas, combinaciones funcionales de genes *cas* o proteínas y repeticiones CRISPR o incluso combinaciones de las mismas.

Adecuadamente, se pueden insertar uno o más espaciadores CRISPR o seudo espaciadores CRISPR en al menos un locus CRISPR.

45 En una realización, la modificación no interrumpe uno o más genes *cas* del al menos un locus CRISPR. En otra realización, los uno o más genes *cas* se mantienen intactos.

En una realización, la modificación no interrumpe una o más repeticiones CRISPR de al menos un locus CRISPR. En una realización, las una o más repeticiones CRISPR se mantienen intactas.

Adecuadamente, se pueden insertar uno o más espaciadores CRISPR o seudo espaciadores CRISPR en o dentro de al menos un locus CRISPR.

Adecuadamente, se pueden insertar uno o más espaciadores CRISPR o seudo espaciadores CRISPR en el extremo 5' de al menos un locus CRISPR.

5 En una realización, la modificación comprende insertar al menos un espaciador CRISPR o seudo espaciadores CRISPR en una célula, tal como una célula receptora. En otra realización, la modificación comprende insertar uno o más espaciadores CRISPR o seudo espaciadores CRISPR en (por ejemplo, modificar o reemplazar) uno o más espaciadores CRISPR de una célula, tal como una célula receptora.

10 En una realización, la modificación comprende insertar al menos un espaciador CRISPR o seudo espaciador CRISPR de un organismo, tal como un organismo donante, en la célula. En otra realización, la modificación comprende insertar uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo, tal como un organismo donante, en (por ejemplo, para modificar o reemplazar) uno o más espaciadores CRISPR o seudo espaciadores CRISPR de una célula.

En una realización, se insertan (por ejemplo, para modificar o reemplazar) uno o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo, tal como un organismo donante, en uno o más espaciadores CRISPR o seudo espaciadores CRISPR de la célula.

15 En una realización, se insertan (por ejemplo, para modificar o reemplazar) uno o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo, tal como un organismo donante, en una o más, preferiblemente, dos o más repeticiones CRISPR de la célula. En esta realización de la invención, se prefiere que al menos una combinación funcional repetición CRISPR-cas se mantenga intacta en la célula.

20 En una realización, se insertan (por ejemplo, para modificar o reemplazar) uno o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo, tal como un organismo donante, en los mismos o diferentes espaciadores CRISPR de la célula.

25 En una realización, se insertan (por ejemplo, para modificar o reemplazar) uno o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo, tal como un organismo donante, adyacente a uno o más espaciadores CRISPR o seudo espaciadores CRISPR de la célula.

30 En el contexto de la presente invención, el término “adyacente” quiere decir “próximo a” en su sentido más amplio e incluye “directamente adyacente”. Por tanto, en una realización, uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo se pueden insertar “directamente adyacente” a uno o más espaciadores CRISPR o seudo espaciadores CRISPR de la célula, es decir, (el)los espaciador(es) CRISPR o seudo espaciador(es) CRISPR se inserta(n) de modo que no hay nucleótidos intermedios entre los espaciadores.

35 En otra realización, el(los) espaciador(es) CRISPR o seudo espaciador(es) CRISPR se insertan de modo que hay al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1.000, 10.000, 100.000 o incluso 1.000.000 o más nucleótidos intermedios entre los espaciadores.

40 En otra realización, el nucleótido intermedio se puede denominar una secuencia líder. Estos términos se usan intercambiamente en la presente memoria. La secuencia líder puede ser de una longitud diferente en bacterias diferentes. Adecuadamente, la secuencia líder es al menos 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400 ó 500 o más nucleótidos de longitud. Adecuadamente, la secuencia líder está entre el último gen cas (en el extremo 3') y la primera repetición CRISPR (en el extremo 5') del locus CRISPR.

En una realización, la secuencia líder puede ser entre aproximadamente 20-500 nucleótidos de longitud.

En una realización, se insertan uno o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo donante, adyacente a una o más, preferiblemente, dos o más repeticiones CRISPR de la célula.

45 En otra realización, se insertan uno o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo donante, adyacente a uno o más genes cas de la célula.

50 En otra realización, se insertan uno o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo donante, adyacente a los mismos o diferentes espaciadores de la célula receptora.

En otra realización, se insertan uno o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo donante, adyacente a las mismas o diferentes repeticiones CRISPR de la célula.

5 En otra realización, se insertan uno o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo donante, adyacente a los mismos o diferentes genes *cas* de la célula receptora.

10 En otra realización, se insertan dos o más espaciadores CRISPR o seudo espaciadores CRISPR, tal como dos o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo donante, adyacentes cada uno a los mismos o diferentes espaciadores CRISPR o seudo espaciadores CRISPR y/o repeticiones CRISPR y/o genes *cas* de la célula receptora.

En otra realización, se modifica la secuencia del espaciador CRISPR, tal como uno o más espaciadores CRISPR de un organismo donante, de la célula receptora de modo que el espaciador CRISPR tiene homología al espaciador CRISPR del organismo donante.

15 En otra realización, se modifica la secuencia del espaciador de la célula de modo que tiene homología al espaciador CRISPR o seudo espaciador CRISPR del organismo.

En una realización, el espaciador CRISPR tiene 100% de homología al espaciador CRISPR del organismo donante.

El(los) espaciador(es) CRISPR o seudo espaciadores CRISPR puede(n) comprender ADN o ARN de origen genómico, sintético o recombinante.

20 El(los) espaciador(es) CRISPR o seudo espaciadores CRISPR puede(n) ser de doble cadena o cadena sencilla si se representa la cadena sentido o antisentido o combinaciones de las mismas.

El(los) espaciador(es) CRISPR o seudo espaciadores CRISPR se puede(n) preparar mediante el uso de técnicas de ADN recombinante (por ejemplo, ADN recombinante), tal como se ha descrito en la presente memoria.

25 La modificación puede comprender la inserción de uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo, tal como un organismo donante, que es sustancialmente resistente a un ácido nucleico diana o un producto de transcripción del mismo en uno o más loci CRISPR de una célula sustancialmente sensible.

La modificación puede comprender la inserción de uno o más espaciadores CRISPR o seudo espaciadores CRISPR de un organismo, tal como un organismo donante, que es sustancialmente resistente a un ácido nucleico diana o un producto de transcripción del mismo en (por ejemplo, entre) una combinación funcional de al menos dos repeticiones CRISPR y al menos un gen *cas* en una célula sustancialmente sensible.

30 La modificación puede incluso comprender la modificación (por ejemplo, mutación) del ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, de modo que uno se crean o más genes *cas* en el ADN de la célula. A modo de ejemplo adicional, los genes *cas* se pueden clonar en una construcción, un plásmido o un vector y similares que, a continuación, se transforma en la célula, usando métodos tales como los descritos en la presente memoria.

35 La modificación puede incluso comprender la modificación (por ejemplo, mutación) del ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, de modo que se crean una o más, preferiblemente, dos o más repeticiones CRISPR en el ADN de la célula. A modo de ejemplo adicional, las repeticiones CRISPR se pueden clonar en una construcción, un plásmido o un vector y similares que, a continuación, se transforma en la célula, usando métodos tales como los descritos en la presente memoria.

40 La modificación puede incluso comprender la modificación (por ejemplo, mutación) del ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, de modo que se crean una o más combinaciones funcionales *cas*-repetición CRISPR en el ADN de la célula. A modo de ejemplo adicional, las combinaciones funcionales *cas*-repetición CRISPR se pueden clonar en una construcción, un plásmido o un vector y similares que, a continuación, se transforma en la célula, usando métodos tales como los descritos en la presente memoria.

45 La modificación puede incluso comprender la modificación (por ejemplo, mutación) del ADN de una célula (por ejemplo, una célula receptora), tal como ADN plásmido o ADN genómico, de modo que se crean uno o más espaciadores CRISPR en el ADN de la célula. A modo de ejemplo adicional, los espaciadores CRISPR se pueden clonar en una construcción, un plásmido o un vector y similares que, a continuación, se transforma en la célula, usando métodos tales como los descritos en la presente memoria.

50

En una realización, un espaciador CRISPR está flanqueado por dos repeticiones CRISPR. En otras palabras, un espaciador CRISPR tiene al menos una repetición CRISPR a cada lado.

5 Adecuadamente, la modificación comprende la inserción de uno o más espaciadores CRISPR (por ejemplo, espaciadores CRISPR heterólogos) en las inmediaciones de (por ejemplo, adyacente a, adecuadamente, directamente adyacente a) uno o más genes *cas* y/o secuencia líder. Adecuadamente, de acuerdo con esta realización de la presente invención, la organización del locus CRISPR que se da de manera natural se mantiene después de la inserción de los uno o más espaciadores CRISPR.

Grupo (“Cluster”)

10 Sorprendentemente, se ha encontrado que no es posible intercambiar simplemente combinaciones repetición CRISPR-*cas* entre cualquier célula (por ejemplo, cualquier cepa, especie o género de células) puesto que se cree que esto no necesariamente dará como resultado combinaciones funcionales repetición CRISPR-*cas*.

Además, para que las combinación(es) repetición CRISPR-*cas* sean funcionales deberían ser compatibles. Por consiguiente, se cree que no es posible cambiar genes *cas* o repeticiones CRISPR entre diferentes loci CRISPR al menos que sean del mismo grupo.

15 Incluso más sorprendentemente es que los grupos no siguen la filogenia de “organismo”. Específicamente, dentro un organismo, puede haber más de un CRISPR. Estos CRISPR(s) pueden pertenecer a diferentes grupos, incluso aunque estén presentes en el mismo organismo. Como consecuencia, se cree que una combinación funcional repetición CRISPR-*cas* requiere que la combinación se cambie dentro de un grupo en lugar de dentro de un organismo.

20 Para evitar la duda, el término “grupo” tal como se usa en la presente memoria no se refiere a un grupo de genes localizados en el mismo locus (generalmente formando un operón) sino al resultado a partir del análisis de comparación de secuencia, tal como el análisis de comparación de secuencia múltiple y/o alineamientos de secuencia múltiple y/o análisis de gráfico de puntos. Por consiguiente, el análisis de grupo de loci CRISPR se puede realizar usando diversos métodos que se conocen en la técnica, tal como el análisis de gráfico de puntos como se enseña a continuación en la presente memoria por ejemplo o alineamiento múltiple seguido de cálculo de dendrograma.

Ventajosamente, el uso de combinación(es) repetición CRISPR-*cas* que se da(n) conjuntamente de manera natural mantiene el intercambio de la combinación tanto dentro como entre una especie dada, haciendo de ese modo posible modificar por ingeniería la resistencia de una cepa que usa la combinación a partir de una cepa diferente.

30 El grupo puede ser una clase, una familia o un grupo de secuencias.

Determinar resistencia

En un aspecto adicional, se ha proporcionado un método para determinar el perfil de resistencia de una célula frente a un ácido nucleico diana. Tal como se usa en la presente memoria, el término “perfil de resistencia” quiere decir que una o más entidades frente a las cuales la célula es sensible o resistente. Por consiguiente, el perfil de resistencia de una célula puede ser que la célula sea resistente a un primer bacteriófago, sensible a un segundo bacteriófago, resistente a un primer elemento genético móvil y sensible a un primer gen de resistencia a antibiótico etc.

Se pueden detectar o secuenciar uno o más genes *cas* o proteínas, y/o una o más, preferiblemente, dos o más repeticiones CRISPR y/o uno o más espaciadores CRISPR, etc. dentro de una célula para predecir/determinar el probable perfil de resistencia de una célula particular.

40 Adecuadamente, se detectan o secuencian uno o más espaciadores CRISPR dentro de una células para predecir/determinar el probable perfil de resistencia de una célula particular.

Los métodos adecuados de detección pueden incluir PCR, hibridación ADN-ADN (o hibridación ADN-ARN, es decir, usando sondas de ADN o ARN que podrían ser sintéticas, oligonucleótidos etiquetados, por ejemplo). También se pueden usar micromatrices (“microarrays”) de ADN.

45 Se pueden detectar o secuenciar uno o más combinaciones funcionales *cas*-repetición CRISPR y/o uno o más espaciadores CRISPR dentro de una célula para predecir/determinar el probable perfil de resistencia de una célula particular. A modo de ejemplo, es posible predecir/determinar el probable perfil de resistencia de una célula bacteriana particular a uno o más bacteriófago que se puede usar como indicador de lisotipo para la selección microbiana.

Se pueden secuenciar uno o más genes Cas y/o una o más repeticiones CRISPR además de uno o más espaciadores CRISPR para verificar la compatibilidad de la combinación gen Cas-repetición CRISPR o incluso identificar nuevos pares de cas/repeticiones compatibles.

Célula receptora

5 Tal como se usa en la presente memoria, el término “célula receptora” se refiere a cualquier célula en la cual la resistencia frente a un ácido nucleico diana o un producto de transcripción del mismo es modulada o es para ser modulada.

En una realización, la célula receptora se refiere a cualquier célula que comprende el ácido nucleico recombinante de acuerdo con la presente invención.

10 La célula receptora puede comprender una o más, preferiblemente, dos o más repeticiones CRISPR y uno o más genes cas o proteínas. Adecuadamente, las repeticiones CRISPR y los genes cas o proteínas forman una combinación funcional en la célula receptora, tal como se describe en la presente memoria.

15 La célula receptora puede comprender una o más repeticiones CRISPR modificadas y/o uno o más genes cas o proteínas modificados. Adecuadamente, las repeticiones CRISPR modificadas y/o los genes cas o proteínas modificados forman una combinación funcional en la célula receptora, tal como se describe en la presente memoria.

La célula receptora puede comprender una o más repeticiones CRISPR modificadas por ingeniería genética y/o uno o más genes cas o proteínas modificados por ingeniería genética. Adecuadamente, las repeticiones CRISPR modificadas por ingeniería genética y/o los genes cas o proteínas modificados por ingeniería genética forman una combinación funcional en la célula receptora, tal como se describe en la presente memoria.

20 La célula receptora puede comprender una o más repeticiones CRISPR recombinantes y/o uno o más genes cas o proteínas recombinantes. Adecuadamente, las repeticiones CRISPR recombinantes y/o los genes cas o proteínas recombinantes forman una combinación funcional en la célula receptora, tal como se describe en la presente memoria.

25 La célula receptora puede comprender uno o más repeticiones CRISPR que se dan de manera natural y uno o más genes cas o proteínas que se dan de manera natural. Adecuadamente, la(s) repetición(es) CRISPR y/o (el)los gen(es) cas o proteínas forma(n) una combinación funcional.

Por “se dan de manera natural” queremos decir que se dan de manera natural en la naturaleza.

30 La célula receptora puede incluso comprender combinaciones de una o más repeticiones CRISPR modificadas, modificadas por ingeniería genética, recombinantes o que se dan de manera natural y uno o más genes cas o proteínas modificados, modificados por ingeniería genética, recombinantes o que se dan de manera natural. Adecuadamente, a los uno o más espaciador(es) CRISPR modificados, modificados por ingeniería genética, recombinantes o que se da(n) de manera natural o los uno o más genes cas o proteínas modificados, modificados por ingeniería genética, recombinantes o que se dan de manera natural forman una combinación funcional.

Adecuadamente, la célula receptora es una célula procariota.

35 Adecuadamente, la célula receptora es una célula bacteriana. Las células bacterianas adecuadas están descritas en la presente memoria.

40 La célula bacteriana se puede seleccionar entre una especie de bacterias de ácido láctico, una especie de *Bifidobacterium*, una especie de *Brevibacterium*, una especie de *Propionibacterium*, una especie de *Lactococcus*, una especie de *Streptococcus*, una especie de *Lactobacillus* incluyendo las especies de *Enterococcus*, especies de *Pediococcus*, una especie de *Leuconostoc* y especies de *Oenococcus*.

Las especies adecuadas incluyen, pero no se limitan a *Lactococcus lactis*, incluyendo *Lactococcus lactis subsp. lactis* y *Lactococcus lactis subsp. cremoris*, *Lactococcus lactis subsp. cremoris*, *Leuconostoc sp.*, *Lactococcus lactis subsp. lactis biovar*, *Streptococcus thermophilus*, *Lactobacillus delbrueckii subsp. bulgaricus* y *Lactobacillus helveticus*, *Bifidobacterium lactis*, *Lactobacillus acidophilus*, *Lactobacillus casei*.

45 La célula en la cual se modula la resistencia puede ser una célula bacteriana usada para la fermentación de carne (incluyendo, ternera, cerdo y aves de corral) incluyendo, pero no limitada a, bacterias de ácido láctico, *Pediococcus cerevisiae*, *Lactobacillus plantarum*, *Lactobacillus brevis*, especies de *Micrococcus*, *Lactobacillus sakei*, *Lactobacillus curvatus*, *Pediococcus pentosaceus*, *Staphylococcus xylosus* y *Staphylococcus vitulinus* y mezclas de las mismas (Food Biotechnology, 538-39 (D. Knorr Ed. 1987); C. Pederson, “Microbiology of Fermented Foods”, 210-34 (2º Ed. 1979); documento US 2.225.783).

50

5 La célula en la cual se modula la resistencia puede ser una célula bacteriana usada para la fermentación de vegetales (por ejemplo, zanahorias, pepinos, tomates, pimientos y repollo) incluyendo, pero no limitada a, *Lactobacillus plantatum*, *Lactobacillus brevis*, *Leuconostoc mesenteroides*, *Pediococcus pentosaceus* y mezclas de las mismas (Food Biotechnology, 540 (D. Knorr Ed. 1987); C. Pederson, "Microbiology of Fermented Foods", 153-209 (2º Ed. 1979); documentos US 3.024.116; US 3.403.032; US 3.932.674; y 3.897.307).

La célula en la cual se modula la resistencia puede ser una célula bacteriana usada para la fermentación de pasta formada a partir de cereales (por ejemplo, trigo, centeno, arroz, avenas, cebada y maíz).

La célula en la cual se modula la resistencia puede ser una célula bacteriana usada para la producción de vino. Generalmente, esto se alcanza mediante la fermentación de zumo de fruta, generalmente zumo de uva.

10 La célula en la cual se modula la resistencia puede ser una célula bacteriana usada para la fermentación de leche para producir queso, tal como *Lactobacillus delbrueckii subsp. bulgaricus*, *Lactobacillus helveticus*, *Streptococcus thermophilus*, *Lactococcus lactis subsp. lactis*, *Lactococcus lactis subsp. cremoris*, *Lactococcus lactis subsp. lactis biovar diacetylactis*, *Bifidobacteria* y *Enterococci* etc y mezclas de las mismas (Food Biotechnology, 530 (D. Knorr Ed. 1987); C. Pederson, "Microbiology of Fermented Foods", 135-51 (2º Ed. 1979)).

15 La célula en la cual se modula la resistencia puede ser una célula bacteriana usada para la fermentación de leche para producir queso, tal como *Lactobacillus bulgaricus*, *Lactobacillus helveticus*, *Streptococcus thermophilus*, *Lactococcus lactis subsp. lactis*, *Lactococcus lactis subsp. cremoris*, *Lactococcus lactis subsp. lactis biovar, Lactococci*, *Bifidobacteria* y *Enterococci* etc y mezclas de las mismas (Food Biotechnology, 530 (D. Knorr Ed. 1987); C. Pederson, Microbiology of Fermented Foods, 135-51 (2º ed. 1979)). La célula en la cual se modula la resistencia puede ser una célula bacteriana usada para la fermentación de huevo, tal como *Pediococcus pentosaceus*, *Lactobacillus plantarum* y mezclas de las mismas (Food Biotechnology, 538-39 (D. Knorr Ed. 1987)).

25 La célula en la cual se modula la resistencia puede ser una bacteria que comprende de manera natural uno o más loci CRISPR. Los loci CRISPR se han identificado en más de 40 procariontes (Jansen et al. 2002b; Mojica et al., 2005; Haft et al., 2005) incluyendo, pero no limitadas a *Aeropyrum*, *Pyrobaculum*, *Sulfolobus*, *Archaeoglobus*, *Halocarcula*, *Methanobacterium*, *Methanococcus*, *Methanosarcina*, *Methanopyrus*, *Pyrococcus*, *Picrophilus*, *Thermoplasma*, *Corynebacterium*, *Mycobacterium*, *Streptomyces*, *Aquifex*, *Porphyromonas*, *Chlorobium*, *Thermus*, *Bacillus*, *Listeria*, *Staphylococcus*, *Clostridium*, *Thermoanaerobacter*, *Mycoplasma*, *Fusobacterium*, *Azarcus*, *Chromobacterium*, *Neisseria*, *Nitrosomas*, *Desulfovibrio*, *Geobacter*, *Myxococcus*, *Campylobacter*, *Wolinella*, *Acinetobacter*, *Erwinia*, *Escherichia*, *Legionella*, *Methylococcus*, *Pasteurella*, *Photobacterium*, *Salmonella*, *Xanthomonas*, *Yersinia*, *Treponema* y *Thermotoga*.

30 La célula en la cual se modula la resistencia puede ser una bacteria para usarse en composiciones cosméticas o farmacéuticas. Tales composiciones pueden comprender un cultivo microbiano y/o bacteria etiquetada y/o un cultivo celular de acuerdo con la presente invención. Por tanto, el cultivo microbiano y/o bacteria etiquetada y/o un cultivo celular de acuerdo con la presente invención pueden ser compatibles en cosméticos o en farmacia o en terapia.

35 **Organismo donante**

En una realización, el término "organismo donante" se refiere a un organismo o célula a partir del cual la repetición CRISPR y/o gen *cas* y/o combinación(es) de los mismos y/o espaciadores CRISPR son derivables (preferiblemente, derivados). Estos pueden ser los mismos o diferentes.

40 En una realización, el término "organismo donante" se refiere a un organismo o célula a partir del cual las una o más, preferiblemente, dos o más repeticiones CRISPR y/o uno o más gen *cas* y/o combinación(es) de los mismos y/o espaciadores CRISPR son derivables (preferiblemente, derivados). Estos pueden ser los mismos o diferentes.

En una realización, el espaciador CRISPR o pseudo espaciador CRISPR es sintéticamente derivado.

En una realización, el organismo o célula donante comprende uno o más espaciadores CRISPR, los cuales confieren la especificidad de inmunidad frente a un ácido nucleico diana o producto de transcripción del mismo.

45 En una realización, el organismo o célula donante a partir del cual el gen *cas* y/o la repetición CRISPR y/o la combinación de los mismos es derivable (preferiblemente, derivado) es también la célula/organismo receptor del locus CRISPR recombinante. Estos pueden ser los mismos o diferentes.

50 En una realización, el organismo o célula donante a partir del cual el espaciador es derivable (preferiblemente derivado) es también la célula/organismo receptor para el locus CRISPR recombinante. Estos pueden ser los mismos o diferentes.

Si es el caso de que el organismo donante es una célula bacteriana entonces el organismo donante comprenderá generalmente un espaciador CRISPR que confiere la inmunidad específica frente al ácido nucleico diana o producto de transcripción del mismo.

El organismo puede ser una célula bacteriana o un bacteriófago.

- 5 Adecuadamente, el organismo es un bacteriófago.

Células hospedantes

Tal como se usa en la presente memoria, el término “célula hospedante” se refiere a cualquier célula que comprende la combinación, la construcción o el vector y similares de acuerdo con la presente invención.

- 10 Las células hospedantes se pueden transformar o transferir con una secuencia de nucleótidos contenida en un vector, por ejemplo, un vector de clonación. Dicha secuencia de nucleótido se puede llevar en un vector para la replicación y/o expresión de la secuencia de nucleótidos. Las células se elegirán para ser compatibles con dicho vector y pueden, por ejemplo, ser células procariotas (por ejemplo, bacterianas).

- 15 Aspectos de la presente invención también se refieren a células hospedantes que comprenden la combinación, construcción o el vector de la presente invención. La construcción o el vector pueden comprender una secuencia de nucleótidos para la replicación y expresión de la secuencia. Las células se elegirán para ser compatibles con el vector y pueden ser, por ejemplo, células procariotas (por ejemplo, bacterianas).

Construcción

En un aspecto adicional, se ha proporcionado una construcción que comprende una o más de las secuencias de ácido nucleico descritas en la presente memoria.

- 20 El término “construcción”, el cual es sinónimo de los términos tales como “conjugado”, “casete” e “híbrido”, incluye una secuencia de nucleótidos directamente o indirectamente acoplada a otra secuencia, tal como una secuencia reguladora (por ejemplo, un promotor). A modo de ejemplo, la presente invención cubre una construcción que comprende una secuencia de nucleótidos enlazada de un modo funcional a dicha secuencia reguladora. El término “enlazada de un modo funcional” se refiere a una yuxtaposición en la que los componentes descritos están en una
25 relación que permite funcionar en su manera deseada. Una secuencia reguladora “enlazada de un modo funcional” a una secuencia codificadora está ligada de tal manera que la expresión de la secuencia codificadora se alcanza bajo condición compatible con las secuencias control.

El término “secuencias reguladoras” incluye promotores y potenciadores (“enhancers”) y otras señales de regulación de expresión.

- 30 El término “promotor” se usa en el sentido normal de la técnica, por ejemplo, un sitio de unión a ARN polimerasa.

La construcción incluso puede contener o expresar un marcador, el cual permite la selección de la construcción de secuencia de nucleótidos, por ejemplo, una bacteria. Existen diversos marcadores que se pueden usar, por ejemplo, aquellos marcadores que proporcionan resistencia a antibiótico, por ejemplo resistencia a antibióticos bacterianos, tales como Eritromicina, Ampicilina, Streptomycin y Tetraciclina.

Vector

La construcción puede ser o estar incluida en un vector (por ejemplo, un plásmido).

Por tanto, en un aspecto adicional se ha proporcionado un vector que comprende una o más de las construcciones o secuencias descritas en la presente memoria.

El término “vector” incluye vectores de expresión y vectores de transformación y vectores de transporte (“shuttle”).

- 40 El término “vector de transformación” quiere decir una construcción capaz de ser transferida desde una entidad a otra entidad, la cual puede ser de la especie o puede ser de una especie diferente. Si la construcción es capaz de ser transferida desde una especie a otra entonces el vector de transformación a veces se denomina un “vector de transporte”.

- 45 Los vectores se pueden transformar en una célula adecuada (por ejemplo, una célula hospedante) tal como se describe a continuación.

Los vectores pueden ser, por ejemplo, vectores de plásmido o fago proporcionados con un origen de replicación, opcionalmente un promotor para la expresión del dicho polinucleótido y opcionalmente un regulador del promotor.

Los vectores pueden contener una o más secuencias de nucleótidos marcador seleccionable. Los sistemas de selección más adecuados para los microorganismos industriales son aquellos formados por el grupo de los marcadores de selección que no requieren una mutación en el organismo hospedante.

5 Los vectores se pueden usar *in vitro*, por ejemplo, para la producción de ARN o usar para someter a transfección o transformar una célula hospedante.

Por tanto, los polinucleótidos se pueden incorporar en un vector recombinante (generalmente un vector replicable), por ejemplo, un vector de clonación o expresión. El vector se puede usar para replicar el ácido nucleico en una célula hospedante compatible.

Transfección

10 La introducción de un ácido nucleico (por ejemplo, una construcción o vector) en una célula se puede efectuar mediante diversos métodos. Por ejemplo, se puede usar transfección de fosfato de calcio, transfección mediada por DEAE-dextrano, transfección mediada por lípido catiónico, electroporación, transducción o infección. Dichos métodos están descritos en muchos manuales de laboratorio estándar, tales como Sambrook et al., "Molecular Cloning: A Laboratory Manual", 2º ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

15 Se pueden seleccionar células que contienen el ácido nucleico (por ejemplo, una construcción o vector) usando, por ejemplo, Eritromicina para células sometidas a transfección con un ácido nucleico (por ejemplo, una construcción o vector) que lleva un marcador elegible a resistencia.

Transformación

20 Las enseñanzas sobre la transformación de las células están bien documentados en la técnica, por ejemplo, véase Sambrook et al. ("Molecular Cloning: A Laboratory Manual", 2º edición, 1989, Cold Spring Harbor Laboratory Press) y Ausubel et al., "Current Protocols in Molecular Biology" (1995), John Wiley & Sons, Inc.

Una célula se puede transformar con un ácido nucleico (por ejemplo, una construcción o vector). Las células transformadas con la secuencia de nucleótidos se pueden cultivar bajo condiciones adecuadas para la replicación o expresión de la secuencia de nucleótidos.

Introducción

En el contexto de la introducción de un ácido nucleico en una célula, en una realización se prefiere que el término "introducir" quiera decir uno o más de transformación, transfección, conjugación o transducción.

Cultivos iniciadores

30 Los cultivos iniciadores se usan ampliamente en la industria alimenticia en la producción de los productos fermentados que incluyen productos de leche, tales como yogurt y queso, productos de carne, productos de panadería, vino y productos vegetales.

35 Los cultivos iniciadores usados en la producción de muchos productos de leche fermentada, queso y mantequilla incluyen cultivos de bacterias, generalmente clasificadas como bacterias de ácido láctico. Dichos cultivos iniciadores bacterianos imparten características específicas a diversos productos lácteos mediante la realización de un número de funciones.

Los cultivos no concentrados comerciales de las bacterias se refieren en la industria como "cultivos madre", y se propagan en el sitio de producción, por ejemplo una lechería, antes de añadirse a un material de inicio comestible, tal como leche, para la fermentación. El cultivo iniciador propagado en el sitio de producción para la inoculación en un material de inicio comestible se denomina el "cultivo iniciador a granel".

40 Los cultivos iniciadores adecuados para usarse en la presente invención pueden incluir cualquier organismo que sea de uso en la industria alimenticia, cosmética o farmacéutica.

45 Por ejemplo, el cultivo iniciador puede ser adecuado para usarse en la industria láctea. Si se usa en la industria láctea el cultivo iniciador se puede seleccionar entre una especie de bacterias de ácido láctico, una especie de *Bifidobacterium*, una especie de *Brevibacterium*, una especie de *Propionibacterium*. Los cultivos iniciadores adecuados del grupo de bacterias de ácido láctico incluyen cepas comúnmente usadas de una especie de *Lactococcus*, una especie de *Streptococcus*, una especie de *Lactobacillus* incluyendo el *Lactobacillus acidophilus*, especies de *Enterococcus*, especies de *Pediococcus*, una especie de *Leuconostoc* y especie de *Oenococcus*.

Los cultivos de bacterias de ácido láctico comúnmente se usan en la producción de productos de leche fermentada, tales como suero de leche, yogurt o crema agria, y en la producción de mantequilla y queso, por ejemplo, Brie o

Harvati. Las especies de *Lactococcus* incluyen el ampliamente usado *Lactococcus lactis*, incluyendo *Lactococcus lactis subsp. lactis* y *Lactococcus lactis subsp. cremoris*.

5 Otras bacterias de ácido láctico incluyen *Leuconostoc sp.*, *Streptococcus thermophilus*, *Lactobacillus delbrueckii subsp. bulgaricus* y *Lactobacillus helveticus*. Además, la cepas probióticas, tales como especies de *Lactococcus*, incluyen el ampliamente usado *Lactococcus lactis*, incluyendo *Lactococcus lactis subsp. lactis* y *Lactococcus lactis subsp. cremoris*. Los cultivos mesófilos de bacterias de ácido láctico comúnmente usados en la producción de productos de leche fermentada tal como suero de leche, yogurt o crema agria, y en la producción de mantequilla y queso, por ejemplo, Brie o Harvati. Otras especies de *Lactococcus* incluyen *Lactococcus lactis subsp. cremoris*,
 10 *Lactococcus lactis*, *Leuconostoc sp.*, *Lactococcus lactis subsp. lactis biovar*, *Streptococcus thermophilus*, *Lactobacillus delbrueckii subsp. bulgaricus* y *Lactobacillus helveticus*. Además, las cepas probióticas tales como *Bifidobacterium lactis*, *Lactobacillus acidophilus*, *Lactobacillus casei* se pueden añadir durante dicha producción para aumentar el sabor o para promover la salud.

15 Los cultivos de bacterias de ácido láctico comúnmente usados en la producción de quesos cheddar y Monterey Jack incluyen *Streptococcus thermophilus*, *Lactococcus lactis subsp. lactis* y *Lactococcus lactis subsp. cremoris* o combinaciones de las mismas.

Los cultivos termófilos de la bacterias de ácido láctico comúnmente usados en la producción de quesos italianos tal como Pasta filata o parmesano, incluyen *Streptococcus thermophilus* y *Lactobacillus delbrueckii subsp. bulgaricus*. Otras especies de *Lactobacillus*, tales como *Lactobacillus helveticus*, se pueden añadir durante la producción para obtener un sabor deseado.

20 Ventajosamente, el organismo del cultivo iniciador puede comprender o consistir en una cepa genéticamente modificada (preparada de acuerdo con los métodos deseados en la presente memoria) de una de las anteriores cepas de bacterias de ácido láctico o cualquiera de otra cepa de cultivo iniciador.

25 La selección de los organismos para el cultivo iniciador de la invención dependerá del tipo particular de productos a preparar y tratar. Por tanto, por ejemplo, para la producción de queso y mantequilla, ampliamente se usan cultivos mesófilos de especies de *Lactococcus*, especies de *Leuconostoc* y especies de *Lactobacillus*, mientras que para yogurt y otros productos de leche fermentada, generalmente se usan cepas termófilas de especies de *Streptococcus* y de especies de *Lactobacillus*.

El cultivo iniciador puede ser incluso un cultivo iniciador seco.

El cultivo iniciador puede ser un cultivo iniciador concentrado.

30 El cultivo iniciador puede ser un cultivo iniciador concentrado usado en la inoculación directa.

El cultivo iniciador puede ser un cultivo iniciador congelado.

El cultivo iniciador puede consistir en una cepa bacteriana, es decir, un cultivo puro. En este caso, sustancialmente todo, o al menos una porción significativa del cultivo iniciador bacteriano comprendería generalmente la misma bacteria.

35 Como alternativa, el cultivo iniciador puede comprender diversas cepas bacterianas, es decir, un cultivo mezclado definido.

Bacterias de ácido láctico

Los cultivos iniciadores particularmente adecuados, en particular cultivos iniciadores secos, para usarse en la presente invención comprenden bacterias de ácido láctico.

40 Tal como se usa en la presente invención, el término "bacterias de ácido láctico" se refiere a bacterias Gram positivo, microaerófilas o anaeróbicas que fermentan el azúcar con la producción de ácidos que incluyen ácido láctico como el ácido predominantemente producido, ácido acético, ácido fórmico y ácido propiónico. La bacterias de ácido láctico industrialmente más útiles se encuentran entre las especies de *Lactococcus*, tales como *Lactococcus lactis*, especies de *Lactobacillus*, especies de *Bifidobacterium*, especies de *Streptococcus*, especies de *Leuconostoc*,
 45 especies de *Pediococcus* y especies de *Propionibacterium*.

Los cultivos iniciadores de la presente invención pueden comprender una o más especies de bacterias de ácido láctico tales como, *Lactococcus lactis*, *Lactobacillus delbrueckii subsp. bulgaricus* y *Streptococcus thermophilus* o combinaciones de las mismas.

50 Los cultivos iniciadores de bacterias de ácido láctico se usan comúnmente en la industria alimenticia como cultivos de cepa mezclada que comprenden una o más especies. Para un número de cultivos de cepa mezclada, tales como

cultivos iniciadores de yogurt que comprenden cepas de *Lactobacillus delbrueckii subsp. Bulgaricus* y *Streptococcus thermophilus*, existe una relación simbiótica entre las especies en las que la producción de ácido láctico es superior en comparación con los cultivos de bacterias de ácido láctico de cepa sencilla (Rajagopal et al. J. Dairy Sci., 73, p.894-899, 1990).

5 Preparación de cultivos iniciadores

Los cultivos iniciadores se pueden preparar mediante técnicas bien conocidas en la técnica tales como las descritas en el documento US 4.621.058. A modo de ejemplo, los cultivos iniciadores se pueden preparar mediante la introducción de un inóculo, por ejemplo una bacteria, a un medio de crecimiento para producir un medio inoculado y madurar el medio inoculado para producir un cultivo iniciador.

10 Preparación de cultivos iniciadores secos

Los cultivos iniciadores secos se pueden preparar mediante técnicas bien conocidas en la técnica, tales como las discutidas en los documentos US 4.423.079 y US 4.140.800.

Los cultivos iniciadores secos para usarse en la presente invención pueden estar en forma de preparaciones sólidas. Ejemplos de preparaciones sólidas incluyen, pero no se limitan a, comprimidos, píldoras, cápsulas, polvos, gránulos y en polvos que pueden ser humectables, secados por atomización, secado por congelación o liofilizado.

Los cultivos iniciadores secos para usarse en la presente invención pueden estar o bien en forma de píldora muy congelada o forma de polvo secado por congelación. Los cultivos iniciadores secos en forma de píldora muy congelada o polvo secado por congelación se pueden preparar de acuerdo con los métodos conocidos en la técnica.

Los cultivos iniciadores para usarse en la presente invención pueden estar en forma de concentrados que comprenden una concentración sustancialmente alta de una o más bacterias. Adecuadamente los concentrados se pueden diluir con agua o resuspender en agua u otros diluyentes adecuados, por ejemplo, un medio de crecimiento adecuado o aceites minerales o vegetales, para usarse en la presente invención. Los cultivos iniciadores secos de la presente invención en forma de concentrados se pueden preparar de acuerdo con los métodos conocidos en la técnica, por ejemplo por centrifugación, filtración o una combinación de tales técnicas.

25 Producto

Los productos adecuados para usarse en la presente invención incluyen, pero no se limitan a, comestibles, productos cosméticos o productos farmacéuticos.

Cualquier producto, el cual esté preparado a partir, o comprenda, un cultivo está contemplado de acuerdo con la presente invención. Estos incluyen, pero no se limitan a, frutas, legumbres, cultivos de forraje y vegetales incluyendo productos derivados, grano y productos derivados del grano, alimentos lácteos y productos derivados de alimentos lácteos, carne, aves de corral, marisco, productos cosméticos y farmacéuticos.

El término "alimento" se usa en un amplio sentido e incluye comidas, comestibles, ingredientes alimenticios, suplementos alimenticios y alimentos funcionales.

Tal como se usa el término "ingrediente alimenticio" incluye una formulación, la cual está o se puede añadir a alimentos e incluye formulaciones que se pueden usar a bajos niveles en una amplia diversidad de productos que requieren, por ejemplo, acidificación o emulsificación.

Tal como se usa en la presente memoria, el término "alimento funcional" quiere decir un alimento que es capaz de proporcionar no solamente un efecto nutricional y/o satisfacción de sabor, sino que también es capaz de liberar un efecto beneficioso adicional al consumidor. Aunque no hay definición legal de un alimento funcional, la mayoría de las partes con un interés en esta área están de acuerdo que hay alimentos comercializados como que tienen efectos específicos para la salud.

El término "alimento" cubre alimento para humanos así como alimento para animales (es decir, una comida). En un aspecto preferido, el alimento es para el consumo humano.

Las células descritas en la presente memoria pueden estar, o pueden añadirse a, un ingrediente alimenticio, un suplemento alimenticio o un alimento funcional.

El alimento puede estar en forma de una disolución o como un sólido, dependiendo del uso y/o el modo de aplicación y/o modo de administración.

Las células descritas en la presente memoria se pueden usar en la preparación de productos alimenticios tales como uno o más de: productos de confitería, productos lácteos, productos cárnicos, productos de ave de corral, productos de pescado y productos de panadería.

5 A modo de ejemplo, la bacteria se puede usar como ingredientes para refrescos, un zumo de fruta o una bebida que comprende proteína de suero, téis saludables, bebidas de cacao, bebidas de leche y bebidas de bacterias de ácido láctico, yogurt, yogurt bebible y vino.

También se ha proporcionado un método de preparación de un alimento, comprendiendo el método mezclar las células de acuerdo con la presente invención con un ingrediente alimenticio (tal como un material de inicio para un alimento). El método para preparar un alimento también es otro aspecto de la presente invención.

10 Adecuadamente un alimento tal como se describe en la presente memoria es un producto lácteo. Más preferiblemente un producto lácteo tal como se describe en la presente memoria es uno o más de los siguientes: un yogurt, un queso (tal como un requesón ácido, un queso fuerte, un queso semi fuerte, un queso cottage), un suero de leche, quark o crema agria, kéfir, una bebida basada en suero fermentado, un Kumis, una bebida de leche y una bebida de yogurt.

15 En la presente memoria, el término "alimento" se usa en un amplio sentido, y cubre alimento para humanos así como alimento para animales (es decir, comida). En un aspecto preferido, el alimento es para consumo humano.

20 El término comida tal como se usa en la presente memoria incluye material de planta y material de no planta en crudo y procesado. La comida puede ser cualquier comida adecuada para consumo por un animal, incluyendo comida de ganado (animal), por ejemplo comida de ave de corral, comida de pescado o comida de crustáceos por ejemplo.

Variantes/homólogos/derivados/fragmentos

25 La presente invención abarca el uso de variantes, homólogos, derivados y fragmentos de los mismos, incluyendo variantes, homólogos, derivados y fragmentos de loci CRISPR, espaciadores CRISPR, seudo espaciadores CRISPR, genes cas o proteínas, repeticiones CRISPR, combinaciones funcionales repetición CRISPR-gen cas y secuencias de ácido nucleico diana o productos de transcripción del mismo.

El término "variante" se usa para querer decir un polipéptido que se da de manera natural o secuencias de nucleótidos que difiere de una secuencia tipo natural.

30 El término "fragmento" indica que una secuencia de polipéptido o nucleótidos comprende una fracción de una secuencia de tipo natural. Puede comprender una o más secciones grandes continuas de secuencia o una pluralidad de pequeñas secciones. La secuencia también puede comprender otros elementos de secuencia, por ejemplo, puede ser una proteína de fusión con otra proteína. Preferiblemente la secuencia comprende al menos el 50%, más preferiblemente al menos el 65%, más preferiblemente al menos el 80%, más preferiblemente al menos el 85%, más preferiblemente al menos el 90%, más preferiblemente al menos el 95%, más preferiblemente al menos el 96%, más preferiblemente al menos el 97%, más preferiblemente al menos el 98%, lo más preferiblemente al menos el 99% de la secuencia tipo natural.

35 Preferiblemente, el fragmento retiene el 50%, más preferiblemente el 60%, más preferiblemente el 70%, más preferiblemente el 80%, más preferiblemente el 85%, más preferiblemente el 90%, más preferiblemente el 95%, más preferiblemente el 96%, más preferiblemente el 97%, más preferiblemente el 97%, más preferiblemente el 98% o lo más preferiblemente el 99% de actividad del polipéptido o secuencia de nucleótidos de tipo natural.

40 Preferiblemente, un espaciador CRISPR o seudo espaciador CRISPR comprende al menos el 50%, más preferiblemente al menos el 65%, más preferiblemente al menos el 80%, más preferiblemente al menos el 85%, más preferiblemente al menos el 90%, más preferiblemente al menos el 95%, más preferiblemente al menos el 96%, más preferiblemente al menos el 97%, más preferiblemente al menos el 98%, lo más preferiblemente al menos el 99% de la secuencia tipo natural. Preferiblemente, un espaciador CRISPR retiene el 50%, más preferiblemente el 60%, más preferiblemente el 70%, más preferiblemente el 80%, más preferiblemente el 85%, más preferiblemente el 90%, más preferiblemente el 95%, más preferiblemente el 96%, más preferiblemente el 97%, más preferiblemente el 98% o lo más preferiblemente el 99% de la actividad del polipéptido o secuencia de nucleótidos de tipo natural.

45 Preferiblemente, un gen cas comprende al menos el 50%, más preferiblemente al menos el 65%, más preferiblemente al menos el 80%, más preferiblemente al menos el 85%, más preferiblemente al menos el 90%, más preferiblemente al menos el 95%, más preferiblemente al menos el 96%, más preferiblemente al menos el 97%, más preferiblemente al menos el 98%, lo más preferiblemente al menos el 99% de la secuencia tipo natural. Preferiblemente, un gen cas retiene el 50%, más preferiblemente el 60%, más preferiblemente el 70%, más preferiblemente el 80%, más preferiblemente el 85%, más preferiblemente el 90%, más preferiblemente el 95%, más

preferiblemente el 96%, más preferiblemente el 97%, más preferiblemente el 98% o lo más preferiblemente el 99% de la actividad del polipéptido o secuencia de nucleótidos de tipo natural.

5 Preferiblemente, una proteína Cas comprende al menos el 50%, más preferiblemente al menos el 65%, más preferiblemente al menos el 80%, más preferiblemente al menos el 85%, más preferiblemente al menos el 90%, más preferiblemente al menos el 95%, más preferiblemente al menos el 96%, más preferiblemente al menos el 97%, más preferiblemente al menos el 98%, lo más preferiblemente al menos el 99% de la secuencia tipo natural. Preferiblemente, una proteína Cas retiene el 50%, más preferiblemente el 60%, más preferiblemente el 70%, más preferiblemente el 80%, más preferiblemente el 85%, más preferiblemente el 90%, más preferiblemente el 95%, más preferiblemente el 96%, más preferiblemente el 97%, más preferiblemente el 98% o lo más preferiblemente el 99% de la actividad del polipéptido o secuencia de nucleótidos de tipo natural.

10 Preferiblemente, una repetición CRISPR comprende al menos el 50%, más preferiblemente al menos el 65%, más preferiblemente al menos el 80%, más preferiblemente al menos el 85%, más preferiblemente al menos el 90%, más preferiblemente al menos el 95%, más preferiblemente al menos el 96%, más preferiblemente al menos el 97%, más preferiblemente al menos el 98%, lo más preferiblemente al menos el 99% de la secuencia tipo natural. Preferiblemente, una repetición CRISPR retiene el 50%, más preferiblemente el 60%, más preferiblemente el 70%, más preferiblemente el 80%, más preferiblemente el 85%, más preferiblemente el 90%, más preferiblemente el 95%, más preferiblemente el 96%, más preferiblemente el 97%, más preferiblemente el 98% o lo más preferiblemente el 99% de la actividad del polipéptido o secuencia de nucleótidos de tipo natural.

15 Preferiblemente, una combinación funcional repetición CRISPR-cas comprende al menos el 50%, más preferiblemente al menos el 65%, más preferiblemente al menos el 80%, más preferiblemente al menos el 85%, más preferiblemente al menos el 90%, más preferiblemente al menos el 95%, más preferiblemente al menos el 96%, más preferiblemente al menos el 97%, lo más preferiblemente al menos el 99% de la secuencia tipo natural. Preferiblemente, la combinación funcional repetición CRISPR-cas retiene el 50%, más preferiblemente el 60%, más preferiblemente el 70%, más preferiblemente el 80%, más preferiblemente el 85%, más preferiblemente el 90%, más preferiblemente el 95%, más preferiblemente el 96%, más preferiblemente el 97%, más preferiblemente el 98% o lo más preferiblemente el 99% de la actividad del polipéptido o secuencia de nucleótidos de tipo natural.

20 Preferiblemente, una secuencia de ácido nucleico diana comprende al menos el 50%, más preferiblemente al menos el 65%, más preferiblemente al menos el 80%, más preferiblemente al menos el 85%, más preferiblemente al menos el 90%, más preferiblemente al menos el 95%, más preferiblemente al menos el 96%, más preferiblemente al menos el 97%, más preferiblemente al menos el 98%, lo más preferiblemente al menos el 99% de la secuencia tipo natural. Preferiblemente, una secuencia de ácido nucleico diana retiene el 50%, más preferiblemente el 60%, más preferiblemente el 70%, más preferiblemente el 80%, más preferiblemente el 85%, más preferiblemente el 90%, más preferiblemente el 95%, más preferiblemente el 96%, más preferiblemente el 97%, más preferiblemente el 98% o lo más preferiblemente el 99% de la actividad del polipéptido o secuencia de nucleótidos de tipo natural.

El fragmento puede ser un fragmento funcional.

40 Por un "fragmento funcional" de una molécula se entiende un fragmento que retiene o pese sustancialmente la misma actividad biológica que la molécula intacta. En todos los ejemplos, un fragmento funcional de una molécula retiene al menos 10%, y al menos 25%, 50%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ó 99% de la actividad biológica de la molécula intacta.

El término "homólogo" quiere decir una entidad que tiene una cierta homología con las secuencias de aminoácidos objeto y las secuencias de nucleótidos objeto. En la presente memoria, el término "homología" se puede igualar con la "identidad".

45 En el presente contexto, una secuencia homóloga se supone que incluye una secuencia de aminoácidos, la cual puede ser al menos 75, 85 ó 90% idéntica, preferiblemente al menos 95%, 96%, 97%, 98% ó 99% idéntica a la secuencia objeto. Aunque la homología también se puede considerar en términos de similitud (es decir, residuos de aminoácidos que tienen propiedades/funciones químicas similares), en el contexto de la presente invención se prefiere expresar homología en término de identidad de secuencia.

50 En el presente contexto, una secuencia homóloga se supone que incluye una secuencia de nucleótidos, la cual puede ser al menos 75, 85 ó 90% idéntica, preferiblemente al menos 95%, 96%, 97%, 98% ó 99% idéntica a la secuencia objeto. Aunque la homología también se puede considerar en términos de similitud (es decir, residuos de aminoácidos que tienen propiedades/funciones químicas similares), en el contexto de la presente invención se prefiere expresar homología en término de identidad de secuencia.

Las comparaciones de homología se pueden conducir a ojo, o más normalmente, con la ayuda de programas de comparación de secuencia fácilmente disponibles. Estos programas informáticos comercialmente disponibles pueden calcular el % de homología entre dos o más secuencias.

5 El % de homología se puede calcular sobre secuencias contiguas, es decir, una secuencia se alinea con la otra secuencia y cada aminoácido en una secuencia se compara directamente con el aminoácido correspondiente en la otra secuencia, un residuo a la vez. Esto se llama un alineamiento "sin huecos" ("ungapped"). Generalmente, tales alineamientos sin huecos se realizan solamente sobre un número relativamente corto de residuos.

10 Aunque esto es un método muy simple y consecuente, no logra tener en consideración que, por ejemplo, en un par de secuencias a pesar de ser idéntico, una inserción o delección causará que los siguientes residuos de aminoácidos se saque fuera del alineamiento, dando así potencialmente como resultado en una gran reducción en % de homología cuando se realiza un alineamiento global. Por lo tanto, la mayoría de los métodos de comparación de secuencia se diseñan para producir alineamientos óptimos que toman en consideración posibles inserciones y delecciones sin penalizar excesivamente el resultado completo de homología. Esto se alcanza mediante la inserción de huecos ("gaps") en el alineamiento de secuencia para tratar maximizar la homología local.

15 Sin embargo, estos métodos más complejos asignan "penalizaciones por hueco" ("gap penalties") a cada hueco que se da en el alineamiento de manera que, para el mismo número de aminoácidos idénticos, un alineamiento de secuencia con tan pocos huecos como sea posible, reflejando mayor relación entre las dos secuencias comparadas, alcanzará un mayor resultado que una con muchos huecos. "Los costes de hueco afín" generalmente se usan que carga un coste relativamente alto para la existencia de un hueco y una penalización menor para cada residuo posterior en el hueco. Esto es el sistema de registro de hueco comúnmente usado. Por su puesto las penalizaciones de hueco altos producirán alineamientos optimizados con menos huecos. La mayoría de los programas de alineamiento permiten las penalizaciones de hueco a modificar. Sin embargo, se prefiere usar los valores por defecto cuando se usa dicho programa informático para comparaciones de secuencia. Por ejemplo, si se usa el paquete GCG Wisconsin Bestfit la penalización de hueco por defecto para las secuencias de aminoácidos es -12 para un hueco y -4 para cada extensión.

20 Por lo tanto, el cálculo del % máximo de homología en primer lugar requiere la producción de un alineamiento óptimo., tomando en consideración las penalizaciones de hueco. Un programa informático adecuado para llevar a cabo dicho alineamiento es el paquete GCG Wisconsin Bestfit (Universidad de Wisconsin, U.S.A.; Devereux et al., 1984, "Nucleic Acids Research" 12:387). Ejemplos de otro programa informático que puede realizar las comparaciones de secuencia incluyen, pero no se limitan a, el paquete BLAST (véase Ausubel et al., 1999 ibid, Capítulo 18), FASTA (Atschul et al., 1990, J. Mol. Biol., 403-410), la serie GENWORKS de herramientas de comparación y CLUSTAL. Tanto BLAST como FASTA están disponibles para la búsqueda desconectada ("offline") y conectada ("online") (véase Ausubel et al., 1999 ibid, páginas 7-58 a 7-60). Sin embargo, para algunas aplicaciones, se prefiere usar el programa GCG Bestfit. Una nueva herramienta, denominada Secuencias BLAST 2 también está disponible para comparar secuencia de proteína y nucleótido (véase FEMS Microbiol Lett 1999 174(2): 274-50; FEMS Microbiol Lett 1999 177(1): 187-8).

30 Aunque el % final de homología se puede medir en términos de identidad, el propio proceso de alineamiento generalmente no se basa en una comparación de par todo-o-nada. En cambio, generalmente se usa una matriz de resultado de similitud graduada que asigna resultados a cada comparación por parejas basado en la similitud química o distancia evolutiva. Un ejemplo de dicha matriz comúnmente usada es la matriz BLOSUM62, la matriz por defecto para la serie BLAST de programas. Los programas GCG Wisconsin generalmente usan o bien los valores por defecto públicos o una tabla de comparación de símbolo a medida si se suministra (véase manual del usuario para más detalles). Para algunas aplicaciones, se prefiere usar los valores por defecto públicos para el paquete GCG, o en el caso de otro programa informático, la matriz por defecto, tal como BLOSUM62.

45 Una vez que el programa informático ha producido un alineamiento óptimo, es posible calcular el % de homología, preferiblemente el % de identidad de secuencia. El programa informático generalmente hace esto como parte de la comparación de secuencia y genera un resultado numérico.

Las penalizaciones de gap se deberían usar si se determina la identidad de secuencia, entonces adecuadamente se usan los siguientes parámetros:

50

PARA BLAST	
HUECO ABIERTO	0
EXTENSIÓN HUECO	0

PARA CLUSTAL	ADN	PROTEÍNA	
TAMAÑO PALABRA	2	1	Triple K
PENALIZACIÓN HUECO	10	10	
HUECO EXTENSIÓN	0,1	0,1	

5 Para la comparación de secuencia de polipéptido se pueden usar los siguientes ajustes: penalización de creación de HUECO de 0,3 y penalización de extensión de HUECO de 0,1. Adecuadamente, el grado de identidad con respecto a una secuencia de aminoácidos se determina sobre al menos 5 aminoácidos contiguos, se determina sobre al menos 10 aminoácidos contiguos, sobre al menos 15 aminoácidos contiguos, sobre al menos 20 aminoácidos contiguos, sobre al menos 30 aminoácidos contiguos, sobre al menos 40 aminoácidos contiguos, sobre al menos 50 aminoácidos contiguos o sobre al menos 60 aminoácidos contiguos.

10 Las secuencias también pueden tener deleciones, inserciones o sustituciones de residuos de aminoácidos, lo cual produce un cambio mudo y da como resultado una sustancia funcionalmente equivalente. Se pueden realizar sustituciones deliberadas de aminoácidos en base a la similitud en polaridad, carga, solubilidad, hidrofobicidad, hidrofiliidad y/o la naturaleza anfifática de los residuos siempre que se retenga la actividad de unión secundaria de la sustancia. Por ejemplo, los aminoácidos negativamente cargados incluyen ácido aspártico y ácido glutámico; los aminoácidos cargados positivamente incluyen lisina y arginina; y los aminoácidos con grupos de cabeza polares no cargados que tienen valores similares de hidrofiliidad incluyen leucina, isoleucina, valina, glicina, alanina, asparagina, glutamina, serina, treonina, fenilalanina y tirosina.

15 Las sustituciones conservadoras se pueden realizar, por ejemplo, de acuerdo con la Tabla de a continuación. Los aminoácidos en el mismo bloque en la segunda columna y adecuadamente en la misma línea en la tercera columna se pueden sustituir entre sí:

ALIFÁTICO	No polar	GAP
		ILV
	Polar-no cargado	CSTM
		NQ
	Polar-cargado	DE
		KR
AROMÁTICO		HFWY

20 La presente invención también abarca que la sustitución homóloga (sustitución y reemplazo se usan ambas en la presente memoria para querer decir que el intercambio de un residuo de aminoácidos existente, con un residuo alternativo) puede darse, es decir, sustitución entre dos cosas parecidas, tal como básico por básico, ácido por ácido, polar por polar, etc. La sustitución no homóloga también se puede dar, es decir, entre una clase de residuo y otra o implicar alternativamente la inclusión de aminoácidos no naturales, tales como ornitina (más adelante referida como Z), ornitina de ácido diaminobutírico (más adelante referido como B), ornitina de norleucina (más adelante referida como O), pirilalanina, tienilalanina, naftilalanina y fenilglicina.

30 También se pueden realizar reemplazos mediante aminoácidos no naturales que incluyen; aminoácidos alfa* y alfa-disustituído*; aminoácidos N-alquilo*, ácido láctico*, derivados de haluro de aminoácidos naturales, tales como trifluorotirosina*, p-Cl-fenilalanina*, p-Br-fenilalanina*, p-I-fenilalanina*, L-alil-glicina*, β-alanina*, ácido L-α-aminobutírico*, ácido L-γ-aminobutírico*, ácido L-α-amino isobutírico*; ácido L-ε-amino caproico*, ácido 7-amino heptanoico*, L-metionina sulfona*, L-norleucina*, L-norvalina*, p-nitro-L-fenilalanina*, L-hidroxiprolina*, L-tioprolina*, derivados de metilo de fenilalanina (Phe)-tales como 4-metil-Phe*, pentametil-Phe*, L-Phe(4-amino)*, L-Tyr(metil)*, L-Phe(4-isopropilo)*, L-Tic(ácido 1,2,3,4-tetrahidroisoquinolina-3-carboxilo)*, ácido L-diaminopropionico* y L-Phe (4-bencilo)*. La anotación * se ha utilizado con el propósito de la anterior discusión (en relación a sustitución homóloga

o no homóloga), para indicar la naturaleza hidrófoba del derivado mientras que # se ha utilizado para indicar la naturaleza hidrófila del derivado, #* indica características anfifáticas.

5 Las secuencias de aminoácidos variantes pueden incluir grupos de espaciador adecuados que se pueden insertar entre cualquiera de los dos residuos de aminoácidos de la secuencia que incluye grupos alquilo, tales como grupos metilo, etilo o propilo, además de los espaciadores de aminoácido, tales como residuos de glicina o β-alanina. Una forma adicional de variación implica que la presencia de uno o más residuos de aminoácidos en forma peptoide será bien entendido por los expertos en la técnica. Para evitar la duda, “la forma peptoide” se usa para referirse a residuos de aminoácidos variantes en los que el grupo sustituto de α-carbono está sobre el átomo de nitrógeno del residuo en lugar del carbono α. Los procesos para la preparación de péptidos en la forma peptoide son conocidos en la técnica, por ejemplo, Simon RJ et al., PNAS (1992) 89(20), 9.367-9.371 y Horwell DC, Trends Biotechnol. (1995) 13(4), 132-134.

15 Las secuencias de nucleótidos para usarse en la presente invención pueden incluir dentro de ellas nucleótidos sintéticos o modificados. Se conocen en la técnica un número de tipos diferentes de modificación a oligonucleótidos. Estos incluyen metil-fosfonato y cadenas principales de fosforotioato y/o la adición de cadenas de acridina o polilisina en los extremos 3' y 5' de la molécula. Con los propósitos de la presente invención, se entiende que las secuencias de nucleótidos se pueden modificar mediante cualquier método disponible en la técnica. Dichas modificaciones se pueden llevar a cabo para aumentar la actividad *in vivo* o vida útil de las secuencias de nucleótidos útiles en la presente invención.

