

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 709 216**

51 Int. Cl.:

C12N 15/62 (2006.01)
C12N 15/82 (2006.01)
C12N 15/85 (2006.01)
A01K 67/027 (2006.01)
C07K 14/415 (2006.01)
C12N 9/16 (2006.01)
C12N 9/22 (2006.01)
C12N 15/90 (2006.01)
C07K 14/195 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **06.06.2012** **PCT/EP2012/060711**
87 Fecha y número de publicación internacional: **13.12.2012** **WO12168304**
96 Fecha de presentación y número de la solicitud europea: **06.06.2012** **E 12725813 (5)**
97 Fecha y número de publicación de la concesión europea: **10.10.2018** **EP 2718440**

54 Título: **Proteínas que tienen actividad nucleasa, proteínas de fusión y usos de estas**

30 Prioridad:

07.06.2011 EP 11004635

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
15.04.2019

73 Titular/es:

**HELMHOLTZ ZENTRUM MÜNCHEN DEUTSCHES
FORSCHUNGSZENTRUM FÜR GESUNDHEIT UND
UMWELT (GMBH) (100.0%)
Ingolstädter Landstrasse 1
85764 Neuherberg, DE**

72 Inventor/es:

KÜHN, RALF

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 709 216 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

Proteínas que tienen actividad nucleasa, proteínas de fusión y usos de estas

5 La presente invención se refiere a las modalidades como se caracterizaron en las reivindicaciones; específicamente a una molécula de ácido nucleico que codifica (I) un polipéptido que tiene la actividad de una endonucleasa, que es (a) una molécula de ácido nucleico que codifica un polipéptido que consiste en la secuencia de aminoácidos de la sec. con núm. de ident.: 1; (b) una molécula de ácido nucleico que consiste en la secuencia de nucleótidos de la sec. con núm. de ident.: 2; (c) una molécula de ácido nucleico que codifica una endonucleasa, cuya secuencia de aminoácidos es al menos 70 %
10 idéntica a la secuencia de aminoácidos de la sec. con núm. de ident.: 1; (d) una molécula de ácido nucleico que consiste en una secuencia de nucleótidos que es al menos 50 % idéntica a la secuencia de nucleótidos de la sec. con núm. de ident.: 2; (e) una molécula de ácido nucleico que es degenerada con respecto a la molécula de ácido nucleico de (d); o (f) una molécula de ácido nucleico correspondiente a la molécula de ácido nucleico de cualquiera de (a) hasta (e) en donde T se reemplaza por U; (II) un fragmento del polipéptido de (I) que tiene la actividad de una endonucleasa. Además, la
15 presente invención se refiere a un vector que comprende la molécula de ácido nucleico y una proteína codificada por dicha molécula de ácido nucleico. Además, la invención se refiere a un método *in vitro* para modificar el genoma de una célula eucariótica y a un método para producir un vertebrado no humano o un mamífero no humano.

20 En esta descripción, se citan un número de documentos que incluyen las solicitudes de patentes y las instrucciones de los fabricantes. La descripción de esos documentos no se considera relevante para la patentabilidad de esta invención.

25 Las nucleasas, desde su descubrimiento a finales de los años 1960, permanecen como una de las herramientas más importantes para los biólogos moleculares. Las nucleasas son enzimas capaces de escindir los enlaces fosfodiéster entre las subunidades de nucleótidos de los ácidos nucleicos. Las enzimas que catalizan la escisión del ADN y del ARN son partes integrales de los principales procesos metabólicos del ADN, tales como la replicación del ADN, la recombinación del ADN, la reparación del ADN, la recombinación en un sitio específico y el empalme del ARN. Además, las actividades nucleasas son esenciales en el procesamiento y la maduración del ARN, en el ARN de interferencia y son componentes de los mecanismos de defensa microbiana.

30 El ARN y el ADN presentan solamente dos tipos de enlaces fosfodiéster para la escisión, 5'- o 3'- de un fosfato escindible y la química fundamental es la sustitución nucleofílica bimolecular. Sin embargo, las estructuras y los mecanismos catalíticos de las nucleasas de ARN y ADN son muy variados y complejos. Las nucleasas pueden ser endonucleasas o exonucleasas, específicas del ADN o del ARN, topoisomerasas, recombinasas, ribozimas, o enzimas de empalme del ARN. Su reacción puede dividirse en las tres etapas del ataque nucleofílico, la formación de un intermedio pentacovalente cargado negativamente y la rotura del enlace escindible. Las nucleasas utilizan una variedad de nucleófilos para escindir
35 un enlace fosfato escindible. Los nucleófilos más comunes son las moléculas de agua desprotonadas por una base general para la hidrólisis directa. Para la escisión del ADN, las cadenas laterales de Ser, Thr e His sirven como nucleófilos para formar un intermedio covalente de ADN fosforil-proteína, que se resuelve posteriormente ya sea mediante la reacción de transferencia del fosforilo nuevamente al ADN durante la recombinación y topoisomerización, o mediante hidrólisis en reacciones de escisión de dos etapas. Para permitir la degradación controlada o el procesamiento del ADN o ARN celular, las actividades de las nucleasas se regulan estrictamente mediante una rigurosa especificidad de sustrato, localización confinada, o mediante inhibidores potentes.

40 Por conveniencia, las nucleasas pueden clasificarse de acuerdo con su mecanismo catalítico en tres clases principales basado en su dependencia de iones metálicos (Yang, W. (2011). Q. Rev. Biophys. 44(1): 1-93). Estas clases de nucleasas dependientes de dos iones metálicos, de un ion metálico e independientes de metales se dividen, además, en familias o superfamilias de acuerdo con la conservación de la secuencia y la estructura y la diversidad funcional.

Endonucleasas de Restricción

50 Entre todas las tres clases catalíticas se encuentran diversas familias de endonucleasas de restricción. Las enzimas de restricción de tipo I, III y IV son maquinarias moleculares complejas y de múltiples subunidades que combinan múltiples actividades que incluyen la restricción, la metilación y la translocación de ADN, requieren cofactores adicionales (AdoMet, ATP o GTP), se unen a más de un sitio diana, y escinden por fuera de la secuencia de reconocimiento, con frecuencia a una distancia aleatoria. Las endonucleasas de restricción de tipo II son enzimas que reconocen secuencias cortas de ADN
55 (generalmente de 4-8 pb de longitud) y escinden la diana en ambas hebras en, o muy cerca de, el sitio de reconocimiento. Las enzimas de restricción tipo II ortodoxas son homodiméricas, escinden dentro de secuencias palindrómicas, requieren iones Mg^{2+} y pueden actuar sobre copias únicas de sus dianas. Debido a su especificidad notablemente alta para reconocer y escindir sus secuencias diana, son de gran interés como las herramientas más frecuentemente usadas para la tecnología del ADN recombinante (Pingoud, A., M. Fuxreiter, y otros (2005). Cell Mol Life Sci 62(6): 685-707; Orłowski, J. y J. M. Bujnicki (2008). Nucleic Acids Res 36(11): 3552-69).

60 En la naturaleza, las REasas tipo II (endonucleasas de restricción) se encuentran en los organismos procarióticos, donde ellas forman sistemas de modificación por restricción con las ADN metiltransferasas de la misma o muy similar especificidad de sustrato. Las ADN metiltransferasas usan S-adenosilmetionina (AdoMet) como un donante del grupo metilo para modificar bases específicas en la secuencia diana, de esta manera las vuelve resistentes a la escisión mediada por la enzima de restricción. Mientras que el ADN del propio sistema de Restricción-Modificación se protege contra la

autodegradación por la nucleasa, cualquier ADN extraño (por ejemplo, de los fagos) que invade la célula huésped y que carece de metilación, puede destruirse eficientemente. Para distinguir los componentes de los sistemas de modificación/restricción los nombres de las metilasas y nucleasas están precedidos de los prefijos 'M'. y 'R.' (por ejemplo, M.FokI y R.FokI).

Muchas endonucleasas de restricción de tipo II usadas comúnmente comparten el motivo conservado PD-(D/E)XK. Dicho motivo se encuentra, generalmente, en proteínas que interactúan con moléculas de ácido nucleico tales como el ADN y no se limita a la presencia en nucleasas. Los tres residuos catalíticos se ubican cercanos entre sí en una horquilla β irregular. La primera D se ubica al inicio de la hebra primera y más corta, y la E y la K, separadas por un residuo hidrófobo x, en el medio de la hebra segunda y más larga. La primera D es la más conservada y coordina ambos iones metálicos, mientras que la segunda E puede reemplazarse por Q, D, N, H o S, y la tercera K puede reemplazarse por E, Q, D, S, N o T. Mediante la variación de las interfaces dimericas y, por lo tanto, de las posiciones relativas de los dos centros catalíticos, las endonucleasas dimericas pueden escindir el ADN para generar extremos romos o extremos escalonados con diversos salientes 5'- o 3'-. El módulo catalítico se aproxima invariablemente al ADN desde el lado del surco menor, y la unión a la secuencia específica se realiza mediante un módulo/subdominio separado en el surco mayor. Los dos primeros carboxilatos del motivo DEK coordinan los iones metálicos. El tercero, que generalmente se une al hidrógeno, tanto con el agua nucleófila como con el módulo de unión al ADN en el surco mayor, acopla el reconocimiento de la secuencia de ADN con la reacción de escisión. Los miembros de esta superfamilia tienen una secuencia primaria muy diversa y, por lo tanto, diferentes estructuras que rodean el núcleo catalítico. Las búsquedas en bases de datos con secuencias de enzimas de restricción, típicamente, no muestran similitud significativa con cualquier proteína, o una similitud muy alta (> 90 % de identidad) con unos pocos isosquizómeros, y ninguna similitud con otras proteínas. Esta distribución fuertemente sesgada de similitudes y disimilitudes dificultó el análisis de secuencia comparativa de todas las enzimas de restricción y planteó una interrogante acerca de si la diversidad de secuencias de aminoácidos de las endonucleasas de restricción indica evolución polifilética (convergencia) o divergencia extrema de un ancestro común.

Mientras que ~70 % de las endonucleasas de restricción pertenecen a la superfamilia PD-(D/E)XK, otros miembros de la superfamilia pueden ser monoméricos o tetraméricos e involucrarse en otros procesos tales como la reparación del ADN y la recombinación homóloga. Además de las endonucleasas, los miembros de esta superfamilia pueden ser, además, exonucleasas 5'- o 3'-. La fuente más completa de información sobre enzimas de restricción es la base de datos REBASE (<http://rebase.neb.com>) que enumera varios miles de enzimas caracterizadas funcionalmente y varios miles de enzimas putativas, deducidas a partir de comparaciones de secuencias o de análisis genómicos. Por lo tanto, existe una gran desproporción entre el número de secuencias conocidas o predichas y el número pequeño de ~50 proteínas caracterizadas experimentalmente con estructuras tridimensionales conocidas. En la actualidad, una gran fracción de enzimas putativas permanece sin ninguna predicción o dato experimental.

Las REasas tipo II se subdividen adicionalmente en numerosos tipos de acuerdo con la simetría de su sitio de reconocimiento, la organización estructural o la necesidad de cofactor. La mayoría de las enzimas de restricción usadas para el trabajo con el ADN recombinante pertenecen al Tipo IIP (P - palindrómico). Las enzimas **Tipo IIA** reconocen secuencias asimétricas, como Bpu10I, un dímero de subunidades no idénticas, cada una de las cuales es responsable de la escisión de una hebra del ADN. Las enzimas **Tipo IIB** escinden el ADN a ambos lados de la secuencia de reconocimiento, un ejemplo es BpII que, en la hebra de arriba escinde 8 nucleótidos antes y 13 nucleótidos después de la secuencia de reconocimiento, mientras que en la hebra de abajo escinde 13 nucleótidos antes y 8 nucleótidos después de la secuencia de reconocimiento. Las enzimas **Tipo IIC** tienen dos dominios, uno de escisión y otro de modificación, dentro de un polipéptido. Las enzimas **Tipo IIE**, para una escisión eficaz, necesitan interactuar con dos copias de sus secuencias de reconocimiento, una es la copia diana para la escisión, la otra copia sirve como un efector alostérico. Las enzimas Tipo IIE, como NaeI, reconocen secuencias de nucleótidos palindrómicas de una manera similar a las enzimas Tipo IIP y escinden el ADN dentro de los límites de sus sitios de reconocimiento; sin embargo, poseen un dominio de unión al ADN separado para realizar una función alostérica. Las enzimas **Tipo IIF** típicamente, son endonucleasas de restricción homotetraméricas que, además, interactúan con dos copias de su sitio de reconocimiento, pero escinden ambas de una manera concertada. Las enzimas **Tipo IIG**, esencialmente un subgrupo de enzimas tipo IIC, tienen dominios de escisión y de modificación dentro de un polipéptido. En general, son estimuladas por AdoMet, pero de cualquier otra manera se comportan como enzimas típicas Tipo II. Las enzimas **Tipo IIH** se comportan como las enzimas Tipo II, pero su organización genética se asemeja a los sistemas de Modificación/Restricción Tipo I. Las enzimas **Tipo IIM** reconocen una secuencia específica metilada y escinden el ADN en un sitio fijo. La enzima representativa más conocida es DpnI que escinde Gm6ATC, Gm6ATm4C y Gm6ATm5C, pero no GATC, GATm4C, GATm5C o los sitios hemimetilados. Muchas otras enzimas de restricción son más o menos tolerantes a la metilación, pero para las enzimas Tipo IIM el grupo metilo es un elemento de reconocimiento esencial. Las enzimas ortodoxas **Tipo IIP** como EcoRI, reconocen secuencias de nucleótidos simétricas y escinden dentro de sus sitios de reconocimiento. Comparten un núcleo estructural común que comprende las cinco hojas- β mezcladas flanqueadas por α -hélices. Sin embargo, los sitios de unión al ADN de las enzimas Tipo IIP son muy diversos y generalmente forman un parche en la superficie de la proteína compuesto por residuos de aminoácidos ubicados en los diferentes elementos estructurales (α -hélices, hojas- β , lazos). Las enzimas ortodoxas Tipo IIP interactúan con el ADN como homodímeros, y cada subunidad contribuye al reconocimiento de la mitad de la secuencia palindrómica. Las enzimas **Tipo IIS** escinden al menos una hebra del ADN diana fuera de la secuencia de reconocimiento. La enzima Tipo IIS mejor conocida es FokI, que al igual que muchas otras enzimas Tipo IIS interactúa con dos sitios de reconocimiento antes de escindir el ADN. Las enzimas Tipo IIS son activas como homodímeros y están compuestas por dos dominios, uno responsable para el reconocimiento de la diana y el otro para la catálisis (que sirve,

además, como el dominio de dimerización). Esto es evidente a partir de la estructura cristalina y de los estudios bioquímicos de FokI (Bitinaite, J., D. A. Wah, y otros (1998). *Proc Natl Acad Sci U S A* 95(18): 10570-5; Wah, D. A., J. Bitinaite, y otros (1998). *Proc Natl Acad Sci U S A* 95(18): 10564-9). El análisis de la estructura cristalina de FokI revela que está compuesta por un módulo de unión específico al ADN fusionado con el dominio de escisión que posee un núcleo catalítico de endonucleasa conservado pero escinde el ADN de una manera no específica. La arquitectura modular es característica, además, para la enzima BfiI de Tipo IIS, que está compuesta por dos dominios de unión al ADN fusionados al núcleo catalítico dímérico similar a la nucleasa no específica que pertenece a la familia de la fosfolipasa D. La presencia de un dominio nucleasa separado se ha informado, además, a partir de la estructura cristalina de la enzima SdaI Tipo IIP (Tamulaitiene, G., A. Jakubauskas, y otros (2006). *Structure* 14(9): 1389-400)

Enzimas de Restricción Modificadas y nucleasas quiméricas como herramientas para la edición del genoma

Las nucleasas que escinden las moléculas de ácido nucleico en sitios específicos en lugar de en sitios aleatorios son cada vez más importantes en las tecnologías emergentes tales como, por ejemplo, en la ingeniería genética y en la transformación génica. La transformación génica es un proceso en el cual una molécula de ADN introducida en una célula reemplaza el segmento cromosómico correspondiente mediante recombinación homóloga y, por lo tanto, presenta una forma precisa de manipular el genoma (Capecchi, M. R. (2005). *Nat Rev Genet* 6(6): 507-12). En el pasado, la aplicación de la transformación génica en células de mamíferos estuvo limitada por su baja eficiencia. Los experimentos en sistemas modelos demostraron que la frecuencia de recombinación homóloga de un vector de transformación génica aumenta fuertemente si se induce una rotura de doble hebra dentro de su secuencia diana cromosómica. Mediante el uso de la endonucleasa de direccionamiento de levadura I-SceI, que escinde el ADN en un sitio de reconocimiento de 18 pares de bases de longitud, se demostró inicialmente que la recombinación homóloga y la transformación génica se estimulan más de 1000 veces en células de mamíferos cuando se inserta un sitio de reconocimiento en un gen diana e I-SceI se expresa en esas células (Rouet, P., Smih, F., Jasin, M.; *Mol Cell Biol* 1994; 14: 8096-8106; Rouet, P., Smih, F. Jasin, M; *Proc Natl Acad Sci USA* 1994; 91: 6064-6068). En ausencia de un vector de transformación génica para la reparación dirigida mediante homología, frecuentemente las células cierran la ruptura de la doble hebra mediante uniones de extremos no homólogos (NHEJ). Dado que este mecanismo es propenso a errores, frecuentemente conduce a la delección o la inserción de múltiples nucleótidos en el sitio de escisión. Si el sitio de escisión se ubica dentro de la región codificante de un gen, de esta manera es posible identificar y seleccionar mutantes que muestren mutaciones con desplazamiento del marco de lectura a partir de una población mutagenizada y que representen alelos inactivados no funcionales del gen diana.

Por lo tanto, las nucleasas de secuencia específica representan una herramienta importante para la biotecnología para modificar el genoma de organismos modelos o líneas celulares. Para generar las nucleasas que reconocen específicamente nuevas secuencias diana dentro de los genes, se siguieron dos enfoques basados en la modificación de las endonucleasas de direccionamiento naturales o en la fusión de un dominio de unión al ADN, natural o modificado por ingeniería genética, a un dominio nucleasa. Tales enzimas de restricción modificadas o nucleasas quiméricas pueden dirigirse a grandes sitios de ADN (hasta 36 pb) y pueden modificarse genéticamente para unirse a secuencias de ADN deseadas.

Las endonucleasas de direccionamiento, tales como I-SceI de levaduras, son elementos genéticos naturales que catalizan su propia duplicación en alelos receptores mediante la generación de DSB en sitios específicos que inician su propia transferencia genética mediante recombinación homóloga. Una característica clave de estas enzimas es que crean roturas de la doble hebra en los sitios de reconocimiento que tienen una longitud de 14 a 40 pb. La principal limitación para el uso de endonucleasas de direccionamiento en la transformación génica es que cada enzima reconoce exclusivamente su secuencia diana natural. Mediante la ingeniería de proteínas se ha intentado modificar las endonucleasas de direccionamiento para reconocer nuevos sitios diana. En este trabajo, podrían hacerse modificaciones que alteren el sitio diana natural dentro de algunos nucleótidos, pero aún no es posible diseñar enzimas específicas para regiones diana completamente nuevas.

Debido a la dificultad para la manipulación de la secuencia de reconocimiento de las endonucleasas de direccionamiento, las nucleasas dedos de zinc (ZFN) son en la actualidad las nucleasas artificiales usadas más comúnmente para la ingeniería genética (Urnov, F. D., E. J. Rebar, y otros *Nat Rev Genet* 11(9): 636-46). Las nucleasas dedos de zinc se desarrollaron mediante la fusión del dominio de escisión, no específico de secuencia, de la endonucleasa de restricción Tipo IIS *FokI* (dominio Fn) a un nuevo dominio de unión al ADN. La ventaja de las nucleasas dedos de zinc es que el dominio dedos de zinc de unión al ADN puede modificarse para reconocer secuencias diana novedosas, que incluyen aquellas en los genes endógenos. Los módulos de proteínas conocidos como dedos de zinc se encuentran en el dominio de unión al ADN de la familia más abundante de factores de transcripción en la mayoría de los genomas eucarióticos. Cada dedo está compuesto por 30 aminoácidos, coordina un ion Zn^{2+} mediante el uso de dos cisteínas y dos residuos de histidina, y contacta principalmente con tres pares de bases de ADN. Dos características críticas de la estructura son que cada dedo une su sitio diana de 3 pb independientemente y que cada nucleótido parece estar en contacto por una única cadena lateral de aminoácidos que se proyecta desde un extremo de la α -hélice hasta el surco mayor del ADN. Se diseñaron dedos individuales para reconocer muchos de los 64 tripletes diana diferentes, pero el mayor éxito ha sido el diseño de dedos de zinc para reconocer los tripletes 5'-GNN-3'. Aunque se han propuesto códigos de reconocimiento de dedos de zinc, actualmente no existe ningún código que resulte consistentemente en dedos de zinc con enlace de alta afinidad. El perfeccionamiento de la especificidad de la unión de los dedos de zinc, tales como mediante el aumento del

número de dedos o mediante la construcción de proteínas de múltiples dedos que usan unidades de dos dedos, permanece como un área activa de investigación.

Mediante el uso de nucleasas dedos de zinc en ausencia de un vector de transformación génica para la reparación dirigida por homología, se generaron alelos inactivados en líneas de células de mamíferos y se obtuvieron ratas y peces cebra portadores de genes inactivados tras la expresión del ARNm de ZFN en embriones unicelulares (Santiago Y, Chan E, Liu PQ, Orlando S, Zhang L, Urnov FD, Holmes MC, Guschin D, Waite A, Miller JC, Rebar EJ, Gregory PD, Klug A, Collingwood TN.; *Proc Natl Acad Sci U S A* 2008; 105:5809-5814; Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Amacher SL.; *Nat Biotechnol* 2008; 26:702-708; Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, Jenkins SS, Wood A, Cui X, Meng X, Vincent A, Lam S, Michalkiewicz M, Schilling R, Foeckler J, Kalloway S, Weiler H, Menoret S, Anegón I, Davis GD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jacob HJ, Buelow R.; *Science* 2009; 325:433). Además, las nucleasas dedo de zinc se usaron en presencia de vectores de transformación génica exógenos que contienen regiones de homología con el gen diana para la reparación dirigida por homología de la ruptura de doble hebra a través de la conversión génica. Esta metodología se aplicó a la ingeniería genética en líneas de células de mamíferos y en la corrección génica en células humanas primarias (Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC.; *Nature* 2005; 435:646-651; Porteus MH, Baltimore D. 2003. *Science* 300:763; Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKaveler RC, Katibah GE, Amora R, Boydston EA, Zeitler B, Meng X, Miller JC, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R.; *Nat Biotechnol* 2009; 27:851-857).

Aunque el uso de las nucleasas dedos de zinc resulta en una mayor frecuencia de recombinación homóloga, se requieren esfuerzos y tiempos considerables para diseñar proteínas dedos de zinc que se unen a una nueva secuencia diana de ADN con alta eficiencia y que actúen como nucleasas de una secuencia específica. Además, durante mucho tiempo se ignoró que la naturaleza del dominio nucleasa de las nucleasas dedos de zinc, y de otras nucleasas quimérica, puede representar un factor de éxito igualmente importante para la actividad general de la proteína de fusión. La razón de esta desatención se basa en el hecho de que hasta la fecha solamente se encontró un único dominio nucleasa que retiene la actividad nucleasa dentro de un dominio de plegamiento de proteínas separado y que puede combinarse con los dominios de unión al ADN, para generar una proteína de fusión nucleasa de secuencia específica. Este dominio nucleasa se deriva de la enzima de restricción FokI, Tipo IIS, que se ha caracterizado en detalle y se conoce que actúa como un dímero obligado (Bitinaite, J., D. A. Wah, y otros (1998). *Proc Natl Acad Sci U S A* 95(18): 10570-5; Wah, D. A., J. Bitinaite, y otros (1998). *Proc Natl Acad Sci U S A* 95(18): 10564-9). En la mayoría de las otras enzimas de restricción, el reconocimiento y la escisión del ADN se combinan en un solo dominio de proteína y no pueden separarse. Una excepción es la enzima SdaI que se ha caracterizado estructuralmente por poseer un dominio nucleasa separado (Tamulaitiene, G., A. Jakubauskas, y otros (2006). *Structure* 14(9): 1389-400). Además, no ha sido posible aislar mutantes que pierden el reconocimiento de ADN pero retienen la actividad de escisión del ADN.

Por lo tanto, debido a la falta de otros dominios nucleasas funcionales comparables, durante mucho tiempo fue esencialmente desconocido si las propiedades enzimáticas del dominio FokI Fn pueden constituir un factor limitante para la actividad nucleasa del dominio Fn de las proteínas de fusión. Por ejemplo, la estructura intrínseca del dominio Fn puede restringir su actividad enzimática o la pequeña interfaz de dimerización de dos dominios Fn puede conducir a una interacción subóptima y a una baja tasa de escisión del sustrato de ADN.

Mediante mutagénesis dirigida, el dominio FokI Fn se modificó genéticamente en las variantes KK y EL que actúan, preferentemente, como heterodímeros (Miller, J. C., M. C. Holmes, y otros (2007). *Nat Biotechnol* 25(7): 778-85). El uso de estas variantes proporciona la especificidad mejorada por la secuencia diana de las nucleasas dedos de zinc y disminuye la toxicidad en las células de mamíferos, ya que se reconocen y procesan menos secuencias genómicas fuera de la diana. Sin embargo, la actividad nucleasa global de las variantes KK y EL es, a lo máximo, comparable a la del dominio Fn de tipo silvestre.

Sólo muy recientemente se encontró que el dominio FokI Fn de tipo silvestre exhibe ciertamente solo una actividad enzimática nucleasa subóptima, que limita el uso de las nucleasas dedos de zinc para la modificación genética del genoma. En un estudio de evolución de proteínas dirigida, el dominio Fn se mutagenizó al azar y se sometió a un ensayo en *E. coli* basado en nucleasa capaz de seleccionar mutantes que exhiben una actividad enzimática aumentada (Guo, J., T. Gaj, y otros (2010), *J Mol Biol* 400(1): 96-107). Mediante este procedimiento, se hizo posible aislar mutantes que exhiben una actividad de nucleasa > 10 veces más alta en comparación con el dominio Fn de tipo silvestre. Después del acoplamiento de estos mutantes a dominios dedos de zinc, tales proteínas de fusión mostraron un procesamiento de sustrato mejorado de tres a seis veces en células de mamíferos. Sin embargo, en la actualidad permanece desconocido si la actividad del dominio Fn puede mejorarse aún más o si la arquitectura intrínseca de la proteína del dominio Fn puede restringir cualquier mejora adicional.

Además de los dominios dedos de zinc de unión al ADN fusionados a los dominios nucleasa, muy recientemente se identificó, además, los dominios de unión al ADN de la proteína efectora TAL. En comparación con los motivos dedos de zinc, los elementos de repetición TAL dentro de las proteínas efectoras TAL proporcionan un nuevo tipo de dominio de unión al ADN que puede combinarse con un dominio nucleasa en nucleasas de secuencia específica. Una característica clave de los elementos peptídicos TAL se proporciona por su naturaleza moduladora. De esta manera, pueden generarse nuevas proteínas de unión al ADN en una secuencia específica mediante la combinación de solo cuatro elementos TAL

básicos que cada uno es específico para los nucleótidos A, C, G o T. Actualmente, solamente el dominio nucleasa de FokI se usa exitosamente en la fusión con los dominios de unión al ADN de la proteína efectora TAL (Miller y otros (2010). Nat. Biotechnol. 29, 143-148).

5 En resumen, existe una necesidad continua de nucleasas que puedan usarse en diversos entornos experimentales, que incluyen su fusión a otras proteínas y la modificación del dominio nucleasa.

El problema técnico que subyace en la presente invención fue la identificación de alternativas y/o medios y métodos mejorados para escindir moléculas de ácido nucleico.

10

La solución a este problema técnico se logra al proporcionar las modalidades caracterizadas en las reivindicaciones.

En consecuencia, la presente invención se refiere en una primera modalidad a una molécula de ácido nucleico que codifica (I) un polipéptido que tiene la actividad de una endonucleasa, que es (a) una molécula de ácido nucleico que codifica un polipéptido que consiste en la secuencia de aminoácidos de la sec. con núm. de ident.: 1; (b) una molécula de ácido nucleico que consiste en la secuencia de nucleótidos de la sec. con núm. de ident.: 2; (c) una molécula de ácido nucleico que codifica una endonucleasa, cuya secuencia de aminoácidos es al menos 70 % idéntica a la secuencia de aminoácidos de la sec. con núm. de ident.: 1; (d) una molécula de ácido nucleico que consiste en una secuencia de nucleótidos que es al menos 50 % idéntica a la secuencia de nucleótidos de la sec. con núm. de ident.: 2; (e) una molécula de ácido nucleico que es degenerada con respecto a la molécula de ácido nucleico de (d); o (f) una molécula de ácido nucleico correspondiente a la molécula de ácido nucleico de cualquiera de (a) hasta (e) en donde T se reemplaza por U; (II) un fragmento del polipéptido de (I) que tiene la actividad de una endonucleasa.

15

20

De acuerdo con la presente invención el término "molécula de ácido nucleico" define una cadena molecular lineal que consiste en al menos (para cada una) 2, 5, 10, 25, 50, 75, 100, 250, 500, tal como al menos 750, 1000, o al menos 2500 o más nucleótidos. El grupo de moléculas designado en la presente descripción como "moléculas de ácido nucleico" comprende, además, genes completos. El término "molécula de ácido nucleico", en la presente descripción se usa indistintamente con el término "polinucleótido".

25

30

El término "molécula de ácido nucleico" de acuerdo con la presente invención incluye ADN, tal como ADNc o ADN genómico, de doble o simple hebra, y ARN. En este sentido, "ADN" (ácido desoxirribonucleico) se refiere a cualquier cadena o secuencia de los bloques de construcción químicos adenina (A), guanina (G), citosina (C) y timina (T), llamadas bases de nucleótidos, que están unidas sobre una cadena principal de azúcar desoxirribosa. El ADN puede tener una hebra de bases de nucleótidos, o dos hebras complementarias que pueden formar una estructura de doble hélice. El "ARN" (ácido ribonucleico) se refiere a cualquier cadena o secuencia de los bloques de construcción químicos adenina (A), guanina (G), citosina (C) y uracilo (U), llamadas bases de nucleótidos, que se unen juntos en una cadena principal de azúcar ribosa. Típicamente, el ARN tiene una hebra de bases de nucleótidos. Además, se incluyen moléculas híbridas monocatenarias y bicatenarias, es decir, ADN-ARN. La molécula de ácido nucleico puede modificarse, además, por muchos medios conocidos en la técnica. Los ejemplos no limitantes de tales modificaciones incluyen la metilación, los "capuchones", la sustitución de uno o más de los nucleótidos de origen natural con un análogo, y las modificaciones entre los nucleótidos, tales como por ejemplo, aquellas con enlaces no cargados (por ejemplo, fosfonatos de metilo, fosfotriésteres, fosforamidatos, carbamatos, etcétera) y con enlaces cargados (por ejemplo, fosforotioatos, fosforoditioatos, etcétera). Los polinucleótidos pueden contener una o más porciones adicionales unidas covalentemente, tales como, por ejemplo, proteínas (por ejemplo, nucleasas, toxinas, anticuerpos, péptidos señal, poli-L-lisina, etcétera), intercaladores (por ejemplo, acridina, psoraleno, etcétera), agentes quelantes (por ejemplo, metales, metales radiactivos, hierro, metales oxidativos, etcétera), y agentes alquilantes. Los polinucleótidos pueden derivatizarse mediante la formación de un metil o etil fosfotriéster o un enlace alquilo fosforamidato. Se incluyen, además, moléculas que imitan el ácido nucleico conocidas en la técnica, tales como los derivados sintéticos o semisintéticos de ADN o ARN y los polímeros mixtos. Tales moléculas que imitan el ácido nucleico o los derivados de ácido nucleico de acuerdo con la invención incluyen ácido nucleico fosforotioato, ácido nucleico fosforamidato, ácido ribonucleico 2'-O-metoxietilo, ácido nucleico morfolino, ácido nucleico hexitol (HNA), ácido nucleico peptídico (PNA) y ácido nucleico bloqueado (LNA) (ver Braasch y Corey, Chem Biol 2001, 8: 1). El LNA es un derivado de ARN en el que el anillo de ribosa se restringe por un enlace metileno entre el oxígeno-2' y el carbono-4'. Además, se incluyen los ácidos nucleicos que contienen bases modificadas, por ejemplo, tiouracilo, tioguanina y fluorouracilo. Típicamente, una molécula de ácido nucleico transporta la información genética, que incluye la información usada por la maquinaria celular para producir proteínas y/o polipéptidos. La molécula de ácido nucleico de la invención puede comprender, adicionalmente, promotores, potenciadores, elementos de respuesta, secuencias de señal, secuencias de poliadenilación, intrones, regiones no codificantes 5' y 3' y similares.

35

40

45

50

55

60

65

El término "polipéptido", como se usa en la presente descripción indistintamente con el término "proteína", describe cadenas moleculares lineales de aminoácidos, que incluye las proteínas de cadena sencilla, que contienen más de 30 aminoácidos, mientras que el término "péptido" describe cadenas moleculares lineales de aminoácidos, que incluye las proteínas de cadena sencilla, que contienen menos de, y hasta 30 aminoácidos. Los polipéptidos pueden formar, además, oligómeros que consisten en al menos dos moléculas idénticas o diferentes. Las correspondientes estructuras de orden superior de tales multímeros son, según el caso, denominados homo o heterodímeros, homo o heterotrímeros, etcétera. Los polipéptidos de la invención pueden formar heteromultímeros u homomultímeros, tales como heterodímeros u homodímeros. Además, los peptidomiméticos de tales proteínas/polipéptidos donde los aminoácidos y/o los enlaces

peptídicos se reemplazaron por análogos funcionales, son además, comprendidos por la invención. Tales análogos funcionales incluyen todos los aminoácidos conocidos distintos de los 20 aminoácidos codificados por genes, tal como la selenocisteína. Los términos "polipéptido" y "proteína" se refieren, además, a polipéptidos y proteínas modificados naturalmente, donde la modificación se efectúa, por ejemplo, mediante glicosilación, acetilación, fosforilación, ubiquitinilación y modificaciones similares que se conocen bien en la técnica.

El término "un polipéptido que tiene la actividad de una endonucleasa", como se usa en la presente descripción, se refiere a un polipéptido que es capaz de escindir los enlaces fosfodiéster entre las subunidades de nucleótidos de ácidos nucleicos dentro de una cadena de polinucleótidos.

De acuerdo con la invención, la actividad enzimática endonucleasa se considera estable cuando, en las condiciones respectivas, la enzima es capaz de durar lo suficiente como para obtener el efecto deseado, específicamente, la escisión de su sustrato. Con respecto a esto, se observa que la actividad endonucleasa puede evaluarse como se describió en los ejemplos de la descripción o mediante métodos bien conocidos en la técnica. Por ejemplo, una molécula de ácido nucleico puede exponerse a una proteína cuya actividad endonucleasa debe evaluarse en condiciones que son adecuadas para la actividad enzimática endonucleasa. Después de la incubación, la composición que comprende la molécula de ácido nucleico (con o sin dicha proteína a evaluar) puede someterse a un ensayo para evaluar la longitud de una molécula de ácido nucleico tal como, por ejemplo, electroforesis en gel, para determinar si la molécula de ácido nucleico fue escindida.

De acuerdo con la presente invención, el término "porcentaje (%) de identidad de secuencia" describe el número de coincidencias ("aciertos") de nucleótidos/aminoácidos idénticos de dos o más secuencias de ácido nucleico o de aminoácido alineadas en comparación a la cantidad de nucleótidos o de residuos de aminoácidos que constituyen la longitud total de las secuencias modelos de ácido nucleico o aminoácidos. En otros términos, mediante el uso de un alineamiento, para dos o más secuencias o subsecuencias, puede determinarse el porcentaje de residuos de aminoácidos o de nucleótidos que son iguales (por ejemplo, 95 % de identidad) cuando se comparan y se alinean las (sub)secuencias para obtener el máximo de correspondencia en una ventana de comparación, o en una región designada según lo medido mediante el uso de un algoritmo de comparación de secuencias como se conoce en la técnica, o cuando se realiza el alineamiento manualmente y se inspecciona visualmente. Esta definición se aplica, además, al complemento de cualquier secuencia para alinearse. El alineamiento y el análisis de la secuencia de aminoácidos en relación con la presente invención se realizaron mediante el uso del algoritmo de la NCBI BLAST (Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, y David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402) y el programa informático del banco de trabajo principal CLC (versión 5.7.1; CLC bio, Aarhus, Dinamarca) que se usan, preferentemente, de acuerdo con esta invención. Preferentemente, se usan los parámetros estándar publicados (Altschul y otros, en el lugar citado). El experto en la técnica conoce la existencia de programas adicionales adecuados para alinear secuencias de ácidos nucleicos. Un programa preferido para el alineamiento de las secuencia de ácido nucleico, de acuerdo con la invención, es el programa informático del banco de trabajo principal CLC que usa los parámetros de alineamientos estándar del programa informático (versión 5.7.1; CLC bio, Aarhus, Dinamarca).

Como se define en las modalidades anteriores en la presente descripción, la invención contempla determinadas identidades de secuencia de aminoácidos. Además, se contempla, con preferencia creciente, las identidades de secuencias de aminoácidos de al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 97,5 %, al menos 98 %, al menos 98,5 %, al menos 99 %, al menos 99,5 %, al menos 99,8 %, y 100 % de identidad con la secuencia de aminoácidos respectiva de acuerdo con la invención.

Como se define en las modalidades anteriores en la presente descripción, la invención contempla determinadas identidades de secuencia de nucleótidos. Además, se contempla, con preferencia creciente, las identidades de secuencias de nucleótidos de al menos 55 %, 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 97,5 %, al menos 98 %, al menos 98,5 %, al menos 99 %, al menos 99,5 %, al menos 99,8 %, y 100 % de identidad con respecto a la secuencia de ácido nucleico respectiva de acuerdo con la invención.

El experto en la técnica apreciará fácilmente que más de una molécula de ácido nucleico puede codificar el mismo polipéptido debido a la degeneración del código genético. La degeneración se debe a que un código de tripletes designa a 20 aminoácidos y un codón de parada. Debido a que existen cuatro bases que se utilizan para codificar la información genética, se necesitan codones de tripletes para producir al menos 21 códigos diferentes. Las 4³ posibilidades posibles para las bases en los tripletes proporcionan 64 codones posibles, lo que significa que alguna degeneración debe existir. Como un resultado, algunos aminoácidos están codificados por más de un triplete, es decir, por hasta seis. La degeneración surge principalmente a partir de alteraciones en la tercera posición en un triplete. Esto significa que las moléculas de ácido nucleico que tienen diferentes secuencias, pero que aún codifican el mismo polipéptido se contemplan y pueden emplearse de acuerdo con el método de la presente invención.

Los fragmentos de acuerdo con la presente invención son polipéptidos que tienen la actividad de una endonucleasa, como se define anteriormente en la presente descripción, y comprenden al menos 90 aminoácidos. Con respecto a esto, se prefiere, con preferencia creciente, que los fragmentos de acuerdo con la presente invención sean polipéptidos de al menos 100, al menos 125, al menos 150, al menos 200 aminoácidos, al menos 300 aminoácidos, al menos 400

aminoácidos. Los fragmentos del polipéptido de la invención, que retienen sustancialmente la actividad endonucleasa, incluyen truncamientos N-terminales, truncamientos C-terminales, sustituciones de aminoácidos, delecciones internas y adición de aminoácidos (ya sea internamente o en cualquier extremo de la proteína). Por ejemplo, las sustituciones de aminoácidos conservativas son conocidas en la técnica y pueden introducirse en la endonucleasa de la invención sin afectar sustancialmente la actividad endonucleasa, es decir, disminuir dicha actividad.

Como es evidente a partir de los ejemplos, el inventor fue capaz de identificar y aislar una nueva nucleasa, en particular el dominio endonucleasa, derivado de una cepa de *Clostridium* como se detalla más abajo. Específicamente, el inventor pudo establecer la utilidad del producto génico de un gen bacteriano putativo sin connotación funcional conocida como una secuencia de nucleasa inespecífica. La nucleasa novedosa puede emplearse en diversos entornos experimentales al igual que cualquier otra nucleasa. Por ejemplo, puede usarse para escindir aleatoriamente moléculas de ácido nucleico o, por ejemplo, en fusión con dominios de unión al ADN, para la escisión en un sitio específico de las moléculas de ácido nucleico. Es importante destacar que, como se describe más abajo y específicamente en los ejemplos, la endonucleasa novedosa puede usarse en combinación con los dominios de unión al ADN de la proteína efectora TAL como parte de una proteína de fusión para la escisión de ácidos nucleicos en una secuencia específica. Con respecto a esto, la nucleasa novedosa muestra su superioridad sobre las endonucleasas del estado de la técnica distintas de FokI, que hasta el momento no pudieron mostrarse activas en las proteínas de fusión correspondientes. Brevemente, los inventores probaron el producto génico de dicho gen microbiano hipotético no caracterizado que designaron como "Clo051" (sec. con núm. de ident.: 17) y que se derivó del genoma de *Clostridium spec. 7_2_43FAA* (Secuencia de Referencia del NCBI: ZP_05132802.1; publicación/base de datos liberada en fecha: 9 de junio, de 2010), más específicamente su dominio nucleasa putativo (ver Figuras 5 y 6), por su actividad endonucleasa en combinación con el dominio de unión al ADN de una proteína efectora TAL. Además, se evaluaron diversas proteínas endonucleasas conocidas en combinación con los dominios de unión al ADN de la proteína efectora TAL, así como también otros dos genes microbianos hipotéticos. Sorprendentemente, solamente pudo demostrarse que el dominio nucleasa de Clo051 estaba activo, mientras que las otras proteínas de fusión no mostraron actividad (ver el Ejemplo 1 para los detalles). Los experimentos comparativos enfatizaron en el significado del descubrimiento de la presente invención en que se identificó una nucleasa novedosa que, además, exhibe actividad cuando se fusiona a los dominios de unión al ADN de las proteínas efectoras TAL. Las proteínas efectoras TAL se expresan por los patógenos de plantas del género *Xanthomonas* y reprograman las células huésped mediante la imitación de los factores de transcripción eucarióticos. Las proteínas efectoras TAL se caracterizan por un dominio central de repeticiones en tándem de 32 a 34 aminoácidos que constituyen un dominio de unión al ADN. El número y el orden de las repeticiones en una proteína efectora TAL determinan su actividad específica de unión al ADN. (Boch, J., y otros 2009 Science 326: 1509-12). Las secuencias de aminoácidos de las repeticiones son conservadas, excepto para dos residuos adyacentes altamente variables (en las posiciones 12 y 13) que determinan la especificidad hacia las bases del ADN A, G, C o T. La unión al ADN se realiza mediante el contacto de un nucleótido de la doble hélice del ADN con los residuos variables en la posición 12 y 13 dentro del motivo efector de TAL, lo que resulta en una correspondencia uno a uno entre las repeticiones secuenciales en las proteínas efectoras TAL y los nucleótidos secuenciales en el ADN diana. La unión a secuencias de ADN más largas se logra mediante la unión en tándem de varios de estos motivos efectoras TAL para formar un "dominio de unión al ADN de una proteína efectora TAL". El uso de tales dominios de unión al ADN de las proteínas efectoras TAL para la creación de proteínas de fusión, motivo efector TAL nucleasa, que reconoce y escinde una secuencia diana específica depende de la creación adecuada de los dominios de unión al ADN de las proteínas efectoras TAL que pueden reconocer específicamente dicha diana particular. La ventaja de los elementos de repetición TAL, en comparación con, por ejemplo, los elementos dedos de zinc, se proporciona por su naturaleza verdaderamente modular. De esta manera, pueden generarse nuevas proteínas de unión al ADN en una secuencia específica a través de la combinación de los cuatro elementos básicos de TAL que son específicos para los nucleótidos A, C, G o T.

Es importante señalar que, en la presente invención, el dominio nucleasa Clo051 fusionado a los dominios de unión al ADN de las proteínas efectoras TAL se evaluó y se encontró que es activo en mamíferos, específicamente en cultivos de células humanas. Por lo tanto, la utilidad de las proteínas de fusión del dominio nucleasa Clo051 para la manipulación del ADN y de los genes, específicamente en células de mamíferos, pero sin limitación, se ha demostrado directamente en el sistema biológico que proporciona aplicaciones importantes para esta tecnología. Este hallazgo es de particular importancia debido a que los estudios sobre la función de las proteínas que se realizan en organismos eucarióticos inferiores, como, por ejemplo, levaduras, no permiten una conclusión definitiva acerca de la utilidad de la proteína en estudio en células de mamíferos. Por ejemplo, una proteína específica puede funcionar de manera óptima a 30° Celsius, la temperatura de crecimiento de la levadura, pero se vuelve inestable o se inactiva a 37° Celsius que es la temperatura corporal típica de los mamíferos. Además, el medio intracelular de, por ejemplo, las células de levadura, como la concentración de iones y de proteínas, la diversidad de proteínas y los mecanismos de degradación de proteínas, se distinguen del medio intracelular de las células de mamíferos.

Mientras que los ejemplos describen solamente el uso del dominio nucleasa de Clo051 (sec. con núm. de ident.: 1), por ejemplo, en combinación con los dominios de unión al ADN, el experto en la técnica apreciará que puede emplearse la secuencia completa de Clo051 como se establece en la sec. con núm. de ident.: 17 o los fragmentos más cortos de esta que tienen actividad endonucleasa y que comprenden la secuencia de aminoácidos de la sec. con núm. de ident.: 1. La secuencia de aminoácidos de la sec. con núm. de ident.: 1 comienza en E389 y termina en Y587 de los aminoácidos de la sec. con núm. de ident.: 17 como se ilustra, además, en la Figura 5.