Técnicas de metodología de ADN recombinante generales

20 La presente invención emplea, al menos que se indique lo contrario, técnicas convencionales de química, biología molecular, microbiología, ADN recombinante e inmunología, las cuales están dentro de las capacidades de una persona de conocimiento normal en la técnica. Dichas técnicas se explican en la bibliografía. Véase, por ejemplo, J. Sambrook, E.F. Fritsch y T. Maniatis, 1989, “Molecular Cloning. A Laboratory Manual” segunda Edición, Libros 1-3, Cold Spring Harbor Laboratory Press; Ausubel, F.M. et al. (1995 y suplementos periódicos; “Current Protocols in Molecular Biology”, ch. 9, 13 y 16, John Wiley & Sons, New York, N.Y.); B. Roe, J. Crabtree y A. Kahn, 1996, “DNA Isolation and Sequencing: Essential Techniques”, John Wiley & Sons; M.J. Gait (Editor), 1984, “Oligonucleotide Synthesis: A Practical Approach”, Irl Press; y D.M.J. Lilley y J.E. Dahlberg, 1992, “Methods of Enzymology: DNA Structure Part A: Synthesis and Physical Analysis of DNA Methods in Enzymology”, Academic Press. Cada uno de estos textos generales está en la presente memoria incorporado como referencia.

30 Ahora se describirá más la invención por medio de Ejemplos, los cuales quiere decir que sirven para ayudar a un experto en la técnica para llevar a cabo la invención y no intentan de ningún modo limitar el alcance de la invención

Ejemplos

Ejemplo 1

35 Inserción de un espaciador específico a fago en un CRISPR funcional existente para proporcionar resistencia al fago correspondiente

Cepa-*Streptococcus thermophilus* ST0089

Fago-2972

40 *Streptococcus thermophilus* ST0089 es una cepa industrialmente importante usada en la producción de yogurt, es genéticamente flexible a manipulación, y susceptible al fago 2972 virulento. Recientemente se determinó la secuencia del genoma completo para el fago 2972.

Los loci CRISPR se determinan en la cepa ST0089. Esto se determina preferentemente mediante la secuenciación del genoma entero de ST0089. Alternativamente, se identifican los loci CRISPR vía PCR usando conjuntos de cebador con secuencias idénticas a elementos CRISPR de *S. thermophilus* previamente identificados.

45 Una vez identificados, se determina la secuencia de loci CRISPR así como las regiones proximales que deberían contener los genes *cas* relevantes.

Al menos se selecciona un locus de CRISPR-*cas* particular para manipulación adicional. La funcionalidad de este locus se establece a través de análisis *silico* de las regiones de espaciador y sus homologías a secuencias de ADN de fago (es decir, la ausencia y/o presencia de las secuencias de espaciador y la correlación a infectividad de fago con la cepa ST0089). En ausencia de esta correlación, la funcionalidad se supone basada en la presencia de todos los elementos documentados (es decir, repeticiones, espaciadores, secuencias líder, genes *cas*, supuestamente codifican proteínas de longitud completa).

Se eligen una(s) secuencia(s) de espaciador adecuada(s) a partir del genoma del fago 2972. Los criterios del espaciador seleccionado se basa en: 1) longitud de los espaciadores dentro del locus CRISPR seleccionado; 2) aproximadamente 100% de identidad a la secuencia del fago; 3) teóricamente se puede seleccionar cualquier secuencia del fago.

- 5 En el ejemplo más simple, se sintetiza químicamente una unidad CRISPR que consiste en una secuencia de espaciador del fago 2972, flanqueada por dos elementos de repetición (idénticos al locus CRISPR seleccionado). Por definición esta "unidad CRISPR" es aproximadamente 100 pb de longitud y es demasiado corta para la consiguiente integración en el locus CRISPR.

- 10 Por lo tanto, se construye ADN flanqueante adicional junto con la unidad CRISPR. Un mínimo de 500 pb de ADN homólogo, idéntico al locus CRISPR dirigido flanquea la unidad CRISPR sintética, para facilitar la integración.

Hay al menos dos enfoques. Una construcción emula la adición de un nuevo espaciador sobre el CRISPR existente. Alternativamente, se reemplaza el locus CRISPR entero con la unidad CRISPR sintética.

El integrante CRISPR resultante se verifica a través de la secuenciación de ADN del locus CRISPR antes del ensayo biológico.

- 15 Se ensaya los patrones de sensibilidad a fago del integrante CRISPR frente al fago 2972 y se compara con la cepa parental.

El integrante CRISPR construido demuestra con éxito la correlación directa entre la presencia de espaciador específico dentro del propio contexto del CRISPR-cas.

Ejemplo 2

- 20 Se inserta un homólogo espaciador a un ADN de fago en una célula, tal como una célula receptora. La célula llega a ser resistente al fago. En un locus CRISPR dentro de la cepa seleccionada, se diseña un nuevo espaciador CRISPR a partir del ADN del fago (con el 100% de identidad al ADN del fago) dentro del gen anti-receptor y se inserta en la célula. El gen antireceptor se dirige debido a que se ha encontrado que los espaciadores CRISPR a partir de otras cepas muestran similitud a genes antireceptor de fago. Cuatro cepas que portan espaciadores que muestran la
25 identidad a los genes antireceptor de fago son resistente al fago particular. El mutante se expone al fago y llega a ser resistente a él.

Ejemplo 3

- 30 Se prepara un plásmido que comprende un espaciador CRISPR, y mostramos que este plásmido no se puede transferir en una célula que contiene el mismo espaciador, mientras que el plásmido sin el espaciador se puede transformar en la célula.

Ejemplo 4

Se inserta un espaciador en un hospedante original, pero no en un locus CRISPR, y el mutante resultante retiene su sensibilidad al fago, mostrando que el espaciador necesita estar en un ambiente particular dentro de un CRISPR y genes cas.

- 35 **Ejemplo 5**

Se inserta una combinación repetición CRISPR-cas completa en una célula, tal como una célula receptora, para proporcionar inmunidad frente al ácido nucleico entrante.

Ejemplo 6

- 40 Para una combinación particular repetición CRISPR-cas presente en dos cepas diferentes, el "intercambio" de espaciadores modifica sus fenotipos (sensibilidad/resistencia al fago).

Ejemplo 7

Se someten a delección uno o más genes cas (a partir de una unidad CRISPR-cas funcional). Los genes cas son necesarios para que se proporcione la inmunidad. Los mutantes Cas son aún sensibles al fago, a pesar de la presencia del espaciador idéntico al ADN de fago.

- 45 **Ejemplo 8**

Se clonan los genes cas sometidos a delección sobre un plásmido. Es posible proporcionar los genes cas en trans al hospedante. Si se bloquea el gen cas, se puede restaurar la inmunidad.

Ejemplo 9

Se preparan diferentes combinaciones cas-repetición CRISPR. No solamente se requieren los genes cas o proteínas, sino también, se requieren para la funcionalidad pares específicos cas-repetición CRISPR. Cuando se proporcionan los genes cas o proteínas a partir de otro locus CRISPR, la cepa se mantiene sensible al fago.

5 **Ejemplo 10**

Cuando se somete a delección un espaciador CRISPR particular a partir de un locus CRISPR que se da de manera natural, esto quita inmunidad frente a un fago dado y el hospedante llega a ser sensible (pérdida resistencia) al fago al cual el espaciador es homólogo.

Ejemplo 11

10 La integración de un espaciador CRISPR en el locus CRISPR de una bacteria proporciona resistencia frente a bacteriófago al que el espaciador CRISPR muestra identidad.

(A) Cepa DGCC7710RH1 de *Streptococcus thermophilus*

Streptococcus thermophilus

15 La cepa DGCC7710 de *Streptococcus thermophilus* (depositada en la "Collection Nationale de Cultures de Microorganismes" Francés bajo el número CNCM I-2423) posee al menos 3 loci CRISPR: CRISPR1, CRISPR2 y CRISPR3. En las cepas CNRZ1066 y LMG18311 para las cuales se conoce la secuencia completa de genoma (Bolotin et al., 2004), CRISPR1 se localiza en el mismo locus cromosómico: entre str0660 (o stu0660) y str0661 (o stu0661).

20 En la cepa DGCC7710, CRISPR1 también se localiza en el mismo locus cromosómico, entre genes altamente similares. CRISPR1 de la cepa DGCC7710 contiene 33 repeticiones (incluyendo la repetición terminal) y por tanto 32 espaciadores.

Todos estos espaciadores son diferentes entre sí. La mayoría de estos espaciadores son nuevos (no todavía descritos dentro del loci CRISPR), pero cuatro espaciadores cerca del tráiler CRISPR1 son idénticos a los espaciadores CRISPR1 ya conocidos:

- 25 - El 28º espaciador de DGCC7710 es 100% idéntico al 31º espaciador CRISPR1 de la cepa CNRZ1575 (Genbank número de entrada DQ072991);
- El 30º espaciador de DGCC7710 es 100% idéntico al 27º espaciador CRISPR1 de la cepa CNRZ703 (Genbank número de entrada DQ072990);
- 30 - El 31º espaciador de DGCC7710 es 100% idéntico al 28º espaciador CRISPR1 de la cepa CNRZ703 (Genbank número de entrada DQ072990);
- El 32º espaciador de DGCC7710 es 100% idéntico al 30º espaciador CRISPR1 de la cepa CNRZ703 (Genbank número de entrada DQ072990).

Bacteriófago virulento

35 D858 es un bacteriófago que pertenece a la familia *Siphoviridae* de virus. Su secuencia de genoma se ha determinado completamente pero todavía no está publicada. Este fago es virulento a la cepa DGCC7710 de *S. thermophilus*.

Mutante resistente a fago

La cepa DGCC7710RH1 de *Streptococcus thermophilus* se ha aislado como un mutante resistente a fago natural usando DGCC7710 como cepa parental, y el fago D858 como fago virulento.

40 CRISPR1 de la cepa DGCC7710-RH1 contiene 34 repeticiones (incluyendo la repetición terminal), y por tanto 33 espaciadores. Cuando se compara con la secuencia CRISPR1 de la cepa DGCC7710 de *Streptococcus thermophilus*, la secuencia CRISPR1 de la cepa DGCC7710-RH1 de *Streptococcus thermophilus* posee un nuevo espaciador adicional (y por supuesto una repetición adicional la cual flanquea el nuevo espaciador) en un extremo del locus CRISPR (es decir, cerca de líder, en el extremo 5' del locus CRISPR).

45 Todos los otros espaciadores del locus CRISPR1 están sin cambios.

La secuencia CRISPR1 (5'-3') de la cepa DGCC7710-RH1 es:

>CRISPR1_DGCC7710-RH1

caaggacagttattgattttataatcaactatgtgggtataaaaacgtcaaaatttcatttgag
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtcaacattgcaacatottataaaccactt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtgtttgacagcaaatcaagattogaattgt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaatsgacaggaggetatggcacaacttaca
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcagatttgacaatctgctgaccactgttate
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcacaacttggcagggttattactcaacagoga
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACctgttcoettgttcttttgttqtatctttt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtctattcttccgtttttgtttgogaactct
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgcttggcagaggaaacgaacaaggcctcaaca
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcatagagtggaaaactagaacagattcaa
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACataatggcgttgaattacaaggcaagggtca..
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgagcagagctcgaataatcttaattacaag
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgcttccgttagcgtcatgtggttaacgtattta
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgagcgtcccaatcctgatttaactactctog
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcaacacagcaagcacaaggagatgatgetatg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcgacacaagaacgtagcaagagttcaag
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcaaatcttctcaccgtaactgtctcaagt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcaataagggcatagaaagggagacaacatg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcgataattttaaatacttctcaactctcat
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgcagtagcagcaagcagctgttagttact
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACataaactatgaaatttataatttttaaga
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaataatttatggtagcttaataatcattg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtgcacagagcaagcttccaggttaccgttcc
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtctatatcagaggtcaactaacaattatgct
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaatcgttcaaatctctgttttaggtacattt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaatcaatacagcaagagttaaaatggtctt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgcttagctgttccaatccaagaaagctggatg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcaaccaacggtaacagctacttttacagt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACataactgaaggataggagcttynaaagtct
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtaactgatcactctcaaggatgatcccaga
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaagtagttagtgacctctacaatggtttat
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACacctagaagcatttgagcgtatattgattg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaatttgcccctctttgccccttgactag
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaccattagcaatcatttgtgcccattgagt
ttgattcaacataaaaagccagttcaattgaaacttggcttt

Leyenda

Secuencia líder:

5' caaggacagttattgattttataatcaactatgtgggtataaaaacgtcaaaatttcatttgag 3'

- 5 Secuencia integrada que comprende una Repetición CRISPR en el caso superior y un espaciador CRISPR (es decir, secuencia de etiquetado) en el caso inferior.

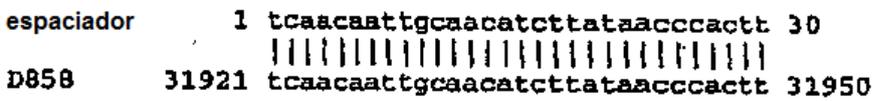
Repeticiones CRISPR

Repetición terminal: 5' gttttgtactctcaagatttaaglaactgtacagt 3'

Secuencia tráiler: 5'ttgattcaacataaaaagccagttcaattgaaacttggcttt3'

- 10 La secuencia del nuevo espaciador existe dentro del genoma del fago D858 y está representada en la presente memoria como SEC. ID. N° 534.

La secuencia del espaciador se encuentra entre las posiciones 31921 y 31950 pb (es decir, en la cadena positiva) del genoma D858's (y tiene el 100% de identidad a la secuencia genómica D858 sobre 30 nucleótidos):



El nuevo espaciador que está integrado en el locus CRISPR1 de la cepa DGCC7710-RH1 de *Streptococcus thermophilus* confiere a esta cepa resistencia al fago D858, tal como se representa en la Figura 5 y Tabla 1.

(B) Cepa DGCC7710-RH2 de *Streptococcus thermophilus*

5 La cepa DGCC7710-RH2 de *Streptococcus thermophilus* se ha aislado como un mutante resistente a fago natural usando la cepa DGCC7710 de *Streptococcus thermophilus* como cepa parental, y el fago D858 como fago virulento.

10 CRISPR1 de la cepa DGCC7710-RH2 de *Streptococcus thermophilus* contiene 34 repeticiones (incluyendo la repetición terminal) y por tanto 33 espaciadores. Cuando se compara con la secuencia CRISPR1 de la cepa DGCC7710 de *Streptococcus thermophilus*, la secuencia CRISPR1 de la cepa DGCC7710-RH2 de *Streptococcus thermophilus* posee un nuevo espaciador adicional (y por supuesto una repetición adicional que flanquea el nuevo espaciador) en un extremo del locus CRISPR (es decir, cerca de líder, en el extremo 5' del locus CRISPR). Todos los otros espaciadores del locus CRISPR1 están sin cambios.

La secuencia CRISPR1 (5'-3') de la cepa DGCC7710-RH2 es:

```
>CRISPR1_DGCC7710-RH2
caaggacagttattgattttataatcactatgtgggtataaaaacgtcaaaatttcatttgag
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACttacgtttgaaaagaatatcaaatcaatga
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtgtttgacagcaaatcaagattcgaattgt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACCaatgacgaggagctattggcacaacttaca
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcgatttgacaatctgctgaccactgttatc
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACacacttggcaggttattactcaacagcga
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACctgttccttgttcttttgttgtatctttc
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtccattcttccgtttttgtttgcaatcct
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgctggcgaggaaacgaacaaggcctcaaca
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcatagagtggaanaactagaacagattcaa
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACataatgcccgttgaattacacggcaagggtca
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgagcggagctcgaataatcttaattacaag
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgcttcgctagcgtcatgtggttaacgtatcta
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACggcgttcccaatcctgattaactactactcg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaacacagcaagacaagaggatgatgctatg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcgacacaagaacgtagcaagagttcaag
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACacaattcttcatccggttaaptgctcaagtg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaatlaagggcatabaaagggagacacaacatg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcgatatttaaaatcattttcataacttcat
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgagcagtagcagcaagcaagctgttagttact
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACataaactatgaaatttataattttaaga...
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaataatttatggtatagcttaatatcattg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtgcatcgagcaggttcgagttaccggttc
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtctatctcgaggtcaactaacaattatgct
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaatcgttcaaatctgcttttaggtacattt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaatcaatagcaagagttaaaatggctctt
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACgcttagctgtccaatccacgaacgtggatg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACataactgaaggataggagcttgtaagctct
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACtaatgctacatctcaaaggatgatccaga
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaagtagttgatgacctctacaatggtttat
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACacctagaagcatttgagcgtatatatgattg
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACcaatttggcccttctttggcccttgactag
GTTTTGTACTCTCAAGATTTAAGTAACTGTACAACaccatttagcaatcatttgtgccattgagtt
ttgattcaacataaaaagccaggttcaattgaacttggtttt
```

Leyenda

15 Secuencia líder:

5' caaggacagttattgattttataatcactatgtgggtataaaaacgtcaaaatttcatttgag 3'

La secuencia integrada que comprende una Repetición CRISPR en el caso superior y un espaciador CRISPR (es decir, secuencia de etiquetado) en el caso inferior.

Repeticiones CRISPR

Repetición terminal:

5 5' gttttgtactctcaagatttaagtaactgtacagt 3'

Secuencia tráiler:

5' ttgattcaacataaaaagccagttcaattgaactggctt3'

Se ha mostrado que la secuencia del nuevo espaciador existe dentro del genoma del fago D858.

10 La secuencia del espaciador (representado en la presente memoria como SEC. ID. N° 535) se encuentra entre las posiciones 17215 y 17244 pb (es decir, sobre la cadena positiva) del genoma de D858 (y tiene 100% de identidad a la secuencia genómica de D858 sobre 30 nucleótidos):

```

espaciador          1 ttacgtttgaaaagaatatcaaataatga 30
                    |||
D858             17215 ttacgtttgaaaagaatatcaaataatga 17244
    
```

15 El nuevo espaciador que está integrado en el locus CRISPR1 de la cepa DGCC7710-RH2 del *Streptococcus thermophilus* confiere a la cepa DGCC7710-RH2 del *Streptococcus thermophilus* una resistencia al fago D858, tal como se representa en la Figura 6 y Tabla 1.

Ejemplo 12

Integración y bloqueo de la construcción

Materiales y Métodos

Cepas y plásmidos

20 Cepa parental DGCC7710 de *Streptococcus thermophilus*, sensible a fagos 858 y 2972

Mutante CRISPR resistente a 858 de *Streptococcus thermophilus* DGCC7778

Streptococcus thermophilus DGCC7778cas1KO

Streptococcus thermophilus DGCC7778cas4KO

Streptococcus thermophilus DGCC7778RT

25 *Streptococcus thermophilus* DGCC7778RT'

Mutante CRISPR resistente a 2972 de *Streptococcus thermophilus* DGCC7714R2

Streptococcus thermophilus DGCC7710R2S1S2

Escherichia coli EC1000 proporciona pORI28 (Russell y Klaenhammer, 2001)

Escherichia coli pCR2.1TOPO proporciona pTOPO (catálogo Invitrogen n° K4500-01)

30 pTOPO es un plásmido usado para la subclonación de las diversas construcciones

pTOPOcas1ko contiene un fragmento integral de cas1

pTOPOcas4ko contiene un fragmento integral de cas4

pTOPO1S2 contiene la construcción de espaciador S1S2

pTOPO RT contiene la construcción de la repetición terminal RT

pORI28 es un plásmido usado para la integración de las diversas construcciones en el cromosoma de las cepas de *Streptococcus thermophilus*

pORicas1ko contiene un fragmento integral de cas1

pORicas4ko contiene un fragmento integral de cas4

5 pORIS1S2 contiene la construcción de espaciador S1S2

purist contiene la construcción de la repetición terminal RT

Cebadores

Cas1

5'-caaatggatagagaaacgc-3' y 5'-ctgataagggttcggtgtcc-3'

10 Cas 4

5'-ggagcagatggaatacaagaaagg-3' y 5'-gagagactagggtgtctcagca-3'

S1S2 y RT

P1 5'-acaacaacagagaagtatctcattg-3'

P2 5'-aacgagtacactcactattgtacg-3'

P3 5'-lccactcacgtacaaatagtgagtactcgttttltattctcaagatttaagtaactgtacagtttgatcaacataaaaag-3'

P4 5'-cttcctcatcctcgttttggt-3'

15 Las cepas y los fagos se obtuvieron a partir de "Danisco Culture Collection", o a partir de material de referencia (Russell y Klaenhammer, "Applied and Environmental Microbiology" 2001, 67:43.691-4.364; Levesque et al., "Applied and Environmental Microbiology" 2005 71:4.057-4.068).

Se llevaron a cabo preparación de fago, purificación y ensayos usando métodos descritos previamente (Duplessis et al., Virology 2005, 340:192-208; Levesque et al., "Applied and Environmental Microbiology" 2005 71:4.057-4.068).

20 Se hicieron crecer las cepas de *Streptococcus thermophilus* a 37°C o 42°C en M17 (Difco Laboratories) complementado con lactosa o sucrosa al 0,5%. Para la infección del fago, se añadieron CaCl₂ 10mM al medio antes de la infección de fago, tal como se ha descrito previamente (Duplessis et al., Virology 2005, 340:192-208; Levesque et al., "Applied and Environmental Microbiology" 2005 71:4.057-4.068).

25 Las enzimas usadas para llevar a cabo las digestiones por restricción y la PCR se adquirieron de Invitrogen y se usaron de acuerdo con las instrucciones del fabricante. Las PCRs se llevaron a cabo sobre un "Eppendorf Mastercycler Gradient Thermocycler" como el descrito previamente (Bairangou et al., 2002 "Applied and Environmental Microbiology 68:2.877-2.884).

30 La inactivación de gen y la inserción de plásmido específico a sitio vía recombinación homóloga en los cromosomas de *Streptococcus thermophilus* se llevaron a cabo mediante subclonación en el sistema pCR2.1TOPO de Invitrogen, posterior clonación en el sistema pORI usando *Escherichia coli* como hospedante y las construcciones por último se purificaron y transformaron en *Streptococcus thermophilus* tal como previamente se ha descrito (Russell y Klaenhammer, "Applied and Environmental Microbiology" 2001, 67:43.691-4.364).

(1) Integración de la construcción

Usando la construcción RT modificada por ingeniería como la mostrada en la Figura 17, se insertó la construcción justo después de cas4, tal como se muestra en la Figura 18.

35 DGCC7778 parental es resistente al fago 858.

La parental tiene dos espaciadores (S1 y D2) que son idénticos al ADN del fago 858.

La cepa resultante (RT) pierde resistencia al fago 858, tal como se muestra en la Tabla 1. Esto demuestra que los genes cas necesitan estar en las inmediaciones cercanas de el(los) espaciador(es) para conferir resistencia.

(2) Bloqueo de Cas1

Tal como se muestra en la Figura 12 DGCC7778 parental se modifica por ingeniería de modo que se altera el gen *cas1*. Tal como se muestra en la Tabla 1, esto da como resultado una pérdida de resistencia, significando que el *cas1* es necesario para conferir resistencia.

5 (3) Bloqueo de Cas4

Tal como se muestra en la Figura 12 DGCC7778 parental se modifica por ingeniería de modo que se altera el gen *cas4*.

(4) Integración de construcción S1S2

Tal como se muestra en las Figuras 14-16 la construcción S1S2 está integrada en DGCC7710 parental.

10 **Compendio**

Las CRISPRs (Repeticiones Palindrómicas Cortas Regularmente Interespaciadas Agrupadas) (también conocidas como SPIDR, repeticiones Directas Intercaladas de Espaciador) constituyen una familia de loci de ADN recientemente descritos extendidos a través de los genomas procarióticos. Están constituidos de repeticiones palindrómicas de ADN altamente conservadas y cortas que están regularmente interespaciadas mediante secuencias altamente polimórficas de aproximadamente la misma longitud. Adicionalmente, los genes *cas* (genes asociados a CRISPR) normalmente están presentes en las inmediaciones de las secuencias CRISPR. En la bibliografía no se ha atribuido todavía función fisiológica clara a las secuencias CRISPR o genes *cas*.

Aquí sugerimos que las secuencias CRISPR en combinación con los genes *cas* se pueden usar para proporcionar resistencia frente al ácido nucleico entrante. Particularmente, proponemos que los espaciadores dentro de los loci CRISPR proporcionan la especificidad para la inmunidad frente al ácido nucleico entrante. Como resultado, sugerimos que los genes *cas* en asociación con las secuencias CRISPR se usen para proporcionar células con resistencia frente a las secuencias particulares de ácido nucleico, tales como de bacteriófagos, de plásmidos, de transposones y secuencias de inserción. Adicionalmente, estos elementos se pueden manipular para generar inmunidad dirigida frente a secuencias particulares de ácido nucleico, tales como componentes de fago, genes de resistencia a antibiótico, factores de virulencia, secuencias novedosas, elementos indeseables y similares. Por tanto, el conocimiento simple de las secuencias de espaciador CRISPR *entre otras* para una cepa bacteriana dada sería una ventaja determinar su lisotipo (el lisotipo define la resistencia/sensibilidad de una bacteria dada a diversos bacteriófagos) y predice su capacidad de sobrevivir a la exposición a secuencias definidas de ácido nucleico. Por lo tanto, la caracterización de los loci CRISPR en las bacterias podría ayudar a determinar, predecir y modificar la interacción hospedante-fago. La aplicación particular de la modificación por ingeniería genética de CRISPR, por adición, delección o modificación de las secuencias de espaciador, podría conducir a variantes bacterianas resistentes a fago.

Las CRISPRs (Repeticiones Palindrómicas Cortas Regularmente Interespaciadas Agrupadas), también conocidas como SPIDR (Repeticiones Directas Intercaladas de Espaciador), forman una nueva familia de secuencias repetidas que se han identificado en secuencias de genoma completo, en numerosos procariotas, mayormente en cromosomas pero también en plásmidos (Mojica et al., 2000; Jansen et al., 2002a). Los loci CRISPR están constituidos de repeticiones de ADN cortas y altamente conservadas (24 a 40 pb, repetidas entre 1 y 140 veces) que son parcialmente palindrómicas. Mientras hay ciertos límites a la degeneración de base entre repeticiones de diferentes loci y especies, no hay secuencia conservada absoluta a través de todas las repeticiones observadas. Además las repeticiones están aparentemente orientadas dentro de un locus particular, con respecto a los genes próximos. Estas secuencias repetidas (normalmente específicas a una especie) están interespaciadas mediante secuencias polimórficas de longitud constante (20 a 58 pb dependiendo del CRISPR) que están diseñados como "espaciadores". Se han encontrado hasta 20 loci CRISPR diferentes dentro de un cromosoma sencillo. La Figura 1 describe una de la CRISPR identificada en *Streptococcus thermophilus* CNRZ1066.

Por ejemplo, el genoma de *S. thermophilus* LMG18311 contiene 3 loci CRISPR. Las secuencias repetidas de 36 pb son diferentes en CRISPR1 (34 repeticiones), CRISPR2 (5 repeticiones) y CRISPR3 (una secuencia sencilla); sin embargo, están perfectamente conservadas dentro de cada locus. Las repeticiones CRISPR1 y CRISPR2 están respectivamente interespaciadas por 33 y 4 secuencias de 30 pb de longitud. Todos estos espaciadores son diferentes entre sí (aparte de las excepciones menores: muy pocos espaciadores pueden estar presentes dos veces dentro de un locus CRISPR dado). También son diferentes de aquellos encontrados en otras cepas, tales como CNRZ1066 (41 espaciadores dentro de CRISPR1) o LMD-9 (16 espaciadores dentro de CRISPR1 y 8 dentro de CRISPR3), las cuales son cepas de *S. thermophilus* que tienen genomas muy similares.

Aunque la función biológica de los loci CRISPR no es conocida se han propuesto algunas hipótesis. Por ejemplo, se ha propuesto que pueden estar implicados en el acoplamiento del cromosoma a una estructura celular, o en la

replicación del cromosoma y partición del replicón, pero no se ha informado de demostración experimental para confirmar estas hipótesis.

Generalmente, los loci CRISPR están inmediatamente adyacentes a un grupo de 4 a 7 genes que se han denominado genes *cas* (del Inglés “CRISPR-associated”, “Asociados a CRISPR”)(Jansen et al., 2002b). Actualmente no se ha atribuido un papel fisiológico claro a las proteínas Cas, pero para algunas de ellas la presencia de motivos de proteína particulares sugiere que podrían actuar como una girasa de ADN o una polimerasa de ADN. Estos grupos de 4 a 7 genes *cas*, que o bien se originan a partir de diferentes loci dentro de un genoma dado o se originan a partir de diferentes microorganismos, se pueden distinguir y agrupar en diferentes tipos en base a la similitud de secuencia. Uno de nuestros principales descubrimientos es que un conjunto dado de genes *cas* siempre está asociado con una secuencia repetida dada dentro de un locus CRISPR particular. En otras palabras, los genes *cas* (o proteínas Cas) parecen ser específicos a una repetición de ADN dada, queriendo decir que los genes *cas* (o proteínas Cas) y la secuencia repetida podrían formar un par funcional. Los análisis de gráfico de puntos indican que los agrupamientos (“cluster”) y los grupos obtenidos cuando se analizan las secuencias de proteína Cas son similares a los obtenidos cuando se analizan las repeticiones CRISPR (tal como se muestra en la figura 2).

En *S. thermophilus*, una especie bacteriana para la cual se han secuenciado varios genomas de fago, los espaciadores de 30 pb con frecuencia son idénticos al ADN del fago (Figura 3). Esta observación también se ha realizado para las secuencias de espaciador de muchos otros géneros y especies bacterianas para las cuales se conocen las secuencias de ADN de fago. Además se ha mencionado previamente en al menos dos publicaciones recientes (Pourcel et al., 2005; Mojica et al., 2005). Por otro lado la ausencia de la similitud de secuencia significativa para las secuencias de espaciador restantes se puede explicar mediante el hecho de que solamente unos pocos genomas de fago están disponibles en este momento. En base a las muy altas similitudes de secuencia de ADN entre algunos espaciadores CRISPR y las secuencias de bacteriófago, proponemos que la especificidad de los espaciadores CRISPR participa en la determinación del lisotipo de cepa. Para soportar el propósito de un efecto de las secuencias del espaciador CRISPR sobre la inmunidad bacteriana frente a bacteriófagos, se encontró que una proporción significativa de parejas para espaciadores en secuencias de genoma de bacteriófago se da dentro de los genes probablemente implicados en la especificidad hospedante (véase figura 3). Otra hipótesis podría ser que las secuencias de espaciador son reconocidos por la bacteria como ADN exógeno. Por tanto, la bacteria eliminaría la molécula de ácido nucleico que porta esta secuencia al entrar en la célula. Un argumento que soporta esta idea es la estructura peculiar deducida de CRISPR. De hecho, proponemos que los elementos de repetición proporcionan una característica estructural mientras los espaciadores contienen la secuencia que proporciona la inmunidad específica frente el ácido nucleico entrante. Las repeticiones palindrómicas tienen el potencial para formar estructuras de “hairpin” (horquilla) muy estables (véase figura 4), y están separados por espaciadores cuyo tamaño corresponde a más o menos 3 vueltas de la hélice de ADN (aunque pueden variar entre 2 y 5). Por tanto, cualquier locus CRISPR podría estar altamente estructurado en una serie de hairpins de ADN regularmente espaciadas.

Ventajosamente, el lisotipo de una cepa bacteriana dada se puede modificar o bien por generación natural de derivados resistentes (Mutantes Insensibles a Bacteriófago), o por ingeniería genética. Específicamente, las soluciones de la ingeniería genética pueden estar diseñadas mediante, por ejemplo, adición, mediante delección o mediante modificación de las secuencias espaciador o incluso un locus CRISPR completo.

Ejemplos de aplicaciones de esta invención incluyen, pero no se limitan a:

(i) Resistencia a fago. Los espaciadores CRISPR particulares derivados del ADN bacteriófago se pueden añadir dentro de un locus CRISPR del hospedante bacteriano para proporcionar resistencia frente a este bacteriófago particular, previniendo así el ataque del fago. Adicionalmente, regiones particulares dentro del genoma del fago (proteínas de especificidad a hospedante) pueden estar dirigidas a proporcionar reconocimiento particular fago-hospedante, o a estar altamente conservadas dentro del ADN del fago, tal como las secuencias de helicasa o genes de primasa, proteínas estructurales de cabeza y cola, o proteínas con dominios conservados (por ejemplo, helicasa, holin, lisina y otros) o secuencias conservadas entre los genes importantes de fago.

(ii) Resistencia a transferencia de plásmido. Los espaciadores CRISPR particulares derivados del ADN plásmido se pueden añadir dentro de un locus CRISPR de bacteria para proporcionar resistencia frente a este plásmido particular, previniendo así la transferencia de ADN exógeno en el microbio. Específicamente, las regiones particulares dentro del ADN plásmido pueden estar dirigidas a proporcionar inmunidad frente a ADN plásmido, tal como las secuencias dentro del origen de replicación del plásmido.

(iii) Resistencia a elementos genéticos móviles. Los espaciadores CRISPR particulares derivados del ADN del elemento genético móvil se pueden añadir dentro de un locus CRISPR de bacteria para proporcionar resistencia frente a elementos genéticos móviles tal como elementos transponibles y secuencias de inserción, previniendo así la transferencia de ADN exógeno y deriva genética. Específicamente, las regiones particulares dentro de los transposones y las secuencias de inserción pueden estar dirigidas a proporcionar inmunidad

frente a elementos genéticos móviles. Por ejemplo, las dianas pueden incluir transposones conjugativos (Tn916), transposones clase II (Tn501) o secuencias de inserciones (IS26).

5 (iv) Resistencia a genes de resistencia a antibiótico. Los espaciadores CRISPR particulares derivados de los genes que codifican resistencia a antibiótico se pueden añadir dentro de un locus CRISPR de bacteria para prevenir la transferencia de los genes que confieren resistencia a antibióticos en el hospedante bacteriano, reduciendo así el riesgo de adquirir marcadores de resistencia a antibiótico. Por ejemplo, las dianas pueden incluir vanR, un gen que confiere resistencia a vancomicina, o teR, un gen que confiere resistencia a tetraciclina, o que fija como objetivo inhibidores beta-lactamasa.

10 (v) Resistencia a genes que codifican factores de virulencia. Los espaciadores CRISPR particulares derivados de genes que codifican factores de virulencia se pueden añadir dentro de un locus CRISPR de bacteria para proporcionar resistencia frente a la transfección de genes que confieren virulencia en la bacteria. Por ejemplo, los factores que comúnmente contribuyen a virulencia en los patógenos microbianos se puede dirigir, tal como toxinas, intermalinas y hemolisinas.

15 (vi) Diagnósticos. Los espaciadores CRISPR dentro de una bacteria particular se pueden detectar o secuenciar para predecir/determinar la probable sensibilidad de microbios particulares a bacteriófago, y usarse así como un indicador de lisotipo para la selección microbiana.

20 (vii) Resistencia a secuencias novedosas. Las secuencias novedosas de espaciador se pueden sintetizar de novo, modificar por ingeniería e integrar en un CRISPR dentro un hospedante bacteriano seleccionado para proporcionar resistencia a una secuencia particular idéntica y novedosa presente en una molécula de ADN infectante.

Puesto que las CRISPR están ampliamente extendidas entre las especies bacterianas, las aplicaciones anteriormente mencionadas se podrían usar en una gran diversidad de organismos. Los loci CRISPR se han descrito en un número de bacterias Gram-positivo (incluyendo bacterias de ácido láctico) y bacterias Gram-negativo. Por tanto, los loci CRISPR en asociación con los genes cas se pueden usar para caracterizar/modificar el lisotipo de cepa y generar resistencia a ácido nucleico en un amplio rango de bacterias. Además de las aplicaciones potenciales para la resistencia a fago, se ha mencionado en la bibliografía que las secuencias CRISPR muestran algo de homología a elementos genéticos móviles tales como plásmidos y transposones (Mojica et al., 2005).

En un aspecto adicional, se ha proporcionado el uso de una combinación de un locus CRISPR y uno o más genes cas para proporcionar resistencia frente a ácido nucleico definido.

30 Adecuadamente, el ácido nucleico es ADN.

Adecuadamente, el ácido nucleico es ARN.

Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de un fago.

Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de un plásmido.

Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de un elemento genético móvil.

35 Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de un transposon (Tn).

Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de una secuencia de inserción (IS).

Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de elementos genéticos dirigidos indeseables.

40 Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de un gen de resistencia a antibiótico.

Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de un factor de virulencia.

Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de una isla de patogenicidad.

45 Adecuadamente, el ácido nucleico es derivable (preferiblemente, derivado) a partir de una secuencia novedosa, para proporcionar resistencia frente a entidades que llevan esta molécula particular.

En un aspecto adicional, se ha proporcionado el uso de CRISPR para la identificación y tipificación.

En un aspecto adicional, se ha proporcionado el uso de uno o más genes *cas* y uno o más elementos CRISPR (por ejemplo, una o más repeticiones CRISPR y/o espaciadores CRISPR) para modular la resistencia en una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.

TABLA 1

Fago 2972				Fago 858	
Cepas	Sobre BIM ¹	Sensibilidad de Fago ²	Homología Espaciador-fago ³	Sensibilidad fago ²	Homología Espaciador-fago ³
DGCC7710	-	S	Ctrl	S	Ctrl
DGCC7778	858	S	>10 SNPs	R	100% (2 espaciadores)
DGCC7710-RH1	858	R	100%	R	100%
DGCC7710-RH2	858	R	100%	R	100%
DGCC7778RT	858	S	>10 SNPs	S	100% pero no próximo a <i>cas</i>
DGCC7778HT'	858	S	>10 SNPs	S	No espaciadores dejados
DGCC7778cas1	858	S	>10 SNPs	S	100% (2 espaciadores) pero <i>cas1</i> KO
DGCC7778cas4	858	S	>10 SNPs	R	100% (2 espaciadores) pero <i>cas4</i> KO
DGCC7710-R2	2972	R	100% (1 espaciador)	S	5 SNPs
DGCC7710-R2S1S2	2972	S	100% pero no próximo a <i>cas</i>	R	S1S2 son 100% idénticos a fago 858

1 Fago usado para generar Mutantes insensibles a Bacteriófago (BIMs, del Inglés "Bacteriophage Insensitive Mutants")

2 Sensibilidad a Fago de la cepa, S=sensible, R=resistente como lo determinado por ensayos de punto y placa

3 Homología entre el nuevo espaciador del mutante, extremo de la secuencia ADN del fago usado para generar los Fagos mutantes retenía la capacidad para adsorber a los mutantes

5

La invención también se refiere a los siguientes aspectos tal como se define en los siguientes párrafos enumerados:

1. El uso de uno o más genes *cas* o proteínas para modular la resistencia en una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.
2. El uso de acuerdo con el párrafo 1, en el que los uno o más genes *cas* o proteínas se usan en combinación con dos o más repeticiones CRISPR.
3. El uso de acuerdo con el párrafo 2, en el que los uno o más genes *cas* o proteínas y/o las dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivados) a partir de la misma célula).
4. El uso de acuerdo con el párrafo 2 ó párrafo 3, en el que los uno o más genes *cas* o proteínas y las dos o más repeticiones CRISPR se dan conjuntamente de manera natural en la misma célula.
5. El uso de acuerdo con cualquiera de los párrafos precedentes, en el que los uno o más genes *cas* o proteínas se usan en combinación con uno o más espaciadores CRISPR.
6. El uso de acuerdo con el párrafo 5, en el que los espaciador(es) CRISPR está(n) o es(son) derivable(s)(preferiblemente, derivado(s)) a partir de un organismo que es diferente a la célula a partir de la cual los uno o más genes *cas* o proteínas y/o las dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivadas).

10

15

20

7. El uso de acuerdo con el párrafo 6, en el que el espaciador se obtiene a partir de una célula que es resistente a un ácido nucleico diana.
8. El uso de acuerdo con el párrafo 5, en el que el espaciador CRISPR es una secuencia de ácido nucleico sintética.
- 5 9. El uso de acuerdo con cualquiera de los párrafos 5-8, en el que el(los) espaciador(es) CRISPR tiene(n) homología al ácido nucleico diana.
10. El uso de acuerdo con el párrafo 9, en el que el(los) espaciador(es) tiene(n) 100% de identidad al ácido nucleico diana sobre al menos la longitud de la zona central del espaciador CRISPR.
- 10 11. El uso de acuerdo con cualquiera de los párrafos precedentes, en el que los uno o más genes *cas* o proteínas se usan en combinación con al menos uno o más espaciadores CRISPR y al menos dos o más repeticiones CRISPR.
12. El uso de acuerdo con cualquiera de los párrafos precedentes, en el que el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de ADN bacteriófago.
- 15 13. El uso de acuerdo con cualquiera de los párrafos precedentes, en el que el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de ADN plásmido.
14. El uso de acuerdo con cualquiera de los párrafos precedentes, en el que el ácido nucleico diana o el producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de un elemento genético móvil.
15. El uso de acuerdo con el párrafo 14, en el que el ácido nucleico diana o el producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de un elemento transponible o una secuencia de inserción.
- 20 16. El uso de acuerdo con cualquiera de los párrafos precedentes, en el que el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de un gen de resistencia a antibiótico.
- 25 17. El uso de acuerdo con cualquiera de los párrafos precedentes, en el que el ácido nucleico diana o producto de transcripción del mismo es o es derivable (preferiblemente, derivado) a partir de un ácido nucleico que codifica un factor de virulencia.
18. El uso de acuerdo con el párrafo 17, en el que el factor de virulencia se selecciona entre el grupo que consiste en un ácido nucleico que codifica una toxina, una internalina y una hemolisina.
- 30 19. El uso de una secuencia de ácido nucleico recombinante que comprende al menos un gen *cas* y al menos dos repeticiones CRISPR junto con al menos un espaciador CRISPR, en el que al menos un espaciador CRISPR es heterólogo a al menos un gen *cas* y/o al menos dos repeticiones CRISPR para modular la resistencia frente a un ácido nucleico diana o producto de transcripción del mismo.
20. Una secuencia de ácido nucleico que consiste esencialmente en al menos un gen *cas*.
- 35 21. Una secuencia de ácido nucleico que consiste esencialmente en al menos un gen *cas* y al menos dos repeticiones CRISPR.
22. Una secuencia de ácido nucleico que consiste esencialmente en al menos un gen *cas* y al menos un espaciador CRISPR.
23. Una secuencia de ácido nucleico que consiste esencialmente en al menos un gen *cas*, al menos un espaciador CRISPR y al menos dos repeticiones CRISPR.
- 40 24. Una secuencia de ácido nucleico recombinante que comprende al menos un gen *cas* y al menos dos repeticiones CRISPR junto con al menos un espaciador CRISPR, en la que el espaciador CRISPR es heterólogo al al menos un gen *cas* y/o al menos dos repeticiones CRISPR.
25. La secuencia de ácido nucleico de acuerdo con cualquiera de los párrafos 20-24, en la que los uno o más genes *cas* y las dos o más repeticiones CRISPR están o son derivables (preferiblemente, derivados) a partir de la misma célula.
- 45 26. La secuencia de ácido nucleico de acuerdo con cualquiera de los párrafos 20-25, en la que los uno o más genes *cas* y las dos o más repeticiones CRISPR se dan conjuntamente de manera natural en la misma célula.

27. La secuencia de ácido nucleico de acuerdo con cualquiera de los párrafos 20-26, en la que los espaciadores CRISPR son o son derivables (preferiblemente, derivados) a partir de un organismo que es diferente a la célula a partir de la cual los uno o más genes *cas* y/o las dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivados).
- 5 28. Una construcción que comprende una o más de las secuencias de ácido nucleico de acuerdo con cualquiera de los párrafos 20-27.
29. Un vector que comprende una o más de las secuencias de ácido nucleico de acuerdo con cualquiera de los párrafos 20-27 o una o más de las construcciones de acuerdo con el párrafo 24.
- 10 30. Una célula que comprende la secuencia de ácido nucleico de acuerdo con cualquiera de los párrafos 20-27, la construcción de acuerdo con el párrafo 28 o el vector de acuerdo con el párrafo 29.
31. La célula de acuerdo con el párrafo 30, en el que la célula es una célula receptora o una célula hospedante.
32. Un método para modular (por ejemplo, conferir o incrementar) la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo que comprende las etapas de :
- 15 (i) identificar una secuencia (por ejemplo, una secuencia conservada) en un organismo (preferiblemente, una secuencia esencial a la función o supervivencia del organismo);
- (ii) preparar un espaciador CRISPR que sea homólogo a la secuencia identificada;
- (iii) preparar un ácido nucleico (por ejemplo, un ácido nucleico recombinante) que comprenda al menos un gen *cas* y al menos dos repeticiones CRISPR junto con el espaciador CRISPR; e
- 20 (iv) introducir dicho ácido nucleico en una célula para así volver a la célula resistente a dicho ácido nucleico diana o producto de transcripción del mismo.
33. Un método para modular (por ejemplo, conferir o incrementar) la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo que comprende las etapas de:
- (i) identificar uno o más espaciadores CRISPR o pseudo espaciadores CRISPR en un organismo resistente al ácido nucleico diana o producto de transcripción del mismo;
- 25 (ii) preparar un ácido nucleico recombinante que comprenda al menos un gen *cas* o proteína y al menos dos repeticiones CRISPR junto con dicho uno o más espaciadores identificados; e
- (iii) introducir dicho ácido nucleico recombinante en una célula para volver así a la célula resistente a dicho ácido nucleico diana o producto de transcripción del mismo.
- 30 34. Un método para modular (por ejemplo, conferir o incrementar) la resistencia de una célula que comprende al menos uno o más genes *cas* o proteínas y dos o más repeticiones CRISPR frente a un ácido nucleico diana o un producto de transcripción del mismo que comprende las etapas de:
- (i) identificar uno o más espaciadores CRISPR en un organismo resistente al ácido nucleico diana o producto de transcripción del mismo; y
- 35 (ii) modificar la secuencia de uno o más espaciador(es) CRISPR en la célula de modo que el(los) espaciador(es) CRISPR tengan homología a el(los) espaciador(es) CRISPR en el organismo.
35. El método de acuerdo con cualquiera de los párrafos 32-34, en el que los uno o más genes *cas* o proteínas y/o las dos o más repeticiones CRISPR son o son derivables (preferiblemente, derivados) a partir de la misma célula.
- 40 36. El método de acuerdo con cualquiera de los párrafos 32-35, en el que los espaciadores son o son derivables (preferiblemente, derivados) a partir de un organismo que es diferente a la célula que comprende los uno o más genes *cas* o proteínas y/o las dos o más repeticiones CRISPR.
37. El método de acuerdo con cualquiera de los párrafos 32-35, en el que los uno o más genes o proteínas y las dos repeticiones CRISPR se dan conjuntamente de manera natural en la misma célula.
- 45 38. El método de acuerdo con cualquiera de los párrafos 32-37, en el que dicha modificación comprende la inserción de uno o más espaciadores CRISPR y/o pseudo espaciadores CRISPR en la célula.

39. El método de acuerdo con el párrafo 38, en el que el espaciador de la célula tiene 100% de homología al espaciador CRISPR o seudo espaciador CRISPR del organismo.
40. El método de acuerdo con el párrafo 38 o el párrafo 39, en el que dicha modificación comprende modificar por ingeniería genética el espaciador CRISPR de la célula.
- 5 41. El método de acuerdo con el párrafo 40, en el que se modifica todo o parte del espaciador en la célula.
42. El método de acuerdo con el párrafo 40 o el párrafo 41, en el que dicha modificación comprende la modificación de un espaciador recombinante.
43. El método de acuerdo con cualquiera de los párrafos 40-42, en el que dicha modificación se da a través de mutación o mutagénesis espontánea.
- 10 44. Un método para modular (por ejemplo, reducir o disminuir) la resistencia de una célula que comprende al menos uno o más genes *cas* o proteínas y dos o más repeticiones CRISPR frente a un ácido nucleico diana o un producto de transcripción del mismo que comprende las etapas de:
- (i) identificar uno o más espaciadores CRISPR en un organismo que es sustancialmente resistente al ácido nucleico diana o producto de transcripción del mismo; y
- 15 (ii) modificar la secuencia de al menos uno o más espaciador(es) CRISPR en la célula de modo que los espaciador(es) CRISPR tengan un grado reducido de homología a el(los) espaciador(es) en el organismo.
45. El método de acuerdo con el párrafo 44, en el que se someten a delección al menos uno o más espaciador(es) CRISPR en la célula.
- 20 46. El método de acuerdo con el párrafo 44 o el párrafo 45, en el que se someten a delección al menos una o más repetición(es) CRISPR en la célula.
47. El método de acuerdo con cualquiera de los párrafos 44-46, en el que se someten a delección uno o más genes *cas*.
48. El método de acuerdo con cualquiera de los párrafos 44-47, en el que se someten a delección CRISPR y/o uno o más genes *cas*.
- 25 49. Un método para modular (por ejemplo, reducir o disminuir) la resistencia de una célula que comprende al menos uno o más genes *cas* o proteínas y dos o más repeticiones CRISPR frente a un ácido nucleico diana o un producto de transcripción del mismo que comprende la modificación de uno o más genes *cas* o proteínas y/o dos o más repeticiones CRISPR en la célula.
- 30 50. El método de acuerdo con el párrafo 49, en el que se somete a delección los uno o más genes *cas* o proteínas y/o dos o más repeticiones CRISPR en la célula.
51. Un método para identificar un espaciador CRISPR o seudo espaciador CRISPR para usarse en la modulación de la resistencia de una célula frente a un ácido nucleico o un producto de transcripción del mismo que comprende las etapas de:
- (i) preparar una célula que comprenda al menos dos repeticiones CRISPR y al menos un gen *cas* o proteína;
- 35 (ii) identificar al menos un espaciador CRISPR o seudo espaciadores CRISPR en un organismo que es sustancialmente resistente al ácido nucleico diana o producto de transcripción del mismo;
- (iii) modificar la secuencia del espaciador CRISPR en la célula de modo que el espaciador CRISPR tenga homología al espaciador del organismo; y
- 40 (iv) determinar si la célula modula resistencia frente al ácido nucleico diana o producto de transcripción del mismo,
- donde la modulación de la resistencia de la célula frente al ácido nucleico diana o producto de transcripción del mismo es indicativo de que el espaciador CRISPR modula la resistencia de la célula.
52. Un método para identificar un gen *cas* para usarse en la modificación de la resistencia de una célula frente a un ácido nucleico diana o producto de transcripción del mismo que comprende las etapas de:
- 45 (i) preparar una célula que comprenda al menos un espaciador CRISPR y al menos dos repeticiones CRISPR;

- (ii) modificar por ingeniería la célula de modo que comprenda al menos un gen *cas*; y
- (iii) determinar si la célula modula resistencia frente al ácido nucleico diana o producto de transcripción del mismo,
- 5 donde la modulación de la resistencia de la célula frente al ácido nucleico diana o producto de transcripción del mismo es indicativo de que el gen *cas* se puede usar para modular la resistencia de la célula.
53. Un método para identificar una repetición CRISPR para usarse en la modulación de la resistencia de una célula frente a un ácido nucleico diana o producto de transcripción del mismo que comprende las etapas de:
- (i) preparar una célula que comprenda al menos un espaciador CRISPR y al menos un gen *cas*;
- (ii) modificar por ingeniería la célula de modo que contenga la repetición CRISPR; y
- 10 (iii) determinar si la célula modula resistencia frente al ácido nucleico o producto de transcripción del mismo, donde la modulación de la resistencia de la célula frente al ácido nucleico diana o producto de transcripción del mismo es indicativo de que la repetición CRISPR se puede usar para modular resistencia.
54. Un método para identificar una combinación funcional de un gen *cas* y una repetición CRISPR que comprende las etapas de:
- 15 (a) determinar las secuencias del gen *cas* y la repetición CRISPR;
- (b) identificar uno o más grupos de los genes *cas* como los determinados por el análisis de comparación de secuencia;
- (c) identificar uno o más grupos de repeticiones CRISPR; y
- (d) combinar aquellas secuencias de gene *cas* y repetición CRISPR que caen dentro del mismo grupo,
- 20 donde la combinación de las secuencias de gen *cas* y repetición CRISPR dentro del mismo grupo es indicativo de que la combinación es una combinación funcional.
55. El método de acuerdo con el párrafo 54, en el que las secuencias de nucleótidos del gen *cas* y la repetición CRISPR son o son derivables (preferiblemente, derivadas) a partir de las mismas o diferentes cepas.
- 25 56. El método de acuerdo con el párrafo 54, en el que las secuencias de nucleótidos del gen *cas* y la repetición CRISPR son o son derivables (preferiblemente, derivadas) a partir de las mismas o diferentes especies.
57. El método de acuerdo con el párrafo 54, en el que las secuencias de nucleótidos del gen *cas* y la repetición CRISPR son o son derivables (preferiblemente, derivadas) a partir de los mismos o diferentes géneros.
58. El método de acuerdo con el párrafo 54, en el que las secuencias de nucleótidos del gen *cas* y la repetición CRISPR son o son derivables (preferiblemente, derivadas) a partir de los mismos o diferentes organismos.
- 30 59. Un método para modular el lisotipo de una célula bacteriana que comprende uno o más genes *cas* o proteínas y dos o más repeticiones CRISPR que comprende las etapas de:
- (i) identificar uno o más pseudo espaciadores CRISPR en la secuencia genómica de un bacteriófago frente al cual se modula la resistencia; y
- 35 (ii) modificar la secuencia de uno o más espaciadores CRISPR de la célula bacteriana de modo que el(los) espaciador(es) de la célula bacteriana tenga(n) homología a el(los) pseudo espaciador(es) CRISPR del bacteriófago frente al cual se modula la resistencia.
60. Un método para modular (por ejemplo, conferir o incrementar) la resistencia de una célula bacteriana frente a un bacteriófago que comprende las etapas de:
- 40 (i) identificar una secuencia (por ejemplo una secuencia conservada) en un bacteriófago (preferiblemente, una secuencia esencial a la función o supervivencia del bacteriófago);
- (ii) preparar un espaciador CRISPR el cual es homólogo a la secuencia identificada;
- (iii) preparar un ácido nucleico que comprenda al menos un gen *cas* y al menos dos repeticiones CRISPR junto con el espaciador CRISPR; e

- (iv) introducir dicho ácido nucleico en la célula bacteriana para así volver a la célula bacteriana resistente a dicho ácido nucleico diana o producto de transcripción del mismo.
- 5 61. Un método para modular (por ejemplo, conferir o incrementar) la resistencia de una célula bacteriana frente a un ácido nucleico diana o producto de transcripción en un bacteriófago del mismo que comprende las etapas de:
- (i) identificar uno o más seudo espaciadores CRISPR en un genoma de bacteriófago que es capaz de proporcionar resistencia al ácido nucleico diana o producto de transcripción del mismo;
- (ii) preparar un ácido nucleico recombinante que comprende al menos un gen *cas* y al menos dos repeticiones junto con dichos uno más seudo espaciadores CRISPR identificados; e
- 10 (iii) introducir dicho ácido nucleico recombinante en dicha célula bacteriana para así volver a la célula bacteriana resistente a dicho ácido nucleico diana o producto de transcripción del mismo.
62. Un método para modular la resistencia de una célula bacteriana que comprende uno o más genes *cas* o proteínas y dos o más repeticiones CRISPR frente a un ácido nucleico diana o producto de transcripción del mismo en un bacteriófago que comprende las etapas de:
- 15 (i) identificar uno o más seudo espaciadores CRISPR en un bacteriófago que es capaz de proporcionar resistencia a un ácido nucleico diana o producto de transcripción del mismo;
- (ii) identificar uno o más espaciadores CRISPR en una célula bacteriana en la cual se modula la resistencia; y
- (iii) modificar la secuencia de el(los) espaciador(es) en la célula bacteriana en la cual se modula la resistencia de modo que el(los) espaciador(es) tengan un grado mayor de homología a el(los) seudo espaciador(es) CRISPR del bacteriófago frente al cual se modula la resistencia.
- 20 63. El método de acuerdo con el párrafo 61 o el párrafo 62, en el que el ácido nucleico diana en el bacteriófago es una secuencia de ácido nucleico altamente conservada.
64. El método de acuerdo con cualquiera de los párrafos 61-63, en el que el ácido nucleico diana en el bacteriófago codifica una proteína de especificidad a hospedante.
- 25 65. El método de acuerdo con cualquiera de los párrafos 61-64, en el que el ácido nucleico diana en el bacteriófago codifica una proteína que es esencial para la supervivencia, replicación o crecimiento del bacteriófago.
66. El método de acuerdo con el párrafo 65, en el que el ácido nucleico en el bacteriófago codifica una helicasa, una primasa, una proteína estructural de cabeza o cola, una proteína con un dominio conservado (por ejemplo, holín, lisina y otros) o una secuencia conservada entre los genes importantes de fago.
- 30 67. Un método para determinar la resistencia de una célula frente a un ácido nucleico diana o producto de transcripción del mismo que comprende identificar una o más combinaciones funcionales repetición CRISPR-*cas* y uno o más espaciadores CRISPR en la célula.
68. El método de acuerdo con el párrafo 67, que comprende la etapa adicional de comparar la secuencia de los uno o más espaciadores CRISPR en la célula con la secuencia al ácido nucleico diana.
- 35 69. El método de acuerdo con el párrafo 67 o el párrafo 68, que comprende la etapa adicional de determinar el perfil de resistencia de la célula.
70. Una célula obtenida u obtenible mediante el método de cualquiera de los párrafos 32-69.
71. Un espaciador CRISPR o seudo espaciador CRISPR obtenido u obtenible mediante el método del párrafo 51.
72. Un gen *cas* obtenido u obtenible mediante el método del párrafo 52
- 40 73. Una repetición CRISPR obtenida u obtenible mediante el método del párrafo 53.
74. Una combinación funcional obtenida u obtenible mediante el método de cualquiera de los párrafos 54-58.
75. Un locus CRISPR recombinante que comprende un espaciador CRISPR o seudo espaciador CRISPR de acuerdo con el párrafo 71, y/o un gen *cas* de acuerdo con el párrafo 72, y/o una repetición CRISPR de acuerdo con el párrafo 73 y/o una combinación funcional de acuerdo con el párrafo 74.

76. El uso de una célula de acuerdo con el párrafo 70, un espaciador CRISPR o seudo espaciador CRISPR de acuerdo con el párrafo 71, un gen *cas* de acuerdo con el párrafo 72, una repetición CRISPR de acuerdo con el párrafo 73 o una combinación funcional de acuerdo con el párrafo 74 para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.
- 5 77. Un cultivo celular que comprende una célula de acuerdo con el párrafo 70, un espaciador CRISPR o seudo espaciador CRISPR de acuerdo con el párrafo 71, un gen *cas* de acuerdo con el párrafo 72, una repetición CRISPR de acuerdo con el párrafo 73 o una combinación funcional de acuerdo con el párrafo 74 para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.
- 10 78. El cultivo celular de acuerdo con el párrafo 77, en el que dicho cultivo es un cultivo iniciador o un cultivo probiótico.
79. Un producto alimenticio o comida que comprende el cultivo de acuerdo con el párrafo 77 o el párrafo 78.
80. Un proceso para preparar un producto alimenticio o comida que comprende el uso del cultivo de acuerdo con el párrafo 77 o el párrafo 78.
81. Un producto alimenticio o comida obtenido u obtenible mediante el proceso del párrafo 80.
- 15 82. El uso del cultivo de acuerdo con el párrafo 77 o el párrafo 78 para preparar un producto alimenticio.
83. Una secuencia de nucleótidos que comprende o consiste en la secuencia descrita en cualquiera de SEC. ID N°s 7-10 y SEC. ID. N°s 359 a 405 o una variante, fragmento, homólogo o derivado de la misma.
84. Una secuencia de aminoácidos codificada por la secuencia de nucleótidos de acuerdo con el párrafo 83.
- 20 85. Una construcción o vector que comprende uno o más de las secuencias de nucleótidos de acuerdo con el párrafo 83.
86. Una célula hospedante en la cual se ha incorporado una o más de las secuencias de nucleótidos de acuerdo con el párrafo 83 o la construcción o vector de acuerdo con el párrafo 85.

REFERENCIAS

Bolotin A, Quinquis B, Sorokin A, Ehrlich SD (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151(8):2551-61.

Groenen PM, Bunschoten AE, van Soolingen D, & JD van Embden (1993). Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Molecular Microbiology* 10:1057-1065.

Hoe N, Nakashima K, Grigsby D, Pan X, Dou SJ, Naidich S, Garcia M, Kahn E, Bergmire-Seat D, & JM Musser (1999). Rapid molecular genetic subtyping of serotype M1 group A *Streptococcus* strains. *Emerging Infectious Diseases* 5:254-263.

Jansen R, Van Embden JDA, Gaastra W, & LM Schouls (2002a). Identification of a novel family of sequence repeats among prokaryotes. *OMICS* 6:23-33.

Jansen R, Van Embden JDA, Gaastra W, & LM Schouls (2002b). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology* 43:1565-1575

Karnerbeek J, Schouls L, Kolk A, Van Agterveld M, Van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, & J Van Embden (1997). Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *Journal of Clinical Microbiology* 35:907-914

Mojica FJM, Diez-Villasenor C, Soria E, & G Juez (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology* 36:244-246

Mojica FJM, Diez-Villasenor C, Garcia-Martinez J, & E Soria (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution* 60:174-182

Pourcel C, Savignol G, & G Vergnaud (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA and provide additional tools for evolutionary studies. *Microbiology* 151:653-663

Saunders NFW, Goodchild A, Raftery M, Guilhaus M, Curmi PMG, & R Cavicchioli (2005). Predicted roles for hypothetical proteins in the low-temperature expressed proteome of the antarctic archaeon *Methanococcus burtonii*. *Journal of Proteome Research* 4:464-472

Mongodin EF, Hance IR, DeBoy RT, Gill SR, Daugherty S, Huber R, Fraser CM, Stetter K, & KE Nelson (2005). Gene transfer and genome plasticity in *Thermotoga maritima*, a model hyperthermophilic species. *Journal of Bacteriology* 187:4935-4944

Peng X, Brugger K, Shen L, She Q, & RA Garrett (2003). Genus-specific protein binding to the large clusters of DNA repeats (Short Regularly Spaced Repeats) present in *Sulfolobus* genomes. *Journal of Bacteriology* 185: 2410-2417

LISTADO DE SECUENCIAS

<110> DANISCO A/S

5 <120> USO

<130> P022907EPAA

<150> US 60/711.396

10 <151> 26-08-2005

<150> US 60/747.683

<151> 19-05-2006

15 <160> 712

<170> PatentIn version 3.3

<210> 1

20 <211> 36

<212> ADN

<213> Streptococcus thermophilus

<400> 1

25 gttttgtac tctcaagatt taagtaactg tacaac 36

<210> 2

<211> 36

30 <212> ADN

<213> Streptococcus thermophilus

<400> 2

35 gttttgtat tctcaagatt taagtaactg tacagt 36

<210> 3

<211> 36

<212> ADN

40 <213> Streptococcus thermophilus

<400> 3

45 gttttgtac tctcaagatt taagtaactg tacagt 36

<210> 4

<211> 36

<212> ADN

<213> Streptococcus thermophilus

50 <400> 4

gtttttgtac tctcaagatt taagtaaccg tacaac 36

55 <210> 5

ES 2 398 918 T3

<211> 36
 <212> ADN
 <213> Streptococcus thermophilus

5 <400> 5

gttttgtac tctcaagatt taagtaactg tgcaac 36

<210> 6
 10 <211> 36
 <212> ADN
 <213> Streptococcus thermophilus

<400> 6
 15 gttttgtac tctcaagatt taagtagctg tacagt 36

<210> 7
 <211> 36
 20 <212> ADN
 <213> Streptococcus thermophilus

<400> 7
 25 gttttgtac tctcaagata taagtaactg tacaac 36

<210> 8
 <211> 36
 <212> ADN
 30 <213> Streptococcus thermophilus

<400> 8
 35 gttttgtac tctcaagatc taagtaactg tacaac 36

<210> 9
 <211> 36
 <212> ADN
 <213> Streptococcus thermophilus

40 <400> 9

gttttgtac tctcaagatg taagtaactg tacaac 36

<210> 10
 <211> 36
 <212> ADN
 <213> Streptococcus thermophilus

50 <400> 10

gtctttgtac tctcaagatt taagtaactg tacaac 36

<210> 11
 55 <211> 36

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 11
 5 aaaaaagtcc cctctcgagg taattaggtt tatatc 36

 <210> 12
 <211> 36
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 12
 15 gtttccgtcc cctctcgagg taattaggtt tatatc 36

 <210> 13
 <211> 36
 <212> ADN
 20 <213> Streptococcus agalactiae

 <400> 13
 25 gttttagagc tgtgtggtt cgaatgggtc caaaac 36

 <210> 14
 <211> 36
 <212> ADN
 <213> Streptococcus agalactiae
 30 <400> 14

 gttttaaagc tgtgctgta ttatgctagg gcacca 36
 35 <210> 15
 <211> 36
 <212> ADN
 <213> Streptococcus agalactiae
 40 <400> 15

 gttttagagc tgtgctggtt cgaatgggtc caaaac 36

 <210> 16
 45 <211> 36
 <212> ADN
 <213> Streptococcus agalactiae

 <400> 16
 50 gttttagagc tgtgctgta ttatgctagg acatca 36

 <210> 17
 <211> 36
 55 <212> ADN

ES 2 398 918 T3

<213> Streptococcus mutans
 <400> 17
 5 gttttagagc catgtagtt actgatttac taaaat 36
 <210> 18
 <211> 36
 <212> ADN
 10 <213> Streptococcus pyogenes
 <400> 18
 15 gttttagagc tatgctgttt tgaatgtcc caaaac 36
 <210> 19
 <211> 36
 <212> ADN
 <213> Streptococcus pyogenes
 20 <400> 19
 gttttagagc tatgctgttt tgaatgtct ccattc 36
 25 <210> 20
 <211> 33
 <212> ADN
 <213> Streptococcus pyogenes
 30 <400> 20
 ctttcaatcc actcacccat gaagggtgag acg 33
 <210> 21
 35 <211> 33
 <212> ADN
 <213> Streptococcus pyogenes
 <400> 21
 40 atttcaatcc actcacccat gaagggtgag act 33
 <210> 22
 <211> 33
 45 <212> ADN
 <213> Streptococcus pyogenes
 <400> 22
 50 atttcaatcc actcacccat gaagggtgag acc 33
 <210> 23
 <211> 30
 <212> ADN
 55 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 23
 5 agaacgtatt ccaaaacctc ttacgatta 30
 <210> 24
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 24
 ttaactgtta tcaaatgat aagatagtct 30
 15 <210> 25
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 25
 cgttgatgft tattcaagta aaataattaa 30
 25 <210> 26
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 26
 tccttcacg ggtagcacac taacatacac 30
 35 <210> 27
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 27
 40 gttggcaatg caaacacct ttatgaaccg 30
 <210> 28
 <211> 30
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 28
 50 tttattcct tgcgataacg ttccacctt 30
 <210> 29
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 55

ES 2 398 918 T3

<400> 29
 agattataag gaacacaacc aactatatag 30
 5 <210> 30
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 30
 acgacatcaa gctgattgtc ttctacataa 30
 <210> 31
 15 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 31
 20 tttggaatac tgaatgttt actgaaaatc 30
 <210> 32
 <211> 30
 25 <212> ADN
 <213> Streptococcus thermophilus
 <400> 32
 30 acaccactat ctttctctcc tgaaaatgaa 30
 <210> 33
 <211> 30
 <212> ADN
 35 <213> Streptococcus thermophilus
 <400> 33
 gtaattccac gaaattatca acctatgca 30
 40 <210> 34
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 45 <400> 34
 ttggaggatt gcccatatt cccaagagt 29
 50 <210> 35
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 55 <400> 35

ES 2 398 918 T3

	gagagggcgtt aaatatagaa atgcaagatt	30
5	<210> 36 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 36	
10	ttttaacgtc atcagtccac cgccttaa	30
15	<210> 37 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 37	
20	cacctcttc gatggaaagg tctcctcta	30
25	<210> 38 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 38	
30	gaccaaagtt tgattataga gctatacacc	30
35	<210> 39 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 39	
40	accatcattc ttaccattac aactgtaatg	30
45	<210> 40 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 40	
50	atacgaattc gggtcgaca attacaattc	30
55	<210> 41 <211> 32 <212> ADN <213> Streptococcus thermophilus <400> 41	

ES 2 398 918 T3

	tatcaacgca atcattacaa caactcaaa ca	32
	<210> 42	
	<211> 30	
5	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 42	
10	atctacgtgt caatacatat cacaaaacag	30
	<210> 43	
	<211> 29	
	<212> ADN	
15	<213> Streptococcus thermophilus	
	<400> 43	
20	atcttagaa atttctgata taataatga	29
	<210> 44	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
25	<400> 44	
	ttgttgaac aaggacgact tggtaaacta	30
30	<210> 45	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
35	<400> 45	
	catattaagc tgactgggcc taatgcttt	30
	<210> 46	
	<211> 30	
40	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 46	
45	ttcatagcat accgtagtg taaaatctat	30
	<210> 47	
	<211> 30	
	<212> ADN	
50	<213> Streptococcus thermophilus	
	<400> 47	
55	aacatttagg gaatgaaatt gataagactg	30

ES 2 398 918 T3

	<210> 48	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
5	<400> 48	
	aacatgagaa actgtagaaa acaagcaata	30
10	<210> 49	
	<211> 31	
	<212> ADN	
	<213> Streptococcus thermophilus	
15	<400> 49	
	tggtgaagat ggcagtcata aatggcacat t	31
20	<210> 50	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
25	<400> 50	
	aagggttgaa aaatgttggat atatcaaacg	30
30	<210> 51	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 51	
35	ttctgtagt ggatttagtc aaacagatgt	30
40	<210> 52	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 52	
45	tccatagagc gtcttaaaca aagaatagtc	30
50	<210> 53	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 53	
	ttatgattga atgacatggt tgtataagta	30
55	<210> 54	

ES 2 398 918 T3

<211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 5 <400> 54
 tttcttagg aataccaggg agttcagctt 30
 <210> 55
 10 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 55
 15 tggcagagat tacacagcaa cggaaacagc 30
 <210> 56
 <211> 30
 20 <212> ADN
 <213> Streptococcus thermophilus
 <400> 56
 25 gggatcatt gtatctagtg atggacctga 30
 <210> 57
 <211> 30
 <212> ADN
 30 3> Streptococcus thermophilus
 <400> 57
 atttgaaaaa tgcacaacag cgttgatag 30
 35 <210> 58
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 58
 gagctaccag ctaccccgta tgcagagag 30
 45 <210> 59
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 50 <400> 59
 cgttcctttt tcaaggtaa tcttgaaag 30
 <210> 60
 55 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 60
 5 aagtcgtaa gcaccagttc caatcgcat 30

 <210> 61
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 61
 15 ttgaatacca atgccagctt ctttaaggc 30

 <210> 62
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 62
 25 aacctcatal atggggaaaa ttgtaagta 30

 <210> 63
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 63

 taacttcatt agttagttg taattagcat 30

 35 <210> 64
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 40 <400> 64

 ttagctacc aaatatcttc tgtttccaa 30

 <210> 65
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 65
 50 gagtttcaa tattggcaca ggagacaatt 30

 <210> 66
 <211> 30
 55 <212> ADN

ES 2 398 918 T3

<213> Streptococcus thermophilus
<400> 66

5 tgatactatt ttagtcagat atgaaatac 30

<210> 67
<211> 31
<212> ADN
10 <213> Streptococcus thermophilus
<400> 67

15 tcatcaatgt ttaaagccca acaatacatg a 31

<210> 68
<211> 30
<212> ADN
20 <213> Streptococcus thermophilus
<400> 68

tagatttaat cagtaatgag ttaggcataa 30

25 <210> 69
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 69

aggaaaatag catgagcgta caacaatcta 30

<210> 70
35 <211> 30
<212> ADN
<213> Streptococcus thermophilus

40 <400> 70

tgtctatcac gcttcctaag tgcataaaa 30

<210> 71
45 <211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 71
50 atgtcaccaa tcactaaaga acctacgctg 30

<210> 72
55 <211> 30
<212> ADN

ES 2 398 918 T3

<213> Streptococcus thermophilus
 <400> 72
 5 aacatctcc tctccgattg caaatagtg 30
 <210> 73
 <211> 29
 <212> ADN
 10 <213> Streptococcus thermophilus
 <400> 73
 15 catatttggg gcccgttcga taaagagta 29
 <210> 74
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 74
 cattaaatcg ctggaagcag acattgaagc 30
 25 <210> 75
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 75
 gacttatctt ggaaggtagt gaaggcactt 30
 <210> 76
 35 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 76
 40 tccttgccat ctgcactgta agccaagca 30
 <210> 77
 <211> 30
 45 <212> ADN
 <213> Streptococcus thermophilus
 <400> 77
 50 tagtacgcat aatcaattca tcaagcttga 30
 <210> 78
 <211> 30
 <212> ADN
 55 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 78
 5 gtatgaccc aaaattctat gaccttgaaa 30
 <210> 79
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 79
 agattggtg gcttacggaa aatccttgt 30
 15 <210> 80
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 80
 tggcaagaag tgtaagagat gcaatggata 30
 25 <210> 81
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 81
 tttattatca ttattctct tccaagcgt 30
 35 <210> 82
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 82
 40 tttatagaa ttggtggtg aacttttca 30
 <210> 83
 <211> 29
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 83
 50 aatgggtcac agattgcat aataaggag 29
 <210> 84
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 55

ES 2 398 918 T3

	<400> 84	
	ccgaggtcac tttagaaccc acaaaataag	30
5	<210> 85 <211> 30 <212> ADN <213> Streptococcus thermophilus	
10	<400> 85	
	atgagagaac acagtataga ccctgataca	30
15	<210> 86 <211> 30 <212> ADN <213> Streptococcus thermophilus	
20	<400> 86	
	cagtattaat gaggtttggg tggcattcc	30
25	<210> 87 <211> 31 <212> ADN <213> Streptococcus thermophilus	
	<400> 87	
30	ccatactctc tatcagttca ttaattctt c	31
35	<210> 88 <211> 30 <212> ADN <213> Streptococcus thermophilus	
	<400> 88	
40	taatatgctg ctctactgat tccaaaacgg	30
45	<210> 89 <211> 30 <212> ADN <213> Streptococcus thermophilus	
	<400> 89	
	atgaattaca ttcattgatt tatcgagtt	30
50	<210> 90 <211> 29 <212> ADN <213> Streptococcus thermophilus	
55	<400> 90	

ES 2 398 918 T3

	cgtgccattg ttcggtcgg acgtgggca	29
5	<210> 91 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 91	
10	ctttctaagt tgaattaaat tcaagtttg	30
15	<210> 92 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 92	
20	tcgctactat ggtaacgat gaggaactct	30
25	<210> 93 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 93	
30	agcaacttta aaactaaaag agctactga	30
35	<210> 94 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 94	
40	aaaaccctac acagtgtgtg agatgtgtca	30
45	<210> 95 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 95	
50	aatgggtcac agattgcat aataaggagg	30
55	<210> 96 <211> 30 <212> ADN <213> Streptococcus thermophilus <400> 96	

ES 2 398 918 T3

	tttttaaaa tccgcatgc tatactatat	30
	<210> 97	
	<211> 30	
5	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 97	
10	aattcaaaact ttctccaata ataccctcca	30
	<210> 98	
	<211> 30	
	<212> ADN	
15	<213> Streptococcus thermophilus	
	<400> 98	
20	catgctttca gtaataaga cgtgggacta	30
	<210> 99	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
25	<400> 99	
	tggaaggggt gtctagtga gaaattgtcg	30
30	<210> 100	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
35	<400> 100	
	ctcgaagcgc ttattgccc tattccttc	30
	<210> 101	
40	<211> 29	
	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 101	
45	atgtctaagg tatccactcg tgaaatcat	29
	<210> 102	
	<211> 30	
50	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 102	
55	atattaatgg aaattcatt caaacgcagt	30

ES 2 398 918 T3

<210> 103
 <211> 30
 <212> ADN
 5 <213> Streptococcus thermophilus

 <400> 103

 tagagagttt atatcctgat ggaatcgatg 30
 10
 <210> 104
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 15
 <400> 104

 tggcgaatta gagagccaat ggcaagcaag 30
 20
 <210> 105
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 25
 <400> 105

 agaagaccaa taaactgag aaaaagcaag 30
 30
 <210> 106
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 35
 <400> 106aaatggtcgt ttaattgta atgtcaaagc 30
 40
 <210> 107
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 45
 <400> 107

 caattgattc taaaatgctt ggtacacgta 30
 50
 <210> 108
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 55
 <400> 108

 tcttcggtt atcacagctt ctacacgttg 30

 <210> 109
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 109
 5 gaaatctcat tgaaccaac ttcaagacca 30

 <210> 110
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 110
 15 tgcttgtag ttgatgcact gcattagtaa 30

 <210> 111
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 111
 25 aatgtaccgg aatagcgta cattgcacat 30

 <210> 112
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 112

 35 ttcataaatt ctcacttttc ctactattc 30

 <210> 113
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 113

 45 tgtcgaaaaa attacctagt cacgacagac 30

 <210> 114
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 50 <400> 114
 caacaattac ttatgcatta ggaacatctg 30

 <210> 115
 <211> 30
 55 <212> ADN

ES 2 398 918 T3

<213> Streptococcus thermophilus
 <400> 115
 5 aattcgtgaa aaacaataaa aacaaaaaaaa 30
 <210> 116
 <211> 30
 <212> ADN
 10 <213> Streptococcus thermophilus
 <400> 116
 taacatttct gtccatttct tccttgatgc 30
 15 <210> 117
 <211> 28
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 117
 caaggcaact caaccaacca aattgacc 28
 25 <210> 118
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 118
 ctaaaatcgt aaatggtaag tgcacgatg 30
 35 <210> 119
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 119
 40 aacgtaagga gttttttat ttcttgta 30
 <210> 120
 <211> 30
 45 <212> ADN
 <213> Streptococcus thermophilus
 <400> 120
 50 gtggaaaatt tcacacccta catatatcaa 30
 <210> 121
 <211> 30
 <212> ADN
 55 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 121
 5 cctctgctaa tgacttaaac ggctcgttt 30
 <210> 122
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 122
 aaaatcaaag tttggggtt gtctacgtg 30
 15 <210> 123
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 123
 atatgtacat acctaaagaa aacacgggca 30
 25 <210> 124
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 124
 cgttgcaaa atatgtgatt acttgtatt 30
 35 <210> 125
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 125
 40 ccatagctgt aatgtgttt gtgactgct 30
 <210> 126
 <211> 30
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 126
 50 cgctaagttt ggcttaagt ataacaagct 30
 <210> 127
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 55

ES 2 398 918 T3

<400> 127
 aaagtacgct tcaaggcacg ttgaagacat 30
 5 <210> 128
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 128
 ctttttaacg tgtagcgtc ttagcttg 30
 <210> 129
 15 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 129
 20 ttggcttcgt gaataatmt taaaacgcat 30
 <210> 130
 <211> 30
 25 <212> ADN
 <213> Streptococcus thermophilus
 <400> 130
 30 tgttgaatca atacgctgaa acacactccc 30
 <210> 131
 <211> 30
 <212> ADN
 35 <213> Streptococcus thermophilus
 <400> 131
 cgttatcagt tgaaagttc aactcgtaag 30
 40 <210> 132
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 45 <400> 132
 taaactagtt ggcactatg ctccaggaag 30
 <210> 133
 50 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 133
 55

ES 2 398 918 T3

tagaccacca tagccgagtt gtcttttcg 30

<210> 134
 <211> 30
 5 <212> ADN
 <213> Streptococcus thermophilus

<400> 134acatcccact ttctgggtt ttagccatg 30

10 <210> 135
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

15 <400> 135

agtatggcta ttgtctgat actcatccac 30

20 <210> 136
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

25 <400> 136

cgctctgac gtggctggtg acatctacgc 30

30 <210> 137
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

<400> 137

35 gagtacatgg agtttctgct agatacacta 30

<210> 138
 <211> 29
 <212> ADN
 40 <213> Streptococcus thermophilus

<400> 138

45 taagttatga aatataaagt tattgtcta 29

<210> 139
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

50 <400> 139

aacgttatga catttaggag ctccaaatt 30

55 <210> 140

ES 2 398 918 T3

<211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 5 <400> 140
 aacacagcaa gacaaaagga tgacacttt 29
 <210> 141
 10 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 141
 15 caaccataac ttacgcatca ggtacatctg 30
 <210> 142
 <211> 29
 20 <212> ADN
 <213> Streptococcus thermophilus
 <400> 142
 25 acacgcgctt acctcgtata tcaaattca 29
 <210> 143
 <211> 30
 <212> ADN
 30 <213> Streptococcus thermophilus
 <400> 143
 tgcccgcaaa ctacgatac acaacagcat 30
 35 <210> 144
 <211> 31
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 144
 ctcaagctct tcatctgtga taggtgttt g 31
 45 <210> 145
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 50 <400> 145
 atcactcttt gatagtatct caaacgctgg 30
 <210> 146
 55 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus
 <400> 146
 5 gaaacagtca gaccagctaa ttcgccaatt 30
 <210> 147
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus
 <400> 147
 15 atatctcgaa agatacaagg acacttacac 30
 <210> 148
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 148
 25 gcggatgaaa cacaactca attgtattca 30
 <210> 149
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 149
 taatgctaca tctcaaagga tgatcccaga 30
 35 <210> 150
 <211> 31
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 150
 acgtctgtct aactggaaag tacctgctaa t 31
 <210> 151
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 151
 50 ctgttctcta atcgagaggc gcgtgattga 30
 <210> 152
 <211> 30
 55 <212> ADN

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 159
 5 acgttgatga atattgttga taaacttta 29

 <210> 160
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 160
 15 caagaagtga acaaagtaca cgctggaagt 30

 <210> 161
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 161
 25 gacagcaaga tacacgtagt tgatgaattg 30

 <210> 162
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30
 <400> 162

 taagaaatca acgcagattt ttagccaaca 30
 35
 <210> 163
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40
 <400> 163

 taaccaata attacagtga agcacaatag 30
 45
 <210> 164
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 164
 50
 caggcgtaag gtatgctaata tataacgat 29

 <210> 165
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 165
 5 gctatcgaac taatagctta gaggaactca 30

 <210> 166
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 166
 15 gtggaatatt aagcccgaat tgttcagca 30

 <210> 167
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 167
 25 tattgcaata ttgctgttg ggaaacctc 30

 <210> 168
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 168

 cgtctgtcta actggaaagt accggctaat 30
 35 <210> 169
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 169

 aaagagatgt accatccat tctaacaggt 30

 <210> 170
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 170
 50 ggggagtga ttcttcat caaacaatg 30

 <210> 171
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 171
 5 catcaaagtt gaaaaggact acaacagccc 30

 <210> 172
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 172
 15 cttaaattta gagcgtggga tcttgaatat 30

 <210> 173
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 173
 25 atataccgat ggcacatctg aaactggctg 30

 <210> 174
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 174

 taactcatat gtatcttgac caactatfff 30
 35 <210> 175
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 175

 aaatagcacc tctaagcgtt aatggattc 30

 <210> 176
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 176
 50 aatatctaca ggctactaca aagctacgct 30

 <210> 177
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 177
 5 gttgggggtg gtttgaacg gcgatgcta 30

 <210> 178
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 178
 15 tcaatcaggt gacggtgatg cttatattaa 30

 <210> 179
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 179
 25 catacatgat agtttgtaa cactttgat 30

 <210> 180
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 180

 tcagcattg gtttcatga cccacgtctg 30
 35 <210> 181
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 181

 caatcaacag gtttgactga ttataacgt 30

 <210> 182
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 182
 50 tagctacaca tgaatttat tacaatggg 30

 <210> 183
 <211> 30

ES 2 398 918 T3

<212> ADN
<213> Streptococcus thermophilus

<400> 183

5 cttacgtttg aaaagaatat caaatcaatg 30

<210> 184
<211> 30
10 <212> ADN
<213> Streptococcus thermophilus

<400> 184

15 ttaaaaaagg gcctttctct aaatcaagta 30

<210> 185
<211> 30
<212> ADN
20 <213> Streptococcus thermophilus

<400> 185

25 tgctgaacgt atctgtccac tgtgtggcca 30

<210> 186
<211> 30
<212> ADN
30 <213> Streptococcus thermophilus

<400> 186

ccgttctca aacgttaaat tccaaggtgt 30

35 <210> 187
<211> 30
<212> ADN
<213> Streptococcus thermophilus

40 <400> 187

gctgcgatta tgacaatgct gtctgtaagg 30

<210> 188
45 <211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 188

50 gaagaattta ttaataaaga tggttctgct 30

<210> 189
<211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 189
 5 aggcagaaaa gaagtatfff ggtaagtatg 30

 <210> 190
 <211> 29
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 190
 15 aaatggttta tcgacaagaa aatgaagct 29

 <210> 191
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 191
 25 ccaaattgac attatacaaa acgctccttc 30

 <210> 192
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 192

 atcctaactg ctttgtaac tacatcatgg 30
 35 <210> 193
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 193

 taacaagata agattagcgt ctcaacat 29

 <210> 194
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 194
 50 aaaagcctat gttgcccac tttgtggaag 30

 <210> 195
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 207
 5 aaattcgaca taagcactac agttatatt 29

 <210> 208
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 208
 15 ctattttcga gagaacgtca gtcattttaa 30

 <210> 209
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 209
 25 gtgctaacta tatcagtcgc atcaataaca 30

 <210> 210
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 210

 ttagcgggta ttggaataga ataagcgaat 30
 35 <210> 211
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 211

 ctctacagc agttaagac acattatcat 30

 <210> 212
 45 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 212
 50 cgtatcgaaa acggcgataa tccaacagt 29

 <210> 213
 <211> 31

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus
 <400> 213
 5 caataccttt ttttaattca tcttgataag t 31
 <210> 214
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus
 <400> 214
 15 ttaagaacaa tatcatcaat acgactttca 30
 <210> 215
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 215
 25 catctatcaa attcaaattc ggataaacta 30
 <210> 216
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 216
 tgagagtgtc tgatggattt attggaacc 30
 35 <210> 217
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 217
 acctcataca tggggaaaac ttgtaagta 29
 <210> 218
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 218
 50 tatttcacga atttctacac tttcaacct 30
 <210> 219
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 219
 5 ctgaaacctt gtttgaagc gcttgaagt 30

 <210> 220
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 220
 15 gtcaattgat actgcaatct cttaacatt 30

 <210> 221
 <211> 31
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 221
 25 acttcaatat ggtcaacatc ttgatcaccg a 31

 <210> 222
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 222

 taaactcgac aaaagcacta catgaatatt 30
 35 <210> 223
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 223
 40 atttttaag gaaaggagga aaataatata 30

 <210> 224
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 224
 50 cgttcaaaac agcgaaaact taaccctaac 30

 <210> 225
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 225
 5 cattaagtcg cttgaggcag acattgaagc 30

 <210> 226
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 226
 15 ccaaactcaa attgtctata ataataaccg 30

 <210> 227
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 227

 25 tatctctatt tcaggtgggt taaaacattc 30

 <210> 228
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 228

 aaacgaagat ggaagcgttg atgtttattc 30

 35 <210> 229
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 40 <400> 229

 gattgcattt gccagtattt ctttgatta 30

 <210> 230
 45 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 230
 50 tgaagacaac ggaacaatc aacctatta 29

 <210> 231
 <211> 30

ES 2 398 918 T3

<212> ADN
<213> Streptococcus thermophilus

<400> 231

5 acttctttt taatgcatc taagacaata 30

<210> 232
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 232

15 gccaatgatg ttcaattcgt taatggaatt 30

<210> 233
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 233

25 tcaacatggg atattcgtt ggcaggatg 30

<210> 234
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 234

30 tatggctctc ttgttgaat aaagatgatt 30

<210> 235
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 235

40 ataacatagc agtctattc ttgctgatg 30

<210> 236
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 236

50 gttaccacgc gccctactgt attagtgagg 30

<210> 237
<211> 30

ES 2 398 918 T3

```

<212> ADN
<213> Streptococcus thermophilus

<400> 243
5      gtgtatttag taatggtgat ttttaaatt                30

<210> 244
<211> 30
10     <212> ADN
       <213> Streptococcus thermophilus

<400> 244
15     cattcatttt ttatatatca ataaaacttt            30

<210> 245
<211> 30
<212> ADN
20     <213> Streptococcus thermophilus

<400> 245
25     ggggattctt attcactgt agttacgatg            30

<210> 246
<211> 30
<212> ADN
<213> Streptococcus thermophilus
30     <400> 246

       caaaaattga tgcacaatt aataaaggtg            30

35     <210> 247
       <211> 30
       <212> ADN
       <213> Streptococcus thermophilus

40     <400> 247

       ctatttctga caatggttga aattgtgtc            30

<210> 248
<211> 29
<212> ADN
<213> Streptococcus thermophilus

<400> 248
50     ctttttttaa attaatttat cgtaagcaa            29

<210> 249
<211> 30

```

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 249
 5 acaaaactta tgagaacggt tgaacggctt 30

 <210> 250
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 250
 15 agcccgctta ttgcttcagt tggttatat 30

 <210> 251
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 251
 25 tggagcaaca agaatgatta actctaagc 30

 <210> 252
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30
 <400> 252

 tttgatgat atcattgata aactatacga 30

 35 <210> 253
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 40 <400> 253

 taacgaaagc aataccaatc gtgctaaagc 30

 <210> 254
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 254
 50 tattcctatg gtcgatattc gaacagtcaa 30

 <210> 255
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 255
 5 caggggacaa ggactttgac ccaacagaag 30

 <210> 256
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 256
 15 agaaacacct aatggtctct tagaaccgga 30

 <210> 257
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 257
 25 aagaagtaa agacaacttt gttaaagact 30

 <210> 258
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30
 <400> 258
 35 gaaaaagcat ccatgatagt gcttagacct 30

 <210> 259
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40
 <400> 259
 45 cggaatgta taaagaatac aaagaaaacg 30

 <210> 260
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 50
 <400> 260
 ccaagtatca cgcaaagaaa tcaacgaga 29

 <210> 261
 <211> 31

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus
 <400> 261
 5 ttgacctggt tatccttggt aactagaata g 31
 <210> 262
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus
 <400> 262
 15 agagcactag catactgttt agtccgaacg 30
 <210> 263
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 263
 25 aggcaaggta ttgatccaa cagaagccaa 30
 <210> 264
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 264
 catgatttac aaccacgcg tagaccaag 29
 35 <210> 265
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 265
 acctagaagc attgagcgt atattgattg 30
 <210> 266
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 266
 50 aatthgccc ctctttgcc ccttgactag 30
 <210> 267
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus
 <400> 267
 5 taatagttta ccaaatcgtc ctgttccaa 30
 <210> 268
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus
 <400> 268
 15 accattagca atcattgtg cccattgagt 30
 <210> 269
 <211> 31
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 269
 25 acgtctgtct aactggaaag tacctgtaa t 31
 <210> 270
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 270
 ttttatact ttggtaatt acaaaatag 29
 35 <210> 271
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 271
 aagaaagaaa tattctagat atagatataa 30
 <210> 272
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 272
 50 caacgaccaa cacaacaact aaagtactg 30
 <210> 273
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 273
 5 tgattatggg tgtaaaca gaggcttatg 30

 <210> 274
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 274
 15 tgagtggtaa gtacaaatac gcaggactga 30

 <210> 275
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 275
 25 ttatttcctc ctttccttaa aaaaattaga 30

 <210> 276
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 276

 ggatgtatct gttgaaagag gtgtgtatat 30
 35 <210> 277
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 277

 aataggtaa aatatgcaa gtcacacaaa 30
 45 <210> 278
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 278
 50 aaaatggcat taaaattaa cataggaata 30

 <210> 279
 <211> 30

ES 2 398 918 T3

<212> ADN
<213> Streptococcus thermophilus

<400> 279

5 tatcagctcg taaatgttcg atagactct 30

<210> 280
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 280

15 attccattaa cgtattgac tcactagct 30

<210> 281
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 281

25 ctgttaccga tccaagagca gacatcatac 30

<210> 282
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 282

aagaagcgg taaatgctc aactgaatag 30

35 <210> 283
<211> 30
<212> ADN
<213> Streptococcus thermophilus

40 <400> 283

aattgctaaa catctaaaag acttaacggg 30

<210> 284
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 284

50 gatgaagatt tgactgatga taaagagaaa 30

<210> 285
<211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 285
 5 gacatcagaa agcagttat aaatattta 30

 <210> 286
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 286
 15 tttgaattta acaacctga tttgatatc 30

 <210> 287
 <211> 31
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 287
 25 tgatacggtc aaagttttc cactaatagc g 31

 <210> 288
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 288

 atggtttca ttcctgaac ccctaagagg 30
 35 <210> 289
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 40 <400> 289

 aagttattga aaaacgcaa catgatgagt 30

 <210> 290
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 290
 50 atataagtc tctattaat atccacaata 30

 <210> 291
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus
 <400> 291
 5 ttgcctcaag agatcctgct tgttgccaag 30
 <210> 292
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus
 <400> 292
 15 tcccatagtt ttaatgagtc ggtaactta 30
 <210> 293
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 293
 25 gtgtactaaa agtgtgctaa gttcataagg 30
 <210> 294
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 294
 35 atatagtgat tgtatccagc tgcggcgtag 30
 <210> 295
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 295
 45 aaaagcaaat cgcgagtata aaggatata 29
 <210> 296
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 296
 50 ttttaattga tctagacacc ctatgaaata 30
 <210> 297
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 297
 5 acagaggaga gaaacctg ctattttaga 30

 <210> 298
 <211> 29
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 298
 15 tggcagcagt gaattcgatg ccgagcaat 29

 <210> 299
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 299
 25 ccaaggaata ccaggtccta aaggtgccga 30

 <210> 300
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 30
 <400> 300

 ctaaatgaac tacaacaaca gcttgatga 29
 35 <210> 301
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 301

 tacctaaca ttttcgatat tttcaaatt 30

 <210> 302
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 302
 50 tttgactgct ttttatctg aattgtaatt 30

 <210> 303
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 303
 5 cagtaaccta aagctctatc aagcctatt 30

 <210> 304
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 304
 15 cgtaacgtg acagaccttg acaacaatc 30

 <210> 305
 <211> 29
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 305
 25 aggcataaat aacattgata accctaaca 29

 <210> 306
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 306

 gccaacgagg tcaaatatgt caacggcatt 30
 35 <210> 307
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 40 <400> 307

 gaaataggaa ctcaaaggt aattcttta 30

 <210> 308
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 308
 50 atttagagca aggaaagcag tacatcatta 30

 <210> 309
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 309
 5 ctgtaatcat ttttaatca ggattatcaa 30

 <210> 310
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 310
 15 ttaaagtgtat cctagtattt ttgtactata 30

 <210> 311
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 311
 25 ccatcagcca actgtatcgg ctactttcta 30

 <210> 312
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 312

 atgctctgg cgactatctc atggagcgtg 30
 35 <210> 313
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 313

 aggaaaaaac ccaaacaacc caaatgtta 30

 <210> 314
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 314
 50 tctaattctg tcaccacgac tatatcgcca 30

 <210> 315
 <211> 30

ES 2 398 918 T3

```

<212> ADN
<213> Streptococcus thermophilus

<400> 321
5      tgttctcata aatgccttt ccttttatg                30

<210> 322
<211> 30
10     <212> ADN
       <213> Streptococcus thermophilus

<400> 322
15     cttatcaaac atcaaggatt gtagatgagg            30

<210> 323
<211> 30
20     <212> ADN
       <213> Streptococcus thermophilus

<400> 323
25     atttcattag tagcttgata aatgtttcta            30

<210> 324
<211> 30
30     <212> ADN
       <213> Streptococcus thermophilus

<400> 32
       4gaaaatacta tactttaaaa gaaatttaa            30

35     <210> 325
       <211> 30
       <212> ADN
       <213> Streptococcus thermophilus

40     <400> 325

       tctcctccga cataatcttt tgtcttccg            30

<210> 326
45     <211> 30
       <212> ADN
       <213> Streptococcus thermophilus

<400> 326
50     acaaaaagcac tgccacctat agaagcatt            30

<210> 327
<211> 30

```


ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus
 <400> 333
 5 ataacagaag gagtagggga cgtaggcgcg 30
 <210> 334
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus
 <400> 334
 15 ttatttgata ggaatgtcag taattttga 30
 <210> 335
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 335
 aacatttcag cgcttactta tcaatctaata 30
 25
 <210> 336
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 30
 <400> 336
 gtattagtag gcatacgatt atggaagta 29
 35
 <210> 337
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40
 <400> 337
 catatatata tatatattta ttttaaataat 30
 45
 <210> 338
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 50
 <400> 338
 ttgtcataat aattaaatcc aataggact 30
 <210> 339
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 339
 5 gaaaattct gttgttct taatattagc 30

 <210> 340
 <211> 30
 10 <212> ADN
 <213> Streptococcus vestibularis

 <400> 340
 15 gtacttcaaa ggttctaact acataacaca 30

 <210> 341
 <211> 30
 <212> ADN
 20 <213> Streptococcus vestibularis

 <400> 341
 25 taaaaccaga tgggtgttct tctgatacta 30

 <210> 342
 <211> 30
 <212> ADN
 <213> Streptococcus vestibularis
 30 <400> 342

 cattttctc agtcaattcg ttctcaagcg 30
 35 <210> 343
 <211> 30
 <212> ADN
 <213> Streptococcus vestibularis
 40 <400> 343

 aaaggacggg ggcaatgaac aaacgacaac 30

 <210> 344
 45 <211> 29
 <212> ADN
 <213> Streptococcus vestibularis

 <400> 344
 50 taatatcatt gatagctca tcaaaggct 29

 <210> 345
 <211> 29

ES 2 398 918 T3

<212> ADN
<213> Streptococcus vestibularis
<400> 345
5 taaattgttc cttgactccg aactgcct 29
<210> 346
<211> 30
10 <212> ADN
<213> Streptococcus vestibularis
<400> 346
15 aaacaatcgt ttatctatcc tcaaaggatg 30
<210> 347
<211> 30
<212> ADN
20 <213> Streptococcus vestibularis
<400> 347
25 ataaaaaac gcctcaaaaa ccgagacaac 30
<210> 348
<211> 30
<212> ADN
30 <213> Streptococcus vestibularis
<400> 348
tggaatccc ttatatcgac aaatacgta 30
35 <210> 349
<211> 30
<212> ADN
<213> Streptococcus vestibularis
40 <400> 349
ttccagtcg tgatttta tgaataccc 30
<210> 350
45 <211> 30
<212> ADN
<213> Streptococcus vestibularis
<400> 350
50 ggacatcgaa caagtcaatg ccgtaagctt 30
<210> 351
<211> 30

ES 2 398 918 T3

```

<212> ADN
<213> Streptococcus vestibularis

<400> 351
5   aatctttaac cggattgtag aaccgttcgg           30

<210> 352
<211> 30
10  <212> ADN
    <213> Streptococcus vestibularis

    <400> 352

15  tgcctttaaa ataactagat ttaccatca           30

    <210> 353
    <211> 29
    <212> ADN
20  <213> Streptococcus vestibularis

    <400> 353

25  gagcaagcac aagcaagctt tactatcct           29

    <210> 354
    <211> 30
    <212> ADN
30  <213> Streptococcus vestibularis

    <400> 354

    cagattggtt tatcgaacaa ggtcgcaagt           30

35  <210> 355
    <211> 31
    <212> ADN
    <213> Streptococcus vestibularis

40  <400> 355

    caaaagctgt tggtaacgg tgcttgggc a           31

45  <210> 356
    <211> 30
    <212> ADN
    <213> Streptococcus vestibularis

    <400> 356

50  cttgttttc ctctggggtc tctgcgact           30

    <210> 357
    <211> 30

```

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus vestibularis
 <400> 357
 5 gaaataaact gcccaaacat tttattttc 30
 <210> 358
 <211> 30
 10 <212> ADN
 <213> Streptococcus vestibularis
 <400> 358
 15 tgagtaagcg acaagctaga aatcaagtca 30
 <210> 359
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 359
 25 atagctaaga tggaagaagc atcaagcacc 30
 <210> 360
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 360
 cagtatctca aacgctggat acaacaagat 30
 35 <210> 361
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 361
 cctactcagt ggacacctgc aattgaagac 30
 <210> 362
 45 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 362
 50 cgattggaac gggtgcttat ggcctaac 29
 <210> 363
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 363
 5 gcgaacaatt gaatttgta gaaaatgctg 30

 <210> 364
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 364
 15 gaagcattta ttaatataga tggttctgct 30

 <210> 365
 <211> 28
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 365
 25 tgctgacgta tctgtccact gtgtgcca 28

 <210> 366
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 366

 tttttatact ttgggtaaatac taaaaatag 30
 35 <210> 367
 <211> 31
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 367

 tcaagggtgc gccttatgga aaagatgctt g 31

 <210> 368
 45 <211> 31
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 368
 50 tgtaaaaatt tctagacggt tagacacttt a 31

 <210> 369
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 369
 5 aaatgatgat tgaatgcttg agatagcagt 30

 <210> 370
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 370
 15 aataagaagt tcttgacgac caaccgacat 30

 <210> 371
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 371

 25 tcgtcaacgt cgatacagaa caacgtgctt 30

 <210> 372
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30
 <400> 372

 35 tgattagcaa atttaaaca ggatatttg 30

 <210> 373
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 373

 45 aaagacaagc ccaaggatt gaactagcaa 30

 <210> 374
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 50 <400> 374

 50 cgaacagtg gcgagaaatc cgctggcgt 30

 <210> 375
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 375
 5 ctacattatt gatcatgitt tttctcctgt 30

 <210> 376
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 376
 15 tagaaggctc tggaaataca aagcaattct 30

 <210> 377
 <211> 31
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 377
 25 tagaaggctc tggtaaatac aaagcaattc t 31

 <210> 378
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 378

 tctgatggct cttggtagg aactggatat 30
 35 <210> 379
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 379

 tttgatggct cttggtagg aactggatat 30

 <210> 380
 45 <211> 31
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 380
 50 ttttgatggc tcttgtagg gaactggata t 31

 <210> 381
 <211> 29

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 381
 5 acagaacaaa atggtagaat atatcatct 29

 <210> 382
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 382
 15 ccctggacaa gctatcagca catatccttg 30

 <210> 383
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 383
 25 cgctgtgat gtaaccgct ttatatatat 30

 <210> 384
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 384

 gaatgaatgt attagagcaa gcactgacc 30
 35 <210> 385
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 385

 tagacgaaaa ggaaggaaaa tagcatgagc 30

 <210> 386
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

 <400> 386
 50 ataactcgat tgctaactta agcaagcagt 30

 <210> 387
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus
 <400> 387
 5 ctgcatgtgt aaccatgact tcttcgtcgt 30
 <210> 388
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus
 <400> 388
 15 cttcgtgga aactcgtag tcatacatac 30
 <210> 389
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 389
 25 aagaccgctg tactggttgg tattcgtacc 30
 <210> 390
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 390
 caaccaagcg aacacagcag tagcaccgca 30
 35 <210> 391
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 391
 atgatgatga agtatcgtca tctactaac 29
 <210> 392
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 392
 50 cttcacctca aatcttagag atgactaaa 30
 <210> 393
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus
 <400> 393
 5 aaaaggtgcg tatgaaactc atcccagcgg 30
 <210> 394
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus
 <400> 394
 15 aagggttaa gtcctcata gagtggaaaa 30
 <210> 395
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 395
 25 cctcaaagct taaaattggg ctgaagtaga 30
 <210> 396
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 396
 35 gcaatttatt cgcttgatgt actcacgttt 30
 <210> 397
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 397
 45 tattattgc aaatggttac catatttta 30
 <210> 398
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 50 <400> 398
 55 tatttagca ctacggatc agcgtatctc 30
 <210> 399
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus
 <400> 399
 5 tgctacgtgc tctggacggg cgctatcagc 30
 <210> 400
 <211> 30
 10 <212> ADN
 <213> Streptococcus thermophilus
 <400> 400
 15 aaatgaacag acaagaagca acagaaattg 30
 <210> 401
 <211> 30
 <212> ADN
 20 <213> Streptococcus thermophilus
 <400> 401
 25 aagttgatcg tatctattta gaatatcgca 30
 <210> 402
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 402
 attcactttg acagatacta atgctacatc 30
 35 <210> 403
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 40 <400> 403
 caagcagtg aaagtggtt tatatgtaa 30
 <210> 404
 45 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 404
 50 catagtatag ccgtcttctt tgattgattg 30
 <210> 405
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

 <400> 405
 5 ccatgggtgc taaagtgat gactaccgct 30

 <210> 406
 <211> 36
 10 <212> ADN
 <213> Streptococcus thermophilus

 <400> 406
 15 ttctaggaa tgggtaatta tagcgagcta gaaagc 36

 <210> 407
 <211> 38
 <212> ADN
 20 <213> Streptococcus thermophilus

 <400> 407
 25 agtgggaag gtctggaaa atctatggca aaaaacct 38

 <210> 408
 <211> 35
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 408

 35 tatatggtc aaatgcgatt caagactat tcaaa 35

 <210> 409
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 40 <400> 409

 45 taattgcaa tgctacaat atctctgca 30

 <210> 410
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 50 <400> 410
 atgttctgaa ttaccttct cgacactccg 30

 <210> 411
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus agalactiae
 <400> 411
 5 accatcaagg ctcttatctg cagattgta 30
 <210> 412
 <211> 30
 10 <212> ADN
 <213> Streptococcus agalactiae
 <400> 412
 15 aaatggttgc caatgacttt ctagagtgat 30
 <210> 413
 <211> 30
 <212> ADN
 20 <213> Streptococcus agalactiae
 <400> 413
 25 acaaaatctt ttgttgctcc tggacgtatt 30
 <210> 414
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 30 <400> 414
 atgtaaggta ttgtaaaact tcttcttgcg 30
 35 <210> 415
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 40 <400> 415
 actgttcta taattaaaat aaaagaggta 30
 <210> 416
 45 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 <400> 416
 50 tgttccagta aaaagtaatt ttaaagcatt 30
 <210> 417
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus agalactiae

 <400> 417
 5 cgctcgattg atgctatcaa ctatattgaa 30

 <210> 418
 <211> 30
 10 <212> ADN
 <213> Streptococcus agalactiae

 <400> 418
 15 ttctcaaga gaactgtag aacagcttca 30

 <210> 419
 <211> 30
 <212> ADN
 20 <213> Streptococcus agalactiae

 <400> 419
 25 aaggctactt tagctgttc ttgtggtgtt 30

 <210> 420
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 30 <400> 420

 acagctactg taaattctgc ttttaggtt 30
 35 <210> 421
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 40 <400> 421

 tagtgcagtt gtcaaggaga ttgtgagcga 30

 <210> 422
 45 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae

 <400> 422
 50 tttaaccttt gaaaatgtga aaggctcgta 30

 <210> 423
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus agalactiae

 <400> 423
 5 gcgatgatgg taagtcacatca tggacagcgt 30

 <210> 424
 <211> 30
 10 <212> ADN
 <213> Streptococcus agalactiae

 <400> 424
 15 ttttacacac gatgtcagat ataatgtaa 30

 <210> 425
 <211> 30
 <212> ADN
 20 <213> Streptococcus agalactiae

 <400> 425
 25 agtactgcac taggaattgt agagatcaaaa 30

 <210> 426
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 30 <400> 426

 cgtaccatct atcaatttac cgcaagctgt 30
 35 <210> 427
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 40 <400> 427

 ttaaagatt taaactatca agcgtcaatt 30

 <210> 428
 45 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae

 <400> 428
 50 ttctaaatgc tggtgactgc ttgcataaaa 30

 <210> 429
 <211> 31

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus agalactiae

 <400> 435
 5 gctggagatt ttacaagcag ttgaatttc 30

 <210> 436
 <211> 30
 10 <212> ADN
 <213> Streptococcus agalactiae

 <400> 436
 15 atcacaccag tcggtatgat ggatgactat 30

 <210> 437
 <211> 30
 <212> ADN
 20 <213> Streptococcus agalactiae

 <400> 437
 25 tgtcaacagt acgtgagacg agtgtgtagg 30

 <210> 438
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 30 <400> 438

 tgaagtgat ggatagtgg atttagagct 30
 35 <210> 439
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 40 <400> 439

 taatcatttt atgagagata ccgcctcaag 30

 <210> 440
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 45 <400> 440

 50 tttaaagaga tatctgttc atctgcgga 30

 <210> 441
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus agalactiae
 <400> 441
 5 aatcactct gcataaatat ctttacttc 30
 <210> 442
 <211> 30
 10 <212> ADN
 <213> Streptococcus agalactiae
 <400> 442
 15 aaacatccgc aacgggataa ataaagctag 30
 <210> 443
 <211> 30
 <212> ADN
 20 <213> Streptococcus agalactiae
 <400> 443
 25 agtttctgt gggtagctt gtccaccgta 30
 <210> 444
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 30 <400> 444
 gaacatgaaa gatttataaa aagaacatt 30
 35 <210> 445
 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 40 <400> 445
 agaggggaaa atatcaatgc cgaatgctga 30
 <210> 446
 45 <211> 30
 <212> ADN
 <213> Streptococcus agalactiae
 <400> 446
 50 gatggtacaa aatcattgt tggactgat 30
 <210> 447
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus mutans

 <400> 447
 5 aaaaggaaac gccattaatt aatatggtga 30

 <210> 448
 <211> 30
 10 <212> ADN
 <213> Streptococcus mutans

 <400> 448
 15 gattgaacca gctagcgag ttagtgctct 30

 <210> 449
 <211> 30
 <212> ADN
 20 <213> Streptococcus mutans

 <400> 449
 25 cgctaaaagc tgtgtgtca tcatagttag 30

 <210> 450
 <211> 30
 <212> ADN
 <213> Streptococcus mutans
 30
 <400> 450

 taaatattt caattagaca atagacaaac 30
 35 <210> 451
 <211> 30
 <212> ADN
 <213> Streptococcus mutans
 40 <400> 451

 tgcctatgta ttcggacatg acttgccaca 30

 <210> 452
 45 <211> 30
 <212> ADN
 <213> Streptococcus mutans

 <400> 452
 50 atgtgaaaag aaagtaacta ctacattga 30

 <210> 453
 <211> 30

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus pyogenes

 <400> 453
 5 tgcgctgggt gatttctct tgcgctttt 30

 <210> 454
 <211> 30
 10 <212> ADN
 <213> Streptococcus pyogenes

 <400> 454
 15 ttatatgaac ataactcaat ttgtaaaaaa 30

 <210> 455
 <211> 30
 <212> ADN
 20 <213> Streptococcus pyogenes

 <400> 455
 25 aggaatatcc gcaataatta attgcgctct 30

 <210> 456
 <211> 30
 <212> ADN
 <213> Streptococcus pyogenes
 30
 <400> 456

 taaattgtt tagcaggtaa accgtgctt 30

 35 <210> 457
 <211> 30
 <212> ADN
 <213> Streptococcus pyogenes

 40 <400> 457

 ttcagcacac tgagactgt tgagtccat 30

 <210> 458
 45 <211> 33
 <212> ADN
 <213> Streptococcus pyogenes

 <400> 458
 50 ctgtgacatt gcgggatgta atcaaagtaa aaa 33

 <210> 459
 <211> 32

ES 2 398 918 T3

<212> ADN

<213> Streptococcus pyogenes

<400> 459

5

aaagcaaacc tagcagaagc agaaaatgac tt 32

<210> 460

<211> 34

10

<212> ADN

<213> Streptococcus pyogenes

<400> 460

15

tgatgtaatt ggtgatttc gtgatatgct tttt 34

<210> 461

<211> 5849

<212> ADN

20

<213> Streptococcus thermophilus

<400> 461

ES 2 398 918 T3

atgagtgact tagttttagg acttgatatac ggtatagggtt ctgttggtgt aggtatcctt 60
 aacaaagtga caggagaaat tatccataaa aactcacgca tcttcccagc agctcaagca 120
 gaaaataacc tagtacgtag aacgaatcgt caaggaagac gcttgacacg acgtaaaaaa 180
 catcgtatag ttcgtttaaa tcgtctatctt gaggaaagtg gattaatcac cgattttacg 240
 aagatttcaa ttaatcttaa cccatatcaa ttacgagtta agggcttgac cgatgaattg 300
 tctaataagag aactgtttat cgctcttaaa aatatgggtga aacaccgtgg gattagttac 360
 ctcgatgatg ctagtgatga cggaaattca tcagtaggag actatgcaca aattgttaag 420
 gaaaatagta aacaattaga aactaagaca cggggacaga tacagttgga acgctaccaa 480
 acatatggtc aattacgtgg tgattttact gttgagaaaag atggcaaaaa acatcgcttg 540
 attaatgtct ttccaacatc agcttatcgt tcagaagcct taaggatact gcaaactcaa 600
 caagaattta atccacagat tacagatgaa tttattaatc gttatctcga aattttaact 660
 ggaaaacgga aatattatca tggaccgga aatgaaaagt cacggactga ttatggtcgt 720
 tacagaacga gtggagaaac tttagacaat atttttggaa ttctaattgg gaaatgtaca 780
 ttttatccag aagagtttag agcagcaaaa gcttcttaca cggctcaaga attcaatttg 840
 ctaaatagatt tgaacaatct aacagttcct actgaaacca aaaagttgag caaagaacag 900
 aagaatcaaa tcattaatta tgtcaaaaat gaaaaggcaa tggggccagc gaaacttttt 960
 aatatatcg ctaagttact ttcttgtgat gttgcagata tcaagggata ccgatatcagc 1020
 aatcaggta aggctgagat tcatactttc gaagcctatc gaaaaatgaa aacgcttgaa 1080
 accttagata ttgaacaaat ggatagagaa acgcttgata aattagccta tgtcttaaca 1140
 ttaaacactg agaggaaggt tattcaagaa gccttagaac atgaatttgc tgatggtagc 1200
 tttagccaga agcaagttga cgaattgggt caattccgca aagcaaatag ttccattttt 1260
 ggaaaaggat ggcataatct ttctgtcaaa ctgatgatgg agttaattcc agaattgtat 1320
 gagacgtcag aagagcaaat gactatcctg acacgacttg gaaaacaaaa acgacttcgt 1380
 cttcaaataa aacaaaatat ttcaaataaa acaaaaatata tagatgagaa actattaact 1440
 gaagaaatct ataatcctgt tgttgctaag tctgttcgcc aggcataaaa aatcgtaaat 1500
 gcggcgatta aagaatacgg agactttgac aatattgtca tcgaaatggc tcgtgaaaca 1560
 aatgaagatg atgaaaagaa agctattcaa aagattcaaa aagccaacaa agatgaaaaa 1620
 gatgcagcaa tgcttaaggc tgctaaccaa tataatggaa aggcctgaatt accacatagt 1680
 gttttccacg gtcataagca attagcgact aaaatccgcc tttggcatca gcaaggagaa 1740
 cgttgccttt atactggtaa gacaatctca atccatgatt tgataaataa tcctaactcag 1800

ES 2 398 918 T3

tttgaagtag atcatattht acctctthtct atcacattcg atgatagcct tgcaataaag 1860
 gttttggtht atgcaactgc taaccaagaa aaaggacaac gaacacctta tcaggcttht 1920
 gatagtatgg atgatgctg gtctthtccgt gaattaaaag cthtthtgtag tgagtcaaaa 1980
 acactthtcaa acaagaaaa agaatacctc cttacagaag aagatatht c aaagthtthgat 2040
 gttcgaaaga aaththattga acgaaatctt gtagatacaa gatacgcctc aagagthtgc 2100
 ctcaatgccc ttcaagaaca cthttagagct cacaagattg atacaaaagt thcctgtht 2160
 cgtggccaat ttacatctca attgagacgc cattggggaa ttgagaagac tcgtgatact 2220
 tatcatcacc atgctgctga tgcattgatt attgccgcct caagtcagtht gaaththtgg 2280
 aaaaaacaaa agaatacctc tgtaagtht tcaagaagac aactcctga thttgaaaca 2340
 ggtgaactta thagtthtga tgagtacaag gaathtctgt tcaaagcccc thtthtcaacat 2400
 thtthtthgata cattgaagag taaagaatht gaagacagta thtthtthtctc atathtcaagtht 2460
 gaththtctagtht thaatctgtaa aaththtctagat gccactatht atgctgacaag acaggctaaa 2520
 gtgggaaaaag athaagaagga tgaaacttht gtctthtggga aaththtcaaaga thtctatact 2580
 caggatgtht atgatgcctt thtgaagatt thataagaagg athaagthtcaa thtctctcatg 2640
 thtctgctcag ccccaacaaac cthtthtggaaa gthtthtctgagc caaththtthtga gaactthtctc 2700
 aathtgaacaaa thaatgaaaa aggaaaagag gthtthtctgta atcctthtctc aaththtataaaa 2760
 gaagaacatg gctaththtctg thaatathtga aaaaaaggca atggthtctga aaththtcaagagtht 2820
 cthtthtatact atgatagtht gctthtthttagtht aaththtcttht ataththtctc agagaathtga 2880
 aaaaaataaag thtthtctthtca gthtthtthttaa cctthtggagaa cagathtthtctc thtthtcaataag 2940
 gctactggaa aaththtctgaaat cctthtggatht aaththtctgctg atctthtcaatht thgagaaggg 3000
 acaggaacat athaagaththt ccaggaaaaa thtcaaththtga thtthtgaaaaa agagggthtga 3060
 gaththtctgatht cagaaththtcaa gthtthtctact thtthtthttht atthtthtthtctc cgtthtthtgaagat 3120
 acagaaacaa aagaacacaa gctthtthtccgt thtthtthtctc gaactthtacc thtthtcaaaaag 3180
 cattaththtgg aaththtthtacc thtthtthtga thtthtthttht aaggagthtga ggcgthttht 3240
 aaagthtthtgg gthtthtctgctg thtthtthtggtht caaththtcthataa aaggactagc aaththtcaatht 3300
 aththtctaththt athaagthtga aacagaththtctc ctaggaaatht agcaththtctc caaththtgaag 3360
 ggtgathtgaagc ctgaagctga thtthtthttht thtthtthtthtga aagthtthtctg aaththtthtga 3420
 atcataththtctc athaaththtctg thtthtthtggagc gcctthtcthatac gthtthtcttht thtthtthtctgaga 3480
 gcthatacaaga thtthtggctthtca thtthtggaaatht aacacctgtht cattthtthtthtga caggthttht 3540
 thtthtthtthtga aagagthtga gaaaththtctc ggagaththtctg acaththtctcagtht caagthtthtga 3600
 agaththtctgctc aagctthtga aactthtthtga thtthtthtthttht gggacaaagag thtthtctgctg 3660
 cacthaagthtga thtthtthtctgata atcgtthtctgag aagthtthtggga thtthtctgtht accctthtctg 3720