En una modalidad preferida de la molécula de ácido nucleico de la invención, en (I)(c) en dicha secuencia de aminoácidos que tiene al menos un 70 % de identidad de secuencia con la sec. con núm. de ident.: 1 los residuos de aminoácidos P66, D67, D84 y/o K86 de la sec. con núm. de ident.: 1 no están modificados.

5 El dominio nucleasa de Clo051, al igual que muchas endonucleasas de restricción Tipo-II y, por ejemplo, la proteína MutH de reparación del ADN, comparte el motivo de secuencia conservada PD-(D/E)XK dentro del núcleo de su dominio catalítico. El núcleo sirve como un andamio para un sitio activo débilmente conservado, que comprende, típicamente, dos o tres residuos ácidos (Asp o Glu) y un residuo de Lys, que juntos forman el motivo catalítico bipartito distintivo [(P)D. Xn. (D/E)XK](donde X es cualquier aminoácido). Este motivo ha llevado a denominar a esta superfamilia de proteínas como

10 'PD-(D/E)XK'. El trabajo sobre las enzimas de restricción y las proteínas de reparación del ADN ha demostrado que los tres residuos catalíticos se ubican cercanos entre sí en una horquilla β irregular. La primera D se ubica al inicio de la hebra primera y más corta, y la E y la K, separadas por un residuo hidrófobo x, en el medio de la hebra segunda y más larga. El módulo catalítico se aproxima invariablemente al ADN desde el lado del surco menor, y la unión a la secuencia específica se realiza mediante un módulo/subdominio separado en el surco mayor. Los dos primeros carboxilatos del motivo DEK coordinan los iones metálicos. La primera D es la más conservada y coordina ambos iones metálicos, mientras que la segunda E puede reemplazarse por Q, D, N, H o S, y la tercera K puede reemplazarse por E, Q, D, S, N o T. El residuo de lisina en el motivo DEK conserva el agua nucleófila junto con el fosfato 3' al enlace escindible; la misma Lisina, además, se une al hidrógeno con un oxígeno carbonilo en el módulo de unión al ADN. Esta lisina, que es conservada en muchas endonucleasas de restricción y se reemplaza por Glu o Gln en BamHI y BglII, se ha propuesto

15 como un sensor para la unión al ADN y un centro que acopla el reconocimiento de bases y la escisión del ADN (Lee y otros (2005). Molecular Cell 20, 155- 166; Orlowski, J. y J. M. Bujnicki (2008). Nucleic Acids Res 36(11): 3552-69).

La secuencia primaria del dominio nucleasa de Clo051 entre las posiciones E389 e Y587 de la secuencia de sec. con núm. de ident.: 17, es decir, la secuencia de la sec. con núm. de ident.: 1, exhibe una distribución única de los residuos de arginina (R) y lisina (K) cargados positivamente y de los residuos de glutamato (E) y aspartato (D) cargados negativamente (Figura 13). Estos residuos constituyen un paisaje tridimensional de cargas dentro del dominio Clo051 que determina la estructura terciaria única de esta nucleasa, como se muestra en el modelo estructural en la Figura 6. Determinados reemplazos de residuos polares versus no polares, o de residuos no polares contra residuos polares, por ejemplo en las posiciones S35 y/o R58 de la sec. con núm. de ident.:1 (o S423 y R446 de la sec. con núm. de ident.: 17),

25 alteran la estructura tridimensional de la cadena proteica y pueden resultar en un aumento de la actividad nucleasa. Tales reemplazos de aminoácidos pueden realizarse mediante ensayos de prueba y error o pueden seguir hipótesis específicas sobre el impacto estructural y funcional en el dominio nucleasa de Clo051. Alternativamente, un gran número de variantes mutagenizadas aleatoriamente de la región codificante del dominio nucleasa de Clo051 pueden ensamblarse en una biblioteca mediante PCR mutagénico, propenso a errores. Esta biblioteca de moléculas mutantes puede analizarse para

30 detectar la presencia de variantes hiperactivas de nucleasa mediante un ensayo de selección fenotípico en células de *E. coli*, levaduras o mamíferos que se acopla a una lectura de nucleasa funcional, por ejemplo, como se describió para la mejora de la recombinasa FLP (Buchholz y otros, Nat. Biotechnol. 16, 657-62, 1998). Tal selección funcional para variantes mejoradas de nucleasas puede resultar en el reemplazo de residuos únicos o múltiples que conduce a un aumento de la actividad nucleasa en comparación con la forma de tipo silvestre de Clo051.

Además, se contemplan modalidades donde más de un residuo de aminoácido P66, D67, D84 y/o K86 de la sec. con núm. de ident.: 1 no están modificados, tal como, por ejemplo, los tramos de aminoácidos, por ejemplo desde al menos P66 hasta al menos K86, al menos R64 hasta al menos Y88, al menos G62 hasta al menos E90, así como también desde L60 hasta al menos Y92 de la sec. con núm. de ident.: 1.

En una modalidad preferida de la invención, la molécula de ácido nucleico codifica, además, un dominio de unión al ADN.

En esta modalidad, la molécula de ácido nucleico de la invención codifica una proteína de fusión que tiene la actividad de una endonucleasa y comprende un dominio de unión al ADN y un dominio de escisión que comprende o consiste en el

50 dominio endonucleasa novedoso. El término "proteína de fusión" es bien conocido en la técnica y tiene el mismo significado en la presente descripción. Específicamente, se refiere a una proteína generada mediante la unión de dos o más secuencias de ácido nucleico diana, por ejemplo, genes que originalmente codifican proteínas separadas, para crear un constructo de fusión. La traducción de dicho constructo de fusión resulta en una proteína única con las propiedades funcionales derivadas de dichas proteínas separadas. Las dos proteínas que dan lugar a la proteína de fusión pueden

55 conectarse mediante un enlazador, tal como, por ejemplo, un péptido enlazador. En otras palabras, el dominio de unión al ADN y el dominio de escisión de las nucleasas pueden fusionarse directamente entre sí o pueden fusionarse a través de un enlazador.

El término "enlazador", como se usa de acuerdo con la presente invención, se refiere a una secuela de aminoácidos (es decir, péptidos enlazadores) así como también a enlazadores no peptídicos.

Los péptidos enlazadores, como se contempla por la presente invención, son péptidos o polipéptidos enlazadores de al menos 1 aminoácido en longitud. Preferentemente, los enlazadores son de 1 hasta 100 aminoácidos en longitud. Con mayor preferencia, los enlazadores son de 5 hasta 50 aminoácidos en longitud, e incluso con mayor preferencia, los enlazadores son de 10 hasta 20 aminoácidos en longitud. Es bien conocido por el experto en la técnica que la naturaleza, es decir, la longitud y/o la secuencia de aminoácidos del enlazador puede modificar o potenciar la estabilidad y/o la

solubilidad de la molécula. Por lo tanto, la longitud y la secuencia de un enlazador dependen de la composición de las porciones respectivas de la proteína de fusión.

El experto en la técnica conoce los métodos para evaluar la idoneidad de diferentes enlazadores. Por ejemplo, las propiedades de la molécula pueden probarse fácilmente mediante la evaluación de la actividad nucleasa, así como también la especificidad de unión al ADN de las porciones respectivas de la proteína de fusión para usarse en el método de la invención.

Se apreciará por el experto en la técnica que cuando la proteína de fusión se proporciona como una molécula de ácido nucleico que codifica la proteína de fusión en forma expresable, el enlazador es un péptido enlazador codificado, además, por dicha molécula de ácido nucleico.

El término "enlazador no peptídico", como se usa de acuerdo con la presente invención, se refiere a grupos de enlace que tienen dos o más grupos reactivos pero que excluyen los péptidos enlazadores como se definió anteriormente. Por ejemplo, el enlazador no peptídico puede ser un polímero que tiene grupos reactivos en ambos extremos, que se unen individualmente a grupos reactivos de las porciones individuales de la proteína de fusión, por ejemplo, un extremo amino, un residuo de lisina, un residuo de histidina o un residuo de cisteína. Los grupos reactivos del polímero incluyen un grupo aldehído, un grupo aldehído propiónico, un grupo aldehído butilo, un grupo maleimida, un grupo cetona, un grupo vinil sulfona, un grupo tiol, un grupo hidrazida, un grupo carbonildimidazol (CDI) un grupo carbonato de nitrofenilo (NPC), un grupo trisilato, un grupo isocianato y derivados de succinimida. Los ejemplos de derivados de succinimida incluyen propionato de succinimidilo (SPA), ácido succinimidil butanoico (SBA), carboximetilato de succinimidilo (SCM), succinimidil succinamida (SSA), succinimidil succinato (SS), succinimidil carbonato y N-hidroxi succinimida (NHS). Los grupos reactivos en ambos extremos del polímero no peptídico pueden ser iguales o diferentes. Por ejemplo, el polímero no peptídico puede tener un grupo maleimida en un extremo y un grupo aldehído en otro extremo. Preferentemente, el enlazador es un péptido enlazador. Con mayor preferencia, el péptido enlazador consiste en siete residuos de glicina.

Además, la proteína de fusión puede estar flanqueada en los extremos N o C terminal por secuencias adicionales no relacionadas a dichas proteínas en la proteína de fusión. De acuerdo con la presente invención, una proteína de fusión de la invención comprende un dominio de unión a ADN. El término "dominio de unión al ADN" tiene el mismo significado como se conoce en la técnica y se refiere a un motivo/conformación de secuencia dentro de una proteína que se une a los motivos del ADN. Los dominios de proteínas que pueden unirse específicamente a una secuencia de ácido nucleico incluyen, por ejemplo, repeticiones de dedos de zinc, el motivo hélice-giro-hélice (HTH) de los homeodominios, y el motivo cinta-hélice-hélice (RHH). La unión específica se refiere a la unión específica a una secuencia y es específica, cuando un dominio de unión al ADN estadísticamente solamente se une a una secuencia particular y no se une, o no se une esencialmente, a una secuencia no relacionada. El experto en la técnica conoce bien las secuencias que codifican los dominios de unión al ADN (Rohs y otros (2010). *Annu. Rev. Biochem.* 79, 233-269; Maeder y otros (2009). *Nat. Protocols* 10, 1471-1501).

En una modalidad más preferida de la molécula de ácido nucleico de la invención, el dominio de unión al ADN es un motivo efector TAL de una proteína efectora TAL.

Esta modalidad se refiere a una molécula de ácido nucleico que codifica, además, una nucleasa TAL. El término "nucleasa TAL", como se usa en la presente descripción, es bien conocido en la técnica y se refiere a una proteína de fusión que comprende un dominio de unión al ADN, en donde el dominio de unión al ADN comprende o consiste en motivos efectores TAL de una proteína efectora TAL y el dominio de escisión no específico de una nucleasa de restricción. La proteína de fusión de la invención que, además, se emplea en el método de la invención más abajo retiene, o retiene esencialmente, la actividad enzimática de la endonucleasa de la invención. De acuerdo con la presente invención, dicha actividad endonucleasa (además, denominada función) se retiene esencialmente si al menos se retiene el 60 % de la actividad biológica de la actividad endonucleasa. Preferentemente, se retiene al menos el 75 % o al menos el 80 % de la actividad endonucleasa. Más preferido es que se retiene al menos 90 % tal como al menos 95 %, incluso más preferido al menos 98 % tal como al menos 99 % de la actividad biológica de la endonucleasa. Lo más preferido es que la actividad biológica se retenga completamente, es decir, hasta el 100 %. Además, de acuerdo con la invención, se contemplan proteínas de fusión que tienen una actividad biológica aumentada en comparación con la endonucleasa cuando no está fusionada a un dominio de unión a ADN, es decir, más del 100 % de la actividad. Los métodos para evaluar la actividad biológica de las endonucleasas (de restricción) se conocen bien por el experto en la técnica e incluyen, sin limitación, la incubación de una endonucleasa con ADN recombinante y el análisis de los productos de la reacción mediante electroforesis en gel (Bloch KD.; *Curr Protoc Mol Biol* 2001; Capítulo 3:Unidad 3.2).

El término "proteína efectora TAL", como se usa en la presente descripción, se refiere a proteínas que pertenecen a la familia de proteínas TAL (similar a activador de la transcripción). Estas proteínas se expresan por los patógenos bacterianos de las plantas del género *Xanthomonas*. Los miembros de la gran familia de efectores TAL son factores clave de la virulencia de *Xanthomonas* y reprograman las células huésped mediante la imitación de los factores de transcripción eucarióticos. La patogenicidad de muchas bacterias depende de la inyección de proteínas efectoras a través de la secreción de tipo III en células eucarióticas para manipular los procesos celulares. Las proteínas efectoras TAL de *Xanthomonas* patogénicas de plantas son factores de virulencia importantes que actúan como activadores transcripcionales en el núcleo de la célula vegetal. PthXo1, una proteína efectora TAL de unas *Xanthomonas* patógena

de arroz, activa la expresión del gen del arroz Os8N3, que permite que *Xanthomonas* colonice las plantas de arroz. Las proteínas efectoras TAL se caracterizan por un dominio central de repeticiones en tándem, es decir, un dominio de unión al ADN, así como también señales de localización nuclear (NLS) y un dominio de activación transcripcional ácido. Los miembros de esta familia de efectores están altamente conservados y difieren principalmente en la secuencia de aminoácidos de sus repeticiones y en el número de repeticiones. El número y el orden de las repeticiones en una proteína efectora TAL determinan su actividad específica. Estas repeticiones son referidas en la presente descripción como "motivos efectores TAL". Un miembro ilustrativo de esta familia de efectores, AvrBs3 de *Xanthomonas campestris* pv. vesicatoria, contiene 17,5 repeticiones e induce la expresión de los genes UPA (regulado positivamente por AvrBs3), que incluye el gen de resistencia Bs3 en plantas de pimiento (Kay, y otros 2005 *Mol Plant Microbe Interact* 18(8): 838-48; Kay, S. y U. Bonas 2009 *Curr Opin Microbiol* 12(1): 37-43). Las repeticiones de AvrBs3 son esenciales para la unión al ADN de AvrBs3 y representan un tipo distinto de dominio de unión al ADN. El mecanismo de reconocimiento del ADN en una secuencia específica se dilucidó mediante estudios recientes de las proteínas AvrBs3, Hax2, Hax3 y Hax4 que revelaron el código de reconocimiento del ADN de los efectores TAL (Boch, J., y otros 2009 *Science* 326: 1509-12).

Los motivos o repeticiones efectores Tal son motivos de secuencia de proteína de 32 a 34 aminoácidos. Las secuencias de aminoácidos de las repeticiones son conservadas, excepto para dos residuos adyacentes altamente variables (en las posiciones 12 y 13) que determinan la especificidad hacia las bases del ADN A, G, C o T. En otras palabras, la unión al ADN se realiza mediante el contacto de un nucleótido de la doble hélice del ADN con los residuos variables en la posición 12 y 13 dentro del motivo efector TAL de una proteína efectora TAL particular (Boch, J., y otros 2009 *Science* 326: 1509-12). Por lo tanto, se encontró una correspondencia uno a uno entre las repeticiones secuenciales en las proteínas efectoras TAL y los nucleótidos secuenciales en el ADN diana. Cada motivo efector TAL reconoce principalmente un solo nucleótido dentro del sustrato de ADN. Por ejemplo, la combinación de histidina en la posición 12 y ácido aspártico en la posición 13 se une específicamente a la citosina; la combinación de asparagina en ambas, la posición 12 y en la posición 13 se une específicamente a la guanina; la combinación de asparagina en la posición 12 e isoleucina en la posición 13 se une específicamente a la adenina y la combinación de asparagina en la posición 12 y la glicina en la posición 13 se une específicamente a la timina. La unión a secuencias de ADN más largas se logra mediante la unión en tándem de varios de estos motivos efectores TAL para formar un "dominio de unión al ADN de una proteína efectora TAL". Por lo tanto, un dominio de unión al ADN de una proteína efectora TAL se refiere a los dominios de unión al ADN que se encuentran en las proteínas efectoras TAL de origen natural, así como también a los dominios de unión al ADN diseñados para unirse a una secuencia de nucleótidos diana específica como se describió en los ejemplos más abajo. El uso de tales dominios de unión al ADN de las proteínas efectoras TAL para la generación de las proteínas de fusión- motivo efector TAL/ nucleasa- que reconocen y escinden una secuencia diana específica depende de la generación adecuada de dominios de unión al ADN de las proteínas efectoras TAL que pueden reconocer específicamente dicha diana particular. Los métodos para la generación de dominios de unión al ADN de las proteínas efectoras TAL se conocen bien en la técnica (Zhang y otros (2011). *Nat Biotechnol.* 29, 149-153; Cermak y otros (2011). *Nucleic Acids Res.* April 14, identificador de PubMed 21493687).

Preferentemente, el dominio de unión al ADN se deriva de los motivos efectores TAL que se encuentran en las proteínas efectoras TAL de origen natural, tales como por ejemplo las proteínas efectoras TAL seleccionadas del grupo que consiste en AvrBs3, Hax2, Hax3 o Hax4 (Bonas y otros 1989. *Mol Gen Genet* 218(1): 127-36; Kay y otros 2005 *Mol Plant Microbe Interact* 18(8): 838-48).

Se contemplan, de acuerdo con la presente invención, proteínas de fusión que se proporcionan como un dominio de unión al ADN de una proteína efectora TAL acoplada con un único dominio nucleasa. Estas proteínas monoméricas pueden combinarse para actuar como un dímero funcional para desarrollar actividad nucleasa a través de la cooperación de dos dominios nucleasa, cada uno de los cuales es parte de una proteína de fusión.

Preferentemente, la nucleasa TAL de acuerdo con la presente invención comprende más de uno, es decir, varios motivos efectores TAL, tal como al menos 12 motivos efectores TAL, tal como, por ejemplo, al menos 14 o al menos 16 motivos efectores TAL. Con mayor preferencia, la nucleasa TAL comprende al menos 18 motivos efectores TAL. En otras palabras, el dominio de unión al ADN de una proteína efectora TAL dentro de dicha proteína de fusión comprende al menos 18 motivos efectores TAL. En el caso de proteínas de fusión que consisten en dímeros, como se describió anteriormente, esto significa que cada monómero de proteína de fusión comprende al menos nueve motivos efectores TAL. Los métodos para evaluar la especificidad de unión al ADN de una proteína de fusión de acuerdo con la presente invención son conocidos por el experto en la técnica e incluyen, sin limitación, los ensayos de genes reporteros de la transcripción y los ensayos de cambio de movilidad electroforética (EMSA).

Preferentemente, el sitio de unión de la proteína de fusión es de hasta 500 nucleótidos, tal como hasta 250 nucleótidos, hasta 100 nucleótidos, hasta 50 nucleótidos, hasta 25 nucleótidos, hasta 10 nucleótidos tal como hasta 5 nucleótidos en el extremo 5' (es decir, 5') o en el extremo 3' (es decir 3') del nucleótido(s) que es/son modificados de acuerdo con el método de la presente invención como se detalla más abajo.

En otra modalidad, la invención se refiere a un vector que codifica la molécula de ácido nucleico de la invención.

El término "vector" de acuerdo con la invención significa, preferentemente, un plásmido, cósmido, virus, bacteriófago u otro vector usado, por ejemplo, convencionalmente en ingeniería genética, que es portador de la molécula de ácido nucleico de la invención que codifica el péptido o la proteína de fusión de la invención. En consecuencia, la molécula de

ácido nucleico de la invención puede insertarse en diversos vectores disponibles comercialmente. Los ejemplos no limitantes incluyen, vectores de plásmidos procarióticos, tales como los de las series pUC, pBluescript (Stratagene), las series pET de vectores de expresión (Novagen) o pCROTOPO (Invitrogen) y vectores compatibles con la expresión en células de mamíferos como pREP (Invitrogen), pcDNA3 (Invitrogen), pCEP4 (Invitrogen), pMC1neo (Stratagene), pXT1 (Stratagene), pSG5 (Stratagene), EBO-pSV2neo, pBPV-1, pdBPVMMTneo, pRSVgpt, pRSVneo, pSV2-dhfr, pLZD35, pLXIN, pSIR (Clontech), pIRES-EGFP (Clontech), pEAK-10 (Edge Biosystems) pTriEx-Hygro (Novagen) y pCINeo (Promega). Los ejemplos para los vectores de plásmidos adecuados para *Pichia pastoris* comprenden, por ejemplo, los plásmidos pAO815, pPIC9K y pPIC3.5K (todos de Invitrogen).

La molécula de ácido nucleico de la presente invención mencionada anteriormente puede insertarse, además, en vectores de manera que se genere una fusión traduccional (adicional) con otra molécula de ácido nucleico. Para este objetivo, puede aplicarse el PCR por extensión de superposición (por ejemplo Wurch, T., Lestienne, F., y Pauwels, P.J., A modified overlap extension PCR method to create chimeric genes in the absence of restriction enzymes, Biotechn. Techn. 12, 9, Sept. 1998, 653-657). Los productos que surgen de este proceso se denominan proteínas de fusión y se describirán adicionalmente más abajo. Las otras moléculas de ácido nucleico pueden codificar una proteína que puede, por ejemplo, aumentar la solubilidad y/o facilitar la purificación de la proteína codificada por la molécula de ácido nucleico de la invención. Los ejemplos no limitantes incluyen pET32, pET41, pET43. Los vectores pueden contener, además, un ácido nucleico expresable adicional que codifica para una o más chaperonas con el propósito de facilitar el plegamiento correcto de la proteína. Los huéspedes para la expresión bacteriana adecuados comprenden, por ejemplo, las cepas derivadas a partir de BL21 (tales como BL21 (DE3), BL21 (DE3)PlysS, BL21 (DE3)RIL, BL21 (DE3)PRARE) o Rosetta®.

Los plásmidos particularmente preferidos que pueden usarse para introducir el ácido nucleico que codifica el polipéptido de la invención, que tiene la actividad de una endonucleasa, en la célula huésped son: pUC18/19 (Roche Biochemicals), pBluescript II (Alting-Mees, y otros (1992). Meth. Enzymol., 216, 483-495), pKK-177-3H (Roche Biochemicals), pBTac2 (Roche Biochemicals), pKK223-3 (Amersham Pharmacia Biotech), pKK-233-3 (Stratagene) y pET (Novagen).

Para las técnicas de modificación de vectores, ver Sambrook y Russel, 2001. Generalmente, los vectores pueden contener uno o más orígenes de replicación (ori) y sistemas de herencia para la clonación o expresión, uno o más marcadores para la selección en el huésped, por ejemplo, de resistencia a antibióticos, y uno o más casetes de expresión. Los orígenes de replicación adecuados incluyen, por ejemplo, los orígenes de replicación Col E1, el SV40 viral y el M13.

Las secuencias codificantes insertadas en el vector pueden, por ejemplo, sintetizarse mediante métodos estándar, o aislarse a partir de fuentes naturales. La ligación de las secuencias codificantes a elementos reguladores de la transcripción y/o a otras secuencias codificantes de aminoácidos puede realizarse mediante el uso de métodos establecidos. Los elementos reguladores de la transcripción (partes de un casete de expresión) que aseguran la expresión en células procarióticas o eucarióticas se conocen bien por los expertos en la técnica. Estos elementos comprenden secuencias reguladoras que garantizan el inicio de la transcripción (por ejemplo, codón de inicio de la traducción, secuencias de terminación de la transcripción, promotores, potenciadores y/o aisladores), sitios internos de entrada ribosomal (IRES) y, opcionalmente, señales poli-A que garantizan la terminación de la transcripción y la estabilidad del transcrito. Los elementos reguladores adicionales pueden incluir potenciadores transcripcionales así como también traduccionales, y/o regiones promotoras heterólogas o asociadas naturalmente. Los elementos reguladores pueden ser elementos reguladores heterólogos. Preferentemente, la molécula de ácido nucleico de la invención se une operativamente a tales secuencias de control de la expresión que permiten la expresión en células procarióticas o eucarióticas. El vector puede comprender, además, secuencias de nucleótidos que codifican señales de secreción como elementos reguladores adicionales. Tales secuencias se conocen bien por el experto en la técnica. Además, en dependencia del sistema de expresión usado, las secuencias líder, capaces de dirigir el polipéptido expresado hacia un compartimento celular, pueden añadirse a la secuencia codificante de la molécula de ácido nucleico de la invención. Tales secuencias líder se conocen bien en la técnica. Los vectores diseñados específicamente permiten la transferencia de ADN entre diferentes huéspedes, tales como entre las células bacterianas-fúngicas o células bacterianas-animales.

La cotransfección con un marcador de selección, tal como kanamicina o genes de resistencia a la ampicilina para el cultivo en *E. coli* y otras bacterias, permite la identificación y el aislamiento de las células transfectadas. Los marcadores de selección para el cultivo de células de mamíferos son los genes de resistencia dhfr, gpt, neomicina, e higromicina. Además, el ácido nucleico transfectado puede amplificarse, para expresar grandes cantidades del polipéptido codificado. El marcador DHFR (dihidrofolato reductasa) es útil para desarrollar líneas celulares que portan varios cientos o incluso varios miles de copias del gen de interés. Otro marcador de selección útil es la enzima glutamina sintasa (GS) (Fisher y otros, Infect Immun. 1991 Oct;59(10):3562-5; Bebbington y otros, Biotechnology (N Y). 1992 Feb;10(2):169-75).

Mediante el uso de tales marcadores, las células se cultivan en medio selectivo y se seleccionan las células con la mayor resistencia.

En otra modalidad, la invención se refiere a una célula huésped que comprende, por ejemplo, como un resultado de la transformación, transducción, microinyección o transfección, la molécula de ácido nucleico o el vector de la invención.

Puede concebirse una diversidad de sistemas de expresión del huésped para expresar la secuencia codificante de la endonucleasa en una célula huésped mediante el uso de un vector adecuado.

La "célula huésped" de acuerdo con la invención puede producirse mediante la introducción de la molécula de ácido nucleico o el vector(es) de la invención en la célula huésped que, en su presencia, preferentemente, media la expresión de la molécula de ácido nucleico de la invención que codifica la endonucleasa de la invención. El huésped a partir del cual se deriva la célula huésped puede ser cualquier célula procariótica o eucariótica.

Una célula huésped eucariótica adecuada puede ser una célula de vertebrado, una célula de anfibio, una célula de pez, una célula de insecto, una célula de hongo/levadura, una célula de nemátodo o una célula vegetal. La célula de insecto puede ser una célula de *Spodoptera frugiperda*, una célula de *Drosophila* S2 o una célula de *Spodoptera* Sf9, la célula de hongo/levadura puede ser una célula de *Saccharomyces cerevisiae*, célula de *Pichia pastoris* o una célula de *Aspergillus*. Se prefiere que la célula de vertebrado sea una célula de mamífero tal como una célula humana, CHO, COS, 293 o células de melanoma de Bowes. La célula vegetal se selecciona, preferentemente, independientemente de una célula de *Anacardium*, *Anona*, *Arachis*, *Artocarpus*, *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Carica*, *Citrus*, *Citrullus*, *Capsicum*, *Carthamus*, *Cocos*, *Coffea*, *Cucumis*, *Cucurbita*, *Daucus*, *Elaeis*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoseyamus*, *Lactuca*, *Linum*, *Lolium*, *Lupinus*, *Lycopersicon*, *Malus*, *Manihot*, *Majorana*, *Medicago*, *Nicotiana*, *Olea*, *Oryza*, *Panicum*, *Pannisetum*, *Passiflora*, *Persea*, *Phaseolus*, *Pistachia*, *Pisum*, *Pyrus*, *Prunus*, *Psidium*, *Raphanus*, *Ricinus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Theobromus*, *Trigonella*, *Triticum*, *Vicia*, *Vitis*, *Vignay* Zea. La célula puede ser una parte de una línea celular. La célula vegetal puede, por ejemplo, derivarse a partir de la raíz, la hoja, la corteza, la aguja, el tronco o el tallo.

Los procariotas adecuados (bacterias) útiles como huésped para la invención son aquellas que se usan generalmente para la clonación y/o la expresión, como *E. coli* (por ejemplo, las cepas de *E. coli* BL21, HB101, DH5a, XL1 Blue, Y1090 y JM101), *Salmonella typhimurium*, *Serratia marcescens*, *Burkholderia glumae*, *Pseudomonas putida*, *Pseudomonas fluorescens*, *Pseudomonas stutzeri*, *Streptomyces lividans*, *Lactococcus lactis*, *Mycobacterium smegmatis*, *Streptomyces* o *Bacillus subtilis*. Los medios y condiciones de cultivo apropiados para las células huésped descritas anteriormente son conocidos en la técnica.

Los ejemplos preferidos para que la célula huésped sea modificada mediante ingeniería genética con la molécula de ácido nucleico o el vector(es) de la invención es una célula de levadura, *E. coli* y/o una especie del género *Bacillus* (por ejemplo, *B. subtilis*). La célula huésped más preferida es *Bacillus spec.*

En una modalidad adicional, la invención se refiere a un método para producir una proteína o fusión que tiene la actividad de una endonucleasa como se definió anteriormente en la presente descripción que comprende las etapas: (a) cultivar la célula huésped de la invención y (b) aislar la proteína que se produce o la proteína de fusión que tiene la actividad de dicha endonucleasa.

Las condiciones adecuadas para el cultivo de un huésped procariótico o eucariótico se conocen bien por el experto en la técnica. Las condiciones adecuadas para cultivar *E. coli* DH18BΔkat E (Invitrogen), *Pichia pastoris* o *Aspergillus niger* son, por ejemplo, proporcionadas en los ejemplos de la invención. En general, las condiciones adecuadas para el cultivo de bacterias son las que facilitan el crecimiento de estas bajo aireación, en medio Luria Bertani (LB). Para aumentar el rendimiento y la solubilidad del producto de expresión, el medio puede tamponarse o complementarse con aditivos adecuados conocidos porque mejoran o facilitan ambos. La *E. coli* puede cultivarse desde 4 a aproximadamente 37 °C, la temperatura exacta o la secuencia de temperaturas depende de la molécula que se va a expresar en exceso. En general, *Aspergillus sp.* puede crecerse sobre agar Sabouraud dextrosa, o agar de patata dextrosa a aproximadamente de 10 °C a aproximadamente 40 °C, y preferentemente, a aproximadamente 25 °C. Se conocen las condiciones adecuadas para los cultivos de levadura, por ejemplo de Guthrie y Fink, "Guide to Yeast Genetics and Molecular Cell Biology" (2002); Academic Pr Inc.. El experto en la técnica es consciente, además, de todas estas condiciones y puede, adaptar estas condiciones aún más a las necesidades de una especie huésped particular y a los requisitos del polipéptido expresado. En el caso de que un promotor inducible controle el ácido nucleico de la invención en el vector presente en la célula huésped, la expresión del polipéptido puede inducirse mediante la adición de un agente inductor apropiado. Los protocolos y estrategias de expresión adecuados son conocidos por el experto en la técnica.

En dependencia del tipo de célula y de sus requisitos específicos, el cultivo de células de mamíferos puede realizarse, por ejemplo, en medio RPMI o DMEM que contiene FCS 10 % (v/v), L-glutamina 2mM y penicilina/estreptomicina 100 U/mL. Las células pueden mantenerse a 37 °C en una atmósfera saturada de agua, al 5 % de CO₂.

Los protocolos de expresión adecuados para células eucarióticas se conocen bien por el experto en la técnica y pueden retomarse, por ejemplo, de Sambrook, 2001.

Los métodos de aislamiento del polipéptido producido se conocen bien en la técnica y comprenden, sin limitación, las etapas de los métodos tales como cromatografía de intercambio iónico, cromatografía de filtración en gel (cromatografía de exclusión por tamaño), cromatografía de afinidad, cromatografía líquida de alta presión (HPLC), HPLC de fase inversa, electroforesis en gel de disco o inmunoprecipitación, ver, por ejemplo, Sambrook, 2001.

La etapa de aislamiento de proteínas es, preferentemente, una etapa de la purificación de proteínas. La purificación de proteínas, de acuerdo con la invención, especifica un proceso o una serie de procesos destinados a aislar más aún el polipéptido de la invención a partir de una mezcla compleja, preferentemente, para la homogeneidad. Las etapas de

purificación, por ejemplo, explotan las diferencias en el tamaño de las proteínas, las propiedades fisicoquímicas y la afinidad de unión. Por ejemplo, las proteínas pueden purificarse de acuerdo con sus puntos isoeléctricos mediante su paso a través de un gel con un gradiente de pH o una columna de intercambio iónico. Además, las proteínas pueden separarse de acuerdo con su tamaño o peso molecular mediante cromatografía de exclusión por tamaño, o mediante el análisis de SDS-PAGE (electroforesis en gel de dodecil sulfato de sodio y poliacrilamida). En la técnica, con frecuencia las proteínas se purifican mediante el uso de 2D-PAGE y después se analizan más aún mediante la huella peptídica para establecer la identidad de la proteína. Esto es muy útil para fines científicos y los límites de detección para las proteínas son muy bajos, y para su análisis son suficientes las cantidades en nanogramos de proteína. Además, las proteínas pueden separarse por su polaridad/hidrofobicidad mediante cromatografía líquida de alto rendimiento o cromatografía de fase inversa. Por lo tanto, los métodos para la purificación de proteínas se conocen bien por el experto en la técnica.

Además, en una modalidad la invención se refiere a una proteína o proteína de fusión que tiene la actividad de una endonucleasa codificada por la molécula de ácido nucleico o vector de la invención.

Las definiciones para las proteínas o proteínas de fusión que tienen la actividad de una endonucleasa codificada por la molécula de ácido nucleico o el vector de la invención, proporcionadas en las modalidades anteriores, pertenecientes a la molécula de ácido nucleico o vector de la invención se aplican explícitamente, además, a esta modalidad.

Como una consecuencia de su actividad endonucleasa, otra modalidad de la invención se refiere al uso de la proteína o proteína de fusión de la invención para escindir una molécula de ácido nucleico, por ejemplo, en uno de los métodos de la invención descritos más abajo.

Adicionalmente, la presente invención se refiere, además, a un kit que comprende la molécula de ácido nucleico, la proteína y/o la proteína de fusión de la invención. Los diversos componentes del kit pueden empaquetarse en uno o más contenedores, tales como uno o más frascos. Los frascos pueden, además de los componentes, comprender conservantes o tampones para almacenamiento. Además, el kit puede contener las instrucciones para usar.

En otra modalidad, la invención se refiere a un método *in vitro* para modificar una secuencia diana en el genoma de una célula eucariótica, el método comprende la etapa: (a) introducir en dicha célula la molécula de ácido nucleico, el vector o la proteína o proteína de fusión como se define en las reivindicaciones.

El término "modificar" como se usa de acuerdo con la presente invención se refiere a manipulaciones genómicas aleatorias y en sitios específicos que resultan en cambios en la secuencia de nucleótidos del genoma del huésped eucariótico. Cuando se introduce la proteína de fusión de la invención, se logra la modificación en un sitio específico de dicha "secuencia diana" en el genoma a través del dominio de unión al ADN. Cuando solamente se introduce la proteína de la invención, la "secuencia diana" no es una secuencia específica, porque la endonucleasa novedosa no es específica de un sitio. Por lo tanto, la proteína de la invención puede usarse para introducir mutaciones aleatorias en un genoma, es decir, la "secuencia diana" se produce múltiples veces en el genoma y no depende de un motivo específico de una secuencia. El material genético que comprende estos cambios en su secuencia de nucleótidos es referido, además, en la presente descripción como la "secuencia diana modificada" cuando la modificación es en un sitio específico como, por ejemplo, en el caso de usar la proteína de fusión de la invención. El término "modificación" incluye, pero no se limita a, sustitución, inserción y delección de uno o más nucleótidos dentro de la secuencia diana. En el proceso de recombinación homóloga, el producto final puede reflejar una delección de secuencias. Como comprenderá el experto en la técnica, una recombinación homóloga, por otra parte, siempre incluye, además, la incorporación de material genético de la secuencia de ADN donante, que en esta modalidad, sin embargo, conduce a una delección total. Se comprenderá por el experto en la técnica que, mediante la simple introducción de roturas de doble hebra en el genoma de una célula, pueden introducirse modificaciones que son el resultado de una recombinación homóloga (en la presencia y ausencia de secuencias donantes exógenas) o una reparación endógena del ADN, mecanismos tales como, por ejemplo, la reparación de ADN mediante la unión de extremos no homólogos (NHEJ), que es propensa a la introducción de pequeñas delecciones en el sitio de la rotura de doble hebra en el curso de la ligadura de los extremos rotos.

El término "sustitución", como se usa en la presente descripción, se refiere al reemplazo de nucleótidos con otros nucleótidos. El término incluye, por ejemplo, el reemplazo de nucleótidos individuales que resultan en mutaciones puntuales. Dichas mutaciones puntuales pueden conducir a un intercambio de aminoácidos en el producto proteico resultante, pero además, pueden no reflejarse a nivel de los aminoácidos. El término "sustitución" contempla, además, las mutaciones que resultan en el reemplazo de múltiples nucleótidos, tales como, por ejemplo, partes de genes, tales como partes de exones o intrones, así como también el reemplazo de genes completos.

El término "inserción" de acuerdo con la presente invención se refiere a la incorporación de uno o más nucleótidos en una molécula de ácido nucleico. El término "inserción" contempla, además, la inserción de partes de genes, tales como partes de exones o intrones, así como también la inserción de genes completos. Cuando el número de nucleótidos insertados no es divisible por tres, la inserción puede resultar en una mutación con desplazamiento del marco de lectura dentro de una secuencia codificante de un gen. Tales mutaciones con desplazamiento del marco de lectura alterarán los aminoácidos codificados por un gen después de la mutación. En algunos casos, tal mutación causará que la traducción activa del gen encuentre un codón de parada prematuro, lo que resulta en una terminación de la traducción y en la producción de una proteína truncada. Cuando el número de nucleótidos insertados es, en cambio, divisible por tres, la

inserción resultante es una "inserción en el marco de lectura". En este caso, el marco de lectura permanece intacto después de la inserción y la traducción probablemente se completará si los nucleótidos insertados no codifican para un codón de parada. Sin embargo, debido a los nucleótidos insertados, la proteína resultante contendrá, en dependencia del tamaño de la inserción, uno o varios aminoácidos nuevos que pueden afectar la función de la proteína.

El término "delección" como se usa de acuerdo con la presente invención se refiere a la pérdida de nucleótidos o parte de genes, tales como exones o intrones, así como también a genes completos. Como se definió con respecto al término "inserción", la delección de una cantidad de nucleótidos que no es divisible de manera equitativa entre tres, conducirá a una mutación con desplazamiento del marco de lectura que causa que todos los codones que se producen después de la delección se lean incorrectamente durante la traducción, lo que produce, potencialmente, una proteína gravemente alterada y, muy probablemente, no funcional. Si una delección no resulta en una mutación con desplazamiento del marco de lectura, es decir, debido a que el número de nucleótidos delecionados puede dividirse en tres, la proteína resultante no obstante se ve alterada, ya que carecerá, en dependencia del tamaño de la delección, de varios aminoácidos que pueden afectar o efectuar la función de la proteína.

Las modificaciones definidas anteriormente no se limitan a las regiones codificantes en el genoma, sino que pueden ocurrir, además, en regiones no codificantes del genoma diana, por ejemplo, en regiones reguladoras tales como elementos promotores, o potenciadores, o en intrones.

Los ejemplos de modificaciones del genoma diana incluyen, sin ser limitantes, la introducción de mutaciones en un gen de tipo silvestre para analizar su efecto sobre la función del gen; el reemplazo de un gen completo con un gen mutado o, alternativamente, si la secuencia diana comprende mutaciones, la alteración de estas mutaciones para identificar que mutación es la causante de un efecto particular; la eliminación de genes completos o proteínas o la eliminación de elementos reguladores de genes o proteínas, así como también la introducción de asociados de fusión, tales como por ejemplo etiquetas de purificación tales como la etiqueta his o la etiqueta tap, etcétera. En el último caso, el término "adición" puede usarse, además, en lugar de "inserción" para describir la adición preferente de una etiqueta a un extremo de un polipéptido más que dentro de la secuencia de un polipéptido.

El término "célula eucariótica", como se usa en la presente descripción, se refiere a cualquier célula de un organismo eucariótico unicelular o multicelular, que incluye las células de animales como los vertebrados y de los hongos y plantas. Preferentemente, pero sin limitación, la célula es una célula de mamífero. El término "célula de mamífero", como se usa en la presente descripción, se conoce bien en la técnica y se refiere a cualquier célula que pertenece a un animal que se agrupa en la clase de mamíferos. El término "célula", como se usa en relación con la presente descripción, puede referirse a una célula única y/o aislada o a una célula que es parte de una entidad multicelular, tal como un tejido, un organismo o a una célula de otro cultivo. En otras palabras, el método puede realizarse *in vivo*, *ex vivo* o *in vitro*. El método de la invención se realiza *in vitro*, como se define en las reivindicaciones. En dependencia del objetivo particular para lograrse a través de la modificación del genoma de una célula de mamífero, pueden usarse células de diferentes subclases de mamíferos, tales como prototerios o terios. Por ejemplo, dentro de la subclase de terios, preferentemente, células de animales de la subclase euteria, con mayor preferencia del orden primates, artiodactyla, perissodactyla, rodentia y lagomorpha se usan en el método de la invención como se detalla más abajo. Además, dentro de una especie, puede elegirse una célula para usar en el método de la invención basado en el tipo de tejido y/o la capacidad para diferenciarse igualmente en dependencia del objetivo para lograrse al modificar el genoma. Tres categorías básicas de células conforman el cuerpo de los mamíferos: células germinales, células somáticas y células madre. Una célula germinal es una célula que da lugar a gametos y, por lo tanto, es continua a través de las generaciones. Las células madre pueden dividirse y diferenciarse en diversos tipos de células especializadas, así como también autorrenovarse para producir más células madre. En los mamíferos hay dos tipos principales de células madre: células madre embrionarias y células madre adultas. Las células somáticas incluyen todas las células que no son gametos, gametocitos o células madre indiferenciadas. Las células de un mamífero pueden agruparse, además, por su capacidad para diferenciarse. Una célula totipotente (conocida además como omnipotente) es una célula que es capaz de diferenciarse en todos los tipos de células de un organismo adulto, que incluyen el tejido placentario, tal como un cigoto (ovocito fecundado) y blastómeros posteriores, mientras que las células pluripotentes, tales como células madre embrionarias, no pueden contribuir al tejido extraembrionario tal como la placenta, pero tienen el potencial de diferenciarse en cualquiera de las tres capas germinales endodermo, mesodermo y ectodermo. Las células progenitoras multipotentes tienen el potencial de dar lugar a células de múltiples, pero número limitado de linajes celulares. Además, hay células oligopotentes que pueden desarrollarse solamente en unos pocos tipos de células y células unipotentes (a veces denominadas células precursoras) que pueden convertirse solamente en un tipo de célula. Hay cuatro tipos básicos de tejidos: tejido muscular, tejido nervioso, tejido conectivo y tejido epitelial del que pueden derivarse una célula que va a usarse en el método de la invención, tal como por ejemplo, células madre hematopoyéticas o células madre neuronales. En la medida en que se prevean células humanas para usar en el método de la invención, se prefiere que tal célula humana no se obtenga a partir de un embrión humano, en particular no a través de métodos que impliquen la destrucción de un embrión humano. Por otra parte, las células madre embrionarias humanas están a disposición de los expertos en la técnica, tal como las que se toman a partir de las líneas de células madre embrionarias existentes, comercialmente disponibles. En consecuencia, en la presente invención puede trabajarse con células madre embrionarias humanas sin necesidad de usar o destruir un embrión humano. Alternativamente, o en lugar de células madre embrionarias humanas, pueden usarse células pluripotentes que se asemejan a las células madre embrionarias tales como las células madre pluripotentes inducidas (iPS), cuya generación se establece como estado de la técnica (Hargus G y otros, Proc Natl Acad Sci U S A 107:15921-15926;

Jaenisch R. y Young R., 2008, Cell 132:567-582; Saha K, y Jaenisch R., 2009, Cell Stem Cell 5:584-595).

El término "moléculas de ácido nucleico que codifican dicha proteína o proteína de fusión en forma expresable" se refiere a una molécula de ácido nucleico que, al expresarse en una célula o un sistema libre de células, resulta en una proteína funcional o proteína de fusión de la invención. Preferentemente, pero sin limitación, dicho ácido nucleico es ARNm. Alternativamente, pueden usarse el ADN que tiene señales de transcripción apropiadas para permitir la expresión, o el ADNc.

La introducción de la proteína, proteína de fusión, o de la molécula de ácido nucleico que codifica dicha proteína, y la proteína de fusión en forma expresable dentro de una célula, pueden lograrse mediante los métodos conocidos en la técnica, y depende de la naturaleza de dichas proteínas o moléculas de ácido nucleico. Por ejemplo, y en el caso de la introducción de moléculas de ácido nucleico, dicha introducción puede lograrse mediante métodos basados en químicos (fosfato de calcio, liposomas, DEAE-dextrán, polietilimina, nucleofección), métodos no químicos (electroporación, sonoporación, transfección óptica, electrotransferencia de genes, administración hidrodinámica), métodos basados en partículas (pistola de genes, magnetofección, impalefacción) y métodos virales. Preferentemente, las moléculas de ácido nucleico son para introducirse en el núcleo mediante métodos tales como, por ejemplo, microinyección o nucleofección. Los métodos para realizar la microinyección se conocen bien en la técnica y se describen, por ejemplo, en Nagy y otros (Nagy A, Gertsenstein M, Vintersten K, Behringer R., 2003. Manipulating the Mouse Embryo. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press) así como también en los ejemplos de la presente descripción más abajo. El experto en la técnica entiende que, en dependencia del método de introducción, puede ser ventajoso adaptar las moléculas de ADN. Por ejemplo, una molécula de ADN lineal puede ser más eficiente en los eventos de recombinación homóloga cuando se usa la electroporación como método para introducir dicha molécula de ADN en, por ejemplo, una célula de mamífero, mientras que una molécula de ADN circular puede ser más ventajosa cuando se inyectan las células.

Todas las definiciones y modalidades preferidas definidas anteriormente con respecto a la molécula de ácido nucleico, proteína o proteína de fusión de la invención se aplican, además, *mutatis mutandis* en el contexto del método de la invención.