ES 2 398 918 T3

tattaagtgc cttaagtaaa tataatattg ccttggtcgt ttgtgataac gaacatttac 3780
caacaggaat ttatcactca caaaatgggc actttagagc gtacaagcgc ttgaaagaac 3840
agctggattg gtctcagaaa caaaaggaca aggcattggca gattgtaact tattataaaa 3900
tcaataacca agaggatggt ctagccatgt ttgaaaaaag tctggacaac attagattac 3960
tttcagacta taaagagcag atagaacctg gtgatagaac gaatagagag ggacatgctg 4020
ccaaggtcta ctttaatgag ctctttggta aacaatttgt cagagtaact cagcaagaag 4080
ctgatgtcat caatgctggt ttaaactatg gctatgctat catgagggct cagatggcta 4140
gaatagtggc gggttatggt ttaaattggc tattaggaat cttccataaa aatgaataca 4200
atcagtttaa tttggttgac gatttgatgg agccatttag acagattgta gatgtttggg 4260
tatatgataa tctacgagat caggaattcc ttaagtatga gtataggttg ggattgacag 4320
atttactcaa tgctaaaatc aaatatggca aagagacttg ctcaagtaca gttgctatgg 4380
acaaatatgt caaaggcttt atcaaatata tttcggaaaa agatagtagt aaatttctact 4440
gcccagtggt atcaagttta gagtggagaa aataagatga ggtatgaagc attgagatta 4500
ttatgttttt ttgatttacc aatggaatcc aaggatgaaa aaagaatata tcgtaatttt 4560
cgtaaagaat taatttcaaa tgggtttgaa atgttacaat tttcggctca ctatcgcact 4620
tgtcctaata gaagctttgc aaataaattt tataagaagt taaagattag caatcttcct 4680
gctgggaatg tgagactttt ggcagttact gaaaaacaat tttcagagat gacattaatt 4740
ataggtggta aaactaagca agaagaaatc gtcagtgata ataagttggt ggttatatga 4800
aatattttgt acaacatcct tacaagaac gtattgaatt aaatattggt gcaatcacac 4860
aaattgfttg tcagaataaa gaactcaaat attatatttg gcaaattttg agctggattt 4920
ttggcggaaa aaaatactca agtgaggact taagtatttt tgattatgag gaacctacta 4980
tacttgatga gtctggagaa atagtgaagc gaagtagcta tcactatata gacatttcaa 5040
gttttaagga tttactggag cagatggaat acaagaaagg aaccttgct cagggttacc 5100
ttagtaaaaat tctcaatcag gttgatattg taggccattt ggagaaaatt aatgaacaag 5160
tagagcttat agaaggagca atgaatcagc atataaactt aaactgtggt caggtggagt 5220
accatttga gaatcacct ctaacactag accaattact ttcaaaaaat tttagtcct 5280
tttttgctat cgagaataag aatttatctt ttgaatgggt ttcaataact gataaacttt 5340
ctctctttct agaaatgta gaccgccttc tgtcaciaaac aacagagaag tatctcattg 5400
tgctaaaaaa tattgatggc tttatctcag aagaatctta tactattttt tataggcaaa 5460
tctgtcatct ggtcaagaag tatccaaatc taacctttat tttgtttcct agtgaccaag 5520
gctatttaaa aattgatgaa gaaaatagta ggctcgtcaa tattttatct gaccaggtgg 5580
agcatttgta tgatgttgag tttatgtatg aaagagtaat gaaatattat ccaagtaatg 5640

ES 2 398 918 T3

at t t t c c g a c g a g a g a a g g t t t t a g g a t g t c t t t a g a a a c t g t g a c a c c t t a t t t a t t g a 5700
c a a a a t g c t g a g a c a a c c t a g t c t c t c a c t t g t t g a t t c a g t a a t a t t g a a t a t c c t a a 5760
a t c a g t t g t t t c a t t t t a g t t a c c g t a t a a g a t a t t c t c a g a c a c c t g a t a a g g a a c t a t 5820
t a c a t a a a t t t t t a g a a a g t a a g g a t t g a 5849

<210> 462

5 <211> 3387

<212> ADN

<213> Streptococcus thermophilus

<400> 462

10

ES 2 398 918 T3

atgagtgact tagttttagg acttgatata ggtatagggt ctgttgggtg aggtatcctt 60
 aacaaagtga caggagaaat tatccataaa aactcacgca tcttcccagc agctcaagca 120
 gaaaataacc tagtacgtag aacgaatcgt caaggaagac gcttgacacg acgtaaaaaa 180
 catcgtatag ttcgtttaaa tcgtctatct gaggaaagtg gattaatcac cgattttacg 240
 aagatttcaa ttaatcttaa cccatatcaa ttacgagtta agggttgac cgatgaattg 300
 tctaataag aactgtttat cgctcttaaa aatatgggtga aacaccgtgg gattagttac 360
 ctcgatgatg ctagtgatga cggaaattca tcagtaggag actatgcaca aattgttaag 420
 gaaaatagta aacaattaga aactaagaca ccgggacaga tacagttgga acgctaccaa 480
 acatatggtc aattacgtgg tgattttact gttgagaaaag atggcaaaaa acatcgcttg 540
 attaatgtct ttccaacatc agcttatcgt tcagaagcct taaggatact gcaaactcaa 600
 caagaattta atccacagat tacagatgaa tttattaatc gttatctcga aattttaact 660
 ggaaaacgga aatattatca tggaccgga aatgaaaagt cacggactga ttatggtcgt 720
 tacagaacga gtggagaaac tttagacaat atttttggaa ttctaattgg gaaatgtaca 780
 tttatccag aagagtttag agcagcaaaa gcttctaca cggctcaaga attcaatttg 840
 ctaaatagatt tgaacaatct aacagttcct actgaaacca aaaagttgag caaagaacag 900
 aagaatcaaa tcattaatta tgtcaaaaat gaaaaggcaa tggggccagc gaaacttttt 960
 aaatatatcg ctaagttact ttcttgtgat gttgcagata tcaagggata ccgatcgcac 1020
 aatcaggta aggctgagat tcatactttc gaagcctatc gaaaaatgaa aacgcttgaa 1080
 acctagata ttgaacaaat ggatagagaa acgcttgata aattagccta tgtcttaaca 1140
 ttaaactctg agaggaagg tattcaagaa gccttagaac atgaatttgc tgatggtagc 1200
 tttagccaga agcaagttga cgaattgggt caattccgca aagcaaatag ttccattttt 1260
 ggaaaaggat ggcataatct ttctgtcaaa ctgatgatgg agttaattcc agaattgtat 1320
 gagacgtcag aagagcaaat gactatcctg acacgacttg gaaaacaaaa acgacttcgt 1380
 cttcaataaa aacaaaatat ttcaataaaa acaaaatata tagatgagaa actattaact 1440
 gaagaaatct ataatcctgt tgttgctaag tctgttcgcc aggctataaa aatcgtaaat 1500

ES 2 398 918 T3

gcggcgatta aagaatacgg agactttgac aatattgtca tcgaaatggc tcgtgaaaca 1560
aatgaagatg atgaaaagaa agctattcaa aagattcaaa aagccaacaa agatgaaaaa 1620
gatgcagcaa tgcttaaggc tgctaaccaa tataatggaa aggctgaatt accacatagt 1680
gttttccacg gtcataagca attagcgact aaaatccgcc tttggcatca gcaaggagaa 1740
cgttgccttt atactggtaa gacaatctca atccatgatt tgataaataa tcctaatacag 1800
tttgaagtag atcatatfff acctctttct atcacattcg atgatagcct tgcaaataag 1860
gttttggttt atgcaactgc taaccaagaa aaaggacaac gaacacctta tcaggcttta 1920
gatagtatgg atgatgcgtg gtctttccgt gaattaaaag cttttgtacg tgagtcaaaa 1980
acactttcaa acaagaaaaa agaatacctc cttacagaag aagatatttc aaagtttgat 2040
gttcgaaaga aatttattga acgaaatctt gtagatacaa gatagccttc aagagttgtc 2100
ctcaatgccc ttcaagaaca ctttagagct cacaagattg atacaaaagt ttccgtgggt 2160
cgtggccaat ttacatctca attgagacgc cattggggaa ttgagaagac tcgtgatact 2220
tatcatcacc atgctgtcga tgcattgatt attgccgcct caagtcagtt gaatttgtgg 2280
aaaaaaciaa agaatacctc tgtaagttat tcagaagaac aactccttga tattgaaaca 2340
ggtgaactta ttagtgatga tgagtacaag gaatctgtgt tcaaagcccc ttatcaacat 2400
tttgttgata cattgaagag taaagaattt gaagacagta tcttattctc atatcaagtg 2460
gattctaagt ttaatcgtaa aatatcagat gccactattt atgcgacaag acaggctaaa 2520
gtgggaaaag ataagaagga tgaaacttat gtcttaggga aatcaaaga tatctatact 2580
caggatgggt atgatgcctt tatgaagatt tataagaagg ataagtcaaa attcctcatg 2640
tategtcacg accacaaaac ctttgagaaa gttatcgagc caattttaga gaactatcct 2700
aataagcaaa tgaatgaaaa aggaaaagag gtaccatgta atcctttcct aaaatataaa 2760
gaagaacatg gctatattcg taaatatagt aaaaaaggca atggctcctga aatcaagagt 2820
cttaaatact atgatagtaa gcttttaggt aatcctattg atattactcc agagaatagt 2880
aaaaataaag ttgtcttaca gtcattaaaa ccttgagaaa cagatgtcta tttcaataag 2940
gctactggaa aatacgaat ccttgatta aaatatgctg atctacaatt tgagaaaggg 3000
acaggaacat ataagatttc ccaggaaaaa tacaatgaca ttaagaaaaa agagggtgta 3060
gattctgatt cagaattcaa gtttacactt tataaaaatg atttgttact cgttaaagat 3120
acagaaacia aagaacaaca gcttttccgt tttctttctc gaactttacc taaacaaaag 3180
cattatgttg aattaaacc ttatgataaa cagaaatttg aaggaggtga ggcgttaatt 3240
aaagtgttgg gtaacgttgc taatggtggt caatgcataa aaggactagc aaaatcaaat 3300
atctctatff ataaagtaag aacagatgtc ctaggaaatc agcatatcat caaaaatgag 3360
ggtgataagc ctaagctaga tttttaa 3387

ES 2 398 918 T3

<210> 463
 <211> 912
 <212> ADN
 5 <213> Streptococcus thermophilus

<400> 463

atgacttgga	gagttgtaca	tgtcagtcaa	agtgagaaga	tgcgcttaaa	gcttgataac	60
ttattagtgc	aaaaaatggg	acaagagttt	acggtgccac	taagtgatat	ttcgataatc	120
gttgcagaag	gtggggatac	agttgttacc	cttcgtctat	taagtgcctt	aagtaaatat	180
aatattgcct	tggtcgtttg	tgataacgaa	catttaccaa	caggaattta	tcactcacia	240
aatgggcact	ttagagcgta	caagcgcttg	aaagaacagc	tggattggtc	tcagaaacia	300
aaggacaagg	catggcagat	tgtaacttat	tataaaatca	ataaccaaga	ggatgttcta	360
gccatgtttg	aaaaaagtct	ggacaacatt	agattacttt	cagactataa	agagcagata	420
gaacctggtg	atagaacgaa	tagagagggg	catgctgcca	aggtctactt	taatgagctc	480
tttggtaaac	aatttgtcag	agtaactcag	caagaagctg	atgtcatcaa	tgctggttta	540
aactatggct	atgctatcat	gagggctcag	atggctagaa	tagtggcggg	ttatggttta	600
aatggcctat	taggaatctt	ccataaaaat	gaatacaatc	agtttaattt	ggttgacgat	660
ttgatggagc	catttagaca	gattgtagat	gtttgggtat	atgataatct	acgagatcag	720
gaattcctta	agtatgagta	taggttggga	ttgacagatt	tactcaatgc	taaaatcaaa	780
tatggcaaag	agacttgctc	agtgacagtt	gctatggaca	aatatgtcaa	aggctttatc	840
aaatatattt	cggaaaaaga	tagtagtaaa	tttactgcc	cagtggtatc	aagtttagag	900
tggagaaaat	aa					912

10 <210> 464
 <211> 324
 <212> ADN
 15 <213> Streptococcus thermophilus

<400> 464

atgaggtatg	aagcattgag	attattatgt	ttttttgatt	taccaatgga	atccaaggat	60
gaaaaaagaa	tatatcgtaa	ttttcgtaaa	gaattaattt	caaatgggtt	tgaaatgtta	120
caattttcgg	tctactatcg	cacttgctct	aatagaagct	ttgcaaataa	attttataag	180
aagttaaaga	ttagcaatct	tcctgctggg	aatgtgagac	ttttggcagt	tactgaaaaa	240
caattttcag	agatgacatt	aattataggt	ggtaaaacta	agcaagaaga	aatcgtcagt	300
gataataagt	tggtggttat	atga				324

ES 2 398 918 T3

<210> 465
 <211> 1053
 <212> ADN
 5 <213> Streptococcus thermophilus

<400> 465

```

atgaaatatt ttgtacaaca tccttacaac gaacgtattg aattaaatat tgggtgcaatc      60
acacaaattg ttggtcagaa taaagaactc aaatattata tttggcaaat tttgagctgg      120
tattttggcg gaaaaaaata ctcaagtgag gacttaagta tttttgatta tgaggaacct      180
actatacttg atgagtctgg agaaatagtg aagcgaagta gctatcacta tatcgacatt      240
tcaagtttta aggatttact ggagcagatg gaatacaaga aaggaacact tgctcagggg      300
taccttagta aaattctcaa tcaggttgat attgtaggcc atttggagaa aattaatgaa      360
caagtagagc ttatagaagg agcaatgaat cagcatataa acttaaactg tggtcagggg      420
gagtaccatt tggagaatca ccctctaaca ctagaccaat tactttcaaa aaatttttagt      480
cccttttttg ctatcgagaa taagaattta tcttttgaat gggtttcaaa tactgataaa      540
ctttctctct ttctagaaat gttagaccgc cttctgtcac aaacaacaga gaagtatctc      600
attgtgctaa aaaatattga tggctttatc tcagaagaat cttatactat tttttatagg      660
caaatctgtc atctgggcaa gaagtatcca aatctaactt ttattttggt tcctagtgcac      720
caaggctatt taaaaattga tgaagaaaat agtaggttcg tcaatatttt atctgaccag      780
gtggagcatt tgtatgatgt tgagtttatg tatgaaagag taatgaaata ttatccaagt      840
aatgattttc cgacgagaga aggttttagg atgtcttttag aaactgtgac accttattta      900
ttgacaaaaa tgctgagaca acctagtctc tcacttgttg attcagtaat attgaatatc      960
ctaaatcagt tgtttcattt tagttaccgt ataagatatt ctcagacacc tgataaggaa     1020
ctattacata aatttttaga aagtaaggat tga                                     1053
    
```

10 <210> 466
 <211> 4123
 <212> ADN
 <213> Streptococcus thermophilus

15 <400> 466

ES 2 398 918 T3

atgagc	gatt	tatata	gtca	aagg	tccaat	tattac	ctgt	ccttat	ctga	acaaga	aatt	60		
atcatta	aaaa	atgata	aataa	agag	attg	tc	aaaga	agtg	ccattt	cact	cgttg	ataat	120	
gtattac	ttt	ttgg	taat	gc	acaact	gacc	accca	actca	tcaa	agcctt	gtcaa	agaac	180	
aagg	tgaat	g	tttact	at	ctcaa	atg	tt	gg	tcaatt	ta	tttct	agtat	240	
aggc	aggac	g	aattc	caaaa	gcaag	ag	ttg	caag	caaag	g	cttatt	ttga	300	
cg	tttag	agg	ttgc	gagg	g	tattg	ctac	g	acca	agg	tga	ggc	accaat	360
agag	ag	ttt	g	at	ac	ggat	g	t	ctac	tagat	ac	ctcag	att	420
gt	caat	gata	ttc	agaa	agc	ttatt	ccatt	ac	agaa	atta	tgg	gttac	ga	480
gc	gaa	atc	ct	at	ttt	tacta	tct	ga	at	ttta	ctc	gttc	cta	540
agg	ag	tag	ac	gg	ct	ggg	ga	gg	att	g	ttt	t	aac	600

ES 2 398 918 T3

ttatattctt gcttaatggg ctgattaaga aaaacgggct aagcttggga tttggggtaa 660
 ttcacaagca tcatcagcat catgcgacct tggccagtga ttaaatggaa gaatggagac 720
 ctatcatcgt cgataatacg cttatggagt tggtagcгаа tggtaaactt cttttaagtc 780
 attttgaaaa taaggatcaa gacttcatac tcacccatga aggagagaa atctttgcac 840
 gggcttttacg ttcaagaata ttagaagtcc atcagtatat tgagttagat aaaaaacgct 900
 attcttttct ttatacagca gataggcaaa tcaagagttt gattagggct ttagagaaac 960
 ttgaccctag tctctatgag acaagttaca caggagggca ttaatgggac tttactttaa 1020
 cctcagcgaa gaagagcgtg agtttgccaa acaaaaaaac catgttttgt ctgattattt 1080
 atgatattcg aagtaacaaa cgtagactta aactctcgaa attacttgag ggttatggcg 1140
 tgagggtgca aaaatcctgt ttcgaagtcg acctgtcaag aatgattat cagtctctcc 1200
 ttaaggatat cgagggcttc tccaaggctg atgaagaaga cagcataata gtgtatgtgc 1260
 caaccaaaga agaggtgact agttttagcc cctaccatag tgctgaaaaa ttagatgaca 1320
 ttctcttccc ctaagccttt atagaccttt aatcatatgg tacactatag atagtgtttc 1380
 cagaggctct taaggaaatc aaagatagag agacacttca aagattttgt agatataatgg 1440
 aagcattagt agcctatttc aagttttatg gaggtaaaga ttaatgacat tcgctaagat 1500
 taaattttca gctcaaattc gtttagagac aggctccat attggtgaa gcgatgcttt 1560
 tgcagccatt ggtgcaatcg attcgctgt tattaagat cctattacca acctaccgat 1620
 cattcctggg tcaagtctca aaggaaaaat gagaacgctt cttgccaagg tttataatga 1680
 aaaggtagct gagaaaccaa gcgatgacag tgatattctt agccgtttat ttgggaatag 1740
 taaagataaa cgattcaaaa tgggacgctt gatttttctg gatgccttct tgtcaaacgc 1800
 tgatgagcta gactctcttg gggtaagaag ttatacagaa gtaaaatttg aaaatacaat 1860
 tgaccgtatc actgccgaag ctaatccaag acaaattgaa cgtgctattc gtaccagtac 1920
 ttttgatttc gagttgattt atgaaattac agatgagaat gaaaatcaag tcgaagaaga 1980
 ttcaaagtg attcgagatg gtttaaaact gcttgaactt gattatcttg gtggttctgg 2040
 atctcgaggt tacggtaagg ttgcttttga aaacctcaaa gctactaccg tatttggtaa 2100
 ttatgatggt aaaacattaa atgaactttt aactgcggag gtctaatatg acctataaac 2160
 tgatattat gacctttcag aatgctcatt ttggttcggg cactcttgat agctcaaaat 2220
 taacattctc agcagaccgt atcttctcag cactagtgtc agaatcccta aaaatgggaa 2280
 aactcgatgc atttcttgcg gaagctaacc aagacaagtt cacgctcaca gatgcctttc 2340
 catttcaatt tggtccttt ttgccgaaac ctattggtta tcccaaacat gaccaaatag 2400
 atcaatcagt tgatgtcaaa gaggttcgcc gtcaagcaaa attgtctaag aaactgcaat 2460
 ttcttgctct agaaaatggt gacgattata tcaatggaga gttatttgaa aatgaagagc 2520

ES 2 398 918 T3

atgcagtcac cgatactgtg acaaaaaatc aaccacataa ggacggcaat ctttatcagg 2580
tagctacaac cagattttca aatgatacgt cgctttacgt catcgcaaac gaatctgatt 2640
tgcttaatga gttgatgtct agtcttcagt attcaggtct tggaggaaag cgttcaagtg 2700
gttttggtcg ttttgagtta gatattcaaa atatcccact agaattgtca gatagactga 2760
ctaagaatca ttcagataaa gtgatgagtc ttacgacagc acttcctgta gatgctgacc 2820
ttgaagaagc aatggaagat ggacattact tattaactaa atcaagtggg tttgcattta 2880
gtcatgccac caatgagaat tatcgtaagc aggatcttta caaatttgct tctgggtcaa 2940
cttttagtaa aacatttgaa ggtcagattg ttgatgtgag accacttgat ttccctcatg 3000
ctgttttaaa ttatgctaaa ccactcttct ttaaattgga ggtataaaaa tgaaaaatga 3060
ctatagaaca tttaaattaa gcctcctgac acttgctcca attcatattg gtaatggaga 3120
gaagtatacc tctagagaat ttatctatga aataaaaaag ttttactttc ctgacatggg 3180
gaaattctat aataaaatgg tggagaagag gcttgctgaa aagtttgaag catttctaata 3240
tcaaactcgt ccaaatgcac gtaataatcg tcttatttcc ttcttaaatg ataaccgaat 3300
tgcagagcgt tcttttgagg gttatagtat ctctgaaaca ggtttagaat cggacaaaaa 3360
tcctgattca accggagcta ttaacgaagt taataaattt attcgagatg cttttggaaa 3420
tcctacatt cctggtagct cactaaaagg tgctattcgt accattttaa tgaatactac 3480
ccctaagtgg aataatgaaa atgctgtaaa tgactttgga agatttccga aagagaataa 3540
gaaccttacc ccttggggac caaaaaaggg aaaagaatac gatgatttgt ttaacgcaat 3600
tcgtgtgagt gatagtaagc cttttgataa taagagtctt atcttagtgc agaaatggga 3660
ttattcagcg aaaacaaata aagctaaacc acttccttg tatagagaat caatctctcc 3720
attaacaaaa attgaatttg agattacaac aaccactgat gaagctggaa gattgattga 3780
agaattaggt aagagagcac aagcgtttta taaagactat aaggcatttt tcctatctga 3840
atctcctgat gataagattc aagccaatct acaataccca atttatttag gtgcggggag 3900
cggtgcttgg acaaagactc tatttaagca agctgatggg atttacaaa gacgatacag 3960
tcgaatgaaa actaaaatgg ttaaaaaagg agttcttaag ctacaaaaag cacctcttaa 4020
aacagttaag attccatctg gtaatcattc attagtcaag aaccacgagt ctttttatga 4080
aatgggaaaa gctaatttca tgattaagga gattgataaa tga 4123

<210> 467

<211> 624

5 <212> ADN

<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 467

atgagc60
 atcatta120
 gtattact180
 aaggt240
 agg300
 cgt360
 ag420
 gt480
 gc540
 ag600
 5 tt624

<210> 468

<211> 396

<212> ADN

10 <213> Streptococcus thermophilus

<400> 468

tt60
 cat120
 gtc180
 aat240
 cgt300
 ctt360
 15 ag396

<210> 469

<211> 273

<212> ADN

20 <213> Streptococcus thermophilus

<400> 469

ES 2 398 918 T3

atgttttgtc tgattattta tgatattcga agtaacaaac gtagacttaa actctcgaaa 60
 ttacttgagg gttatggcgt gagggtgcaa aaatcctggt tcgaagtcga cctgtcaaga 120
 aatgattatc agtctctcct taaggatatc gagggcttct ccaaggctga tgaagaagac 180
 agcataatag tgtatgtgcc aaccaaaaga gaggtgacta gttttagccc ctaccatagt 240
 gctgaaaaat tagatgacat tctcttcccc taa 273

<210> 470

<211> 663

5 <212> ADN

<213> Streptococcus thermophilus

<400> 470

atgacattcg ctaagattaa attttcagct caaattcggt tagagacagg cctccatatt 60
 ggtggaagcg atgcttttgc agccattggt gcaatcgatt cgcttggtat taaagatcct 120
 attaccaacc taccgatcat tcctgggtca agtctcaaag gaaaaatgag aacgcttctt 180
 gccaaaggtt ataatgaaaa ggtagctgag aaaccaagcg atgacagtga tattcttagc 240
 cgtttatttg ggaatagtaa agataaacga ttcaaaatgg gacgcttgat ttttcgtgat 300
 gccttcttgt caaacgctga tgagctagac tctcttgggg taagaagtta tacagaagta 360
 aaatttgaaa atacaattga ccgtatcact gccgaagcta atccaagaca aattgaacgt 420
 gctattcgta ccagtacttt tgatttcgag ttgatttatg aaattacaga tgagaatgaa 480
 aatcaagtcg aagaagattc caaagtgatt cgagatgggt taaaactgct tgaactgat 540
 tatcttggtg gttctggatc tcgaggttac ggtaagggtg cttttgaaaa cctcaaagct 600
 actaccgtat ttggtaatta tgatgttaaa acattaaatg aacttttaac tgcggaggtc 660
 10 taa 663

<210> 471

<211> 900

<212> ADN

15 <213> Streptococcus thermophilus

<400> 471

ES 2 398 918 T3

atgacctata aactgtatat tatgaccttt cagaatgctc attttggttc gggcactcct	60
gatagctcaa aattaacatt ctccagcagac cgtatcttct cagcactagt gctagaatcc	120
ctaaaaatgg gaaaactcga tgcatttctt gcggaagcta accaagacaa gttcacgctc	180
acagatgcct ttccatttca atttggtccc tttttgccga aacctattgg ttatcccaaa	240
catgaccaa tagatcaatc agttgatgtc aaagaggttc gccgtcaagc aaaattgtct	300
aagaaactgc aatttcttgc tctagaaaat gttgacgatt atatcaatgg agagttattt	360
gaaaatgaag agcatgcagt catcgatact gtgacaaaaa atcaaccaca taaggacggc	420
aatctttatc aggtagctac aaccagatth tcaaatgata cgtcgcttta cgtcacgca	480
aacgaatctg atttgcttaa tgagttgatg tctagtcttc agtattcagg tcttggtgga	540
aagcgttcaa gtggttttgg tcgttttgag ttagatattc aaaatatccc actagaattg	600
tcagatagac tgactaagaa tcattcagat aaagtgatga gtcttacgac agcacttct	660
gtagatgctg accttgaaga agcaatggaa gatggacatt acttattaac taaatcaagt	720
ggttttgcat ttagtcatgc caccaatgag aattatcgta agcaggatct ttacaaatth	780
gcttctggtt caacttttag taaaacatth gaaggtcaga ttgttgatgt gagaccactt	840
gatttcctc atgctgtttt aaattatgct aaaccactct tctttaaatt ggaggtataa	900

<210> 472

<211> 1074

5 <212> ADN

<213> Streptococcus thermophilus

<400> 472

ES 2 398 918 T3

atgaaaaatg actatagaac atttaaatta agcctcctga cacttgctcc aattcatatt 60
 ggtaatggag agaagtatac ctctagagaa tttatctatg aaaataaaaa gttttacttt 120
 cctgacatgg ggaaattcta taataaaatg gtggagaaga ggcttgctga aaagtttgaa 180
 gcatttctaa ttcaaactcg tccaaatgca cgtaataatc gtcttatttc cttcttaaat 240
 gataaccgaa ttgcagagcg ttcttttgga ggttatagta tctctgaaac aggtttagaa 300
 tcggacaaaa atcctgattc aaccggagct attaacgaag ttaataaatt tattcgagat 360
 gcttttgga atccctacat tcttggtagc tcaactaaaag gtgctattcg taccatttta 420
 atgaacta cccctaagtg gaataatgaa aatgctgtaa atgactttgg aagatttccg 480
 aaagagaata agaaccttat cccttgggga ccaaaaaagg gaaaagaata cgatgatttg 540
 tttaacgcaa ttcgtgtgag tgatagtaag ccttttgata ataagagtct tatcttagtg 600
 cagaaatggg attattcagc gaaaacaaat aaagctaac cacttccctt gtatagagaa 660
 tcaatctctc cattaacaaa aattgaattt gagattacaa caaccactga tgaagctgga 720
 agattgattg aagaattagg taagagagca caagcgtttt ataaagacta taaggcattt 780
 ttctatctg aatttcctga tgataagatt caagccaatc tacaataccc aatttattta 840
 ggtgcgggga gcggtgcttg gacaaagact ctatttaagc aagctgatgg tattttacaa 900
 agacgataca gtcgaatgaa aactaaaatg gttaaaaaag gagttcttaa gctcacaaaa 960
 gcacctctta aacagttaa gattccatct ggtaatcatt cattagtcaa gaaccacgag 1020
 tccttttatg aaatgggaaa agctaatttc atgattaagg agattgataa atga 1074

<210> 473

<211> 5832

5 <212> ADN

<213> Streptococcus thermophilus

<400> 473

atgagtgact tagttttagg acttgatatc ggtatagggt ctggttgggtg aggtatcctt 60
 aacaaagtga caggagaaat tatccataaa aactcacgca tcttcccagc agctcaagca 120
 gaaaataacc tagtacgtag aacgaatcgt caaggaagac gcttgacacg acgtaaaaaa 180
 catcgtatag ttcgtttaaa tcgtctattt gaggaaagtg gattaatcac cgattttacg 240
 aagatttcaa ttaatcttaa cccatatcaa ttacgagtta agggcttgac cgatgaattg 300
 tctaatgaag aactgtttat cgctcttaaa aatatgggtga aacaccgtgg gattagttac 360
 ctcgatgatg ctagtgatga cggaaattca tcagtaggag actatgcaca aattgttaag 420
 10 gaaaatagta aacaattaga aactaagaca ccgggacaga tacagttgga acgctaccaa 480

ES 2 398 918 T3

acatatggtc aattacgtgg tgatcttact gttgagaaag atggcaaaaa acatcgcttg 540
 attaatgtct ttccaacatc agcttatcgt tcagaagcct taaggatact gcaaactcaa 600
 caagaattta attcacagat tacagatgaa tttattaatc gttatctcga aattttaact 660
 ggaaaacgga aatattatca tggaccocgga aatgaaaagt cacggactga ttatggtcgt 720
 tacagaacga atggagaaac tttagacaat atttttggaa ttctaattgg gaaatgtaca 780
 ttttatccag acgagtttag agcagcaaaa gcttctaca cggctcaaga attcaatttg 840
 ctaaagatt tgaacaatct aacagttcct actgaaacca aaaagttgag caaagaacag 900
 aagaatcaaa tcattaatta tgtcaaaaat gaaaaggtaa tggggccagc gaaacttttt 960
 aatatatcg ctaaattact ttcttgtgat gttgcagata tcaagggaca ccgatcgcac 1020
 aatcaggta aggctgagat tcatactttc gaagcctatc gaaaaatgaa aacgcttgaa 1080
 acctagata ttgagcaaat ggatagagaa acgcttgata aattagccta tgtcttaaca 1140
 ttaaacactg agaggggaagg tattcaagaa gctttagaac atgaatttgc tgatggtagc 1200
 tttagccaga agcaagttga cgaattgggt caattccgca aagcaaatag ttccattttt 1260
 ggaaaaggat ggcataaatt ttctgtcaaa ctgatgatgg agttaattcc agaattgtat 1320
 gagacgtcag aagagcaaat gactatcctg acacgacttg gaaaacaaaa aacaacttcg 1380
 tcttcaaata aaacaaaata tatagatgag aaactattaa ctgaagaaat ctataatcct 1440
 gttgttgcta agtctgttcg ccaggctata aaaatcgtaa atgcggcgat taaagaatac 1500
 ggagactttg acaatattgt catcgaaatg gctcgtgaaa caaatgaaga tgatgaaaag 1560
 aaagctattc aaaagattca aaaagccaac aaagatgaaa aagatgcagc aatgcttaag 1620
 gctgctaacc aatataatgg aaaggctgaa ttaccacata gtgttttcca cggtcataag 1680
 caattagcga ctaaaatccg cctttggcat cagcaaggag aacgttgcct ttatactggt 1740
 aagacaatct caatccatga tttgataaat aatcctaatc agtttgaagt agatcatatt 1800
 ttacctctt ctatcacatt cgatgatagc cttgcaaata aggttttggg ttatgcaact 1860
 gctaaccaag aaaaaggaca acgaacacct tatcaggctt tagatagtat ggatgatgcg 1920
 tggctcttcc gtgaattaaa agcttttgta cgtgagtcaa aaacactttc aaacaagaaa 1980
 aaagaatacc tccttacaga agaagatatt tcaaagtttg atgttcgaaa gaaatttatt 2040
 gaacgaaatc ttgtagatac aagatacgtc tcaagagttg tcctcaatgc cttcaagaa 2100
 cactttagag ctcaacaagat tgatacaaaa gtttccgtgg ttcgtggcca atttacatct 2160
 caattgagac gccattgggg aattgagaag actcgtgata cttatcatca ccatgctgtc 2220
 gatgcattga ttattgccgc ctcaagtcag ttgaatttgt ggaaaaaaca aaagaatacc 2280
 cttgtaagtt attcagaaga acaactcctt gatattgaaa caggtgaact tattagtgat 2340
 gatgagtaca agaatctgt gttcaaagcc ccttatcaac attttgttga tacattgaag 2400

ES 2 398 918 T3

agtaaagaat ttgaagacag tatcttattc tcatatcaag tggattctaa gtttaatcgt 2460
 aaaatatcag atgccactat ttatgcgaca agacaggcta aagtgggaaa agataagaag 2520
 gatgaaactt atgtcttagg gaaaatcaaa gatatctata ctcaggatgg ttatgatgcc 2580
 tttatgaaga tttataagaa ggataagtca aaattcctca tgtatcgtca cgacccacaa 2640
 acctttgaga aagttatcga gccaatTTTA gagaactatc ctaataagga aatgaatgaa 2700
 aaagggaaaag aagtaccatg taatcctttc ctaaaatata aagaagaaca tggctatatt 2760
 cgtaaataata gtaaaaaagg caatggtcct gaaatcaaga gtcttaaata ctatgatagt 2820
 aagcttttag gtaatcctat tgatattact ccagagaata gtaaaaaataa agttgtctta 2880
 cagtcattaa aaccttggag aacagatgtc tatttcaata aaaatactgg taaatatgaa 2940
 attttaggac tgaaatatgc tgatttaciaa tttgaaaaga agacaggaac atataagatt 3000
 tcccaggaaa aatacaatgg cattatgaaa gaagaggggtg tagattctga ttcagaattc 3060
 aagtttacac tttataaaaa tgatttgTTA ctcgTTaaag atacagaaac aaaagaacaa 3120
 cagcttttcc gttttctttc tcgaactatg cctaattgtga aatattatgt agagTTaaag 3180
 ccttattcaa aagataaatt tgagaagaat gagtcaactta ttgaaatTTT aggttctgca 3240
 gataagtcag gacgatgtat aaaagggcta ggaaaatcaa atatttctat ttataaggta 3300
 agaacagatg tcctaggaaa tcagcatatc atcaaaaatg agggtgataa gcctaagcta 3360
 gatttttaat attaattgTT aaaaaagtgt tgcaattata gttatcatat gctataataa 3420
 tcgtgtaagg gacgccttac acagttactt aaatcttgca gaagctaciaa agataaggct 3480
 tcatgccgaa atcaacaccc tgtcattTTA tggcaggggtg ttttcgTTat ttaaagagga 3540
 gaagaaatga cttggagagt tgtacatgtc agtcaaagtg agaagatgcg cttaaagctt 3600
 gataacttat tagtgcaaaa gatgggaciaa gagtttacgg tgccactaag tgatatttCG 3660
 ataatcgTTg cagaaggTgg ggatacagtt gttacccttc gtctattaag tgcttaagt 3720
 aaatataata ttgccttggT cgTTtTgtgat aacgaacatt taccaacagg aatttatcac 3780
 tcacaaaatg ggcactTTtag agcgtaciaaag cgcttgaaag aacagctgga ttggTctcag 3840
 aaacaaaagg aaaaggcatg gcagattgta acttattata aaatcaataa ccaagaggat 3900
 gtcttagcca tgTTtgaaaa aagtctggac aacattagat tactttcaga ctataaagag 3960
 cagatagaac ctggTgatag aacgaataga gagggacatg ctgccaaggt ctactttaat 4020
 gagctctTTg gtaaacaatt tgtcagagta actcagcaag aagctgatgt catcaatgct 4080
 ggtttaaact atggctatgc tatcatgagg gctcagatgg ctagaatagt ggcgggTTat 4140
 ggtttaaatg gcctattagg aatcttccat aaaaatgaat acaatcagtt taatttggTT 4200
 gacgattTga tggagccatt tagacagatt gtagatgTTT gggTatatga taatctacga 4260
 gatcaggaat tccttaagta tgagtatagg ttgggattga cagatttact caatgctaaa 4320

ES 2 398 918 T3

atcaaataatg gcaaagagac ttgctcagtg acagttgcta tggacaaata tgtcaaaggc 4380
 tttatcaaat atatttcgga aaaagatagt agtaaatttc actgcccagt ggtatcaagt 4440
 ttagagtgga gaaaataaga tgaggtatga agcattgaga ttattatggt tttttgattt 4500
 accaatggaa tccaaggatg aaaaaagaat atatcgtaat tttcgtaaag aattaatttc 4560
 aatggggtt gaaatgttac aattttcggt ctactatcgc acttgtccta atagaagctt 4620
 tgcaaataaa ttttataaga agttaaagat gagcaatctt cctgctggga atgtgagact 4680
 tttggcagtt actgaaaaac aattttcaga gatgacatta attataggtg gtaaaaactaa 4740
 gcaagaagaa atcgtcagtg ataataagtt ggtgatcata tgaaattttt tgtacaacat 4800
 ccttaciaag aacgtattga attaaatatt ggtgcaatca cacaaattgt tggtcagaat 4860
 aatgaactca aatattatac ttggcagatt ttgagctggt attttgggtg aaaaaaatac 4920
 tcaagtgagg acttaagtat ttttgattat gaggagccta ccatacttga tgaggccaga 4980
 gaaatagtga aacgaagtag ctatcactat atcgacattt caagttttaa ggatttactg 5040
 gagcagatgg aatacaagaa aggaacactt gctcagggtt accttcgtaa aattgtcaat 5100
 caagttgata ttgtaggcca tttggagaaa attaataaac aagtagagct tattgaagaa 5160
 gctatgaatc ggcatataaa cttaaactgt ggacaggtag aataccattt ggagaatctc 5220
 cctctaacac tagaccaact actcacaaaa aatttttagcc cattttttgc cattgagaac 5280
 aagaatctat cttttgaatg ggtttctaata attgataaac tatccctctt tttagaaatg 5340
 ttagaccatc ttctttcaca aacaacagag aagtatctca ttgtgctaaa aaatattgat 5400
 ggctttatct cagaagaatc ttatactatt ttttatagtc aaatctgtca tctggtaag 5460
 aagtatccaa atctaactt tattttggtt cctagtgacc aaggctattt aaaaattgat 5520
 gaagaaaata gtaggttcgt caatatttta tctgaccagg tggacattt gtatgatggt 5580
 gagtttatgt atgaaagggt aatgaaatat tatccaagta atgattttcc gacgagagaa 5640
 ggtttttagga tgtctttaga aactgtgaca ccttatttat tgacaaaaat gctgagacaa 5700
 cctagtctct cacttgttga ttcagtaata ttgaatatcc taaatcagct gtttcatttt 5760
 agttaccgta taagatgttc tcagacacct gataaggaac tattacagaa attttttagaa 5820
 agtaaggatt ga 5832

<210> 474

<211> 3369

5 <212> ADN

<213> Streptococcus thermophilus

<400> 474

ES 2 398 918 T3

```
atgagtgact tagttttagg acttgatata ggtatagggt ctggtgggtg aggtatcctt    60
aacaagtga caggagaaat tatccataaa aactcacgca tcttcccagc agctcaagca    120
gaaaataacc tagtacgtag aacgaatcgt caaggaagac gcttgacacg acgtaaaaaa    180
```

ES 2 398 918 T3

catcgtatag ttcgtttaaa tcgtctattt gaggaagtg gattaatcac cgatttttacg 240
aagatttcaa ttaatcttaa cccatatcaa ttacgagtta agggcttgac cgatgaattg 300
tctaataaag aactgtttat cgctcttaaa aatatggtga aacaccgtgg gattagttac 360
ctcgatgatg ctagtgatga cggaaattca tcagtaggag actatgcaca aattgttaag 420
gaaaatagta aacaattaga aactaagaca ccgggacaga tacagttgga acgctaccaa 480
acatatggtc aattacgtgg tgattttact gttgagaaag atggcaaaaa acatcgcttg 540
attaatgtct ttccaacatc agcttatcgt tcagaagcct taaggatact gcaaactcaa 600
caagaattta attcacagat tacagatgaa tttattaatc gttatctcga aattttaact 660
ggaaaacgga aatattatca tggaccggga aatgaaaagt cacggactga ttatggctgt 720
tacagaacga atggagaaaac tttagacaat atttttggaa ttctaattgg gaaatgtaca 780
ttttatccag acgagtttag agcagcaaaa gcttctaca cggctcaaga attcaatttg 840
ctaaatgatt tgaacaatct aacagttcct actgaaacca aaaagttgag caaagaacag 900
aagaatcaaa tcattaatta tgtcaaaaat gaaaaggtaa tggggccagc gaaacttttt 960
aaatatacgc ctaaattact ttcttgtgat gttgcagata tcaagggaca ccgtatcgac 1020
aatcaggta aggctgagat tcatacttcc gaagcctatc gaaaaatgaa aacgcttgaa 1080
accttagata ttgagcaaat ggatagagaa acgcttgata aattagccta tgtcttaaca 1140
ttaaacactg agaggggaagg tattcaagaa gctttagaac atgaatttgc tgatggtagc 1200
tttagccaga agcaagttga cgaattgggt caattccgca aagcaaatag ttccattttt 1260
ggaaaaggat ggcataattt ttctgtcaaa ctgatgatgg agttaattcc agaattgtat 1320
gagacgtcag aagagcaaat gactatcctg acacgacttg gaaaacaaaa aacaacttcg 1380
tcttcaata aaacaaaata tatagatgag aaactattaa ctgaagaaat ctataatcct 1440
gttgttgcta agtctgttcg ccaggctata aaaatcgtaa atgcgcgat taaagaatac 1500
ggagactttg acaatattgt catcgaaatg gctcgtgaaa caaatgaaga tgatgaaaag 1560
aaagctattc aaaagattca aaaagccaac aaagatgaaa aagatgcagc aatgcttaag 1620
gctgctaacc aatataatgg aaaggctgaa ttaccacata gtgttttcca cggtcataag 1680
caattagcga ctaaaatccg cctttggcat cagcaaggag aacgttgcct ttatactggt 1740
aagacaatct caatccatga tttgataaat aatcctaate agtttgaagt agatcatatt 1800
ttacctttt ctatcacatt cgatgatagc cttgcaaaata aggttttggg ttatgcaact 1860
gctaaccaag aaaaaggaca acgaacacct tatcaggctt tagatagtat ggatgatgag 1920
tggtctttcc gtgaattaaa agcttttgta cgtgagtcaa aaacactttc aaacaagaaa 1980
aaagaatacc tccttacaga agaagatatt tcaaagtttg atgttcgaaa gaaatttatt 2040
gaacgaaatc ttgtagatac aagatacgtt tcaagagttg tctcaatgc cttcaagaa 2100

ES 2 398 918 T3

cacttttagag ctcaagaat tgatacaaaa gtttccgtgg ttcgtggcca atttacatct 2160
 caattgagac gccattgggg aattgagaag actcgtgata cttatcatca ccatgctgtc 2220
 gatgcattga ttattgccgc ctcaagtcag ttgaatttgt ggaaaaaaca aaagaatacc 2280
 cttgtaagtt attcagaaga acaactcctt gatattgaaa caggtgaact tattagtgat 2340
 gatgagtaca aggaatctgt gttcaaagcc ccttatcaac attttgttga tacattgaag 2400
 agtaaagaat ttgaagacag tatcttattc tcatatcaag tggattctaa gtttaatcgt 2460
 aaaatatcag atgccactat ttatgcgaca agacaggcta aagtgggaaa agataagaag 2520
 gatgaaaactt atgtcttagg gaaaatcaaa gatatctata ctccaggatgg ttatgatgcc 2580
 tttatgaaga tttataagaa ggataagtca aaattcctca tgtatcgtca cgaccacaaa 2640
 acctttgaga aagttatcga gccaatttta gagaactatc ctaataagga aatgaatgaa 2700
 aaagggaaag aagtaccatg taatcctttc ctaaaatata aagaagaaca tggctatatt 2760
 cgtaaatata gtaaaaaagg caatggctct gaaatcaaga gtcttaata ctatgatagt 2820
 aagcttttag gtaatcctat tgatattact ccagagaata gtaaaaaata agttgtctta 2880
 cagtcattaa aaccttggag aacagatgtc tatttcaata aaaatactgg taaatatgaa 2940
 attttaggac tgaaatatgc tgatttaciaa tttgaaaaga agacaggaac atataagatt 3000
 tcccaggaaa aatacaatgg cattatgaaa gaagagggtg tagattctga ttcagaattc 3060
 aagtttacac tttataaaaa tgatttgta ctcgttaaag atacagaaac aaaagaacaa 3120
 cagcttttcc gttttctttc tcgaactatg cctaattgta aatattatgt agagttaaag 3180
 ccttattcaa aagataaatt tgagaagaat gagtactta ttgaaatfff aggttctgca 3240
 gataagtcag gacgatgtat aaaagggcta ggaaaatcaa atatttctat ttataaggta 3300
 agaacagatg tcctaggaaa tcagcatatc atcaaaaatg agggtgataa gcctaagcta 3360
 gatttttaa 3369

<210> 475

<211> 912

5 <212> ADN

<213> Streptococcus thermophilus

<400> 475

ES 2 398 918 T3

atgacttggg gagttgtaca tgtcagtcaa agtgagaaga tgcgcttaaa gcttgataac 60
 ttattagtgc aaaagatggg acaagagttt acggtgccac taagtgatat ttcgataatc 120
 gttgcagaag gtggggatac agttgttacc ctctcgtctat taagtgcctt aagtaaatat 180
 aatattgcct tggtcgtttg tgataacgaa catttaccaa caggaattta tcactcacia 240
 aatgggcact ttagagcgta caagcgcttg aaagaacagc tggattggtc tcagaaacia 300
 aaggaaaagg catggcagat tgtaacttat tataaaatca ataaccaaga ggatgtccta 360
 gccatgtttg aaaaaagtct ggacaacatt agattacttt cagactataa agagcagata 420
 gaacctggtg atagaacgaa tagagagggg catgctgcca aggtctactt taatgagctc 480
 tttggtaaac aatttgtcag agtaactcag caagaagctg atgtcatcaa tgctggttta 540
 aactatggct atgctatcat gagggtcag atggctagaa tagtggcggg ttatggttta 600
 aatggcctat taggaatctt ccataaaaat gaatacaatc agtttaattt ggttgacgat 660
 ttgatggagc catttagaca gattgtagat gtttgggtat atgataatct acgagatcag 720
 gaattcctta agtatgagta taggttggga ttgacagatt tactcaatgc taaaatcaaa 780
 tatggcaaag agacttgctc agtgacagtt gctatggaca aatatgtcaa aggctttatc 840
 aaatatattt cggaaaaaga tagtagtaaa tttcactgcc cagtggatc aagtttagag 900
 tggagaaaat aa 912

5 <210> 476
 <211> 324
 <212> ADN
 <213> Streptococcus thermophilus

10 <400> 476
 atgaggtatg aagcattgag attattatgt ttttttgatt taccaatgga atccaaggat 60
 gaaaaaagaa tatatcgtaa ttttcgtaa gaattaattt caaatgggtt tgaaatgtta 120
 caattttcgg tctactatcg cacttgtcct aatagaagct ttgcaaataa attttataag 180
 aagttaaaga tgagcaatct tcctgctggg aatgtgagac ttttggcagt tactgaaaaa 240
 caattttcag agatgacatt aattataggt ggtaaaacta agcaagaaga aatcgtcagt 300
 gataataagt tggatgatcat atga 324

15 <210> 477
 <211> 1053
 <212> ADN
 <213> Streptococcus thermophilus

20 <400> 477

ES 2 398 918 T3

atgaaat ttt ttgtacaaca tctttacaaa gaacgtattg aattaaatat tgggtgcaatc 60
 acacaaattg ttggtcagaa taatgaactc aatattata cttggcagat tttgagctgg 120
 tattttgggtg gaaaaaaata ctcaagtgag gacttaagta tttttgatta tgaggagcct 180
 accatacttg atgaggccag agaaatagtg aaacgaagta gctatcacta tatcgacatt 240
 tcaagtttta aggatttact ggagcagatg gaatacaaga aaggaacact tgctcagggt 300
 taccttcgta aaattgtcaa tcaagttgat attgtaggcc atttggagaa aattaatgaa 360
 caagtagagc ttattgaaga agctatgaat cggcatataa acttaaactg tggacaggta 420
 gaataccatt tggagaatct ccctctaaca ctagaccaac tactcacaaa aaat ttttagc 480
 ccattttttg ccattgagaa caagaatcta tcttttgaat gggtttctaa tattgataaa 540
 ctatccctct ttttagaaat gttagaccat cttctttcac aaacaacaga gaagtatctc 600
 attgtgctaa aaaatattga tggctttatc tcagaagaat cttatactat tttttatagg 660
 caaatctgtc atctgggtcaa gaagtatcca aatctaacct ttattttggt tccatagtgac 720
 caaggctatt taaaaattga tgaagaaaat agtaggttcg tcaatatttt atctgaccag 780
 gtggaacatt tgtatgatgt tgagtttatg tatgaaaggg taatgaaata ttatccaagt 840
 aatgattttc cgacgagaga aggttttagg atgtctttag aaactgtgac accttattta 900
 ttgacaaaaa tgctgagaca acctagtctc tcacttgttg attcagtaat attgaatatc 960
 ctaaatacagc tgtttcattt tagttaccgt ataagatggt ctcagacacc tgataaggaa 1020
 ctattacaga aat ttttaga aagtaaggat tga 1053

- 5 <210> 478
- <211> 7900
- <212> ADN
- <213> Streptococcus thermophilus

- 10 <400> 478

ES 2 398 918 T3

atgagcgatt	tatatagtca	aaggccaat	tattacctgt	ccttatctga	acaagaatt	60
atcattaata	atgataataa	agagattgtc	aaagaagtgt	ccatttcact	cgttgataat	120
gtattacttt	ttggtaatgc	acaactgacc	acccaactca	tcaaagcctt	gtcaaagaac	180
aaggatgaatg	tttactat	ctcaaagtgt	ggccaattta	tttctagtat	tgaacccac	240
aggcaggacg	aattccaaaa	gcaagagttg	caagcaaagg	cctattttga	agaggatttc	300
cgtttagagg	ttgcgaggag	tattgctacg	accaaggtga	ggcacccaat	tgccttactt	360
agagagtttg	atacggatgg	tctactagat	acctcagatt	attctagggt	tgaagatagt	420
gtcaatgata	ttcagaaagc	ttattccatt	acagaaatta	tgggttacga	aggctgcctt	480
gcgaaatcct	atTTTTacta	tctgaattta	ctcgttctca	atgactttca	ttttaatggt	540
aggagtagac	ggcctgggga	ggattgtttt	aacagtgcc	tcaatTTTgg	ctatagtatc	600
ttatattcct	gcttaatggg	ctgattaaga	aaaacgggct	aagcttggga	tttggggtaa	660
ttcacaagca	tcatcagcat	catgcgacct	tggccagtga	tttaatggaa	gaatggagac	720
ctatcatcgt	cgataatacg	cctatggagt	tggtagaaa	tggtaaactt	cttttaagtc	780
atTTTgaaaa	taaggatcaa	gacttcatac	tcacccatga	aggcagagaa	atctttgcac	840
gggctttacg	ttcaagaata	ttagaagtcc	atcagtatat	tgagttagat	aaaaaacgct	900
attcttttct	ttatacagca	gataggcaaa	tcaagagttt	gattagggct	tttagagaac	960
ttgaccctag	tctctatgag	acaagttaca	caggagggca	ttaatgggac	tttactttaa	1020
cctcagcgaa	gaagagcgtg	agtttgccaa	acaaaaaac	atgTTTTgtc	tgattattta	1080
tgatattcga	agtaacaaac	gtagacttaa	actctcgaaa	ttacttgagg	gttatggcgt	1140
gagggtgcaa	aatcctggt	tcgaagtcaa	cctgtcaaga	aatgattatc	agtctctcct	1200

ES 2 398 918 T3

taaggatatac gagggcttct acaaggctga tgaagaagac agcataatag tgtatgtgac 1260
 aaccaaagaa gaggtgacta gttttagccc ctaccatagt gctgaaaaat tagatgacat 1320
 tctcttcttc taagccttta tagaccttta atcatatggt aactataga tagtgtttcc 1380
 agtaggtcct acatcttggt cctctagcaa ctgcctagag cacaagatat ggggatataa 1440
 acctaattac ctcgagaggg gacggaaacg ctttctagct cgctataatt acccattcct 1500
 agaaagatat aaacctaatt acctcgagag gggacggaaa ctttgaatag tctttgaatc 1560
 gcatttgaac catatagata taaacctaat tacctcgaga ggggacggaa acaggttttt 1620
 tgccatagat tttccaagac cttcccaact gatataaacc taattacctc gagaggggac 1680
 ggaaacgctt tctagctcgc tataattacc cattcctaga aagatataaa cctaattacc 1740
 tcgagagggg acttttttga aaattttgaa aacagtattg ataccgcttc cagaaagtgt 1800
 tagactaaaa gcacattaag ggcgccccaa tgagttgaaa agtactttca gcttttgggg 1860
 ttttttcata caaagatgaa ggagtcgaat gaaaaaatta gtatttactt ttaaaaggat 1920
 cgaccatcct gcacaagatt tggctgtaa atttcatggc ttcttgatgg agcagttgga 1980
 tagtgactat gttgattatc tgcacagca gcaaacaat ccctatgcca ccaaggtaat 2040
 ccaagggaaa gaaaacacgc agtgggttgt acatctgctc acagacgaca tcgaggataa 2100
 ggtttttatg accttattac agattaaaga ggtgtcctta aacgatctgc ctaaactcag 2160
 tgtcgaaaaa gttgagattc aggagttggg ggcagataaa ctgtagaga ttttcaatag 2220
 tgaggaaaaat caaacctatt tttcaattat ttttgagact ccaacagggt ttaaatctca 2280
 aggttcctac gtcaccttcc cgtctatgcg tttgattttt caaagtttga tgcaaaagta 2340
 tggaaggttg gttgaaaatc aacctgaaat tgaagaggat accttagatt acctatctga 2400
 acacagcact atcacgaatt atcgcttggg gacgagttat ttcagggtgc acaggcaacg 2460
 aattcctgcc tttagaggaa agttaacctt taaagtacaa ggcgccccaa ctctaaaagc 2520
 ttatgtcaaa atgcttctaa cattcgggtga atattcagggt cttggcatga aaacgagtct 2580
 cggtatggga gggataaagc ttgaagaaaag aaaagattga tttattttac ggagctcttt 2640
 tgcattgatat cggttaagggtc attcaaaggg cgacaggaga acgaaaaaaaa cacgccttgg 2700
 taggcgcgga ttggtttgat gagattgctg ataatacaagt tatttccgat caaattagat 2760
 atcacatggc taactaccag agtgataaac ttggaaatga ccatcttgct tacataactt 2820
 atatcgctga taacattgcc tctggtgtcg acagaagaca gtcaaatgag gagagtgacg 2880
 aggatacatc agctaagatt tgggatacct atacaaacca ggctgatatt tttaacgttt 2940
 ttggggcaca aacggataaa cgctacttta aaccgacggt tctaaacttg aaatctaaac 3000
 ctaactttgc gtcggcaaca tatgaacctt tctcaaaagg tgattatgcg gcaattgcca 3060
 ctcgatcaa aatgaattg gcagaatttg agtttaatca agtacaatt gactctttgt 3120

ES 2 398 918 T3

taaatctggt cgaagcaacc ctctcttttg tgccttcttc gactaatact aaagaaatcg 3180
 ctgatatttc acttgctgat catagtcgtc tgacagcagc ttttgctcta gccatctatg 3240
 attacttgga agacaaaggt cgtcataact ataaggagga cttgtttact aaagcatcag 3300
 ccttttatga ggaagaagct tttctcctag ctagctttga cttatcaggg attcaagact 3360
 ttatctataa tattaatatt gcgacgaatg gtgctgctaa acaattgaag gctagatctt 3420
 tataatctga ctttatgagc gagtatatag cagacagttt acttgataaa ctaggcctca 3480
 atcgggctaa tatgctctat gtcggtgggg gacatgctta ctttgtccta gccaatactg 3540
 aaaaaacggt agaaacactc gttcaatttg aaaaagattt caatcaattt ttattggcaa 3600
 atttccaaac cagattatat gttgcctttg gttggggaag ctttgcggct aaggatatca 3660
 tgagcgaact gaactcacct gaaagctata gacaggtcta tcaaaaggct agtcgcatga 3720
 tttctgagaa aaaaatctca aggtatgatt atcaaaccct tatgttggtg aacaggggcg 3780
 gtaaactctc tgaaagagag tgcgagattt gtcattccgt tgagaattta gttgcttctc 3840
 atgaccaaaa agtgtgtgac atttgtcgag gcttgtatca attttctaaa gagattgccc 3900
 atgaccattt cattatcact gaaaatgaag ggcttcctat tgggccgaac gcatgtctta 3960
 aggggtgtgc atttgaaaag ctgagccaag aagctttttc ccgtgtctat gtcaaaaatg 4020
 actataaggc tggtagcgtt aaggcaacc atgtttttgt tggagattac cagtatgatg 4080
 aaatatacaa ttatgctgcc ttatctaaaa acgaaaatgg gtttaggtatt aaacgttttag 4140
 ctgttgtagc tcttgacgtg gatgatttgg gagcagcctt tatggctggc ttctcccaac 4200
 aaggaaatgg gcaatatagt actctatcac gctcagccac tttctctcga agcatgagtc 4260
 ttttcttcaa ggtttatatt aaccagtttg ctagtgataa gaagctctct atcatctatg 4320
 ccggtgggga tgatgttttt gctattggct cttggcaaga tattattgcc tttactgttg 4380
 aacttcgtga gaacttcatt aaatggacaa atggaaaact aacactatca gctggtatcg 4440
 gtctgtttgc tgataagacc cctattagct taatggcaca tcaaacaggg gagctagaag 4500
 aaacagctaa aggcaatgag aaagatagta tttcactctt tagttccgac tataccttta 4560
 aatttgatcg gtttatcact aatgtttacg acgataagtt agagcagatt cgctatttct 4620
 ttaatcacca agatgaacga ggcaagaatt tcatttataa attgattgaa ttgcttcgaa 4680
 attatgatcg tatgaatatg gcacgttttag cttattattt aacacgactt gaagaattga 4740
 cgcgtgaaac agacagggat aaatttaaaa catttaaaaa tttattctat tcttgggtaca 4800
 caaataagga tgataaggat agaaaagaag cagagtttagc cttgcttctc tataatctatg 4860
 agattagaaa ggattaggat atgacaatct tgactgatga gaattacggt gatattgcag 4920
 aaaaagcaat tctaaaacta gaaagaata ctaggaacag aaagaatcct gatgccttct 4980
 ttcttacaac aagtaagctc agaaacttgc tgagcttaac tagtacactt tttgatgaga 5040

ES 2 398 918 T3

gtaagggtcaa agaatatgat gctctccttg atcgtattgc ttatttaaga gtacaatttg 5100
 tctaccaagc aggtagagag attgcagtaa aagatctgat agaaaaggct caaattcctg 5160
 aggctcttaa ggaaatcaaa gatagagaga cacttcaaag attttgtaga tatatggaag 5220
 cattagtagc ctatttcaag ttttatggag gtaaagatta atgacattcg ctaagattaa 5280
 attttcagct caaattcggt tagagacagg cctccatatt ggtggaagcg atgcttttgc 5340
 agccattggt gcaatcgatt cgcctgttat taaagatcct attaccaacc taccgatcat 5400
 tcctggttca agtctcaaag gaaaaatgag aacgcttctt gccaaaggttt ataataaaaa 5460
 ggtagctgag aaaccaagcg atgacagtga tattcttagc cgtttatttg ggaatagtaa 5520
 agataaacga ttcaaatgg gacgcttgat ttttcgtgat gccttcttgt caaacgctga 5580
 tgagctagac tctcttgggg taagaagtta tacagaagta aaatttgaaa atacaattga 5640
 ccgtatcact gccgaagcta atccaagaca aattgaacgt gctattcgtta ccagtacttt 5700
 tgatttcgag ttgatttatg aaattacaga tgagaatgaa aatcaagtcg aagaagattt 5760
 caaagtgatt cgagatgggt taaaactgct tgaacttgat tatcttgggt gttctggatc 5820
 tcgaggttac ggtaagggtg cttttgaaaa actcaaagct actaccgtat ttggtaatta 5880
 tgatgttaaa acattaaatg aacttttaac tgccggaggtc taatatgacc tataaactgt 5940
 atattatgac ctttcagaat gctcattttg gttcggggcac tcttgatagc tcaaaattaa 6000
 cattctcagc agaccgtatc ttctcagcac tagtgctaga atccctaaaa atgggaaaaac 6060
 tcgatgcatt tcttgccgaa gctaaccaag acaagttcac gctcacagat gcctttccat 6120
 ttcaatttgg tccctttttg ccgaaaccga ttggttatcc caaacatgac caaatagatc 6180
 aatcagttga tgtcaaagag gttcgcctgc aagcaaaatt gtctaagaaa ctgcaatttc 6240
 ttgctctaga aaatggtgac gattatctca atggagagtt atttgaaaat gaagagcatg 6300
 cagtcacga tactgtgaca aaaaatcaac cacataagga cggcaatctt tatcaggtag 6360
 ctacaaccag attttcaa at gatacgtcgc tttacgtcat cgcaaacgaa tctgatttgc 6420
 ttaatgagtt gatgtctagt cttcagtatt caggtcttgg tggaaagcgt tcaagtgggt 6480
 ttggtcgttt tgagttagat attcaaaata tcccactaga attgtcagat agactgacta 6540
 agaatcattc agataaagtg atgagtctta cgacagcact tcctgtagat gctgaccttg 6600
 aagaagcaat ggaagatgga cttacttat taactaaatc aagtggtttt gcatttagtc 6660
 atgctaccaa tgagaattat cgtaagcagg atctttacaa atttgcttct ggttcaactt 6720
 ttagtaaaac atttgaaggt cagattggtg atgtgagacc acttgatttc cctcatgctg 6780
 ttttaattaa tgctaaacca ctcttcttta aattggagggt ataaaaatga aaaatgacta 6840
 tagaacattt aaattaagcc tcctgacact tgctccaatt catattggta atggagagaa 6900
 gtatacctct agagaattta tctatgaaaa taagaagttt tactttcctg acatggggaa 6960

ES 2 398 918 T3

attctataat aaaatggtgg agaagaggct tgctgaaaag tttgaagcat ttctaattca 7020
aactcgtcca aatgcacgta ataatcgtct tatttccttc ttaaatagata accgaattgc 7080
agagcgttct tttggagggt atagtatctc tgaaacagggt ttagaatcgg acaaaaatcc 7140
tgattcagcc ggagctatta acgaagttaa taaatttatt cgagatgctt ttggaaatcc 7200
ctacattcct ggtagctcac taaaagggtgc tattcgtacc attttaatga atactacccc 7260
taagtggaat aatgaaaatg ctgtaaaatga ctttgggaaga tttccgaaag agaataagaa 7320
ccttatccct tggggaccaaa aaaagggaaa agaatacgat gatttggtta acgcaattcg 7380
tgtgagtgat agtaagcctt ttgataataa gagtcttatac ttagtgcaga aatgggatta 7440
ttcagcgaaa acaataaaag ctaaaccact tccttgtat agagaatcaa tctctccatt 7500
aacaaaaatt gaatttgaga ttacaacaac cactgatgaa gctggaagat tgattgaaga 7560
attaggtaag agagcacaag cgttttataa agactataag gcatttttcc tatctgaatt 7620
tcctgatgat aagattcaag ccaatctaca ataccaatt tatttaggtg cggggagcgg 7680
tgcttgaca aagactctat ttaagcaagc tgatggtatt ttacaaagac gatacagtcg 7740
aatgaaaact aaaatggtta aaaaaggagt tcttaagctc acaaaagcac ctcttaaac 7800
agttaagatt ccatctggta atcattcatt agtcaagaac cacgagtcct tttatgaaat 7860
gggaaaagct aatttcatga ttaaggagat tgataaatga 7900

<210> 479

<211> 624

5 <212> ADN

<213> Streptococcus thermophilus

<400> 479

atgagcgtt tatatagtca aaggccaat tattacctgt ccttatctga acaagaatt 60
atcattaata atgataataa agagattgtc aaagaagtgt ccatttcact cgttgataat 120
gtattacttt ttggtaatgc acaactgacc acccaactca tcaaagcctt gtcaaagaac 180
aaggatgaatg tttactatct ctcaaatggt ggtcaattta tttctagat tgaaacccac 240
aggcaggacg aattccaaaa gcaagagttg caagcaaagg cttatcttga agaggatttc 300
cgttttagagg ttgagaggag tattgctacg accaaggatga ggcaccaat tgccttactt 360
agagagtttg atacggatgg tctactagat acctcagatt attctagggt tgaagatagt 420
gtcaatgata ttcagaaagc ttattccatt acagaaatta tgggttacga aggtcgcctt 480
gcgaaatcct atttttacta tctgaattta ctgcttcta atgactttca ttttaatggt 540
aggagtagac ggccctgggga ggattgtttt aacagtgcc tcaattttgg ctatagtatc 600
10 ttatattcct gcttaatggg ctga 624

ES 2 398 918 T3

<210> 480
 <211> 396
 <212> ADN
 5 <213> Streptococcus thermophilus

 <400> 480

 ttgcttaatg ggctgattaa gaaaaacggg ctaagcttgg gatttggggg aattcacaag 60
 catcatcagc atcatgcgac cttggccagt gatttaatgg aagaatggag acctatcatc 120
 gtcgataata cgcttatgga gttggtacga aatggtaaac ttcttttaag tcattttgaa 180
 aataaggatc aagacttcat actcacccat gaaggcagag aaatctttgc acgggcttta 240
 cgttcaagaa tattagaagt ccatcagtat attgagttag ataaaaaacg ctattctttt 300
 ctttatacag cagataggca aatcaagagt ttgattaggg cttttagaga acttgaccct 360
 agtctctatg agacaagtta cacaggaggg cattaa 396
 10
 <210> 481
 <211> 330
 <212> ADN
 <213> Streptococcus thermophilus
 15
 <400> 481

 atgggacttt actttaacct cagcgaagaa gagcgtgagt ttgccaaaca aaaaaccatg 60
 ttttgtctga ttatttatga tattcgaagt aacaaacgta gacttaaact ctcgaaatta 120
 cttgagggtt atggcgtgag ggtgcaaaaa tcctgtttcg aagtcaacct gtcaagaaat 180
 gattatcagt ctctccttaa ggatcgcgag ggcttctaca aggctgatga agaagacagc 240
 ataatagtgt atgtgacaac caaagaagag gtgactagtt ttagccccta ccatagtgct 300
 gaaaaattag atgacattct cttcttctaa 330
 20
 <210> 482
 <211> 732
 <212> ADN
 <213> Streptococcus thermophilus
 25 <400> 482

ES 2 398 918 T3

atgaaaaaat tagtatttac ttttaaaagg atcgaccatc ctgcacaaga tttggctggt 60
aaatttcattg gcttcttgat ggagcagttg gatagtgact atgttgatta tctgcatcag 120
cagcaaacia atccctatgc gaccaaggta atccaaggga aagaaaacac gcagtggggt 180
gtacatctgc tcacagacga catcgaggat aaggttttta tgaccttatt acagattaaa 240
gaggtgtcct taaacgatct gcctaaactc agtgtcgaaa aagttgagat tcaggagttg 300
ggggcagata aactggttaga gattttcaat agtgaggaaa atcaaacctt tttttcaatt 360
atttttgaga ctccaacagg ttttaaatct caaggttcct acgtcatctt cccgtctatg 420
cgtttgattt ttcaaagttt gatgcaaaaag tatggaagggt tggttgaaaa tcaacctgaa 480
attgaagagg ataccttaga ttacctatct gaacacagca ctatcacgaa ttatcgcttg 540
gagacgagtt atttcagggt gcacaggcaa cgaattcctg ccttttagagg aaagttaacc 600
tttaaagtac aaggcgccca aactctaaaa gcttatgtca aaatgcttct aacattcggg 660
gaatattcag gtcttggcat gaaaacgagt ctcggtatgg gagggataaa gcttgaagaa 720
agaaaagatt ga 732

<210> 483

5 <211> 2277

<212> ADN

<213> Streptococcus thermophilus

10 <400> 483

ES 2 398 918 T3

ttgaagaaag	aaaagattga	tttattttac	ggagctcttt	tgcatgatat	cggttaaggtc	60
attcaaaggg	cgacaggaga	acgaaaaaaaa	cacgccttgg	taggcgcgga	ttggtttgat	120
gagattgctg	ataatcaagt	tatttccgat	caaattagat	atcacatggc	taactaccag	180
agtgataaac	ttggaaatga	ccatcttgct	tacataactt	atatacgtga	taacattgcc	240
tctggtgtcg	acagaagaca	gtcaaatgag	gagagtgacg	aggatacatc	agctaagatt	300
tgggatacct	atacaaacca	ggctgatatt	tttaacgttt	ttggggcaca	aacggataaa	360
cgctacttta	aaccgacggt	tctaaacttg	aaatctaaac	ctaactttgc	gtcggcaaca	420
tatgaacctt	tctcaaaag	tgattatgcg	gcaattgCGA	ctcgtatcaa	aatgaattg	480
gcagaatttg	agtttaata	agtacaaatt	gactctttgt	taaatctggt	cgaagcaacc	540
ctctcttttg	tgccctcttc	gactaatact	aaagaaatcg	ctgatatttc	acttgctgat	600
catagtcgtc	tgacagcagc	ttttgctcta	gccatctatg	attacttgga	agacaaaggt	660
cgtcataact	ataaggagga	cttgtttact	aaagcatcag	ccttttatga	ggaagaagct	720
tttctcctag	ctagctttga	cttatcaggg	attcaagact	ttatctataa	tattaatatt	780
gcgacgaatg	gtgctgctaa	acaattgaag	gctagatctt	tatatcttga	ctttatgagc	840
gagtatatag	cagacagttt	acttgataaa	ctaggcctca	atcgggctaa	tatgctctat	900
gtcggtgggg	gacatgctta	ctttgtccta	gccaatactg	aaaaaacggt	agaaacactc	960
gttcaatttg	aaaaagattt	caatcaattt	ttattggcaa	atctccaaac	cagattatat	1020
gttgcctttg	gttggggaag	ctttgCGGct	aaggatatca	tgagcgaact	gaactcacct	1080
gaaagctata	gacaggtcta	tcaaaaggct	agtcgcatga	tttctgagaa	aaaaatctca	1140
aggtatgatt	atcaaaccct	tatgttggtg	aacaggggCG	gtaaactctc	tgaaagagag	1200
tgcgagattt	gtcattccgt	tgagaattta	gttgcttatac	atgaccaaaa	agtgtgtgac	1260
atgtgtcgag	gcttgatca	atcttctaaa	gagattgccc	atgaccattt	cattatcact	1320
gaaaatgaag	ggcttcctat	tggtccgaac	gcatgtctta	agggtgttgc	atgtgaaaag	1380
ctgagccaag	aagctttttc	ccgtgtctat	gtcaaaaatg	actataaggc	tggtacagtt	1440
aaggcaaccc	atgtttttgt	tggagattac	cagtatgatg	aaatatacaa	ttatgctgcc	1500

ES 2 398 918 T3

ttatctaaaa acgaaaatgg gttaggtatt aaacgtttag ctgttgtacg tcttgacgtg 1560
gatgatttgg gagcagcctt tatggctggc ttctcccaac aaggaaatgg gcaatatagt 1620
actctatcac gctcagccac tttctctcga agcatgagtc ttttcttcaa ggtttatatt 1680
aaccagtttg ctagtgataa gaagctctct atcatctatg ccggtgggga tgatgttttt 1740
gctattggct cttggcaaga tattattgcc tttactgttg aacttcgtga gaacttcatt 1800
aatggacaa atggaaaact aacactatca gctggatcg gtctgtttgc tgataagacc 1860
cctattagct taatggcaca tcaaacaggg gagctagaag aacagctaa aggcaatgag 1920
aaagatagta tttcactctt tagttccgac tataccttta aatttgatcg gtttatcact 1980
aatgtttacg acgataagtt agagcagatt cgctatttct ttaatcacca agatgaacga 2040
ggcaagaatt tcatttataa attgattgaa ttgcttcgaa attatgatcg tatgaatatg 2100
gcacgtttag cttattattt aacacgactt gaagaattga cgcgtgaaac agacagggat 2160
aaatttaaaa catttaaaaa tttattctat tcttggtaga caaataagga tgataaggat 2220
agaaaagaag cagagttagc cttgcttctc tatatctatg agattagaaa ggattag 2277

<210> 484

<211> 381

5 <212> ADN

<213> Streptococcus thermophilus

<400> 484

atgacaatct tgactgatga gaattacgtt gatattgcag aaaaagcaat tctaaaacta 60
gaaagaaata ctaggaacag aaagaatcct gatgccttct ttcttacaac aagtaagctc 120
agaaacttgc tgagcttaac tagtacactt tttgatgaga gtaagggtcaa agaatatgat 180
gctctccttg atcgtattgc ttatttaaga gtacaatttg tctaccaagc aggtagagag 240
attgcagtaa aagatctgat agaaaaggct caaattcttg aggctcttaa ggaaatcaaa 300
gatagagaga cacttcaaag attttgtaga tatatggaag cattagtagc ctatttcaag 360
10 ttttatggag gtaaagatta a 381

<210> 485

<211> 663

<212> ADN

15 <213> Streptococcus thermophilus

<400> 485

ES 2 398 918 T3

atgacattcg ctaagattaa attttcagct caaatcgtt tagagacagg cctccatatt 60
 ggtggaagcg atgcttttgc agccattggt gcaatcgatt cgctgttat taaagatcct 120
 attaccaacc taccgatcat tcctggttca agtctcaaag gaaaaatgag aacgcttctt 180
 gccaaaggttt ataatgaaaa ggtagctgag aaaccaagcg atgacagtga tattcttagc 240
 cgtttatttg ggaatagtaa agataaacga ttcaaatgg gacgcttgat ttttcgtgat 300
 gccttcttgt caaacgctga tgagctagac tctcttgggg taagaagtta tacagaagta 360
 aaatttgaaa atacaattga ccgtatcact gccgaagcta atccaagaca aattgaacgt 420
 gctattcgta ccagtacttt tgatttcgag ttgatttatg aaattacaga tgagaatgaa 480
 aatcaagtcg aagaagattht caaagtgatt cgagatgggt taaaactgct tgaacttgat 540
 tatcttgggt gttctggatc tcgaggttac ggtaagggtg cttttgaaaa actcaaagct 600
 actaccgtat ttggttaatta tgatgttaaa acattaaatg aacttttaac tgcggaggtc 660
 taa 663

5 <210> 486
 <211> 900
 <212> ADN
 <213> Streptococcus thermophilus

10 <400> 486

ES 2 398 918 T3

atgacctata aactgtatat tatgaccttt cagaatgctc attttggttc gggcactcctt 60
gatagctcaa aattaacatt ctccagcagac cgtatcttct cagcactagt gctagaatcc 120
ctaaaaatgg gaaaactcga tgcatttctt gcggaagcta accaagacaa gttcacgctc 180
acagatgcct ttccatttca atttgggtccc tttttgccga aaccgattgg ttatcccaaa 240
catgacccaaa tagatcaatc agttgatgtc aaagaggttc gccgtcaagc aaaattgtct 300
aagaaactgc aatttcttgc tctagaaaat gttgacgatt atctcaatgg agagttatct 360
gaaaatgaag agcatgcagt catcgatact gtgacaaaaa atcaaccaca taaggacggc 420
aatctttatc aggtagctac aaccagattt tcaaatgata cgtcgcttta cgtcacgca 480
aacgaatctg atttgcttaa tgagttgatg tctagtcttc agtattcagg tcttgggtgga 540
aagcgttcaa gtggttttgg tcgttttgag ttagatattc aaaatatccc actagaattg 600
tcagatagac tgactaagaa tcattcagat aaagtgatga gtcttaacgac agcacttcct 660
gtagatgctg accttgaaga agcaatggaa gatggacatt acttattaac taaatcaagt 720
ggttttgcat ttagtcatgc taccaatgag aattatcgta agcaggatct ttacaaatct 780
gcttctgggt caacttttag taaaacattt gaaggtcaga ttggtgatgt gagaccactt 840
gatttccttc atgctgtttt aaattatgct aaaccactct tctttaaatt ggaggtataa 900

<210> 487

<211> 1074

5 <212> ADN

<213> Streptococcus thermophilus

<400> 487

atgaaaaatg actatagaac atttaaatta agcctcctga cacttgctcc aattcatatt 60
ggtaatggag agaagtatac ctctagagaa tttatctatg aaaataagaa gttttacttt 120
10 cctgacatgg ggaaattcta taataaaatg gtggagaaga ggcttgctga aaagtttgaa 180

ES 2 398 918 T3

gcatttctaa ttcaaactcg tccaaatgca cgtaataatc gtcttatttc cttcttaaat 240
gataaccgaa ttgcagagcg ttcttttggg ggttatagta tctctgaaac aggttttagaa 300
tcggacaaaa atcctgattc agccggagct attaacgaag ttaataaatt tattcgagat 360
gcttttggaa atccctacat tcttggtagc tcaactaaaag gtgctattcg taccatttta 420
atgaatacta cccctaagtg gaataatgaa aatgctgtaa atgactttgg aagatttccg 480
aaagagaata agaaccttat cccttgggga ccaaaaaagg gaaaagaata cgatgatttg 540
tttaacgcaa ttcgtgtgag tgatagtaag ccttttgata ataagagtct tatcttagtg 600
cagaaatggg attattcagc gaaaacaaat aaagctaaac cacttccctt gtatagagaa 660
tcaatctctc cattaacaaa aattgaattt gagattacaa caaccactga tgaagctgga 720
agattgattg aagaattagg taagagagca caagcgtttt ataaagacta taaggcattt 780
ttcctatctg aatttcctga tgataagatt caagccaatc tacaataccc aatttattta 840
ggtgcgggga gcggtgcttg gacaaagact ctatttaagc aagctgatgg tattttacaa 900
agacgatata gtcgaatgaa aactaaaatg gttaaaaaag gagttcttaa gctcacaaaa 960
gcaccttta aaacagttaa gattccatct ggtaatcatt cattagtcaa gaaccacgag 1020
tccttttatg aaatgggaaa agctaatttc atgattaagg agattgataa atga 1074

<210> 488

<211> 5974

5 <212> ADN

<213> Streptococcus agalactiae

<400> 488

ES 2 398 918 T3

atgaataagc	catattcaat	aggccttgac	atcggtacta	attccgtcgg	atggagcatt	60
attacagatg	attataaagt	acctgctaag	aagatgagag	ttttagggaa	cactgataaa	120
gaatatatta	agaagaatct	cataggtgct	ctgctttttg	atggcgggaa	tactgctgca	180
gatagacgct	tgaagcgaac	tgctcgtcgt	cgttatacac	gtcgtagaaa	tcgtattcta	240
tatttacaag	aaatTTTTgc	agaggaaatg	agtaaagttg	atgatagttt	ctttcatcga	300
ttagaggatt	cttttctagt	tgaggaagat	aagagagggg	gcaagtatcc	tatctttgca	360
acattgcagg	aagagaaaga	ttatcatgaa	aaatTTTcga	caatctatca	tttgagaaaa	420
gaattagctg	acaagaaaga	aaaagcagac	cttcgtctta	tttatattgc	tctagctcat	480
atcattaaat	ttagagggca	tttcctaatt	gaggatgata	gctttgatgt	caggaataca	540
gacatttcaa	aacaatatca	agatTTTTta	gaaatcttta	atacaacttt	tgaaaataat	600
gatttgttat	ctcaaaacgt	tgacgtagag	gcaatactaa	cagataagat	tagcaagtct	660
gcgaagaaag	atcgtatTTT	agcgcagtat	cctaaccaaa	aatctactgg	cattTTTTgca	720
gaatTTTTga	aattgattgt	cggaaatcaa	gctgacttca	agaaatattt	caatttgag	780
gataaaacgc	cgcttcaatt	cgctaaggat	agctacgatg	aagatttaga	aaatcttctt	840

ES 2 398 918 T3

ggacagattg gtgatgaatt tgcagactta ttctcagcag cgaaaaagtt atatgatagt 900
 gtccttttgt ctggcattct tacagtaatc gacctcagta ccaaggcgcc actttcagct 960
 tctatgattc agcgttatga tgaacataga gaggacttga aacagttaaa acaattcgta 1020
 aaagcttcat tgccggaaaa atatcaagaa atatttgctg attcatcaaa agatggctac 1080
 gctggttata ttgaaggtaa aactaatcaa gaagcttttt ataaatacct gtcaaaaattg 1140
 ttgaccaagc aagaagatag cgagaatfff cttgaaaaaa tcaagaatga agatttcttg 1200
 agaaaacaaa ggacctttga taatggctca attccacacc aagtccattt gacagagctg 1260
 aaagctatta tccgcogtca atcagaatac tatcccttct tgaaagagaa tcaagatagg 1320
 attgaaaaaa tccttacctt tagaattcct tattatatcg ggccactagc acgtgagaag 1380
 agtgattttg catggatgac tcgcaaaaaca gatgacagta ttcgaccttg gaattttgaa 1440
 gacttggttg ataaagaaaa atctgcggaa gcttttatcc atcgatgac caacaatgat 1500
 ttttatcttc ctgaagaaaa agttttacca aagcatagtc ttatttatga aaaatttacg 1560
 gtctataatg agttgactaa ggtagatat aaaaatgagc aaggtgagac ttattttttt 1620
 gatagcaata ttaaacaaga aatctttgat ggagtattca aggaacatcg taaggtatcc 1680
 aagaagaagt tgctagatft tctggctaaa gaatatgagg agtttaggat agtagatgft 1740
 attggtctag ataaagaaaa taaagctttc aacgcctcat tgggaactta ccacgatctc 1800
 gaaaaaatac tagacaaaga ttttctagat aatccagata atgagtctat tctggaagat 1860
 atcgccaaa ctctaacatt atttgaagac agagaaatga ttaagaagcg tcttgaaaac 1920
 tataaagatc tttttacaga gtcacaacta aaaaaactct atcgtcgtca ctatactggc 1980
 tggggacgat tgtctgctaa gttaatcaat ggtattcgag ataaagagag tcaaaaaaca 2040
 atcttgact atcttattga tgatggtaga tctaactgca actttatgca gttgataaat 2100
 gatgatggtc tatctttcaa atcaattatc agtaaggcac aggctggtag tcattcagat 2160
 aatctaaaag aagttgtagg tgagcttga ggtagccctg ctattaaaaa ggaatttcta 2220
 caaagtttga aaattggtga tgagcttgtt aaagtcatgg gatacgaacc tgaacaaatt 2280
 gtggttgaga tggcgcgtga gaatcaaaaca acaaatcaag gtcgtcgtaa ctctcgacaa 2340
 cgctataaac ttcttgatga tggcgtaag aatctagcta gtgacttgaa tggcaatatt 2400
 ttgaaagaat atcctacgga taatcaagcg ttgcaaaatg aaagactttt cctttactac 2460
 ttacaaaacg gaagagatat gtatacaggg gaagctctag atattgacaa tttaaagtcaa 2520
 tatgatattg accacattat tcctcaagct ttcataaaaag atgattctat tgataatcgt 2580
 gttttggtat catctgctaa aaatcgtgga aagtcagatg atgttcctag ccttgaaatt 2640
 gtaaaagatt gtaaagtttt ctggaaaaaa ttacttgatg ctaagttaat gagtcagcgt 2700
 aaqtatgata atttgactaa ggcagagcgc ggaggcctaa cttccgatga taaggcaaga 2760

ES 2 398 918 T3

tttatccaac gtcagttggt tgagacacga caaattacca agcatggtgc ccgtatcttg 2820
 gatgaacgct ttaataatga gcttgatagt aaaggtagaa ggatccgcaa agttaaatt 2880
 gtaaccttga agtcaaattt ggtttcaa atccgaaaag aatttgatt ctataaatt 2940
 cgtgaagtta acaattatca ccatgcacat gatgcctatc ttaatgcagt agttgctaaa 3000
 gctattctaa ccaaatatcc tcagttagag ccagaatttg tctacggcga ctatccaaa 3060
 tataatagtt acaaacgcg taaatccgct acagaaaagc ttttttcta ttcaaatatt 3120
 atgaacttct ttaaaactaa ggtaacttta gcggatggaa ccgttggtgt aaaagatgat 3180
 attgaagtta ataatgatac gggtgaaatt gtttgggata aaaagaaaca ctttgcgaca 3240
 gttagaaaag tcttgcata ccctcagaac aatatcgtga agaagacaga gattcagaca 3300
 ggtggtttct ctaaggaatc aatcttggcg catggtaact cagataagtt gattccaaga 3360
 aaaacgaagg atatttattt agatcctaag aaatatggag gttttgatag tccgatagta 3420
 gcttactctg ttttagttgt agctgatatc aaaaagggta aagcacaaaa actaaaaaca 3480
 gttacggaac ttttaggaat taccatcatg gagaggtcca gatttgagaa aatccatca 3540
 gctttccttg aatcaaaaagg ctatttaaat attagggctg ataaactaat tttttgccc 3600
 aagtatagtc tgttcgaatt agaaaatggg cgtcgtcgtat tacttgctag tgctggtgaa 3660
 ttacaaaaag gtaatgagct agccttacca acacaattta tgaagttctt atacctgca 3720
 agtcgttata atgagtcaaa aggtaaacca gaggagattg agaagaaaca agaatttgta 3780
 aatcaacatg tctcttattt tgatgacatc cttcaattaa ttaatgattt ttcaaacga 3840
 gttattctag cagatgctaa ttagagaaa atcaataagc tttaccaaga taataaggaa 3900
 aatatatcag tagatgaact tgctaataat attatcaatc ttttacttt taccagtcta 3960
 ggagctccag cagcttttaa attttttgat aaaatagttg atagaaaacg ctatacatca 4020
 actaaagaag tacttaattc taccetaatt catcaatcta ttactggact ttatgaaaca 4080
 cgtattgatt tgggtaagtt aggagaagat tgatatggca ggttggcga cgttggtgt 4140
 aaatacacat tctaagctct cttataaaaa taatcatctg atttttaag attcttatca 4200
 gacggaaatg attcatctat cagagattga cattctaate atggaaaca cagatatcgt 4260
 tttgtcgacc atgctgatta aacgtttggg tgatgaaaat attttagtta tttttgtga 4320
 cgataaacgc ttccaacag ctatgttaat gccgtactat gccagacatg attcgagttt 4380
 acaattatct aggcagatgt catggattga agatgtcaaa gcagatgttt ggacatcaat 4440
 tattgcacaa aaaattttga atcagtcttt ttatctcggg gagggttctt tctttgaaaa 4500
 atcccagtct attatgaatc tctacatga cttagaacct tttgatcctt ctaatcgtga 4560
 ggggcatgct gctaggattt atttcaatac actttttgga aatgattttt caagagagca 4620
 ggataatcca ataatgctg gtttagacta cggatattca ttgcttttga gtatggttgc 4680

ES 2 398 918 T3

gcgtgaagtt gttaagtgtg gttgcatgac acaatttggc ttgaagcatg ctaatcaatt 4740
 taatcagttc aacctagcaa gcgatattat ggaaccattt cgcccaatcg ttgataggat 4800
 tatttatgaa aataggcaga gtgattttgt caaaatgaaa agagaactct tttctatggt 4860
 ttcagagaca tacagctaca atggtaaaga aatgtatctc tcaaatttg tcagcgacta 4920
 taccaaaaaa gttattaagt cgctaaatag tgatgggaat ggaattccgg agtttaggat 4980
 atgagttatc ggtatatgcg aatgatttta atgtttgata tgcctactga aacagcagaa 5040
 gaacggaagg cgtatcgtaa gtttagaaaag tttctcttga gcgaaggctt tatcatgcat 5100
 cagttttctg tttatagtaa attattactc aataatacag ctaataatgc tatgataggt 5160
 cggcttaaag tgaataatcc taaaaagggg aatatcacac tcttaacagt tacggaaaaa 5220
 caatttgcgga gaatgggtta cctccatgga gaacgcaaca caagtgttgc caactctgat 5280
 agtcgcttgg ttttcctagg agattcttat gatcaagatt aattttccaa ttttagatga 5340
 accattagtg ttaagtaatg ctacgatttt aacgatagaa gatgtttcag tttattcttc 5400
 attggtgaaa catttttatc aatatgacgt agatgaacat ttgaaattat ttgatgataa 5460
 gcagaaaagt ctgaaggcaa cagagttaat gctggttaca gatatcttag gatagatgt 5520
 caactcagca cctattctaa agttgataca tggtgactta gaaaatcaat tcaacgaaaa 5580
 gccagaagtg aatcaatgg tagaaaaatt agcagctact attacagAAC ttatcgcatt 5640
 tgagtgtcta gagaatgagc ttgattttaga atacgatgaa attaagattt tagaactcat 5700
 taaggcactg ggagtcaaaa ttgagacaca gagcgacact atctttgaaa aatgttttga 5760
 aattatacaa gtttaccatt atttaacgaa aaagaatctc ttggtttttg ttaatagcgg 5820
 agcttatctt accaaagatg aagttataaa attatgtgaa tacatcaatt taatgcaaaa 5880
 gtcagtactc tttctagaac ctagaagact ctatgattta ccgcaatatg ttattgataa 5940
 ggattatttc ttgataggcg aaaatatggt ataa 5974

<210> 489

<211> 4113

5 <212> ADN

<213> Streptococcus agalactiae

<400> 489

ES 2 398 918 T3

atgaataagc	catattcaat	aggccttgac	atcggtacta	attccgtcgg	atggagcatt	60
attacagatg	attataaagt	acctgctaag	aagatgagag	ttttagggaa	cactgataaa	120
gaatatatta	agaagaatct	cataggtgct	ctgctttttg	atggcgggaa	tactgctgca	180
gatagacgct	tgaagcgaac	tgctcgtcgt	cgttatacac	gtcgtagaaa	tcgtattcta	240
tatttacaag	aaatTTTTgc	agaggaaatg	agtaaagttg	atgatagttt	ctttcatcga	300
ttagaggatt	cttttctagt	tgaggaagat	aagagagggg	gcaagtatcc	tatctttgca	360

ES 2 398 918 T3

acattgcagg aagagaaaga ttatcatgaa aaattttcga caatctatca tttgagaaaa 420
 gaattagctg acaagaaaga aaaagcagac cttcgtctta tttatattgc tctagctcat 480
 atcattaat ttagagggca tttcctaatt gaggatgata gctttgatgt caggaataca 540
 gacatttcaa aacaatatca agatTTTTta gaaatcttta atacaacttt tgaaaataat 600
 gatttgttat ctcaaaacgt tgacgtagag gcaatactaa cagataagat tagcaagtct 660
 gcgaagaaag atcgtatTTT agcgcagtat cctaaccaaa aatctactgg catttttgca 720
 gaatTTTTga aattgattgt cggaaatcaa gctgacttca agaaatattt caatttgag 780
 gataaaacgc cgcttcaatt cgctaaggat agctacgatg aagattttaga aaatcttctt 840
 ggacagattg gtgatgaatt tgcagactta ttctcagcag cgaaaagtt atatgatagt 900
 gtcctTTTgt ctggcattct tacagtaatc gacctcagta ccaaggcgcc actttcagct 960
 tctatgattc agcgttatga tgaacataga gaggacttga aacagttaa acaattcgta 1020
 aaagcttcat tgccggaaaa atatcaagaa atatTTTgctg attcatcaaa agatggctac 1080
 gctggttata ttgaaggtaa aactaatcaa gaagctTTTT ataaatacct gtcaaaaattg 1140
 ttgaccaagc aagaagatag cgagaatttt cttgaaaaaa tcaagaatga agatttcttg 1200
 agaaaaacaaa ggacctttga taatggctca attccacacc aagtccattt gacagagctg 1260
 aaagctatta tccgcgtca atcagaatac tatcccttct tgaagagaa tcaagatagg 1320
 attgaaaaaa tccttacctt tagaattcct tatttatatcg ggccactagc acgtgagaag 1380
 agtgattttg catggatgac tcgcaaaaca gatgacagta ttcgaccttg gaattttgaa 1440
 gacttggttg ataaagaaaa atctgcggaa gcttttatcc atcgtatgac caacaatgat 1500
 ttttatcttc ctgaagaaaa agttttacca aagcatagtc ttatttatga aaaatttacg 1560
 gtctataatg agttgactaa ggtagatat aaaaatgagc aaggtgagac ttattTTTT 1620
 gatagcaata ttaaacaaga aatctttgat ggagtattca aggaacatcg taaggtatcc 1680
 aagaagaagt tgctagattt tctggctaaa gaatatgagg agtttaggat agtagatggt 1740
 attggtctag ataaagaaaa taaagctttc aacgcctcat tgggaactta ccacgatctc 1800
 gaaaaaatac tagacaaaga ttttctagat aatccagata atgagtctat tctggaagat 1860
 atcgtccaaa ctctaacatt atttgaagac agagaaatga ttaagaagcg tcttgaaaac 1920
 tataaagatc tttttacaga gtcacaacta aaaaaactct atcgtcgtca ctatactggc 1980
 tggggacgat tgtctgctaa gttaatcaat ggtattcgag ataaagagag tcaaaaaaca 2040
 atcttgact atcttattga tgatggtaga tctaategca actttatgca gttgataaat 2100
 gatgatggtc tatctttcaa atcaattatc agtaaggcac aggctggtag tcattcagat 2160
 aatctaaaag aagttgtagg tgagcttgca ggtagccctg ctattaaaaa gggatttcta 2220
 caaagtttga aaattgttga tgagcttgtt aaagtcatgg gatacgaacc tgaacaaatt 2280

ES 2 398 918 T3

gtggttgaga tggcgcgtga gaatcaaca acaaatcaag gtcgtcgtaa ctctcgacaa 2340
 cgctataaac ttcttgatga tggcgttaag aatctagcta gtgacttgaa tggcaatatt 2400
 ttgaaagaat atcctacgga taatcaagcg ttgcaaaatg aaagactttt cctttactac 2460
 ttacaaaacg gaagagatat gtatacaggg gaagctctag atattgacaa ttaagtcaa 2520
 tatgatattg accacattat tcctcaagct ttcataaaaag atgattctat tgataatcgt 2580
 gttttggtat catctgctaa aaatcgtgga aagtcagatg atgttcctag ccttgaaatt 2640
 gtaaaaagatt gtaaagtttt ctggaaaaaa ttacttgatg ctaagttaat gagtcagcgt 2700
 aagtatgata atttgactaa ggagagcgc ggaggcctaa cttccgatga taaggcaaga 2760
 tttatccaac gtcagttggt tgagacacga caaattacca agcatgttgc ccgatcttg 2820
 gatgaacgct ttaataatga gcttgatagt aaaggtagaa ggatccgcaa agttaaatt 2880
 gtaaccttga agtcaaattt ggtttcaaat ttccgaaaag aatttgatt ctataaatt 2940
 cgtgaagtta acaattatca ccatgcacat gatgcctatc ttaatgcagt agttgctaaa 3000
 gctattctaa ccaaatacc tcagttagag ccagaatttg tctacggcga ctatccaaaa 3060
 tataatagtt acaaaacgcg taaatccgct acagaaaagc ttttttcta ttcaaatatt 3120
 atgaacttct ttaaaactaa ggtaacttta gcggatggaa ccgttgttgt aaaagatgat 3180
 attgaagtta ataataatgac ggggtgaaatt gtttgggata aaaagaaaca ctttgcgaca 3240
 gttagaaaag tcttgcata ccctcagaac aatatcgtga agaagacaga gattcagaca 3300
 ggtggtttct ctaaggaatc aatcttggcg catggtaact cagataagtt gattccaaga 3360
 aaaacgaagg atatttattt agatcctaag aaatatggag gttttgatag tccgatagta 3420
 gcttactctg ttttagttgt agctgatatc aaaaagggtg aagcacaaaa actaaaaaca 3480
 gttacggaac ttttaggaat taccatcatg gagaggtcca gatttgagaa aaatccatca 3540
 gctttccttg aatcaaaagg ctattttaat attagggctg ataaactaat tttttgccc 3600
 aagtatagtc tgttcgaatt agaaaatggg cgtcgtcgat tacttgctag tgctggtgaa 3660
 ttacaaaaag gtaatgagct agccttacca acacaattta tgaagttctt ataccttgca 3720
 agtcgttata atgagtcaaa aggtaaacca gaggagattg agaagaaaca agaatttgta 3780
 aatcaacatg tctcttattt tgatgacatc cttcaattaa ttaatgattt ttcaaacga 3840
 gttattctag cagatgctaa ttagagaaa atcaataagc tttaccaaga taataaggaa 3900
 aatatatcag tagatgaact tgctaataat attatcaatc ttttacttt taccagtcta 3960
 ggagctccag cagcttttaa atttttgat aaaatagttg atagaaaacg ctatacatca 4020
 actaaagaag tacttaattc taccctaatt catcaatcta ttactggact ttatgaaaca 4080
 cgtattgatt tgggtaagtt aggagaagat tga 4113

<210> 490

ES 2 398 918 T3

<211> 870
 <212> ADN
 <213> Streptococcus agalactiae

5 <400> 490

atggcagggtt ggcgaaccgt tgttgtaa atcacattcta agctctctta taaaaataat 60
 catctgattt ttaaagattc ttatcagacg gaaatgattc atctatcaga gattgacatt 120
 ctaatcatgg aaacaacaga tatcgttttg tcgacatgc tgattaaacg tttggttgat 180
 gaaaatattt tagttatatt ttgtgacgat aaacgcttgc caacagctat gttaatgccg 240
 tactatgcca gacatgattc gagtttacia ttatctagggc agatgtcatg gattgaagat 300
 gtcaaagcag atgtttgac atcaattatt gcacaaaaaa ttttgaatca gtctttttat 360
 ctcggtgagt gttctttctt tgaaaaatcc cagtctatta tgaatctcta ccatgactta 420
 gaaccttttg atccttctaa tcgtgagggg catgctgcta ggatttattt caatacactt 480
 tttggaaatg atttttcaag agagcaggat aatccaataa atgctggttt agactacgga 540
 tattcattgc ttttgagtat gtttgcgct gaagttgtta agtgtggttg catgacacaa 600
 tttggcttga agcatgctaa tcaattta atcagttcaacc tagcaagcga tattatggaa 660
 ccatttcgcc caatcgttga taggattatt tatgaaaata ggcagagtga ttttgtcaaa 720
 atgaaaagag aactcttttc tatgttttca gagacataca gctacaatgg taaagaaatg 780
 tatctctcaa atattgtcag cgactatacc aaaaaagtta ttaagtctct aatagtgat 840
 gggaatggaa ttccggagtt taggatatga 870

10 <210> 491
 <211> 327
 <212> ADN
 <213> Streptococcus agalactiae

15 <400> 491

atgcgaatga ttttaatggt tgatatgcct actgaaacag cagaagaacg gaaggcgtat 60
 cgtaagttta gaaagtttct cttgagcga ggctttatca tgcacagtt ttctgtttat 120
 agtaaattat tactcaataa tacagcta atgctatga taggtcggct taaagtgaat 180
 aatcctaaaa agggtaatat cacactctta acagttacgg aaaaacaatt tgcgagaatg 240
 gtttacctcc atggagaacg caacacaagt gttgccaact ctgatagtcg cttggttttc 300
 ctaggagatt cttatgatca agattaa 327

20 <210> 492
 <211> 666
 <212> ADN
 <213> Streptococcus agalactiae

ES 2 398 918 T3

<400> 492

atgatcaaga ttaattttcc aatttttagat gaaccattag tgtaagtaa tgctacgatt 60
 ttaacgatag aagatgtttc agtttattct tcattggtga aacattttta tcaatatgac 120
 5 gtagatgaac atttgaaatt atttgatgat aagcagaaaa gtctgaaggc aacagagtta 180
 atgctgggta cagatatctt aggatacgat gtcaactcag cacctattct aaagttgata 240
 catggtgact tagaaaatca attcaacgaa aagccagaag tgaaatcaat ggtagaaaaa 300
 ttagcagcta ctattacaga acttatcgca tttgagtgtc tagagaatga gcttgattta 360
 gaatacgatg aaattaagat tttagaactc attaaggcac tgggagtcaa aattgagaca 420
 cagagcgaca ctatctttga aaaatgtttt gaaattatac aagtttacca ttatttaacg 480
 aaaaagaatc tcttgggttt tgtaatagc ggagcttatac ttaccaaaaga tgaagttata 540
 aaattatgtg aatacatcaa tttaatgcaa aagtcagtac tctttctaga acctagaaga 600
 ctctatgatt taccgcaata tgttattgat aaggattatt tcttgatagg cgaaaatatg 660
 gtataa 666

<210> 493

<211> 5995

10 <212> ADN

<213> Streptococcus agalactiae

<400> 493

ES 2 398 918 T3

atgaataagc catattcaat aggccttgac atcgg tacta attccg tccg atggagcatt 60
 attacagatg attataaagt acctgctaag aagatgagag ttttagggaa cactgataaa 120
 gaatatatta agaagaatct cataggtgct ctgctttttg atggcgggaa tactgctgca 180
 gatagacgct tgaagcgaac tgctcgtcgt cgttatacac gtcgtagaaa tcgtattcta 240
 tatttacaag aaatTTTTgc agaggaaatg agtaaagttg atgatagttt ctttcatcga 300
 ttagaggatt cttttctagt tgaggaagat aagagaggta gcaagtatcc tatctttgca 360
 acaatgcagg aggagaaata ttatcatgaa aaatTTccga caatctatca tttgagaaaa 420
 gaattggctg acaagaaaga aaaagcagac cttcgtcttg tttatctggc tctagctcat 480
 atcattaat tccagagggca tttcctaatt gaggatgata gatttgatgt gaggaatacc 540
 gatattcaaa aacaatatca agcctTTTTa gaaatTTTTg atactacctt tgaaaataat 600
 catttgttat ctcaaatgt agatgtagaa gcaattctaa cagataagat tagcaagtct 660
 gcgaagaagg atcgcactct agcgcagtat cctaaccaaa aatctactgg tatttttgca 720
 gaatTTTTga aattgattgt cggaaatcaa gctgacttca agaaacattt caatttgag 780
 gataaaacac cgcttcaatt cgctaaggat agctacgatg aagatttaga aaatcttctt 840
 ggacagattg gtgatgaatt tgcagactta ttctcagtag cgaaaagct atatgatagt 900
 gttcttttat ctggcattct tacagtaact gatctcagta ccaaggcgcc actttctgcc 960
 tctatgattc agcgttatga tgaacatcat gaggacttaa agcatctaaa acaattcgta 1020
 aaagcttcat tacctgaaaa ttatcgggaa gtatttgctg attcatcaaa agatggctac 1080

ES 2 398 918 T3

gctggctata ttgaaggcaa aactaatcaa gaagcttttt ataaatatct gttaaaattg 1140
ttgaccaaac aagaaggtag cgagtatttt cttgagaaaa ttaagaatga agattttttg 1200
agaaaacaga gaacctttga taatggctca atcccgcac aagtccattt gacagaattg 1260
agggtatta ttcgacgtca atcagaatac tatccattct tgaagagaa tcaagatagg 1320
attgaaaaaa tccttacctt tagaattcct tattatgtcg ggccactagc acgtgagaag 1380
agtgattttg catggatgac tcgcaaaaaca gatgacagta ttcgaccttg gaattttgaa 1440
gacttggttg ataaagaaaa atctgcggaa gcttttatcc atcgcacgac caacaatgac 1500
ctctatcttc cagaagaaaa agttttacca aagcatagtc ttatttatga aaaatttact 1560
gtttacaatg aattaacgaa ggtagattt ttggcagaag gctttaaaga ttttcaattt 1620
ttaaatagga agcaaaaaga aactatcttt aacagcttgt ttaaggaaaa acgtaaagta 1680
actgaaaagg atattattag ttttttgaat aaagttgatg gatatgaagg aattgcaatc 1740
aaaggaattg agaaacagtt taacgctagc ctttcaacct atcatgatct taaaaaata 1800
cttggaagg atttccttga taatacagat aacgagctta ttttgaaga ttcgctcaa 1860
actctaacct tatttgaaga tagagaaatg attaagaagt gtcttgacat ctataaagat 1920
ttttttacag agtcacagct taaaaagctc tatcgccgct actatactgg ctggggacga 1980
ttgtctgcta agctaataaa tggcatccga aataaagaga atcaaaaaac aatcttgac 2040
tatcttattg atgatggaag tgcaaacgca aacttcatgc agttgataaa tgatgatgat 2100
ctatcattta aaccaattat tgacaaggca cgaactggta gtcattcgga taatctgaaa 2160
gaagttgtag gtgaacttgc tggtagccct gctattaaaa aagggttctt acaaagtttg 2220
aaaatagttg atgagctggt taaagtcacg ggctatgaac ctgaacaaat cgtggttgaa 2280
atggcacgtg agaaccaaac gacagcaaaa ggattaagtc gttcacgaca acgcttgaca 2340
accttgagag aatctcttgc taatttgaag agtaatattt tggaagagaa aaagcctaag 2400
tatgtgaaag atcaagttga aatcatcat ttatctgatg accgtctttt cctttactac 2460
ttacaaaacg gaagagatat gtatacaaaa aaggctctgg atattgataa ttttaagtcaa 2520
tatgatattg accacattat tcctcaagct ttcataaaag atgattctat tgataatcgt 2580
gttttggtat catctgctaa aaatcgtgga aaatcagatg atgttcctag cattgaaatt 2640
gtaaaagctc gcaaaatggt ctggaaaaat ttactggatg ctaagttaat gagtcagcgt 2700
aagtatgata atttgactaa ggcagagcgc ggaggcctaa cttccgatga taaggcaaga 2760
tttatccaac gtcagttggt tgagactcga caaattacca agcatgtagc tcgtatcttg 2820
gatgaacgct tcaataatga agttgataat ggtaaaaaga tttgcaaggt taaaattgta 2880
accttgaagt caaatttgggt ttcaaatttc cgaaaagaat ttggattcta taaaattcgt 2940
gaagttaatg attatcacca tgcacacgat gcttatctta atgcagtagt tgccaaagct 3000

ES 2 398 918 T3

attctaacca aatatccaca gttagagcca gagtttgtct acggaatgta tagacagaaa 3060
 aaactttcga aaatcgttca tgaggataag gaagaaaaat atagtgaagc aaccaggaaa 3120
 atgtttttct actccaactt gatgaatatg ttcaaaagag ttgtgagggt agcagatggg 3180
 tctattggtg taagaccagt aatagaaact ggtagatata tgagaaaaac tgcattgggat 3240
 aaaaagaaac actttgcgac agttagaaaa gtcttgtcat accctcagaa caatatcgtg 3300
 aagaagacag agattcagac aggtgggttc tctaaggaat caatcttggc gcatggtaac 3360
 tcagataagt tgattccaag aaaaacgaag gatatttatt tagatcctaa gaaatatgga 3420
 ggttttgata gtccgatagt agcttactct gtttttagttg tagctgatat caaaaaagggt 3480
 aaagcacaaa aactaaaaac agttacggaa cttttaggaa ttaccatcat ggagagggtcc 3540
 agatttgaga aaaatccatc agctttcctt gaatcaaaag gttattttaa tattagggac 3600
 gataaattaa tgattttacc gaagtatagt ctgttcgaat tagaaaatgg gcgtcgtcga 3660
 ttacttgcta gtgctgggga attacaaaaa ggtaacgagc tagccttacc aacacaattt 3720
 atgaagtctt tataccttgc aagtcggtat aatgagtcaa aaggtaaacc agaggagatt 3780
 gagaagaaac aagaatttgt aatcaacat gtctcttatt ttgatgacat ccttcaatta 3840
 attaatgatt tttcaaaacg agttattcta gcagatgcta atttagagaa aatcaataag 3900
 ctttaccagg ataataagga aatatacca gtatagtaac ttgctaataa tattatcaat 3960
 ctatttactt ttaccagtct aggagctcca gcagctttta aattttttga taaaatagtt 4020
 gatagaaaac gctatacatc aactaaagaa gtacttaatt ctactctaatt ccatcaatct 4080
 attactggac tttatgaaac acgtattgat ttgggtaaat taggagaaga ttgatatggc 4140
 aggttggcga actgttggtg taaatacaca ttctaagctc tcttataaaa ataatcatct 4200
 gatttttaa gattcttacc agacggaaat gattcatctt tcagagattg atattctaatt 4260
 catggaaaac acagatatg ttttgtcgac tatgctgatt aaacgtttgg ttgatgaaaa 4320
 tatttttagtc atattttgtg atgataaacg cttgccaaca gctatgtaa tgccgtacta 4380
 tgctagacat gattcgagtt tacaattatc taggcagatg tcatggattg aggatgtcaa 4440
 agcggatggt tggacatcaa ttattgcaca aaaaattttg aatcagtcct tttatctcgg 4500
 tgagtgttct ttctttgaaa aatcccagtc tattatgaat ctctatcatg atttagaatc 4560
 ttttgaccct tccaatcgtg aaggatcatg agctaggatt tatttcaata cactttttgg 4620
 aatgatattt tcaagagagc aggataatcc aataaatgct ggtttagact atggatattc 4680
 tctgattttg agtatgtttg cgcgtgaagt tgttaagtgt ggttgcatga cacaatttgg 4740
 cttaaagcat gctaatcaat ttaatcagtt caacctagca agcgatatta tggaaaccatt 4800
 tcgccaatc gttgatagga ttatttatga aatagggcag agtgattttg tcaaaatgaa 4860
 aagagaactc ttttctatgt tttcagagac atacagctac aacggtaaag aatgtatct 4920

ES 2 398 918 T3

ttcaaatatt gtcagcgatt acaccaaaaa agttattaag tcgctaaata gtgatgggaa 4980
 tgggaattccg gagtttagga tatgagttat cggtatatga gaatgatttt aatgtttgat 5040
 atgcctactg aaacagtaga agaacgtaag gcgtatcgta agtttagaaa gtttctgttg 5100
 agcgaagggt ttattatgca tcagttctct gtttatagta aattattgct caataatata 5160
 gctaataatg ccatgatagg tcggcttaaa gtgaataatc ctaagaaagg gagtataact 5220
 cttttgacag ttaccgagaa gcagtttgca aggatggttt atctacatgg tgaacataat 5280
 atgagtgttg ccaactctga tagtcgcttg gttttcctag gagattctta tgatcaagat 5340
 taatttcca attttagatg aaccattagt gttaagtaat gctacgattt taacgataga 5400
 agatgtttca gtttattctt cattggtgaa acatTTTTat caatatgacg tagatgaaca 5460
 tttgaaatta tttgatgata agcagaaaag tctgaaggca acggagttaa tgtagttac 5520
 agatatctta ggatacgatg tcaactcagc acctattcta aagttgatac atggtgactt 5580
 agaaaatcaa ttcaacgaaa agccagaagt gaaatcaatg gtagaaaaat tagcagctac 5640
 tattacagaa cttatcgcat ttgagtgtct agagaatgag cttgatttag aatacgatga 5700
 aattacgatt ttagaactca ttaaggcact gggagtcaaa attgagacac agagcgacac 5760
 tatctttgaa aaatgttttg aaattataca agtttaccat tatttaacga aaaagaatct 5820
 cttagttttt gttaatagcg gagcttatct taccaaagat gaagttataa aattatgtga 5880
 atacatcaat ttaatgcaaa agtcagtact ctttctagaa cctagaagac tctatgattt 5940
 accgcaatat gttattgata aggattattt cttgataggc gaaaatatgg tataa 5995

<210> 494

<211> 4134

5 <212> ADN

<213> Streptococcus agalactiae

<400> 494

ES 2 398 918 T3

atgaataagc catattcaat aggccttgac atcggtaacta attccgtcgg atggagcatt	60
attacagatg attataaagt acctgctaag aagatgagag ttttagggaa cactgataaa	120
gaatatatta agaagaatct cataggtgct ctgctttttg atggcgggaa tactgctgca	180
gatagacgct tgaagcgaac tgctcgtcgt cgttatacac gtcgtagaaa tcgtattcta	240
tatttacaag aaatttttgc agaggaaatg agtaaagttg atgatagttt ctttcatcga	300
ttagaggatt cttttctagt tgaggaagat aagagaggta gcaagtatcc tatctttgca	360
acaatgcagg aggagaaata ttatcatgaa aaatttccga caatctatca tttgagaaaa	420
gaattggctg acaagaaaga aaaagcagac cttcgtcttg tttatctggc tctagctcat	480
atcattaaat tcagagggca tttcctaatt gaggatgata gatttgatgt gaggaatacc	540
gatattcaaa aacaatatca agccttttta gaaatttttg atactacctt tgaaaataat	600
catttgttat ctcaaaatgt agatgtagaa gcaattctaa cagataagat tagcaagtct	660

ES 2 398 918 T3

gcgaagaagg atcgcacatt agcgcagtat cctaaccaaa aatctactgg tttttttgca 720
gaatTTTTga aattgattgt cggaaatcaa gctgacttca agaaacattt caatttggag 780
gataaaacac cgcttcaatt cgctaaggat agctacgatg aagatttaga aaatcttctt 840
ggacagattg gtgatgaatt tgcagactta ttctcagtag cgaaaaagct atatgatagt 900
gttcttttat ctggcattct tacagtaact gatctcagta ccaaggcgcc actttctgcc 960
tctatgattc agcgttatga tgaacatcat gaggacttaa agcatctaaa acaattcgta 1020
aaagcttcat tacctgaaaa ttatcgggaa gtatttgctg attcatcaaa agatggctac 1080
gctggctata ttgaaggcaa aactaatcaa gaagcttttt ataaatatct gttaaaattg 1140
ttgaccaaac aagaaggtag cgagtatttt cttgagaaaa ttaagaatga agattttttg 1200
agaaaacaga gaacctttga taatggctca atcccgcac aagtccattt gacagaattg 1260
agggtatta ttcgacgtca atcagaatac tatccattct tgaaagagaa tcaagatagg 1320
attgaaaaaa tccttacctt tagaattcct tattatgtcg ggccactagc acgtgagaag 1380
agtgattttg catggatgac tcgcaaaaca gatgacagta ttcgaccttg gaatTTTgaa 1440
gacttggttg ataaagaaaa atctgcgaa gcttttatcc atcgcacgac caacaatgac 1500
ctctatcttc cagaagaaaa agttttacca aagcatagtc ttatttatga aaaatttact 1560
gtttacaatg aattaacgaa ggtagattt ttggcagaag gctttaaaga ttttcaattt 1620
ttaaatagga agcaaaaaga aactatcttt aacagcttgt ttaaggaaaa acgtaaagta 1680
actgaaaagg atattattag ttttttgaat aaagttgatg gatatgaagg aattgcaatc 1740
aaaggaattg agaaacagtt taacgctagc ctttcaacct atcatgatct taaaaaata 1800
cttggaagg atttccttga taatacagat aacgagctta ttttgaaga tatcgtccaa 1860
actctaacct tatttgaaga tagagaaatg attaagaagt gtcttgacat ctataaagat 1920
ttttttacag agtcacagct taaaaagctc tatcgccgtc actatactgg ctggggacga 1980
ttgtctgcta agctaataaa tggcatccga aataaagaga atcaaaaaac aatcttggac 2040
tatcttattg atgatggaag tgcaaacgga aacttcatgc agttgataaa tgatgatgat 2100
ctatcattta aaccaattat tgacaaggca cgaactggta gtcattcgga taatctgaaa 2160
gaagttgtag gtgaacttgc tggtagccct gctattaaaa aagggttctt acaaagtttg 2220
aaaatagttg atgagctggt taaagtcacg ggctatgaac ctgaacaaat cgtggttgaa 2280
atggcacgtg agaaccaaac gacagcaaaa ggattaagtc gttcacgaca acgcttgaca 2340
acctgagag aatctcttgc taatttgaag agtaatatTT tggaagagaa aaagcctaag 2400
tatgtgaaag atcaagttga aaatcatcat ttatctgatg accgtctttt cctttactac 2460
ttacaaaacg gaagagatat gtatacaaaa aaggctctgg atattgataa ttttaagtcaa 2520
tatgatattg accacattat tcctcaagct ttcataaaag atgattctat tgataatcgt 2580

ES 2 398 918 T3

gttttggtat catctgctaa aaatcgtgga aaatcagatg atgttcctag cattgaaatt 2640
 gtaaaagctc gcaaaatggt ctggaaaaat ttactggatg ctaagttaat gagtcagcgt 2700
 aagtatgata atttgactaa ggcagagcgc ggaggcctaa ctccgatga taaggcaaga 2760
 tttatccaac gtcagttggt tgagactcga caaattacca agcatgtagc tcgtatcttg 2820
 gatgaacgct tcaataatga agttgataat ggtaaaaaga tttgcaaggt taaaattgta 2880
 accttgaagt caaatttggt ttcaaatttc cgaaaagaat ttggattcta taaaattcgt 2940
 gaagttaatg attatcacca tgcacacgat gcttatctta atgcagtagt tgccaaagct 3000
 attctaacca aatatccaca gttagagcca gagtttgtct acggaatgta tagacagaaa 3060
 aaactttcga aaatcgttca tgaggataag gaagaaaaat atagtgaagc aaccaggaaa 3120
 atgtttttct actccaactt gatgaatatg ttcaaaagag ttgtgaggtt agcagatggt 3180
 tctattgttg taagaccagt aatagaaact ggtagatata tgagaaaaac tgcattggat 3240
 aaaaagaaac actttgcgac agttagaaaa gtcttgtcat accctcagaa caatatcgtg 3300
 aagaagacag agattcagac aggtggtttc tctaaggaat caatcttggc gcatggtaac 3360
 tcagataagt tgattccaag aaaaacgaag gatatttatt tagatcctaa gaaatatgga 3420
 ggttttgata gtccgatagt agcttactct gtttttagttg tagctgatat caaaaaaggt 3480
 aaagcacaaa aactaaaaac agttacggaa ctttttaggaa ttaccatcat ggagaggtcc 3540
 agatttgaga aaaatccatc agctttcctt gaatcaaaag gttatttaaa tattagggac 3600
 gataaattaa tgattttacc gaagtatagt ctgttcgaat tagaaaatgg gcgtcgtcga 3660
 ttacttgcta gtgctggtga attacaaaaa ggtaacgagc tagccttacc aacacaattt 3720
 atgaagttct tataccttgc aagtcgttat aatgagtcaa aaggtaaacc agaggagatt 3780
 gagaagaaac aagaatttgt aaatcaacat gtctcttatt ttgatgacat ccttcaatta 3840
 attaatgatt tttcaaaacg agttattcta gcagatgcta atttagagaa aatcaataag 3900
 ctttaccagg ataataagga aaatatacca gtagatgaac ttgctaataa tattatcaat 3960
 ctatttactt ttaccagtct aggagctcca gcagctttta aattttttga taaaatagtt 4020
 gatagaaaac gctatacatc aactaaagaa gtacttaatt ctactctaatt ccatcaatct 4080
 attactggac tttatgaaac acgtattgat ttgggtaaat taggagaaga ttga 4134