De acuerdo con la presente invención, el término "secuencia diana en el genoma" se refiere a la ubicación genómica que se va a modificar por el método de la invención. La "secuencia diana en el genoma" comprende, pero no se restringe a los nucleótidos sujetos a la modificación particular. Además, y preferentemente con respecto a la proteína de fusión de la invención, el término "secuencia diana en el genoma" comprende, además, las regiones para la unión de secuencias homólogas de una segunda molécula de ácido nucleico. En otras palabras, el término "secuencia diana en el genoma" comprende, además, la secuencia que flanquea/rodea el nucleótido(s) pertinente que va a modificarse. En algunos casos, el término "secuencia diana" puede referirse, además, al gen completo que va a modificarse.

La unión específica se definió anteriormente en la presente descripción y garantiza que solamente las roturas de doble hebra se introduzcan dentro de dicha secuencia diana.

En una modalidad más preferida del método de la invención, la modificación de dicha secuencia diana es mediante recombinación homóloga con una secuencia de ácido nucleico donante, que comprende además la etapa: (b) introducir una molécula de ácido nucleico en dicha célula, en donde dicha molécula de ácido nucleico comprende dicha secuencia de ácido nucleico donante, en donde dicha secuencia de ADN donante está flanqueada hacia el extremo 5' por un primer elemento flanqueante y hacia el extremo 3' por un segundo elemento flanqueante, en donde dicho primer y segundo elementos flanqueantes son diferentes y en donde cada uno de dichos primer y segundo elementos flanqueantes son homólogos a una secuencia continua de ADN a cada lado de la rotura de la doble hebra introducida en (a) del método de la invención dentro de dicha secuencia diana en el genoma de dicha célula eucariótica.

El término "recombinación homóloga", se usa de acuerdo con las definiciones proporcionadas en la técnica. Por lo tanto, se refiere a un mecanismo de recombinación genética en el que dos hebras de ADN que comprenden secuencias de nucleótidos similares intercambian material genético. Las células usan la recombinación homóloga, durante la meiosis, donde esta sirve para reorganizar el ADN para crear un conjunto totalmente único de cromosomas haploides, pero además, para la reparación del ADN dañado, en particular para la reparación de roturas de doble hebra. El mecanismo de recombinación homóloga se conoce bien por el experto en la técnica y se ha descrito, por ejemplo, por Paques y Haber (Paques F, Haber JE.; Microbiol Mol Biol Rev 1999; 63:349-404). En el método de la presente invención, la recombinación homóloga de la secuencia donante se permite por la presencia de dicho primer y dicho segundo elemento flanqueante que se ubican hacia el extremo 5' (5') y hacia el extremo 3' (3'), respectivamente, de dicha secuencia de ADN donante cada uno de los cuales es homólogo a una secuencia de ADN continua dentro de dicha secuencia diana.

De acuerdo con la presente invención, el término "secuencia de ADN donante" se refiere a una secuencia de ADN que sirve como un molde en el proceso de recombinación homóloga y que porta la modificación que va a introducirse dentro la secuencia diana. Mediante el uso de esta secuencia de ADN donante como un molde, la información genética, que incluye las modificaciones, se copia en la secuencia diana dentro del genoma de la célula por medio de recombinación homóloga. En los ejemplos no limitantes, la secuencia de ácido nucleico donante puede ser esencialmente idéntica a la parte de la secuencia diana que va a reemplazarse, con la excepción de un nucleótido que difiere y resulta en la introducción de una mutación puntual después de la recombinación homóloga o esta puede consistir en un gen adicional

previamente no presente en la secuencia diana. Es concebible que la naturaleza de la secuencia de ADN donante, es decir, su longitud, composición de bases, similitud con la secuencia diana, dependa de cómo se modificará la secuencia diana, así como también el objetivo particular que se logrará mediante la modificación de la secuencia diana. Los expertos en la técnica entenderán que dicha secuencia de ADN donante está flanqueada por secuencias que son homólogas a las secuencias dentro de la secuencia diana para permitir que tenga lugar una recombinación homóloga que conduce a la incorporación de la secuencia de ADN donante en el genoma de dicha célula. Además de ser homólogos a una secuencia continua de ADN dentro del ADN genómico, el primer y el segundo elementos flanqueantes son diferentes para permitir que tenga lugar una recombinación homóloga dirigida.

El término "homólogo a una secuencia continua de ADN a cada lado de la rotura de doble hebra introducida en (a) del método de la invención dentro de dicha secuencia diana", de acuerdo con la presente invención, se refiere a regiones que tienen suficiente identidad de secuencia para asegurar la unión específica a las secuencias diana que se encuentran hacia los extremos 5' y 3' de la ubicación de la rotura de doble hebra. El término "homólogo" como se usa en la presente descripción puede intercambiarse con el término "idéntico" como se describe en la presente descripción en otra parte con respecto a los niveles variables de identidad de secuencia. Los métodos para evaluar el nivel de identidad entre dos secuencias de ácido nucleico se conocen bien en la técnica y se describieron anteriormente en la presente descripción. Estos métodos involucran programas, además de proporcionar un alineamiento de secuencia por pares, informan, además, el nivel de identidad de la secuencia (generalmente en porcentaje de identidad) y la probabilidad de que ocurra el alineamiento por azar (valor de P) y pueden usarse, además, para predecir la aparición de la unión específica.

Preferentemente, dicho primer y segundo elementos flanqueantes son "homólogos a una secuencia de ADN continua dentro de dicha secuencia diana" (denominados en la técnica, además, "brazos de homología") tienen una identidad de secuencia con la parte correspondiente de la secuencia diana de al menos 95 %, más preferido al menos 97 %, más preferido al menos 98 %, más preferido al menos 99 %, incluso más preferido al menos 99,9 % y el más preferido 100 %. Las identidades de secuencia definidas anteriormente se definen solamente con respecto a aquellas partes de la secuencia diana que sirven como sitios de unión para los brazos de homología, es decir, dicho primer y dicho segundo elementos flanqueantes. Por lo tanto, la identidad de secuencia global entre la secuencia diana completa y las regiones homólogas de la molécula de ácido nucleico de la etapa (b) del método de modificación de una secuencia diana de la presente invención puede diferir de las identidades de secuencia definidas anteriormente, debido a la presencia de la parte de la secuencia diana que va a reemplazarse por la secuencia de ADN donante.

Los elementos flanqueantes homólogos a la secuencia diana comprendida en la molécula de ADN tienen una longitud de al menos 170 pb cada uno. Preferentemente, cada uno de los elementos tienen una longitud de al menos 250 nucleótidos, al menos 300 nucleótidos, al menos 400 nucleótidos, al menos 500 nucleótidos, tal como al menos 600 nucleótidos, al menos 750 pb de nucleótidos, con mayor preferencia al menos 1000 nucleótidos, tal como al menos 1500 nucleótidos, incluso con mayor preferencia al menos 2000 nucleótidos y con la máxima preferencia al menos 2500 nucleótidos. La longitud máxima de los elementos homólogos a la secuencia diana comprendida en la molécula de ácido nucleico depende del tipo de vector de clonación usado y puede tener una longitud de hasta 20 000 nucleótidos en plásmidos de alto número de copias de E. coli que usan el origen de replicación col EI (por ejemplo, pBluescript) o hasta una longitud de 300 000 nucleótidos cada uno en plásmidos que usan el origen del factor F (por ejemplo, en vectores BAC, tal como, por ejemplo, pTARBAC1).

Las moléculas de ADN que comprenden la secuencia de ADN donante y los elementos flanqueantes se modifican, necesariamente, si el sitio de unión a la nucleasa específica de un sitio (proteína de fusión) está contenido sin interrupciones dentro de uno de los elementos flanqueantes y, preferentemente, si el sitio de unión a la nucleasa específica de un sitio (proteína de fusión) se interrumpe por la secuencia donante, es decir, una parte en cada uno de los elementos flanqueantes, de manera que la proteína de fusión no introduzca una rotura de doble hebra en la secuencia del ADN donante como parte de una molécula de ADN. Cuando la proteína de fusión es una TAL o una nucleasa dedo de zinc, esto puede lograrse, por ejemplo, mediante una modificación ya sea en el motivo de unión o de escisión (ver Ejemplo 2, Figura 12).

Se apreciará por el experto en la técnica que dicha molécula de ADN para introducirse en la célula en el ítem (b) del método de la invención puede comprender toda una molécula de ácido nucleico (secuencia) que codifica dicha proteína de fusión en forma expresable y la molécula de ácido nucleico que comprende la secuencia de ácido nucleico donante y los elementos flanqueantes homólogos a la secuencia diana. Alternativamente, la molécula de ácido nucleico del ítem (b) puede ser una molécula de ácido nucleico distinta, para introducirse adicionalmente a las moléculas de ácido nucleico que codifican dicha proteína de fusión en forma expresable del ítem (a).

En una modalidad preferida del método de la invención se contempla, además, que dicha célula se analiza para la modificación exitosa de dicha secuencia diana en el genoma.

Los métodos para analizar la presencia o ausencia de una modificación se conocen bien en la técnica e incluyen, sin limitación, ensayos basados en la separación física de moléculas de ácido nucleico, ensayos de secuenciación, así como también ensayos de escisión y digestión, y análisis del ADN mediante la reacción en cadena de la polimerasa (PCR).

Los ejemplos de ensayos basados en la separación física de moléculas de ácido nucleico incluyen, sin limitación, MALDI-

TOF, electroforesis en gel de gradiente desnaturizante y otros de tales métodos conocidos en la técnica, ver por ejemplo Petersen y otros, Hum. Mutat. 20 (2002) 253-259; Hsia y otros, Theor. Appl. Genet. 111 (2005) 218-225; Tost and Gut, Clin. Biochem. 35 (2005) 335-350; Palais y otros, Anal. Biochem. 346 (2005) 167-175.

Los ejemplos de ensayos de secuenciación comprenden, sin limitación, los métodos de análisis de secuencia mediante secuenciación directa, SSCP fluorescente en un secuenciador de ADN automatizado y pirosecuenciación. Estos procedimientos son comunes en la técnica, ver, por ejemplo, Adams y otros (Ed.), "Automated DNA Sequencing and Analysis", Academic Press, 1994; Alpey, "DNA Sequencing: From Experimental Methods to Bioinformatics", Springer Verlag Publishing, 1997; Ramon y otros, J. Transl. Med. 1 (2003) 9; Meng y otros., J. Clin. Endocrinol. Metab. 90 (2005) 3419-3422.

Los ejemplos de ensayos de escisión y digestión incluyen, sin limitación, ensayos de digestión de restricción, tales como ensayos de polimorfismo de longitud de fragmentos de restricción (ensayos de RFLP), ensayos de protección de RNasa, ensayos basados en métodos de escisión química y ensayos de escisión de desapareamientos mediante enzimas, ver por ejemplo Youil y otros, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 87-91; Todd y otros, J. Oral Maxil. Surg. 59 (2001) 660-667; Amar y otros, J. Clin. Microbiol. 40 (2002) 446-452.

Alternativamente, en vez de analizar las células para detectar la presencia o ausencia de la modificación deseada, en particular, en el caso de la modificación en una secuencia específica, las células modificadas exitosamente pueden seleccionarse mediante la incorporación de marcadores de selección apropiados. Los marcadores de selección incluyen marcadores de selección positivos y negativos, que se conocen bien en la técnica y se emplean habitualmente por el experto en la técnica. Los ejemplos no limitantes de marcadores de selección incluyen dhfr, gpt, neomicina, higromicina, dihidrofolato reductasa, G418 o glutamina sintasa (GS) (Murphy y otros, Biochem J. 1991, 227:277; Bebbington y otros, Bio/Technology 1992, 10:169). Mediante el uso de estos marcadores, las células se cultivan en medio selectivo y se seleccionan las células con la mayor resistencia. Se contemplan, además, los marcadores de selección positiva/negativa combinados, que pueden incorporarse en el genoma diana mediante recombinación homóloga o integración aleatoria. Después de la selección positiva, el primer casete que comprende el marcador de selección positiva flanqueado por sitios de reconocimiento de recombinasa se intercambia mediante intercambio de casete mediado por recombinasa contra un segundo casete sin marcador. Después, los clones que contienen el casete de intercambio deseado se obtienen mediante selección negativa.

En una modalidad preferida del método de la invención, la célula se selecciona a partir del grupo que consiste en una célula de mamífero o de vertebrado, una célula vegetal o una célula fúngica.

En otra modalidad preferida del método de la invención, la célula es un ovocito no humano.

Como se usa en la presente descripción, el término "ovocito" se refiere a la célula germinal femenina involucrada en la reproducción, es decir, el óvulo o la célula huevo. De acuerdo con la presente invención, el término "ovocito" comprende ambos, los ovocitos antes de la fecundación, así como también los ovocitos fecundados que, además, se denominan cigotos. Por lo tanto, el ovocito antes de la fecundación comprende solamente los cromosomas maternos, mientras que un ovocito después de la fecundación comprende los cromosomas maternos y paternos. Después de la fecundación, el ovocito permanece en un estado doble haploide durante varias horas, en ratones, por ejemplo, hasta 18 horas después de la fecundación. De acuerdo con la invención, el ovocito puede ser no humano.

Además, se describe un método en donde el ovocito no humano es un ovocito fecundado. El término "ovocito fecundado", como se usa en la presente descripción, se refiere a un ovocito después de la fusión con el esperma fecundante. Durante un período de muchas horas (tal como hasta 18 horas en ratones) después de la fecundación, el ovocito se encuentra en un estado doble haploide, que comprende un pronúcleo haploide materno y un pronúcleo haploide paterno. Después de la migración de los dos pronúcleos juntos, sus membranas se rompen y los dos genomas se condensan en cromosomas, de esta manera se reconstituye un organismo diploide. Preferentemente, el ovocito de mamífero o aviar no humano usado en el método descrito es un ovocito de mamífero o aviar fecundado en el estado doble haploide.

En el caso de los ovocitos para usarse como células en el método de la invención, la proteína, la proteína de fusión o la molécula de ácido nucleico que codifica dicha proteína o proteína de fusión se introducen en el ovocito mediante microinyección. La microinyección en el ovocito puede realizarse mediante inyección en el núcleo (antes de la fecundación), en el pronúcleo (después de la fecundación) y/o mediante la inyección en el citoplasma (tanto antes como después de la fecundación). Cuando se emplea un ovocito fecundado, la inyección en el pronúcleo se realiza, ya sea para un pronúcleo, o para ambos pronúcleos. La inyección de la nucleasa dedos-Tal o de un ADN que codifica la nucleasa dedos-Tal de la etapa (a) del método para modificar una secuencia diana de la presente invención es, preferentemente, dentro del núcleo/pronúcleo, mientras que la inyección de un ARNm que codifica la nucleasa dedos-Tal de la etapa (a) es, preferentemente, en el citoplasma. La inyección de la molécula de ácido nucleico de la etapa (b) es, preferentemente, dentro del núcleo/pronúcleo. Sin embargo, la inyección de la molécula de ácido nucleico de la etapa (b) puede realizarse, además, en el citoplasma cuando dicha molécula de ácido nucleico se proporciona como una secuencia de ácido nucleico que tiene una señal de localización nuclear, para asegurar su suministro dentro del núcleo/pronúcleo. Preferentemente, la microinyección se realiza mediante inyección tanto dentro del núcleo/pronúcleo como en el citoplasma. Por ejemplo, la aguja puede introducirse dentro del núcleo/pronúcleo y se inyecta una primera cantidad de la nucleasa dedos-Tal y/o la

molécula de ácido nucleico dentro del núcleo/pronúcleo. Mientras se retira la aguja del ovocito, se inyecta una segunda cantidad de la nucleasa dedos-Tal y/o la molécula de ácido nucleico en el citoplasma.

5 Los métodos para realizar la microinyección se conocen bien en la técnica y se describen, por ejemplo, en Nagy y otros (Nagy A, Gertsenstein M, Vintersten K, Behringer R., 2003. *Manipulating the Mouse Embryo*. Cold Spring Harbour, Nueva York: Cold Spring Harbour Laboratory Press) así como también en los ejemplos de la presente descripción más abajo. Se prefiere, además, que la molécula de ácido nucleico de la etapa (b) del método de la invención se introduzca (además) dentro de la célula mediante microinyección.

10 En otra modalidad, la invención se refiere a un método para producir un vertebrado no humano o un mamífero no humano que porta una secuencia diana modificada en su genoma, el método comprende producir una célula no humana de acuerdo con los métodos descritos en la presente descripción. La presente descripción describe, además, la transferencia de una célula que se produce por el método de la invención dentro de un huésped femenino pseudo embarazada.

15 El término "transferencia de una célula que se produce por el método de la invención dentro de un huésped femenino pseudo embarazada" incluye la transferencia de un ovocito fecundado pero, además, la transferencia de embriones de preimplantación de, por ejemplo, la etapa de 2 células, 4 células, 8 células, 16 células y blastocisto (70 a 100 células). Dichos embriones de preimplantación pueden obtenerse mediante el cultivo de la célula en condiciones apropiadas para que se convierta en un embrión de preimplantación. Además, la inyección o fusión de la célula con un blastocisto son
20 métodos apropiados para obtener un embrión de preimplantación. Cuando la célula que se produce por el método de la invención es una célula somática, se requiere la derivación de células madre pluripotentes inducidas antes de transferir la célula a un huésped femenino tal como, por ejemplo, antes de cultivar la célula o inyección o fusión de la célula con un embrión de preimplantación. Los métodos para transferir un ovocito o un embrión preimplantado a una hembra pseudo embarazada se conocen bien en la técnica y se describen, por ejemplo, en Nagy y otros, (Nagy A, Gertsenstein M, Vintersten K, Behringer R., 2003. *Manipulating the Mouse Embryo*. Cold Spring Harbour, Nueva York: Cold Spring Harbour Laboratory Press).
25

Se contempla además, de acuerdo con el método para producir un vertebrado no humano o un mamífero no humano que porta una secuencia diana modificada en su genoma, que se realice una etapa de análisis de la modificación genómica exitosa antes del trasplante en la hembra huésped. Como ejemplo no limitante, el ovocito puede cultivarse hasta la etapa de 2 células, 4 células o 8 células y una célula puede eliminarse sin destruir o alterar el embrión resultante. El análisis de la constitución genómica, por ejemplo, la presencia o ausencia de la modificación genómica, puede realizarse después mediante el uso, por ejemplo, de PCR o de técnicas de transferencia southern o cualquiera de los métodos descritos anteriormente en la presente descripción. Tales métodos de análisis de genotipificación exitosa antes del trasplante se conocen en la técnica y se describen, por ejemplo, en Peippo y otros (Peippo J, Viitala S, Virta J, Raty M, Tammiranta N, Lamminen T, Aro J, Myllymaki H, Vilkkki J.; *Mol Reprod Dev* 2007; 74:1373-1378).
30
35

Cuando la célula es un ovocito, el método para producir un vertebrado no humano o un mamífero no humano que porta una secuencia diana modificada en su genoma comprende (a) modificar la secuencia diana en el genoma de un ovocito de un vertebrado o mamífero de acuerdo con el método de la invención. La presente descripción describe, además, (b) transferir el ovocito que se obtiene en (a) a un huésped hembra pseudo embarazada; y, opcionalmente, (c) analizar la descendencia suministrada por el huésped hembra para detectar la presencia de la modificación.
40

Para el método descrito para producir un vertebrado no humano o un mamífero no humano, se requiere la fecundación del ovocito. Dicha fecundación puede ocurrir antes de la modificación de la secuencia diana en la etapa (a) de acuerdo con el método descrito para producir un vertebrado no humano o un mamífero no humano, es decir, puede usarse un ovocito fecundado para el método de modificación de una secuencia diana. La fecundación puede realizarse, además, después de la modificación de la secuencia diana en la etapa (a), es decir, puede usarse un ovocito no fecundado para el método de modificación de una secuencia diana, en donde el ovocito se fecunda posteriormente antes de transferirlo a la hembra huésped pseudo embarazada.
45
50

La etapa de análisis para detectar la presencia de la modificación en la descendencia suministrada por el huésped hembra proporciona la información necesaria acerca de si el vertebrado no humano o el mamífero no humano producido porta, o no, la secuencia diana modificada en su genoma. Por lo tanto, la presencia de la modificación es indicativa de que dicha descendencia es portadora de una secuencia diana modificada en su genoma, mientras que la ausencia de la modificación es indicativa de que dicha descendencia no es portadora de la secuencia diana modificada en su genoma. Los métodos de análisis para detectar la presencia o ausencia de una modificación se detallaron anteriormente.
55

El vertebrado no humano o el mamífero no humano producido por el método descrito es, entre otras cosas, útil para estudiar la función de los genes de interés y la expresión/resultado fenotípico de las modificaciones del genoma en tales animales. Se contempla, además, que los mamíferos no humanos pueden emplearse como modelos de enfermedades y para evaluar agentes/composiciones terapéuticas. Adicionalmente, el vertebrado no humano o el mamífero no humano puede usarse, además, para la cría de ganado.
60

El método descrito para producir un vertebrado no humano o un mamífero no humano comprende, además, cultivar la célula para formar un embrión de preimplantación o introducir la célula en un blastocisto antes de transferirla al huésped
65

hembra pseudo embarazada. Los métodos para cultivar la célula para formar un embrión de preimplantación o introducir la célula en un blastocisto se conocen bien en la técnica y, por ejemplo, se describen en Nagy y otros, en el lugar citado.

El término "introducir la célula en un blastocisto" como se usa en la presente descripción, abarca la inyección de la célula en un blastocisto así como también la fusión de una célula con un blastocisto. Los métodos para introducir una célula en un blastocisto se describen en la técnica, por ejemplo en Nagy y otros, en el lugar citado.

La presente descripción se refiere, además, a un animal vertebrado no humano o a un mamífero no humano que se obtienen mediante los métodos descritos anteriormente.

En una modalidad preferida de los métodos de la invención, la célula es de un mamífero no humano seleccionado del grupo que consiste en roedores, perros, félidos, primates, conejos, cerdos, o vacas o la célula es de un ave seleccionada del grupo que consiste en pollos, pavos, faisanes, patos, gansos, codornices y aves no voladoras, que incluyen avestruces, emús y casuarios, o la célula es de un pez tal como, por ejemplo, un pez cebra, salmón, trucha, carpa común o carpa koi.

Todos los mamíferos, aves y peces descritos en la presente descripción se conocen bien por los expertos en la técnica y se definen taxonómicamente de acuerdo con la técnica anterior y el conocimiento general común de los expertos en la técnica.

Los ejemplos no limitantes de "roedores" son ratones, ratas, ardillas, ardillas ralladas, ardilla de la tierra, puercoespines, castores, hámster, jerbos, cobayas, degús, chinchillas, perritos de las praderas y marmotas.

Los ejemplos no limitantes de "perros" incluyen los miembros de la subespecie *canis lupus familiaris*, así como también lobos, zorros, chacales, y coyotes.

Los ejemplos no limitantes de "félidos" incluyen miembros de las dos subfamilias: las pantherinas, que incluyen leones, tigres, jaguares y leopardos y las felinas, que incluyen pumas, guepardos, gato cerval, lince, caracales, ocelotes y gatos domésticos.

El término "primates", como se usa en la presente descripción, se refiere a todos los monos, que incluyen, por ejemplo, cercopithecoid (mono del viejo mundo) o platyrrhine (mono del nuevo mundo), así como también lemures, tarseros, simios y monos tití (*Callithrix jacchus*).

Con respecto a las modalidades caracterizadas en esta descripción, en particular en las reivindicaciones, se pretende que cada modalidad mencionada en una reivindicación dependiente se combine con cada modalidad de cada reivindicación (independiente o dependiente) de la que depende dicha reivindicación dependiente. Por ejemplo, en el caso de una reivindicación independiente 1 que enumera 3 alternativas A, B y C, una reivindicación dependiente 2 que enumera 3 alternativas D, E y F y una reivindicación dependiente 3 de conformidad con las reivindicaciones 1 y 2 que enumera 3 alternativas G, H y I, debe entenderse que la descripción describe de manera inequívoca las modalidades correspondientes a las combinaciones A, D, G; A, D, H; A, D, I; A, E, G; A, E, H; A, E, I; A, F, G; A, F, H; A, F, I; B, D, G; B, D, H; B, D, I; B, E, G; B, E, H; B, E, I; B, F, G; B, F, H; B, F, I; C, D, G; C, D, H; C, D, I; C, E, G; C, E, H; C, E, I; C, F, G; C, F, H; C, F, I, a menos que se mencione específicamente de cualquier otra manera.

De manera similar, y, además, en aquellos casos donde las reivindicaciones independientes y/o dependientes no enumeren alternativas, se entiende que si las reivindicaciones dependientes se refieren nuevamente a una pluralidad de reivindicaciones anteriores, se considera que cualquier combinación de la materia cubierta de esta manera se describe explícitamente. Por ejemplo, en el caso de una reivindicación independiente 1, una reivindicación dependiente 2 que se refiere nuevamente a la reivindicación 1 y una reivindicación dependiente 3 que se refiere nuevamente a las reivindicaciones 2 y 1, se deduce que la combinación de la materia de las reivindicaciones 3 y 1 se describe de manera clara e inequívoca, al igual que la combinación de la materia de las reivindicaciones 3, 2 y 1. En caso de que esté presente, además, una reivindicación dependiente 4 que se refiere a cualquiera de las reivindicaciones 1 a 3, se deduce que la combinación de la materia de las reivindicaciones 4 y 1, de las reivindicaciones 4, 2 y 1, de las reivindicaciones 4, 3 y 1, así como también de las reivindicaciones 4, 3, 2 y 1, se describen de manera clara e inequívoca.

Las Figuras muestran:

Figura 1: Vectores de expresión de Nucleasa-TAL.

La figura muestra la estructura y función de las proteínas de fusión Nucleasa-TAL, que consiste en un dominio de unión al ADN en una secuencia específica y un dominio de escisión del ADN (nucleasa) no específico. El dominio de unión al ADN puede ensamblarse a partir de los cuatro tipos de elementos peptídicos de TAL, de 34 aminoácidos, que exhiben especificidad de unión contra uno de los nucleótidos del ADN a través de las posiciones de los aminoácidos 12 y 13 (NI - A; HD - C; NG - T; NN - G). Después de la unión del dominio del elemento TAL a la secuencia del ADN diana seleccionada, el dominio nucleasa de la proteína de fusión entra en contacto estrecho con la doble hebra del ADN, pero no escinde el ADN como un monómero de nucleasa. Solamente después de la unión de una segunda proteína de fusión Nucleasa-TAL

a una segunda secuencia diana de ADN ubicada hacia el extremo 3' con respecto al sitio de unión de la primera proteína de fusión, la doble hebra de ADN se escinde a través de la cooperación de los dos dominios nucleasa que están en contacto estrecho.

5 Figura 2: Modificación de las secuencias genómicas inducida por la Nucleasa-TAL.

La figura muestra un par de proteínas de fusión Nucleasa-TAL que se unen en los extremos 5' y 3' con respecto al sitio diana seleccionado dentro de un gen diana genómico. Después de la creación de una rotura de doble hebra del ADN dentro del sitio diana, dos mecanismos competidores de reparación del ADN se activan fuertemente en las células: i) mediante recombinación homóloga, en presencia de un vector de transformación génica introducido externamente que comprende dos regiones de homología con el gen diana y una modificación/mutación genética prediseñada, la modificación planeada previamente se copia a partir del vector de transformación en el genoma; por esta vía, cualquier modificación génica dirigida (por ejemplo, inactivación génica, reemplazo génico) puede colocarse en el genoma, ii) mediante la vía de reparación por la unión de extremos no homólogos (NHEJ), los extremos de ADN libres se unen por ligación sin un molde de reparación; por esta vía, con frecuencia se pierde un número variable de nucleótidos (símbolo del cuchillo) antes de la ligación final y frecuentemente resulta en un alelo inactivado del gen diana.

Figura 3: Uso de Nucleasas-TAL para la transformación génica en líneas celulares de mamíferos y cigotos.

20 **A:** Para la generación de modificaciones génicas en líneas celulares de mamíferos, los vectores de transformación de Nucleasa-TAL pueden transfectarse junto con, o sin, un vector específico de transformación génica, en células en cultivo. Después de la expresión de la nucleasa y la reparación del ADN, una fracción de las células tratadas contiene la alteración génica deseada. Estas células pueden aislarse y cultivarse adicionalmente como una línea celular modificada genéticamente pura. **B:** Después de la microinyección del ARNm de Nucleasa-TAL, junto con o sin un vector de transformación génica específico, dentro de ovocitos de mamíferos fecundados (cigotos, aislados de hembras de tipo silvestre, por ejemplo, ratones) puede introducirse directamente un alelo inactivado (KO) o un gen de reemplazo (KI) en el genoma del embrión unicelular. Las hembras pseudo embarazadas suministran la descendencia a partir de ovocitos microinyectados. A la descendencia se le realiza una genotipificación para detectar la presencia de la modificación génica inducida. Los animales positivos se seleccionan para su reproducción ulterior con el propósito de establecer una cepa modificada genéticamente.

Figura 4: Vectores de expresión de Nucleasa-TAL.

El vector de expresión Nucleasa-TAL pCAG-Tal-nucleasa contiene una región del promotor CAG y una unidad transcripcional que comprende, el extremo 5' de un par central de sitios de restricción BsmBI, un codón de inicio de la transcripción ATG (flecha), una secuencia de localización nuclear (NLS), una secuencia de etiqueta FLAG (FLAG), una secuencia enlazadora, un segmento codificante para 110 aminoácidos de la proteína TAL AvrBs3 (AvrN) y sus repeticiones invariables Tal en el N-terminal (r0.5). En el extremo 3' de los sitios de restricción BsmBI la unidad transcripcional contiene una repetición Tal invariable en el extremo C-terminal (rx.5), un segmento codificante para 44 aminoácidos derivados de la proteína Tal AvrBs3, unos sitios de restricción PmeI y MluI para la inserción de las regiones codificantes de la nucleasa y una secuencia señal de poliadenilación (pA). Los segmentos de ADN codificantes para los elementos de repetición TAL pueden insertarse en los sitios BsmBI del pCAG-Tal-nucleasa para la expresión de proteínas de fusión Nucleasa-TAL variables. Para crear los vectores de expresión μ ArtTal1-nucleasa el arreglo ArtTal1 de los elementos de repetición TAL, que reconocen la secuencia diana especificada de 12 pb, se insertaron en los sitios BsmBI del pCAG-TAL-nucleasa. Cada repetición Tal de 34 aminoácidos se representa como un cuadrado que indica el código de los aminoácidos de las repeticiones en las posiciones 12/13 que confieren la unión a uno de los nucleótidos del ADN de la secuencia diana (NI > A, NG > T, HD > C, NN > G) mostrada anteriormente. A continuación, las regiones codificantes del dominio nucleasa sintética se insertaron en los sitios PmeI y MluI del pCAG-ArtTal1-nucleasa para obtener los vectores de expresión: **A:**pCAG-ArtTal1-Alw que incluye el dominio nucleasa de la endonucleasa de restricción AlwI, **B:** pCAG-ArtTal1-CleDORF que incluye el dominio nucleasa del gen CleDORF, **C:** pCAG-ArtTal1-Clo051 que incluye el dominio nucleasa del gen Clo051, **D:**pCAG-ArtTal1-Mly que incluye el dominio nucleasa de la endonucleasa de restricción MlyI, **E:** pCAG-ArtTal1-Pept071 que incluye el dominio nucleasa del gen Pept071, **F:** pCAG-ArtTal1-Sbf que incluye el dominio nucleasa de la endonucleasa de restricción SbfI, **G:** pCAG-ArtTal1-Sdal que incluye el dominio nucleasa de la endonucleasa de restricción SdaI **H:** pCAG-ArtTal1-Sst que incluye el dominio nucleasa de la endonucleasa de restricción SstI, e **I:** pCAG-ArtTal1-Fok que incluye el dominio nucleasa de la endonucleasa de restricción FokI

Figura 5: Secuencia de aminoácidos de la proteína Clo051

Secuencia de los 587 aminoácidos de la proteína Clo051 en el código de una letra. Se indican la metionina en la posición 1 (M1), la tirosina en la posición 587 (Y587) y el residuo 199 del dominio nucleasa entre las posiciones E389 e Y587. Además, se destacan las posiciones D455, D472 y K474 que son características del sitio activo conservado de la superfamilia de enzimas 'PD-(D/E)XK' que interaccionan con el ADN.

Figura 6: Estructura predicha de la proteína Clo051 y su dominio nucleasa.

La estructura terciaria de la proteína Clo051 se predijo a partir de su secuencia de aminoácidos (Figura 5) mediante el uso del programa informático I-TASSER. Las estructuras secundarias se muestran como regiones alfa-helicoidales y de hojas beta. Se destacan la metionina en la posición 1 (M1), el residuo de glutamato 389 (E389) y la tirosina 587 (Y587). La cadena de proteínas entre E389 e Y587 forma un dominio de plegamiento separado que actúa como una nucleasa.

Figura 7: Plásmidos reporteros de Nucleasa-TAL-y ensayo reportero de nucleasa.

A: Los plásmidos reporteros Nucleasa-TAL contienen una región del promotor de CMV, una secuencia codificante de 400 pb para el segmento del extremo N-terminal de β -galactosidasa y un codón de parada. Esta unidad es seguida por una región diana de unión a TAL que consiste en dos secuencias de reconocimiento orientadas inversas (subrayadas), separadas por una región espaciadora de 15 pb (NNN..), para el arreglo ArtTal1 (a), el arreglo TalRab1 (b), el arreglo TalRab2 (c), o una región de unión híbrida compuesta por una secuencia de reconocimiento ArtTal1 y una TalRab2 (d). La región diana Nucleasa-TAL es seguida por la región codificante completa para β -galactosidasa y una señal de poliadenilación (pA). Para evaluar la actividad nucleasa contra la secuencia diana un vector de expresión Nucleasa-TAL (Figura 4) se cotransfectó transitoriamente con su plásmido reportero correspondiente en células HEK 293. Después de la expresión de la proteína Nucleasa-TAL el plásmido reportero se abre mediante una rotura de doble hebra inducida por nucleasa dentro de la secuencia diana Nucleasa-TAL (símbolo de tijeras).

B: Las regiones de ADN adyacentes a la rotura de doble hebra son idénticas en 400 pb y pueden alinearse y recombinarse (X) mediante la reparación del ADN por recombinación homóloga. **C:** La recombinación homóloga de un plásmido reportero abierto resulta en un vector de expresión de β -galactosidasa funcional que produce la enzima β -galactosidasa. Después de dos días, las células transfectadas se lisaron y la actividad de la enzima en el lisado se determinó con un ensayo reportero de quimioluminiscencia. Los niveles de la emisión de luz catalizada por el reportero se miden e indican la actividad de la Nucleasa-TAL en comparación con las muestras que se transfectaron con el plásmido reportero solo.

Figura 8: Actividad de las proteínas de fusión Nucleasa-TAL en células HEK 293.

Para evaluar la actividad nucleasa de las proteínas de fusión de los dominio nucleasa-TAL, los vectores de expresión para las proteínas ArtTal1-Alwl, -CleDORF, - Clo051, -Mlyl, -FokI, -Pept071, -Sbfl, -Sdal, y -Stsl (Figura 4) se transfectaron juntos con el plásmido reportero ArtTal1 (Figura 7) en células HEK 293. La actividad nucleasa específica contra la secuencia diana del plásmido reportero conduce a la recombinación homóloga y a la expresión de β -galactosidasa. Dos días después de la transfección, se lisaron las poblaciones celulares y se determinó la actividad de la β -galactosidasa con un ensayo reportero de quimioluminiscencia. Los niveles de emisión de luz se normalizaron en relación con la actividad de un plásmido de expresión de luciferasa cotransfectado (pLuciferasa) y se muestran en comparación con la actividad de un vector de expresión de β -galactosidasa control positivo. La barra para cada muestra transfectada representa el valor medio y la SD derivada de tres pocillos de cultivo transfectados a la par. **A:** La transfección del plásmido reportero ArtTal1 sin vector de expresión de nucleasa resulta en un bajo nivel de fondo de β -galactosidasa. La cotransfección de pCAG-ArtTal1-Alwl, -CleDORF, y -Mlyl con el plásmido reportero ArtTal1 no condujo a un aumento significativo de la expresión del reportero, lo que indica que las proteínas de fusión ArtTal1-Alwl, -CleDORF, y -Mlyl no exhiben actividad nucleasa. Por el contrario, la cotransfección del plásmido reportero ArtTal1 y el plásmido pCAG-ArtTal1-Clo051 resultó en un fuerte aumento de la expresión del reportero, lo que indica que la proteína de fusión ArtTal1-Clo051 exhibe actividad nucleasa específica de la secuencia diana en las células 293. **B:** En un experimento de transfección independiente, la cotransfección de pCAG-ArtTal1-Pept071, -Sbfl, -Sdal y -Sst con el plásmido reportero ArtTal1 no condujo a un aumento significativo de la expresión del reportero, en comparación con el plásmido reportero ArtTal1 solo, lo que indica que las proteínas de fusión ArtTal1-Pept071, -Sbfl, - Sdal, y -Stsl no exhiben actividad nucleasa. Por el contrario, la cotransfección de los plásmidos reportero ArtTal1 y pCAG-ArtTal1-FokI resultó en un aumento de la expresión del reportero, lo que indica la actividad nucleasa de la proteína de fusión ArtTal1-FokI en células 293.

Figura 9: Especificidad de la secuencia diana de la nucleasa ArtTal1-Clo051.

Para evaluar la especificidad de la nucleasa ArtTal1-Clo051 contra la secuencia diana prediseñada en comparación con las secuencias de ADN no relacionadas, el vector de expresión pCAG-ArtTal1-Clo051 se cotransfectó con el plásmido reportero ArtTal1 correspondiente o con los plásmidos reporteros TalRab1 o TalRab2 (Figura 7), que contienen secuencias diana no relacionadas, en células HEK 293. La actividad nucleasa fuerte se desarrolló solamente en la combinación específica del vector de expresión ArtTal1-Clo051 junto con el plásmido reportero ArtTal1, lo que indica que la nucleasa ArtTal1-Clo051 actúa específicamente contra la secuencia diana prediseñada.

Figura 10: Caracterización de la cooperatividad de proteínas de fusión nucleasa TAL-Clo051

A: Para evaluar la cooperatividad de los dominios nucleasa Clo051 de un par de proteínas de fusión TAL-Clo051, los vectores de expresión para las proteínas de fusión ArtTal1-Clo051 o TalRab2-Clo051 se cotransfectaron con el correspondiente plásmido reportero ArtTal1 o TalRab2 (Figura 7) y se comparó con la cotransfección con el plásmido reportero ArtTal1/TalRab2, que contiene una región diana híbrida (Figura 7). La actividad nucleasa significativa se desarrolló solamente en la combinación de vectores de expresión de Nucleasa-TAL con plásmidos reporteros que contienen dos copias idénticas e inversas de la secuencia diana del arreglo TAL correspondiente, pero no con el plásmido reportero ArtTal1/TalRab2 que contiene solamente una secuencia de unión única de las proteínas de fusión ArtTal1-

Clo051 y TalRab2-Clo051. Este resultado indica que dos dominios nucleasa Clo051 deben cooperar para inducir una rotura de doble hebra del ADN, mientras que un único dominio nucleasa Clo051 no actúa como una nucleasa. **B:** La cotransfección del plásmido reportero ArtTal1/TalRab2 con ambos vectores de expresión para ArtTal1-Clo051 y TalRab2-Clo051, pero no con ArtTal1-Clo051 o -Fok solo, resulta en una actividad nucleasa fuerte, en comparación con la transfección del plásmido reportero ArtTal1/TalRab2. Este resultado indica que la actividad nucleasa y la inducción de roturas de doble hebra en la región diana se producen solamente después de la unión de dos proteínas de fusión TAL-Clo051 y de la interacción de un par de dominios nucleasa Clo051.

Figura 11: Diseño de un par de proteínas de fusión TAL-Clo051 de acuerdo con la presente invención, que reconoce el gen Rab38 de ratón.

Nucleasa TAL que reconoce una secuencia diana dentro del exón 1 del gen de ratón *Rab38*. El trinucleótido que representa el codón 19 está subrayado. Se indica cada secuencia de 14 nucleótidos que se reconoce por una de las proteínas de fusión TAL-Clo051 indicadas, RabChtTal1- y RabChtTal2-Clo051. Las dos secuencias diana de 14 pb flanquean una secuencia espaciadora central de 15 pb que se escinde por los dominios nucleasa Clo051.

Figura 12: Estrategia para la modificación del gen de ratón *Rab38* en células ES y cigotos mediante el uso de proteínas de fusión TAL-Clo051.

Dentro del exón 1 del gen de tipo silvestre *Rab38* (Rab38 WT) se indica la posición de los sitios de unión para el par Nucleasa TAL RabChtTal1- y RabChtTal2-Clo051. El vector de transformación Rab38-cht contiene una región de homología 5' de 942 pb y una región de homología 3' de 2788 pb que flanquean los sitios de reconocimiento TAL de Rab38. Dentro del exón 1 el cambio de dos nucleótidos dentro del codón 19 (GTA) de *Rab38* produce una mutación chocolate (cht) sin sentido que codifica para valina (Val) en lugar de la glicina (Gly) de tipo silvestre (WT), y elimina un sitio de restricción BsaI. En cada uno de los sitios adyacentes de reconocimiento TAL de Rab38 se introdujeron numerosas mutaciones silentes para evitar la unión de las proteínas TAL de Rab38 al vector de transformación. La inducción de una rotura de doble hebra dentro del gen de tipo silvestre *Rab38* por el par de proteínas RabChtTal estimula la recombinación homóloga con el vector de transformación Rab38-cht e integra la mutación chocolate sin sentido y las mutaciones silentes en el genoma.

Figura 13: Aislamiento de mutantes hiperactivos de nucleasa Clo051.

La figura muestra la secuencia primaria del dominio nucleasa Clo051 entre las posiciones E389 e Y587. Se indica la distribución de los residuos arginina (R) y lisina (K) (cuadrados sólidos) cargados positivamente y de los residuos glutamato (E) y aspartato (D) (círculos abiertos) cargados negativamente. Los triángulos indican las posiciones S423 y R446. Estos residuos constituyen un marco tridimensional de cargas dentro del dominio Clo051 que determina la estructura terciaria única de esta nucleasa, como se modeló en la estructura de la Figura 6. Determinados reemplazos de residuos polares versus no polares o de residuos no polares contra residuos polares, por ejemplo en las posiciones 423 y 446, cambian la estructura tridimensional de la cadena proteica y resulta en una actividad nucleasa que trabaja más eficientemente.

Figura 14: Actividad de la nucleasa ArtTal1-Clo051 sobre un reportero genómico en células HEK 293 Las células HEK293 que poseen copias genómicas integradas del constructo reportero pCMV-Rab-Reporter(hygro) se transfectaron con pBluescript o pCAG-ArtTal1-Clo051. La actividad nucleasa específica contra la secuencia diana del reportero conduce a la recombinación homóloga y a la expresión de 3-galactosidasa. Dos días después de la transfección, se fijaron las poblaciones celulares y se determinó la fracción de células que expresaban 3-galactosidasa mediante tinción histoquímica con X-Gal. **A:** Cultivo de células reporteras teñidas con X-Gal después de la transfección con pBluescript. **B:** Cultivo de células reporteras teñidas con X-Gal después de la transfección con el vector de expresión de nucleasa pCAG-ArtTal1-Clo051.

Ejemplo 1: Construcción de vectores de expresión y reporteros para Nucleasa Tal y detección de la actividad nucleasa específica

Construcción de vectores de expresión Nucleasa-TAL

Para la expresión de nucleasa-TAL en células de mamíferos, diseñamos el vector de expresión genérico pCAG-TAL-nucleasa (sec. con núm. de ident.: 3) (Figura 4), que contiene una región del promotor híbrida CAG y una unidad transcripcional que comprende una secuencia codificante para un péptido N-terminal de 176 aminoácidos (sec. con núm. de ident.: 4) de las proteínas de fusión Nucleasa TAL, ubicada en el extremo 5' de un par de sitios de restricción BsmBI. Esta región N-terminal incluye un codón de inicio de la transcripción ATG, una secuencia de localización nuclear, una secuencia de etiqueta FLAG, una secuencia enlazadora rica en glicina, un segmento codificante para 110 aminoácidos de la proteína Tal AvrBs3 y la repetición Tal invariable del extremo N-terminal del efector TAL Hax3. En el extremo 3' de los sitios centrales BsmBI, la unidad transcripcional contiene 78 codones (**sec. con núm. de ident.: 5**) que incluyen una repetición TAL invariable del extremo C-terminal (34 aminoácidos) y 44 residuos derivados de la proteína TAL AvrBs3, seguido por un sitio de restricción PmeI y MluI para la inserción de una región codificante de nucleasa y por una secuencia señal de poliadenilación (pA). Los segmentos de ADN codificantes para los arreglos de repeticiones TAL, diseñados para

unir una secuencia diana de nucleasa TAL pueden insertarse en los sitios BsmBI de pCAG-Tal-nucleasa en el marco de lectura con las regiones codificantes de los extremos 5' y 3' para la expresión de proteínas nucleasa TAL prediseñadas.

Para generar los vectores TAL-nucleasa para la expresión en células de mamíferos insertamos un segmento de ADN sintético con la región codificante de un arreglo de 12 repeticiones Tal, designado ArtTal1 (**sec. con núm. de ident.: 6**), en los sitios BsmBI de pCAG-TAL-nucleasa, para derivar el plásmido pCAG-ArtTal1-nucleasa (**sec. con núm. de ident.: 7**). El elemento TAL en el arreglo ArtTal1 reconoce la secuencia diana de ADN artificial 5'-ATTCTGGGACGT-3' (Figura 4). En otro ejemplo insertamos un segmento de ADN sintético con la región codificante de un arreglo de 14 repeticiones Tal, designado TalRab2 (**sec. con núm. de ident.: 8**), en los sitios BsmBI de pCAG-TAL-nucleasa, para derivar el plásmido pCAG-TalRab2-nucleasa (**sec. con núm. de ident.: 9**). El elemento TAL en el arreglo TalRab2 reconoce la secuencia diana 5'-GGTGGCCCCGGTAGT-3' (Figura 7) que ocurre dentro del gen de ratón *Rab38*. Las secuencias diana TAL se seleccionaron de manera que las regiones de unión de las proteínas TAL fueran precedidas por un nucleótido T. Siguiendo la secuencia hacia el extremo 3' de la T inicial en dirección 5'>3', los dominios de unión al ADN específicos de TAL se combinaron juntos en arreglos de 12 (ArtTal1) (Figura 4), o 14 (TalRab2) elementos TAL. Cada motivo de elemento TAL consiste en 34 aminoácidos, la posición 12 y 13 de los cuales determina la especificidad hacia el reconocimiento de A, G, C o T dentro de la secuencia diana. Para derivar los dominios de unión al ADN del elemento TAL usamos el motivo efector TAL (repetición) #11 de la proteína Hax3 de *Xanthomonas* (Núm. de acceso del GenBank AY993938.1 (LTPEQVVAIASNIGGKQALETVQRLLPVLCQAHG) con los aminoácidos N12 e I13 para reconocer A, el motivo efector TAL (repetición) #5 (LTPQQVVAIASHDGGKQALETVQRLLPVLCQAHG) derivado de la proteína Hax3 con los aminoácidos H12 y D13 para reconocer C, y el motivo efector TAL (repetición) #4 (LTPQQVVAIASNIGGKQALETVQRLLPVLCQAHG) de la proteína Hax4 de *Xanthomonas* (Núm. de acceso del GenBank: AY993939.1) con los aminoácidos N12 y G13 para reconocer T. Para reconocer un nucleótido diana G usamos el motivo efector TAL (repetición) #4 de la proteína Hax4 con el reemplazo de los aminoácidos 12 por N y 13 por N (LTPQQVVAIASNIGGKQALETVQRLLPVLCQAHG).

A continuación, construimos proteínas de fusión del dominio de unión al ADN ArtTal1 con dominios de proteínas derivado de nucleasas conocidas o putativas y evaluamos si esas proteínas de fusión nucleasa-TAL son capaces de inducir una rotura en la doble hebra después de unirse al ADN mediante la región de reconocimiento TAL. Para este propósito, insertamos segmentos de ADN sintético que comprenden las regiones codificantes de ocho dominios de nucleasa putativos y el dominio conocido de nucleasa de FokI (**sec. con núm. de ident.: 10**), en los sitios PmeI y MluI del plásmido pCAG-ArtTal1-nucleasa. Entre los ocho dominios putativos de nucleasa seleccionamos dominios de las cinco enzimas de restricción conocidas AlwI (**sec. con núm. de ident.: 11**), MlyI (**sec. con núm. de ident.: 12**), SbfI (**sec. con núm. de ident.: 13**), SdaI (**sec. con núm. de ident.: 14**) y StsI (**sec. con núm. de ident.: 15**). Además, seleccionamos dominios putativos de nucleasa de tres genes microbianos hipotéticos, aún no caracterizados, designados en la presente como 'CleDORF' (**sec. con núm. de ident.: 16**) (Secuencia de Referencia del NCBI: ZP_02080987.1, derivado a partir del genoma de *Clostridium leptum* DSM753), 'Clo051' (**sec. con núm. de ident.: 17**) (Secuencia de Referencia del NCBI: ZP_05132802.1, derivado a partir del genoma de *Clostridium spec. 7_2_43FAA*) y 'Pept071' (**sec. con núm. de ident.: 18**) (Secuencia de Referencia del NCBI: ZP_07399918.1, derivado a partir del genoma de *Peptoniphilus duerdenii* ATCC BAA-1640). Estas proteínas se seleccionaron por los elementos de secuencia característicos que son compatibles con el sitio activo conservado de la superfamilia de enzimas 'PD-(D/E)XK' (Kosinski, J., y otros (2005). BMC Bioinformatics, 6,172) que interaccionan con el ADN (ver Figura 6 para la proteína Clo051). En particular, la proteína Clo051 de 587 residuos puede clasificarse como un miembro de la familia de proteína PD-(D/E)XK por la localización de los pares de aminoácidos P454/ D455 (motivo PD) y D472/ K474 (motivo DXK) (Figura 5). Para dilucidar si la proteína Clo051 contiene un dominio nucleasa separado, realizamos una predicción estructural tridimensional a partir de su secuencia de aminoácidos primaria mediante el uso del programa informático I-TASSER (Roy, A. y otros (2010). Nat Protoc., 5(4):725-38). Como se muestra en la Figura 6 la proteína Clo051 está compuesta por dos dominios de proteína. El dominio C-terminal de Clo051, que comienza aproximadamente con el residuo E389, contiene el motivo consenso de la familia PD-(D/E)XK y aparece como un dominio nucleasa no específico.

Para la expresión de estos dominios de proteínas en células de mamíferos usamos regiones codificantes sintéticas optimizadas de acuerdo con el uso de codones en mamíferos e insertamos segmentos que comprenden los dominios nucleasa putativos de AlwI (**sec. con núm. de ident.: 19**), CleDORF (**sec. con núm. de ident.: 20**), Clo051 (**sec. con núm. de ident.: 1**), MlyI (**sec. con núm. de ident.: 21**), Pept071 (**sec. con núm. de ident.: 22**), SbfI (**sec. con núm. de ident.: 23**), SdaI (**sec. con núm. de ident.: 24**), StsI (**sec. con núm. de ident.: 25**) y el dominio nucleasa conocido de FokI (**sec. con núm. de ident.: 26**) dentro de los sitios PmeI y MluI del plásmido pCAG-ArtTal1-nucleasa, para derivar los vectores de expresión pCAG-ArtTal1-AlwI (**sec. con núm. de ident.: 27**) (Figura 4A), pCAG-ArtTal1-CleDORF (**sec. con núm. de ident.: 28**) (Figura 4B), pCAG-ArtTal1-Clo051 (**sec. con núm. de ident.: 29**) (Figura 4C), pCAG-ArtTal1-MlyI (**sec. con núm. de ident.: 30**) (Figura 4D), pCAG-ArtTal1-Pept071 (**sec. con núm. de ident.: 31**) (Figura 4E), pCAG-ArtTal1-SbfI (**sec. con núm. de ident.: 32**) (Figura 4F), pCAG-ArtTal1-SdaI (**sec. con núm. de ident.: 33**) (Figura 4G), pCAG-ArtTal1-StsI (**sec. con núm. de ident.: 34**) (Figura 4H), y pCAG-ArtTal1-FokI (**sec. con núm. de ident.: 35**) (Figura 4I). Estos vectores de expresión codifican para las proteínas de fusión-TAL designadas como ArtTal1-AlwI (**sec. con núm. de ident.: 36**), ArtTal1-CleDORF (**sec. con núm. de ident.: 37**), ArtTal1-Clo051 (**sec. con núm. de ident.: 38**), ArtTal1-MlyI (**sec. con núm. de ident.: 39**), ArtTal1-Pept071 (**sec. con núm. de ident.: 40**), ArtTal1-SbfI (**sec. con núm. de ident.: 41**), ArtTal1-SdaI (**sec. con núm. de ident.: 42**), ArtTal1-StsI (**sec. con núm. de ident.: 43**), y ArtTal1-FokI (**sec. con núm. de ident.: 44**).

Construcción de plásmidos reporteros de nucleasa TAL

Para determinar la actividad y la especificidad del dominio nucleasa TAL de las proteínas de fusión en células de mamíferos, construimos plásmidos reporteros de nucleasa TAL que contienen dos copias de una secuencia diana de ADN TAL en orientación inversa, separadas por una región espaciadora de 15 nucleótidos (**Figuras 7a-d**). Esta configuración permite medir la actividad de un solo tipo de nucleasa TAL que interacciona como un homodímero de dos moléculas de proteínas que se unen al par inverso de las secuencias diana del plásmido reportero. Después de la unión al ADN y la interacción de dos dominios nucleasas, el ADN del plásmido reportero se escinde dentro de la región espaciadora de 15 pb y exhibe una rotura de doble hebra.

Los plásmidos reporteros de nucleasa TAL contienen una región promotora de CMV, una secuencia de 400 pb que codifica el segmento N-terminal de β -galactosidasa y un codón de parada. Esta unidad es seguida por la región diana de nucleasa TAL (que consiste en dos secuencias de reconocimiento de orientación inversa separadas por una región espaciadora de 15 pb) para las proteínas de fusión ArtTal1 en el plásmido reportero ArtTal1 (**sec. con núm. de ident.: 45**) (**Figura 7 a**), por la secuencia diana no relacionada TalRab1 en el plásmido reportero TalRab1 (**sec. con núm. de ident.: 46**) (**Figura 7 b**), por la región diana para las proteínas de fusión TalRab2 en el plásmido reportero TalRab2 (**sec. con núm. de ident.: 47**) (**Figura 8 c**), o una región diana híbrida que contiene una copia de las secuencias de reconocimiento ArtTal1 y TalRab2 en el plásmido reportero ArtTal1/TalRab2 (**sec. con núm. de ident.: 48**) (**Figura 8 d**).

Dentro de estos plásmidos reporteros, las regiones diana de nucleasa TAL están seguidas por la región codificante completa de la β -galactosidasa y una señal de poliadenilación (pA). Para evaluar la actividad nucleasa contra la secuencia diana específica, un vector de expresión de nucleasa (**Figura 4**) se cotransfectó transitoriamente con su correspondiente plásmido reportero en células de mamíferos. Después de la expresión de la proteína nucleasa TAL, el plásmido reportero se abre mediante una rotura de doble hebra inducida por la nucleasa dentro de la secuencia diana de la nucleasa TAL (**Figura 7 A**). Las regiones de ADN adyacentes a la rotura de doble hebra son idénticas en 400 pb y pueden alinearse y recombinarse mediante la recombinación homóloga del proceso de reparación del ADN (**Figura 7 B**). La recombinación homóloga de un plásmido reportero abierto resultará posteriormente en una región codificante de β -galactosidasa funcional transcrita a partir del promotor de CMV que conduce a la producción de la proteína β -galactosidasa (**Figura 7 C**). En los lisados de las células transfectadas, puede determinarse la actividad enzimática de la β -galactosidasa mediante quimioluminiscencia y referir la actividad nucleasa de las proteínas de fusión TAL.

Medición de la actividad y especificidad de la Nucleasa-TAL en células 293 humanas

Para determinar la actividad y especificidad de las nucleasas TAL en células de mamíferos, electroporamos un millón de células HEK 293 (ATCC #CRL-1573) (Graham FL, Smiley J, Russell WC, Nairn R., J. Gen. Virol. 36, 59-74, 1977) con 5 μ g de ADN plasmídico de uno de los vectores de expresión de Nucleasa TAL (**Figura 4**) junto con 5 μ g de uno de los plásmidos reporteros de Nucleasa TAL (**Figura 7**). Además, cada muestra recibió 5 μ g del plásmido de expresión de luciferasa de luciérnaga pCMV-hLuc (**sec. con núm. de ident.: 49**) y se ajustó a una cantidad total de ADN de 20 μ g con el ADN plasmídico pBluescript (pBS) (**sec. con núm. de ident.: 50**). Después de la transfección, las células se sembraron por triplicado en pocillos de una placa de cultivo de tejidos de 6 pocillos y se cultivaron durante dos días antes de comenzar el análisis. Para el análisis, las células transfectadas de cada pocillo se lisaron y se determinaron individualmente las actividades de las enzimas β -galactosidasa y luciferasa de los lisados mediante el uso de los ensayos reporteros quimioluminiscentes de acuerdo con las instrucciones del fabricante (Roche Applied Science, Alemania) en un luminómetro (Berthold Centro LB 960). Como control positivo transfectamos 5 μ g del plásmido de expresión de β -galactosidasa pCMV β (**sec. con núm. de ident.: 51**) con 15 μ g de pBS, como control negativo se transfectaron 5 μ g de pCMV-hLuc con 15 μ g de pBS o 5 μ g de pCMV-hLuc junto con 5 μ g de un plásmido reportero de nucleasa TAL y 10 μ g de pBS. Los valores de β -galactosidasa por triplicado de cada muestra se normalizaron en relación con los niveles de actividad de luciferasa y el valor medio y la desviación estándar de la actividad de β -galactosidasa se calcularon y expresaron en comparación con el control positivo pCMV β . En este tipo de ensayo de recombinación, el nivel de emisión de luz catalizada por la β -galactosidasa refleja la escisión y reparación de los plásmidos reporteros y, de esta manera, indica la actividad de las nucleasas TAL.

Como se muestra en la **Figura 8** la transfección del plásmido reportero ArtTal1 solo resultó en niveles de fondo de β -galactosidasa. La cotransfección del plásmido reportero ArtTal1 con los vectores de expresión pCAG-ArtTal1-AlwI, -CleDORF, -MylI, -Pept071, -SbfI, -SdaI, y -StsI no revelaron ninguna actividad nucleasa significativa de las proteínas de fusión que codifican TAL (**Figura 8**), lo que indica que los dominios nucleasas seleccionados son incapaces de operar en combinación con los elementos TAL de unión al ADN. En contraste, la cotransfección del plásmido reportero ArtTal1 con los vectores de expresión pCAG-ArtTal1-Clo051 (**Figura 8A**) y -FokI (**Figura 8B**) resultó en un aumento significativo de la actividad reportera, lo que indica que los dominios de proteínas seleccionados FokI y Clo051 son capaces de funcionar como nucleasas en fusión con los elementos TAL de unión al ADN.

Dado que en ensayos repetidos, las fusiones de TAL con el dominio Clo051 parecía más activo en comparación a las fusiones con el dominio nucleasa de FokI, creemos que el dominio Clo051 es el más adecuado para la construcción de Nucleasas TAL altamente activas.

Para definir si la nucleasa ArtTal1-Clo051 reconoce específicamente su secuencia diana dentro del plásmido reportero ArtTal1 (**Figura 7a**), el pCAG-ArtTal1-Clo051 se cotransfectó con el correspondiente plásmido reportero ArtTal1 o con los

plásmidos reporteros no relacionados TalRab1 o TalRab2 (**Figura 7b,c**) en células HEK 293. Como se muestra en la Figura 9 se detectó actividad reportera aumentada significativamente solamente a partir de la combinación específica de la nucleasa ArtTal-Clo051 con su correspondiente promotor, mientras que la cotransfección con plásmidos reporteros no relacionados no exhibe actividad nucleasa significativa. Estos resultados indican que el dominio nucleasa Clo051 en fusión con los elementos TAL de unión al ADN actúa en una secuencia diana de manera específica y que las secuencias diana no relacionadas no se procesan.

A continuación, caracterizamos si el dominio nucleasa Clo051, como un monómero, induce roturas de doble hebra recombinogénicas o si es necesario la interacción de dos dominios nucleasa como un dímero. Para este propósito, construimos el plásmido reportero híbrido ArtTal1/TalRab2-reportero (sec. con núm. de ident.: 48) (Figura 7d) que contiene una secuencia de reconocimiento ArtTal1 en el extremo 5' de la región espaciadora y una secuencia de reconocimiento TalRab2 en el extremo 3' de la región espaciadora. El arreglo TalRab2 (sec. con núm. de ident.: 8) está compuesto por 14 elementos TAL que reconocen la secuencia diana 5'-GGTGGCCCGGTAGT-3'. El dominio nucleasa Clo051 se clonó como una región codificante sintética en los sitios Pmel y MluI del plásmido pCAG-TalRab2-nucleasa (sec. con núm. de ident.: 9) para derivar el vector de expresión pCAG-TalRab2-Clo051 (sec. con núm. de ident.: 52) para la expresión de la proteína TalRab2-Clo051 (sec. con núm. de ident.: 53). Como se muestra en la Figura 10A la cotransfección de pCAG-ArtTal1-Clo051 junto con el plásmido reportero ArtTal1 resultó en una expresión génica significativa del reportero, lo que indica la actividad nucleasa específica de la proteína de fusión ArtTal1-Clo051. Dado que los plásmidos reporteros ArtTal1 contienen dos secuencias de unión ArtTal1 inversas, la actividad nucleasa de ArtTal1-Clo051 puede resultar a partir de la acción de una proteína de fusión única o de la acción combinada de dos moléculas. Para distinguir entre estas posibilidades el pCAGArtTal1-Clo051 se cotransfectó con el plásmido reportero ArtTal1/TalRab2 que contiene solamente una secuencia de unión ArtTal1. Como se muestra en la Figura 10A la nucleasa ArtTal1-Clo051 no exhibe actividad nucleasa significativa sobre el reportero ArtTal1/TalRab2, lo que indica que para inducir una rotura en la doble hebra de ADN los dos dominios nucleasa Clo051 deben interactuar como un dímero. Estos resultados se confirmaron con la nucleasa TalRab2-Clo051 que actúa sobre su correspondiente plásmido reportero TalRab2 pero no sobre el plásmido reportero híbrido ArtTal1/TalRab2 (Figura 10A). Como se esperaba, la proteína de fusión ArtTal1-FokI igualmente no exhibe actividad nucleasa sobre el reportero ArtTal1/TalRab2 (Figura 10B).

A continuación, estudiamos si dos dominios nucleasa Clo051, que se fusionan a diferentes arreglos de los elementos TAL de unión al ADN, son capaces, además, de interactuar e inducir roturas en la doble hebra. Para este propósito, los vectores de expresión pCAG-ArtTal1-Clo051 y pCAG-TalRab2-Clo051 se cotransfectaron junto con el plásmido reportero ArtTal1/TalRab2 y los resultados se compararon con la cotransfección del pCAG-ArtTal1-Clo051 junto con el reportero ArtTal1/TalRab2. Como se muestra en la Figura 10B, la actividad nucleasa significativa sobre el reportero ArtTal1/TalRab2 se desarrolló solamente mediante la coexpresión de las nucleasas ArtTal1- y TalRab2-Clo051, lo que indica que los dominios nucleasa Clo051 fusionados con diferentes arreglos de TAL son capaces de interactuar e inducir una rotura en la doble hebra del ADN dentro de una región diana híbrida que contiene las secuencias de reconocimiento de dos arreglos TAL de unión al ADN distinguidos.

Ejemplo 2: Transformación del gen de ratón *Rab38* en células ES y cigotos con nucleasas TAL-Clo051

Construcción de nucleasas TAL-Clo051 específicas de *Rab38* y un vector de transformación

Para demostrar la funcionalidad de las proteínas de fusión nucleasa-dominio efector TAL de unión al ADN en células de mamíferos diseñamos un par de proteínas de fusión que reconocen una secuencia diana de ADN dentro del gen de ratón *Rab38* (Figura 11). Las dos proteínas de fusión nucleasa-dominio efector TAL de unión al ADN son destinadas a unirse junto a la región del ADN diana bipartita e inducir una rotura en la doble hebra en la región espaciadora para estimular la recombinación homóloga en el locus diana en células de mamíferos.

El gen de ratón *Rab38* codifica la proteína RAB38 que es un miembro de una familia de proteínas conocidas por tener un papel crucial en el transporte vesicular. En ratones mutantes chocolate (cht) el intercambio de un solo nucleótido en la posición 146 (mutación G > T) dentro del primer exón de *Rab38* conduce al reemplazo de glicina por valina en el codón 19 (Loftus, S.K., y otros, Proc Natl Acad Sci U S A, 2002. 99(7): p. 4471-6). Este reemplazo de aminoácido se ubica dentro del dominio de unión a GTP conservado de RAB38 y afecta la clasificación de la proteína relacionada con la tirosinasa 1 (TYRP1) en los melanosomas de los melanocitos *Rab38^{cht}/Rab38^{cht}*. La TYRP1 es una glicoproteína de membrana melanosomal, que funciona como una enzima oxidasa 5,6-Dihidroxiindol-2-ácido carbónico para producir melanina y como un proveedor de estabilidad estructural para la tirosinasa en el complejo enzimático melanogénico. Se cree que TYRP1 transita desde la red trans-Golgi hasta los melanosomas en etapa II por medio de vesículas recubiertas de clatrina. La cantidad reducida de TYRP1 ubicada correctamente conduce a un deterioro de la producción de pigmento y al cambio del color del pelaje, de negro a marrón similar al chocolate, en ratones *Rab38^{cht}/Rab38^{cht}*. Dado que se conoce que las mutaciones de los genes necesarios para la función de los melanocitos causan el albinismo oculocutáneo (OCD), tal como el síndrome de Hermansky-Pudlak en el hombre, el gen *Rab38* es un locus candidato en los pacientes con OCD.

Nos propusimos introducir una fenocopia de la mutación chocolate en el codón 19 de *Rab38* mediante el uso de un par de nucleasas TAL (RabChtTal1- y RabChtTal2-Clo051) que cada una reconoce una secuencia diana de 14 pb ubicada en el extremo 5' y en el extremo 3' de una secuencia espaciadora de 15 pb dentro del exón 1 del gen *Rab38* (**Figura 11**). Para derivar los vectores de expresión para las nucleasas sintéticas RabChtTal1 y RabChtTal2-Clo051, las regiones

codificantes sintéticas para los dominios de unión al ADN RabChtTal1 y RabChtTal2 compuestos de 14 elementos TAL y el dominio nucleasa Clo051 se insertaron en el vector pCAG-TAL-nucleasa. El plásmido resultante pCAG-RabChtTal1-Clo051 (sec. con núm. de ident.: 54) codifica la proteína de fusión RabChtTal1-Clo051 (sec. con núm. de ident.: 55), y el plásmido pCAG-RabChtTal2-Clo051 (sec. con núm. de ident.: 56) codifica la proteína de fusión RabChtTal2-Clo051 (sec. con núm. de ident.: 57).

Para la modificación del gen *Rab38* mediante recombinación homóloga en ovocitos fecundados, construimos el vector de transformación génica pRab38-chtTAL (Figura 12) (sec. con núm. de ident.: 58), que comprende dos regiones de homología que abarcan 942 y 2788 pb de la secuencia genómica que flanquean el exón 1 del gen de ratón *Rab38* (sec. con núm. de ident.: 59). Para este propósito, los vectores con los brazos de homología 5' y 3' se amplificaron a partir del clon genómico BAC, RPCI-421G2 (derivado a partir del genoma de C57BL/6J, Imagenes GmbH, Berlín), mediante el uso de cebadores específicos de PCR. Dentro de la secuencia del codón 19, introdujimos dos cambios de nucleótidos que modifican el codón 19 de la secuencia de tipo silvestre GGT, que codifica para glicina, en GTA, que codifica para valina. Esta nueva mutación chocolate puede distinguirse de la mutación chocolate de origen natural, que exhibe un solo intercambio de nucleótidos dentro del codón 19 (GTT) que codifica para valina (Loftus, S.K., y otros, Proc Natl Acad Sci U S A, 2002. 99(7): p. 4471-6). Ambos alelos mutantes chocolate pueden distinguirse adicionalmente del alelo de tipo silvestre mediante análisis de restricción, ya que las mutaciones en el codón 19 eliminan un sitio de reconocimiento para la endonucleasa de restricción BsaJI (Figura 12). La región de reconocimiento de las nucleasas TAL se ubica en el extremo 3' del codón 19 (Figura 11). Para la construcción de la región de homología 3' del vector de transformación, cada secuencia de reconocimiento de 14 pb de la proteína de fusión TAL se modificó adicionalmente mediante la introducción de cambios de nucleótidos silentes que no alteran la secuencia de la proteína RAB38 (Figura 12), para evitar el procesamiento potencial del vector de transformación por las nucleasas TAL específicas de *Rab38*.

Para la modificación del gen *Rab38* mediante recombinación homóloga en células ES de ratón, modificamos el vector de transformación génica pRab38-chtTAL (Figura 12) mediante la inserción de un gen de resistencia a neomicina, como un marcador de selección, en la región espaciadora de la región de reconocimiento de la nucleasa-TAL, para derivar el vector de transformación pRab38-chtTAL-neo (sec. con núm. de ident.: 60).

Transformación del gen *Rab38* en células ES y cigotos

Para demostrar la utilidad de las proteínas RabChtTal1- y RabChtTal2-Clo051 para la transformación génica en células de mamíferos (Figura 3) introdujimos los vectores de expresión o el ARNm codificante de las proteínas junto con el vector de transformación pRab38-chtTAL-neo en células ES de ratón o con el vector pRab38-chtTAL en ovocitos fecundados de ratón.

Para la transformación en células ES transfectamos las células ES IDG3.2 (Hitz, C. y otros Nucleic Acids Res. 35, e90, 2007) con el vector de transformación linealizado pRab38-chtTAL-neo junto con, o sin, los plásmidos de expresión de nucleasa TAL pCAG-RabChtTal1- y pCAG-RabChtTal2-Clo051. La transfección, selección, expansión y genotipificación de los clones de células ES resistentes a neomicina se realizó de acuerdo con los procedimientos estándar de detección de genes diana como se describió (Nagy A, Gertsenstein M, Vintersten K, Behringer R., 2003. Manipulating the Mouse Embryo. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press). El análisis de los clones de células ES resistentes reveló que la expresión de las nucleasas TAL conduce a una tasa significativamente aumentada de recombinación homóloga en el gen *Rab38* en células ES. Para la microinyección en ovocitos de ratón fecundados, el vector de ADN circular pRab38-chtTAL se mezcló *in vitro* con el transcrito de ARNm codificante de las proteínas RabChtTal1- y RabChtTal2-Clo051 en el tampón de inyección como se describió (Meyer, M., y otros, Proc Natl Acad Sci U S A. 107(34): p. 15022-6). El ARNm de TAL-nucleasa se prepara a partir de los plásmidos de expresión linealizados pCAG-RabChtTal1- y pCAG-RabChtTal2-Clo051.

Mediante la transcripción *in vitro* a partir del promotor T7 mediante el uso del kit mMessage mMachine (Ambion) de acuerdo con las instrucciones del fabricante. El ARNm se modificó adicionalmente mediante la adición de una cola de poli-A mediante el uso del kit de adición de Poli(A) y se purificó con las columnas MegaClear de Ambion. Finalmente, el ARNm se precipitó y disolvió en tampón de inyección.

Para aislar los ovocitos fecundados, los machos de la cepa C57BL/6 se aparean con hembras superovuladas de la cepa FVB. Para inducir la superovulación las hembras FVB de tres semanas de edad, 2 días antes del apareamiento, se trataron con suero de yeguas preñadas 2.5 IU (PMS), y con gonadotropina coriónica humana 2.5 IU (hCG) el día del apareamiento. Los ovocitos fecundados se aislaron de los oviductos de las hembras con tapón positivo y se microinyectaron en medio M2 (Sigma-Aldrich Inc Núm. Cat. M7167) con el ARNm de nucleasa TAL y la preparación del vector de transformación pRab38-chtTAL, en uno de los pronúcleos y el citoplasma de acuerdo con los procedimientos estándar (Nagy A, Gertsenstein M, Vintersten K, Behringer R., 2003. Manipulating the Mouse Embryo. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press).

Después de la microinyección, los ARNm de nucleasa TAL se traducen en proteínas que inducen una rotura de doble hebra en uno o en ambos alelos de *Rab38* en una o más células del embrión en desarrollo. Este evento estimula la recombinación del vector de transformación pRab38-chtTAL con un alelo *Rab38* a través de las regiones de homología presentes en el vector y conduce a la inserción en un sitio específico del codón mutante 19 en el genoma, lo que resulta

- en un alelo *Rab38^{cht}* portador de la mutación chocolate (Figura 12). Los cigotos microinyectados se transfirieron a hembras pseudopreñadas para permitir su desarrollo ulterior en ratones vivos (Nagy A, Gertsenstein M, Vintersten K, Behringer R., 2003. *Manipulating the Mouse Embryo*. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press). A partir de la descendencia resultante, se extrajo ADN genómico de las puntas de la cola para analizar la presencia del evento de recombinación homóloga deseado en el locus *Rab38* mediante PCR. Este análisis se realizó mediante la amplificación por PCR de la región genómica que abarca el exón 1. La presencia de un alelo *Rab38^{cht}* puede reconocerse después de la digestión de los productos del PCR con BsaJI, ya que la mutación *Rab38^{cht}* en el codón 19 conduce a la eliminación de un sitio de restricción BsaJI que está presente en la secuencia de tipo silvestre.
- En un experimento de este tipo, se analizaron ratones derivados de cigotos microinyectados mediante un ensayo de PCR de *Rab38*. Entre este grupo, la mayoría de los ratones exhibieron los dos alelos del genotipo normal *Rab38* de tipo silvestre, mientras que algunos individuos albergaban un alelo de la mutación chocolate *Rab38* prediseñada, como lo indica la ausencia del sitio de restricción BsaJI en el exón 1.
- En conjunto, fue posible introducir una modificación prediseñada en la región codificante del gen *Rab38* mediante recombinación homóloga asistida por la nucleasa TAL-Clo051 en células ES de ratón y ovocitos fecundados.

Ejemplo 3: Aislamiento de mutantes hiperactivos de nucleasa Clo051.

- Como se muestra en la Figura 13 la secuencia primaria del dominio nucleasa Clo051 entre las posiciones E389 e Y587 exhibe una distribución única de residuos de arginina (R) y lisina (K) cargados positivamente y de residuos de glutamato (E) y aspartato (D) cargados negativamente. Estos residuos componen un panorama tridimensional de cargas dentro del dominio Clo051 que determina la estructura terciaria única de esta nucleasa, como se muestra en el modelo estructural en la Figura 6. Determinados reemplazos de residuos polares versus no polares o de residuos no polares contra residuos polares, por ejemplo en las posiciones 423 y 446, alteran la estructura tridimensional de la cadena proteica y puede resultar en un aumento de la actividad nucleasa.

- Tales reemplazos de aminoácidos pueden realizarse mediante ensayos de prueba y error o pueden seguir hipótesis específicas sobre el impacto estructural y funcional en el dominio nucleasa de Clo051. Alternativamente, un gran número de variantes mutagenizadas aleatoriamente de la región codificante del dominio nucleasa Clo051 pueden ensamblarse en una biblioteca mediante PCR mutagénica. Esta biblioteca de moléculas mutantes puede analizarse para detectar la presencia de variantes de nucleasa hiperactiva mediante un ensayo de selección fenotípica en células de levadura, de mamíferos o en *E. coli* que se acopla a una lectura de nucleasa funcional, por ejemplo, como se describió para la mejora de la recombinasa FLP (Buchholz y otros, *Nat. Biotechnol.* 16, 657-62, 1998).

- Tal selección funcional para variantes mejoradas de nucleasas puede resultar en el reemplazo de, por ejemplo, el residuo 423 de una serina por una prolina y del residuo 446 de una arginina por un glutamato. Tales moléculas variantes pueden demostrar una actividad nucleasa superior en comparación con la forma de tipo silvestre Clo051.

- Ejemplo 4: Recombinación inducida por nucleasa Clo051 de sustratos genómicos en células humanas

- La acción de la nucleasa Clo051 se evaluó adicionalmente en células HEK293 humanas sobre un constructo reportero genómico integrado. Para este propósito, el plásmido reportero ArtTal1 (Figura 7) se modificó mediante la inserción de un gen de resistencia a higromicina en la cadena plasmídica. Además, el marco de lectura de la β -galactosidasa se fusionó con la región codificante del gen de resistencia a la neomicina, lo que resulta en el plásmido reportero pCMV-Rab-Reportero(hygro) (sec. con núm. de ident.: 61). Para generar una línea celular que albergara el constructo reportero en su genoma, se electroporó ADN plasmídico reportero linealizado en células HEK 293 humanas (ATCC #CRL-1573) (Graham FL, Smiley J, Russell WC, Nairn R., *J. Gen. Virol.* 36, 59-74, 1977) y se seleccionaron y aislaron los clones resistentes a higromicina. Uno de los clones resistentes, que no mostró actividad de fondo del gen reportero, 293ArtTal-Rep#2, se escogió para el trabajo posterior.

- A continuación, se transfectaron un millón de células reporteras con 5 μ g de ADN plasmídico del vector de expresión de nucleasa Tal, pCAG-ArtTal1-Clo051 (Figura 4) o con 5 μ g del vector de clonación no relacionado pBluescript como control negativo. Después de la transfección, las células se sembraron en pocillos duplicados de una placa de cultivo de tejidos de 6 pocillos y se cultivaron durante dos días antes de comenzar el análisis. Para el análisis, las células transfectadas de cada pocillo se fijaron durante 10 minutos con formaldehído al 4 % y se incubaron durante 4 horas con la solución de tinción de X-Gal (K3(Fell(CN)6) 5 mM, K4(Fell(CN)6) 5 mM, MgCl₂ 2 mM, X-Gal (5-bromo-cloro-3-indoil- β -D-galactopiranosido) 1 mg/mL. Las células recombinadas que expresan el gen reportero se visualizaron mediante una tinción azul intracelular y se cuantificaron en imágenes fotográficas mediante el uso de la función de recuento de células del programa informático ImageJ (disponible en el sitio web con la dirección <http://imagej.nih.gov/ij>). Como se muestra en la Figura 14A, la transfección con el plásmido de control pBluescript no resultó en células reporteras positivas (> 0,1 %, 0 células positivas de 1076 células contadas). En contraste, la transfección de pCAG-ArtTal-1 resultó en una fracción sustancial de células que recombinaron la construcción reportera y expresaron β -galactosidasa (Figura 14 B). Según se cuantificó a partir de imágenes fotográficas, 42,7 % de las células reporteras (227 células positivas de 531 células contadas) mostraron una recombinación exitosa como lo indica la expresión del gen reportero. En conclusión, este

ES 2 709 216 T3

resultado indica que la proteína nucleasa ArtTal1-Clo051 puede procesar eficientemente una secuencia diana localizada dentro del ADN genómico de mamíferos.

LISTADO DE SECUENCIAS

5 <110> Helmholtz Zentrum München - Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)

<120> Proteína que tiene actividad de nucleasa, proteínas de fusión y usos de estas

10 <130> T2026 PCT
<150> EP 11 00 4635.6

<151> 2011-06-07
<160> 61

15 <170> PatentIn versión 3.5
<210> 1

<211> 199

20 <212> PRT

<213> Clostridium spec. 7_2_43 FAA
<400> 1

25 Glu Gly Ile Lys Ser Asn Ile Ser Leu Leu Lys Asp Glu Leu Arg Gly
 1 5 10 15

30 Gln Ile Ser His Ile Ser His Glu Tyr Leu Ser Leu Ile Asp Leu Ala
 20 25 30

 Phe Asp Ser Lys Gln Asn Arg Leu Phe Glu Met Lys Val Leu Glu Leu
 35 40 45

35 Leu Val Asn Glu Tyr Gly Phe Lys Gly Arg His Leu Gly Gly Ser Arg
 50 55 60

40 Lys Pro Asp Gly Ile Val Tyr Ser Thr Thr Leu Glu Asp Asn Phe Gly
 65 70 75 80

45 Ile Ile Val Asp Thr Lys Ala Tyr Ser Glu Gly Tyr Ser Leu Pro Ile
 85 90 95

 Ser Gln Ala Asp Glu Met Glu Arg Tyr Val Arg Glu Asn Ser Asn Arg
 100 105 110

50 Asp Glu Glu Val Asn Pro Asn Lys Trp Trp Glu Asn Phe Ser Glu Glu
 115 120 125

55 Val Lys Lys Tyr Tyr Phe Val Phe Ile Ser Gly Ser Phe Lys Gly Lys
 130 135 140

60 Phe Glu Glu Gln Leu Arg Arg Leu Ser Met Thr Thr Gly Val Asn Gly
 145 150 155 160

65

ES 2 709 216 T3

Ser Ala Val Asn Val Val Asn Leu Leu Leu Gly Ala Glu Lys Ile Arg
165 170 175

5 Ser Gly Glu Met Thr Ile Glu Glu Leu Glu Arg Ala Met Phe Asn Asn
180 185 190

10 Ser Glu Phe Ile Leu Lys Tyr
195

<210> 2

<211> 597

15 <212> ADN

<213> Clostridium spec. 7_2_43 FAA

<400> 2

20 gaagggcatca aaagcaacat ctccctcctg aaagacgaac tccgggggca gattagccac 60
attagtcacg aatacctctc cctcatcgac ctggctttcg atagcaagca gaacaggctc 120
tttgagatga aagtgtctgga actgctcgtc aatgagtacg ggttcaaggg tcgacacctc 180
25 ggcggatcta ggaaaccaga cggcatcgtg tatagtacca cactggaaga caactttggg 240
atcattgtgg ataccaaggc atactctgag ggttatagtc tgcccatttc acaggccgac 300
gagatggaac ggtacgtgcg cgagaactca aatagagatg aggaagtcaa ccctaacaag 360
30 tgggtgggaga acttctctga ggaagtgaag aaatactact tcgtctttat cagcgggtcc 420
ttcaagggtta aatttgagga acagctcagg agactgagca tgactaccgg cgtgaatggc 480
35 agcggcgtca acgtgggtcaa tctgctcctg ggcgctgaaa agattcggag cggagagatg 540
accatcgaag agctggagag ggcaatgttt aataatagcg agtttatcct gaaatac 597

<210> 3

40 <211> 5866

<212> ADN

<213> Secuencia artificial

<220>

45 <223> pCAG-TAL-nucleasa

<400> 3

ggcgcgccgg attcgacatt gattattgac tagttattaa tagtaatcaa ttacggggtc 60
50 attagttcat agcccatata tggagttccg cgttacataa cttacggtaa atggcccgcc 120
tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg ttcccatagt 180
aacgccaata gggactttcc attgacgtca atgggtggag tatttacggt aaactgccc 240
55 cttggcagta catcaagtgt atcatatgcc aagtacgcc cctattgacg tcaatgacgg 300
taaattggccc gcctggcatt atgcccagta catgacctta tgggactttc ctacttggca 360
60 gtacatctac gtattagtca tcgctattac catggctcag gtgagcccca cgttctgctt 420

65

ES 2 709 216 T3

	cactctcccc atctcccccc cctccccacc cccaattttg tatttattta ttttttaatt	480
	attttgtgca gcgatggggg cggggggggg gggggggcgc gcgccaggcg gggcggggcg	540
5	gggcgagggg cggggcgggg cgaggcggag aggtgcggcg gcagccaatc agagcggcgc	600
	gctccgaaag tttcctttta tggcgaggcg gcggcgggcg cgccctata aaaagcgaag	660
	cgcgcggcgg gcgggagtcg ctgcgcgctg ccttcgcccc gtgccccgct ccgcgcgcgc	720
10	ctcgcgcgcg ccgccccggc tctgactgac cgcgttactc ccacaggaga gcggggggga	780
	cggcccttct cctccgggct gtaattagcg cttggtttaa tgacggcttg tttcttttct	840
	gtggctgctg gaaagccttg aggggctccg ggagggccct ttgtgcgggg gggagcggct	900
15	cggggggtgc gtgcgtgtgt gtgtgcgtgg ggagcgcgcg gtgcggctcc gcgctgcccg	960
	gcggctgtga gcgctgcggg cgcggcgcgg ggctttgtgc gctccgcagt gtgcgcgagg	1020
	ggagcgcggc cgggggcggg gccccgcggg gcgggggggg ctgcgagggg aacaaaggct	1080
20	gcgtgcgggg tgtgtgcgtg ggggggtgag cagggggtgt gggcgcgctg gtcgggctgc	1140
	aacccccctt gcacccccct ccccgagttg ctgagcacgg cccggcttcg ggtgcggggc	1200
	tccgtacggg gcgtggcggg gggctcgcgg tgccgggcgg ggggtggcgg caggtggggg	1260
25	tgccgggcgg ggcggggccg cctcgggcgg ggaggggctc gggggagggg cgcggcggcc	1320
	cccgagcgcg cggcggtgtg cgaggcgcgg cgagccgcag ccattgcctt ttatggtaat	1380
	cgtgcgagag ggcgagggga cttcctttgt cccaaatctg tgcggagccg aaatctggga	1440
30	ggcgccgcgg caccctctct agcgggcggg gggcgaagcg gtgcggcgcc ggcaggaagg	1500
	aaatgggcgg ggagggcctt cgtgcgtcgc cgcgcgcggg tcccccttc cctctccagc	1560
35	ctcggggctg tocgcggggg gacggctgcc ttccgggggg acggggcagg gcggggttcg	1620
	gcttctggcg tgtgaccggc ggctctagag cctctgctaa ccatgttcat gccttcttct	1680
	ttttctaca gatccttaat taataatacg actcactata ggggccgcca ccatgggacc	1740
40	taagaaaaag aggaaggtgg cggccgctga ctacaaggat gacgacgata aaccaggtgg	1800
	cggaggtagt ggcggaggtg gggtagccgc cagtccagca gccaggtgg atctgagaac	1860
	cctcggctac agccagcagc agcaggagaa gatcaaacca aaggtgcggg ccaccgtcgc	1920
45	tcagcaccat gaagcactgg tggggcacgg tttcacacac gcccatattg tggtctctgc	1980
	tcagcatccc gctgcaactg ggactgtggc cgtcaaatat caggacatga tcgccgctct	2040
	gcctgaggca acccacgaag ccattgtggg cgtcggaag cagtggagcg gtgccagagc	2100
50	actcgaagca ctctcaccg tcgccgggga actgcggggg ccaccactcc agtccggact	2160
	ggacactgga cagctgctga agatcgctaa acgcccggga gtgacagctg tggaagctgt	2220
	gcacgcttgg aggaatgctc tgacaggagc cccactgaat cttatgagac gacgtctcac	2280
55	ggcctgaccc cacagcaggg cgtcgctatt gcttctaatt gcggagggcg gcctgctctg	2340
60		
65		

ES 2 709 216 T3

	gagagcattg	tggtcagct	gtccaggccc	gacccctgcc	tggttagatc	cgcactcact	2400
	aacgatcatc	tggtcgtctc	cgcttgccct	ggtggacggc	ccgctctgga	cgcagtcaaa	2460
5	aagggtctcc	cccatgctcc	cgcactgac	aagagaacca	acaggagaat	tcctgaggga	2520
	tcgatcggtt	taaacgatca	cgcgtaaatg	attgcagatc	cactagtctt	agaattccag	2580
	ctgagcgccg	gtcgctacca	ttaccagttg	gtctggtgtc	aaaaataata	ataaccgggc	2640
10	aggggggac	tgcatggatc	tttgtgaagg	aaccttactt	ctgtggtgtg	acataattgg	2700
	acaaactacc	tacagagatt	taaagctcta	aggtaaatat	aaaattttta	agtgtataat	2760
	gtgttaaact	actgattcta	attgtttgtg	tatttttagat	tccaacctat	ggaactgatg	2820
15	aatgggagca	gtggtggaat	gccagatcca	gacatgataa	gatacattga	tgagtttgga	2880
	caaaccacaa	ctagaatgca	gtgaaaaaaa	tgctttatatt	gtgaaatttg	tgatgctatt	2940
	gctttatatt	taaccattat	aagctgcaat	aaacaagtta	acaacaacaa	ttgcattcat	3000
20	tttatgtttc	aggttcaggg	ggaggtgtgg	gaggtttttt	aaagcaagta	aaacctctac	3060
	aatgtggta	tggtcgatta	tgatctcgcg	ccgccactgg	ccgtcgcttt	acaacgtcgt	3120
	gactgggaaa	accctggcgt	tacccaactt	aatcgccctg	cagcacatcc	ccctttcgcc	3180
25	agctggcgta	atagcgaaga	ggcccgccac	gacgcgccct	cccaacagtt	gcgcagcctg	3240
	aatggcgaat	ggaacgcgcc	ctgtagcggc	gcattaagcg	cggcggtgtg	ggtggttacg	3300
	cgcagcgta	ccgctacact	tgccagcgcc	ctagcgcccg	ctcctttcgc	tttcttccct	3360
30	tcctttctcg	ccacgttcgc	cggctttccc	cgtcaagctc	taaatcgggg	gtccctttta	3420
	gggttcgat	ttagtgtttt	acggcacctc	gaccccaaaa	aacttgatta	gggtgatggt	3480
	tcacgtagt	ggccatcgcc	ctgatagacg	gtttttcgcc	ctttgacgtt	ggagtccacg	3540
35	ttctttaata	gtggactctt	gttccaaact	ggaacaacac	tcaaccctat	ctcggcttat	3600
	tcctttgatt	tataagggat	tttgccgatt	tcggcctatt	ggttaaaaaa	tgagctgatt	3660
40	taacaaaaat	ttaacgcgaa	ttttaacaaa	atattaacgc	ttacaattta	ggtggcactt	3720
	ttcggggaaa	tgtgcgcgga	acccctatatt	gtttattttt	ctaaatacat	tcaaatatgt	3780
	atccgctcat	gagacaataa	ccctgataaa	tgcttcaata	atattgaaaa	aggaagagta	3840
45	tgagtattca	acatttcctg	gtcgccctta	ttcccttttt	tgcggcattt	tgccctcctg	3900
	tttttgctca	cccagaaacg	ctggtgaaag	taaaagatgc	tgaagatcag	ttgggtgcac	3960
	gagtgggtta	catcgaactg	gatctcaaca	gcggtaaag	ccttgagagt	tttcgccccg	4020
50	aagaacgttt	tccaatgatg	agcactttta	aagttctgct	atgtggcgcg	gtattatccc	4080
	gtattgacgc	cgggcaagag	caactcggtc	gccgcataca	ctattctcag	aatgacttgg	4140
55	ttgagtactc	accagtcaca	gaaaagcatc	ttacggatgg	catgacagta	agagaattat	4200
60							
65							