<210> 495

<211> 870

5 <212> ADN

<213> Streptococcus agalactiae

<400> 495

ES 2 398 918 T3

atggcagggtt ggcgaactgt tgttgtaaata acacattcta agctctctta taaaaataat 60
 catctgattt ttaaagattc ttatcagacg gaaatgattc atctttcaga gattgatatt 120
 ctaatcatgg aaacgacaga tattgttttg tcgactatgc tgattaaacg tttggttgat 180
 gaaaatattt tagtcatatt ttgtgatgat aaacgcttgc caacagctat gttaatgccg 240
 tactatgcta gacatgattc gagtttacia ttatctaggc agatgtcatg gattgaggat 300
 gtcaaagcgg atgtttggac atcaattatt gcacaaaaaa ttttgaatca gtccttttat 360
 ctcggtgagt gttctttctt tgaaaaatcc cagtctatta tgaatctcta tcatgattta 420
 gaatcttttg acccttccaa tcgtgaaggc catgcagcta ggatttattt caatacactt 480
 tttggaaatg atttttcaag agagcaggat aatccaataa atgctggttt agactatgga 540
 tattctctga ttttgagtat gtttgcgcgt gaagttgtta agtgtggttg catgacacaa 600
 tttggcttaa agcatgctaa tcaatttaat cagttcaacc tagcaagcga tattatggaa 660
 ccatttcgcc caatcgttga taggattatt tatgaaaata ggcagagtga ttttgtcaaa 720
 atgaaaagag aactcttttc tatgttttca gagacataca gctacaacgg taaagaaatg 780
 tatctttcaa atattgtcag cgattacacc aaaaaagtta ttaagtcgct aaatagtgat 840
 gggaatggaa ttccggagtt taggatatga 870

<210> 496
 5 <211> 342
 <212> ADN
 <213> Streptococcus agalactiae

<400> 496

10 atgagttatc ggtatatgag aatgatttta atgtttgata tgcctactga aacagtagaa 60
 gaacgtaagg cgtatcgtaa gtttagaaag tttctgttga gcgaaggttt tattatgcat 120
 cagttctctg tttatagtaa attattgctc aataatacag ctaataatgc catgataggt 180
 cggcttaaag tgaataatcc taagaaaggg agtataactc ttttgacagt taccgagaag 240
 cagtttgcaa ggatggttta tctacatggg gaacataata tgagtgttgc caactctgat 300
 agtcgcttgg ttttcttagg agattcttat gatcaagatt aa 342

<210> 497
 15 <211> 666
 <212> ADN
 <213> Streptococcus agalactiae

<400> 497

ES 2 398 918 T3

```

atgatcaaga ttaattttcc aatttttagat gaaccattag tgtaagtaa tgctacgatt      60
ttaacgatag aagatgtttc agttttattct tcattggtga aacattttta tcaatatgac      120
gtagatgaac atttgaaatt atttgatgat aagcagaaaa gtctgaaggc aacggagtta      180
atggttagtta cagatatctt aggatacgat gtcaactcag cacctattct aaagttgata      240
catggtgact tagaaaatca attcaacgaa aagccagaag tgaaatcaat ggtagaaaaa      300
ttagcagcta ctattacaga acttatcgca tttgagtgtc tagagaatga gcttgattta      360
gaatacgatg aaattacgat tttagaactc attaaggcac tgggagtcaa aattgagaca      420
cagagcgaca ctatctttga aaaatgtttt gaaattatac aagtttacca ttatttaacg      480
aaaaagaatc tcttagtttt tgtaaatagc ggagcttatac ttaccaaaga tgaagttata      540
aaattatgtg aatacatcaa tttaatgcaa aagtcagtac tctttctaga acctagaaga      600
ctctatgatt taccgcaata tgttattgat aaggattatt tcttgatagg cgaaaatag      660
gtataa                                          666

```

5 <210> 498
 <211> 6580
 <212> ADN
 <213> Streptococcus mutans

10 <400> 498

ES 2 398 918 T3

atgaaaaaac cttactctat tggacttgat attggaacca attctgttgg ttgggctggt 60
 gtgacagatg actacaaagt tcctgctaag aagatgaagg ttctgggaaa tacagataaa 120
 agtcatatcg agaaaaattt gcttggcgct ttattatttg atagcgggaa tactgcagaa 180
 gacagacggt taaagagaac tgctcgccgt cgttacacac gtcgcagaaa tcgtatttta 240
 tatttgcaag agattttttc agaagaaatg ggcaaggtag atgatagttt ctttcatcgt 300
 ttagaggatt cttttcttgt tactgaggat aaacgaggag agcgccatcc catttttggg 360
 aatcttgaag aagaagttaa gtatcatgaa aattttccaa ccatttatca tttgcggcaa 420
 tatcttgcgg ataatccaga aaaagttgat ttgcgtttag tttatttggc tttggcacat 480
 ataattaagt ttagagggtca ttttttaatt gaaggaaagt ttgatacacg caataatgat 540
 gtacaaaagac tgtttcaaga atttttagca gtctatgata atacttttga gaatagtctg 600
 cttcaggagc aaaatgttca agttgaagaa attctgactg ataaaatcag taaatctgct 660
 aagaaaagata gagttttgaa actttttcct aatgaaaagt ctaatggccg ctttgcagaa 720
 tttctaaaac taattgttgg taatcaagct gattttaaaa agcattttga attagaagag 780
 aaagcaccat tgcaattttc taaagatact tatgaagaag agttagaagt actattagct 840
 caaattggag ataattacgc agagctcttt ttatcagcaa agaaactgta tgatagtatc 900
 cttttatcag ggattttaac agttactgat gttggtacca aagcgctttt atctgcttcg 960
 atgattcagc gatataatga acatcagatg gatttagctc agcttaaaca attcattcgt 1020
 cagaaattat cagataaata taacgaagtt ttttctgatg tttcaaaaga cggctatgcg 1080
 ggttatattg atgggaaaac aatcaagaa gctttttata aataccttaa aggtctatta 1140
 aataagattg agggaagtgg ctatttcctt gataaaattg agcgtgaaga ttttctaaga 1200
 aagcaacgta cctttgacaa tggctctatt ccacatcaga ttcacttca agaaatgcgt 1260
 gctatcattc gtagacaggc tgaattttat cggtttttag cagacaatca agataggatt 1320
 gagaaattat tgactttccg tattccctac tatgttggtc cattagcgcg cggaaaaagt 1380

ES 2 398 918 T3

gattttgctt ggttaagtcg gaaatcggct gataaaatta caccatggaa ttttgatgaa 1440
atcgttgata aagaatcctc tgcagaagct tttatcaatc gtatgacaaa ttatgatttg 1500
tacttgccaa atcaaaaagt tcttcctaaa catagtttat tatacgaaaa atttactggt 1560
tacaatgaat taacaaaggt taaatataaa acagagcaag gaaaaacagc atttttgat 1620
gccaatatga agcaagaaat ctttgatggc gtatttaagg tttatcgaaa agtaactaaa 1680
gataaattaa tggatttcct tgaaaaagaa tttgatgaat ttcgtattgt tgatttaaca 1740
ggctctggata aagaaaataa agtatttaac gcttcttatg gaacttatca tgatttgtgt 1800
aaaattttag ataaagattt tctcgataat tcaaagaatg aaaagatttt agaagatatt 1860
gtgttgacct taacgttatt tgaagataga gaaatgatta gaaaacgtct agaaaattac 1920
agtgatttat tgaccaaaaga acaagtgaaa aagctggaaa gacgtcatta tactggttgg 1980
ggaagattat cagctgagtt aattcatggt attcgcaata aagaaagcag aaaaacaatt 2040
cttgattatc tcattgatga tggcaatagc aatcggaact ttatgcaact gattaacgat 2100
gatgctcttt ctttcaaaga agagattgct aaggcacaag ttattggaga aacagacaat 2160
ctaaatcaag ttgttagtga tattgctggc agccctgcta ttaaaaaagg aattttaca 2220
agcttgaaga ttgttgatga gcttgtcaaa attatgggac atcaacctga aaatatcgtc 2280
gtggagatgg cgcgtgaaaa ccagtttacc aatcagggac gacgaaattc acagcaacgt 2340
ttgaaagggt tgacagattc tattaagaa tttggaagtc aaattcttaa agaacatccg 2400
gttgagaatt cacagttaca aaatgataga ttgtttctat attatttaca aaacggcaga 2460
gatatgtata ctggagaaga attggatatt gattatctaa gccagtatga tatagacat 2520
attatcccgc aagcttttat aaaggataat tctattgata atagagtatt gactagctca 2580
aaggaaaatc gtggaaaatc ggatgatgta ccaagtaaag atgttgttcg taaaatgaaa 2640
tcctattgga gtaagctact ttcggcaaag cttattacac aacgtaaatt tgataatttg 2700
acaaaagctg aacgaggtgg attgaccgac gatgataaag ctggattcat caagcgtcaa 2760
ttagtagaaa cacgacaaat taccaacat gtagcacgta ttctggacga acgatttaat 2820
acagaaacag atgaaaaca caagaaaatt cgtcaagtaa aaattgtgac cttgaaatca 2880
aatcttgttt ccaatttccg taaagagttt gaactctaca aagtgcgtga aattaatgac 2940
tatcatcatg cacatgatgc ctatctcaat gctgtaattg gaaaggcttt actaggtggt 3000
taccacaat tggaacctga atttgtttat ggtgattatc ctcattttca tggacataaa 3060
gaaaataaag caactgctaa gaaatttttc tattcaaata ttatgaactt ctttaaaaaa 3120
gatgatgtcc gtactgataa aaatggtgaa attatctgga aaaaagatga gcataatttct 3180
aatattaanaa aagtgctttc ttatccaca gttaatattg ttaagaaagt agaggagcaa 3240
acgggaggat tttctaaaga atctatcttg ccgaaaggta attctgacaa gcttattcct 3300

ES 2 398 918 T3

cgaaaaacga agaaatthta ttgggatacc aagaaatatg gaggatttga tagcccgatt 3360
 gttgcttatt ctatthtagt tattgctgat attgaaaaag gtaaactaa aaaattgaaa 3420
 acagtcaaag ccttagttgg tgtcactatt atggaaaaga tgactthtga aagggatcca 3480
 gttgctthtc ttgagcgaag aggctatcga aatgthcaag aagaaaatat tataaagtta 3540
 ccaaaatata gthtattta actagaaaac ggacgaaaaa ggctattggc aagtgctagg 3600
 gaacttcaaa agggaaatga aatcgthttg ccaaatcatt taggaacctt gctthtacc 3660
 gctaaaaata ttcataaagt tgatgaacca aagcatttgg actatgthga taaacataaa 3720
 gatgaattta aggagttgct agatgthtg tcaaaactth ctaaaaata tactthtagca 3780
 gaaggaaatt tagaaaaat caaagaatta tatgcacaaa ataatgthga agatcttaaa 3840
 gaattagcaa gthcatttht caacttatta acatthactg ctataggagc accggctact 3900
 thtaattct ttgataaaaa tattgatcga aaacgatata ctthcaactac tgaaattctc 3960
 aacgctaccc tcatccacca atccatcacc ggtctthtatg aaacgcggat tgatctcaat 4020
 aagthaggag gagactaatg ggctggcggg cagthgthgt taatacgcatt tccaagthgt 4080
 ctataagaa caaccactg atththtaag atgcttattca gacagagatg atthcatctg 4140
 ctgagattga catcttatta ctthgagacaa cagatattgt thtgthcaact atgctaatca 4200
 aacgcttggg tgatgagaat atththgtca thththtgthga tgacaaactg ctgccaacag 4260
 ccatgctcat gccttactat gcgcgtcacg atthccagctt gcagctgagth catcagattth 4320
 ctthggacaga agaagthgaaa tgcgatgthct ggacaacaat catcgctcaa aagattthtga 4380
 atcagthcatg thattthggg gaatgththth atththgaaaa atctcagthca atthattgatt 4440
 tatatcatga ctthagagct ththgacctt gthaatcgaga aggacattct gcgcggattth 4500
 atthcaatac ctthattthgga aatgthththth ccagagaaca agataatgatt atthaatgcag 4560
 gthctthgacta thgthtatacgt ctgctgthta gthattgththgc gcgthgagthg gththgattctg 4620
 gctgthgac acaattthgg ctcaagcagth ccaaccaatt caatcagthth aactthgcca 4680
 gthgattatth ggagcctthth cgtccaattg thgacctgatt thththattgaa aatcgaaaata 4740
 actctththt taaaataaaaa cgtgagctatt thcagcagthth thcagacacc thctthtata 4800
 ataataagga gattgattthg acaaatattg thcagcgatta taccaaaaag gthaatcaagg 4860
 cgctgaataa tgatgggaaa ggagthctctg agththaggatt atgagthtacc gatattgthg 4920
 aatgattthta atgththgata thgccaacaga tactgctgag gaacgcaaag ctthattcgthaa 4980
 atthcggaag thththactg gcaagththth catcagthcat cagthththcag thattcagcaa 5040
 gctgctththg aataactctg ccaatacagc catgattgct cgctthgagga agaataatcc 5100
 aaagaagggc aatattcacct thgthgacctg gactgaaaag cagththgccc gthattgattta 5160
 cctgaaatggg gagcgtgata ctgacattgc thaatcgatt thcagactgg thththctagg 5220

ES 2 398 918 T3

ggaggctttt cctgatgaaa cttaattttc ctatattgga tgaaccaata actcttgaaa 5280
aatctacgat tttggtatta gaagatgtgc aagtttttgc tcaaattggtg agaaatcttt 5340
atcaatatga tgaagatagt gaacttaaat tttttaatag aaaatttaag agtctgaaac 5400
catctgagtt aatgcttgtg acagatattt taggttatga tgtcaatgcc ccgtccttgc 5460
tgaagttggt tcacgctgat ttagaaaatc agtttaatga aaaaccagag gttaagtcta 5520
tggttgaaaa actggcaaat accattacgg aattaattgc ttatgaatgt ttagaaaatg 5580
aattggactt agaatatgat gagattacta ttttagagtt aatcaaagct ttaggcgtca 5640
aaattgaaac acaaagtgat accatTTTTg aaaaaatggt tgaagtcctt caagtttata 5700
agtatctaaa taaaagaag cttctcgttt ttatcaatac tttatcctat tttaaaagag 5760
aagaaatcgc gcaaattcta gaatatattc acttatccga tatggttggt ttatttcttg 5820
aaccctgtaa aattgatggt ttgctcaat atattttaga tgaagattat ttcttgataa 5880
cagaaagcaa caactaaata cgaataataa gatagtttct aaatcagggg ctgtctttta 5940
ttatggattg acaaatgcgt ataatgcgta taaaataaaa agagaaatgt tatttgccat 6000
taacagggaa agaattagct aaattagcga taacaatgg atgggaagaa gttcgggtga 6060
gaggaagtca tcatcatttc aagaaagatg gagtatctta tattgtgacg attcctattc 6120
atggaaataa agtgcttaaa attggtcttg aaaagaaact cttaagggat ttaaatttat 6180
tatgatagag gaggaagtcg tcatgttaaa atcatatcct gtaatttttc ataaggaaga 6240
ggaagggat tgggttgaat ttctgaatt tggcgggtg acgcaagggg aagatttgga 6300
agaagccatg aagaacgctc gtcagatggt agaaagtgtg ttggcatctt atcttgatga 6360
agggttggtt ctaccattt caagcगतat tcagaaaata tctggtgaag atggttttgc 6420
gaccatgatt caagctgatc ctagtcctta tctcaaaaat aacaaagcta ttcggaaaaa 6480
tgttaccgtg cctgagtggt tgatacgatt agcagaccgt gaccgagtaa attattctga 6540
agtattaaca aaggctttgg aaaagaaact acaattataa 6580

<210> 499

<211> 4038

5 <212> ADN

<213> Streptococcus mutans

<400> 499

atgaaaaaac cttactctat tggacttgat attggaacca attctggttg tgggctggt 60
gtgacagatg actacaaagt tcctgctaag aagatgaagg ttctgggaaa tacagataaa 120
agtcatatcg agaaaaattt gcttggcgtt ttattatttg atagcgggaa tactgcagaa 180
gacagacggt taaagagaac tgctcgccgt cgttacacac gtcgcagaaa tcgtatttta 240
10 tatttgcaag agattttttc agaagaaatg ggcaaggtag atgatagttt ctttcatcgt 300

ES 2 398 918 T3

ttagaggatt cttttcttgt tactgaggat aaacgaggag agcgccatcc ctttttggg 360
 aatcttgaag aagaagttaa gtatcatgaa aatthtccaa ccatttatca tttgcygcaa 420
 tatcttgcyg ataatccaga aaaagttgat ttgcygtttag tttatttggc tttggcacat 480
 ataattaagt ttagagggtca ttttttaatt gaaggaaagt ttgatacacg caataatgat 540
 gtacaaagac tgtttcaaga attttttagca gtctatgata atacttttga gaatagtctg 600
 cttcaggagc aaaatgttca agttgaagaa attctgactg ataaaatcag taaatctgct 660
 aagaaagata gagttttgaa actttttcct aatgaaaagt ctaatggccg ctttgcagaa 720
 tttctaaaaac taattgttgg taatcaagct gatttttaaaa agcattttga attagaagag 780
 aaagcaccat tgcaatthtc taaagatact tatgaagaag agttagaagt actattagct 840
 caaattggag ataattacgc agagctcttt ttatcagcaa agaaactgta tgatagtatc 900
 cttttatcag ggatttttaac agttactgat gttggtacca aagcyctttt atctgcttgc 960
 atgattcagc gatataatga acatcagatg gatttagctc agcttaaaca attcattcgt 1020
 cagaaattat cagataaata taacgaagtt ttttctgatg tttcaaaaga cggctatgcy 1080
 ggttatattg atgggaaaac aatcaagaa gctttttata aataccttaa aggtctatta 1140
 aataagattg agggaaagtgg ctatttcctt gataaaattg agcytgaaga ttttctaaga 1200
 aagcaacgta ctttgacaa tggctctatt ccacatcaga ttcatcttca agaaatgcgt 1260
 gctatcattc gtagacaggc tgaatthtat ccgtttttag cagacaatca agataggatt 1320
 gagaaattat tgactttccg tattccctac tatgttggtc cattagcygc cggaaaaagt 1380
 gattttgctt ggtaagtgc gaaatcggct gataaaatta caccatggaa ttttgatgaa 1440
 atcgttgata aagaatcctc tgcagaagct tttatcaatc gtatgacaaa ttatgatttg 1500
 tacttgccaa atcaaaaagt tcttctaaa catagtttat tatacgaaaa atttactgtt 1560
 tacaatgaat taacaaaggt taaatataaa acagagcaag gaaaaacagc attttttgat 1620
 gccaatatga agcaagaaat ctttgatggc gtatttaagg tttatcgaaa agtaactaaa 1680
 gataaattaa tggatttcct tgaaaaagaa tttgatgaat ttcgtattgt tgatttaaca 1740
 ggtctggata aagaaaataa agtatttaac gcttcttatg gaacttatca tgatttgtgt 1800
 aaaatthtag ataaagattt tctcgataat tcaaagaatg aaaagatttt agaagatatt 1860
 gtgttgacct taacgttatt tgaagataga gaaatgatta gaaaacgtct agaaaattac 1920
 agtgatttat tgaccaaaaga acaagtgaaa aagctggaaa gacgtcatta tactggttgg 1980
 ggaagattat cagctgagtt aattcatggt attcgcaata aagaaagcag aaaaacaatt 2040
 cttgattatc tcattgatga tggcaatagc aatcggaact ttatgcaact gattaacgat 2100
 gatgctcttt ctttcaaaaga agagattgct aaggcacaag ttattggaga aacagacaat 2160
 ctaaatcaag ttgttagtga tattgctggc agccctgcta ttaaaaaagg aatthtcaa 2220

ES 2 398 918 T3

agcttgaaga ttgttgatga gcttgtcaaa attatgggac atcaacctga aaatatcgtc 2280
 gtggagatgg cgcgtgaaaa ccagtttacc aatcagggac gacgaaattc acagcaacgt 2340
 ttgaaagggt tgacagattc tattaaagaa tttggaagtc aaattcttaa agaacatccg 2400
 gttgagaatt cacagttaca aaatgataga ttgtttctat attatttaca aaacggcaga 2460
 gatatgtata ctggagaaga attggatatt gattatctaa gccagtatga tatagacat 2520
 attatcccgc aagcttttat aaaggataat tctattgata atagagtatt gactagctca 2580
 aaggaaaatc gtggaaaatc ggatgatgta ccaagtaaag atgttggtcg taaaatgaaa 2640
 tcctattgga gtaagctact ttcggcaaaag cttattacac aacgtaaatt tgataatttg 2700
 acaaaagctg aacgaggtgg attgaccgac gatgataaag ctggattcat caagcgtcaa 2760
 ttagtagaaa cacgacaaat taccaaacat gtagcacgta ttctggacga acgatttaat 2820
 acagaaacag atgaaaacaa caagaaaatt cgtcaagtaa aaattgtgac cttgaaatca 2880
 aatcttgttt ccaatttccg taaagagttt gaactctaca aagtgcgtga aattaatgac 2940
 tatcatcatg cacatgatgc ctatctcaat gctgtaattg gaaaggcttt actaggtggt 3000
 taccacaat tggaacctga atttgtttat ggtgattatc ctcattttca tggacataaa 3060
 gaaaataaag caactgctaa gaaattttcc tattcaaata ttatgaactt ctttaaaaaa 3120
 gatgatgtcc gtactgataa aaatggtgaa attatctgga aaaaagatga gcatattttct 3180
 aatattaata aagtgccttc ttatccacaa gttaatattg ttaagaaagt agaggagcaa 3240
 acgggaggat tttctaaaga atctatcttg ccgaaaggta attctgacaa gcttattcct 3300
 cgaaaaacga agaaatttta ttgggatacc aagaaatatg gaggatttga tagccccgatt 3360
 gttgcttatt ctattttagt tattgctgat attgaaaaag gtaaatctaa aaaattgaaa 3420
 acagtcaaag ccttagttgg tgtcactatt atggaaaaga tgacttttga aagggatcca 3480
 gttgcttttc ttgagcgaag aggctatcga aatgttcaag aagaaaatat tataaagtta 3540
 ccaaaatata gtttatttaa actagaaaac ggacgaaaaa ggctattggc aagtgcctagg 3600
 gaacttcaaa agggaaatga aatcgttttg ccaaatcatt taggaacctt gctttatcac 3660
 gctaaaaata ttcataaagt tgatgaacca aagcatttgg actatgttga taaacataaa 3720
 gatgaattta aggagttgct agatgttgtg tcaaactttt ctaaaaaata tacttttagca 3780
 gaaggaaatt tagaaaaaat caaagaatta tatgcacaaa ataatggtga agatcttaaa 3840
 gaattagcaa gttcatttat caacttatta acatttactg ctataggagc accggctact 3900
 tttaaattct ttgataaaaa tattgatcga aaacgatata cttcaactac tgaaattctc 3960
 aacgctaccc tcatccacca atccatcacc ggtctttatg aaacgcggat tgatctcaat 4020
 aagttaggag gagactaa 4038

<210> 500

<211> 867

ES 2 398 918 T3

<212> ADN
<213> Streptococcus mutans

<400> 500

5
atgggctggc ggacagtggc tgttaatacg cattccaagt tgtcttataa gaacaaccac 60
ttgattttta aagatgctta tcagacagag atgattcatc tgtctgagat tgacatctta 120
ttacttgaga caacagatat tgttttgcact actatgctaa tcaaacgctt ggttgatgag 180
aatatthttgg tcattthtttg tgatgacaaa cgtctgccaa cagccatgct catgccttac 240
tatgctgctc acgattccag cttgcagctg agtcatcaga tttcttggac agaagaagtg 300
aatgctgatg tctggacaac aatcatcgtc caaaagattt tgaatcagtc atgttatttg 360
ggagaatgth tttattthtga aaaatctcag tcaattatgg atthtatca tgacttagag 420
cctthttgacc ctagtaatcg agaaggacat tctgctgctg tttatttcaa taccttattt 480
ggaaatgtht tttccagaga acaagataat gatattaatg caggtcttga ctatggtht 540
acgtgctgtg taagtatgth tgcctgctgaa gtggttgtat ctggctgtat gacacaattt 600
ggtctcaagc atgccaacca attcaatcag thtaacttht ccagtgatat tatggagctt 660
thtctgctcaa thgttgaccg tathgttht gaaaatctgaa ataactctth tathaaaata 720
aaactgagc thttcagcat gthttcagac acctatctth athaataata ggagatgtat 780
thgacaaata thgtcagcga thataccaaa aaggthaatca aggcctgaa thatgatggg 840
aaaggagtht ctgagthttag gatatga 867

<210> 501
<211> 330
10 <212> ADN
<213> Streptococcus mutans

<400> 501

atgctgaatga ththtaatgth tgatatgcca acagatactg ctgaggaacg caaagcttht 60
cgtaaattht ggaaatthtt actgagctgaa gthttcatca tgcatcagth thcagthatac 120
agcaagctgc thttgaataa ctctgccaat acagccatga thgcccgtth gaaggagaat 180
aatccaaaga agggcaatath caccttgtht accgtgactg aaaagcagth thgcccgtatg 240
atthacctga atggtgagctg tgatactagc atthctaat cggattcacg actggtctth 300
15 ctaggggagg cthttctgta tgaaactth 330

<210> 502
<211> 573
20 <212> ADN
<213> Streptococcus mutans

ES 2 398 918 T3

<400> 502

atggtgagaa atctttatca atatgatgaa gatagtgaac ttaaattttt taatagaaaa 60
 tttaaagagtc tgaaaccatc tgagttaatg cttgtgacag atatttttagg ttatgatgtc 120
 5 aatgccccgt ccttgctgaa gttggttcac gctgatttag aaaatcagtt taatgaaaaa 180
 ccagagggta agtctatggt tgaaaaactg gcaaatacca ttacggaatt aattgcttat 240
 gaatgttttag aaaatgaatt ggacttagaa tatgatgaga ttactatttt agagttaatc 300
 aaagcttttag gcgtcaaaat tgaaacacaa agtgatacca tttttgaaaa aatgtttgaa 360
 gtccttcaag tttataagta tctaaataaa aagaagcttc tcgtttttat caatacttta 420
 tcctatttta aaagagaaga aatcgcgcaa attctagaat atattcactt atccgatatg 480
 gttgttttat ttcttgaacc ccgtaaaatt gatggttttg ctcaatatat tttagatgaa 540
 gattatttct tgataacaga aagcaacaac taa 573

<210> 503

<211> 378

10 <212> ADN

<213> Streptococcus mutans

<400> 503

atgttaaaat catatcctgt aatttttcat aaggaagagg aagggatttg ggttgaattt 60
 cctgaatttg gcggtggtac gcaaggggaa gatttgaag aagccatgaa gaacgctcgt 120
 cagatgtag aaagtgtggt ggcattctat cttgatgaag ggttggttct accatttca 180
 agcgatattc agaaaatatc tgttgaagat ggttttgca ccatgattca agctgacct 240
 agtccttatc tcaaaaataa caaagctatt cggaaaaatg ttaccgtgcc tgagtggttg 300
 atacgattag cagaccgtga ccgagtaaat tattctgaag tattaacaaa ggctttggaa 360
 15 aagaaactac aattataa 378

<210> 504

<211> 5966

<212> ADN

20 <213> Streptococcus pyogenes

<400> 504

ES 2 398 918 T3

atggataaga	aataactcaat	aggcttagat	atcggcacia	atagcgtcgg	atgggcggg	60
atcactgatg	aatataaggt	tccgtctaaa	aagttcaagg	ttctgggaaa	tacagaccgc	120
cacagtatca	aaaaaaatct	tataggggct	cttttatttg	acagtggaga	gacagcggaa	180
gcgactcgtc	tcaaacggac	agctcgtaga	aggtatacac	gtcggaagaa	tcgtatttgt	240
tatctacagg	agatTTTTTc	aatgagatg	gcgaaagtag	atgatagttt	ctttcatcga	300
cttgaagagt	cttttttgg	ggaagaagac	aagaagcatg	aacgtcatcc	tatttttgg	360
aatatagtag	atgaagttgc	ttatcatgag	aaatatccaa	ctatctatca	tctgcaaaaa	420
aaattggtag	attctactga	taaagcggat	ttgcgcttaa	tctatttggc	cttagcgc	480
atgattaagt	ttcgtgggtca	ttttttgatt	gagggagatt	taaatacctga	taatagtgat	540
gtggacaaac	tatttatcca	gttggtaaaa	acctacaatc	aattatttga	agaaaaccct	600

ES 2 398 918 T3

attaacgcaa gtggagtaga tgctaaagcg attccttctg cacgattgag taaatcaaga 660
 cgattagaaa atctcattgc tcagctcccc ggtgagaaga aaaatggctt atttgggaat 720
 ctccattgctt tgtcattggg tttgaccctt aattttaaat caaattttga tttggcagaa 780
 gatgctaaat tacagctttc aaaagatact tacgatgatg atttagataa tttattggcg 840
 caaattggag atcaatatgc tgatttgttt ttggcagcta agaatttatc agatgctatt 900
 ttactttcag atatcctaag agtaaatact gaaataacta aggctcccct atcagcttca 960
 atgattaaac gctacgatga acatcatcaa gacttgactc ttttaaagc tttagtctga 1020
 caacaacttc cagaaaagta taaagaaatc ttttttgatc aatcaaaaaa cggatagca 1080
 ggttatattg atgggggagc tagccaagaa gaattttata aatttatcaa accaatttta 1140
 gaaaaaatgg atggtactga ggaattattg gtgaaactaa atcgtgaaga tttgctgcgc 1200
 aagcaacgga cctttgacaa cggctctatt ccccatcaaa ttcacttggg tgagctgcat 1260
 gctattttga gaagacaaga agacttttat ccatttttaa aagacaatcg tgagaagatt 1320
 gaaaaaatct tgacttttgc aattccttat tatggtggtc cattggcgcg tggcaatagt 1380
 cgttttgcat ggatgactcg gaagtctgaa gaaacaatta ccccatggaa ttttgaagaa 1440
 gttgtcgata aagggtgctc agctcaatca tttattgaac gcatgacaaa ctttgataaa 1500
 aatcttccaa atgaaaaagt actaccaaaa catagtttgc tttatgagta ttttacggtt 1560
 tataacgaat tgacaaaggc caaatatggt actgaaggaa tgcgaaaacc agcatttctt 1620
 tcagggtgaa agaagaaagc cattgttgat ttactcttca aaacaaatcg aaaagtaacc 1680
 gttaagcaat taaaagaaga ttatttcaaa aaaatagaat gttttgatag tgttgaaatt 1740
 tcaggagtgg aagatagatt taatgcttca ttaggtacct accatgattt gctaaaaatt 1800
 attaaagata aagatttttt ggataatgaa gaaaatgaag atatcttaga ggatattggt 1860
 ttaacattga ccttatttga agataggag atgattgagg aaagacttaa aacatagct 1920
 cacctctttg atgataaggc gatgaaacag cttaaacgct gccggttatac tggttgggga 1980
 cgtttgtctc gaaaattgat taatggtatt aggataagc aatctggcaa aacaatatta 2040
 gattttttga aatcagatgg ttttgccaat cgcaatttta tgcagctgat ccatgatgat 2100
 agtttgacat ttaaagaaga cattcaaaaa gcacaagtgt ctggacaagg cgatagttta 2160
 catgaacata ttgcaaattt agctggtagc cctgctatta aaaaaggat tttacagact 2220
 gtaaaagtgg ttgatgaatt ggtcaaagta atggggcggc ataagccaga aatatcggt 2280
 attgaaatgg cacgtgaaaa tcagacaact caaaagggcc agaaaaattc gcgagagcgt 2340
 atgaaacgaa tcgaagaagg tatcaaagaa ttaggaagtc agattcttaa agagcatcct 2400
 gttgaaaata ctcaattgca aatgaaaag ctctatctct attatctcca aatggaaga 2460
 cacatgatg tggaccaaga attagatatt aatcgtttaa gtgattatga tgtcgatcac 2520

ES 2 398 918 T3

attgttccac aaagtttccct taaagacgat tcaatagaca ataaggtcctt aacgcgttct 2580
 gataaaaatc gtggtaaate ggataacggt ccaagtgaag aagtagtcaa aaagatgaaa 2640
 aactattgga gacaacttct aaacgccaag ttaatcactc aacgtaagtt tgataattta 2700
 acgaaagctg aacgtggagg tttgagtga cttgataaag ctggttttat caaacgccaa 2760
 ttggttgaaa ctgcaccaat cactaagcat gtggcacaaa ttttgatag tcgcatgaat 2820
 actaaatagc atgaaaatga taaacttatt cgagaggtta aagtgattac cttaaaatct 2880
 aaattagttt ctgacttccg aaaagatttc caattctata aagtacgtga gattaacaat 2940
 taccatcatg cccatgatgc gtatctaaat gccgtcgttg gaactgcttt gattaagaaa 3000
 tatccaaaac ttgaatcgga gtttgtctat ggtgattata aagtttatga tgttcgtaaa 3060
 atgattgcta agtctgagca agaaataggc aaagcaaccg caaaatattt cttttactct 3120
 aatatcatga acttcttcaa aacagaaatt acacttgcaa atggagagat tcgcaaaccg 3180
 cctctaactg aaactaatgg ggaaactgga gaaattgtct gggataaagg gcgagatttt 3240
 gccacagtgc gcaaagtatt gtccatgcc caagtcaata ttgtcaagaa aacagaagta 3300
 cagacaggcg gattctccaa ggagtcaatt ttaccaaaaa gaaattcgga caagcttatt 3360
 gctcgtaaaa aagactggga tccaaaaaaa tatggtgggt ttgatagtcc aacggtagct 3420
 tattcagtcc tagtggttgc taaggtggaa aaagggaaat cgaagaagtt aaaatccggt 3480
 aaagagttac tagggatcac aattatggaa agaagttcct ttgaaaaaaa tccgattgac 3540
 tttttagaag ctaaaggata taaggaagtt aaaaaagact taatcattaa actacctaaa 3600
 tatagtcttt ttgagttaga aaacggtcgt aaacggatgc tggctagtgc cggagaatta 3660
 caaaaaggaa atgagctggc tctgccaaagc aaatatgtga attttttata tttagctagt 3720
 cattatgaaa agttgaaggg tagtccagaa gataacgaac aaaaacaatt gtttgtggag 3780
 cagcataagc attatttaga tgagattatt gagcaaatca gtgaattttc taagcgtggt 3840
 attttagcag atgccaattt agataaagtt cttagtgcac ataacaaca tagagacaaa 3900
 ccaatacgtg aacaagcaga aatattatt catttattta cgttgacgaa tcttgagct 3960
 cccgctgctt ttaaataattt tgatacaaca attgatcgta aacgatatac gtctacaaaa 4020
 gaagttttag atgccactct tatccatcaa tccatcactg gtctttatga aacacgcatt 4080
 gatttgagtc agctaggagg tgactgatgg ctggttggcg tactgttgtg gtaaatacc 4140
 actcgaaatt atcctataag aataatcatc tgatttttaa ggatgcctat aaaacggagc 4200
 tgatccattt atcagaaatt gatattttgt tattagaaac gaccgatatt gtcttgccta 4260
 ctatgctggt aaaacggcta gtggatgaga atgtccttgt catattctgt gatgataaac 4320
 gattaccaac agctatgctg atgccttttt atggtcgtca tgattcgagt ttacagcttg 4380
 ggaacaaaat gtccctgtca gaaacagtca aatcgcaggt ttggacgacq attattgctc 4440

ES 2 398 918 T3

aaaagat ttt gaatcaatct tgctatctag gagcatgctc ctat ttttgaa aaatcccaat 4500
 ctattatgga tttatatcat ggtttggaaa at tttgatcc gagtaatcga gaagggcatg 4560
 cagcgagaat ttat ttttaac acact tttttg ggaacgattt ctcaagagat ttggagcatc 4620
 caatcaatgc aggtctggat tatggttata ctttattatt gagtatgttt gcgctggaag 4680
 tggttgtgtc tggatgtatg actcagtttg ggcttaaaca cgctaactcag tttaatcagt 4740
 tcaat ttttg tagcgatatt atggaacat ttaggccttt agtggataag attgtttatg 4800
 aaaatcgaaa tcagcctttt cccaaaataa agagagagtt atttactttg ttttcagata 4860
 cattttcata taatggtaaa gagatgtatc tcacgaatat tattagcgat tataactaaa 4920
 aagttgtcaa agctctgaat aatgaagggg aaggagttcc tgaatttagg atatgagtta 4980
 tagatatatg agaatgatac ttatg tttga tatgccgacg gacaccgctg aggaacgaaa 5040
 agcctatcga aaatttcgga aatttttact tagtgaaggg tttatcatgc atcaat tttc 5100
 tatttatagt aagttgctgt tgaataatac agctaacaat gccatgattg gtcggctgag 5160
 ggagcataat cctaataaag gaaatattac attactaacg gtcacggaaa aacagtttgc 5220
 acgaatgatt tattttacatg gtgaaagaaa taattgtatt gcaaactccg atgaaagact 5280
 tgtat tttctt ggggaggctt ttgatgaatc ttaatt tttc ctactagat gaaccgattc 5340
 cattaagagg cggtaacaatt cttgtgctcg aagatgtctg tgtat tttca aaaatagttc 5400
 aatattgtta ccaatatgag gaagattctg aacttaaatt ttttgatcac aagatgaaaa 5460
 caatcaaaga atcagaaatc atgcttgtaa cagatatttt aggatttgat gttaaactcct 5520
 caaccat tttt aaaattgatt catgcagatt tagaatctca atttaatgag aaaccggaag 5580
 tgaatc gat gattgacaaa ttggttgcta cgattacaga actgattgtc tttgaaatgct 5640
 tagaaaatga attagattta gagtatgatg aaatcacaat cctggaattg attaagttcct 5700
 taggagtaaa agtagaaacg caaagtgata ctat tttttga aaaatgtcta gagatacttc 5760
 aaat tttcaa atatctcact aagaaaaagt tgcttatttt tgtcaatagc ggagct tttc 5820
 taacaaagga tgaagtggct agtttacaag agtatatatc attgacaaat ttaacagttc 5880
 tct ttttaga accacgtgaa ctatatgatt ttccgcagta tatttttagat gaagattatt 5940
 tcttaataac taaaaatattg gtataa 5966

<210> 505

<211> 4107

5 <212> ADN

<213> Streptococcus pyogenes

<400> 505

ES 2 398 918 T3

atggataaga aataactcaat aggcttagat atcggcacia atagcgtcgg atgggcgggtg 60
atcactgatg aatataaggt tccgtctaaa aagttcaagg ttctgggaaa tacagaccgc 120

ES 2 398 918 T3

cacagtatca aaaaaaatct tataggggct cttttatttg acagtggaga gacagcggaa 180
gcgactcgtc tcaaacggac agctcgtaga aggtatacac gtcggaagaa tcgtatttgt 240
tatctacagg agattttttc aatgagatg gcgaaagtag atgatagttt ctttcatcga 300
cttgaagagt ctttttttgg ggaagaagac aagaagcatg aacgtcatcc tatttttggga 360
aatatagtag atgaagttgc ttatcatgag aaatatccaa ctatctatca tctgcgaaaa 420
aaattggtag attctactga taaagcggat ttgcgcttaa tctatttggc cttagcgcac 480
atgattaagt ttcgtggtca ttttttgatt gagggagatt taaatcctga taatagtgat 540
gtggacaaac tatttatcca gttggtacaa acctacaatc aattatttga agaaaaccct 600
attaacgcaa gtggagtaga tgctaaagcg attctttctg cacgattgag taaatcaaga 660
cgattagaaa atctcattgc tcagctcccc ggtgagaaga aaaatggctt atttgggaat 720
ctcattgctt tgcattggg tttgaccct aattttaaat caaattttga tttggcagaa 780
gatgctaaat tacagctttc aaaagatact tacgatgatg atttagataa tttattggcg 840
caaattggag atcaatatgc tgatttgttt ttggcagcta agaatttatc agatgctatt 900
ttactttcag atatcctaag agtaaatact gaaataacta aggctccccct atcagcttca 960
atgattaaac gctacgatga acatcatcaa gacttgactc ttttaaaagc tttagttcga 1020
caacaacttc cagaaaagta taaagaaatc ttttttgatc aatcaaaaaa cggatatgca 1080
ggttatattg atgggggagc tagccaagaa gaattttata aatttatcaa accaatttta 1140
gaaaaaatgg atggtactga ggaattattg gtgaaactaa atcgtgaaga tttgctgcgc 1200
aagcaacgga cctttgacaa cggctctatt ccccatcaaa ttcacttggg tgagctgcat 1260
gctattttga gaagacaaga agacttttat ccatttttaa aagacaatcg tgagaagatt 1320
gaaaaaatct tgacttttcg aattccttat tatgttggtc cattggcgcg tggcaatagt 1380
cgttttgcat ggatgactcg gaagtctgaa gaaacaatta ccccatggaa ttttgaagaa 1440
gttgtcgata aagggtgctc agctcaatca tttattgaac gcatgacaaa ctttgataaa 1500
aatcttccaa atgaaaaagt actaccaaaa catagtttgc tttatgagta ttttacggtt 1560
tataacgaat tgacaaaggt caaatatggt actgaaggaa tgcgaaaacc agcatttctt 1620
tcaggtgaac agaagaaagc cattgttgat ttiactcttca aaacaaatcg aaaagtaacc 1680
gttaagcaat taaaagaaga ttatttcaaa aaaatagaat gttttgatag tgttgaat 1740
tcaggagtgt aagatagatt taatgcttca ttaggtacct accatgattt gctaaaaatt 1800
attaagata aagatttttt ggataatgaa gaaaatgaag atatcttaga ggatattggt 1860
ttaacattga ccttatttga agataggag atgattgagg aaagacttaa aacatagct 1920
cacctctttg atgataaggt gatgaaacag cttaaacgtc gccgttatac tggttgggga 1980
cgtttgcctc gaaaattgat taatggtatt agggataagc aatctggcaa aacaatatta 2040

ES 2 398 918 T3

gattttttga aatcagatgg ttttgccaat cgcaatttta tgcagctgat ccatgatgat 2100
agtttgacat ttaaagaaga cattcaaaaa gcacaagtgt ctggacaagg cgatagttta 2160
catgaacata ttgcaaattt agctggtagc cctgctatta aaaaagggtat tttacagact 2220
gtaaaagttg ttgatgaatt ggtcaaagta atggggcggc ataagccaga aaatatcgtt 2280
attgaaatgg cacgtgaaaa tcagacaact caaaagggcc agaaaaattc gcgagagcgt 2340
atgaaacgaa tcgaagaagg tatcaaagaa ttaggaagtc agattcttaa agagcatcct 2400
gttgaaaata ctcaattgca aaatgaaaag ctctatctct attatctcca aaatggaaga 2460
gacatgtatg tggaccaaga attagatatt aatcgtttta gtgattatga tgtcgcacac 2520
attgttccac aaagtttctt taaagacgat tcaatagaca ataaggcttt aacgcgttct 2580
gataaaaaatc gtggtaaatc ggataacggt ccaagtgaag aagtagtcaa aaagatgaaa 2640
aactattgga gacaacttct aaacgccaag ttaatcactc aacgtaagtt tgataattta 2700
acgaaaagctg aacgtggagg tttgagtgaa cttgataaag ctggttttat caaacgccaa 2760
ttggttgaaa ctgcgcaaat cactaagcat gtggcacaaa ttttggatag tcgcatgaat 2820
actaaatagc atgaaaatga taaacttatt cgagagggta aagtgattac cttaaaatct 2880
aaattagttt ctgacttccg aaaagatttc caattctata aagtacgtga gattaacaat 2940
taccatcatg cccatgatgc gtatctaaat gccgtcgttg gaactgcttt gattaagaaa 3000
tatccaaaac ttgaatcgga gtttgtctat ggtgattata aagtttatga tgttcgtaaa 3060
atgattgcta agtctgagca agaaataggc aaagcaaccg caaaatattt cttttactct 3120
aaatcatgaa acttcttcaa aacagaaatt acaactgcaa atggagagat tcgcaaacgc 3180
cctctaactg aaactaatgg ggaaactgga gaaattgtct gggataaagg gcgagatttt 3240
gccacagtgc gcaaagtatt gtccatgccc caagtcaata ttgtcaagaa aacagaagta 3300
cagacaggcg gattctcaa ggagtcaatt ttacaaaaa gaaattcgga caagcttatt 3360
gctcgtaaaa aagactggga tccaaaaaaa tatggtgggt ttgatagtcc aacggtagct 3420
tattcagtcc tagtggttgc taaggtggaa aaagggaaat cgaagaagtt aaaatccgtt 3480
aaagagttac tagggatcac aattatggaa agaagttcct ttgaaaaaaa tccgattgac 3540
tttttagaag ctaaaggata taaggaagtt aaaaaagact taatcattaa actacctaaa 3600
tatagtcttt ttgagttaga aaacggtcgt aaacggatgc tggctagtgc cggagaatta 3660
caaaaaggaa atgagctggc tctgccaagc aaatatgtga atttttata tttagctagt 3720
cattatgaaa agttgaaggg tagtccagaa gataacgaac aaaaacaatt gtttgtggag 3780
cagcataagc attatttaga tgagattatt gagcaaatca gtgaattttc taagcgtgtt 3840
attttagcag atgccaattt agataaagtt cttagtgcac ataacaaca tagagacaaa 3900
ccaatacgtg aacaagcaga aaatattatt catttattta cgttgacgaa tcttgagct 3960

ES 2 398 918 T3

cccgctgctt ttaaataatt tgatacaaca attgatcgta aacgatatac gtctacaaaa 4020
 gaagttttag atgccactct tatccatcaa tccatcactg gtctttatga aacacgcatt 4080
 gatttgagtc agctaggagg tgactga 4107

<210> 506
 <211> 870
 5 <212> ADN
 <213> Streptococcus pyogenes

<400> 506

atggctgggtt ggcgtactgt tgtggtaaata acccactcga aattatccta taagaataat 60
 catctgattt ttaaggatgc ctataaaacg gagctgatcc atttatcaga aattgatatt 120
 ttgttattag aaacgaccga tattgtcttg tccactatgc tggtaaaacg gctagtggat 180
 gagaatgtcc ttgtcatatt ctgtgatgat aaacgattac caacagctat gctgatgcct 240
 ttttatggtc gtcatgattc gagtttacag cttgggaaac aaatgtcctg gtcagaaaca 300
 gtcaaatcgc aggtttggac gacgattatt gctcaaaaga ttttgaatca atcttgctat 360
 ctaggagcat gtcctattt tgaaaaatcc caatctatta tggatttata tcatggtttg 420
 gaaaattttg atccgagtaa tcgagaaggg catgcagcga gaatttattt taatacactt 480
 tttgggaacg atttctcaag agatttggag catccaatca atgcaggtct ggattatggc 540
 tatactttat tattgagat gtttgcgcgt gaagtgggtg tgtctggatg tatgactcag 600
 tttgggctta aacacgctaa tcagtttaata cagttcaatt ttgctagcga tattatggaa 660
 ccatttaggc ctttagtgga taagattggt tatgaaaatc gaaatcagcc ttttcccaaa 720
 ataaagagag agttatttac tttgttttca gatacatttt catataatgg taaagagatg 780
 tatctcacga atattattag cgattatact aaaaaagttg tcaaagctct gaataatgaa 840
 10 gggaaaggag ttcctgaatt taggatatga 870

<210> 507
 <211> 342
 <212> ADN
 15 <213> Streptococcus pyogenes

<400> 507

ES 2 398 918 T3

atgagttata gatatatgag aatgatactt atgtttgata tgccgacgga caccgctgag 60
 gaacgaaaag cctatcgaaa atttcggaaa tttttactta gtgaagggtt tatcatgcat 120
 caatthttcta tttatagtaa gttgctggtg aataatacag ctaacaatgc catgattggt 180
 cggctgaggg agcataatcc taataaagga aatattacat tactaacggt cacggaaaaa 240
 cagtttgcac gaatgattta tttacatggt gaaagaaata attgtattgc aaactccgat 300
 gaaagacttg tatttcttgg ggaggctttt gatgaatctt aa 342

<210> 508

<211> 663

5 <212> ADN

<213> Streptococcus pyogenes

<400> 508

atgaatctta atthtttctt actagatgaa cggattccat taagaggcgg tacaattctt 60
 gtgctcgaag atgtctgtgt atthttcaaaa atagtgcaat attgttacca atatgaggaa 120
 gattctgaac ttaaathttt tgatcacaag atgaaaacaa tcaaagaatc agaatcatg 180
 cttgtaacag atathtttagg atttgatgtt aactcctcaa ccathtttaa attgattcat 240
 gcagatttag aatctcaatt taatgagaaa cccgaagtga aatcgatgat tgacaaattg 300
 gttgctacga ttacagaact gattgtcttt gaatgcttag aaaatgaatt agatttagag 360
 tatgatgaaa tcacaatcct ggaattgatt aagtccttag gagtaaaagt agaaacgcaa 420
 agtgatacta tthttgaaaa atgtctagag atacttcaaa tthttcaaata tctcactaag 480
 aaaaagttgc ttathtttgt caatagcggg gctthttctaa caaaggatga agtggctagt 540
 ttacaagagt atatatcatt gacaaattta acagttctct tthttagaacc acgtgaacta 600
 tatgathttc cgcagtatat tthtagatgaa gathathttct taataactaa aatatggta 660
 10 taa 663

<210> 509

<211> 8020

<212> ADN

15 <213> Streptococcus pyogenes

<400> 509

ES 2 398 918 T3

atgagaatga	ttttagcaca	ctatgactgt	aaaaaagata	aaaagcaatc	tttagatgag	60
catttatggc	atgtggcctg	ttctagtcga	caggaagcat	ctataattgg	tcaaggagat	120
gtgctttttt	taattggtct	ttaccacgac	ctgggcaaag	ctgatcgaac	ctttcaagat	180
aaattattaa	ataatccaaa	tcggcatggt	gatcactctt	atgcaggggc	aaaatactta	240
tgttctatta	ttgggcctca	tctaaaaaac	cgagggggtg	ataaaaatga	gagaatgaca	300
ttcaacgaaa	tgggtggggta	tgtcatctct	gctcatcatg	ggatgtatga	tttatgctac	360
tattttgacg	atgctgaata	ttatggcttt	aataagttta	aaaatcgtat	caatagagac	420
ttagatggtt	atcactatca	tgaagatatt	aaagggtagc	ctctaaaatt	agaaaaaaaa	480
ttatgtgatt	atggctacaa	agatttaagg	gagcttattg	ataaagcttt	tgataattac	540
caacaagcca	tgtcttcctt	aaactggcaa	gataagagtg	agtgggatta	ttatcagtct	600
tgatggatga	gactttactt	gtcactctta	aaaaacgctg	atattttgga	cacagtaaat	660
gcctatggcc	ttaagataag	tcctatggat	aaaacagagc	gatcctttct	aaaacactcc	720
tatttagcgg	ccattgaaca	aaaatatgct	agctttggac	agccaaacaa	tcagttgaac	780
actattcgga	cagaaatcgc	tgagcgtggt	aaagaaagag	gtaaacgaga	ttccaagggg	840

ES 2 398 918 T3

atttatoctt tagatttacc gacaggagct ggcaagacta atcttagtat gcgttatgcg 900
 tttaccaat tagttcatca cgacaaatca aggttttttt acataactcc ctttctttcg 960
 gttcttgagc aaaatgcttc cgaaattaga aaagttacag gtgaccttg cgttctagaa 1020
 caccattcca atgtggtgaa acaggcta at gaagatgatg atgataagga cagtttattg 1080
 tcagcttacc ttagtgatag ctgggacagt caagtagtct tgacttctat gtttcaattt 1140
 ttccaaacac ttttcaaac aaaatcagct aatctgagac gtttttcaag tttgattaat 1200
 agtgttgtga ttctagatga agttcaatcc ctgcctattg aagtcaccac tttgtttaat 1260
 ttaacgatga attttttaaa taaagttatg gatacaacca tcgttctttg cacagcgaca 1320
 caacctgctt atgattcttc agagattgac catcgtatct gttatggagg gaacttggga 1380
 gaattagctg aaatagttga gttaacgatt gaagaaaaac agattttttc aaggacagag 1440
 cttagaaaat ttgatgatag tgatcagaaa gttcacttga ctgatgttat taaccttatt 1500
 ctaggtgagg aaaactcagt tcttgctatt ttttaatacga aaaaaacggt tcataactgc 1560
 tatactatgc taaaagacat gactgataga ccggctctatc agctttcgcac aaatatgtgt 1620
 gcgcagcata gacttgactt gattgctaag atcaaacgg agttacaaaa taatatccct 1680
 attatttgta ttagcacgca attaattgaa gcagggttag atggtgattt tcatcgcgctc 1740
 attcgttctt actcagggat tgattctatt gttcaggctg ctggacggtg taaccgagaa 1800
 ggcaaacgag ataaagggca agtcactctt gtcaatctga ccaatgaaga ggaaaatatt 1860
 tctaggctga cagaaataaa aactaaaaaa gaagccacag aatctattct tcataagatt 1920
 gggctctcaa ttgatattct aactttaaac cgtgactttt ttgagtatta ttatgccaat 1980
 aatcagggac tgatggatta tcctttggaa gacaacctat caatctacga ctatttaagc 2040
 cttaatattt atcagacggc aaataaaaag ttcaaaggta agttaaaca agcttttaaa 2100
 acagcaggag ccaaaatgaa cctcatcaat aatgatatga taggaattct cgtaccttat 2160
 ggcaagctg agaaaaaatt ggcttattta gaagaattag gtgtgtcaca tttttatca 2220
 gcaaaagatt atcaaacgat aaaatcatta ctaaaagagt tacaaccttt tacggttaat 2280
 gtccgcgaga acgatcctct ctttgagaca acaaaatctt atctaaatgg tcagattctg 2340
 gttttgacgt cggagtatta tgacacggaa agaggagtta aatacattc agctagcttt 2400
 tacttctaac tcaaacgaa agaagattaa caaaagggtg ttagaggacc ttgttaacct 2460
 gccaatcatc attagtaatt attatcaatt tagactattt aataaaatta gattacaaaa 2520
 aaacagaagg aggaaagtag cttgtacaga tctagagact tctacgtgag agtaagtgg 2580
 cagcgagctc tttttacaaa tccagccaca aaaggggat cggaacgctc atcctattcg 2640
 gttccgacta gacaggcact gaatggtatc gttgatgcca tctattataa gccgaccttt 2700
 actaatatcg tcacagagg taaggttatt aaccagattc aaaccgaatt acagggtgct 2760

ES 2 398 918 T3

agggctctgt tacatgatta tagtgcagat ttaagttatg tadcctatth gagtgatgth 2820
 gtttatctga tcaagtttca ttttgthtg aatgaagata gaaaagatth gaactcagat 2880
 agacttccag ctaaacaatga agccattatg gagcgttcta ttcgtaaagg gggacgtcga 2940
 gatgtgthth tgggtacaag agaaththta gggctttag atgatatcag ccaagaagag 3000
 tatgagacta ctgtgtcga ttataatggt gtcaatcag acttggaat catgttccat 3060
 tcctthgcct atccgaagga caaaaagaca ccattaaaat catactthac aaagactgtg 3120
 atgaaaaatg gagtcattac gthtaagca cagtctgaat gcgatathgt taacacgctt 3180
 tctagthtatg cththaaagc accagaggag ataaaatcgg thaacgatga atgcatggag 3240
 tatgatgcca tggagaaagg agaaaactga tggaththth tacttctctc thgaagactt 3300
 atgaaaaagc agagctagca gacttggtg atcatcaaaa aagaaataat gagccgthth 3360
 tactgccgat ttatcatag agththaaagt caaatggtaa aaatathcatt tcagtgaaac 3420
 thgacaaaga tggccagthh cacaaggcag aaththatggc agataagcaa atgaththth 3480
 thcctgthaac ggctgathct gthgctaggt caggtagtca thcctgcaccg catcccctag 3540
 tcgataaath thgctththth agtgctgaaa tggggcagat tcagththth thctththcata 3600
 agcaactgaa thactggath gathththth agggaggthga thgthcaagaaa thththaacct 3660
 thgththcagca gththththth aagccagaath thththaacath gathththth thctththathth 3720
 gthctgatha thcaacataath caaththaaag thcaaththth thgathththth ggaaaagaaa 3780
 aaththathth thththththth thgctththth aathththcaath thgathththth cagggctthth 3840
 aaaaatgathc ggthththgaca thththaaagcct thaccaathc thththththth ththgththgag 3900
 ccaathththth aaththththth atththththth thththththth agaggaththth cththththth 3960
 agcathththth thththththth aaththththth thththththth thththththth agagaththth 4020
 athththththth thththththth thththththth thththththth thththththth actthththth 4080
 agathththth aaththththth thctthththth aaththththth thththththth thththththth 4140
 thththththth thththththth thththththth atgathththth aaththththth cggthththth 4200
 thththththth aaththththth gaththththth aagathththth thththththth thththththth 4260
 thththththth agcaactgaa gacaaththth gaththththth atcathththth aagggaththth 4320
 aaththththth thththththth actthththth thththththth gaththththth agcaathththth 4380
 ggaththththth aaththththth thththththth aagcgtththth atththththth aaththththth 4440
 agththththth aacathththth thththththth gaththththth thththththth thththththth 4500
 cctthththth thththththth thththththth thththththth thththththth thththththth 4560
 aaththththth thththththth aaththththth aaththththth thththththth agththththth 4620
 atgththththth gaththththth thththththth aaththththth thththththth aagathththth 4680

ES 2 398 918 T3

atcgttaccg taagcactgg tatcaagttg agcaggtctg cttagcaatt ttacacaaac 4740
 aaaatgggga ggaattttca ccgatgctag atcataccea tcaaaatcgt tcctatcttt 4800
 ttggacgatt attagcaatt tttgaattaa tcgagacctt gcgttatggc ttggatggaa 4860
 acaataacga ccgtattacc aatgctgaac gttattggac agcctatact ggacaaccaa 4920
 caaaattgat gatgttattg gaaaataaaa ttaagcctta cgaagaacca ttgaaattaa 4980
 atcgtcgtgg cagttggatg aaattagaaa aagaaaaaga agagatttta gaactgtaa 5040
 atcctctggt agaaacagaa acaatggaaa aacccttaga ttaccgcttt atttttgggt 5100
 attatgctga gaaaaactat tactatacaa aacaaaacac ggaagtaaca gaaagtgagg 5160
 agtaaaaaga tgttggaaca caaaattgat tttatggtaa ctottgaagt gaaagaagca 5220
 aatgcaaagtg gtgatccctt aaatggaaac atgcctcgta cagatgcaa aggatatggt 5280
 gtgatgagtg atgtctccat taaacgtaag attcgtaatc gtttgcaaga tatggggaag 5340
 tctatTTTTg tgcaagctaa tgagcgtatt gaagatgatt ttcgttcact ggaaaaacgc 5400
 ttttcgcaac attttacagc taagacacct gacaaagaaa ttgaagaaaa agcaaagca 5460
 ttatggtttg atgttcgtgc ttttggacaa gtttttactt atctgaaaaa atcaattggg 5520
 gtgcgtggac cagtttccat cagtatggct aagtccttgg agccaattgt catttccagc 5580
 cttcaaatta cgcgtagtac caatggtagt gaagctaaga ataatagtgg ccgctcttct 5640
 gatacgatgg ggacaaaaca tttttagat tatgggtgtgt atgtacttaa aggttctatc 5700
 aatgcttatt ttgctgaaaa gactggtttt tctcaggaag atgctgaggc tattaagaa 5760
 gttttggtta gcttgtttga aaatgatgcg tcgtctgcac gtccggaagg ctctatgcga 5820
 gtttgtgaag tcttttggtt tacgcattca agcaaattgg gaaatgtttc aagtgcgcgt 5880
 gtctttgact tgtagagta tcatcaatca atagaagaaa aaagcactta tgacgcttat 5940
 cagattcatc taaatcaaga aaaattggct aaatatgaag cgaaagggtt aacgcttgaa 6000
 atcctagaag gactctagta tggctatgc cgaagatgat tattaatgc tgtcaggtat 6060
 tcagcatttc caattttgta aacgtcaatg ggcggtgatc catattgagc aacaatggct 6120
 tgataatgaa gcgacagcgc atggacaggt ttacatact aaagcagata acccttacat 6180
 taaagaaaaa cgaaaagagc ttttggcttc acgtgctatg cccatttctt ctgcagaact 6240
 tggactttca ggaattatgg atgttgtgga attttataaa gatgatcaag gtgtgtcttt 6300
 gaggggaaaa cgtgggaaat ggttaccaa agttgtggaa tacaagcgcg gaaaacctaa 6360
 aaaagatacc agagatattg tccagttggt ggctcagacc atgtgtttag aagaaacgct 6420
 agactgcgac attaacgaag gttgtcttta ttaccatagt gtcaatcaa gagtgattgt 6480
 tcctatgaca tcagctttgc gtcaagaagt gaaggaatta gccgcagaga tgcagaggt 6540
 ttatcagagt caaatgctac ctaaagcagc ttattttaaa aactgtcagc tttgttcttt 6600