ES 2 709 216 T3

	gcagtgtctgc cataacctatg agtgataaca ctgcgggccaa cttactttctg acaacgatcg	4260
	gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg	4320
5	atcgtttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc	4380
	ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt	4440
	cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct	4500
10	cggcccttcc ggctggctgg tttatttctg ataaatctgg agccggtgag cgtgggtctc	4560
	gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca	4620
	cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgtgag ataggtgcct	4680
15	cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt	4740
	taaaaacttca tttttaattht aaaaggatct aggtgaagat cctttttgat aatctcatga	4800
20	ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgtg gaaaagatca	4860
	aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac	4920
	caccgctacc agcgggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg	4980
25	taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag	5040
	gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac	5100
	cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt	5160
30	taccggataa ggcgcagcgg tcgggctgaa cgggggggtc gtgcacacag ccagccttg	5220
	agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcggccacgc	5280
35	ttcccgaagg gagaaaggcg gacaggtatc cggttaagcgg cagggtcgga acaggagagc	5340
	gcacgagggg gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc	5400
	acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa	5460
40	acgccagcaa cggggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt	5520
	tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgtgctg	5580
	ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcag gaagcggaag	5640
45	agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc	5700
	acgacagggt tcccactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc	5760
	tcactcatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa	5820
50	ttgtgagcgg ataacaattht cacacaggaa acagctatga ccatga	5866

<210> 4

<211> 176

55 <212> PRT

<213> Secuencia artificial

<220>

<223> Péptido N-terminal

60

<400> 4

65

ES 2 709 216 T3

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

<210> 5
 <211> 78
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Péptido C-terminal

<400> 5

ES 2 709 216 T3

Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu
65 70 75

5

<210> 6
<211> 408
<212> PRT
<213> Secuencia artificial

10

<220>
<223> ArtTall

15

<400> 6

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
1 5 10 15

20

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
20 25 30

25

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
35 40 45

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
50 55 60

30

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
65 70 75 80

35

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
85 90 95

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
100 105 110

40

Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
115 120 125

45

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
130 135 140

50

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
145 150 155 160

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
165 170 175

55

Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val
180 185 190

60

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
195 200 205

Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu
210 215 220

65

ES 2 709 216 T3

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
 225 230 235 240
 Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala
 245 250 255
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
 260 265 270
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
 275 280 285
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
 290 295 300
 His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser His Asp Gly
 305 310 315 320
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
 325 330 335
 Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
 340 345 350
 Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
 355 360 365
 Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
 370 375 380
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
 385 390 395 400
 |Pro Val Leu Cys Gln Ala His Gly

40 405

<210> 7

<211> 7067

45 <212> ADN

<213> Secuencia artificial

<220>

<223> pCAG-ArtTall-nucleasa

50 <400> 7

gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 60
 catatatgga gttccgcgtt acataactta cggtaaattgg ccgcctggc tgaccgccca 120
 acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga 180
 cttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 240
 aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa tgacggtaaa tggccgcct 300
 ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 360
 tagtcatgcg tattaccatg gtgcaggtga gccccacgtt ctgcttcact ctccccatct 420
 cccccccctc cccacccccca attttgtatt tatttatattt ttaattattt tgtgcagcga 480

ES 2 709 216 T3

	tgggggccccg gggggggggg gggcgccgcg caggcggggc gggcggggc gaggggccccg	540
	gcggggcgag gcggagaggt gcggcgccag ccaatcagag cggcgccgc cgaaggttc	600
5	cttttatggc gaggcggcg cggcgccgc cctataaaaa gcgaagcgcg cggcgggcg	660
	gagtcgctgc gcgctgcctt cggcccgctgc cccgctccgc cggcgccgc cggcgccgc	720
	ccgggctctg actgaccgcg ttactccac aggtgagcg gcgggacgc ccttctctc	780
10	cgggctgtaa ttagcgcttg gtttaatgac ggcttggttc tttctgtg ctgctgaaa	840
	gccttgaggg gctccgggag gggcctttgt gcggggggga gcggctcggg gggcgctgc	900
	gtgtgtgtgt gcgtggggag cggcgccgc ggctccgcg tgcccgggcg ctgtgagcg	960
15	tgccggcgcg gcgggggct ttgtgcgctc cgcagtgtgc gcgagggag cggcgccgg	1020
	ggcggtgccc cggcgctgcg ggggggctgc gaggggaaca aaggctgcgt gcggggtgtg	1080
	tgctggggg ggtgagcag ggggtgtggc gcgtcggtcg ggctgcaacc cccctgcac	1140
20	ccccctccc gagttgctga gcacggcccg gcttcgggtg cggggctccg tacggggcg	1200
	ggcgccgggc tcgccgtgc gggcggggg tgccggcag tgggggtgcc gggcgggcg	1260
	ggcgccctc gggcgggga gggctcggg gagggcgcg gcggccccc gagcgccgc	1320
25	ggctgtcgag gcggcgcg cgcagccat tgccttttat ggtaatcgt cgagagggcg	1380
	cagggaactt ctttgtccca aatctgtgc gagccgaaat ctgggagcg ccgcccacc	1440
	ccctctagc ggcgggggc gaagcggtgc ggcccgcca ggaaggaaat gggcgggag	1500
30	ggccttcgtg cgtcgccgc cggcgctcc cttctccctc tccagcctc gggtgtccg	1560
	cggggggacg gctgccttc ggggggacg ggcaggcg ggctcggtt ctggcgtgtg	1620
35	accggcggt ctagagctc tgctaaccat gttcatgct tcttctttt cctacagatc	1680
	cttaattaat aatacgact actataggg ccgccaccat gggacctag aaaaagagga	1740
	aggtggcggc cgctgactac aaggatgac acgataaac aggtggcgga ggtagtggc	1800
40	gaggtgggt acccgccagt ccagcagccc aggtggatct gagaacctc ggctacagcc	1860
	agcagcagca ggagaagatc aaaccaagg tgccgtccac cgtcgtcag caccatgaag	1920
	cactgggtgg gcacgggttc acacacgccc atattgtggc tctgtctcag catcccgctg	1980
45	cactcgggac tgtggccgtc aaatatcagg acatgatgc cgtctgcct gaggaaccc	2040
	acgaagccat tgtggcgctc ggaagcagt ggagcggtgc cagagcactc gaagcactc	2100
	tcaccgtcgc cggggaactg cggggccac cactccagtc cggactggac actggacagc	2160
50	tgctgaagat cgctaaacgc ggcggagtga cagctgtgga agctgtgcac gcttgagga	2220
	atgctctgac aggagcccca ctgaatctta ctccagaaca ggtcgtcgca atcgcaagta	2280
	acatcgccgcg aaaacaggcc ctcgaaaccg tccagagact cctccccgtg ctgtgccagg	2340
55	cccacggact gacccacag caggtggtc ccatcgctag caacggcgga gggaaagcgg	2400
	ctctggagac cgtgcagagg ctgctccccg tcctgtgcca ggcacatgg ctcacacctc	2460
	agcaggtggt cgcaattgcc tccaatggtg gcggaaaaca ggcctggaa actgtgcaga	2520
60	gactgctccc cgtgctgtgc caggctcacg gtctcacacc ccagcaggtg gtcgctatcg	2580
	catctcatga cgggggcaag caggcactgg agacagtga gcggctgctc cctgtcctgt	2640
	gccaggccca cggactcact cctcagcagg tcgtcgccat tgctagtaac ggcggaggga	2700
65	aacaggctct ggaaccgtg cagcgctgc tccccgtgt gtgccaagcc cagggcctga	2760

ES 2 709 216 T3

	ccccccagca ggtggtcgca atcgctcaa acaatggtgg caagcaggcc ctggagactg	2820
	tgcagcgact gctcccagtg ctgtgccagg cccatggact cacaccacag caggctcgctg	2880
5	ctattgcaag caacaatgga gggaaacagg cactggaaac agtccagagg ctgctccccg	2940
	tgtgtgtcca agcgcatgga ctcactcccc agcaggctcgt cgccatcgct tccaataacg	3000
	gcggcaagca ggccctggag accgtccaga gactgctccc cgtgctgtgc caagctcacg	3060
10	gactcacacc tgagcaggtc gtggcaatcg cctctaacat tggagggaaa caggccctgg	3120
	aaactgtaca gcggctgctc cccgtgctgt gccaaacaca cggactcact ccacagcagg	3180
	tcgtggccat tgcaagtcac gacggaggca agcaggccct ggaaacagtg cagcgccctgc	3240
15	tcctctgtgt gtgccaggct catggtctga ctctcagca ggtggtggcc atcgcttcca	3300
	acaatggagg gaagcaggcc ctggagaccg tacagagact gctccccgtg ctgtccaag	3360
	cgcacggtct gacccctcag caggctcgct caatcgccag caatggcggg ggcaagcagg	3420
20		
	ctctcgaaac cgtccagcgg ctctcccag tctctgtca ggctcacggc ctgacccac	3480
	agcaggctgt cgctattgct tctaattggc gagggcggcc tgctctggag agcattgttg	3540
25	ctcagctgtc caggcccgat cctgccctgg ctgatccgc actcactaac gatcatctgg	3600
	tcgctctcgc ttgcctcggg ggacggcccg ctctggacgc agtcaaaaag ggtctcccc	3660
	atgctccgc actgatcaag agaaccaaca ggagaattcc tgagggatcc gatcgtttaa	3720
30	acgatcacgc gtaaatgatt gcagatccac tagttctaga attccagctg agcgcgggtc	3780
	gctaccatta ccagttggtc tgggtgtcaa aataataata accgggcagg ggggatctgc	3840
	atggatcttt gtgaaggaa cttacttctg tgggtgtgaca taattggaca aactacctac	3900
35	agagatttaa agctctaagg taaatataaa atttttaagt gtataatgtg ttaaactact	3960
	gattctaatt gtttgtgtat tttagattcc aacctatgga actgatgaat gggagcagtg	4020
	gtggaatgcc agatccagac atgataagat acattgatga gtttgacaa accacaacta	4080
40	gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct ttatttgtaa	4140
	ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt atgtttcagg	4200
	ttcaggggga ggtgtgggag gtttttttaa gcaagtaaaa cctctacaaa tgtggtatgg	4260
45	ctgattatga tctgcccgg ccaactggcg tcgttttaca acgtcgtgac tgggaaaacc	4320
	ctggcggtac ccaacttaat cgccttgacg cacatcccc ttccgccagc tggcgtaata	4380
	gcgaagaggc ccgcaccgat cgccttccc aacagttgag cagcctgaat ggcgaatgga	4440
50	acgcgccctg tagcggcgca ttaagcggg cgggtgtggt ggttacgcgc agcgtgaccg	4500
	ctacacttgc cagcgccta gcgccgctc ctttcgcttt ctcccttcc tttctcgcca	4560
	cgttcgcggg ctttcccgt caagctctaa atcgggggct ccttttaggg ttccgattta	4620
55	gtgctttacg gcacctcgac ccaaaaaaac ttgattaggg tgatggttca cgtagtgggc	4680
	catcgccctg atagacgggt ttccgccctt tgacgttggg gtccacgttc tttaatagt	4740
	gactcttgtt ccaaaactgga acaacactca accctatctc ggtctattct tttgatttat	4800
60	aagggttttt gcgattttcg gcctattggt taaaaaatga gctgatttaa caaaaattta	4860

65

ES 2 709 216 T3

	acgcgaattt taacaaaata ttaacgctta caatttaggt ggcacttttc ggggaaatgt	4920
	gcgcggaacc cctatttgtt tatttttcta aatacattca aatatgtatc cgctcatgag	4980
5	acaataaacc tgataaatgc ttcaataata ttgaaaaagg aagagtatga gtattcaaca	5040
	tttcogtgc gccottatcc ccttttttgc ggcattttgc cttcctgttt ttgctcacc	5100
	agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag tgggttacat	5160
10	cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgcctcgaag aacgttttcc	5220
	aatgatgagc acttttaaag ttctgctatg tggcgcggtt ttatcccgta ttgacgcgg	5280
	gcaagagcaa ctcggtgcc gcatacacta ttctcagaat gacttggttg agtactcacc	5340
15	agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgccat	5400
	aacctagagt gataaactcg cgcccaactt actttctgaca acgatcggag gaccgaagga	5460
	gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc gttgggaacc	5520
20	ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg tagcaatggc	5580
	aacaacgctg cgcaactat taactggcga actacttact ctagcttccc ggcaacaatt	5640
	aatagactgg atggaggcgg ataaagtgc aggaccactt ctgcgctcgg cccttcgggc	5700
25	tggctggttt attgctgata aatctggagc cggtagcgtt ggtctcgcg gtatcattgc	5760
	agcactgggg ccagatggtg agccctcccg tatcgtagtt atctacacga cggggagtca	5820
	ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca	5880
30	ttggttaactg tcagaccaag ttactcata tatactttag attgatttaa aacttcattt	5940
	ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca aaatccctta	6000
35	acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg	6060
	agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac cgctaccagc	6120
	ggtggtttgt ttgccgcatc aagagctacc aactctttt ccgaaggtaa ctggcttcag	6180
40	cagagcgagc ataccaaata ctgtccttct agttagcgg tagttaggcc accacttcaa	6240
	gaactctgta gcaccgccta catacctcgc tctgctaate ctgttaccag tggctgctgc	6300
	cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac cggataaggc	6360
45	gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttgagc gaacgacct	6420
	caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc ccgaaggag	6480
	aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca cgaggagct	6540
50	tccaggggga aacgcctggt atctttatag tcctgtcggg ttctgccacc tctgacttga	6600
	gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc	6660
	ggccttttta cggttcctgg ccttttctg gccttttct cecatgttct ttctgcgtt	6720
55	atcccctgat tctgtggata accgtattac cgcctttgag tgagctgata ccgctcggcg	6780
	cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc gcccaatacg	6840
	caaaccgcct ctcccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc	6900
60	cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca ctcattaggc	6960
	accccgagct ttacacttta tgcttcgggc tcgtatgttg tgtggaattg tgagcggata	7020
65	acaatttcac acaggaaaca gctatgacca tgaggcgcg cggtatc	7067

ES 2 709 216 T3

<210> 8
 <211> 476
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> TalRab2

<400> 8

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

ES 2 709 216 T3

	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	225	230	235	240
5	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	245	250	255	
	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	260	265	270	
10	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	275	280	285	
	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	290	295	300	
15	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	305	310	315	320
20	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	325	330	335	
	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	340	345	350	
25	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	355	360	365	
30	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	370	375	380	
	Ser	Asn	Ile	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	385	390	395	400
35	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	405	410	415	
	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	420	425	430	
40	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	435	440	445	
45	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	450	455	460	
	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly					465	470	475	

50
 <210> 9
 <211> 7271
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <223> pCAG-TalRab2-nucleasa
 <400> 9

60

65

ES 2 709 216 T3

	ggcgcgccgg attcgacatt gattattgac tagttattaa tagtaatcaa ttacggggtc	60
	attagttcat agcccatata tggagttccg cgttacataa cttacggtaa atggcccggc	120
5	tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg ttcccatagt	180
	aacgccaata gggactttcc attgacgtca atgggtggag tatttacggt aaactgccc	240
	cttggcagta catcaagtgt atcatatgcc aagtacgcc cctattgacg tcaatgacgg	300
10	taaatggccc goctggcatt atgcccagta catgacctta tgggactttc ctacttgcca	360
	gtacatctac gtattagtca tcgctattac catggctcgag gtgagcccca cgttctgctt	420
	cactctcccc atctcccccc cctccccacc cccaattttg tatttattta ttttttaatt	480
15	attttgtgca gcgatggggg cggggggggg gggggggcgc gcgccaggcg gggcggggcg	540
	gggcgagggg cggggcgggg cgaggcgagg aggtgcggcg gcagccaatc agagcgggcg	600
	gctccgaaag tttcctttta tggcgaggcg gcggcgggcg cgccctata aaaagcgaag	660
20	cgcgcggcgg gcgggagtcg ctgcgcgctg ccttcgcccc gtgccccgct ccgcccgcgc	720
	ctcgcgcgcg ccgccccggc tctgactgac cgcgttactc ccacaggaga gcggcgggga	780
	cggcccttct cctccgggct gtaattagcg cttggtttaa tgacggcttg tttcttttct	840
25	gtggctgctg gaaagccttg aggggctccg ggagggccct ttgtgcgggg gggagcgctt	900
	cggggggtgc gtgctgtgtg gtgtgcgtgg ggagcgccgc gtgcggtccc gcgctgcccg	960
	gcggctgtga gcgctgcggg cgcggcgcgg ggctttgtgc gctccgcagt gtgcgcgagg	1020
30	ggagcgcggc cggggcgggg gcccccggtg gcgggggggg ctgcgagggg aacaaaggct	1080
	gcgtgcgggg tgtgtgcgtg ggggggtgag cagggggtgt gggcgcgctg gtccggctgc	1140
	aacccccctt gcacccccct ccccgagttg ctgagcacgg cccggcttcg ggtgcggggc	1200
35	tccgtacggg gcgtggcgcg gggctcgccg tgccggggcg ggggtggcgg cagggtgggg	1260
	tgccggggcg ggcggggcg cctcggggcg gggagggctc gggggagggg cgcggcgggc	1320
	cccggagcgc cggcggtgtg cgaggcgcg cgagccgag ccattgcctt ttatggtaat	1380
40	cgtgcgagag ggcgcaggga cttcctttgt cccaaatctg tcgggagccg aaatctggga	1440
	ggcgccgcgc caccctctct agcgggcgcg gggcgaagcg gtgcggcgcc ggcagggaag	1500
45	aaatggggcg ggagggcctt cgtgcgtcgc cgcgccgcgc tccccttctc cctctccagc	1560
	ctcggggctg tccgcggggg gacggctgcc ttccgggggg acggggcagg gcgggggttcg	1620
	gcttctggcg tgtgaccggc ggctctagag cctctgctaa ccatgttcat gccttcttct	1680
50	ttttctaca gatccttaat taataatacg actcactata gggcccgcca ccatgggacc	1740
	taagaaaaag aggaaggttg cggccgctga ctacaaggat gacgacgata aaccaggttg	1800
	cggaggtagt ggcggaggtg gggatccgcg cagtccagca gccaggttg atctgagaac	1860
55	cctcggctac agccagcagc agcaggagaa gatcaaacca aaggtgcggt ccaccgtcgc	1920
	tcagcaccat gaagcactgg tggggcacgg ttccacacac gccatattg tggctctgtc	1980
	tcagcatccc gctgcactcg ggactgtggc cgtcaaatat caggacatga tcgccgctct	2040
60	gcctgaggca acccacgaag ccattgtggg cgtcggaaaag cagtggagcg gtgccagagc	2100
	actcgaagca ctccctaccg tcgccgggga actgcggggg ccaccactcc agtcgggact	2160
	ggacactgga cagctgctga agatcgctaa acgcggcgga gtgacagctg tggaagctgt	2220
65	gcacgcttgg aggaatgctc tgacaggagc cccactgaat ctgacacccc agcaggtggt	2280

ES 2 709 216 T3

	ggccattgct agcaacaatg ggggcaagca ggctctggag acagtgcagc gcctgctgcc	2340
	tgtgctgtgc caggctcacg gactgactcc acagcagggt gtggccatcg cttccaacaa	2400
5	tggagggaaa caggctctgg aaacagtgc gaggctgctg cccgtgctgt gccaggetca	2460
	tggactgaca cctcagcagg tcgtcgccat tgcttctaac ggcgaggga agcaggctct	2520
	ggagactgtg cagagactgc tgccagtgt gtgccaggcc catggactga cccctcagca	2580
10	ggtcgtggct atcgctagta acaatggcgg aaaacaggct ctggaaactg tgcagcggct	2640
	gctccccgtg ctgtgccagg cccacggcct cactccacag caggctcgtc ctatcgctc	2700
	taataacggg ggcaagcagg ctctggagac agtacagcgc ctgttaccgc tgctgtgcc	2760
15	ggcacacggc ctcacacctc agcaggctgt ggcaatcgct tcccatgacg gagggaaaca	2820
	ggctctggaa acggtccaga ggctgctccc cgtgctgtgc caagctcacg gcctacccc	2880
	tcagcagggt gtcgctattg cttctcatga tggcggaag caggctctgg agaccgtgca	2940
20	gagactgtc cctgtgctgt gccaaagcca cggcctgact ccacagcagg tcgtggccat	3000
	cgctagtcat gacgggggca aacaggctct ggaaacagta cagcggtgt taccgtgct	3060
	gtgccaagcc catggcctca cacctcagca agtcgtcgt atcgctagca acaatggagg	3120
25	gaagcaggct ctggagacgg tgcagcgct gctcccagtg ctgtgccaag ctcatggct	3180
	cacccctcag caagtcgtc caattgcttc caataacggc ggaaaacagg ctctggaaac	3240
	cgtccagagg ctgctgccc tgctgtgcc agcacatggc ttaactccac agcaagtgg	3300
30	ggccattgct tctaattggg gcggaaagca ggccctggag acagtcacga gactgttgc	3360
	cgtgctgtgc caagcgcat gactgacacc tgaacaggte gtcgctatcg ctagtaatat	3420
	tgggggcaaa caggccctgg aaacagtgc gcggctgct cccgtgctgt gccaggcgca	3480
35	tggactcaca cccagcagg tcgtcgcaat cgcctctaata acggaggga agcaggccct	3540
	ggaaaccgtg cagagactgt tacctgtgct gtgccaggca catggtctga caccacagca	3600
40	ggtggtcgca attgctagca atggcgagg gaagcaggcc ctggagactg tccagagact	3660
	gctaccgtg ctgtgccaag cgcacggcct gacccacag caggctcgtc ctattgcttc	3720
	taatggcgga gggcggcctg ctctggagag cattgtggct cagctgtcca ggcccgatcc	3780
45	tgccttggt agatccgcac tactaaca tcatctggtc gctctcgct gcctcggtgg	3840
	acggcccgct ctggacgcag tcaaaaagg tctccccat gctccgcac tgatcaagag	3900
	aaccaacagg agaattcctg agggatccga tcgtttaaac gatcacggt aatgattgc	3960
50	agatccacta gttctagaat tccagctgag cgcggctgc taccattacc agttggtctg	4020
	gtgtcaaaaa taataataac cgggcagggg ggatctgcat ggatctttgt gaaggaacct	4080
	tacttctgtg gtgtgacata attggacaaa ctacctacag agatttaaag ctctaaggta	4140
55	aatataaaat ttttaagtgt ataatgtgt aaactactga ttctaattgt ttgtgtat	4200
	tagattccaa cctatggaac tgatgaatgg gagcagtggt ggaatgccag atccagacat	4260
	gataagatac attgatgagt ttggacaaac cacaactaga atgcagtga aaaaatgctt	4320
60	tatttgtgaa atttgtgatg ctattgcttt atttgaacc attataagct gcaataaaca	4380
	agttaacaac aacaattgca ttcattttat gtttcagggt cagggggagg tgtgggagg	4440
	tttttaagc aagtaaaacc tctacaaatg tggatggct gattatgatc tgcggccgcc	4500
65		

ES 2 709 216 T3

	actggcgcgc gttttacaac gtcgtgactg ggaaaaccct ggcgttaccc aacttaacgc	4560
	ccttgacgca catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg	4620
5	cccttcccaa cagttgcgca gcctgaatgg cgaatggaac gcgcctgtga ggggcgcatt	4680
	aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgcctagc	4740
	gcccgcctct ttcgctttct tcccttcctt tctcgccacg ttgcgcggct ttccccgtca	4800
10	agctctaaat cgggggctcc ctttaggggt ccgatttagt gctttacggc acctcgaccc	4860
	caaaaaactt gattaggggt atggttcacg tagtgggcca tcgcctgat agacggtttt	4920
	tcgccctttg acgttgaggt ccacgttctt taatagtga ctctgttcc aaactggaac	4980
15	aacactcaac cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc	5040
	ctattggtta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt	5100
	aacgcttaca atttaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta	5160
20	tttttctaaa tacattcaaa tatgtatcgg ctcatgagac aataaccctg ataaatgctt	5220
	caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc	5280
	ttttttgcgg cattttgcct tcctgttttt gctcaccag aaacgctggt gaaagtaaaa	5340
25	gatgctgaag atcagttggg tgacagagtg ggttacatcg aactggatct caacagcgg	5400
	aagatccttg agagttttcg cccgaagaa cgttttccaa tgatgagcac ttttaaagtt	5460
	ctgctatgtg gcgcggtatt atcccgatt gaogccgggc aagagcaact cggtcgcgc	5520
30	atacactatt ctcaaatga cttggttgag tactcaccag tcacagaaaa gcattctacg	5580
	gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg	5640
	gccaaacttac ttctgacaac gatcggagga cgaaggagc taaccgcttt ttgacacaac	5700
35	atgggggac atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca	5760
	aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta	5820
	actggcgaac tacttactct agcttcccg caacaattaa tagactggat ggaggcggat	5880
40	aaagttgcag gaccacttct gcgctcggcc cttccgctg gctggtttat tgctgataaa	5940
	tctggagccg gtgagcgtgg gtctcgcgg atcattgcag cactggggcc agatggttaag	6000
	ccctcccgt tctgagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat	6060
45	agacagatcg ctgagatagg tgccctcactg attaagcatt ggtaactgtc agaccaagtt	6120
	tactcatata tacttttagat tgatttaaaa cttcattttt aattttaaag gatctaggtg	6180
50	aagatccttt ttgataatct catgaccaa atccctaac gtgagttttc gttccactga	6240
	gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta	6300
	atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa	6360
55	gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact	6420
	gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca	6480
	tacctcgtc tgctaatect gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt	6540
60	accgggttg actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg	6600
	ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag	6660
	cgtgagctat gagaaagcgc cacgcttccc gaaggagaa aggcggacag gtatccggta	6720
65	agcggcaggg tcggaacagg agagcgcacg agggagcttc caggggaaa cgcctggtat	6780

ES 2 709 216 T3

```

ctttatagtc ctgtcgggtt tggccacctc tgacttgagc gtcgattttt gtgatgctcg      6840
tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc      6900
5  ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac      6960
cgtattaccg cctttgagtg agctgatacc gctcgcgcga gccgaacgac cgagcgcagc      7020
gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgcgt      7080
10  tggcgcattc attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag      7140
cgcaacgcaa ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg      7200
cttcgcgctc gtatgttgtg tggaattgtg agcggataac aatttcacac aggaacagc      7260
15  tatgaccatg a                                                    7271

```

<210> 10

<211> 583

20 <212> PRT

<213> Flavobacterium okeanoites

<400> 10

```

25  Met Phe Leu Ser Met Val Ser Lys Ile Arg Thr Phe Gly Trp Val Gln
    1             5             10             15

30  Asn Pro Gly Lys Phe Glu Asn Leu Lys Arg Val Val Gln Val Phe Asp
    20             25             30

35  Arg Asn Ser Lys Val His Asn Glu Val Lys Asn Ile Lys Ile Pro Thr
    35             40             45

40  Leu Val Lys Glu Ser Lys Ile Gln Lys Glu Leu Val Ala Ile Met Asn
    50             55             60

45  Gln His Asp Leu Ile Tyr Thr Tyr Lys Glu Leu Val Gly Thr Gly Thr
    65             70             75             80

50  Ser Ile Arg Ser Glu Ala Pro Cys Asp Ala Ile Ile Gln Ala Thr Ile
    85             90             95

55  Ala Asp Gln Gly Asn Lys Lys Gly Tyr Ile Asp Asn Trp Ser Ser Asp
    100            105            110

60  Gly Phe Leu Arg Trp Ala His Ala Leu Gly Phe Ile Glu Tyr Ile Asn
    115            120            125

65  Lys Ser Asp Ser Phe Val Ile Thr Asp Val Gly Leu Ala Tyr Ser Lys
    130            135            140

70  Ser Ala Asp Gly Ser Ala Ile Glu Lys Glu Ile Leu Ile Glu Ala Ile
    145            150            155            160

75  Ser Ser Tyr Pro Pro Ala Ile Arg Ile Leu Thr Leu Leu Glu Asp Gly
    165            170            175

80  Gln His Leu Thr Lys Phe Asp Leu Gly Lys Asn Leu Gly Phe Ser Gly
    180            185            190

```

ES 2 709 216 T3

	Glu Ser Gly Phe Thr Ser Leu Pro Glu Gly Ile Leu Leu Asp Thr Leu
	195 200 205
5	Ala Asn Ala Met Pro Lys Asp Lys Gly Glu Ile Arg Asn Asn Trp Glu
	210 215 220
10	Gly Ser Ser Asp Lys Tyr Ala Arg Met Ile Gly Gly Trp Leu Asp Lys
	225 230 235 240
15	Leu Gly Leu Val Lys Gln Gly Lys Lys Glu Phe Ile Ile Pro Thr Leu
	245 250 255
20	Gly Lys Pro Asp Asn Lys Glu Phe Ile Ser His Ala Phe Lys Ile Thr
	260 265 270
25	Gly Glu Gly Leu Lys Val Leu Arg Arg Ala Lys Gly Ser Thr Lys Phe
	275 280 285
30	Thr Arg Val Pro Lys Arg Val Tyr Trp Glu Met Leu Ala Thr Asn Leu
	290 295 300
35	Thr Asp Lys Glu Tyr Val Arg Thr Arg Arg Ala Leu Ile Leu Glu Ile
	305 310 315 320
40	Leu Ile Lys Ala Gly Ser Leu Lys Ile Glu Gln Ile Gln Asp Asn Leu
	325 330 335
45	Lys Lys Leu Gly Phe Asp Glu Val Ile Glu Thr Ile Glu Asn Asp Ile
	340 345 350
50	Lys Gly Leu Ile Asn Thr Gly Ile Phe Ile Glu Ile Lys Gly Arg Phe
	355 360 365
55	Tyr Gln Leu Lys Asp His Ile Leu Gln Phe Val Ile Pro Asn Arg Gly
	370 375 380
60	Val Thr Lys Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu
	385 390 395 400
65	Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile
	405 410 415
70	Glu Ile Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val
	420 425 430
75	Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly
	435 440 445
80	Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile
	450 455 460

ES 2 709 216 T3

	Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn	
	465	470 475 480
5	Leu Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn	
		485 490 495
10	Gln Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr	
		500 505 510
15	Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe	
		515 520 525
20	Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn	
		530 535 540
25	Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu	
		545 550 555 560
30	Met Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe	
		565 570 575
35	Asn Asn Gly Glu Ile Asn Phe	
		580
40	<210> 11	
	<211> 558	
	<212> PRT	
	<213> Acinetobacter lwoffii	
45	<400> 11	
50	Met Ser Thr Trp Leu Leu Gly Asn Thr Thr Val Arg Ser Pro Phe Arg	
		1 5 10 15
55	Leu Ile Asp Gly Leu Lys Val Phe Ala Leu Thr Asn Gly Asp Ile Arg	
		20 25 30
60	Gly Thr Lys Glu Lys Glu Leu Val Phe Cys Lys Ala Leu Val Glu Gly	
		35 40 45
65	Gly Ile Ile Ser Ala Ser Phe Glu Ala Glu Asp Thr Ser Gly Phe Ser	
		50 55 60
70	Asp Thr Thr Tyr Ser Val Gly Arg Lys Trp Arg Ser Ala Leu Glu Lys	
		65 70 75 80
75	Leu Gly Phe Ile Glu Gln Phe Asn Gln Ile Tyr Ile Leu Thr Glu Asn	
		85 90 95
80	Gly Arg Asn Leu Leu Asn Ser Gln Thr Leu Gln Ser Asp Gln Glu Cys	
		100 105 110
85	Tyr Leu Arg Ser Leu Ile Leu Tyr Ser Tyr Lys Ala Glu Asn Ser Asp	
		115 120 125

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

ES 2 709 216 T3

5

10

15

20

25

30

35

40

45

50

55

60

65

Gly Gly Ala Asp Val Ile Phe Glu Tyr Glu Asn Phe Lys Ile Leu Gly
420 425 430

Glu Val Thr Leu Thr Ser Asn Ser Arg Gln Glu Ala Ala Glu Gly Glu
435 440 445

Pro Val Arg Arg His Ile Ala Val Glu Thr Val Asn Thr Pro Asp Lys
450 455 460

Asp Val Tyr Gly Leu Phe Leu Ala Leu Thr Ile Asp Thr Asn Thr Ala
465 470 475 480

Glu Thr Phe Arg His Gly Ala Trp Tyr His Gln Glu Glu Leu Met Asp
485 490 495

Val Lys Ile Leu Pro Leu Thr Leu Glu Ser Phe Lys Lys Tyr Leu Glu
500 505 510

Ser Leu Arg Lys Lys Asn Gln Val Glu Thr Gly Ile Phe Asp Leu Lys
515 520 525

Lys Met Met Asp Glu Ser Leu Lys Leu Arg Glu Thr Leu Thr Ala Pro
530 535 540

Gln Trp Lys Asn Glu Ile Thr Asn Lys Phe Ala Arg Pro Ile
545 550 555

<210> 12
<211> 556
<212> PRT
<213> Micrococcus lylae
<400> 12

Met Ala Ser Leu Ser Lys Thr Lys His Leu Phe Gly Phe Thr Ser Pro
1 5 10 15

Arg Thr Ile Glu Lys Ile Ile Pro Glu Leu Asp Ile Leu Ser Gln Gln
20 25 30

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

Ile Phe Asn Ser Asp Phe Tyr Glu Gly Thr Thr Tyr Pro Gln Asp Pro
50 55 60

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

Phe Ile Gln Leu Lys Pro Val Ile Gln Leu Thr Lys Ala Gly Asn Gln
85 90 95

Leu Val Asn Gln Lys Arg Leu Pro Glu Leu Phe Thr Lys Gln Leu Leu
100 105 110

ES 2 709 216 T3

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

ES 2 709 216 T3

Pro Asp Ile Glu Ile Asn Tyr Gly Ser Phe Ser Cys Ile Val Glu Val
420 425 430

5 Thr Met Ser Ser Gly Glu Thr Gln Phe Asn Met Glu Gly Ser Ser Val
435 440 445

10 Pro Arg His Tyr Gly Asp Leu Val Arg Lys Val Asp His Asp Ala Tyr
450 455 460

Cys Ile Phe Ile Ala Pro Lys Val Ala Pro Gly Thr Lys Ala His Phe
465 470 475 480

15 Phe Asn Leu Asn Arg Leu Ser Thr Lys His Tyr Gly Gly Lys Thr Lys
485 490 495

20 Ile Ile Pro Met Ser Leu Asp Asp Phe Ile Cys Phe Leu Gln Val Gly
500 505 510

Ile Thr His Asn Phe Gln Asp Ile Asn Lys Leu Lys Asn Trp Leu Asp
515 520 525

25 Asn Leu Ile Asn Phe Asn Leu Glu Ser Glu Asp Glu Glu Ile Trp Phe
530 535 540

30 Glu Glu Ile Ile Ser Lys Ile Ser Thr Trp Ala Ile
545 550 555

<210> 13
<211> 323
<212> PRT
<213> Streptomyces spec. Bf-61

40 Met Asn Ser Ser Asp Gly Ile Asp Gly Thr Val Ala Ser Ile Asp Thr
1 5 10 15

45 Ala Arg Ala Leu Leu Lys Arg Phe Gly Phe Asp Ala Gln Arg Tyr Asn
20 25 30

Val Arg Ser Ala Val Thr Leu Leu Ala Leu Ala Gly Leu Lys Pro Gly
35 40 45

50 Asp Arg Trp Val Asp Ser Thr Thr Pro Arg Leu Gly Val Gln Lys Ile
50 55 60

55 Met Asp Trp Ser Gly Glu His Trp Ala Lys Pro Tyr Ala Thr Gly Ser
65 70 75 80

Arg Glu Asp Phe Arg Lys Lys Thr Leu Arg Gln Trp Val Asp Asn Gly
85 90 95

60 Phe Ala Val Leu Asn Ala Asp Asn Leu Asn Ile Ala Thr Asn Ser Gln
100 105 110

65 Leu Asn Glu Tyr Cys Leu Ser Asp Glu Ala Leu Gln Ala Leu Arg Ala
115 120 125

ES 2 709 216 T3

	Tyr Gly Thr Glu Gly Phe Glu Glu Ser Leu Val Val Phe Leu Asp Glu	
	130 135 140	
5	Ala Ser Lys Ala Val Lys Ala Arg Ala Glu Ala Leu Gln Ala Ala Met	
	145 150 155 160	
10	Ile Ser Val Asp Leu Pro Gly Gly Glu Glu Phe Leu Leu Ser Pro Ala	
	165 170 175	
	Gly Gln Asn Pro Leu Leu Lys Lys Met Val Glu Glu Phe Val Pro Arg.	
15	180 185 190	
	Phe Ala Pro Arg Ser Thr Val Leu Tyr Leu Gly Asp Thr Arg Gly Lys	
	195 200 205	
20	His Ser Leu Phe Glu Arg Glu Ile Phe Glu Glu Val Leu Gly Leu Thr	
	210 215 220	
25	Phe Asp Pro His Gly Arg Met Pro Asp Leu Ile Leu His Asp Glu Val	
	225 230 235 240	
	Arg Gly Trp Leu Phe Leu Met Glu Ala Val Lys Ser Lys Gly Pro Phe	
30	245 250 255	
	Asp Glu Glu Arg His Arg Ser Leu Gln Glu Leu Phe Val Thr Pro Ser	
	260 265 270	
35	Ala Gly Leu Ile Phe Val Asn Cys Phe Glu Asn Arg Glu Ser Met Arg	
	275 280 285	
40	Gln Trp Leu Pro Glu Leu Ala Trp Glu Thr Glu Ala Trp Val Ala Glu	
	290 295 300	
	Asp Pro Asp His Leu Ile His Leu Asn Gly Ser Arg Phe Leu Gly Pro	
45	305 310 315 320	
	Tyr Glu Arg	
50	<210> 14 <211> 323 <212> PRT <213> Streptomyces diastaticus	
55	<400> 14	
	Met Thr Asn Ser Asn Asp Ile Asp Glu Thr Ala Ala Thr Ile Asp Thr	
	1 5 10 15	
60	Ala Arg Ala Leu Leu Lys Ser Phe Gly Phe Glu Ala Gln Arg His Asn	
	20 25 30	
65	Val Arg Ser Ala Val Thr Leu Leu Ala Leu Ala Gly Leu Lys Pro Gly	
	35 40 45	

ES 2 709 216 T3

Asp His Trp Ala Asp Ser Thr Thr Pro Arg Leu Gly Val Gln Lys Ile
50 55 60

Met Asp Trp Ser Gly Ala Tyr Trp Ala Lys Pro Tyr Ala Thr Gly Ser

65 70 75 80

Arg Glu Asp Phe Arg Lys Lys Thr Leu Arg Gln Trp Val Asp Asn Gly
85 90 95

Phe Ala Val Leu Asn Pro Asp Asn Leu Asn Ile Ala Thr Asn Ser Gln
100 105 110

Leu Asn Glu Tyr Cys Leu Ser Asp Glu Ala Ala Gln Ala Ile Arg Ser
115 120 125

Tyr Gly Thr Asp Ala Phe Glu Ser Ala Leu Val Asp Phe Leu Ser Lys
130 135 140

Ala Ser Asp Thr Val Arg Ala Arg Ala Glu Ala Leu Arg Ala Ala Met
145 150 155 160

Ile Ser Val Asp Leu Ala Asp Gly Asp Glu Phe Leu Leu Ser Pro Ala
165 170 175

Gly Gln Asn Pro Leu Leu Lys Lys Met Val Glu Glu Phe Met Pro Arg
180 185 190

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

His Thr Arg Phe Glu Lys Arg Ile Phe Glu Glu Thr Leu Gly Leu Thr
210 215 220

Phe	Asp	Pro	His	Gly	Arg	Met	Pro	Asp	Leu	Val	Leu	His	Asp	Lys	Val
225					230					235					240

Arg Lys Trp Leu Phe Leu Met Glu Ala Val Lys Ser Lys Gly Pro Phe
245 250 255

Asp Glu Glu Arg His Arg Thr Leu Arg Glu Leu Phe Ala Thr Pro Val
260 265 270

Ala Gly Leu Val Phe Val Asn Cys Phe Glu Asn Arg Glu Ala Met Arg
275 280 285

Gln Trp Leu Pro Glu Leu Ala Trp Glu Thr Glu Ala Trp Val Ala Asp
290 295 300

Asp Pro Asp His Leu Ile His Leu Asn Gly Ser Arg Phe Leu Gly Pro
305 310 315 320

Tyr Glu Arg

ES 2 709 216 T3

<210> 15

<211> 602

<212> PRT

5 <213> Streptococcus sanguis

<400> 15

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

Trp Leu Ala Gly Val Leu Val Thr Ala Gly Leu Ala Thr Ser Thr Thr
245 250 255

5 Lys Lys Val Thr His Thr Tyr Lys Asn Arg Lys Phe Glu Leu Thr Leu
260 265 270

10 Gly Thr Val Tyr Gln Ile Thr Ala Lys Gly Leu Thr Ala Leu Lys Glu
275 280 285

15 Val Asn Gly Lys Ser Arg Tyr Pro Arg Ser Arg Lys Arg Val Met Trp
290 295 300

20 Glu Phe Leu Ala Thr Lys Asp Lys Glu Ala Ile Ala Lys Lys Thr Ser
305 310 315 320

25 Arg Ser Leu Met Leu Lys His Leu Thr Glu Lys Lys Asn Pro Ile Gln
325 330 335

30 Ala Glu Val Ile Ala Thr Leu Ile Asn Thr Asp Tyr Pro Thr Leu Glu
340 345 350

35 Ile Thr Pro Glu Glu Val Ile Asp Asp Cys Ile Gly Leu Asn Arg Ile
355 360 365

40 Gly Ile Glu Ile Leu Ile Asp Gly Asp Lys Leu Thr Leu Asn Asp Lys
370 375 380

45 Leu Phe Asp Phe Glu Ile Pro Val Gln Lys Asp Val Val Leu Glu Lys
385 390 395 400

50 Ser Asp Ile Glu Lys Phe Lys Asn Gln Leu Arg Thr Glu Leu Thr Asn
405 410 415

55 Ile Asp His Ser Tyr Leu Lys Gly Ile Asp Ile Ala Ser Lys Lys Lys
420 425 430

60 Thr Ser Asn Val Glu Asn Thr Glu Phe Glu Ala Ile Ser Thr Lys Ile
435 440 445

65 Phe Thr Asp Glu Leu Gly Phe Ser Gly Lys His Leu Gly Gly Ser Asn
450 455 460

Lys Pro Asp Gly Leu Leu Trp Asp Asp Asp Cys Ala Ile Ile Leu Asp
465 470 475 480

Ser Lys Ala Tyr Ser Glu Gly Phe Pro Leu Thr Ala Ser His Thr Asp
485 490 495

Ala Met Gly Arg Tyr Leu Arg Gln Phe Thr Glu Arg Lys Glu Glu Ile
500 505 510

Lys Pro Thr Trp Trp Asp Ile Ala Pro Glu His Leu Asp Asn Thr Tyr
515 520 525

ES 2 709 216 T3

Phe Ala Tyr Val Ser Gly Ser Phe Ser Gly Asn Tyr Lys Glu Gln Leu
 530 535 540
 5
 Gln Lys Phe Arg Gln Asp Thr Asn His Leu Gly Gly Ala Leu Glu Phe
 545 550 555 560
 10
 Val Lys Leu Leu Leu Leu Ala Asn Asn Tyr Lys Thr Gln Lys Met Ser
 565 570 575
 15
 Lys Lys Glu Val Lys Lys Ser Ile Leu Asp Tyr Asn Ile Ser Tyr Glu
 580 585 590
 20
 Glu Tyr Ala Pro Leu Leu Ala Glu Ile Glu
 595 600
 <210> 16
 <211> 593
 <212> PRT
 <213> Clostridium leptum
 25
 <400> 16
 Met Ile His Leu Ile Pro Thr Glu Ala Lys Arg Phe Arg Thr Phe Gly
 1 5 10 15
 30
 Trp Val Gln Asp Pro Ser Asp Phe Arg Ser Leu Cys Asp Val Val Ala
 20 25 30
 35
 Ile Phe Asp Glu Thr Ser Leu Lys His Gln Glu Leu Ala Gly Gln Val
 35 40 45
 40
 Ile Pro Ala Leu Val Glu Glu Arg Asp Gly Arg Gln Arg Leu Leu Asp
 50 55 60
 Ala Leu Asn Gln Arg Pro Leu Arg Ile Ser Tyr Thr Asp Leu Val Gly
 65 70 75 80
 45
 Thr Ser Phe Thr Pro Arg Ser Ala Ala Arg Cys Asn Gly Ile Val Gln
 85 90 95
 50
 Ala Ala Val Arg Gly Gln Val Arg Pro Phe Ile Gly Asp Trp Pro Ala
 100 105 110
 Asp Asn Phe Val Arg Trp Ala His Ala Leu Gly Phe Leu Arg Tyr Gly
 115 120 125
 55
 Tyr Gln Gly Asp Ala Phe Glu Leu Thr Glu Thr Gly Lys Ala Leu Ala
 130 135 140
 60
 Gln Ala Arg Thr Gln Gly Glu Glu Leu Asn Ser Gln Glu Lys Glu Leu
 145 150 155 160
 65
 Leu Thr Ser Ala Val Leu Ala Tyr Pro Pro Ala Val Arg Ile Leu Ser
 165 170 175

5 Leu Leu Gly Glu Gly Glu Gly Ala His Leu Thr Lys Phe Glu Leu Gly
 180 185 190
 Lys Gln Leu Gly Phe Val Gly Glu Asp Gly Phe Thr Ser Leu Pro Gln
 195 200 205
 10 Thr Val Leu Val Arg Ser Leu Ala Ser Ser Lys Asp Ala Lys Glu Lys
 210 215 220
 Asn Lys Met Lys Thr Asp Trp Asp Gly Ser Ser Asp Lys Tyr Ala Arg
 225 230 235 240
 15 Met Ile Ala Lys Trp Leu Glu Lys Leu Gly Leu Val Lys Gln Glu Ala
 245 250 255
 20 Lys Pro Val Thr Val Thr Leu Ala Gly Arg Lys Tyr Thr Glu Ser Ile
 260 265 270
 25 Gly Gln Ser Tyr Val Ile Thr Gly Leu Gly Ile Thr Ala Leu Asn Arg
 275 280 285
 Thr Leu Gly Lys Ser Arg His Lys Arg Ile Pro Lys Asn Val Ser Phe
 290 295 300
 30 Glu Met Met Ala Thr Lys Gly Asp Asp Arg Glu Tyr Leu Arg Thr Arg
 305 310 315 320
 35 Arg Thr Cys Val Leu Lys Ala Val Ser Glu Gly Lys Gly Arg Val Ser
 325 330 335
 Tyr Thr Glu Ile Gln Lys Tyr Leu Glu Ala Leu Gly Leu Gln Glu Asp
 340 345 350
 40 Glu Ala Thr Ile Arg Asp Asp Val Gln Gly Leu Ile His Ile Gly Leu
 355 360 365
 45 Asn Ile Ala Ala Gly Glu Arg Glu Cys Val Trp Lys Asp Glu Ile Asn
 370 375 380
 Asp Leu Ile Leu Pro Val Pro Lys Lys Leu Ala Lys Ser Ser Gln Ser
 385 390 395 400
 50 Glu Thr Lys Glu Lys Leu Arg Glu Lys Leu Arg Asn Leu Pro His Glu
 405 410 415
 55 Tyr Leu Ser Leu Val Asp Leu Ala Tyr Asp Ser Lys Gln Asn Arg Leu
 420 425 430
 Phe Glu Met Lys Val Ile Glu Leu Leu Thr Glu Glu Cys Gly Phe Gln
 435 440 445
 60 Gly Leu His Leu Gly Gly Ser Arg Arg Pro Asp Gly Val Leu Tyr Thr
 450 455 460
 65 Ala Gly Leu Thr Asp Asn Tyr Gly Ile Ile Leu Asp Thr Lys Ala Tyr
 465 470 475 480

ES 2 709 216 T3

Ser Ser Gly Tyr Ser Leu Pro Ile Ala Gln Ala Asp Glu Met Glu Arg
 485 490 495
 5 Tyr Val Arg Glu Asn Gln Thr Arg Asp Glu Leu Val Asn Pro Asn Gln
 500 505 510
 10 Trp Trp Glu Asn Phe Glu Asn Gly Leu Gly Thr Phe Tyr Phe Leu Phe
 515 520 525
 15 Val Ala Gly His Phe Asn Gly Asn Val Gln Ala Gln Leu Glu Arg Ile
 530 535 540
 20 Ser Arg Asn Thr Gly Val Leu Gly Ala Ala Ala Ser Ile Ser Gln Leu
 545 550 555 560
 25 Leu Leu Leu Ala Asp Ala Ile Arg Gly Gly Arg Met Asp Arg Glu Arg
 565 570 575
 30 Leu Arg His Leu Met Phe Gln Asn Glu Glu Phe Leu Leu Glu Gln Glu
 580 585 590
 35 Leu
 <210> 17
 <211> 587
 <212> PRT
 <213> Clostridium spec. 7_2_43 FAA
 40 Met Ile Asn Ile Ile Asp Val Asn Asn Lys Thr Ile Arg Thr Phe Gly
 1 5 10 15
 45 Trp Val Gln Asn Pro Ser Asn Phe Glu Ser Leu Lys Lys Val Val Ala
 20 25 30
 50 Ile Phe Asp Asn Thr Ser Lys Thr Tyr Asn Glu Leu Lys Asp Lys Lys
 35 40 45
 55 Ile Lys Lys Leu Val Asp Glu Arg Asp Gly Gln Lys Glu Leu Leu Asn
 50 55 60
 60 Ala Leu Asn Ala Asn Pro Leu Lys Ile Lys Tyr Cys Asn Leu Val Gly
 65 70 75 80
 65 Thr Ser Phe Thr Pro Arg Ser Ser Ala Arg Cys Asn Gly Ile Val Gln
 85 90 95
 70 Ala Thr Val Lys Gly Gln Arg Lys Glu Phe Ile Asp Asp Trp Ser Ser
 100 105 110
 75 Asp Asn Phe Val Arg Trp Ala His Ala Leu Gly Phe Ile Lys Tyr Asn
 115 120 125
 80

ES 2 709 216 T3

Tyr Asp Thr Asp Thr Phe Glu Ile Thr Asp Val Gly Arg Lys Tyr Val
 130 135 140
 5
 Gln Ser Glu Asp Asp Ser Asn Glu Glu Ser Thr Ile Leu Glu Glu Ala
 145 150 155 160
 10
 Met Leu Ser Tyr Pro Pro Val Ala Arg Val Leu Thr Leu Leu Ser Asn
 165 170 175
 Gly Glu His Leu Thr Lys Tyr Glu Ile Gly Lys Lys Leu Gly Phe Val
 180 185 190
 15
 Gly Glu Ala Gly Phe Thr Ser Leu Pro Leu Asn Val Leu Ile Met Thr
 195 200 205
 20
 Leu Ala Thr Thr Asp Glu Pro Lys Glu Lys Asn Lys Ile Lys Thr Asp
 210 215 220
 25
 Trp Asp Gly Ser Ser Asp Lys Tyr Ala Arg Met Ile Ser Gly Trp Leu
 225 230 235 240
 30
 Val Lys Leu Gly Leu Leu Val Gln Arg Pro Lys Leu Val Thr Val Asp
 245 250 255
 Phe Gly Gly Glu Leu Tyr Ser Glu Thr Ile Gly His Ala Tyr Met Ile
 260 265 270
 35
 Thr Asp Arg Gly Leu Lys Ala Val Arg Arg Leu Leu Gly Ile Asn Lys
 275 280 285
 40
 Val Ala Arg Val Ser Lys Asn Val Phe Trp Glu Met Leu Ala Thr Lys
 290 295 300
 45
 Gly Ile Asp Lys Asn Tyr Ile Arg Thr Arg Arg Ala Tyr Ile Leu Lys
 305 310 315 320
 Ile Leu Ile Glu Ser Asn Lys Val Leu Thr Leu Glu Asp Ile Lys Gly
 325 330 335
 50
 Lys Leu Lys Leu Ala Ser Ile Asn Glu Ser Ile Asn Thr Ile Lys Asp
 340 345 350
 55
 Asp Ile Asn Gly Leu Ile Asn Thr Gly Ile Asn Ile Lys Ser Glu Thr
 355 360 365
 60
 Thr Gly Tyr Lys Ile Tyr Asp Ser Ile Asn Asp Phe Ile Ile Pro Lys
 370 375 380
 65

ES 2 709 216 T3

Thr Gly Asp Thr Glu Gly Ile Lys Ser Asn Ile Ser Leu Leu Lys Asp
 385 390 395 400
 5
 Glu Leu Arg Gly Gln Ile Ser His Ile Ser His Glu Tyr Leu Ser Leu
 405 410 415
 10
 Ile Asp Leu Ala Phe Asp Ser Lys Gln Asn Arg Leu Phe Glu Met Lys
 420 425 430
 15
 Val Leu Glu Leu Leu Val Asn Glu Tyr Gly Phe Lys Gly Arg His Leu
 435 440 445
 20
 Gly Gly Ser Arg Lys Pro Asp Gly Ile Val Tyr Ser Thr Thr Leu Glu
 450 455 460
 25
 Asp Asn Phe Gly Ile Ile Val Asp Thr Lys Ala Tyr Ser Glu Gly Tyr
 465 470 475 480
 30
 Ser Leu Pro Ile Ser Gln Ala Asp Glu Met Glu Arg Tyr Val Arg Glu
 485 490 495
 35
 Asn Ser Asn Arg Asp Glu Glu Val Asn Pro Asn Lys Trp Trp Glu Asn
 500 505 510
 40
 Phe Ser Glu Glu Val Lys Lys Tyr Tyr Phe Val Phe Ile Ser Gly Ser
 515 520 525
 45
 Phe Lys Gly Lys Phe Glu Glu Gln Leu Arg Arg Leu Ser Met Thr Thr
 530 535 540
 50
 Gly Val Asn Gly Ser Ala Val Asn Val Val Asn Leu Leu Leu Gly Ala
 545 550 555 560
 55
 Glu Lys Ile Arg Ser Gly Glu Met Thr Ile Glu Glu Leu Glu Arg Ala
 565 570 575
 60
 Met Phe Asn Asn Ser Glu Phe Ile Leu Lys Tyr
 580 585
 65
 Met Ala Glu Arg Thr Leu Gly Trp Ile Gln Asn Pro Ser Ser Phe Glu
 1 5 10 15
 Asn Leu Lys Asn Val Val Ser Val Phe Asp Lys Asn Ser Asp Ile Tyr
 20 25 30
 Lys Glu Ile Leu Asn Thr Lys Leu Pro Lys Leu Val Lys Asp Leu Asp
 35 40 45

<210> 18

<211> 589

<212> PRT

<213> Peptoniphilus duerdenii

<400> 18

ES 2 709 216 T3

5

Leu Gln Asn Lys Leu Ile Ser Glu Leu Glu Lys Asp Pro Leu Glu Met
50 55 60

10

Asp Tyr Val Leu Leu Lys Gly His Gly Ile Lys Ser Gly Gln Lys Arg
65 70 75 80

15

Ala Asp Ala Glu Cys Ser Gly Ile Val Gln Ala Ala Ile Thr Thr Gln
85 90 95

20

Gly Gly Arg Ala Tyr Thr Asp Asp Trp Thr Ala Asp Gly Phe Leu Arg
100 105 110

25

Trp Gly Ile Ser Ile Gly Leu Leu Asp Tyr Asp Thr Glu Lys Asp Thr
115 120 125

30

Val Ser Ile Thr Lys Leu Gly Glu Lys Phe Val Lys Ser Asn Ser Glu
130 135 140

35

Asp Ser Asp Lys Glu Ile Leu Ile Ser Ala Phe Leu Ser Tyr Pro Pro
145 150 155 160

40

Ala Val Arg Ile Leu Thr Leu Leu Glu Asn Gly Asp His Leu Thr Lys
165 170 175

45

Phe Glu Leu Gly Lys Gln Leu Gly Gly Leu Gly Glu Ala Gly Phe Thr
180 185 190

50

Ser Ile Pro Gln Asp Leu Tyr Ile Gln Ala Ile Glu Leu Ala Ala Asp
195 200 205

55

Lys Asp Lys Ala Ser Ile Arg Ser Asn Thr Glu Gly Ser Ala Asp Lys
210 215 220

60

Tyr Ala Arg Met Ile Ser Gly Trp Leu Ser Lys Val Gly Leu Ile Gln
225 230 235 240

65

Arg Ile Gly Lys Glu Val Ser Thr Lys Ile Gly Asp Val Glu Tyr Lys
245 250 255

Val Asn Ile Gly His Ser Phe Arg Ile Thr Leu Asn Gly Ile Lys Glu
260 265 270

Leu Lys Arg Ala Met Gly Leu Ser Ser Tyr Pro Lys Thr Asp Lys Ile
275 280 285

Val Tyr Trp Gln Met Leu Ala Thr Lys Gly Lys Asp Arg Asp Tyr Ile
290 295 300

ES 2 709 216 T3

	Ser	Ile	Thr	Thr	Ile	Glu	Asp	Glu	Leu	Lys	Val	Ile	Glu	Ala	Met	Gly	
				340					345					350			
5	Leu	Ser	Phe	Lys	His	Ser	Arg	Asn	Gly	Tyr	Val	Ile	Asp	Asp	Asn	Ile	
			355					360					365				
	Ile	Lys	Leu	Glu	Ile	Pro	Arg	Thr	Lys	Ile	Ser	Lys	Thr	Asn	Val	Leu	
10		370					375					380					
	Glu	Leu	Lys	Asp	Lys	Val	Arg	Asp	Lys	Leu	Lys	Tyr	Val	Asp	His	Arg	
	385					390					395					400	
15	Tyr	Leu	Ala	Leu	Ile	Asp	Leu	Ala	Tyr	Asp	Gly	Thr	Ala	Asn	Arg	Asp	
					405					410					415		
	Phe	Glu	Ile	Gln	Thr	Ile	Asp	Leu	Leu	Ile	Asn	Glu	Leu	Lys	Phe	Lys	
20				420					425					430			
	Gly	Val	Arg	Leu	Gly	Glu	Ser	Arg	Lys	Pro	Asp	Gly	Ile	Ile	Ser	Tyr	
			435					440					445				
25	Asn	Ile	Asn	Gly	Val	Ile	Ile	Asp	Asn	Lys	Ala	Tyr	Ser	Thr	Gly	Tyr	
		450					455						460				
	Asn	Leu	Pro	Ile	Asn	Gln	Ala	Asp	Glu	Met	Ile	Arg	Tyr	Ile	Glu	Glu	
30		465				470					475					480	
	Asn	Gln	Thr	Arg	Asp	Glu	Lys	Ile	Asn	Ser	Asn	Lys	Trp	Trp	Glu	Ser	
				485						490					495		
35	Phe	Asp	Asp	Lys	Val	Lys	Asp	Phe	Asn	Tyr	Leu	Phe	Val	Ser	Ser	Phe	
				500					505					510			
	Phe	Lys	Gly	Asn	Phe	Lys	Asn	Asn	Leu	Lys	His	Ile	Ala	Asn	Arg	Thr	
40			515					520					525				
	Gly	Val	Ser	Gly	Gly	Ala	Ile	Asn	Val	Glu	Asn	Leu	Leu	Tyr	Phe	Ala	
45		530					535					540					
	Glu	Glu	Leu	Lys	Ala	Gly	Arg	Leu	Ser	Tyr	Val	Asp	Ser	Phe	Lys	Met	
	545					550					555					560	
50	Tyr	Asp	Asn	Asp	Glu	Ile	Tyr	Val	Gly	Asp	Phe	Ser	Asp	Tyr	Ser	Tyr	
					565					570					575		
	Val	Lys	Phe	Ala	Ala	Glu	Glu	Glu	Gly	Glu	Tyr	Leu	Thr				
55				580					585								

<210> 19
 <211> 279
 <212> PRT
 <213> Acinetobacter lwoffii

<400> 19

ES 2 709 216 T3

	Lys	Glu	Thr	Asn	Ile	Leu	Leu	Val	Glu	Gln	Leu	Glu	Glu	Thr	Leu	Asn	
	1				5					10					15		
5	Arg	Asn	Arg	Ile	Leu	Phe	Glu	Lys	Asn	Ser	Ser	Ile	Ala	Gln	Ala	Pro	
				20					25					30			
10	Ile	Gly	Glu	Ile	Lys	Asn	Tyr	Arg	Tyr	His	Leu	Glu	Glu	Leu	Leu	Phe	
			35					40					45				
15	Glu	Asn	Asn	Glu	Lys	Lys	Phe	Ala	Glu	Asn	Gln	Lys	Asn	Glu	Trp	Asp	
		50					55					60					
20	Glu	Ile	Leu	Ala	Tyr	Met	Asp	Leu	Leu	Ile	Ser	Pro	Lys	Pro	Ile	Ser	
	65					70					75				80		
25	Ile	Glu	Ile	Ala	Asp	Lys	Glu	Ile	Ser	Ile	Pro	Ser	Gly	Glu	Arg	Pro	
					85					90					95		
30	Ala	Tyr	Phe	Glu	Trp	Val	Leu	Trp	Arg	Ala	Phe	Leu	Ala	Leu	Asn	His	
				100					105					110			
35	Leu	Ile	Ile	Glu	Pro	Gln	Gln	Cys	Arg	Arg	Phe	Lys	Val	Asp	Gln	Asp	
			115					120					125				
40	Phe	Lys	Pro	Ile	His	Asn	Ala	Pro	Gly	Gly	Gly	Ala	Asp	Val	Ile	Phe	
		130					135					140					
45	Glu	Tyr	Glu	Asn	Phe	Lys	Ile	Leu	Gly	Glu	Val	Thr	Leu	Thr	Ser	Asn	
	145					150					155				160		
50	Ser	Arg	Gln	Glu	Ala	Ala	Glu	Gly	Glu	Pro	Val	Arg	Arg	His	Ile	Ala	
					165					170					175		
55	Val	Glu	Thr	Val	Asn	Thr	Pro	Asp	Lys	Asp	Val	Tyr	Gly	Leu	Phe	Leu	
				180					185					190			
60	Ala	Leu	Thr	Ile	Asp	Thr	Asn	Thr	Ala	Glu	Thr	Phe	Arg	His	Gly	Ala	
			195				200						205				
65	Trp	Tyr	His	Gln	Glu	Glu	Leu	Met	Asp	Val	Lys	Ile	Leu	Pro	Leu	Thr	
		210					215					220					
70	Leu	Glu	Ser	Phe	Lys	Lys	Tyr	Leu	Glu	Ser	Leu	Arg	Lys	Lys	Asn	Gln	
	225					230					235				240		
75	Val	Glu	Thr	Gly	Ile	Phe	Asp	Leu	Lys	Lys	Met	Met	Asp	Glu	Ser	Leu	
				245						250					255		
80	Lys	Leu	Arg	Glu	Thr	Leu	Thr	Ala	Pro	Gln	Trp	Lys	Asn	Glu	Ile	Thr	
				260					265					270			
85	Asn	Lys	Phe	Ala	Arg	Pro	Ile										
				275													

ES 2 709 216 T3

<210> 20
 <211> 201
 <212> PRT
 <213> Clostridium leptum

<400> 20

```

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

```

<210> 21
 <211> 250
 <212> PRT
 <213> Micrococcus lylae

<400> 21

ES 2 709 216 T3

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

ES 2 709 216 T3

<210> 22
 <211> 213
 <212> PRT
 <213> Peptoniphilus duerdenii

<400> 22

5

10

15

20

25

30

35

40

45

50

55

60

65

```

Lys Ile Ser Lys Thr Asn Val Leu Glu Leu Lys Asp Lys Val Arg Asp
1      5      10      15

Lys Leu Lys Tyr Val Asp His Arg Tyr Leu Ala Leu Ile Asp Leu Ala
20      25      30

Tyr Asp Gly Thr Ala Asn Arg Asp Phe Glu Ile Gln Thr Ile Asp Leu
35      40      45

Leu Ile Asn Glu Leu Lys Phe Lys Gly Val Arg Leu Gly Glu Ser Arg
50      55      60

Lys Pro Asp Gly Ile Ile Ser Tyr Asn Ile Asn Gly Val Ile Ile Asp
65      70      75      80

Asn Lys Ala Tyr Ser Thr Gly Tyr Asn Leu Pro Ile Asn Gln Ala Asp
85      90      95

Glu Met Ile Arg Tyr Ile Glu Glu Asn Gln Thr Arg Asp Glu Lys Ile
100     105     110

Asn Ser Asn Lys Trp Trp Glu Ser Phe Asp Asp Lys Val Lys Asp Phe
115     120     125

Asn Tyr Leu Phe Val Ser Ser Phe Phe Lys Gly Asn Phe Lys Asn Asn
130     135     140

Leu Lys His Ile Ala Asn Arg Thr Gly Val Ser Gly Gly Ala Ile Asn
145     150     155     160

Val Glu Asn Leu Leu Tyr Phe Ala Glu Glu Leu Lys Ala Gly Arg Leu
165     170     175

Ser Tyr Val Asp Ser Phe Lys Met Tyr Asp Asn Asp Glu Ile Tyr Val
180     185     190

Gly Asp Phe Ser Asp Tyr Ser Tyr Val Lys Phe Ala Ala Glu Glu Glu
195     200     205

Gly Glu Tyr Leu Thr
210
  
```

ES 2 709 216 T3

<210> 23
 <211> 163
 <212> PRT
 <213> Streptomyces spec. Bf-61

5

<400> 23

10

Ile Ser Val Asp Leu Pro Gly Gly Glu Glu Phe Leu Leu Ser Pro Ala
 1 5 10 15

Gly Gln Asn Pro Leu Leu Lys Lys Met Val Glu Glu Phe Val Pro Arg
 20 25 30

15

Phe Ala Pro Arg Ser Thr Val Leu Tyr Leu Gly Asp Thr Arg Gly Lys
 35 40 45

20

His Ser Leu Phe Glu Arg Glu Ile Phe Glu Glu Val Leu Gly Leu Thr
 50 55 60

25

Phe Asp Pro His Gly Arg Met Pro Asp Leu Ile Leu His Asp Glu Val
 65 70 75 80

Arg Gly Trp Leu Phe Leu Met Glu Ala Val Lys Ser Lys Gly Pro Phe
 85 90 95

30

Asp Glu Glu Arg His Arg Ser Leu Gln Glu Leu Phe Val Thr Pro Ser
 100 105 110

35

Ala Gly Leu Ile Phe Val Asn Cys Phe Glu Asn Arg Glu Ser Met Arg
 115 120 125

Gln Trp Leu Pro Glu Leu Ala Trp Glu Thr Glu Ala Trp Val Ala Glu
 130 135 140

40

Asp Pro Asp His Leu Ile His Leu Asn Gly Ser Arg Phe Leu Gly Pro
 145 150 155 160

45

Tyr Glu Arg

<210> 24
 <211> 163
 <212> PRT
 <213> Streptomyces diastaticus

50

<400> 24

55

Ile Ser Val Asp Leu Ala Asp Gly Asp Glu Phe Leu Leu Ser Pro Ala
 1 5 10 15

Gly Gln Asn Pro Leu Leu Lys Lys Met Val Glu Glu Phe Met Pro Arg
 20 25 30

60

Phe Ala Pro Gly Ala Lys Val Leu Tyr Ile Gly Asp Trp Arg Gly Lys
 35 40 45

65

ES 2 709 216 T3

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

<210> 25
 <211> 208
 <212> PRT
 35 <213> Streptococcus sanguis
 <400> 25

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

ES 2 709 216 T3

5 Glu Arg Lys Glu Glu Ile Lys Pro Thr Trp Trp Asp Ile Ala Pro Glu
 115 120 125

 His Leu Asp Asn Thr Tyr Phe Ala Tyr Val Ser Gly Ser Phe Ser Gly
 130 135 140

10 Asn Tyr Lys Glu Gln Leu Gln Lys Phe Arg Gln Asp Thr Asn His Leu
 145 150 155 160

15 Gly Gly Ala Leu Glu Phe Val Lys Leu Leu Leu Ala Asn Asn Tyr
 165 170 175

20 Lys Thr Gln Lys Met Ser Lys Lys Glu Val Lys Lys Ser Ile Leu Asp
 180 185 190

 Tyr Asn Ile Ser Tyr Glu Glu Tyr Ala Pro Leu Leu Ala Glu Ile Glu
 195 200 205

25 <210> 26
 <211> 196
 <212> PRT
 <213> Flavobacterium okeanoikoites

30 <400> 26

 Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His
 1 5 10 15

35 Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala
 20 25 30

40 Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe
 35 40 45

45 Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg
 50 55 60

50 Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly
 65 70 75 80

55 Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile
 85 90 95

 Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr Arg
 100 105 110

 Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser
 115 120 125

60 Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn
 130 135 140

65 Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly
 145 150 155 160

ES 2 709 216 T3

Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys
165 170 175

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

Glu Ile Asn Phe
195

<210> 27
<211> 7903
<212> ADN
<213> Secuencia artificial

<220>
<223> pCAG-ArtTail-AlwI

<400> 27

gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca	120
acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagggg	180
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgctt	300
ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat	360
tagtcatcgc tattaccatg gtcgagggtga gcccacggtt ctgcttcaact ctccccatct	420
cccccccctc cccacccccca attttgtatt tattttatctt ttaattatctt tgtgcagcga	480
tgggggcggg gggggggggg gggcgcgcg caggcggggc ggggcggggc gaggggcggg	540
gcggggcgag gcggagaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtctc	600
cttttatggc gaggcggcg cggcggcggc cctataaaaa gcgaagcgcg cggcgggcg	660
gagtcgctgc gcgctgcctt cgcgccgtgc cccgctcgc cgcgcctcg cgcgcgccgc	720
cccggtctg actgaccgcg ttactccac aggtgagcgg gcgggacggc ccttctcctc	780
cgggctgtaa ttagcgcttg gtttaatgac ggcttgcttc tttctgtgg ctgcgtgaaa	840
gccttgaggg gctccgggag ggcctttgt gcggggggga gcggctcggg gggcgctgc	900
gtgtgtgtgt gcgtggggag cgcgcgtgc ggctccgcgc tgcccggcg ctgtgagcgc	960
tgccggcgcg gcgcggggct ttgtgcgctc cgcagtgtgc gcgaggggag cgcggccggg	1020
ggcgggtgcc cgcggtgcgg ggggggctgc gaggggaaca aaggctgcgt gcggggtgtg	1080
tgcgtagggg ggtgagcagg ggggtgtggc gcgtcggtcg ggctgcaacc cccctgcac	1140
ccccctccc gagttgctga gcacggcccg gcttcgggtg cggggtccg tacggggcgt	1200
ggcgcggggc tcgcccgtgc gggcgggggg tggcggcagg tgggggtgcc gggcgggcg	1260
gggcgcctc gggcgggga gggctcggg gagggcgcg gcggccccg gagcgccgc	1320
ggctgtcgag gcgcggcgag cgcagccat tgccctttat ggtaatcgt cgagagggcg	1380
cagggacttc ctttgtccca aatctgtgcg gagccgaaat ctgggaggcg ccgcgcacc	1440
ccctctagcg ggcgcggggc gaagcgggtc ggcgcgggca ggaaggaaat gggcggggag	1500

ES 2 709 216 T3

5	ggccttcgtg cgctgcgcgc cgcgcgtccc cttctccctc tccagcctcg gggtgtccg	1560
	cggggggacg gctgccttcg ggggggacgg ggcagggcgg ggttcggctt ctggcgtgtg	1620
	accggcggct cttaggcctc tgctaaccat gttcatgcct tcttcttttt cctacagatc	1680
10	cttaattaat aatacgactc actatagggg cgcgcccat gggaacctag aaaaagagga	1740
	aggtggcggc cgctgactac aaggatgacg acgataaacc aggtggcgga ggtagtggcg	1800
	gaggtggggt acccgccagt ccagcagccc aggtggatct gagaaccctc ggctacagcc	1860
15	agcagcagca ggagaagatc aaaccaaagg tgcggtccac cgctgcctcag caccatgaag	1920
	cactgggtggg gcacggtttc acacacgccc atattgtggc tctgtctcag catcccgctg	1980
	cactcgggac tgtggccgtc aaatatcagg acatgatcgc cgctctgcct gaggaacccc	2040
20	acgaagccat tgtgggcgtc ggaaagcagt ggagcgggtc cagagcactc gaagcactcc	2100
	tcaccgctgc cggggaactg cggggtccac cactccagtc cggactggac actggacagc	2160
	tgctgaagat cgctaaacgc ggcggagtga cagctgtgga agctgtgcac gcttgaggga	2220
25	atgctctgac aggagcccca ctgaatctta ctccagaaca ggtcgtcgca atcgcaagta	2280
	acatcggcgg aaaaacaggcc ctcgaaaccg tccagagact cctccccgtg ctgtgccagg	2340
	cccacggact gacccacag caggtggctg ccacgcctag caacggcgga gggaagcagg	2400
30	ctctggagac cgtgcagagg ctgtccccg tccctgtgcca ggcacatggg ctcacacctc	2460
	agcaggtggt cgcaattgcc tccaatggtg gcggaacaca ggccctggaa actgtgcaga	2520
	gactgtctcc cgtgtgtgtc caggtctacg gtctcacacc ccagcaggtg gtcgtctatc	2580
35	catctcatga cgggggcaag caggcactgg agacagtga gcggtgtct cctgtctgt	2640
	gccaggccca cggactcact cctcagcagg tcgtcgccat tgctagtaac ggcggaggga	2700
	aacaggctct ggaaccgtg cagcgccctg tccccgtgct gtgccaagcc cacggcctga	2760
40	ccccccagca ggtggtcgca atcgccctaa acaatggtgg caagcaggcc ctggagactg	2820
	tgacgcgact gctcccagtg ctgtgccagg cccatggact cacaccacag caggtcgtcg	2880
	ctattgcaag caacaatgga gggaaacagg cactggaaac agtccagagg ctgctccccg	2940
45	tgctgtgcca agcgcattga ctactcccc agcaggtcgt cgccatcgt tccaataacg	3000
	gcggcaagca ggccctggag accgtccaga gactgtctcc cgtgtgtgtc caagctcacg	3060
	gactcacacc tgagcaggtc gtggcaatcg cctctaacat tggagggaaa caggccctgg	3120
50	aaactgtaca gcggtgtct cccgtgtgtg gccaaacaca cggactcact ccacagcagg	3180
	tcgtggccat tgcaagtcat gacggaggca agcaggccct ggaaacagtg cagcgccctg	3240
	tcctgtgtct gtgccaggct catggtctga ctctcagca ggtggtggcc atcgcttcca	3300
55	acaatggagg gaagcaggcc ctggagaccg tacagagact gctccccgtg ctgtgccaa	3360
	cgcacggctc gaccctcag caggtcgtcg caatcgccag caatggcggg ggcaagcagg	3420
	ctctcgaaac cgtccagcgg ctctctccag tccctgtgca ggctcacggc ctgacccac	3480
60	agcaggtcgt cgctattgct tctaattggc gagggcgccc tgctctggag agcattgtgg	3540
	ctcagctgtc caggcccgat cctgccctgg ctagatccgc actcactaac gatcatctgg	3600
	tcgtctctgc ttgcctcggt ggacggcccc ctctggacgc agtcaaaaag ggtctcccc	3660
65	atgctccgc actgatcaag agaaccaaca ggagaattcc tgagggatcc gatcgtttaa	3720
	aaaagagac taatatcctc ctctgcagc agctggaaga gaccctcaat cgcaatcgca	3780
	ttctgtttga aaagaactcc tcaatgcac agggcccaat tggcgagatc aagaactacc	3840

ES 2 709 216 T3

	ggtatcacct	ggaggaactg	ctcttcgaga	acaatgaaaa	gaaatttgca	gagaaccaga	3900
	aaaatgagtg	ggacgaaatt	ctggcctaca	tggatctgct	catctcacc	aagcctatca	3960
5	gcattgagat	cgctgacaaa	gaaatttcta	tcccaagtgg	ggagcgaccc	gcatatttcg	4020
	aatgggtgct	gtggagggca	tttctggccc	tcaaccacct	gatcattgag	ccccagcagt	4080
	gcaggagatt	caaggtcgac	caggacttca	agcctatcca	taatgctcca	ggcggagggg	4140
10	cagatgtgat	tttcgagtac	gaaaacttta	agatcctggg	cgaggtcacc	ctcacaagca	4200
	attcccagaca	ggaagcagct	gagggagaac	ccgtgcggcg	ccatattgcc	gtggagacag	4260
	tcaacactcc	tgacaaggat	gtctatggac	tgttcctcgc	tctgaccatc	gacactaata	4320
15	ccgccgagac	atttcgacac	ggggcttgg	atcaccagga	ggaactgatg	gatgtgaaga	4380
	ttctccccct	gactctcgag	tccttcaaga	agtatctgga	atctctcaga	aagaaaaatc	4440
	aggtggagac	aggaatcttt	gacctgaaga	aaatgatgga	tgaaagcctg	aagctccggg	4500
20	aaaccctgac	cgcacccag	tggaanaatg	aatcacaaa	caaattcgcc	agaccaatct	4560
	gaacgcgtaa	atgattgcag	atccactagt	tctagaattc	cagctgagcg	ccggctcgcta	4620
25	ccattaccag	ttggtctggt	gtcaaaaata	ataataaccg	ggcagggggg	atctgcatgg	4680
	atctttgtga	aggaacctta	cttctgtggt	gtgacataat	tggacaaact	acctacagag	4740
	atttaaagct	ctaaggtaaa	tataaaat	ttaagtgtat	aatgtgttaa	actactgatt	4800
30	ctaattgttt	gtgtatttta	gattccaacc	tatggaactg	atgaatggga	gcagtgtgtg	4860
	aatgccagat	ccagacatga	taagatacat	tgatgagttt	ggacaaacca	caactagaat	4920
	gcagtgaaaa	aaatgcttta	tttgtgaaat	ttgtgatgct	attgctttat	ttgtaaccat	4980
35	tataagctgc	aataaacaag	ttaacaacaa	caattgcatt	cattttatgt	ttcaggttca	5040
	gggggaggtg	tgggaggttt	tttaaagcaa	gtaaaacctc	tacaaatgtg	gtatggctga	5100
40	ttatgatctg	cgcccgccac	tggccgtcgt	tttacaacgt	cgtgactggg	aaaaccctgg	5160
	cgttacccaa	cttaatcgcc	ttgcagcaca	tccccctttc	gccagctggc	gtaatagcga	5220
	agaggcccg	accgatcgcc	cttccaaca	gttgcgcagc	ctgaatggcg	aatggaacgc	5280
45	gccctgtagc	ggcgcattaa	gcgcggcg	tgtggtggtt	acgcgcagcg	tgaccgctac	5340
	acttgccagc	gccctagcgc	ccgctccttt	cgctttcttc	ccttcctttc	tcgccacgtt	5400
	cgccggcttt	ccccgtcaag	ctctaaatcg	ggggctccct	ttagggttcc	gatttagtgc	5460
50	tttacggcac	ctcgaccca	aaaaacttga	ttaggtgat	ggttcacgta	gtgggccatc	5520
	gccctgatag	acggtttttc	gccctttgac	gttgagtc	acgttcttta	atagtggact	5580
	cttgttccaa	actggaacaa	cactcaaccc	tatctcggtc	tattcttttg	atttataagg	5640
55	gattttgccg	atttcggcct	attggttaaa	aatgagctg	atttaacaaa	aatttaacgc	5700
60							
65							

ES 2 709 216 T3

	gaattttaac	aaaatattaa	cgcttacaat	ttaggtggca	cttttcgggg	aaatgtgcgc	5760
	ggaaccccta	tttgtttatt	tttctaaata	cattcaaata	tgtatccgct	catgagacaa	5820
5	taaccctgat	aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	tcaacatttc	5880
	cgtgtcgcgc	ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	tcaccagaa	5940
	acgctggtga	aagtaaaaga	tgctgaagat	cagttgggtg	cacgagtggg	ttacatcgaa	6000
10	ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	ccgaagaacg	ttttccaatg	6060
	atgagcactt	ttaaagttct	gctatgtggc	gcggtattat	cccgtattga	cgccgggcaa	6120
	gagcaactcg	gtcgcgcgat	acactattct	cagaatgact	tggttgagta	ctcaccagtc	6180
15	acagaaaagc	atcttacgga	tggtcatgaca	gtaagagaat	tatgcagtgc	tgccataacc	6240
	atgagtgata	acactgcggc	caacttactt	ctgacaacga	tcggaggacc	gaaggagcta	6300
20	accgcttttt	tgcacaacat	gggggatcat	gtaactcgcc	ttgatcgttg	ggaaccggag	6360
	ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	tgccctgtagc	aatggcaaca	6420
	acgttgcgca	aactattaac	tggtgaacta	cttactctag	cttcccgcca	acaattaata	6480
25	gactggatgg	aggcggataa	agttgcagga	ccacttctgc	gctcggccct	tccggctggc	6540
	tggtttattg	ctgataaatc	tggtgacggg	gagcgtgggt	ctcgcggtat	cattgcagca	6600
	ctggggccag	atggtaagcc	ctcccgatc	gtagtattct	acacgacggg	gagtcaggca	6660
30	actatggatg	aacgaaatag	acagatcgct	gagatagggt	cctcactgat	taagcattgg	6720
	taactgtcag	accaagttta	ctcatatata	ctttagattg	atttaaaact	tcatttttaa	6780
	tttaaaagga	tctaggtgaa	gatccttttt	gataatctca	tgacccaaat	cccttaacgt	6840
35	gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	ttcttgagat	6900
	cctttttttc	tgcgcgtaat	ctgctgcttg	caaacaaaaa	aaccaccgct	accagcgggtg	6960
	gtttgtttgc	cggtacaaga	gctaccaact	ctttttccga	aggtaactgg	cttcagcaga	7020
40	gcgcagatac	caaatactgt	ccttctagtg	tagccgtagt	taggccacca	cttcaagaac	7080
	tctgtagcac	cgctacata	cctcgctctg	ctaactcctgt	taccagtggc	tgctgccagt	7140
45	ggcgataagt	cgtgtcttac	cgggttggac	tcaagacgat	agttaccgga	taaggcgag	7200
	cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	tggtgcgaac	gacctacacc	7260
	gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	cgcttcccga	agggagaaag	7320
50	gcggacaggt	atccggtaag	cggcagggtc	ggaacaggag	agcgacagag	ggagcttcca	7380
	gggggaaacg	cctgggtatct	ttatagtcct	gtcgggtttc	gccacctctg	acttgagcgt	7440
	cgatttttgt	gatgctcgtc	aggggggcgg	agcctatgga	aaaacgccag	caacggggcc	7500
55	tttttacggt	tcttgccctt	ttgctggcct	tttctcaca	tggtctttcc	tgcttatcc	7560
	cctgattctg	tggtataaccg	tattaccgcc	tttgagttag	ctgataccgc	tcgcccagc	7620

60

65

ES 2 709 216 T3

	cgaacgaccg agcgcgagcga gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa	7680
5	ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac	7740
	tggaaagcgg gcagtgagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc	7800
	caggctttac actttatgct tccggctcgt atgttggtg gaattgtgag cggataacaa	7860
10	tttcacacag gaaacagcta tgaccatgag gcgcgcggga ttc	7903
	<210> 28	
	<211> 7669	
	<212> ADN	
15	<213> Secuencia artificial	
	<220>	
	<223> pCAG-ArtTall-CLEDORF	
20	<400> 28	
	gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
25	catatatgga gttccgcgtt acataactta cggtaaattgg cccgcctggc tgaccgccc	120
	acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggg	180
	ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
30	aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggccgcct	300
	ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat	360
	tagtcatcgc tattaccatg gtcgaggtga gcccacgtt ctgcttact ctccccatct	420
35	cccccccctc ccacccccca attttgtatt tattttatttt ttaattattt tgtgcagcga	480
	tggggggcggg gggggggggg gggcgcgcg caggcggggc gggcggggc gaggggcggg	540
40	gcgggggcgag gcgggagagt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc	600
	cttttatggc gaggcggcgg cggcgcgggc cctataaaaa gcgaagcgcg cggcgggcgg	660
	gagtcgctgc gcgctgcctt cgccccgtgc cccgctccgc cgccgcctcg cgcgcgccgc	720
45	ccgggctctg actgaccgcg ttactcccac aggtgagcgg gcgggacggc ccttctcctc	780
	cgggctgtaa ttagcgcttg gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa	840
	gccttgaggg gctccgggag gggcctttgt gcggggggga gcggctcggg gggcgctgc	900
50	gtgtgtgtgt gcgtggggag cggcgctgc ggctccgcgc tgcccggcgg ctgtgagcgc	960
	tgcgggcgcg gcgcggggct ttgtgcgctc cgcagtgtgc gcgaggggag cgcggccggg	1020
	ggcgggtgcc cgcggtgcgg ggggggctgc gaggggaaca aaggctgcgt gcggggtgtg	1080
55	tgcgtggggg ggtgagcagg ggggtgtggc gcgtcggtcg ggctgcaacc cccctgcac	1140
	ccccctccc gagttgtga gcacggcccg gcttcgggtg cggggctccg tacggggcgt	1200
60	ggcgcggggc tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcggggcg	1260

65

ES 2 709 216 T3

	gggcccctc	gggcccggga	gggctcgggg	gaggggcccgcg	gcggcccccg	gagcgcgggc	1320
5	ggctgtcgag	gcgccgcgag	ccgcagccat	tgccttttat	ggtaatcgtg	cgagagggcg	1380
	cagggacttc	ctttgtccca	aatctgtgcg	gagccgaaat	ctgggagggcg	ccgcgcgacc	1440
	ccctctagcg	ggcgccgggc	gaagcgggtgc	ggcgccggca	ggaaggaaat	gggcggggag	1500
10	ggccttcgtg	cgtcgccgcg	ccgccgtccc	cttctccctc	tccagcctcg	gggctgtccg	1560
	cggggggacg	gctgccttcg	ggggggacgg	ggcagggcgg	ggttcggctt	ctggcgtgtg	1620
	accggcggtc	ctagagcctc	tgctaaccat	gttcatgcct	tcttcttttt	cctacagatc	1680
15	cttaattaat	aatacgaactc	actatagggg	ccgccaccat	gggacctaat	aaaaagagga	1740
	aggtggcggc	cgctgactac	aaggatgacg	acgataaacc	aggtggcgga	ggtagtggcg	1800
	gaggtggggg	acccgccagt	ccagcagccc	aggtggatct	gagaaccctc	ggctacagcc	1860
20	agcagcagca	ggagaagatc	aaaccaaagg	tgcgggtccac	cgtcgctcag	caccatgaag	1920
	cactggtggg	gcacggtttc	acacacgccc	atattgtggc	tctgtctcag	catcccgtcg	1980
	cactcgggac	tgtggccgtc	aaatatcagg	acatgatcgc	cgtctgcctc	gaggcaaccc	2040
25	acgaagccat	tgtgggcgtc	ggaaagcagt	ggagcgggtg	cagagcactc	gaagcactcc	2100
	tcaccgtcgc	cgggggaactg	cgggggtccac	cactccagt	cggactggac	actggacagc	2160
	tgctgaagat	cgctaaacgc	ggcggagtga	cagctgtgga	agctgtgcac	gcttggagga	2220
30	atgctctgac	aggagcccca	ctgaatctta	ctccagaaca	ggctgtcgca	atcgcaagta	2280
	acatcggcgg	aaaacaggcc	ctcgaaaccg	tccagagact	cctccccgtg	ctgtgccagg	2340
	cccacggact	gacccacag	caggtggtcg	ccatcgctag	caacggcgga	gggaagcagg	2400
35	ctctggagac	cgtgcagagg	ctgctccccg	tcctgtgcc	ggcacatggg	ctcacacctc	2460
	agcaggtggg	cgcaattgcc	tccaatggtg	gcggaaaaca	ggccctggaa	actgtgcaga	2520
	gactgctccc	cgtgtctgtg	caggctcacg	gtctcacacc	ccagcaggtg	gtcgctatcg	2580
40	catctcatga	cgggggcaag	caggcactgg	agacagtga	gcggctgctc	cctgtcctgt	2640
	gccaggccca	cggactcact	cctcagcagg	tcgtcgccat	tgctagtaac	ggcggagggg	2700
	aacaggctct	ggaaaccgtg	cagcgccctg	tccccgtgct	gtgccaagcc	cacggcctga	2760
45	ccccccagca	ggtggtcgca	atcgccctca	acaatggtgg	caagcaggcc	ctggagactg	2820
	tgcagcgact	gctcccagtg	ctgtgccagg	cccatggact	cacaccacag	caggtcgtcg	2880
	ctattgcaag	caacaatgga	gggaaacagg	cactggaaac	agtccagagg	ctgctccccg	2940
50	tgctgtgcca	agcgcatgga	ctcactcccc	agcaggtcgt	cgccatcgct	tccaataacg	3000
	gcggcaagca	ggccctggag	accgtccaga	gactgctccc	cgtgctgtgc	caagctcacg	3060
55	gactcacacc	tgagcaggtc	gtggcaatcg	cctctaacat	tgaggggaaa	caggccctgg	3120
	aaactgtaca	gcggctgctc	cccgtgctgt	gccaaagcaca	cggactcact	ccacagcagg	3180

60

65

ES 2 709 216 T3

	tcgtggccat	tgcaagtc	catgacggaggca	agcaggccct	ggaacagtg	cagcgccctgc	3240
5	tccctgtgct	gtgccaggct	catggtctga	ctcctcagca	ggtggtggcc	atcgcttcca	3300
	acaatggagg	gaagcaggcc	ctggagaccg	tacagagact	gctccccgtg	ctgtgccaag	3360
	cgcacggtct	gacccctcag	caggtcgtcg	caatcgccag	caatggcggg	ggcaagcagg	3420
10	ctctcgaaac	cgtccagcgg	ctcctcccag	tcctctgtca	ggctcacggc	ctgacccac	3480
	agcaggtcgt	cgtattgct	tctaattggc	gagggcgcc	tgctctggag	agcattgtgg	3540
	ctcagctgtc	caggcccgat	cctgccctgg	ctagatccgc	actcactaac	gatcatctgg	3600
15	tcgctctcgc	ttgcctcgg	ggacggcccg	ctctggacgc	agtcaaaaag	ggtctcccc	3660
	atgctcccgc	actgatcaag	agaaccaaca	ggagaattcc	tgagggatcc	gatcgtttaa	3720
	acaagctcgc	aaagtcaagc	cagtcgaaa	caaaggaaaa	actcagagaa	aaactcagaa	3780
20	acctgcccc	tgaatacctg	tccctcgtcg	acctggccta	cgattcaaa	cagaaccgcc	3840
	tctttgagat	gaaagtgatc	gaactgctca	cagaggaatg	cgggttccag	ggtctgcacc	3900
	tcggcggaag	caggagacca	gacggcgtcc	tgtacaccgc	cggactcaca	gacaactatg	3960
25	ggatcattct	ggataactaag	gcttacagct	ccggatattc	cctgccatt	gccagcgtg	4020
	acgagatgga	acggtacgtg	cgcgagaatc	agactagaga	tgaactggtc	aaccctaatac	4080
30	agtgggtgga	gaactttgaa	aatggcctgg	gaaccttcta	ttttctcttc	gtggctgggc	4140
	atttcaacgg	taatgtccag	gcacagctgg	agcgaatcag	taggaatacc	ggcgtgctgg	4200
	gagccgctgc	atctatcagt	cagctgctcc	tgctcgcaga	cgccattaga	gggggtcgga	4260
35	tggaatagaga	gagactgcgg	cacctcatgt	ttcagaacga	agagtttctg	ctggaacagg	4320
	agctgtgaac	gcgtaaatga	ttgcagatcc	actagttcta	gaattccagc	tgagcgccgg	4380
	tcgctaccat	taccagttgg	tctggtgtca	aaaataataa	taaccgggca	ggggggtct	4440
40	gcatggatct	ttgtgaagga	accttacttc	tgtggtgtga	cataattgga	caaactacct	4500
	acagagattt	aaagctctaa	ggtaaatata	aaatttttaa	gtgtataatg	tgttaaacta	4560
	ctgattctaa	ttgtttgtgt	attttagatt	ccaacctatg	gaactgatga	atgggagcag	4620
45	tggtggaatg	ccagatccag	acatgataag	atacattgat	gagtttgac	aaaccacaac	4680
	tagaatgcag	tgaaaaaaat	gctttatttg	tgaaatttgt	gatgctattg	ctttatttgt	4740
	aaccattata	agctgcaata	aacaagttaa	caacaacaat	tgcattcatt	ttatgtttca	4800
50	ggttcagggg	gaggtgtggg	aggtttttta	aagcaagtaa	aacctctaca	aatgtggtat	4860
	ggctgattat	gatctgcggc	cgccactggc	cgtcgtttta	caacgtcgtg	actgggaaaa	4920
55	ccctggcgtt	acccaactta	atcgccttgc	agcacatccc	cctttcgcca	gctggcgtaa	4980
	tagcgaagag	gcccgaccg	atcgccttcc	ccaacagttg	cgcagcctga	atggcgaatg	5040