ES 2 398 918 T3

agtcgatatt tgtaagccca ggttgagtaa aaaaacaagg agtgtgtcgc gttacatcaa 6660
 tgaggctatg accagtgagg agatggacct atgaagaagt tgctaaatac cttgtatttg 6720
 acgcaagaag atttttatgt cactaaagag ggcgataaca ttgttatcaa gcaagaaggt 6780
 aaggttctca aacggtttcc gtttcggatt attgacggta ttgtctgttt ttcttatttg 6840
 ggtgtgtcgt ctgctttggt gaagttatgt acggagaatc agattaatth atcgtttcat 6900
 acaccacaag ggcgtttttg tggtcgctat attggttcaa ccaatgggaa tgtgttggtg 6960
 cgtagagaac attatcgtht atctgatcgt gaggaatctt tggaaatcgc aaagcggtht 7020
 attttggtca aaatttccaa ctcaaggaaa tacttgctac gctttaaacy agatcatcgt 7080
 caacagattg ataccaagct ttttgaggct gttaatgacy aattgatatg ggctttagag 7140
 atggttcagg cagcagataa taaagactct ttaagagggg ttgaaggcca agctgctaat 7200
 cagtattttc gcatatttaa tgacctggtg ttgacggaca aaaaaacytt ttacttccaa 7260
 ggtcggagta aacgaccacc cttagattgt gttaatgcc tcttgtcttt tggttacagt 7320
 ttactgacct ttgaatgtca atctgccttg gaagctgtcg gattagacag ttacgttgggt 7380
 ttctttcaca cggatcgtcc tgggcgtgct agtttagcgc ttgatttagt tgaagagttc 7440
 cgctcatata ttgtagatcg ttttgtcttt tcattaatta ataaaggaca acttcagaaa 7500
 aaacactttg aggttaaaga aatggtagt attttattga cggaaaatgg cagagctatt 7560
 tttattgatt tgtggcagaa gcgtaagcat actgaggtag aacatccttt taaaaagag 7620
 aaagtaaaac ttatgttatt accctatgta caagcgcgac ttttagctaa ggctatacga 7680
 ggagatttag aaagctatcc accttttatg gtttaggaga tgttatatga tggtttttagt 7740
 cacttatgat gtaaatacgg aaacacctgc tggtagaaaa agattgcgtc atgttgccaa 7800
 actctgtgtg gactatgggc aacgtgttca aaattctgth tttgaatgth ctgtgacacc 7860
 cgcagaatth gtggatataa agcaccgctt aacacaaatc attgatgaga aaactgatag 7920
 tattcgctth tattttattgg ggaaaaattg gcagaggcgt gtggaaacac ttggtcgctc 7980
 agacagctat gaccagata aaggtgtctt attattgtaa 8020

<210> 510

<211> 2409

5 <212> ADN

<213> Streptococcus pyogenes

<400> 510

ES 2 398 918 T3

atgagaatga	ttttagcaca	ctatgactgt	aaaaaagata	aaaagcaatc	tttagatgag	60
catttatggc	atgtggcctg	ttctagtcga	caggaagcat	ctataattgg	tcaaggagat	120
gtgctttttt	taattggtct	ttaccacgac	ctgggcaaag	ctgatcgaac	ctttcaagat	180
aaattattaa	ataatccaaa	tcggcatggt	gatcactctt	atgcaggggc	aaaatactta	240
tgttctatta	ttgggcctca	tctaaaaaac	cgagggggtg	ataaaaatga	gagaatgaca	300

ES 2 398 918 T3

ttcaacgaaa tgggtgggta tgtcatctct gctcatcatg ggatgtatga tttatgctac 360
 tattttgacg atgctgaata ttatggcttt aataagttta aaaatcgtat caatagagac 420
 ttagatgggt atcactatca tgaagatatt aaagggtagc ctctaaaatt agaaaaaaaa 480
 ttatgtgatt atggctacaa agatttaagg gagcttattg ataaagcttt tgataattac 540
 caacaagcca tgtcttcctt aaactggcaa gataagagtg agtgggatta ttatcagtct 600
 tgtatgggta gactttactt gtcactctta aaaaacgctg atattttggg cacagtaaat 660
 gcctatggcc ttaagataag tcctatggat aaaacagagc gatcctttct aaaacactcc 720
 tatttagcgg ccattgaaca aaaatatgct agctttggac agccaaacaa tcagttgaac 780
 actattcgga cagaaatcgc tgagcgtggt aaagaaagag gtaaacgaga ttccaagggg 840
 atttatcgct tagatttacc gacaggagct ggcaagacta atcttagtat gcgttatgcg 900
 tttaccaat tagttcatca cgacaaatca aggttttttt acataactcc ctttctttcg 960
 gttcttgagc aaaatgcttc cgaaattaga aaagttacag gtgaccttgg cgttctagaa 1020
 caccattcca atgtggtgaa acaggctaata gaagatgatg atgataagga cagtttattg 1080
 tcagcttata ttagtgatag ctgggacagt caagtagtct tgacttctat ggttcaattt 1140
 ttccaaacac ttttcaaac aaaatcagct aatctgagac gtttttcaag tttgattaat 1200
 agtgttgta ttctagatga agttcaatcc ctgcctattg aagtcaccac tttgtttaat 1260
 ttaacgatga attttttaaa taaagttatg gatacaacca tcgttctttg cacagcgaca 1320
 caacctgctt atgattcttc agagattgac catcgtatct gtatggagg gaacttggga 1380
 gaattagctg aaatagttga gttaacgatt gaagaaaaac agattttttc aaggacagag 1440
 cttagaaaat ttgatgatag tgatcagaaa gttcacttga ctgatgttat taaccttatt 1500
 ctaggtgagg aaaactcagt tcttgctatt ttaatacga aaaaaacggt tcataactgc 1560
 tatactatgc taaaagacat gactgataga ccggtctatc agctttcgac aaatatgtgt 1620
 gcgcagcata gacttgactt gattgctaag atcaaaacgg agttacaaaa taatatccct 1680
 attatgtgta ttagcacgca attaattgaa gcaggtgtag atggtgattt tcatcgcgctc 1740
 attcgttctt actcagggat tgattctatt gttcaggctg ctggacggtg taaccgagaa 1800
 ggcaaacgag ataaagggca agtcactctt gtcaatctga ccaatgaaga ggaaaatatt 1860
 tctaggctga cagaaataaa aactaaaaaa gaagccacag aatctattct tcataagatt 1920
 gggctccaa ttgatatctc aactttaaac cgtgactttt ttgagtatta ttatgccaat 1980
 aatcagggac tgatggatta tcctttggaa gacaacctat caatctacga ctatttaagc 2040
 cttaatatth atcagacggc aaataaaaag ttcaaaggta agttaaaca agcttttaaa 2100
 acagcaggag ccaaaatgaa cctcatcaat aatgatatga taggaattct cgtaccttat 2160
 ggcgaagctg agaaaaaatt ggcttattta gaagaattag gtgtgtcaca ttttttatca 2220

ES 2 398 918 T3

gcaaaagatt atcaaacgat aaaatcatta ctaaaagagt tacaaccttt tacggttaat 2280
 gtccgcgaga acgatcctct ctttgagaca acaaaatcctt atctaaatgg tcagattctg 2340
 gttttgacgt cggagtatta tgacacggaa agaggagtta aatacgattc agctagcttt 2400
 tacttctaa 2409

<210> 511

<211> 729

5 <212> ADN

<213> Streptococcus pyogenes

<400> 511

ttgtacagat ctagagactt ctacgtgaga gtaagtggtc agcgagctct ttttaciaat 60
 ccagccacaa aagggggatc ggaacgctca tcctattcgg ttccgactag acaggcactg 120
 aatggatcgc ttgatgccat ctattataag ccgaccttta ctaatatcgt cacagaggtt 180
 aaggttatta accagattca aaccgaatta cagggtgtca gggctctggt acatgattat 240
 agtgcagatt taagttatgt atcctatctg agtgcgcttg tttatctgat caagtttcat 300
 tttgtttggg atgaagatag aaaagatttg aactcagata gacttccagc taaacatgaa 360
 gccattatgg agcgttctat tcgtaaaggg ggacgtcgag atgtgttttt ggggtacaaga 420
 gaatgtttag ggctttaga tgatcagc caagaagagt atgagactac tgtgtcgtat 480
 tataatggcg tcaatatcga cttgggaatc atgttccatt cctttgccta tccgaaggac 540
 aaaaagacac cattaataatc atactttaca aagactgtga tgaaaaatgg agtcattacg 600
 tttaaagcac agtctgaatg cgatattggt aacacgcttt ctagttatgc ttttaaagca 660
 ccagaggaga taaaatcggg taacgatgaa tgcatggagt atgatgccat ggagaaagga 720
 10 gaaaactga 729

<210> 512

<211> 1896

<212> ADN

15 <213> Streptococcus pyogenes

<400> 512

ES 2 398 918 T3

atggatTTTT ttacttctct cttgaagact tatgaaaaag cagagctagc agacttggtt	60
gatcatcaaa aaagaaataa tgagccggtt ttactgccga tttatcatac gagtttaaag	120
tcaaattggtt aaaatatcat ttcagtgaaa cttgacaaaag atggccagtt tcacaaggca	180
gaatttatgg cagataagca aatgattatt tttcctgtaa cggctgattc tgttgctagg	240
tcaggtagtc atcctgcacc gcatccccta gtcgataaat ttgcttatta tagtgctgaa	300
atggggcaga ttcagtatga ttcttttcat aagcaactga ataactggat tgattattgt	360
gaggaggggtg atgtcaagaa atttttaacc tttgttcagc agttcatttt gaagccagaa	420
tttctaacat tgattcttga ttctttaatt ggtcctgatt atcaacataa tcaattaaaa	480

ES 2 398 918 T3

gtcacatttt gtgatgccac tggaaaagaa aaattaattg atttatcagc ttgcttttta 540
 gaattttcaa ttgatcagtt ccagggtctt aaaaatgaat cggtttcgac atttaaagcc 600
 ttacaccaat cctatatttc ttttgttgaa gccaatcgtg aaaatctcgg tatttgtaat 660
 attagtggac gagaggaaca gcttaccgat aagcatagag gtttgatggg gaatgctaaa 720
 atcatctctg ttagtaataa aagagaagct tataaaggac gttttagaga acgcgaagac 780
 gtttttagtg ttggctatga aacttccgaa aagattcatt taatgctcaa gtaccttita 840
 gaaaataaaa ataccagtac ttggttaggg tcttctcaat atttaatcaa ctggttcagc 900
 gatgatttaa caaatgatag tcggttggat attgtatcac caatcttga tgatggactt 960
 gaagaagatg atgatgacga tacgcctcct gttataacat tagcaactga agacaataaa 1020
 agaattggta aatcattcat caagggacaa aaattatttg ctaatgatgc cacttactac 1080
 gttgctatth tgaataaaac cagcaatggg cggattgctt taaaatattt tcgtcagctt 1140
 caagcgtccc aattactcac caatcttaac aagtggcagg aaacatacag ttgggagtcg 1200
 cgatctaagt ttgggaaaag tcgcttaaga acccctactt ttcatgacat ccttaatgtg 1260
 tcctacgggg ttgatagggg tcgcttcctt gaattagata atgataactt caaaagtgat 1320
 caaattcaaa agttagtggc aagtttgatt gatggtaaac cgatgccaca gtccattgtc 1380
 aaaaagttag gtaacaatgt taaagaacga catcgttacc gtaagcactg gtatcaagtt 1440
 gagcaggtct gcttagcaat tttacacaaa caaaatgggg aggaattttc accgatgcta 1500
 gatcatacca atcaaaatcg ttcctatctt tttggacgat tattagcaat ttttgaatta 1560
 atcgagacct tgcgttatgg cttggatgga aacaataacg accgtattac caatgctgaa 1620
 cgttattgga cagcctatac tggacaacca acaaaattga tgatgttatt ggaaaataaa 1680
 attaagcctt acgaagaacc attgaaatta aatcgtcgtg gcagttggat gaaattagaa 1740
 aaagaaaaag aagagattht agaactgtta aatcctctgt tagaaacaga aacaatggaa 1800
 aaacccttag attaccgctt tathtttggg tattatgctg agaaaaacta ttactataca 1860
 aaacaaaaca cggaagtaac agaaagtgag gagtaa 1896

<210> 513

<211> 849

5 <212> ADN

<213> Streptococcus pyogenes

<400> 513

ES 2 398 918 T3

atggttgaac acaaaattga ttttatggta actccttgaag tgaaagaagc aaatgcaaat 60
 ggtgatccct taaatggaaa catgcctcgt acagatgcca aaggatatgg tgtgatgagt 120
 gatgtctcca ttaaacgtaa gattcgtaat cgtttgcaag atatggggaa gtctatTTTT 180
 gtgcaagcta atgagcgtat tgaagatgat tttcgttcac tggaaaaacg cttttcgcaa 240
 cattttacag ctaagacacc tgacaaagaa attgaagaaa aagcaaatgc attatgggtt 300
 gatgttcgtg cttttggaca agtttttact tatctgaaaa aatcaattgg ggtgcgtgga 360
 ccagtttcca tcagtatggc taagtccttg gagccaattg tcatttccag ctttcaaatt 420
 acgcgtagta ccaatgggat ggaagctaag aataatagtg gccgctcttc tgatacgatg 480
 gggacaaaac atttttaga ttatgggtg gatgtactta aaggttctat caatgcttat 540
 tttgctgaaa agactgggtt ttctcaggaa gatgctgagg ctattaaaga agttttgggt 600
 agcttgtttg aaaatgatgc gtcgtctgca cgtccggaag gctctatgcg agtttgtaa 660
 gtcttttgggt ttacgcattc aagcaaattg ggaaatgttt caagtgcgcg tgtctttgac 720
 ttgttagagt atcatcaatc aatagaagaa aaaagcactt atgacgctta tcagattcat 780
 ctaaatacaag aaaaattggc taaatatgaa gcgaaagggt taacgcttga aatcctagaa 840
 ggactctag 849

- 5 <210> 514
- <211> 675
- <212> ADN
- <213> Streptococcus pyogenes

10 <400> 514

atggtctatg ccgaagatga ttatttaatg ctgtcaggta ttcagcattt ccaattttgt 60
 aaacgtcaat gggcgttgat ccatattgag caacaatggc ttgataatga agcgacagcg 120
 catggacagg ttttacatac taaagcagat aacccttaca ttaaagaaaa acgaaaagag 180
 cttttggctc cacgtgctat gcccatcttct tetgcagaac ttggactttc aggaattatg 240
 gatgttgtgg aattttataa agatgatcaa ggtgtgtcct tgaggggaaa acgtgggaaa 300
 tggttaccaa aagttgtgga atacaagcgc ggaaaaccta aaaaagatac cagagatatt 360
 gtccagttgg tggctcagac catgtgttta gaagaaacgc tagactgcca cattaacgaa 420
 ggttgtcttt attaccatag tgtcaatcaa agagtgattg ttcctatgac atcagctttg 480
 cgtcaagaag tgaaggaatt agccgcagag atgcatgagg tttatcagag tcaaagtcta 540
 cctaaagcag cttattttaa aaactgtcag ctttgttctt tagtcgatat ttgtaagccc 600
 aggttgagta aaaaaacaag gagtgtgtcg cgttacatca atgaggctat gaccagtgag 660
 gagatggacc tatga 675

ES 2 398 918 T3

<210> 515
 <211> 1026
 <212> ADN

5 <213> Streptococcus pyogenes

<400> 515

```

atgaagaagt tgctaaatac cttgtatctg acgcaagaag atttttatgt cactaaagag      60
ggcgataaca ttgttatcaa gcaagaaggt aaggttctca aacggtttcc gtttcggatt      120
attgacggta ttgtctgttt ttcttatctg ggtgtgtcgt ctgctttggt gaagttatgt      180
acggagaatc agattaatct atcgtttcat acaccacaag ggcgtttttg tggtcgctat      240
attggttcaa ccaatgggaa tgtggtggtg cgtagagaac attatcgttt atctgatcgt      300
gaggaatctt tggaatacgc aaagcggttt attttggtca aaatttccaa ctcaaggaaa      360
tacttgctac gctttaaacg agatcatcgt caacagattg ataccaagct ttttgaggct      420
gttaatgacg aattgatatg ggctttagag atggttcagg cagcagataa taaagactct      480
ttaagagggg ttgaaggcca agctgctaag cagtatcttc gcatatttaa tgacctggtg      540
ttgacggaca aaaaaacggt ttacttccaa ggtcggagta aacgaccacc cttagattgt      600
gttaatgccc tcttgtcttt tggttacagt ttactgacct ttgaatgtca atctgccttg      660
gaagctgtcg gattagacag ttacgttggg ttctttcaca cggatcgtcc tgggcgtgct      720
agtttagcgc ttgatttagt tgaagagttc cgctcatata ttgtagatcg ttttgtcttt      780
tcattaatta ataaaggaca acttcagaaa aaacactttg aggttaaaga aatggtagt      840
attttattga cggaaaatgg cagagctatt tttattgatt tgtggcagaa gcgtaagcat      900
actgaggtag aacatccttt tacaaaagag aaagtaaaac ttatgttatt accctatgta      960
caagcgcagc ttttagctaa ggctatacga ggagatttag aaagctatcc accttttatg     1020
gtttag                                             1026
    
```

10

<210> 516
 <211> 294
 <212> ADN

15 <213> Streptococcus pyogenes

<400> 516

ES 2 398 918 T3

atgatggttt tagtcactta tgatgtaa atcgaaacac ctgctggtag aaaaagattg 60
cgatcatgtt ccaaactctg tgggactat gggcaactg ttcaaaattc tgtttttgaa 120
tgttctgtga cacccgcaga atttgtggat ataaagcacc gcttaacaca aatcattgat 180
gagaaaactg atagtattcg cttttattta ttggggaaaa attggcagag gcgtgtggaa 240
acacttggtc gctcagacag ctatgacca gataaaggtg tcttattatt gtaa 294

<210> 517

<211> 5966

5 <212> ADN

<213> Streptococcus pyogenes

<400> 517

atggataaga aatactcaat aggcttagat atcggcacia atagcgtcgg atgggcggtg 60
atcactgatg attataaggt tccgtctaaa aagctcaagg gtctgggaaa tacagaccgc 120
cacggtatca aaaaaaatct tataggggct cttttatttg acagtggaga gacagcggaa 180
10 gcgactcgtc tcaaacggac agctcgtaga aggtatacac gtcggaagaa tcgtatttgt 240

ES 2 398 918 T3

tatctacagg agatTTTTTC aaatgagatg gcgaaagtag atgatagttt ctttcatcga 300
cttgaagagt cttttttggt ggaagaagac aagaagcatg aacgtcatcc tatttttggga 360
aatatagtag atgaagttgc ttatcatgag aaatatccaa ctatctatca tctgcgaaaa 420
aaattggcag attctactga taaagtggat ttgcgcttaa tctatTTTggc cttagcgcag 480
atgattaagt ttcgtggTca ttttttgatt gagggagatt taaatcctga taatagtgat 540
gtggacaaac tatttatcca gttggTacaa acctacaatc aattatTTTga agaaaaccct 600
attaacgcaa gtagagtaga tgctaaagcg attcTTTctg cacgattgag taaatcaaga 660
cgattagaaa atctcattgc tcagctcccc ggtgagaaga aaaatggatt gtttgggaat 720
ctcattgctt tgtcattggg attgaccct aatTTTaaat caaatTTTga tttggcagaa 780
gatgctaaat tacagctTtc aaaagatact tacgatgatg atttagataa tttattggcg 840
caaattggag atcaatatgc tgatttgtt ttggcagcta agaattTatc agatgctact 900
ttactTtcag atatcctaag agtaaatagt gaaataacta aggcTcccct atcagctTca 960
atgattaagc gctacgatga acatcatcaa gacttgactc tttTaaaagc tttagTtcga 1020
caacaacttc cagaaaagta taaagaaatc ttttttgatc aatcaaaaaa cggatatgca 1080
ggttatattg atgggggagc tagccaagaa gaattTtata aatttatcaa accaatTTta 1140
gaaaaaatgg atggtactga ggaattattg gcgaaactaa atcgtgaaga tttgctgcgc 1200
aagcaacgga cttttgacaa cggctctatt ccctatcaaa ttcactTggg tgagctgcat 1260
gctatTTTga gaagacaaga agactTttat ccattTttaa aagacaatcg tgagaagatt 1320
gaaaaaatct tgactTttcg aattccttat tatgTtggtc cattggcgcg tggcaatagt 1380
cgTTTtgcag ggatgactcg gaagtctgaa gaaacaatta ccccatggaa ttttgaagaa 1440
gttgtcgata aaggTgcttc agctcaatca tttattgaac gcatgacaaa ctttgataaa 1500
aatcttccaa atgaaaaagt actaccaaaa catagTttgc tttatgagta ttttacggtt 1560
tataacgaat tgacaaaagt caaatatgTt actgagggaa tgcgaaaacc agcattTctt 1620
tcaggTgaac agaagaaagc cattgTtgat ttactcttca aaacaaatcg aaaagtaacc 1680
gttaagcaat taaaagaaga ttatTTTcaaa aaaatagaat gttttgatag tgttgaaatt 1740
tcaggagTtg aagatagatt taatgcttca ttaggtacct accatgattt gctaaaaatt 1800
attaaagata aagattTTTt ggataatgaa gaaaacgaag atatcttaga ggatattgTt 1860
ttaacattga ccttattTga agatagggag atgattgagg aaagacttaa aacatatgct 1920
cacctctTtg atgataaggt gatgaaacag cttaaacgTc gccgttatac tggTtgggga 1980
cgTTTgtctc gaaaattgat taatggtatt agggataagc aatctggcaa aacaatatta 2040
gattTTTtga aatcagatgg ttttgccaat cgcaattTta tgcagctgat ccatgatgat 2100
agTTTgacat ttaaagaaga cattcaaaaa gcacaagtgt ctggacaagg cgatagTtta 2160

ES 2 398 918 T3

catgaacata ttgcaaattt agctggtagc cctgctatta aaaaaggat tttacagact 2220
gtaaaagttg ttgatgaatt ggtcaaagta atggggcggc ataagccaga aaatatcgtt 2280
attgaaatgg cacgtgaaaa tcagacaact caaaagggcc agaaaaattc gcgtgagcgt 2340
atgaaacgta ttgaagaagg aataaaagaa ctaggaagtg atattctaaa ggagtatcct 2400
gttgaaaaca ctcaattaca aatgaaaag ctctatctct attatctcca aaatggaaga 2460
gacatgtatg tggaccaaga attagatatt aatcgtttaa gtgattatga tgtcgatcac 2520
attgttccac aaagtttctt taaagacgat tcaatagaca ataaggcttt aacgcgttct 2580
gataaaaatc gtggtaaattc ggataacggt ccaagtgaag aagtagtcaa aaagatgaaa 2640
aactattgga gacaacttct aaacgccaag ttaatcactc aacgtaagtt tgataattta 2700
acaaaagctg aacgtggagg tttgagtga cttgataaag ttggttttat caaacgccaa 2760
ttggttgaaa ctgcgcaaat cactaagcat gtggcacaaa ttttgatag tcgcatgaat 2820
actaaatcgc atgaaaatga taaacttatt cgagaggtta gagtgattac cttaaaatct 2880
aaattagttt ctgacttccg aaaagatctc caattctata aagtacgtga gattaacaat 2940
taccatcatg cccatgatgc gtatcttaat gccgtcgttg gaactgcttt gattaagaaa 3000
tatccaaaac ttgaatcgga gttgtctat ggtgattata aagtttatga tgttcgtaaa 3060
atgattgcta agtctgagca ggaaataggc aaagcaaccg caaaatattt cttttactct 3120
aatatcatga acttcttcaa aacagaaatt acacttgcaa atggagagat tcgcaaaccg 3180
cctctaactg aaactaatgg ggaaactgga gaaattgtct gggataaagg gcgagatttt. 3240
gccacagtgc gcaaagtatt gtccatgcc caagtcaata ttgtcaagaa aacagaagta 3300
cagacaggcg gattctccaa ggagtcaatt ttaccaaaaa gaaattcgga caagcttatt 3360
gctcgtaaaa aagactggga tccaaaaaaa tatggtgggt ttgatagtcc aacggtagct 3420
tattcagtcc tagtggttgc taaggtggaa aaagggaaat cgaagaagtt aaaatccgtt 3480
aaagagttac tagggatcac aataatggaa agaagctctt ttgaaaaga tccgattgac 3540
tttttagaag ctaaaggata taaggaagtt agaaaagact taatcattaa actacctaaa 3600
tatagtcttt ttgagttaga aaacggtcgt aaacggatgc tggctagtgc cggagaattg 3660
caaaaaggaa atgagctagc tctgccaagc aaatatgtga atttttata tttagctagt 3720
cattatgaaa agttgaaggg tagtccagaa gataacgaac aaaaacaatt gtttgtggag 3780
cagcataagc attatttaga tgagattatt gagcaaatca gtgaattttc taagcgtggt 3840
attttagcag atgccaattt agataaagtt cttagtgc ataaacaaaca tagagacaaa 3900
ccaatcgtg aacaagcaga aatattatt catttattta cgttgacgaa tcttggagct 3960
cccgtgctt ttaaattttt tgatacaaca attgatcgta aacgatatac gtctacaaaa 4020
gaagttttag atgccactct tatccatcaa tccatcactg gtctttatga aacacgcatt 4080

ES 2 398 918 T3

gatttgagtc agctaggagg tgactgatgg ctggttggcg tactgttggtg gtaaatatccc 4140
actcgaaatt atcctataag aataatcatc tgatttttaa ggatgcctat aaaacggagc 4200
tgatccatth atcagaaatt gatatthttgt tattagaaac gaccgatatt gtcttgtcca 4260
ctatgctggg aaaacggcta gtggatgaga atgtccttgt catattctgt gatgataaac 4320
gattaccaac agctatgctg atgcctthtt atggctcgtca tgattcaggt ttacagcttg 4380
ggaaacaaat gtcctgggtca gaaacaggtca aatcgcaggt ttggacgacg attattgctc 4440
aaaagattht gaatcaatct tgctatctag gagcatgctc ctatthttgaa aaatcccaat 4500
ctattatgga thttatatcat ggthttggaaa atthttgatcc gagtaatcga gaagggtcatg 4560
cagcgagaat thattthtaat aactthtttg ggaacgattt ctcaagagat ttggagcatc 4620
caatcaatgc aggtctggat tatggttata cthttattatt gagtatgtht gcgctgaaag 4680
tggttggtgc tggatgtatg actcaattht gactcaaaca cgccaatcag thtaatcagt 4740
tcaatthttg tagcgatatt atggaacct thaggcttht ggtggataag atthtttatg 4800
aaaatcgaaa tcagctthtt cccaaaataa agagagagtt atthacttht thttcagata 4860
cathttcata taatggtaaa gagatgtatc tcacgaatat tattagcgat tatactaaaa 4920
aagthgtcaa agctctgaaat aatgaaggga aaggagthcc tgaatthagg atatgagtha 4980
tagatatatg agaatgatac thtatgthtga tatgcccagc gacctgctg aggaacgaaa 5040
agcttatcga aaatthcgga aatthttact tagtgaaggg thtatcatgc atcaatthtc 5100
tathttatag aagthtactgt tgaataatac agctaacaac gccatgattg gtcggctgag 5160
ggagcataat cctcataaag gaaatattac atthactaaca gtcacagaaa aacagthttg 5220
acgaatgatt thtttacatg gtgaaagaaa taatthgtatt gcaactccg atgagagact 5280
tgtatthctt ggggaggctt thgatgaaat thaatthtcc cthattagat gaaccgattc 5340
cattaagagg cgttacaatt cthgtgctcg aagatgtctg tgtatthtca aaaatagthc 5400
aatatthgta caaatatgag gaagatthctg aactthaaat thttgatcac aagatgaaaa 5460
ccatcaaaga atcagaaatc atgctthgta cagatathtt aggaththgat gthtaactcct 5520
caaccathtt aaaatthgatt catgcagatt tagaatctca atthaatgag aaaccgaaag 5580
tgaaatcgat gattgacaaa thggttgcta cgattacaga actgattgct thtgaatgct 5640
tagaaaatga atthagatth gagtatgatg aatcacaaat cthggaatthg atthagthct 5700
taggagtaaa agtagaaacg caaagtgata ctathtttga aaaatgctca gagatacttc 5760
aaatthtcaa atatctcact aagaaaaagt thctthathtt thtcaatagc ggagctthtc 5820
taacaaagga tgaagthgct agthttacaag agtatatac atthgacaaat thaacagthc 5880
thttthttag accacgtgaa ctatatgatt thccgcagta thttthtagat gaagatthatt 5940
thtthaataac thaaaatath gtataa 5966

ES 2 398 918 T3

<210> 518
 <211> 4107
 <212> ADN
 5 <213> Streptococcus pyogenes

<400> 518

```

atggataaga aataactcaat aggcttagat atcggcacia atagcgtcgg atgggcgggtg      60
atcactgatg attataaggt tccgtctaaa aagctcaagg gtctgggaaa tacagaccgc      120
cacggtatca aaaaaaatct tataggggct cttttatttg acagtggaga gacagcggaa      180
gcgactcgtc tcaaacggac agctcgtaga aggtatacac gtcggaagaa tcgtatttgt      240
tatctacagg agattttttc aaatgagatg gcgaaagtag atgatagttt ctttcatcga      300
cttgaagagt cttttttggg ggaagaagac aagaagcatg aacgtcatcc tatttttggg      360
aatatagtag atgaagttgc ttatcatgag aaatatccaa ctatctatca tctgcgaaaa      420
aaattggcag attctactga taaagtggat ttgcgcttaa tctatttggc cttagcgcag      480
atgattaagt ttcgtggtca ttttttgatt gagggagatt taaatcctga taatagtgat      540
gtggacaaac tatttatcca gttggtacaa acctacaatc aattatttga agaaaaccct      600
attaacgcaa gtagagtaga tgctaaagcg attctttctg cacgattgag taaatcaaga      660
cgattagaaa atctcattgc tcagctcccc ggtgagaaga aaaatggatt gtttgggaat      720
ctcattgctt tgtcattggg attgaccctt aattttaaat caaattttga tttggcagaa      780
gatgctaaat tacagctttc aaaagatact tacgatgatg atttagataa tttattggcg      840
caaattggag atcaatatgc tgatttgttt ttggcagcta agaatttatc agatgctact      900
ttactttcag atatcctaag agtaaatagt gaaataacta aggctcccct atcagcttca      960
atgattaagc gctacgatga acatcatcaa gacttgactc ttttaaaagc tttagttcga     1020
caacaacttc cagaaaagta taaagaaatc ttttttgatc aatcaaaaaa cggatatgca     1080
ggttatattg atgggggagc tagccaagaa gaattttata aatttatcaa accaatttta     1140
gaaaaaatgg atgggtactga ggaattattg gcgaaactaa atcgtgaaga tttgctgcgc     1200
aagcaacgga ctttgacaa cggctctatt ccctatcaaa ttcacttggg tgagctgcat     1260
gctattttga gaagacaaga agacttttat ccatttttaa aagacaatcg tgagaagatt     1320
gaaaaaatct tgacttttcg aattccttat tatgttggtc cattggcgcg tggcaatagt     1380
cgttttgcat ggatgactcg gaagtctgaa gaaacaatta ccccatggaa ttttgaagaa     1440
gttgcgata aagggtcttc agctcaatca tttattgaac gcatgacaaa ctttgataaa     1500
aatcttccaa atgaaaaagt actaccaaaa catagtttgc tttatgagta ttttacggtt     1560
tataacgaat tgacaaaagt caaatatggt actgagggaa tgcgaaaacc agcatttctt     1620
tcaggtgaac agaagaaagc cattgttgat ttactcttca aaacaaatcg aaaagtaacc     1680
gttaagcaat taaaagaaga ttatttcaaa aaaatagaat gttttgatag tgttgaaatt     1740
    
```

ES 2 398 918 T3

tcaggagttg aagatagatt taatgcttca ttaggtacct accatgattt gctaaaaatt 1800
 attaaagata aagatTTTTT ggataatgaa gaaaacgaag atatcttaga ggatattggt 1860
 ttaacattga ccttatttga agatagggag atgattgagg aaagacttaa aacatagct 1920
 cacctctttg atgataaggt gatgaaacag cttaaacgtc gccgttatac tggttgggga 1980
 cgtttgtctc gaaaattgat taatggtatt agggataagc aatctggcaa aacaatatta 2040
 gatTTTTTga aatcagatgg ttttgccaat cgcaatttta tgcagctgat ccatgatgat 2100
 agtttgacat ttaaagaaga cattcaaaaa gcacaagtgt ctggacaagg cgatagttta 2160
 catgaacata ttgcaaatTT agctggtagc cctgctatta aaaaaggtat tttacagact 2220
 gtaaaagttg ttgatgaatt ggtcaaagta atggggcggc ataagccaga aaatatcgtt 2280
 attgaaatgg cacgtgaaaa tcagacaact caaaagggcc agaaaaattc gcgtgagcgt 2340
 atgaaacgta ttgaagaagg aataaaaagaa ctaggaagtg atattctaaa ggagtatcct 2400
 gttgaaaaca ctcaattaca aaatgaaaag ctctatctct attatctcca aaatggaaga 2460
 gacatgtatg tggaccaaga attagatatt aatcgtttaa gtgattatga tgtcgatcac 2520
 attgttccac aaagtttctt taaagacgat tcaatagaca ataaggtctt aacgcgttct 2580
 gataaaaatc gtggtaaatc ggataacggt ccaagtgaag aagtagtcaa aaagatgaaa 2640
 aactattgga gacaacttct aaacgccaag ttaatcactc aacgtaagtt tgataattta 2700
 acaaaagctg aacgtggagg tttgagtga cttgataaag ttggTTTTat caaacgcaa 2760
 ttggttgaaa ctcgccaaat cactaagcat gtggcacaaa ttttgatag tcgcatgaat 2820
 actaaatcag atgaaaatga taaacttatt cgagaggtta gagtgattac cttaaaatct 2880
 aaattagttt ctgacttccg aaaagatttc caattctata aagtacgtga gattaacaat 2940
 taccatcatg cccatgatgc gtatcttaat gccgtcgttg gaactgcttt gattaagaaa 3000
 tatccaaaac ttgaatcggg gtttgtctat ggtgattata aagtttatga tgttcgtaaa 3060
 atgattgcta agtctgagca ggaaataggc aaagcaaccg caaaatattt cttttactct 3120
 aatatcatga acttcttcaa aacagaaatt acacttgcaa atggagagat tcgcaaacgc 3180
 cctctaactg aaactaatgg ggaaactgga gaaattgtct gggataaagg gcgagatTTT 3240
 gccacagtgc gcaaagtatt gtccatgcc caagtcaata ttgtcaagaa aacagaagta 3300
 cagacaggcg gattctccaa ggagtcaatt ttaccaaaaa gaaattcgga caagcttatt 3360
 gctcgtaaaa aagactggga tccaaaaaaa tatggtggtt ttgatagtcc aacggtagct 3420
 tattcagtcc tagtggttgc taaggtggaa aaagggaaat cgaagaagtt aaaatccggt 3480
 aaagagttac tagggatcac aataatggaa agaagctctt ttgaaaaaga tccgattgac 3540
 tttttagaag ctaaaggata taaggaagtt agaaaagact taatcattaa actacctaaa 3600
 tatagtcttt ttgagttaga aaacggtcgt aaacggatgc tggctagtgc cggagaattg 3660

ES 2 398 918 T3

caaaaaggaa atgagctagc tctgccaagc aaatatgtga attttttata tttagctagt 3720
cattatgaaa agttgaaggg tagtccagaa gataacgaac aaaaacaatt gtttgtggag 3780
cagcataagc attattttaga tgagattatt gagcaaatca gtgaattttc taagcgtggt 3840
attttagcag atgccaattt agataaagtt cttagtgcot ataacaaaca tagagacaaa 3900
ccaatacgtg aacaagcaga aaatattatt catttattta cgttgacgaa tcttggagct 3960
cccgtgctt ttaaatattt tgatacaaca attgatcgta aacgatatac gtctacaaaa 4020
gaagttttag atgccactct tatccatcaa tccatcactg gtctttatga aacacgcatt 4080
gatttgagtc agctaggagg tgactga 4107

<210> 519

<211> 870

5 <212> ADN

<213> Streptococcus pyogenes

<400> 519

atggctggtt ggcgtactgt tgtggtaaata acccactcga aattatccta taagaataat 60
catctgattt ttaaggatgc ctataaaacg gagctgatcc atttatcaga aattgatatt 120
ttgttattag aaacgaccga tattgtcttg tccactatgc tggtaaaacg gctagtggat 180
gagaatgtcc ttgtcatatt ctgtgatgat aaacgattac caacagctat gctgatgcct 240
ttttatggtc gtcattgattc gagtttacag cttgggaaac aaatgtcctg gtcagaaaca 300
gtcaaatcgc aggtttggac gacgattatt gctcaaaaga ttttgaatca atcttgctat 360
ctaggagcat gctcctattt tgaaaaatcc caatctatta tggatttata tcatggtttg 420
gaaaattttg atccgagtaa tgcagaaggg catgcagcga gaatttattt taatacactt 480
tttgggaacg atttctcaag agatttggag catccaatca atgcaggctt ggattatggt 540
tatactttat tattgagtat gtttgccgct gaagtgggtg tgtctggatg tatgactcaa 600
tttgactca aacacgcaa tcagttaat cagttcaatt ttgctagcga tattatggaa 660
ccatttaggc ctttgggtgga taagattggt tatgaaaatc gaaatcagcc ttttccaaa 720
ataaagagag agttatttac tttgttttca gatacatttt catataatgg taaagagatg 780
tatctcacga atattattag cgattatact aaaaaagttg tcaaagctct gaataatgaa 840
10 gggaaaggag ttcctgaatt taggatatga 870

<210> 520

<211> 327

15 <212> ADN

<213> Streptococcus pyogenes

<400> 520

ES 2 398 918 T3

atgagaatga tactttatggt tgatatgccg acggacactg ctgaggaacg aaaagcttat 60
 cgaaaatttc ggaaatTTTT acttagtgaa gggtttatca tgcatacaatt ttctatttat 120
 agtaagttac tgttgaataa tacagctaac aacgccatga ttggtcggct gagggagcat 180
 aatcctcata aaggaaatat tacattacta acagtcacag aaaaacagtt tgcacgaatg 240
 atttatttac atggtgaaag aaataattgt attgcaaact ccgatgagag acttgtattt 300
 cttggggagg cttttgatga atcttaa 327
 5 <210> 521
 <211> 663
 <212> ADN
 <213> Streptococcus pyogenes
 10 <400> 521
 atgaatctta attttccctt attagatgaa ccgattccat taagaggcgg tacaattctt 60
 gtgctcgaag atgtctgtgt attttcaaaa atagtgcaat attgttacaa atatgaggaa 120
 gattctgaac ttaaattttt tgatcacaag atgaaaacca tcaaagaatc agaaatcatg 180
 cttgtaacag atatttttagg atttgatggt aactcctcaa ccattttaa attgattcat 240
 gcagatttag aatctcaatt taatgagaaa cccgaagtga aatcgatgat tgacaaattg 300
 gttgctacga ttacagaact gattgtcttt gaatgcttag aaaatgaatt agatttagag 360
 tatgatgaaa tcacaatcct ggaattgatt aagtccttag gagtaaaagt agaaacgcaa 420
 agtgatacta tttttgaaaa atgtctagag atacttcaa ttttcaaata tctcactaag 480
 aaaaagttgc ttatttttgt caatagcggg gcttttctaa caaaggatga agtggctagt 540
 ttacaagagt atatatcatt gacaaattta acagttctct ttttagaacc acgtgaacta 600
 tatgattttc cgcagtatat ttagatgaa gattatttct taataactaa aaatattgta 660
 taa 663
 15 <210> 522
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 522
 caacacattc aacagattaa tgaagaatac 30
 25 <210> 523
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 523

5 tccactcacg tacaaatagt gagtgactc 30

<210> 524
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 524

gcccttctaa ttgattacc ttccgaggtg 30

15 <210> 525
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 525

ctcagtcggt actggtgaac cagtttcaat 30

25 <210> 526
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 526

attgtctatt acgacaacat ggaagatgat 30

35 <210> 527
<211> 31
<212> ADN
<213> Streptococcus thermophilus

<400> 527

40 gagtttcttt gtcagactct aacacagccg c 31

<210> 528
<211> 30
<212> ADN

45 <213> Streptococcus thermophilus

<400> 528

50 ttactagagc gtgtcgtaa ccactttaa 30

<210> 529
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 529

5 ttcgttaaag tcacctcgtg ctagcgttg 30

<210> 530
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 530

ataacggtag caaatataaa cctgttactg 30

15 <210> 531
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 531

gaagtagcca tacaagaaga tggatcagca 30

25 <210> 532
<211> 29
<212> ADN
<213> Streptococcus thermophilus

30 <400> 532

atgtcactga gtgtctaagc attgcgtac 29

35 <210> 533
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 533

40 tgaataagca gttcttgacg accaaccgac 30

<210> 534
<211> 30
<212> ADN

45 <213> Streptococcus thermophilus

<400> 534

50 tcaacaattg caacatctta taaccact 30

<210> 535
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 535

5 ttacgttga aaagaatc aaatcaatga 30

<210> 536
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 536

gctctacgac ttctccacg agttcctgcc 30

15 <210> 537
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 537

aacacagcaa gacaagagga tgatgctatg 30

25 <210> 538
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 538

aagtattga tgacctctac aatggttat 30

35 <210> 539
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 539

40 aataattat ggtatagctt aatatcattg 30

<210> 540
<211> 30
<212> ADN
<213> Streptococcus thermophilus

45 <400> 540

50 aatcaatacg acaagagtta aaatggtcct 30

<210> 541
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 541

5 aatcgtcaa attctgtttt aggtacatt 30

<210> 542
<211> 30
<212> ADN
10 <213> Streptococcus thermophilus

<400> 542

15 aatgacgagg agctattggc acaactaca 30

<210> 543
<211> 30
<212> ADN
20 <213> Streptococcus thermophilus

<400> 543

aattaagggc atagaaaggg agacaacatg 30

25 <210> 544
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 544

acaattcttc atccggtaac tgctcaagt 30

<210> 545
35 <211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 545

40 acacttgga ggcttattac tcaacagcga 30

<210> 546
<211> 30
45 <212> ADN
<213> Streptococcus thermophilus

<400> 546

50 ataaactatg aaattttata attttaaga 30

<210> 547
<211> 30
<212> ADN
55 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 547

5 ataactgaag gataggagct tgtaaagtct 30

 <210> 548
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus

10 <400> 548

 ataatgccgt tgaattacac ggcaagtca 29

15 <210> 549
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

20 <400> 549

 caaccaacgg taacagctac ttttacagt 30

25 <210> 550
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

30 <400> 550

 catagagtgg aaaactagaa acagattcaa 30

35 <210> 551
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus

40 <400> 551

40 cgacacaaga acgtatgcaa gaggtaag 29

 <210> 552
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

45 <400> 552

50 cgatatttaa aatcatttcc ataacttcat 30

 <210> 553
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 553
 5 cgattgaca atctgctgac cactgttatc 30
 <210> 554
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 554
 ctgttccttg ttcttttgt gtatctttc 30
 15 <210> 555
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 555
 gagcgagctc gaaataatct taattacaag 30
 25 <210> 556
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 556
 gcagtatcag caagcaagct gtagttact 30
 35 <210> 557
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 557
 40 gctggcgagg aaacgaacaa ggcctcaaca 30
 <210> 558
 <211> 30
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 558
 50 gcttagctgt ccaatccacg aacgtggatg 30
 <210> 559
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 559

5 ggcgtcccaa tcctgattaa tacttactcg 30

<210> 560
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 560

gttcgctagc gtcattggt aacgtattta 30

15 <210> 561
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 561

tctatatcga ggcaactaa caattatgct 30

25 <210> 562
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 562

tgcatcgagc acgttcgagt ttaccgtttc 30

35 <210> 563
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 563

40 tgttgacag caaatcaaga ttcgaattgt 30

<210> 564
<211> 30
<212> ADN
<213> Streptococcus thermophilus

45 <400> 564

ttcattcttc cgttttggt tgccaatcct 30

50 <210> 565
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 565

5 tgacttagcg aatttaatcg ctaagatat 30

<210> 566
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 566

tttatacttt atcttttaa agaattgatt 30

15 <210> 567
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 567

cctaaatca tttcaacga gttgcgatac 30

25 <210> 568
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 568

aataaattgc tatgatacag cgtaccgata 30

35 <210> 569
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 569

40 tgctctctat gcgattggac gtctgtctaa 30

<210> 570
<211> 30
<212> ADN
<213> Streptococcus thermophilus

45 <400> 570

aagaaagata agaaaaaagt aacactactt 30

50 <210> 571
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 571
 5 tctctttcca tcggtactgg tataatctcat 30
 <210> 572
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 572
 attggtagcc aagtaaataat caccattgat 30
 15 <210> 573
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 573
 ttcttcaaat tcaccgactg caaaattaca 30
 25 <210> 574
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 574
 gcttcctaag tgcatgaaaa tcgcaaacgg 30
 35 <210> 575
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 575
 40 tataacctgtc tatgtaaggg aatttaactc 30
 <210> 576
 <211> 30
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 576
 50 ggtgtaggtg ctgttggtaa gttgtttaat 30
 <210> 577
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 577

5 gtgaaacagg ttatcaaaaa acgtatattg 30

<210> 578
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 578

ttattcttgg aattattaca gaccctacta 30

15 <210> 579
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 579

gcttcatta tatcacttac tcataaatct 30

25 <210> 580
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 580

taatcacccc ttttctagc tcttgattga 30

35 <210> 581
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 581

40 caagcagtg aaagtggtt taaatgtaa 30

<210> 582
<211> 30
<212> ADN
<213> Streptococcus thermophilus

45 <400> 582

aacccgctg gttatggct tgaggagtgt 30

50 <210> 583
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 583

5 atattaatag cgattctatg ctacaacgtg 30

<210> 584
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 584

tcattctcta agtaaatacc actgtcaggg 30

15 <210> 585
<211> 29
<212> ADN
<213> Streptococcus thermophilus

20 <400> 585

ttttcgaaa gtaagcgaag ctctacgtg 29

25 <210> 586
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 586

ttctgtagcc actccgtgga tgccttcagc 30

35 <210> 587
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 587

40 ttcttagtt cggacaccct caacacctat 30

<210> 588
<211> 30
<212> ADN
<213> Streptococcus thermophilus

45 <400> 588

gctttgattg gacggaaaat ggtatccctg 30

50 <210> 589
<211> 31
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

	<400> 589	
5	ttctcatct ttctccgctt ttgctagact t	31
	<210> 590	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
10	<400> 590	
	ttagaccaga tggacagata ttctcatcg	30
15	<210> 591	
	<211> 31	
	<212> ADN	
	<213> Streptococcus thermophilus	
20	<400> 591	
	tcatcagagt caacaatcac gggaaagacc t	31
	<210> 592	
25	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 592	
30	acactcatcc ttatcctgta gttcaaaaca	30
	<210> 593	
	<211> 30	
35	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 593	
40	cagcactagc cgcaagccct tgtatattaa	30
	<210> 594	
	<211> 30	
	<212> ADN	
45	<213> Streptococcus thermophilus	
	<400> 594	
50	tagaaatcaa ggaacttga tgaaaagtaa	30
	<210> 595	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	

ES 2 398 918 T3

	<400> 595	
5	atatgaaagg gaaatgatat gaagaatgaa	30
	<210> 596	
	<211> 31	
	<212> ADN	
	<213> Streptococcus thermophilus	
10	<400> 596	
	ttttgggata caacacgcag tcgttgactt g	31
15	<210> 597	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
20	<400> 597	
	gtttgagatg ccaatgtttt tcaatccttg	30
	<210> 598	
25	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 598	
30	gtatcaaaag acgattcat gaagcgagct	30
	<210> 599	
	<211> 30	
35	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 599	
40	aaaacaatt gaaattcata atcagcgctt	30
	<210> 600	
	<211> 29	
	<212> ADN	
45	<213> Streptococcus thermophilus	
	<400> 600	
	gctttaacg ttttaagaga ataccctct	29
50	<210> 601	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	

ES 2 398 918 T3

<400> 601
 5 gtgacgctgc aatgactgc catagtaatt 30
 <210> 602
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 602
 atactggat atagtaattc atacttcac 30
 15 <210> 603
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 603
 ttggtttcat atttactcct ttgtgtttg 30
 <210> 604
 25 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 604
 30 ctgatttggc cttgttcttt tgcctcttt 30
 <210> 605
 <211> 30
 35 <212> ADN
 <213> Streptococcus thermophilus
 <400> 605
 40 gcagcagtg agaacttag cgccagtg 30
 <210> 606
 <211> 30
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 606
 tgctactatg aaggacgctg ttgatactt 30
 50 <210> 607
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 607

5 tcttctttaa tctttttaa cgtaacggt 30

<210> 608
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 608

gtatccatta atagtagc attctatca 30

15 <210> 609
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 609

attcattaat atctgcaagg atgtctggt 30

<210> 610
<211> 30
<212> ADN
<213> Streptococcus thermophilus

25 <400> 610

30 gagaaagtag cccattcggc ccattcgggg 30

<210> 611
<211> 30
<212> ADN
<213> Streptococcus thermophilus

35 <400> 611

40 tacttgagtt agctctggaa gtcattatc 30

<210> 612
<211> 30
<212> ADN
<213> Streptococcus thermophilus

45 <400> 612

ctgcattgt aaccatgact tcttcgctgt 30

50 <210> 613
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 613

5 aattgtcat cgacatctac caacgccag 30

<210> 614
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 614

ataaaattat gccacgtttt ggcactagat 30

15 <210> 615
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 615

atgtctctga ggctgtagta attactgt 30

<210> 616
<211> 30
<212> ADN
<213> Streptococcus thermophilus

25 <400> 616

30 cttaaagag ttgattaagt gcgttactgt 30

<210> 617
<211> 30
<212> ADN
<213> Streptococcus thermophilus

35 <400> 617

40 aaatgggta tgctgtcaa tatgcgtccc 30

<210> 618
<211> 30
<212> ADN
<213> Streptococcus thermophilus

45 <400> 618

aaactgaaaa caacacagac aattcaaca 30

50 <210> 619
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 619
 5 gccc aaaatg ctgacg ttt gaatgacggc 30
 <210> 620
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 620
 atgaagaacg tgattcacct acggtatgct 30
 15 <210> 621
 <211> 31
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 621
 gctttgcag aattgtctcc agtgccgatt t 31
 25 <210> 622
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 622
 tgactctat tgattgcttc atctttatta 30
 35 <210> 623
 <211> 31
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 623
 40 cttcaagat actcatcaac cattgatgctc a 31
 <210> 624
 <211> 30
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 624
 50 ctatgtcttt actgttcttc caaaaccacc 30
 <210> 625
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 625
 5 tgctacgtgc tctgtacggg cgctatcagc 30
 <210> 626
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 626
 cgtggcagcg tggcgggtt taatagcccg 30
 15 <210> 627
 <211> 29
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 627
 aagcccaagt cagagcatcc gtccaagcc 29
 25 <210> 628
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 628
 attgggttc ggtaagaact aaacatacca 30
 35 <210> 629
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 629
 40 cacaaaataa ttcggtagt ttactaact 30
 <210> 630
 <211> 29
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 630
 50 tttgaccgtt tatttagacg tgctaaagt 29
 <210> 631
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 631
 5 cttcacctca aatcttagag ctggactaaa 30
 <210> 632
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 632
 atgtctgaaa aataaccgac catcattact 30
 15 <210> 633
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 633
 gaagctcatc atgtaagc taaaacctat 30
 25 <210> 634
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 634
 tagtctaaat agattcttg caccattgta 30
 35 <210> 635
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 635
 40 attcgtgaaa aaatcgtg aaataggcaa 30
 <210> 636
 <211> 30
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 636
 50 tctaggctca tctaaagata aatcagtagc 30
 <210> 637
 <211> 31
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 637
 5 taaaaacatg gggcggcggg aatagtgtaa g 31
 <210> 638
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 638
 acaaccagca aagagagcgc cgacaacatt 30
 15 <210> 639
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 639
 tataacacag gtttagagga tgttatact 30
 25 <210> 640
 <211> 31
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 640
 ctagaagctc aagcggtaaa agtggatggc g 31
 35 <210> 641
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 641
 40 cttgagggc aagccctcgc cgttcatt 30
 <210> 642
 <211> 30
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 642
 50 aactaccaag caaatcagca atcaataagt 30
 <210> 643
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 643

5 ctataagtga caatcagcgt agggaatcag 30

<210> 644
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 644

atcagtgcgg tatattacc ctagacgcta 30

15 <210> 645
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 645

aacagttact attaatcacg attccaacgg 30

25 <210> 646
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 646

aattagggcg tcttcctta ttccgtggtt 30

35 <210> 647
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 647

40 atagctcat tgcgctttt aattgacct 30

<210> 648
<211> 30
<212> ADN
<213> Streptococcus thermophilus

45 <400> 648

aacaacaaag caaatacaac agtaacaacc 30

50 <210> 649
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 649

5 ctaaactacg ttgaaggtc tcaactccgt 30

<210> 650
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 650

gaggttgaat agtgagtgca ccatgtttgt 30

15 <210> 651
<211> 30
<212> ADN
<213> Streptococcus thermophilus

20 <400> 651

agtagagaga ccagcacact actgtactac 30

25 <210> 652
<211> 30
<212> ADN
<213> Streptococcus thermophilus

30 <400> 652

cttcgcacga aagtttatta gacaactcgc 30

35 <210> 653
<211> 30
<212> ADN
<213> Streptococcus thermophilus

<400> 653

40 tgatagagct agaattgtct ttttaccga 30

<210> 654
<211> 30
<212> ADN
<213> Streptococcus thermophilus

45 <400> 654

agatactctt gctcgcctct gaacaaccag 30

50 <210> 655
<211> 30
<212> ADN
<213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 655
 5 ggtgaaaaag gttcactgta cgagtactta 30
 <210> 656
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 10 <400> 656
 tcaatgagtg gtatccaaga cgaaaactta 30
 15 <210> 657
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 20 <400> 657
 ccttgctg gctctccata cgcccatata 30
 25 <210> 658
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus
 30 <400> 658
 tgttgggaa accgcagtag ccatgattaa 30
 35 <210> 659
 <211> 32
 <212> ADN
 <213> Streptococcus thermophilus
 <400> 659
 40 acagagtaca atattgctt cattggagac ac 32
 <210> 660
 <211> 30
 <212> ADN
 45 <213> Streptococcus thermophilus
 <400> 660
 50 ctcatattcg ttagtgctt ttgcataaa 30
 <210> 661
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

ES 2 398 918 T3

<400> 661

5 agaactttat caagataaaa ctactttaa 30

<210> 662
<211> 30
<212> ADN
<213> Streptococcus thermophilus

10 <400> 662

atagtattaa ttccattgaa aaataattgt 30

15 <210> 663
<211> 36
<212> ADN
<213> Streptococcus thermophilus

20 <400> 663

gcttctagc tcgctataat taccattcc tagaaa 36

25 <210> 664
<211> 40
<212> ADN
<213> Streptococcus thermophilus

<400> 664

30 tcaaaatag ttattacct gtattcata attcaattaa 40

<210> 665
<211> 39
<212> ADN
<213> Streptococcus thermophilus

<400> 665

40 ccacttgctg tgtacatcct accagtccg cctatgatg 39

<210> 666
<211> 63
<212> ADN
<213> Streptococcus thermophilus

45 <400> 666

caaggacagt tattgatttt ataactacta tgggggtata aaaacgtcaa aatttcattt 60

50 gag 63

<210> 667
<211> 41

ES 2 398 918 T3

<212> ADN
 <213> Streptococcus thermophilus

<400> 667

5 ttgattcaac ataaaaagcc agtcaattg aactggctt t 41

<210> 668
 <211> 2317
 <212> ADN
 <213> Streptococcus thermophilus

10

<400> 668

caaggacagt tattgatttt ataatcacta tgtgggtata aaaacgtcaa aatttcattt 60
 gaggtttttg tactctcaag atttaagtaa ctgtacaact caacaattgc aacatcttat 120
 aaccacttg tttttgtact ctcaagattt aagtaactgt acaactgttt gacagcaaat 180
 caagattcga atttgttttt tgtactctca agatttaagt aactgtacaa caatgacgag 240
 gagctattgg cacaacttac agtttttgta ctctcaagat ttaagtaact gtacaaccga 300
 tttgacaatc tgctgaccac tgttatcggt tttgtactct caagatttaa gtaactgtac 360
 aacacacttg gcaggcttat tactcaacag cgagtttttg tactctcaag atttaagtaa 420
 ctgtacaacc tgttccttgt tcttttgttg tatcttttcg tttttgtact ctcaagattt 480
 aagtaactgt acaacttcat tcttccggtt ttgtttgcca atcctgtttt tgtactctca 540
 agatttaagt aactgtacaa cgctggcgag gaaacgaaca aggcctcaac agtttttgta 600
 ctctcaagat ttaagtaact gtacaacat agagtggaaa actagaaaca gattcaagtt 660
 tttgtactct caagatttaa gtaactgtac aacataatgc cgttgaatta cacggcaagg 720
 tcagtttttg tactctcaag atttaagtaa ctgtacaacg agcgagctcg aaataatctt 780
 aattacaagg tttttgtact ctcaagattt aagtaactgt acaacgttcg ctagcgtcat 840
 gtggtaacgt atttagtttt tgtactctca agatttaagt aactgtacaa cggcgtccca 900
 atcctgatta atacttactc ggtttttgta ctctcaagat ttaagtaact gtacaacaac 960
 acagcaagac aagaggatga tgctatgggt tttgtactct caagatttaa gtaactgtac 1020
 aaccgacaca agaacgtatg caagagttca aggtttttgt actctcaaga ttttaagtaac 1080
 tgtacaacac aattcttcat ccggtaactg ctcaagtgggt tttgtactc tcaagattta 1140
 agtaactgta caacaattaa gggcatagaa aggagacaa catggttttt gtactctcaa 1200
 gatttaagta actgtacaac cgatatttaa aatcattttc ataacttcat gttttgtac 1260
 tctcaagatt taagtaactg tacaacgcag tatcagcaag caagctgtta gttactgttt 1320
 15 ttgtactctc aagatttaag taactgtaca acataaacta tgaaatttta taatttttaa 1380