60

65

ES 2 709 216 T3

	gaacgcgccc	tgtagcggcg	cattaagcgc	ggcgggtgtg	gtggttacgc	gcagcgtgac	5100
5	cgctacactt	gccagcgcgc	tagcgcgcgc	tcctttcgc	ttcttccctt	cctttctcgc	5160
	cacgttcgcc	ggctttcccc	gtcaagctct	aaatcggggg	ctccctttag	ggttcgcatt	5220
	tagtgcttta	cggcacctcg	acccccaaaa	acttgattag	ggtgatgggt	cacgtagtgg	5280
10	gccatcgccc	tgatagacgg	tttttcgccc	tttgacgttg	gagtcacagt	tctttaatag	5340
	tggactcctg	ttccaaactg	gaacaacact	caaccctatc	tcgggtctatt	cttttgattt	5400
	ataagggtt	ttgccgattt	cggcctattg	gttaaaaaat	gagctgattt	aacaaaaatt	5460
15	taacgcgaat	tttaacaaaa	tattaacgct	tacaatttag	gtggcacttt	tcggggaaat	5520
	gtgcgcggaa	cccctatttg	tttatttttc	taaatacatt	caaatatgta	tccgctcatg	5580
	agacaataac	cctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	5640
20	catttcogtg	tcgcccttat	tccctttttt	gcggcathtt	gccttcctgt	ttttgctcac	5700
	ccagaaacgc	tgggtgaaagt	aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	5760
	atcgaactgg	atctcaacag	cggtaagatc	cttgagagtt	ttcgccccga	agaacgtttt	5820
25	ccaatgatga	gcacttttaa	agttctgcta	tgtggcgcg	tattatcccg	tattgacgcc	5880
	gggcaagagc	aactcggtcg	ccgcatacac	tattctcaga	atgacttggt	tgagtactca	5940
	ccagtcacag	aaaagcatct	tacggatggc	atgacagtaa	gagaattatg	cagtgcgtgcc	6000
30	ataaccatga	gtgataaac	tgcgggcaac	ttacttctga	caacgatcgg	aggaccgaag	6060
	gagctaaccg	cttttttgca	caacatgggg	gatcatgtaa	ctcgccttga	tcgttgggaa	6120
	ccggagctga	atgaagccat	accaaacgac	gagcgtgaca	ccacgatgcc	tgtagcaatg	6180
35	gcaacaacgt	tcgcgcaaact	attaactggc	gaactactta	ctctagcttc	ccggcaacaa	6240
	ttaatagact	ggatggaggc	ggataaagtt	gcaggaccac	ttctgcgctc	ggcccttccg	6300
	gctggctggg	ttattgctga	taaatctgga	gccgggtgagc	gtgggtctcg	cggatatcatt	6360
40	gcagcactgg	ggccagatgg	taagccctcc	cgtatcgtag	ttatctacac	gacggggagt	6420
	caggcaacta	tggatgaacg	aaatagacag	atcgctgaga	taggtgcctc	actgattaag	6480
45	cattggtaac	tgtcagacca	agtttactca	tatatacttt	agattgatth	aaaacttcat	6540
	ttttaattta	aaaggatcta	ggtgaagatc	ctttttgata	atctcatgac	caaaatccct	6600
	taacgtgagt	tttcgttcca	ctgagcgtca	gaccccgtag	aaaagatcaa	aggatcttct	6660
50	tgagatcctt	tttttctgcg	cgtaaatctgc	tgcttgcaaa	caaaaaaacc	accgctacca	6720
	gcgggtgggtt	gtttgcccga	tcaagagcta	ccaactcttt	ttccgaagg	aactggcttc	6780
	agcagagcgc	agataccaaa	tactgtcctt	ctagtgtagc	cgtagttagg	ccaccacttc	6840
55	aagaactctg	tagcaccgcc	tacatacctc	gctctgctaa	tcctgttacc	agtggctgct	6900
	gccagtggcg	ataagtcgtg	tcttaccggg	ttggactcaa	gacgatagtt	accggataag	6960
60							
65							

ES 2 709 216 T3

	gcgcagcggg	cggtctgaac	gggggggttcg	tgcacacagc	ccagcttgga	gcgaacgacc	7020
5	tacaccgaac	tgagatacct	acagcgtgag	ctatgagaaa	gcgccacgct	tcccgaaggg	7080
	agaaaggcgg	acaggtatcc	ggtaagcggc	agggtcggaa	caggagagcg	cacgagggag	7140
	cttccagggg	gaaacgcctg	gtatctttat	agtcctgtcg	ggtttcgcca	cctctgactt	7200
10	gagcgtcgat	ttttgtgatg	ctcgtcaggg	gggcggagcc	tatggaaaaa	cgccagcaac	7260
	gcggcctttt	tacggttcct	ggccttttgc	tggccttttg	ctcacatgtt	ctttcctgcg	7320
	ttatcccctg	attctgtgga	taaccgtatt	accgcctttg	agtgagctga	taccgctcgc	7380
15	cgcagccgaa	cgaccgagcg	cagcgagtca	gtgagcgagg	aagcgggaaga	gcgccaata	7440
	cgcaaaccgc	ctctccccgc	gcgttgcccg	attcattaat	gcagctggca	cgacaggttt	7500
	cccgaactga	aagcgggcag	tgagcgcaac	gcaattaatg	tgagttagct	cactcattag	7560
20	gcaccccagg	ctttacactt	tatgcttccg	gctcgtatgt	tgtgtggaat	tgtgagcgga	7620
	taacaatttc	acacaggaaa	cagctatgac	catgaggcgc	gccggattc		7669
25	<210> 29						
	<211> 7663						
	<212> ADN						
	<213> Secuencia artificial						
30	<220>						
	<223> pCAG-ArtTall-Clo051						
	<400> 29						
35	gacattgatt	attgactagt	tattaatagt	aatcaattac	ggggtcatta	gttcatagcc	60
	catatatgga	gttccgcgtt	acataactta	cggtaaatgg	cccgcctggc	tgaccgcca	120
	acgacccccg	cccattgacg	tcaataatga	cgatgtgtcc	catagtaacg	ccaataggga	180
40	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	tgcccacttg	gcagtacatc	240
	aagtgtatca	tatgccaaat	acgcccccta	ttgacgtcaa	tgacggtaaa	tggcccgccct	300
	ggcatttatgc	ccagtacatg	accttatggg	actttcctac	ttggcagtac	atctacgtat	360
45	tagtcatcgc	tattaccatg	gtcagaggtga	gccccacgtt	ctgcttcaact	ctccccatct	420
	ccccccccctc	cccacccccca	attttgtatt	tattttatttt	ttaattattt	tgtgcagcga	480
	tggggggcggg	ggggggggggg	gggcgcgcgc	caggcggggc	ggggcggggc	gaggggcggg	540
50	gcggggcgag	gcggagaggt	gcggcggcag	ccaatcagag	cggcgcgctc	cgaaagtttc	600
	cttttatggc	gaggcggcgg	cggcggcggc	cctataaaaa	gcgaagcgcg	cggcgggcgg	660
	gagtcgctgc	gcgctgcctt	cgccccgtgc	cccgcctcgc	cgccgcctcg	cgccgcctcg	720
55	cccggtctctg	actgaccgcg	ttactcccac	aggtgagcgg	gcgggacggc	ccttctcctc	780
	cgggctgtaa	ttagcgcttg	gtttaatgac	ggcttggttc	ttttctgtgg	ctgcgtgaaa	840

ES 2 709 216 T3

5	gccttgaggg gctccgggag ggcctttgt gcggggggga gcggctcggg ggggtgcgtgc	900
	gtgtgtgtgt gcgtggggag cgcgcgtgc ggctccgcgc tgcccggcgg ctgtgagcgc	960
	tgcgggcgcg gcgcggggct ttgtgcgtc cgcagtgtgc gcgaggggag cgcggccggg	1020
10	ggcggtgccc cgcgggtgcg ggggggctgc gaggggaaca aaggctgcgt gcggggtgtg	1080
	tgctggggg ggtgagcagg ggggtgtggc gcgtcggctg ggctgcaacc cccctgcac	1140
	ccccctccc gagttgctga gcacggcccg gcttcgggtg cggggctccg tacggggcgt	1200
15	ggcgcggggc tcgcctgcc gggcggggg tggcggcagg tgggggtgcc gggcggggcg	1260
	gggcccctc gggccgggga gggctcggg gaggggcgcg gcggccccc gagcgcggc	1320
	ggctgtcag gcgcggcgag ccgcagccat tgccttttat ggtaatcgtg cgagaggcg	1380
20	cagggacttc ctttgtcca aatctgtgc gagccgaat ctgggaggcg ccgccgcacc	1440
	ccctctagc ggcgcggggc gaagcgtgc ggcgcggca ggaaggaaat gggcggggag	1500
	ggccttcgtg cgtgcgcgcg ccgcctccc cttctccctc tccagcctcg gggctgtccg	1560
25	cggggggacg gctgccttcg ggggggacgg ggcagggcgg ggttcggctt ctggcgtgtg	1620
	accggcggct cttaggcctc tgctaaccat gttcatgcct tcttcttttt cctacagatc	1680
	cttaattaat aatacgactc actatagggg ccgccaccat gggacctaa aaaaagagga	1740
30	aggtggcggc cgctgactac aaggatgacg acgataaacc aggtggcgga ggtagtggcg	1800
	gaggtggggt acccgccagt ccagcagccc aggtggatct gagaaccctc ggctacagcc	1860
	agcagcagca ggagaagatc aaaccaaagg tgcggtccac cgtcgtcag caccatgaag	1920
35	cactggtggg gcacggtttc acacacgccc atattgtggc tctgtctcag catcccgtg	1980
	cactcgggac tgtggccgtc aaatatcagg acatgatcgc cgtctgcct gaggcaaccc	2040
	acgaagccat tgtgggcgtc ggaaagcagt ggagcggcgc cagagcactc gaagcactcc	2100
40	tcaccgtcgc cggggaactg cgggggccac cactccagtc cggactggac actggacagc	2160
	tgctgaagat cgctaaacgc ggcggagtga cagctgtgga agctgtgcac gcttgaggga	2220
	atgctctgac aggagcccca ctgaatctta ctccagaaca ggtcgtcgca atcgcaagta	2280
45	acatcggcgg aaaacaggcc ctcgaaaccg tccagagact cctccccgtg ctgtgccagg	2340
	cccacggact gacccacag cagggtgctg ccacgcctag caacggcgga gggaaagcagg	2400
	ctctggagac cgtgcagagg ctgctccccg tcctgtgcca ggcacatggg ctcacacctc	2460
50	agcaggtggt cgcaattgcc tccaatggtg gcggaaca ggcctggaa actgtgcaga	2520
	gactgctccc cgtgctgtgc caggctcacg gtctcacacc ccagcaggtg gtcgctatcg	2580
	catctcatga cgggggcaag caggcactgg agacagtgc cggcgtgctc cctgtcctgt	2640
55	gccagggcca cggactcact cctcagcagg tcgtcgccat tgctagtaac ggcggaggga	2700
	aacaggctct ggaaaccgtg cagcgcctgc tccccgtgct gtgccaagcc cacggcctga	2760
60		
65		

ES 2 709 216 T3

	ccccccagca ggtggtcgca atcgccctcaa acaatggtgg caagcaggcc ctggagactg	2820
5	tgcagcgact gctcccagtg ctgtgccagg cccatggact cacaccacag caggtcgctcg	2880
	ctattgcaag caacaatgga gggaaacagg cactggaaac agtccagagg ctgctccccg	2940
	tgctgtgcca agcgcatgga ctcaactccc agcaggtcgt cgccatcgct tccaataacg	3000
10	gcggaagca ggcctggag accgtccaga gactgctccc cgtgctgtgc caagctcacg	3060
	gactcacacc tgagcaggtc gtggcaatcg cctctaacat tggagggaaa caggccctgg	3120
	aaactgtaca gcggtgctc cccgtgctgt gccaaagcaca cggactcact ccacagcagg	3180
15	tcgtggccat tgcaagtcac gacggaggca agcaggccct ggaaacagtg cagcgccctgc	3240
	tccctgtgct gtgccaggct catggtctga ctctcagca ggtggtggcc atcgcttcca	3300
20	acaatggagg gaagcaggcc ctggagaccg tacagagact gctccccgtg ctgtgccaaag	3360
	cgcacggtct gaccctcag caggtcgctc caatcgccag caatggcggg ggcaagcagg	3420
	ctctcgaaac cgtccagcgg ctccctccag tcctctgtca ggctcacggc ctgacccac	3480
25	agcaggtcgt cgctattgct tctaattggc gagggcggcc tgctctggag agcattgtgg	3540
	ctcagctgtc caggcccgat cctgccctgg ctagatccgc actcactaac gatcatctgg	3600
	tcgctctcgc ttgcctcggg ggacggcccg ctctggacgc agtcaaaaag ggtctcccc	3660
30	atgctccgc actgatcaag agaaccaaca ggagaattcc tgagggatcc gatcgtttaa	3720
	acgaaggcat caaaagcaac atctccctcc tgaaagacga actccggggg cagattagcc	3780
	acattagtca cgaatacctc tccctcatcg acctggcttt cgatagcaag cagaacaggc	3840
35	tctttgagat gaaagtgtc gaactgctc tcaatgagta cgggttcaag ggtcgacacc	3900
	tcggcggatc taggaaacca gacggcatcg tgtatagtac cacactggaa gacaactttg	3960
40	ggatcattgt ggataccaag gcatactctg agggttatag tctgcccatt tcacaggccg	4020
	acgagatgga acggtacgtg cgcgagaact caaatagaga tgaggaaagtc aaccctaaca	4080
	agtggtgga gaacttctct gaggaagtga agaaatacta cttcgtcttt atcagcgggt	4140
45	ccttcaaggg taaatttgag gaacagctca ggagactgag catgactacc ggcgtgaatg	4200
	gcagcgccgt caacgtggtc aatctgctcc tgggcgctga aaagattcgg agcggagaga	4260
	tgaccatcga agagctggag agggcaatgt ttaataatag cgagtttatc ctgaaatact	4320
50	gaacgcgtaa atgattgcag atccactagt tctagaattc cagctgagcg ccggtcgcta	4380
	ccattaccag ttggtctggt gtcaaaaata ataataaccg ggcagggggg atctgcatgg	4440
55	atctttgtga aggaacctta cttctgtggt gtgacataat tggacaaact acctacagag	4500
	atttaaagct ctaaggtaaa tataaaatth ttaagtgtat aatgtgttaa actactgatt	4560
60	ctaattgttt gtgtatttta gattccaacc tatggaactg atgaatggga gcagtgggtg	4620
65		

ES 2 709 216 T3

	aatgccagat ccagacatga taagatacat tgatgagttt ggacaaacca caactagaat	4680
5	gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat	4740
	tataagctgc aataaacaag ttaacaacaa caattgcatt cattttatgt ttcagggttca	4800
	gggggaggtg tgggaggttt tttaaagcaa gtaaaacctc taaaaatgtg gtatggctga	4860
10	ttatgatctg cggccgccac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg	4920
	cgttacccaa cttaatcgcc ttgcagcaca tcccccttc gccagctggc gtaatagcga	4980
	agaggcccg cccgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggaacgc	5040
15	gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac	5100
	acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt	5160
	cgccggcttt ccccgctcaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc	5220
20	tttacggcac ctcgacccca aaaaacttga ttaggggtgat gggtcacgta gtgggccatc	5280
	gccctgatag acggtttttc gcccttgac gttggagtcc acgttcttta atagtggact	5340
	cttgttccaa actggaacaa cactcaacct tatctcggtc tattcttttg atttataagg	5400
25	gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc	5460
	gaattttaac aaaatattaa cgcttacaat ttaggtggca cttttcgggg aaatgtgcgc	5520
	ggaacccta tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa	5580
30	taaccctgat aaatgcttca ataatttga aaaaggaaga gtatgagtat tcaacatttc	5640
	cgtgtcgccc ttattccctt ttttgcgga ttttgccctc ctgtttttgc tcaccagaa	5700
35	acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa	5760
	ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg	5820
	atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa	5880
40	gagcaactcg gtcgcgcgat acactattct cagaatgact tggttgagta ctcaccagtc	5940
	acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc	6000
	atgagtgata aactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta	6060
45	accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag	6120
	ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca	6180
	acgttgcgca aactattaac tggcgaacta cttactctag cttcccgga acaattaata	6240
50	gactggatgg aggcggataa agttgcagga ccacttctgc gtcggccct tccggctggc	6300
	tggtttattg ctgataaatc tggagccggt gagcgtgggt ctcgcggtat cattgcagca	6360
	ctggggccag atgtaagcc ctcccgatc gtagttatct acacgacggg gagtcaggca	6420
55	actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg	6480
	taactgtcag accaagttta ctcatatata ctttagattg atttaaaaact tcatttttaa	6540
60		
65		

ES 2 709 216 T3

	tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaaat cccttaacgt	6600
5	gagtttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat	6660
	ccttttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcgggtg	6720
	gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga	6780
10	gcgagatac caaatactgt ccttctagt tagccgtagt taggccacca cttcaagaac	6840
	tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt	6900
	ggcgataagt cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgag	6960
15	cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc	7020
	gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag	7080
	gcggacaggt atccgtaag cggcagggc ggaacaggag agcgacagag ggagcttcca	7140
20	gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt	7200
	cgatttttgt gatgctcgtc aggggggagg agcctatgga aaaacgccag caacgcggcc	7260
25	tttttacggt tcttgccctt ttgctggcct tttgctcaca tggtctttcc tgcgttatcc	7320
	cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc	7380
	cgaacgaccg agcgacagca gtcagttagc gaggaagcgg aagagcgccc aatacgcaaa	7440
30	ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac	7500
	tggaaagcgg gcagttagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc	7560
	caggctttac actttatgct tccggtcgt atgtgtgtg gaattgtgag cggataacaa	7620
35	tttcacacag gaaacagcta tgaccatgag gcgcgccgga ttc	7663

<210> 30
 <211> 7816
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> pCAG-ArtTail-Myl
 <400> 30

50	gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
	catatatgga gttccgcgtt acataactta cggtaaattg cccgcctggc tgaccgcca	120
	acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagggg	180
55	ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
	aagtgtatca tatgccaaat acgcccccta ttgacgtcaa tgacggtaaa tggccccgct	300
	ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtag atctacgtat	360
60	tagtcatcgc tattaccatg gtcgaggtga gcccacggt ctgcttact ctccccatct	420

65

ES 2 709 216 T3

	ccccccctc	cccaccccca	atthttgtatt	tattttatthtt	ttaattatthtt	tgtgcagcga	480
5	tgggggcggg	gggggggggg	gggcgcgcgc	caggcggggc	ggggcggggc	gaggggcggg	540
	gcggggcgag	gcggagaggt	gcggcggcag	ccaatcagag	cggcgcgctc	cgaaggtttc	600
	cttttatggc	gagggcgcg	cggcgcgcg	cctataaaaa	gcgaagcgcg	cggcgggcg	660
10	gagtcgctgc	gcgctgcctt	cgccccgtgc	cccgctccgc	cgcgcgctcg	cgcgcggcg	720
	cccgctctcg	actgaccgcg	ttactcccac	aggtgagcgg	gcgggacggc	ccttctcctc	780
	cgggctgtaa	ttagcgcttg	gtttaatgac	ggcttgcttc	ttttctgtgg	ctgcgtgaaa	840
15	gccttgaggg	gctccgggag	ggccctttgt	gcggggggga	gcggctcggg	gggtgcgtgc	900
	gtgtgtgtgt	gcgtggggag	cgcgcgctgc	ggctccgcgc	tgcccgcgcg	ctgtgagcgc	960
	tgccggcgcg	gcgcggggct	ttgtgcgctc	cgcagtgctc	gcgaggggag	cgcggccggg	1020
20	ggcggtgccc	cgcgtgcggg	ggggggctgc	gaggggaaca	aaggctgcgt	gcggggtgtg	1080
	tgcggtgggg	ggtgagcagg	gggtgtgggc	gcgtcggtcg	ggctgcaacc	ccccctgcac	1140
	ccccctcccc	gagttgctga	gcacggcccc	gcttcgggtg	cggggctccg	tacggggcgt	1200
25	ggcgcggggc	tcgccgtgcc	ggcggggggg	tgccggcagg	tgggggtgcc	ggcgggggcg	1260
	gggcgcctc	gggcggggga	gggctcgggg	gaggggcgcg	gcggcccccg	gagcgccggc	1320
	ggctgtcgag	gcgcggcgag	ccgcagccat	tgccctttat	ggtaatcgtg	cagagggcg	1380
30	cagggacttc	ctttgtccca	aatctgtgcg	gagccgaaat	ctgggagcg	ccgccgcacc	1440
	ccctctagcg	ggcgcggggc	gaagcggtgc	ggcgccggca	ggaaggaaat	ggcgggggag	1500
35	ggccttcgtg	cgtcgccgcg	ccgcgctccc	cttctccctc	tccagcctcg	gggtgtccg	1560
	cggggggacg	gctgccttcg	ggggggacgg	ggcagggcgg	ggttcggctt	ctggcgtgtg	1620
	accgcgggct	ctagagcctc	tgctaaccat	gttcatgcct	tcttcttttt	cctacagatc	1680
40	cttaattaat	aatacgactc	actatagggg	ccgccacccat	gggacctaat	aaaaagagga	1740
	aggtggcggc	cgtcgactac	aaggatgacg	acgataaacc	aggtggcgga	gtagtgccg	1800
	gaggtgggg	acccgccagt	ccagcagccc	aggtggatct	gagaaccctc	ggctacagcc	1860
45	agcagcagca	ggagaagatc	aaaccaaagg	tgccgtccac	cgtcgctcag	caccatgaag	1920
	cactgggtgg	gcacggtttc	acacacgccc	atattgtggc	tctgtctcag	catcccgcgt	1980
	cactcgggac	tgtggccgct	aaatatcagg	acatgatcgc	cgtctgcctc	gaggcaaccc	2040
50	acgaagccat	tgtgggcgct	ggaaagcagt	ggagcggtgc	cagagcactc	gaagcactcc	2100
	tcaccgtcgc	cgggggaactg	cgggggtccac	cactccagtc	cggactggac	actggacagc	2160
	tgctgaagat	cgtctaaacg	ggcgagtgta	cagctgtgga	agctgtgcac	gcttgaggga	2220
55	atgctctgac	aggagcccca	ctgaatctta	ctccagaaca	ggctgtcgca	atcgcaagta	2280
	acatcggcgg	aaaacaggcc	ctcgaaaccg	tccagagact	cctccccgtg	ctgtgccagg	2340

ES 2 709 216 T3

	cccacggact gacccacag caggtggtcg ccatcgctag caacggcgga gggaagcagg	2400
5	ctctggagac cgtgcagagg ctgctccccg tcctgtgcca ggcacatggg ctcacacctc	2460
	agcaggtggt cgcaattgcc tccaatggtg gcggaaaaca ggccctggaa actgtgcaga	2520
	gactgctccc cgtgctgtgc caggctcacg gtctcacacc ccagcaggtg gtcgctatcg	2580
10	catctcatga cgggggcaag caggcactgg agacagtgca gcggtgctc cctgtcctgt	2640
	gccaggccca cggactcact cctcagcagg tcgtcgccat tgctagtaac ggcgaggga	2700
	aacaggctct ggaaaccgtg cagcgctgc tccccgtgct gtgccaagcc caccgcctga	2760
15	ccccccagca ggtggtcgca atcgctcaa acaatggtgg caagcaggcc ctggagactg	2820
	tgcagcgact gctcccagtg ctgtgccagg cccatggact cacaccacag caggtcgctg	2880
20	ctattgcaag caacaatgga gggaaacagg cactggaaac agtccagagg ctgctccccg	2940
	tgctgtgcca agcgcagga ctactcccc agcaggtcgt cgccatcgct tccaataacg	3000
	gcggcaagca ggccctggag accgtccaga gactgctccc cgtgctgtgc caagctcacg	3060
25	gactcacacc tgagcaggtc gtggcaatcg cctctaactat tggagggaaa caggccctgg	3120
	aaactgtaca gcggtgctc cccgtgctgt gccaaagcaca cggactcact ccacagcagg	3180
	tcgtggccat tgcaagtcag gacggaggca agcaggccct ggaaacagtg cagcgctgc	3240
30	tccctgtgct gtgccaggct catggtctga ctctcagca ggtggtggcc atcgcttcca	3300
	acaatggagg gaagcaggcc ctggagaccg tacagagact gctccccgtg ctgtgccaa	3360
	cgcacggtct gacccctcag caggtcgctg caatcgccag caatggcggg ggcaagcagg	3420
35	ctctcgaaac cgtccagcgg ctctctccag tcctctgtca ggctcacggc ctgacccac	3480
	agcaggtcgt cgctattgct tctaattggc gagggcgcc tgctctggag agcattgtgg	3540
	ctcagctgtc caggcccgat cctgccctgg ctagatccgc actcactaac gatcatctgg	3600
40	tcgctctgc ttgcctcggg ggacggcccg ctctggacgc agtcaaaaag ggtctcccc	3660
	atgctccgc actgatcaag agaaccaaca ggagaattcc tgagggatcc gatcgtttaa	3720
45	acatcaatag caagatcaag cagctggacg atagcatcaa cgtggagtcc ctgaagattg	3780
	acgatccaa agatctgctg aatgacctgg agatccagcg gaaggctaaa accattgaag	3840
	atacagtga ccacctgaag ctgcgtccg acatcgagga tattctggac gtgttcgcca	3900
50	aaatcaagaa aagggatgtg cccgacgtgc ctctgttcct ggagtggaat atctggcggg	3960
	cctttgccgc tctgaatcat acccaggcta tcgaaggga ctttattgtg gacctggatg	4020
	gcatgcccc gaatacagct ccaggaaaga aaccgatat cgagattaac tacggaagct	4080
55	tctcctgcat cgtggaagt actatgagct ccggggagac ccagttaac atggaaggct	4140
	ctagtgtgcc taggcactac ggagacctgg tgagaaaggt ggaccatgat gcctattgta	4200
60		
65		

ES 2 709 216 T3

	tcttcattgc ccctaaggtg gctccagga ctaaagctca cttctttaac ctgaataggc	4260
5	tgtctacaaa gcattatggc ggaaagacta agatcattcc aatgagtctg gacgatttca	4320
	tctgctttct gcaagtgggc attaccacaca actttcagga tatcaacaag ctgaaaaatt	4380
	ggctggacaa cctgattaac ttcaatctgg agtctgaaga cgaggaaatc tggtttgagg	4440
10	aaatcatttc taagatcagt acatgggcca tttgaacgcg taaatgattg cagatccact	4500
	agtcttagaa ttccagctga gcgcccgtcg ctaccattac cagttggtct ggtgtcaaaa	4560
15	ataataataa ccgggcaggg gggatctgca tggatctttg tgaaggaacc ttacttctgt	4620
	ggtgtgacat aattggacaa actacctaca gagatttaaa gctctaaggc aaatataaaa	4680
	tttttaagtg tataatgtgt taaactactg attctaattg tttgtgtatt ttagattcca	4740
20	acctatggaa ctgatgaatg ggagcagtgg tggaatgcca gatccagaca tgataagata	4800
	cattgatgag tttggacaaa ccacaactag aatgcagtga aaaaaatgct ttatttgtga	4860
	aatttgtgat gctattgctt tatttctaac cattataagc tgcaataaac aagttaacaa	4920
25	caacaattgc attcatttta tgtttcaggt tcagggggag gtgtgggagg ttttttaag	4980
	caagtaaac ctctacaaat gtggtatggc tgattatgat ctgcggccgc cactggccgt	5040
	cgttttacaa cgctcgtgact gggaaaaccc tggcggtacc caacttaatc gccttgacgc	5100
30	acatccccct ttgccagct ggcgtaatag cgaagaggcc cgcaccgacg gcccttccca	5160
	acagttgcgc agcctgaatg gcgaatggaa cgcgccctgt agcggcgcat taagcgcggc	5220
35	gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcggccctag cgcccgctcc	5280
	tttcgctttc ttcccttcct ttctcgccac gttcgccggc tttccccgct aagctctaaa	5340
	tcgggggctc ccttttaggt tccgatttag tgccttacgg cacctcgacc ccaaaaaact	5400
40	tgattagggc gatggttcac gtagtgggcc atcgccctga tagacggttt ttcgcccttt	5460
	gacgttggag tccacgttct ttaatagtgg actcttgctt caaactggaa caacactcaa	5520
	ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg cctattgggt	5580
45	aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt aacaaaatat taacgcttac	5640
	aatttaggtg gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa	5700
	atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat	5760
50	tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc cttttttgcg	5820
	gcattttgcc ttctgtttt tgctcaccga gaaacgctgg tgaaagtaaa agatgctgaa	5880
55	gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt	5940
	gagagttttc gccccgaaga acgttttcca atgatgagca cttttaaggt tctgctatgt	6000
	ggcgcggtat tatcccgat tgacgcggg caagagcaac tcggtcgccg catacactat	6060
60	tctcagaatg acttggttga gtactacca gtcacagaaa agcatcttac ggatggcatg	6120

ES 2 709 216 T3

	acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta	6180
5	cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgacaaa catgggggat	6240
	catgtaactc gccttgatcg ttggaaccg gagctgaatg aagccatacc aaacgacgag	6300
	cgtgacacca cgatgcctgt agcaatggca acaacgttgc gcaaactatt aactggcgaa	6360
10	ctacttactc tagcttcccc gcaacaatta atagactgga tggaggcgga taaagttgca	6420
	ggaccacttc tgcgctcggc ccttccggct ggctgggtta ttgctgataa atctggagcc	6480
15	ggtgagcgtg ggtctcgcgg tatcattgca gcaactgggc cagatggtaa gccctccgct	6540
	atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc	6600
	gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat	6660
20	atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt	6720
	tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac	6780
	cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc	6840
25	ttgcaaacia aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca	6900
	actctttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgccttcta	6960
	gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct	7020
30	ctgctaatac tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg	7080
	gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc	7140
35	acacagccca gcttgagcgg aacgacctac accgaactga gatacctaca gcgtgagcta	7200
	tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg	7260
	gtcggaaacag gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt	7320
40	cctgtcgggt ttccgccact ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg	7380
	cggagcctat ggaaaaacgc cagcaacgcg gcctttttac gggtcctggc cttttgctgg	7440
45	ccttttgctc acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc	7500
	gcctttgagt gagctgatac cgctcgccgc agccgaacga ccgagcgag cgagtcagtg	7560
	agcgaggaag cgggaagagcg cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt	7620
50	cattaatgca gctggcacga caggtttccc gactggaaaag cgggcagtgga gcgcaacgca	7680
	attaatgtga gttagctcac tcattaggca cccaggctt tacactttat gcttccggct	7740
	cgtatgttgt gtggaattgt gagcggataa caatttcaca caggaaacag ctatgaccat	7800
55	gaggcgcgcc ggattc	7816

<210> 31

<211> 7705

<212> ADN

60 <213> Secuencia artificial

<220>

<223> pCAG-ArtTal1-Pept071

65 <400> 31

ES 2 709 216 T3

	gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
5	catatatgga gttccgcgtt acataactta cggtaaattg cccgcctggc tgaccgcccc	120
	acgacccccg ccatttgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	180
	ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
10	aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgctt	300
	ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat	360
15	tagtcatcgc tattaccatg gtcgaggtga gcccacggt ctgcttcaact ctccccatct	420
	cccccccctc cccacccccca attttgtatt tattttatttt ttaattatttt tgtgcagcga	480
	tggggggcggg gggggggggg gggcgcgcg caggcggggc gggcgggggc gaggggcggg	540
20	gcggggcgag gcggagaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc	600
	cttttatggc gaggcggcgg cggcgggcgg cctataaaaa gcgaagcgcg cggcgggcgg	660
	gagtcgctgc gcgctgcctt cggcccgctc cccgctccgc cggcgctcgc cggcgccgc	720
25	cccggtctctg actgaccgcg ttactccac aggtgagcgg gcgggacggc ccttctcctc	780
	cgggctgtaa ttagcgcttg gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa	840
	gccttgaggg gctccgggag ggccctttgt gcggggggga gcggctcggg gggcgctgc	900
30	gtgtgtgtgt gcgtggggag cggcgctgc ggctccgcgc tgcccgggcg ctgtgagcgc	960
	tgcgggcgcg gcgcggggct ttgtgcgctc cgcagtgtgc gcgagggag cgcggccggg	1020
35	ggcgggtgcc cgcgggtgcg ggggggctgc gaggggaaca aaggctgcgt gcgggggtgtg	1080
	tgcgtggggg ggtgagcagg ggggtgtggc gcgtcggctg ggctgcaacc cccctgcac	1140
	ccccctcccc gagttgctga gcacggcccg gcttcgggtg cggggctccg tacggggcgt	1200
40	ggcgcggggc tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcggggcg	1260
	gggcccctc gggccgggga gggctcgggg gaggggcgcg gcggcccccg gagcgccggc	1320
	ggctgtcgag gcgcggcgag ccgcagccat tgccctttat ggtaatcgtg cgagagggcg	1380
45	cagggacttc ctttgtccca aatctgtgcg gagccgaaat ctgggagggc ccgccgcacc	1440
	ccctctagcg ggcgcggggc gaagcgggtgc ggcgcgggca ggaaggaaat gggcggggag	1500
50	ggccttcgtg cgtcgccgcg ccgcgctccc cttctccctc tccagcctcg gggctgtccg	1560
	cggggggacg gctgccttcg ggggggacgg ggcagggcg ggttcggctt ctggcgtgtg	1620
	accggcggt ctagagcctc tgctaaccat gttcatgcct tcttcttttt cctacagatc	1680
55	cttaattaat aatacgactc actatagggg ccgccaccat gggacctaag aaaaagagga	1740
60		
65		

ES 2 709 216 T3

	aggtggcggc	cgctgactac	aaggatgacg	acgataaacc	aggtggcggg	ggtagtggcg	1800
5	gaggtggggg	acccgccagt	ccagcagccc	aggtggatct	gagaaccctc	ggctacagcc	1860
	agcagcagca	ggagaagatc	aaaccaaagg	tgcgggtccac	cgctcgctcag	cacccatgaag	1920
	cactgggtgg	gcacgggttc	acacacgccc	atattgtggc	tctgtctcag	catcccgtg	1980
10	cactcgggac	tgtggccgtc	aaatatcagg	acatgatcgc	cgctctgcct	gaggcaaccc	2040
	acgaagccat	tgtgggctgc	ggaaagcagt	ggagcgggtc	cagagcactc	gaagcactcc	2100
15	tcaccgtcgc	cggggaactg	cggggtccac	cactccagtc	cggactggac	actggacagc	2160
	tgctgaagat	cgctaaacgc	ggcggagtga	cagctgtgga	agctgtgcac	gcttgaggga	2220
	atgctctgac	aggagcccca	ctgaatctta	ctccagaaca	ggtcgtcgca	atcgcaagta	2280
20	acatcggcgg	aaaacaggcc	ctcgaaccg	tccagagact	cctccccgtg	ctgtgccagg	2340
	cccacggact	gacccacag	caggtgggtc	ccatcgctag	caacggcgga	gggaagcagg	2400
	ctctggagac	cgctcagagg	ctgctccccg	tcctgtgcca	ggcacatggg	ctcacacctc	2460
25	agcaggtggt	cgcaattgcc	tcgaatggtg	gcggaaaaca	ggccctggaa	actgtgcaga	2520
	gactgctccc	cgctgtgtgc	caggctcacg	gtctcacacc	ccagcaggtg	gtcgctatcg	2580
	catctcatga	cgggggcaag	caggcactgg	agacagtgc	gcggctgctc	cctgtcctgt	2640
30	gccaggccca	cggactcact	cctcagcagg	tcgtcgccat	tgctagtaac	ggcggaggga	2700
	aacaggctct	ggaaccgtg	cagcgctgc	tccccgtgct	gtgccaagcc	cacggcctga	2760
35	ccccccagca	ggtggtcgca	atcgctcaa	acaatggtgg	caagcaggcc	ctggagactg	2820
	tgcagcgact	gtccccagt	ctgtgccagg	cccatggact	cacaccacag	caggtcgctg	2880
	ctattgcaag	caacaatgga	gggaaacagg	cactggaaac	agtccagagg	ctgctccccg	2940
40	tgctgtgcca	agcgcatgga	ctcactcccc	agcaggtcgt	cgccatcgct	tccaataacg	3000
	gcggcaagca	ggccctggag	accgtccaga	gactgctccc	cgctgtgtgc	caagctcacg	3060
	gactcacacc	tgagcaggtc	gtggcaatcg	cctctaacat	tggagggaaa	caggccctgg	3120
45	aaactgtaca	gcggctgctc	cccgtgctgt	gccaagcaca	cggactcact	ccacagcagg	3180
	tcgtggccat	tgcaagtc	gacggaggca	agcaggccct	ggaacagtg	cagcgctgc	3240
	tcctgtgct	gtgccaggct	catggtctga	ctcctcagca	ggtggtggcc	atcgcttcca	3300
50	acaatggagg	gaagcaggcc	ctggagaccg	tacagagact	gctccccgtg	ctgtgccaag	3360
	cgcacggtct	gacccctcag	caggtcgctg	caatcgccag	caatggcggg	ggcaagcagg	3420
	ctctcgaaac	cgccagcgg	ctcctcccag	tcctctgtca	ggctcacggc	ctgacccac	3480
55	agcaggtcgt	cgctattgct	tctaattggc	gagggcggcc	tgctctggag	agcattgtgg	3540
	ctcagctgtc	caggcccgat	cctgccctgg	ctagatccgc	actcactaac	gatcatctgg	3600
60							
65							

ES 2 709 216 T3

	tcgctctcgc ttgcctcggg ggacggcccg ctctggacgc agtcaaaaag ggtctcccc	3660
5	atgctccccg actgatcaag agaaccaaca ggagaattcc tgagggatcc gatcgttta	3720
	acaagatcag caaaaccaat gtgctggagc tcaaggacaa agtccgagat aagctgaaat	3780
10	acgtggacca caggtatctg gcactcatcg acctcgcta tgatgggacc gctaacaggg	3840
	acttcgaaat ccagacaatt gatctgctca ttaatgagct gaagtttaaa ggggtcaggc	3900
	tcggtgaaag tagaaagccc gacggcatca ttccatacaa catcaatgga gtgatcattg	3960
15	ataacaagggc ttactctact gggtataacc tgcctattaa tcaggccgac gagatgatcc	4020
	ggatatattga ggaaaatcag acccgcatg aaaaaatcaa ctccaataag tgggtgggagt	4080
	ctttcgacga taaggtcaaa gacttcaact acctgtttgt gagctccttc ttttaagggga	4140
20	actttaaaaa caatctgaag catatcgcta acagaacagg tgtcagcggc ggagcaatta	4200
	acgtggagaa tctgctctac ttgcgagagg aactgaaagc cgcccggtc tcatatgtgg	4260
	atagctttaa gatgtacgac aacgatgaga tctatgtcgg cgacttctct gattacagtt	4320
25	atgtgaagtt tgccgctgag gaagagggag aatacctgac ttgaacgcgt aaatgattgc	4380
	agatccacta gttctagaat tccagctgag cgccggtcgc taccattacc agttggtctg	4440
30	gtgtcaaaaa taataataac cgggcagggg ggatctgcat ggatctttgt gaaggaacct	4500
	tacttctgtg gtgtgacata attggacaaa ctacctacag agatttaaaag ctctaaggta	4560
	aataaaaaat ttttaagtgt ataatgtgtt aaactactga ttctaattgt ttgtgtattt	4620
35	tagattccaa cctatggaac tgatgaatgg gagcagtggt ggaatgccag atccagacat	4680
	gataagatac attgatgagt ttggacaaac cacaactaga atgcagtgaa aaaaatgctt	4740
	tatttgtgaa atttgtgatg ctattgcttt atttgaacc attataagct gcaataaaca	4800
40	agttaacaac aacaattgca ttcatthttat gtttcagggt cagggggagg tgtgggagg	4860
	tttttaagc aagtaaaacc tctacaaatg tggatggct gattatgatc tgcggccgcc	4920
	actggccgct gttttacaac gtcgtgactg ggaaaaccct ggcgttacct aacttaatcg	4980
45	ccttgcagca catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg	5040
	cccttcccaa cagttgcgca gcctgaatgg cgaatggaac gcgccctgta gcggcgcat	5100
50	aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc	5160
	gcccgcctct ttcgctttct tcccttcctt tctcgccacg ttcgccggct tccccgtca	5220
	agctctaaat cgggggctcc ctttaggggt ccgatttagt gctttacggc acctcgacct	5280
55	caaaaaactt gattaggggt atggttcacg tagtgggcca tcgccctgat agacggtttt	5340
	tcgccctttg acgttggagt ccacgttctt taatagtgga ctcttggtcc aaactggaac	5400
	aactactcaac cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc	5460
60	ctattggtta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt	5520

65

ES 2 709 216 T3

	aacgcttaca atttaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta	5580
5	tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt	5640
	caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc	5700
10	ttttttgogg cattttgcct tcctgttttt gctcaccag aaacgctggt gaaagtaaaa	5760
	gatgctgaag atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggg	5820
	aagatccttg agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt	5880
15	ctgctatgtg gcgcggtatt atcccgatt gacgcggggc aagagcaact cggtcgccc	5940
	atacactatt ctcagaatga cttggttgag tactcaccag tcacagaaaa gcactctacg	6000
	gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga taacactcgg	6060
20	gccaaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt tttgcacaac	6120
	atgggggatc atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca	6180
	aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta	6240
25	actggcgaac tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat	6300
	aaagttgcag gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa	6360
30	tctggagccg gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtgag	6420
	ccctcccgtg tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat	6480
	agacagatcg ctgagatagg tgccctactg attaagcatt ggtaactgtc agaccaagtt	6540
35	tactcatata tacttttagat tgatttaaaa cttcattttt aatttaaaaag gatctaggtg	6600
	aagatccttt ttgataatct catgaccaa atcccttaac gtgagttttc gttccactga	6660
	gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta	6720
40	atctgctgct tgcaaaaaaaa aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa	6780
	gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact	6840
45	gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca	6900
	tacctcgctc tgctaatacct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt	6960
	accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg	7020
50	ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag	7080
	cgtgagctat gagaaagcgc cacgcttccc gaaggagaa aggcggacag gtatccggta	7140
	agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat	7200
55	ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg	7260
	tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg ccttttttacg gttcctggcc	7320
60	ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac	7380

65

ES 2 709 216 T3

5	cgtattaccg cctttgagtg agctgatacc gctcgcgcga gccgaacgac cgagcgcagc	7440
	gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgcggt	7500
	tggccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag	7560
10	cgcaacgcaa ttaatgtgag ttagctcact cattaggcac ccagggcttt acactttatg	7620
	cttccggctc gtatgttggtg tggaattgtg agcggataac aatttcacac aggaacagc	7680
	tatgaccatg aggcgcgcgc gattc	7705
15	<210> 32 <211> 7555 <212> ADN <213> Secuencia artificial	
20	<220> <223> pCAG-ArtTall-Sbfl <400> 32	
25	gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
	catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca	120
	acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	180
30	ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
	aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa tgacggtaaa tggccgcgct	300
	ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat	360
35	tagtcatcgc tattaccatg gtcgaggtga gcccacggt ctgcttcact ctccccatct	420
	cccccccctc cccacccccca attttgtatt tattttatttt ttaattattt tgtgcagcga	480
	tgggggcggg gggggggggg gggcgcgcg caggcggggc gggcgggggc gagggcgggg	540
40	gcggggcgag gcggagaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc	600
	cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcggcgcg	660
	gagtcgctgc gcgctgcctt cgccccgtgc cccgctccgc cgccgcctcg cgccgccgcg	720
45	cccggtcttg actgaccgcg ttactccac aggtgagcgg gcgggacggc ccttctctc	780
	cgggtgtaa ttagcgcttg gtttaatgac ggcttggttc ttttctgtgg ctgcgtgaaa	840
	gccttgaggg gctccgggag ggccttttgt gcggggggga gggctcggg ggggtgcgtgc	900
50	gtgtgtgtgt gcgtggggag cgccgcgtgc ggctccgcgc tgcccggcgg ctgtgagcgc	960
	tgcgggcgcg gcgcggggct ttgtgcgtc cgagtgctgc gcgaggggag cgcgccggg	1020
	ggcggtgccc cgcggtgcgg ggggggctgc gaggggaaca aaggctgcgt gcggggtgtg	1080
55	tgcgtggggg ggtgagcagg ggggtggggc gcgtcggtcg ggctgcaacc cccctgcac	1140
	ccccctcccc gagttgctga gcacggcccg gcttcgggtg cggggctccg tacggggcgt	1200
	ggcgcggggc tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcggggcg	1260
65		