ES 2 398 918 T3

gagtttttgt actctcaaga ttttaagtaac tgtacaacaa taatttatgg tatagcttaa 1440
 tatcattggg ttttgtactc tcaagattta agtaactgta caactgcatc gagcacgttc 1500
 gagtttaccg tttcgttttt gtactctcaa gatttaagta actgtacaac tctatatcga 1560
 ggtcaactaa caattatgct gtttttgtac tctcaagatt taagtaactg tacaacaatc 1620
 gttcaaattc tgtttttaggt acatttgttt ttgtactctc aagatttaag taactgtaca 1680
 acaatcaata cgacaagagt taaaatggtc ttgtttttgt actctcaaga ttttaagtaac 1740
 tgtacaacgc ttagctgtcc aatccacgaa cgtggatggg ttttgtactc tcaagattta 1800
 agtaactgta caaccaacca acggtaacag ctacttttta cagtgttttt gtactctcaa 1860
 gatttaagta actgtacaac ataactgaag gataggagct tgtaaagtct gtttttgtac 1920
 tctcaagatt taagtaactg tacaactaat gctacatctc aaaggatgat cccagagttt 1980
 ttgtactctc aagatttaag taactgtaca acaagtagtt gatgacctct acaatggttt 2040
 atgtttttgt actctcaaga ttttaagtaac tgtacaacac ctagaagcat ttgagcgtat 2100
 attgattggg ttttgtactc tcaagattta agtaactgta caacaatttt gcccttctt 2160
 tgccccttga ctaggttttt gtactctcaa gatttaagta actgtacaac accattagca 2220
 atcatttggtg cccattgagt gtttttgtac tctcaagatt taagtaactg tacagtttga 2280
 ttcaacataa aaagccagtt caattgaact tggcttt 2317

<210> 669

<211> 30

5 <212> ADN

<213> Streptococcus thermophilus

<400> 669

10 tcaacaattg caacatctta taaccactt

30

<210> 670

<211> 2317

<212> ADN

15 <213> Streptococcus thermophilus

<400> 670

ES 2 398 918 T3

caaggacagt	tattgatttt	ataatcacta	tgtgggtata	aaaacgtcaa	aatttcattt	60
gaggtttttg	tactctcaag	atttaagtaa	ctgtacaact	tacgtttgaa	agaatatca	120
aatcaatgag	tttttgact	ctcaagattt	aagtaactgt	acaactgttt	gacagcaaat	180
caagattcga	attgtgtttt	tgtactctca	agatttaagt	aactgtacaa	caatgacgag	240
gagctattgg	cacaacttac	agtttttgta	ctctcaagat	ttaagtaact	gtacaaccga	300
tttgacaatc	tgctgaccac	tgttatcgtt	tttgactct	caagatttaa	gtaactgtac	360
aacacacttg	gcaggcttat	tactcaacag	cgagtttttg	tactctcaag	atttaagtaa	420
ctgtacaacc	tgttccttgt	tcttttgttg	tatcttttcg	tttttgact	ctcaagattt	480

ES 2 398 918 T3

aagtaactgt acaacttcat tcttccgttt ttgtttgcga atcctgtttt tgtactctca 540
agatttaagt aactgtacaa cgctggcgag gaaacgaaca aggcctcaac agtttttgta 600
ctctcaagat ttaagtaact gtacaacat agagtggaaa actagaaaca gattcaagtt 660
tttgactct caagatttaa gtaactgtac aacataatgc cgttgaatta cacggcaagg 720
tcagtttttg tactctcaag atttaagtaa ctgtacaacg agcgagctcg aaataatctt 780
aattacaagg tttttgtact ctcaagattt aagtaactgt acaacgttcg ctagecgtcat 840
gtggtaacgt atttagtttt tgtactctca agatttaagt aactgtacaa cggcgtccca 900
atcctgatta atacttactc ggtttttgta ctctcaagat ttaagtaact gtacaacaac 960
acagcaagac aagaggatga tgctatggtt tttgtactct caagatttaa gtaactgtac 1020
aaccgacaca agaacgtatg caagagttca aggtttttgt actctcaaga ttttaagtaac 1080
tgtacaacac aattcttcat ccggttaactg ctcaagtggg ttttgtactc tcaagattta 1140
agtaactgta caacaattaa gggcatagaa agggagacaa catggttttt gtactctcaa 1200
gatttaagta actgtacaac cgatatttaa aatcattttc ataacttcat gttttgtac 1260
tctcaagatt taagtaactg tacaacgcag tatcagcaag caagctgtta gttactgttt 1320
ttgtactctc aagatttaag taactgtaca acataaacta tgaattttta taatttttaa 1380
gagttttgt actctcaaga ttttaagtaac tgtacaacaa taatttatgg tatagcttaa 1440
tatcattggg ttttgtactc tcaagattta agtaactgta caactgcac gagcacgttc 1500
gagtttaccg tttcgttttt gtactctcaa gatttaagta actgtacaac tctatatcga 1560
ggccaactaa caattatgct gttttgtac tctcaagatt taagtaactg tacaacaatc 1620
gttcaaattc tgttttaggt acatttggtt ttgtactctc aagatttaag taactgtaca 1680
acaatcaata cgacaagagt taaaatggc ttgtttttgt actctcaaga ttttaagtaac 1740
tgtacaacgc ttagctgtcc aatccacgaa cgtggatggg ttttgtactc tcaagattta 1800
agtaactgta caaccaacca acggtaacag ctacttttta cagtgttttt gtactctcaa 1860
gatttaagta actgtacaac ataactgaag gataggagct tgtaaagtct gttttgtac 1920
tctcaagatt taagtaactg tacaactaat gctacatctc aaaggatgat cccagagttt 1980
ttgtactctc aagatttaag taactgtaca acaagtagtt gatgacctct acaatggttt 2040
atgtttttgt actctcaaga ttttaagtaac tgtacaacac ctagaagcat ttgagcgtat 2100
attgattggg ttttgtactc tcaagattta agtaactgta caacaatttt gcccttctt 2160
tgccccttga ctaggttttt gtactctcaa gatttaagta actgtacaac accattagca 2220
atcatttggt cccattgagt gttttgtac tctcaagatt taagtaactg tacagtttga 2280
ttcaacataa aaagccagtt caattgaact tggcttt 2317

ES 2 398 918 T3

	<210> 671	
	<211> 30	
	<212> ADN	
5	<213> Streptococcus thermophilus	
	<400> 671	
10	ttacgttga aaagaatataaatcaatga	30
	<210> 672	
	<211> 19	
	<212> ADN	
15	<213> Secuencia Artificial	
	<220>	
	<223> Cebador Sintético	
	<400> 672	
20	caaatggata gagaaacgc	19
	<210> 673	
	<211> 21	
25	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Cebador Sintético	
30	<400> 673	
	ctgataaggt gtcgtgtc c	21
35	<210> 674	
	<211> 24	
	<212> ADN	
	<213> Secuencia Artificial	
40	<220>	
	<223> Cebador Sintético	
	<400> 674	
45	ggagcagatg gaatacaaga aagg	24
	<210> 675	
	<211> 22	
	<212> ADN	
50	<213> Secuencia Artificial	
	<220>	
	<223> Cebador Sintético	
55	<400> 675	

gagagactag gttgtctcag ca 22

5 <210> 676
 <211> 26
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <223> Cebador Sintético

<400> 676

15 acaacaaca gagaagtatc tcattg 26

20 <210> 677
 <211> 25
 <212> ADN
 <213> Secuencia Artificial

25 <220>
 <223> Cebador Sintético

<400> 677

30 aacgagtaca ctactattt gtacg 25

35 <210> 678
 <211> 24
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador Sintético

40 <400> 678

45 cttcctca tcctcgctt gggt 24

50 <210> 679
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

55 <400> 679

ttctgtagt ggtttagtc aaacagatg 30

60 <210> 680
 <211> 30
 <212> ADN
 <213> Streptococcus thermophilus

65 <400> 680

ES 2 398 918 T3

	ttctgtagt ggtttagtc aaacagatgt	30
	<210> 681	
	<211> 30	
5	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 681	
10	ttctgtagt ggatttagtc aaacagatgt	30
	<210> 682	
	<211> 30	
	<212> ADN	
15	<213> Streptococcus thermophilus	
	<400> 682	
20	ttctgtagt ggtttagtc aaacagatgt	30
	<210> 683	
	<211> 30	
	<212> ADN	
	<213> Streptococcus thermophilus	
25	<400> 683	
30	ttctgtagt ggtttagtc aaacagatgt	30
	<210> 684	
	<211> 23	
	<212> ADN	
	<213> Streptococcus thermophilus	
35	<400> 684	
40	tctgtagtg gatttagtca aac	23
	<210> 685	
	<211> 23	
	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 685	
45	tctgtagtg gatttagtca aac	23
	<210> 686	
	<211> 26	
50	<212> ADN	
	<213> Streptococcus thermophilus	
	<400> 686	
55	ggtagtggat ttagtcaaac agatgt	26

ES 2 398 918 T3

<210> 687
 <211> 26
 <212> ADN
 5 <213> Streptococcus thermophilus

 <400> 687

 10 ggtagtggt ttagtcaaac agatgt 26

 <210> 688
 <211> 28
 <212> ADN
 <213> Streptococcus thermophilus
 15
 <400> 688

 ttctgtagt ggatttagtc aaacagat 28

 20 <210> 689
 <211> 28
 <212> ADN
 <213> Streptococcus thermophilus

 25 <400> 689

 ttctggcagt ggtttagtc aaacagat 28

 <210> 690
 30 <211> 36
 <212> ARN
 <213> Streptococcus thermophilus

 <400> 690
 35
 guuuuagagc uguguuuuu cgaauuguuc caaac 36

 <210> 691
 <211> 222
 40 <212> ADN
 <213> Secuencia Artificial

 <220>
 <223> Constructo nucleótido sintético
 45
 <400> 691

 caggaaacag ctatgaccat gattacgcca agcttggtac cgagctcgga tccactagta 60
 acggccgcca gtgtgctgga attcgccctt aagggcgaat tctgcagata tccatcacac 120
 tggcggccgc tcgagcatgc atctagaggg cccaattcgc cctatagtga gtcgtattac 180
 aattcactgg ccgtcgtttt acaacgctcgt gactgggaaa ac 222

 50 <210> 692

ES 2 398 918 T3

<211> 84
<212> ADN
<213> Secuencia Artificial

5 <220>
<223> Cebador Sintético

<400> 692

tccactcacg tacaaatagt gagtgactc gtttttgtat tctcaagatt taagtaactg 60

10 tacagtttga ttcaacataa aaag 84

<210> 693
<211> 8
<212> PRT
15 <213> Streptococcus thermophilus

<400> 693

Asp Cys Phe Cys Thr Leu Lys Ile
1 5

20 <210> 694
<211> 21
<212> PRT
<213> Streptococcus thermophilus

25 <400> 694

Val Thr Val Gln Leu Leu Thr Ser Ser Val His Arg Leu Lys Cys Phe
1 5 10 15

Cys Thr Leu Lys Ile
20

30 <210> 695
<211> 16
<212> PRT
<213> Streptococcus thermophilus

35 <400> 695

Val Thr Val Gln Pro Pro Leu Ser Met Glu Arg Tyr Pro Ser Ser Phe
1 5 10 15

40 <210> 696
<211> 10
<212> PRT
<213> Streptococcus thermophilus

45 <400> 696

Lys Ile Val Phe Val Leu Ser Arg Phe Lys
1 5 10

ES 2 398 918 T3

<210> 697
<211> 4
<212> PRT
5 <213> Streptococcus thermophilus

<400> 697

Leu Tyr Asn Phe
1
10
<210> 698
<211> 16
<212> PRT
<213> Streptococcus thermophilus
15
<400> 698

Arg His Gln Ser Thr Ala Leu Asn Val Phe Val Leu Ser Arg Phe Lys
1 5 10 15
20 <210> 699
<211> 15
<212> PRT
<213> Streptococcus thermophilus
25 <400> 699

Leu Tyr Asn His Leu Phe Arg Trp Lys Gly Ile Leu Leu Val Phe
1 5 10 15
30 <210> 700
<211> 22
<212> PRT
<213> Streptococcus thermophilus
35 <400> 700

Arg Leu Phe Leu Tyr Ser Gln Asp Leu Ser Asn Cys Thr Thr Phe Asn
1 5 10 15

Val Ile Ser Pro Pro Pro
 20
40 <210> 701
<211> 22
<212> PRT
<213> Streptococcus thermophilus

<400> 701

ES 2 398 918 T3

Met Phe Leu Tyr Ser Gln Asp Leu Ser Asn Cys Thr Thr Thr Ser Phe
 1 5 10 15

Asp Gly Lys Val Ser Phe
 20

<210> 702
 <211> 143
 5 <212> ADN
 <213> Streptococcus thermophilus

<400> 702

agattgtttt tgtactctca agatttaagt aactgtacaa cttttaacgt catcagttcca 60

ccgccttaaa tgtttttgta ctctcaagat ttaagtaact gtacaaccac ctctttcgat 120

10 ggaaaggtat ccttctagtt ttt 143

<210> 703
 <211> 143
 <212> ADN
 15 <213> Streptococcus thermophilus

<400> 703

tctaacaaaa acatgagagt tctaaattca ttgacatggt gaaaattgca gtagtcaggt 60

ggcgggaattt acaaaaaacat gagagttcta aattcattga catgttggtg gagaaagcta 120

20 cctttccata ggaagatcaa aaa 143

<210> 704
 <211> 6
 <212> PRT
 25 <213> Streptococcus thermophilus

<400> 704

Leu Asn Asn Lys Tyr Glu
 1 5

<210> 705
 <211> 15
 <212> PRT
 30 <213> Streptococcus thermophilus

<400> 705

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

<210> 706
 40 <211> 5

ES 2 398 918 T3

<212> PRT
<213> Streptococcus thermophilus

<400> 706

5

Ile Asn Lys Tyr Glu
1 5

<210> 707

<211> 16

10

<212> PRT
<213> Streptococcus thermophilus

<400> 707

Ser Lys Leu Leu Gln Val Val Val Glu Lys Ser Pro Phe Thr Asp Lys
1 5 10 15

<210> 708

<211> 8

20

<212> PRT
<213> Streptococcus thermophilus

<400> 708

Ser Gln Lys Gln Val Arg Leu Ile
1 5

25

<210> 709

<211> 21

<212> PRT

30

<213> Streptococcus thermophilus

<400> 709

Thr Val Thr Cys Ser Lys Val Asp Asp Thr Trp Arg Arg Leu His Lys
1 5 10 15

Gln Val Arg Leu Ile
20

35

<210> 710

<211> 17

<212> PRT

<213> Streptococcus thermophilus

40

<400> 710

Thr Val Thr Cys Gly Gly Arg Glu Ile Ser Leu Tyr Gly Glu Leu Lys
1 5 10 15

Gln

<210> 711

ES 2 398 918 T3

<211> 14
<212> PRT
<213> Streptococcus thermophilus

5 <400> 711

Ile Thr Lys Thr Ser Glu Leu Asn Leu Tyr Ser Tyr Leu Lys
1 5 10

10 <210> 712
<211> 29
<212> PRT
<213> Streptococcus thermophilus

15 <400> 712

Asp Val Ala Lys Phe Thr Lys Thr Ser Glu Leu Asn Leu Tyr Ser Tyr
1 5 10 15

Leu Trp Arg Lys Arg His Phe Pro Ile Arg Arg Thr Lys
20 25

REIVINDICACIONES

1. Un método para modular la resistencia de una célula bacteriana frente a un ácido nucleico diana o producto de transcripción de un bacteriófago del mismo que comprende las etapas de:
- 5 (i) identificar uno o más pseudo espaciadores CRISPR en un genoma de bacteriófago que sea capaz de proporcionar resistencia al ácido nucleico diana o producto de transcripción del mismo;
- (ii) preparar un ácido nucleico recombinante que comprenda un gen *cas1* y al menos dos repeticiones CRISPR junto con dichos uno o más pseudo espaciadores CRISPR identificados; e
- (iii) introducir dicho ácido nucleico recombinante en dicha célula bacteriana para así volver a la célula bacteriana resistente a dicho ácido nucleico diana o producto de transcripción del mismo.
- 10 2. El método de acuerdo con la reivindicación 1, en el que el ácido nucleico diana en el bacteriófago es una secuencia de ácido nucleico altamente conservada.
3. El método de acuerdo con la reivindicación 1 ó reivindicación 2, en el que el ácido nucleico diana en el bacteriófago codifica un proteína de especificidad a hospedante.
- 15 4. El método de acuerdo con cualquiera de las reivindicaciones precedentes, en el que el ácido nucleico diana en el bacteriófago codifica una proteína que es esencial para la supervivencia, la replicación o el crecimiento del bacteriófago.
5. El método de acuerdo con la reivindicación 4, en el que el ácido nucleico diana en el bacteriófago codifica una helicasa, una primasa, una proteína estructural de cabeza o cola, una proteína con un dominio conservado (por ejemplo, holin, lisina y otras) o una secuencia conservada entre genes importantes de fago.
- 20 6. Una célula obtenida u obtenible mediante un método de cualquiera de las reivindicaciones precedentes.
7. El uso de una célula de acuerdo con la reivindicación 6, para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.
8. Un cultivo celular que comprende una célula de acuerdo con la reivindicación 6, para modular la resistencia de una célula frente a un ácido nucleico diana o un producto de transcripción del mismo.
- 25 9. El cultivo celular de acuerdo con la reivindicación 8, en el que dicho cultivo es un cultivo iniciador o un cultivo probiótico.
10. Un producto alimenticio o comida que comprende el cultivo de acuerdo con la reivindicación 8 o reivindicación 9.
- 30 11. Un proceso para preparar un producto alimenticio o comida que comprende el uso del cultivo de acuerdo con la reivindicación 8 o reivindicación 9.
12. El uso del cultivo de acuerdo con la reivindicación 8 o reivindicación 9 para preparar un producto alimenticio.

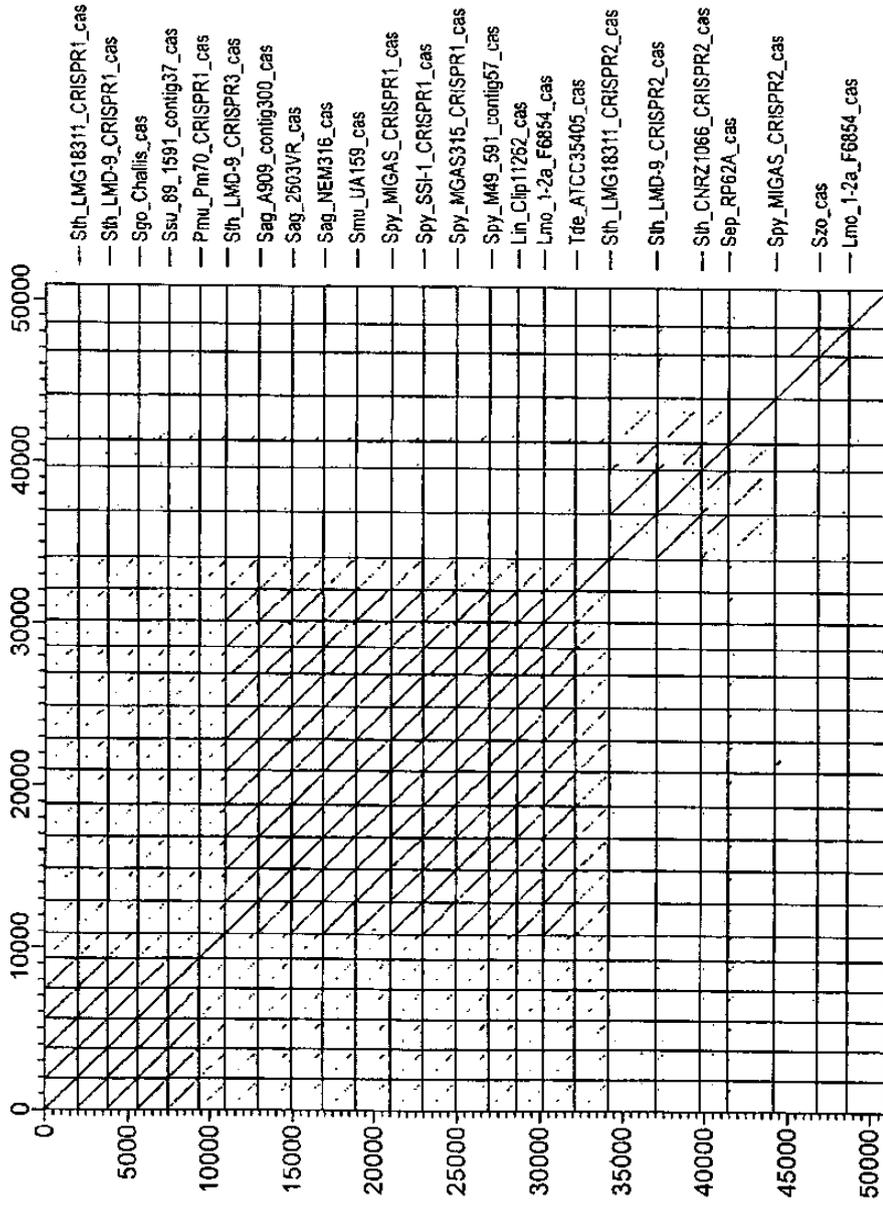


FIG. 2A

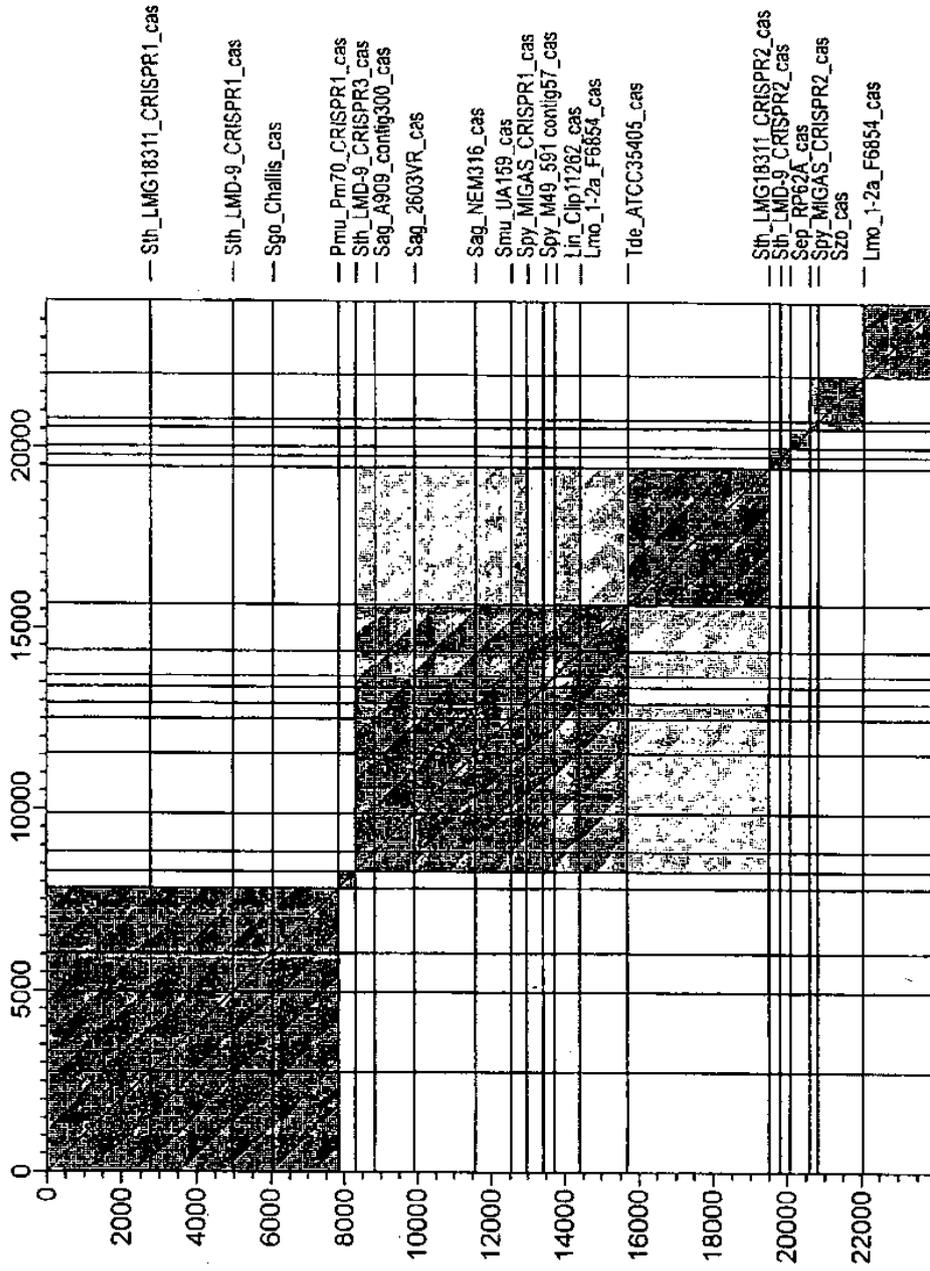
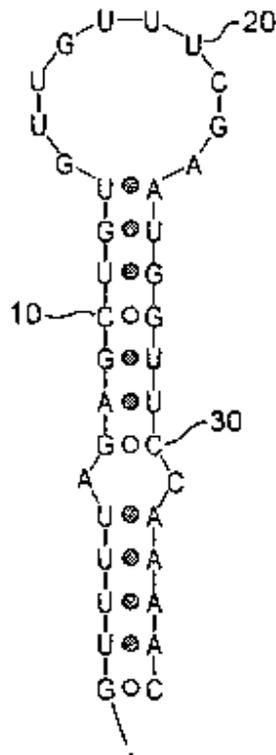


FIG. 2B

Número secuencia interespiadora	Fagos													Número de parejas			
	Sfi 19	Sfi 21	Sfi 11	DT1	MD2	Sfi 18	Sfi 21	Sf3	O1205	Dt2	2972	Dt4	MD1		Q5	Kappa3	7201
1																	3
3																	3
5 y 31																	3
6																	2
8																	1
11																	2
13																	6
17																	2
20																	4
22																	7
23																	2
25																	2
27																	1
28																	2
29																	8
32																	2
34																	2
36																	1
37																	1
41																	2

FIG. 3A



dG = -7.4 (inicialmente -7.4) sp

FIG. 4

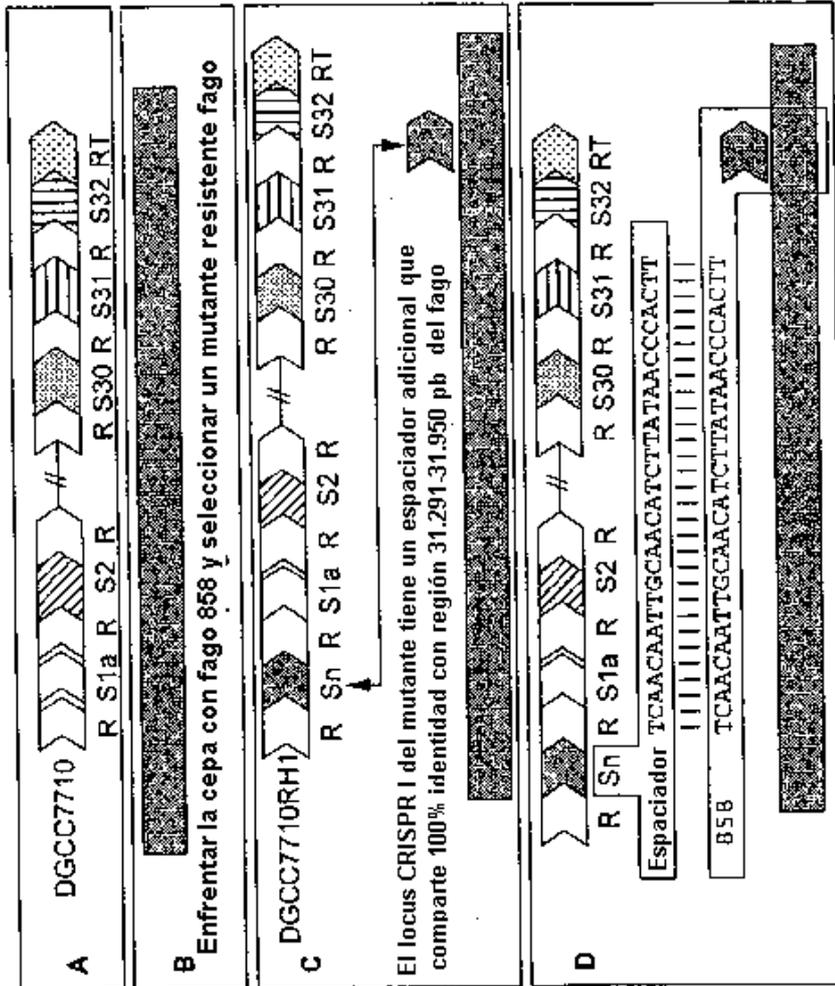


FIG. 5

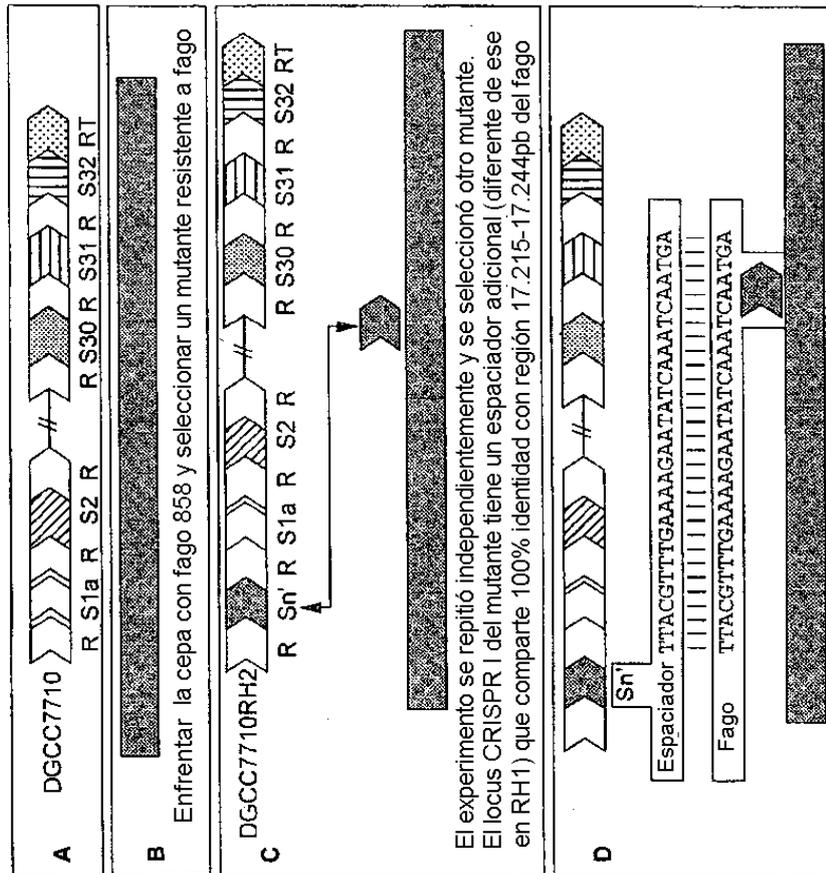


FIG. 6

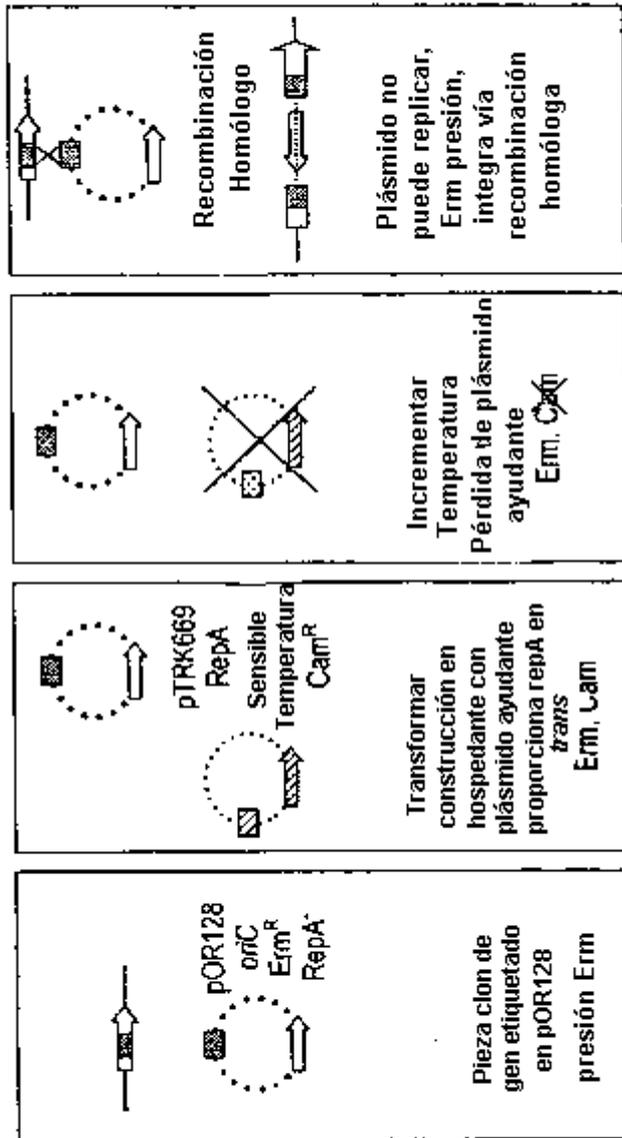


FIG. 9

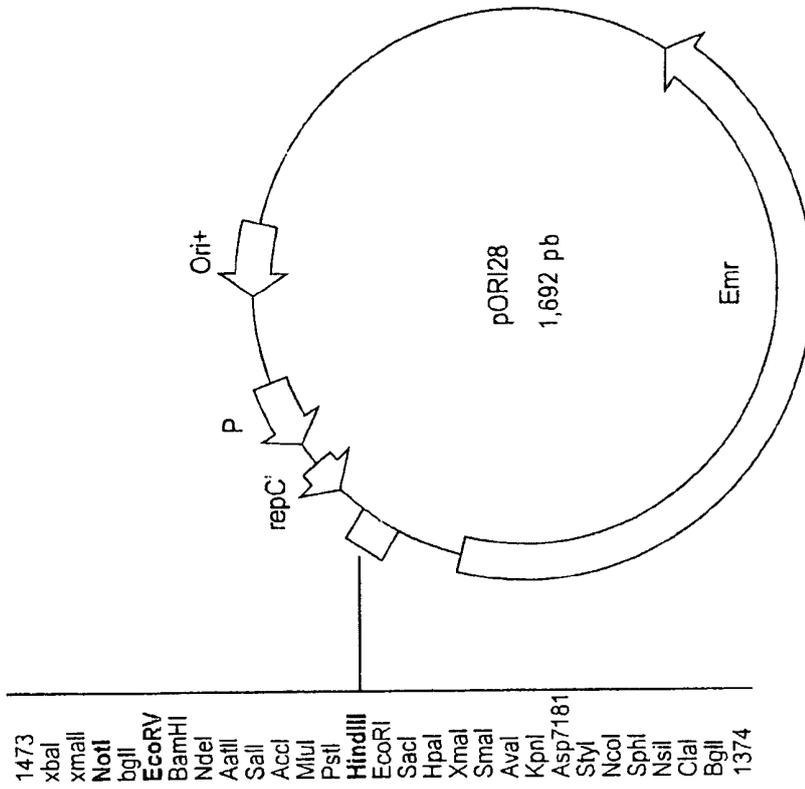


FIG. 11

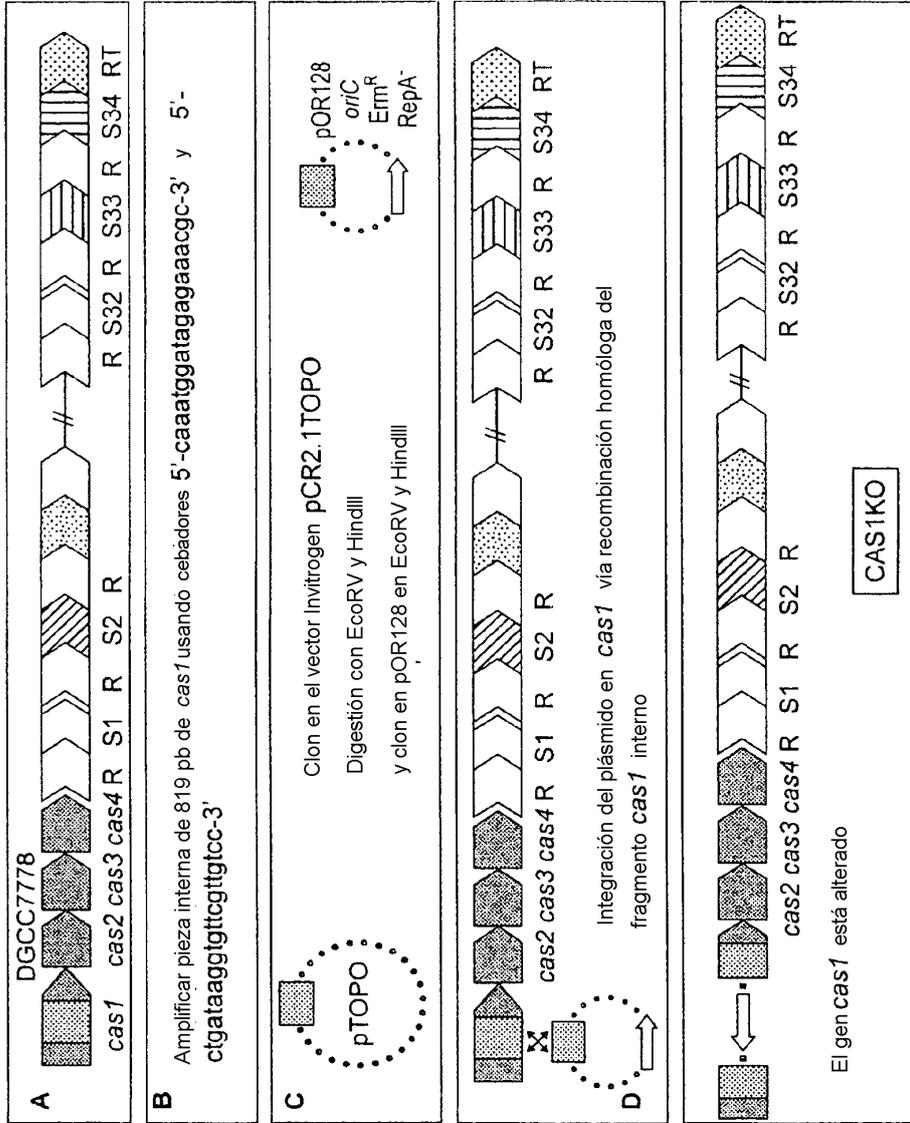


FIG. 12

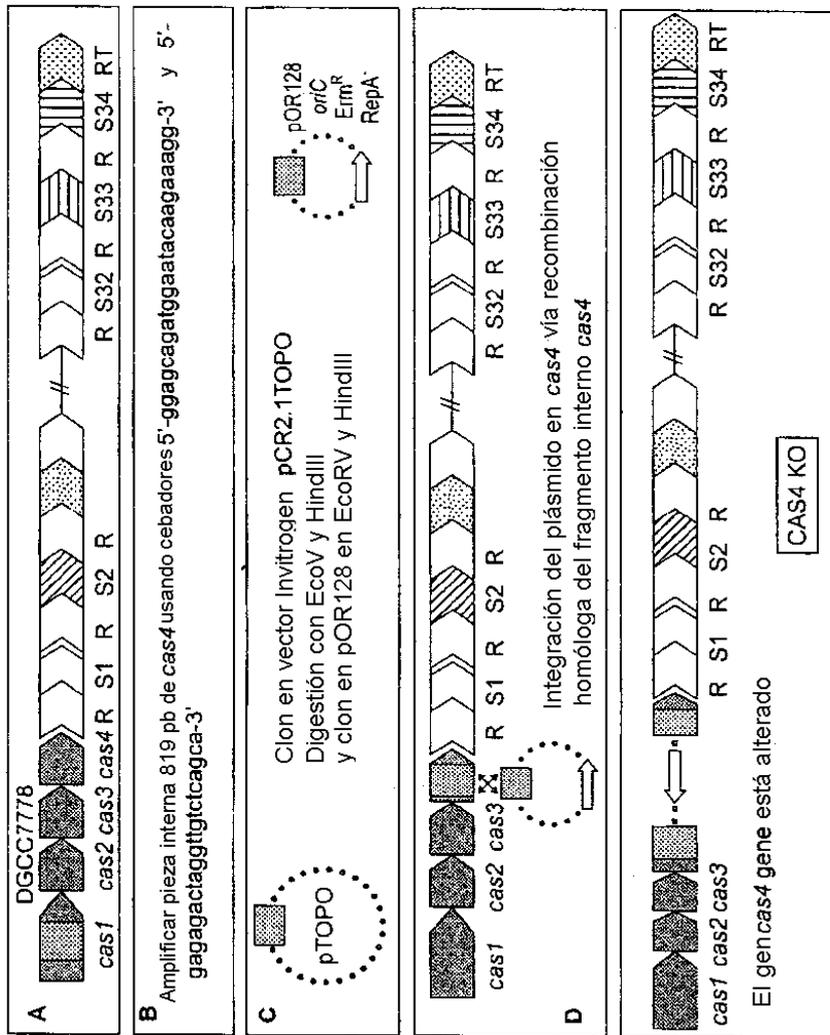


FIG. 13

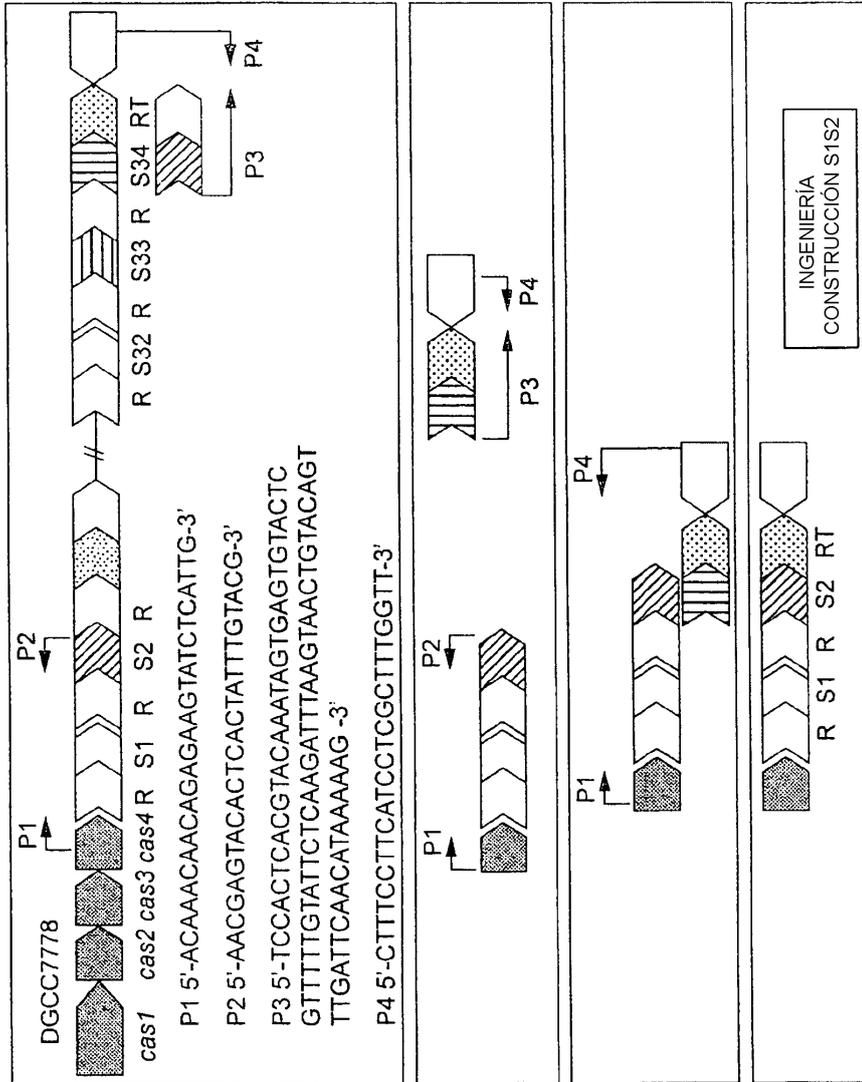


FIG. 14

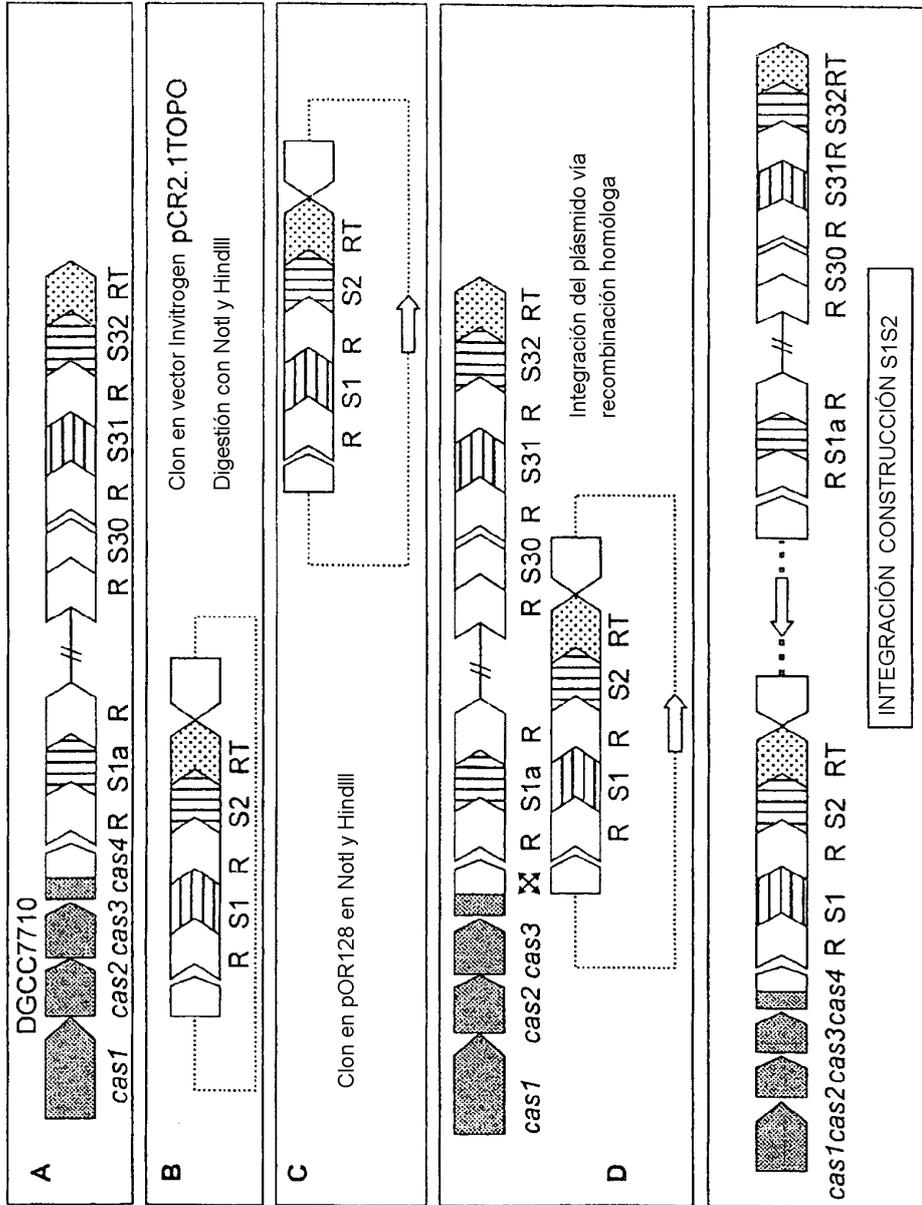


FIG. 16

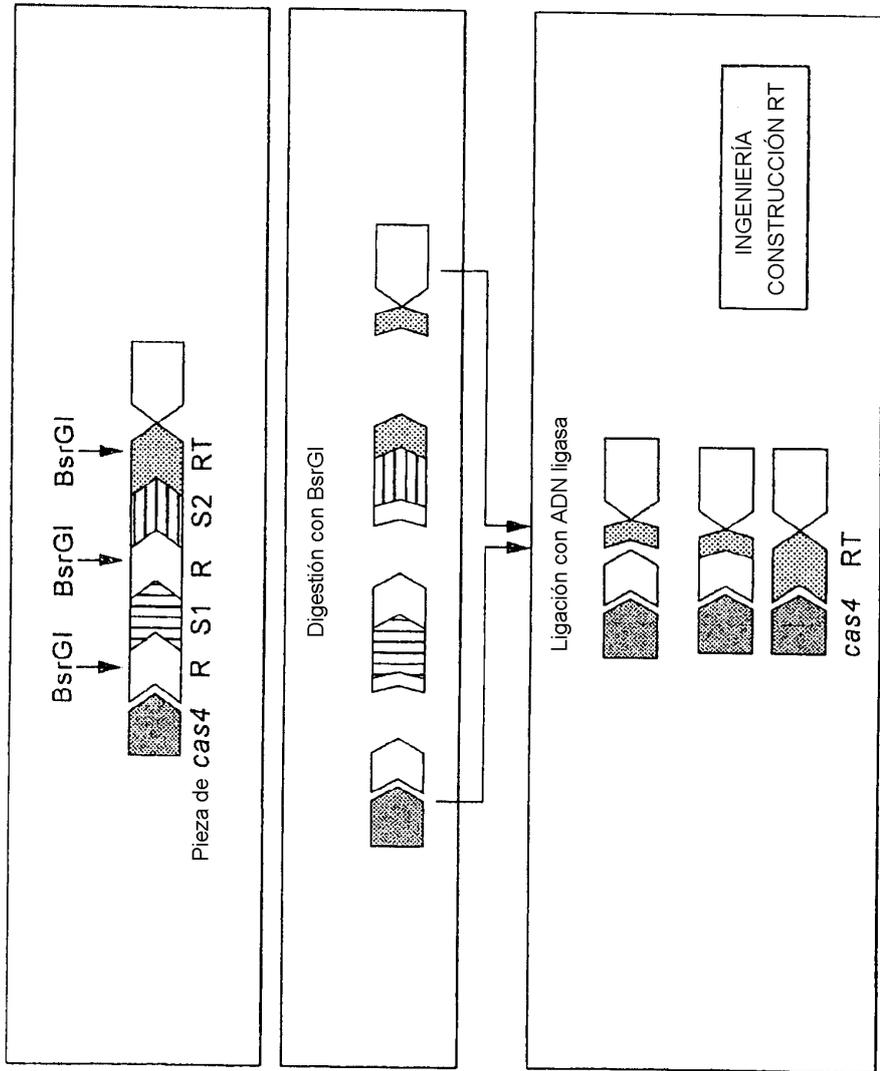


FIG. 17

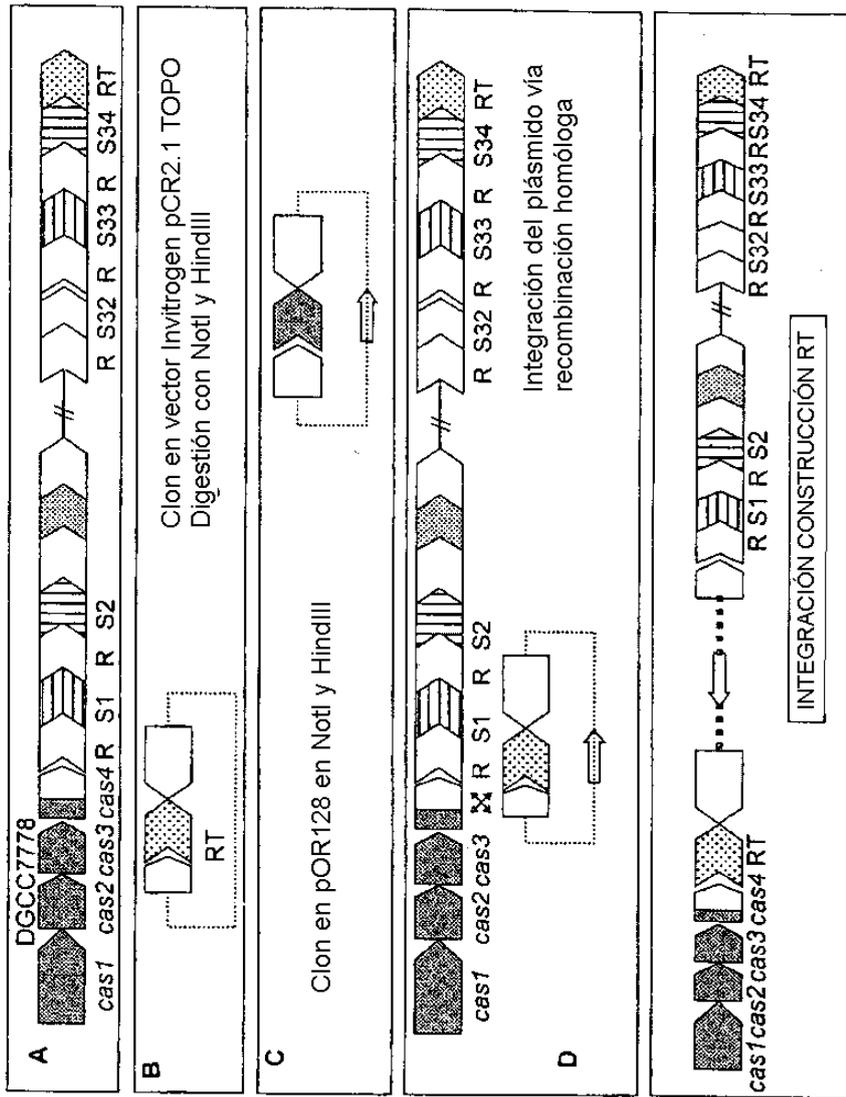


FIG. 18

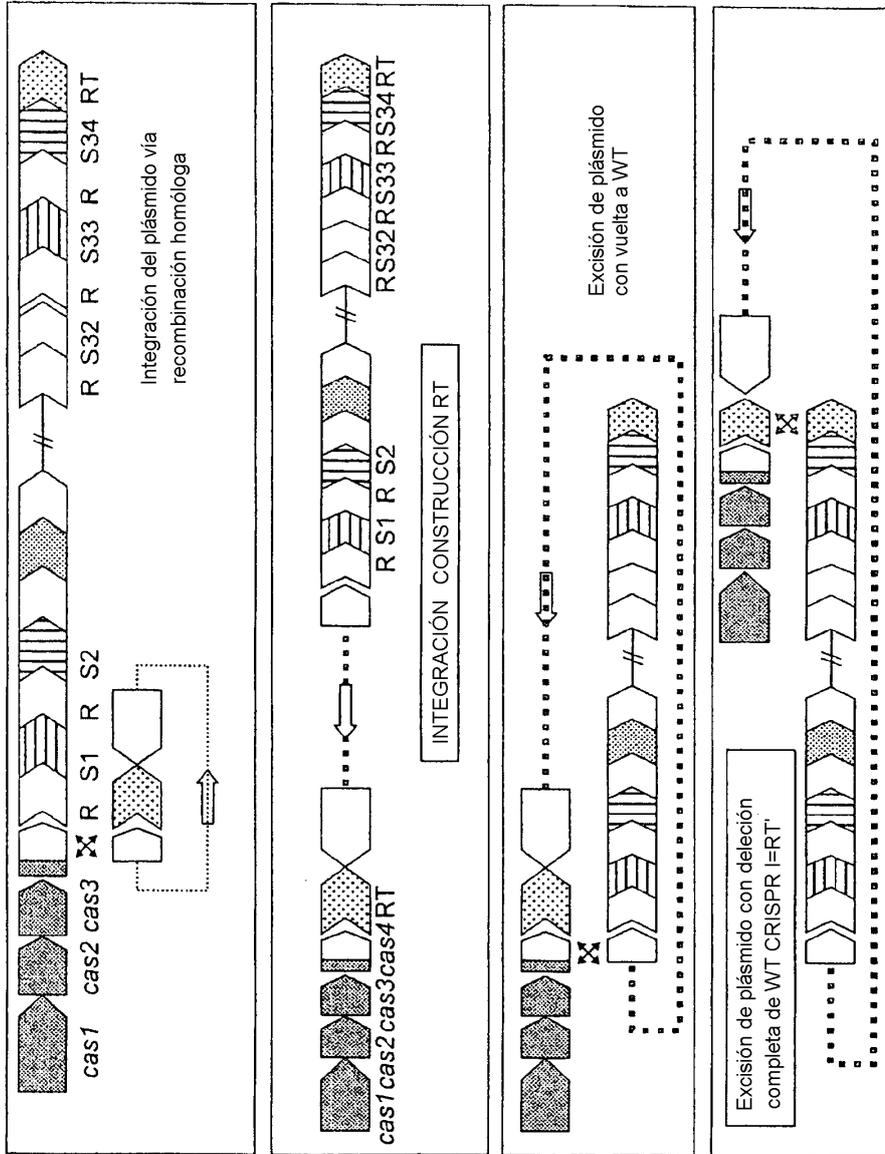


FIG. 19

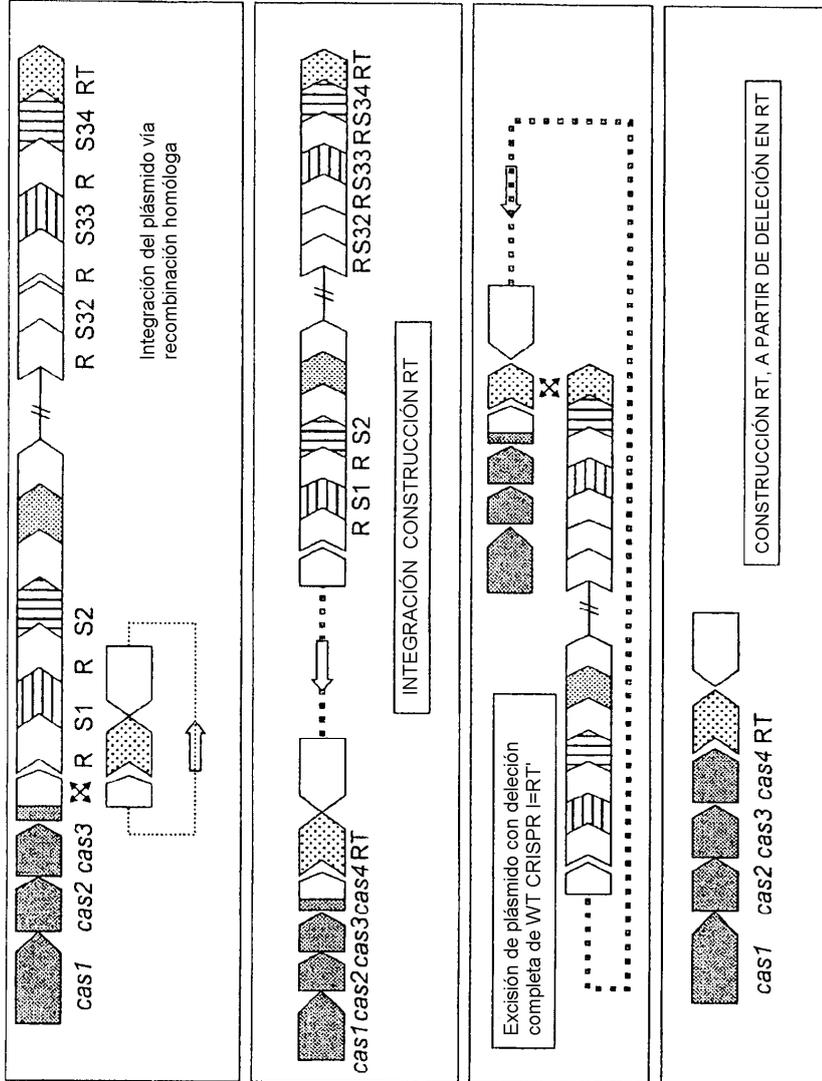


FIG. 20