ES 2 709 216 T3

	gggccgcctc	gggccgggga	gggctcgggg	gaggggcgcg	gcggcccccg	gagcgccggc	1320
5	ggctgtcgag	gcgcggcgag	ccgcagccat	tgccttttat	ggtaatcgtg	cgagagggcg	1380
	cagggacttc	ctttgtccca	aatctgtgcg	gagccgaaat	ctgggagggc	ccgccgcacc	1440
10	ccctctagcg	ggcgcggggc	gaagcggg	ggcgccggca	ggaaggaaat	gggcggggag	1500
	ggccttcgtg	cgtcgcccg	ccgccgtccc	cttctccctc	tccagcctcg	gggctgtccg	1560
	cggggggacg	gctgccttcg	ggggggacgg	ggcagggcg	ggttcggctt	ctggcgtgtg	1620
15	accggcggt	ctagagcctc	tgctaaccat	gttcatgcct	tcttcttttt	cctacagatc	1680
	cttaattaat	aatacgactc	actatagggg	ccgccaccat	gggacctaat	aaaaagagga	1740
20	aggtggcggc	cgctgactac	aaggatgacg	acgataaacc	aggtggcgga	ggtagtggcg	1800
	gaggtgggg	acccgccagt	ccagcagccc	aggtggatct	gagaaccctc	ggctacagcc	1860
	agcagcagca	ggagaagatc	aaaccaaagg	tgcgggtccac	cgtcgctcag	caccatgaag	1920
25	cactggtggg	gcacggtttc	acacacgccc	atattgtggc	tctgtctcag	catcccgtcg	1980
	cactcgggac	tgtggccg	aaatatcagg	acatgatcgc	cgctctgcct	gaggcaaccc	2040
	acgaagccat	tgtgggcg	ggaaagcagt	ggagcggg	cagagcactc	gaagcactcc	2100
30	tcaccgtcgc	cggggaactg	cgggggtccac	cactccagtc	cggactggac	actggacagc	2160
	tgctgaagat	cgctaaacgc	ggcggagtga	cagctgtgga	agctgtgcac	gcttgagga	2220
35	atgctctgac	aggagcccca	ctgaatctta	ctccagaaca	ggtcgtcgca	atcgcaagta	2280
	acatcgggcg	aaaacaggcc	ctcgaaaccg	tccagagact	cctccccgtg	ctgtgccagg	2340
	cccacggact	gaccccacag	caggtgggtc	ccatcgctag	caacggcgga	gggaagcagg	2400
40	ctctggagac	cgtgcagagg	ctgctccccg	tcctgtgcca	ggcacatggg	ctcacacctc	2460
	agcaggtgg	cgcaattgcc	tccaatgg	gcggaaaaca	ggccctggaa	actgtgcaga	2520
45	gactgctccc	cgtgctgtgc	caggctcacg	gtctcacacc	ccagcaggtg	gtcgctatcg	2580
	catctcatga	cgggggcaag	caggcactgg	agacagtgca	gcggctgctc	cctgtcctgt	2640
	gccaggccca	cggactcact	cctcagcagg	tcgtcgccat	tgctagtaac	ggcggagggg	2700
50	aacaggctct	ggaaaccgtg	cagcgcctgc	tccccgtgct	gtgccaagcc	cacggcctga	2760
	ccccccagca	ggtggtcgca	atcgccctca	acaatgggtg	caagcaggcc	ctggagactg	2820
55	tgcagcgact	gctcccagtg	ctgtgccagg	cccatggact	cacaccacag	caggtcgtcg	2880
	ctattgcaag	caacaatgga	gggaaacagg	cactggaaac	agtccagagg	ctgctccccg	2940
	tgctgtgcca	agcgcagtg	ctcactcccc	agcaggtcgt	cgccatcgct	tccaataacg	3000
60	gcggcaagca	ggccctggag	accgtccaga	gactgctccc	cgtgctgtgc	caagctcacg	3060
	gactcacacc	tgagcaggtc	gtggcaatcg	cctctaaccat	tgaggggaaa	caggccctgg	3120

65

ES 2 709 216 T3

	aaactgtaca gcggtgctc cccgtgctgt gccaaagcaca cggactcact ccacagcagg	3180
5	tcgtggccat tgcaagtcat gacggaggca agcaggccct ggaaacagtg cagcgcctgc	3240
	tccctgtgct gtgccaggct catggtctga ctccctcagca ggtggtggcc atcgcttcca	3300
10	acaatggagg gaagcaggcc ctggagaccg tacagagact gctccccgtg ctgtgccaaag	3360
	cgcacggctc gacccctcag caggctcgtcg caatcgccag caatggcggg ggcaagcagg	3420
	ctctcgaaac cgtccagcgg ctccctccag tcctctgtca ggctcacggc ctgacccac	3480
15	agcaggctcg cgtattgct tctaattggcg gagggcgcc tgcctctggag agcattgtgg	3540
	ctcagctgtc caggcccgat cctgccctgg ctatagccgc actcactaac gatcatctgg	3600
	tcgctctcgc ttgcctcggg ggacggcccc ctctggacgc agtcaaaaag ggtctcccc	3660
20	atgctccgc actgatcaag agaaccaaca ggagaattcc tgagggatcc gatcgttta	3720
	acatctctgt ggacctgcca ggcgagagg aattcctgct gagtccagcc ggacagaacc	3780
25	ccctgctgaa gaaaaagggtg gaggaattcg tgccccggtt tgctcctcgc agcaccgtgc	3840
	tgtacctggg ggacacaagg ggcaagcact ccctgttcga gagagaaatc tttgaggaag	3900
	tgctggccct gaccttcgat cctcacggac ggatgccaga cctgattctg catgatgagg	3960
30	tgagggggtg gctgttcctg atggaagccg tgaagtctaa aggcccttt gatgaggaaa	4020
	ggcatagaag cctgcaggag ctgtttgtga ctccctccgc cggcctgac ttcgtgaact	4080
	gctttgagaa tagggaatct atgagacagt ggctgccga gctggcttgg gagaccgaag	4140
35	cctgggtggc tgaagacct gatcacctga ttcatctgaa tggaagtcgg tttctggggc	4200
	catatgagcg ctgaacgcgt aaatgattgc agatccacta gttctagaat tccagctgag	4260
40	cgcgggtcgc taccattacc agttggtctg gtgtcaaaaa taataataac cgggcagggg	4320
	ggatctgcat ggatctttgt gaaggaacct tacttctgtg gtgtgacata attggacaaa	4380
	ctacctacag agatttaaag ctctaaggta aatataaaat ttttaagtgt ataatgtgtt	4440
45	aaactactga ttctaattgt ttgtgtatct tagattccaa cctatggaac tgatgaatgg	4500
	gagcagtggt ggaatgccag atccagacat gataagatac attgatgagt ttggacaaac	4560
	cacaactaga atgcagtga aaaaatgctt tatttgtgaa atttgtgatg ctattgcttt	4620
50	atttctaacc attataagct gcaataaaca agttaacaac aacaattgca ttcattttat	4680
	gtttcagggt cagggggagg tgtgggagg tttttaagc aagtaaaacc tctacaaatg	4740
55	tggtatggct gattatgac tgccggccgc actggccgct gttttacaac gtcgtgactg	4800
	ggaaaaccct ggcgttacc aacttaatcg ccttgacgca catccccct tccgagctg	4860
	gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca gcctgaatgg	4920
60	cgaatggaac gcgcctgta gcggcgcat aagcgcgcg ggtgtggtgg ttacgcgcag	4980
	cgtgaccgct acacttgcca gcgcctagc gcccgctcct ttcgctttct tcccttcctt	5040

65

ES 2 709 216 T3

	tctcgccacg	ttcgccggct	ttccccgtca	agctctaaat	cgggggctcc	ctttagggtt	5100
5	ccgatttagt	gctttacggc	acctcgaccc	caaaaaactt	gattaggggtg	atggttcacg	5160
	tagtgggcca	tcgccctgat	agacggtttt	tcgccctttg	acgttggagt	ccacgttctt	5220
10	taatagtgga	ctcttggtcc	aaactggaac	aacactcaac	cctatctcgg	tctattcttt	5280
	tgatttataa	gggattttgc	cgatttcggc	ctatttggtta	aaaaatgagc	tgatttaaca	5340
	aaaatttaac	gcgaatttta	acaaaatatt	aacgcttaca	atthaggtgg	cacttttcgg	5400
15	ggaaatgtgc	gcggaacccc	tatttggtta	tttttctaaa	tacattcaaa	tatgtatccg	5460
	ctcatgagac	aataaccctg	ataaatgctt	caataatatt	gaaaaaggaa	gagtatgagt	5520
20	attcaacatt	tccgtgtcgc	ccttattccc	ttttttgcgg	cattttgcct	tcctgttttt	5580
	gctcaccag	aaacgctggt	gaaagtaaaa	gatgctgaag	atcagttggg	tgacagagtg	5640
	ggttacatcg	aactggatct	caacagcggc	aagatccttg	agagttttcg	ccccgaagaa	5700
25	cgttttccaa	tgatgagcac	ttttaaaagt	ctgctatgtg	gcgcggtatt	atcccgtatt	5760
	gacgccgggc	aagagcaact	cggtcgccgc	atacactatt	ctcagaatga	cttggttgag	5820
30	tactcaccag	tcacagaaaa	gcattcttac	gatggcatga	cagtaagaga	attatgcagt	5880
	gctgccataa	ccatgagtga	taacactgcg	gccaacttac	ttctgacaac	gatcggagga	5940
	ccgaaggagc	taaccgcttt	tttgcacaa	atgggggatc	atgtaactcg	ccttgatcgt	6000
35	tggaaccgg	agctgaatga	agccatacca	aacgacgagc	gtgacaccac	gatgcctgta	6060
	gcaatggcaa	caacgttgcg	caaactatta	actggcgaac	tacttactct	agcttcccgg	6120
40	caacaattaa	tagactggat	ggaggcggat	aaagttgcag	gaccacttct	gcgctcggcc	6180
	cttcoggctg	gctggtttat	tgctgataaa	tctggagccg	gtgagcgtgg	gtctcgcggt	6240
	atcattgcag	cactggggcc	agatggtaag	ccctcccgtg	tcgtagtatt	ctacacgacg	6300
45	gggagtcagg	caactatgga	tgaacgaaat	agacagatcg	ctgagatagg	tgccctactg	6360
	attaagcatt	ggtaactgtc	agaccaagtt	tactcatata	tacttttagat	tgatttaaaa	6420
	cttcattttt	aattttaaaag	gatctaggtg	aagatccttt	ttgataatct	catgacccaa	6480
50	atcccttaac	gtgagttttc	gttccactga	gcgtcagacc	ccgtagaaaa	gatcaaagga	6540
	tcttcttgag	atcctttttt	tctgcgcgta	atctgctgct	tgcaaacaaa	aaaaccaccg	6600
55	ctaccagcgg	tggtttgttt	gccggatcaa	gagctaccaa	ctctttttcc	gaaggtaact	6660
	ggcttcagca	gagcgcagat	accaaatact	gtccttctag	tgtagccgta	gttaggccac	6720
	cacttcaaga	actctgtagc	accgcctaca	tacctcgctc	tgctaatact	gttaccagtg	6780
60	gctgctgcc	gtggcgataa	gtcgtgtctt	accgggttg	actcaagacg	atagttaccg	6840
	gataaggcgc	agcggtcggg	ctgaacgggg	ggttcgtgca	cacagcccag	cttgagcgca	6900

65

ES 2 709 216 T3

	acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc	6960
5	gaagggagaa aggcggacag gtatccggtg agcggcaggg tcggaacagg agagcgcacg	7020
	agggagcttc cagggggaaa cgcctggtat ctttatagtc ctgtcgggtt tcgccacctc	7080
	tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc	7140
10	agcaacgcgg cctttttacg gttcctggcc ttttgctggc cttttgtcga catgttcttt	7200
	cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg agctgatacc	7260
15	gctcgccgca gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc	7320
	ccaatacgca aaccgcctct ccccgcgctg tggccgattc attaatgcag ctggcacgac	7380
	aggtttcccg actggaagc gggcagtgag cgcaacgcaa ttaatgtgag ttagctcact	7440
20	cattaggcac ccagggcttt acactttatg cttccggctc gtatgttgtg tggaattgtg	7500
	agcggataac aatttcacac aggaaacagc tatgaccatg aggcgcgccg gattc	7555
25	<210> 33 <211> 7555 <212> ADN <213> Secuencia artificial	
30	<220> <223> pCAG-ArtTall-Sda <400> 33	
35	gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
	catatatgga gttccgcggt acataactta cggtaaattg cccgcctggc tgaccgcccc	120
	acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagggg	180
40	ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
	aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct	300
	ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat	360
45	tagtcatcgc tattaccatg gtogaggatga gcccacggt ctgcttcact ctccccatct	420
	cccccccctc cccaccccc aattttgtatt tatttatattt ttaattattt tgtgcagcga	480
50	tgggggaggg gggggggggg gggcgcgcg caggcggggc gggcgggggc gaggggaggg	540
	gcggggcgag gcggagaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc	600
	cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcgggcgg	660
55	gagtcgctgc gcgctgcctt cgccccgtgc cccgctccgc cgccgcctcg cgccgcccgc	720
	cccgctctg actgaccgcg ttactccac aggtgagcgg gcgggacggc ctttctcctc	780
	cgggctgtaa ttagcgcttg gtttaatgac ggcttgtttc tttctgtgg ctgctgaaa	840
60	gccttgaggg gctccgggag ggccctttgt gcggggggga gcggctcggg gggcgctgc	900
	gtgtgtgtgt gcgtggggag cgccgcgtgc ggctccgcgc tgcccgggcg ctgtgagcgc	960

ES 2 709 216 T3

5	tgccgggacgcg gcgcgggggct ttgtgcgcctc cgcagtgtgc gcgaggggag cgcggccggg	1020
	ggcgggtgcc cgcgggtgcg ggggggctgc gaggggaaca aaggctgcgt gcgggggtgtg	1080
	tgctgtgggg ggtgagcagg ggggtgtggc gcgtcggtcg ggctgcaacc cccctgcac	1140
10	ccccctccc gagttgctga gcacggcccg gcttcgggtg cggggctccg tacggggcgt	1200
	ggcgcggggc tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcggggcg	1260
	gggcccgcctc gggccgggga gggctcgggg gaggggagcg gcggcccccg gagcgccggc	1320
15	ggctgtcgag gcgcggcgag ccgcagccat tgccttttat ggtaatcgtg cgagagggcg	1380
	cagggacttc ctttgtccca aatctgtgcg gagccgaaat ctgggaggcg ccgccgcacc	1440
	ccctctagcg ggcgcggggc gaagcgggtgc ggcgcgggca ggaaggaaat gggcggggag	1500
20	ggccttcgtg cgtgcgcgcg ccgccgtccc cttctccctc tccagcctcg gggctgtccg	1560
	cggggggacg gctgccttcg ggggggacgg ggcaggggcg ggttcggctt ctggcgtgtg	1620
	accggcggct ctagagcctc tgctaaccat gttcatgcct tcttcttttt cctacagatc	1680
25	cttaattaat aatacgactc actatagggg ccgccaccat gggacctag aaaaagagga	1740
	aggtggcggc cgctgactac aaggatgacg acgataaacc aggtggcgga ggtagtggcg	1800
	gaggtggggt acccgccagt ccagcagccc aggtggatct gagaaccctc ggctacagcc	1860
30	agcagcagca ggagaagatc aaaccaaagg tgcgggtccac cgtcgctcag caccatgaag	1920
	cactggtggg gcacggtttc acacacgccc atattgtggc tctgtctcag catcccgtg	1980
	cactcgggac tgtggccgtc aaatatcagg acatgatcgc cgctctgcct gaggcaaccc	2040
35	acgaagccat tgtgggcgtc ggaaagcagt ggagcgggtc cagagcactc gaagcactcc	2100
	tcaccgtcgc cggggaactg cggggtccac cactccagtc cggactggac actggacagc	2160
	tgctgaagat cgctaaacgc ggcggagtga cagctgtgga agctgtgcac gcttgagga	2220
40	atgctctgac aggagcccca ctgaatctta ctccagaaca ggtcgtcgca atcgcaagta	2280
	acatcggcgg aaaacaggcc ctcgaaaccg tccagagact cctccccgtg ctgtgccagg	2340
	cccacggact gacccacag cagggtggtc ccacgcctag caacggcgga gggaagcagg	2400
45	ctctggagac cgtgcagagg ctgctccccg tcctgtgcca ggcacatggg ctcacacctc	2460
	agcaggtggt cgcaattgcc tccaatggtg gcggaaaaca ggcctggaa actgtgcaga	2520
	gactgtctcc cgtgctgtgc caggctcacg gtctcacacc ccagcagggtg gtcgctatcg	2580
50	catctcatga cgggggcaag caggcactgg agacagtgc cgggctgctc cctgtcctgt	2640
	gccaggccca cggactcact cctcagcagg tcgtcgccat tgctagtaac ggcggaggga	2700
	aacaggctct ggaaaccgtg cagcgcctgc tccccgtgct gtgccaagcc cacggcctga	2760
55	ccccccagca ggtggtcgca atcgctcaa acaatggtgg caagcaggcc ctggagactg	2820
60		
65		

ES 2 709 216 T3

	tgcagcgact gctcccagtg ctgtgccagg cccatggact cacaccacag caggctcgctg	2880
5	ctattgcaag caacaatgga gggaaacagg cactggaaac agtccagagg ctgctccccg	2940
	tgtgtgtcca agcgcgatgga ctccactcccc agcaggctcgt cgccatcgct tccaataacg	3000
	gcggaagca ggccctggag accgtccaga gactgctccc cgtgctgtgc caagctcacg	3060
10	gactcacacc tgagcagggtc gtggcaatcg cctctaacat tggagggaaa caggccctgg	3120
	aaactgtaca ggggtgctc cccgtgctgt gccaaagcaca cggactcact ccacagcagg	3180
	tctgtggccat tgcaagtcac gacggaggca agcaggccct ggaaacagtg cagcgcctgc	3240
15	tccctgtgct gtgccaggct catggtctga ctctcagca ggtggtggcc atcgcttcca	3300
	acaatggagg gaagcaggcc ctggagaccg tacagagact gctccccgtg ctgtgccaag	3360
	cgcacgggtct gacccctcag caggctcgctg caatcgccag caatggcggg ggcaagcagg	3420
20	ctctcgaaac cgtccagcgg ctccctcccag tcctctgtca ggctcacggc ctgaccccac	3480
	agcaggctcgt cgctattgct tctaattggcg gagggcgggc tgctctggag agcattgtgg	3540
	ctcagctgtc caggcccgat cctgccctgg ctagatccgc actcactaac gatcatctgg	3600
25	tcgctctcgc ttgcctcggg ggacggcccg ctctggacgc agtcaaaaag ggtctcccc	3660
	atgctccgc actgatcaag agaaccaaca ggagaattcc tgagggatcc gatcgtttaa	3720
	acattagcgt ggacctcgcc gatggagatg agttcctgct gagccccgct ggacagaatc	3780
30	ctctgctgaa aaagatgggtg gaagaattta tgccacgatt cgcacctgga gctaagggtc	3840
	tgtacatcgg cgactggcga ggaaagcaca cacgggtcga gaaacgcatt tttgaggaaa	3900
35	ccctggggct cacatttgat ccacacggta gaatgccga cctggtgctc catgataagg	3960
	tccggaaatg gctgttcctc atggaggccg tgaagagcaa aggccctttt gacgaggaaa	4020
	ggcatagaac tctgcgggaa ctcttcgcta cccagtcgc aggactggtg ttcgtcaact	4080
40	gctttgagaa tcgagaagcc atgaggcagt ggctgccga gctcgcttg gagaccgaag	4140
	catgggtggc cgacgacct gaccacctga tccacctcaa cgggagcaga ttcctgggac	4200
	cctatgaaag atgaacgcgt aaatgattgc agatccacta gttctagaat tccagctgag	4260
45	cgccggtcgc taccattacc agttggtctg gtgtcaaaaa taataataac cgggcagggg	4320
	ggatctgcat ggatctttgt gaaggaacct tacttctgtg gtgtgacata attggacaaa	4380
	ctacctacag agatttaaag ctctaaggta aatataaaat ttttaagtgt ataagtgtt	4440
50	aaactactga ttctaattgt ttgtgtatth tagattccaa cctatggaac tgatgaatgg	4500
	gagcagtggt ggaatgccag atccagacat gataagatac attgatgagt ttggacaaac	4560
	cacaactaga atgcagtga aaaaatgctt tatttgtgaa atttgtgatg ctattgcttt	4620
55	atttgaacc attataagct gcaataaaca agttaacaac aacaattgca ttcattttat	4680
	gtttcagggt cagggggagg tgtgggagggt tttttaagc aagtaaaacc tctacaaatg	4740
60		
65		

ES 2 709 216 T3

5	tggtatggct gattatgatc tgcggccgcc actggccgtc gttttacaac gtcgtgactg	4800
	ggaaaaccct ggcgttaccc aacttaatcg ccttgacagca catccccctt tcgccagctg	4860
	gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca gcctgaatgg	4920
10	cgaatggaac gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg ttacgcgcag	4980
	cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct tcccttcctt	5040
	tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc ctttagggtt	5100
15	ccgatttagt gctttacggc acctcgaccc caaaaaactt gattaggggtg atgggtcacg	5160
	tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttgaggt ccacgttctt	5220
	taatagtgga ctcttgttcc aaactggaac aacactcaac cctatctcgg tctattcttt	5280
20	tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc tgatttaaca	5340
	aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttaggtgg cacttttcgg	5400
	ggaaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatcgg	5460
25	ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt	5520
	attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cattttgcct tccgtgtttt	5580
	gctcaccacg aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg	5640
30	ggttacatcg aactggatct caacagcggg aagatccttg agagttttcg ccccgaaaga	5700
	cgttttccaa tgatgagcac ttttaagtt ctgctatgtg gcgcgggtatt atcccgattt	5760
	gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga cttggttgag	5820
35	tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt	5880
	gctgccataa ccatgagtga taacactgcg gccaaacttac ttctgacaac gatcggagga	5940
	ccgaaggagc taaccgcttt ttgacacaac atgggggatc atgtaactcg ccttgatcgt	6000
40	tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta	6060
	gcaatggcaa caacgttgcg caaactatta actggcgaaac tacttactct agcttcccgg	6120
	caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc	6180
45	cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggg	6240
	atcattgcag cactggggcc agatggtaag ccctcccgtg tcgtagttat ctacacgacg	6300
	gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg	6360
50	attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat tgatttaaaa	6420
	cttcattttt aatttaaaaag gatctaggtg aagatccttt ttgataatct catgacaaaa	6480
	atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga	6540
55	tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccacgg	6600
60		
65		

ES 2 709 216 T3

	ctaccagcgg	tggtttgttt	gccggatcaa	gagctaccaa	ctctttttcc	gaaggtaact	6660
5	ggcttcagca	gagcgcagat	accaaatact	gtccttctag	tgtagccgta	gttaggccac	6720
	cacttcaaga	actctgtagc	accgcctaca	tacctcgctc	tgctaatact	gttaccagtg	6780
	gctgctgcc	gtggcgataa	gtcgtgtctt	accgggttgg	actcaagacg	atagttaccg	6840
10	gataaggcgc	agcggtcggg	ctgaacgggg	ggttcgtgca	cacagcccag	cttgagcgca	6900
	acgacctaca	ccgaactgag	atacctacag	cgtgagctat	gagaaagcgc	cacgcttccc	6960
15	gaagggagaa	aggcggacag	gtatccggta	agcggcaggg	tcggaacagg	agagcgcacg	7020
	agggagcttc	cagggggaaa	cgcttggtat	ctttatagtc	ctgtcgggtt	tcgccacctc	7080
	tgacttgagc	gtcgattttt	gtgatgctcg	tcaggggggc	ggagcctatg	gaaaaacgcc	7140
20	agcaacgcgg	cctttttacg	gttcctggcc	ttttgctggc	cttttgctca	catgttcttt	7200
	cctgcgttat	cccctgattc	tgtggataac	cgtattaccg	cctttgagtg	agctgatacc	7260
	gctcgccgca	gccgaacgac	cgagcgcagc	gagtcagtga	gcgaggaagc	ggaagagcgc	7320
25	ccaatacgca	aaccgcctct	ccccgcgcgt	tggccgattc	attaatgcag	ctggcacgac	7380
	aggtttcccg	actggaaagc	gggcagtgag	cgcaacgcaa	ttaatgtgag	ttagctcact	7440
30	cattaggcac	cccagccttt	acactttatg	cttcggctc	gtatgttgtg	tggaattgtg	7500
	agcggataac	aatttcacac	aggaacagc	tatgaccatg	aggcgcgccg	gattc	7555
35	<210> 34 <211> 7690 <212> ADN <213> Secuencia artificial						
40	<220> <223> pCAG-ArtTall-Sisl						
	<400> 34						
45	gacattgatt	attgactagt	tattaatagt	aatcaattac	ggggtcatta	gttcatagcc	60
	catatatgga	gttccgcggt	acataactta	cggtaaattg	ccgcctggc	tgaccgccca	120
	acgacccccg	ccattgacg	tcaataatga	cgtatgttcc	catagtaacg	ccaataggga	180
50	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	tgcccacttg	gcagtacatc	240
	aagtgtatca	tatgccaagt	acgcccccta	ttgacgtcaa	tgacggtaaa	tgccccgcct	300
	ggcattatgc	ccagtacatg	accttatggg	actttcctac	ttggcagtac	atctacgtat	360
55	tagtcatcgc	tattaccatg	gtcgagggtga	gccccacgtt	ctgcttcact	ctccccatct	420
	cccccccctc	cccaccccca	attttgtatt	tatttatattt	ttaattattt	tgtgcagcga	480
60	tgggggcggg	gggggggggg	gggcgcgcgc	caggcggggc	ggggcggggc	gaggggcggg	540
	gcggggcgag	gcggagaggt	gcggcggcag	ccaatcagag	cggcgcgctc	cgaaagtttc	600
	cttttatggc	gaggcggcgg	cggcggcggc	cctataaaaa	gcgaagcgcg	cggcgggcgg	660

ES 2 709 216 T3

5	gagtcgctgc gcgctgcctt cgccccgtgc ccgctccgc cgccgcctcg cgccgcccgc	720
	cccggctctg actgaccgcg ttactccac aggtgagcgg gcgggacggc ccttctcctc	780
	cgggctgtaa ttagcgcttg gtttaatgac ggcttgtttc ttttctgtgg ctgctgaaa	840
10	gccttgaggg gctccgggag ggcctttgt gcggggggga gcgctcggg ggggtcgtgc	900
	gtgtgtgtgt gcgtggggag cgccgcgtgc ggctccgcgc tgcccgggcg ctgtgagcgc	960
	tgccggcgcg gcgcggggct ttgtgcgtc cgcaagtgtc gcgaggggag cgcgccggg	1020
15	ggcgggtccc cgcggtgcgg ggggggctgc gaggggaaca aaggctgcgt gcgggggtgtg	1080
	tgctggtggg ggtgagcagg ggggtgtggc gcgtcggtcg ggctgcaacc cccctgcac	1140
	ccccctccc gagttgctga gcacggccc gcttcgggtg cggggctccg tacggggcgt	1200
20	ggcgcggggc tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcgggcg	1260
	gggcccgcctc gggccgggga gggctcgggg gaggggcgcg gcggccccc gagcgccggc	1320
	ggctgtcgag gcgcggcgag ccgcagccat tgccctttat ggtaatcgtg cgagaggcg	1380
25	cagggaactc ctttgtccca aatctgtgcg gagccgaaat ctgggaggcg ccgccgcacc	1440
	ccctctagcg ggcgcggggc gaagcgggtc ggcgccggca ggaaggaaat gggcggggag	1500
	ggccttcgtg cgctgcgcgc ccgcgtccc cttctccctc tccagcctcg gggctgtccg	1560
30	cggggggacg gctgccttcg ggggggacgg ggcaggcgcg ggttcggctt ctggcgtgtg	1620
	accggcggt ctagagcctc tgctaaccat gttcatgcct tcttcttttt cctacagatc	1680
	cttaattaat aatacgactc actatagggg ccgccaccat gggacctaa gaaaagagga	1740
35	aggtggcggc cgctgactac aaggatgacg acgataaacc aggtggcgga ggtagtgcg	1800
	gaggtggggt acccgccagt ccagcagccc aggtggatct gagaaccctc ggctacagcc	1860
	agcagcagca ggagaagatc aaaccaaagg tgcgggtccac cgctcgctcag caccatgaag	1920
40	cactggtggg gcacgggttc acacacgccc atattgtggc tctgtctcag catcccgtg	1980
	cactcgggac tgtggccgtc aaatatcagg acatgatcgc cgctctgcct gaggcaaccc	2040
	acgaagccat tgtgggcgtc ggaaagcagt ggagcgggtc cagagcactc gaagcactcc	2100
45	tcaccgtcgc cggggaactg cggggtccac cactccagtc cggactggac actggacagc	2160
	tgctgaagat cgctaaacgc ggcggagtga cagctgtgga agctgtgcac gcttgagga	2220
	atgctctgac aggagcccca ctgaatctta ctccagaaca ggtcgtcgca atcgcaagta	2280
50	acatcggcgg aaaacaggcc ctcgaaaccg tccagagact cctccccgtg ctgtgccagg	2340
	cccacggact gacccacag caggtgggtc ccatcgctag caacggcgga gggaagcagg	2400
	ctctggagac cgtgcagagg ctgctccccg tcctgtgcca ggcacatggg ctcacacctc	2460
55	agcaggtggt cgcaattgcc tccaatggtg gcggaaaaca ggcctggaa actgtgcaga	2520
60		
65		

ES 2 709 216 T3

	gactgctccc cgtgctgtgc caggctcacg gtctcacacc ccagcaggtg gtcgctatcg	2580
5	catctcatga cgggggcaag caggcactgg agacagtgcg gcggctgctc cctgtcctgt	2640
	gccaggccca cggactcact cctcagcagg tcgtcgccat tgctagtaac ggcgaggga	2700
	aacaggctct ggaaaccgtg cagcgcctgc tccccgtgct gtgccaagcc cacggcctga	2760
10	ccccccagca ggtggtcgca atcgccctcaa acaatggtgg caagcaggcc ctggagactg	2820
	tgcagcgact gctcccagtg ctgtgccagg cccatggact cacaccacag caggctcgctg	2880
	ctattgcaag caacaatgga gggaaacagg cactggaaac agtccagagg ctgctccccg	2940
15	tgctgtgcca agcgcattgga ctactcccc agcaggctgt cgccatcgct tccaataacg	3000
	gcggcaagca ggccctggag accgtccaga gactgctccc cgtgctgtgc caagctcacg	3060
20	gactcacacc tgagcaggtc gtggcaatcg cctctaacat tggagggaaa caggccctgg	3120
	aaactgtaca gcggctgctc cccgtgctgt gccaaagcaca cggactcact ccacagcagg	3180
	tcgtggccat tgcaagtcat gacggaggca agcaggccct ggaaacagtg cagcgcctgc	3240
25	tccctgtgct gtgccaggct catggtctga ctctcagca ggtggtggcc atcgcttcca	3300
	acaatggagg gaagcaggcc ctggagaccg tacagagact gctccccgtg ctgtgccaag	3360
	cgcacggctc gacccctcag caggctcgct caatcgccag caatggcggg ggcaagcagg	3420
30	ctctcgaaac cgtccagcgg ctccctccag tcctctgtca ggctcacggc ctgacccccc	3480
	agcaggctgt cgctattgct tctaattggc gagggcgccc tgctctggag agcattgtgg	3540
	ctcagctgtc caggcccgat cctgccctgg ctagatccgc actcactaac gatcatctgg	3600
35	tcgctctcgc ttgcctcggg ggacggcccg ctctggacgc agtcaaaaag ggtctcccc	3660
	atgctcccg c actgatcaag agaaccaaca ggagaattcc tgagggatcc gatcggttaa	3720
	acgatgtggt gctggagaaa agcgacatcg aaaaattcaa gaaccagctg aggaccgagc	3780
40	tgacaaatat tgatcactcc tacctgaagg gaatcgacat tgccctcaag aaaaagacct	3840
	ctaactgga gaatacagag tttgaagcta tctctactaa gattttcacc gatgaactgg	3900
	gcttcagcgg gaaacatctg ggcggaagca ataagccaga tggcctgctg tgggacgatg	3960
45	actgcgccat cattctggac agtaaggctt acagcgaggg gttccccctg acagcctccc	4020
	acactgacgc tatgggcagg tatctgagac agtttactga gcggaaagag gaaatcaagc	4080
	ccacctggtg ggatattgcc cctgaacatc tggacaacac ctacttcgct tatgtgagcg	4140
50	gctccttttc tggaaattat aaagagcagc tgcagaagtt ccgccaggat acaaaccacc	4200
	tggggggcgc cctggaattt gtgaagctgc tgctgctggc taacaattac aaaactcaga	4260
	agatgtccaa aaaggagggtg aaaaagtcta tcctggacta taacattagt tacgaggaat	4320
55	atgccccct gctggctgag atcgaatgaa cgcgtaaatg attgcagatc cactagttct	4380
	agaattccag ctgagcgccg gtcgctacca ttaccagttg gtctggtgtc aaaaataata	4440
60		
65		

ES 2 709 216 T3

	ataaccgggc agggggggtc tgcattggatc tttgtgaagg aaccttactt ctgtgggtgtg	4500
5	acataattgg acaaactacc tacagagatt taaagctcta aggtaaatat aaaattttta	4560
	agtgtataat gtgttaaact actgattcta attgtttgtg tatttttagat tocaacctat	4620
	ggaactgatg aatgggagca gtggtggaat gccagatcca gacatgataa gatacattga	4680
10	tgagtttggc caaaccacaa ctagaatgca gtgaaaaaaa tgctttattt gtgaaatttg	4740
	tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagtta acaacaacaa	4800
15	ttgcattcat tttatgtttc aggttcaggg ggaggtgtgg gaggtttttt aaagcaagta	4860
	aaacctctac aaatgtggta tggctgatta tgatctgcgg ccgccactgg ccgtcgtttt	4920
	acaacgtcgt gactgggaaa accctggcgt taccctactt aatcgccctg cagcacatcc	4980
20	ccctttcggc agctggcgta atagcgaaga ggcccgacc gatcgccctt cccaacagtt	5040
	gcgcagcctg aatggcgaat ggaacgcgcc ctgtagcggc gcattaagcg cggcgggtgt	5100
	ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg ctcccttcgc	5160
25	tttcttcctt tcttttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg	5220
	gctcccttta ggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgatta	5280
	gggtgatggg tcacgtatg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt	5340
30	ggagtccacg ttctttaata gtgactctt gttccaaact ggaacaacac tcaaccctat	5400
	ctcggctcat tcttttgatt tataagggat tttgccgatt tcggcctatt ggttaaaaaa	5460
35	tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc ttacaattta	5520
	ggtggcaact ttcggggaaa tgtgcgcgga acccctattt gtttatTTTT ctaaatacat	5580
	tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa	5640
40	aggaagagta tgagtattca acatttcctg gtgcgcctta ttcccttttt tgcggcattt	5700
	tgccctcctg tttttgctca ccagaaaacg ctggtgaaag taaaagatgc tgaagatcag	5760
	ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaaagt ccttgagagt	5820
45	tttcgccccg aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg	5880
	gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag	5940
	aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta	6000
50	agagaattat gcagtgtgct cataacctatg agtgataaca ctgcggccaa cttacttctg	6060
	acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta	6120
	actcgccctg atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac	6180
55	accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt	6240
	actctagctt cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca	6300

60

65

ES 2 709 216 T3

	cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag	6360
5	cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta	6420
	gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag	6480
	ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt	6540
10	tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat	6600
	aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgtg	6660
	gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa	6720
15	acaaaaaac caccgctacc agcgggtggt tgtttgccgg atcaagagct accaactctt	6780
	tttccgaag taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag	6840
20	ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta	6900
	atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca	6960
	agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag	7020
25	cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa	7080
	agcgccacgc ttcccgaagg gagaaaggcg gacaggatc cggtaagcgg cagggtcggg	7140
	acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc	7200
30	gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc	7260
	ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt	7320
	gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt	7380
35	gagtgaactg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag	7440
	gaagcggaag agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa	7500
40	tgcagctggc acgacagggt tcccgaactg aaagcgggca gtgagcgcaa cgcaattaat	7560
	gtgagttagc tcaactatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg	7620
	ttgtgtgtaa ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgaggcg	7680
45	cgccggattc	7690

<210> 35

<211> 7654

50 <212> ADN

<213> Secuencia artificial

<220>

<223> pCAG-ArtTal1-FokI

55 <400> 35

	gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
60	catatatgga gttccgcgtt acataactta cggtaaattg cccgcctggc tgaccgcca	120
	acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	180

65

ES 2 709 216 T3

5	ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
	aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgctt	300
	ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat	360
10	tagtcatcgc tattaccatg gtcgaggtga gccccacgtt ctgcttcact ctccccatct	420
	cccccccctc cccacccccca attttgtatt tatttatttt ttaattatct tgtgcagcga	480
	tgggggaggg gggggggggg gggcgcgcg caggcggggc gggcgggggc gaggggaggg	540
15	gcggggaggg gcgggagagt gcggcgagc ccaatcagag cggcgagctc cgaaagtctc	600
	cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcggcggg	660
	gagtcgctgc gcgctgcctt cgccccgtgc ccgctccgc cgccgctcgc cgccgcccgc	720
20	cccggtctgc actgaccgcg ttactccac aggtgagcgg gcgggagcgc ccttctcctc	780
	cgggctgtaa ttagcgcttg gtttaatgac ggcttggttc ttttctgtgg ctgctgaaa	840
	gccttgaggg gctccgggag ggcctttgt gcggggggga gcggtcggg gggcgctgc	900
25	gtgtgtgtgt gcgtggggag cgccgctgc ggctccgcgc tgcccgggcg ctgtgagcgc	960
	tgcggcgcg gcgcggggct ttgtgcgctc cgcagtgtgc gcgaggggag cgcgccggg	1020
	ggcggtgcc cgcggtgcg ggggggctgc gaggggaaca aagctgcgt gcggggtgtg	1080
30	tgctggggg ggtgagcagg ggggtgtggc gcgtcggtcg ggctgcaacc cccctgcac	1140
	ccccctccc gagttgctga gcacggccc gcttcgggtg cggggctccg tacggggcgt	1200
	ggcgcggggc tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcgggcg	1260
35	gggcccctc gggccgggga gggctcggg gaggggcgcg gcggcccccg gaggcgggc	1320
	ggctgtcag gcgcggcag ccgcagccat tgccttttat ggtaatcgt cgagagggcg	1380
	cagggacttc ctttgtcca aatctgtgc gagccgaaat ctgggaggcg ccgccgacc	1440
40	ccctctagc ggcgcggggc gaagcgggtgc ggcgccggca ggaaggaaat gggcggggag	1500
	ggccttcgt gcgtgccgc ccgcgctccc cttctccctc tccagcctcg gggctgtccg	1560
	cggggggac gctgccttc ggggggacgg ggcagggcg ggttcggctt ctggcgtgtg	1620
45	accggcggt ctagagcctc tgctaaccat gttcatgcct tcttcttttt cctacagatc	1680
	cttaattaat aatacagctc actataggg ccgccaccat gggacctaag aaaaagagga	1740
	aggtggcggc cgctgactac aaggatgac acgataaacc aggtggcgga ggtagtggcg	1800
50	gaggtggggt acccgccagt ccagcagccc aggtggatct gagaaccctc ggctacagcc	1860
	agcagcagca ggagaagatc aaaccaaagg tgcgggtccac cgtcgctcag caccatgaag	1920
	cactggtggg gcacgggttc acacacgccc atattgtggc tctgtctcag catcccgtg	1980
55	cactcgggac tgtggccgtc aaatatcagg acatgatcgc cgctctgcct gaggcaaccc	2040
60		
65		

ES 2 709 216 T3

	acgaagccat	tgtgggcgtc	ggaaagcagt	ggagcgggtgc	cagagcactc	gaagcactcc	2100
5	tcaccgtcgc	cggggaactg	cgggggtccac	cactccagtc	cggactggac	actggacagc	2160
	tgctgaagat	cgctaaacgc	ggcggagtga	cagctgtgga	agctgtgcac	gcttggagga	2220
	atgctctgac	aggagcccca	ctgaatctta	ctccagaaca	ggtcgtcgca	atcgcaagta	2280
10	acatcggcgg	aaaacaggcc	ctcgaaaccg	tccagagact	cctccccgtg	ctgtgccagg	2340
	cccacggact	gaccccacag	caggtggtcg	ccatcgctag	caacggcgga	gggaagcagg	2400
	ctctggagac	cgtgcagagg	ctgctccccg	tcctgtgcc	ggcacatggg	ctcacacctc	2460
15	agcaggtggt	cgcaattgcc	tccaatggtg	gcggaaaaca	ggccctggaa	actgtgcaga	2520
	gactgctccc	cgtgctgtgc	caggctcacg	gtctcacacc	ccagcaggtg	gtcgctatcg	2580
	catctcatga	cgggggcaag	caggcactgg	agacagtgc	gcggctgctc	cctgtcctgt	2640
20	gccaggccca	cggactcact	cctcagcagg	tcgtcgccat	tgctagtaac	ggcggaggga	2700
	aacaggctct	ggaaaccgtg	cagcgcctgc	tccccgtgct	gtgccaaagg	cacggcctga	2760
	ccccccagca	ggtggtcgca	atcgcctcaa	acaatggtgg	caagcaggcc	ctggagactg	2820
25	tgcagcgact	gctcccagtg	ctgtgccagg	cccatggact	cacaccacag	caggctcgctg	2880
	ctattgcaag	caacaatgga	gggaaacagg	cactggaaac	agtccagagg	ctgctccccg	2940
	tgctgtgcc	agcgcattga	ctcactcccc	agcaggctcg	cgccatcgct	tccaataacg	3000
30	gcggcaagca	ggccctggag	accgtccaga	gactgctccc	cgtgctgtgc	caagctcacg	3060
	gactcacacc	tgagcaggtc	gtggcaatcg	cctctaacat	tggagggaaa	caggccctgg	3120
35	aaactgtaca	gcggctgctc	cccgtgctgt	gccaagcaca	cggactcact	ccacagcagg	3180
	tcgtggccat	tgcaagtcat	gacggaggca	agcaggccct	ggaaacagtg	cagcgcctgc	3240
	tccctgtgct	gtgccaggct	catggtctga	ctcctcagca	ggtggtggcc	atcgcttcca	3300
40	acaatggagg	gaagcaggcc	ctggagaccg	tacagagact	gctccccgtg	ctgtgccaag	3360
	cgcacggctc	gacccctcag	caggctcgctg	caatcgccag	caatggcggg	ggcaagcagg	3420
	ctctcgaaac	cgtccagcgg	ctcctcccag	tcctctgtca	ggctcacggc	ctgacccccc	3480
45	agcaggctcg	cgctattgct	tctaattggc	gagggcgggc	tgctctggag	agcattgtgg	3540
	ctcagctgtc	caggcccgat	cctgccctgg	ctagatccgc	actcactaac	gatcatctgg	3600
	tcgctctcgc	ttgcctcggt	ggacggcccg	ctctggacgc	agtcaaaaag	ggtctcccc	3660
50	atgctccgc	actgatcaag	agaaccaaca	ggagaattcc	tgagggatcc	gatcggttaa	3720
	accagctcgt	gaaaagcgaa	ctcgaagaaa	agaaaagtga	actgcggcac	aaactgaaat	3780
	acgtcccaca	tgaatacatt	gagctgatcg	agattgctag	gaactccacc	caggacagaa	3840
55	tcctcgagat	gaaagtgatg	gaattcttta	tgaaagtcta	cgggtatcgg	ggcaagcacc	3900
	tgggcgggatc	tcgcaaacca	gatggggcaa	tctacactgt	gggtagtccc	atcgactatg	3960

60

65

ES 2 709 216 T3

	gcgtgattgt cgataccaag gcctacagtg ggggttataa tctgcccatt ggacaggctg	4020
5	acgagatgca gcgatacgtg gaggaaaacc agacaagaaa taagcatatc aaccccaatg	4080
	agtgggtggaa agtgtatcct agctccgtca ctgaattcaa gtttctcttc gtgtcaggcc	4140
	actttaaggg aaactacaaa gcacagctga ccaggctcaa tcatattaca aactgcaatg	4200
10	gcgccgtgct gagcgtcgag gaactgctca tcggcggaga gatgatcaag gccggcacac	4260
	tcaccctgga ggagggtccgc cgaaaattca ataacgggga aatcaacttc tgaacgcgta	4320
15	aatgattgca gatccactag ttctagaatt ccagctgagc gccggtcgct accattacca	4380
	gttggtctgg tgtcaaaaat aataataacc gggcaggggg gatctgcatg gatccttctg	4440
	aaggaacctt acttctgtgg tgtgacataa ttggacaaac tacctacaga gatttaaagc	4500
20	tctaaggtaa atataaaatt tttaagtgtg taatgtgtta aactactgat tctaattggt	4560
	tgtgtatctt agattccaac ctatggaact gatgaatggg agcagtgggtg gaatgccaga	4620
	tccagacatg ataagataca ttgatgagtt tggacaaacc acaactagaa tgcagtga	4680
25	aaaatgcttt atttgtgaaa tttgtgatgc tattgcttta tttgtaacca ttataagctg	4740
	caataaacia gttaacaaca acaattgcat tcattttatg tttcaggttc agggggaggt	4800
	gtgggaggtt ttttaagca agtaaaacct ctacaaatgt ggtatggctg attatgatct	4860
30	gcggccgcca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca	4920
	acttaatcgc cttgcagcac atcccccttt cgccagctgg cgtaaatagcg aagaggcccg	4980
35	caccgatcgc ccttcccaac agttgcgcag cctgaatggc gaatggaacg cgccctgtag	5040
	cggcgcatga agcgcggcgg gtgtgggtgg tacgcgcagc gtgaccgcta cacttgccag	5100
	cgccctagcg cccgctcctt tcgctttctt cccttccttt ctgccacgt tcgccggtt	5160
40	tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg ctttacggca	5220
	cctcgacccc aaaaaacttg attagggtga tgggtcacgt agtgggcat cgccctgata	5280
	gacgggtttt cgccctttga cgttggagtc cacgttcttt aatagtggac tcttgttcca	5340
45	aactggaaca acactcaacc ctatctcggc ctattctttt gatttataag ggattttgcc	5400
	gatttcggcc tattgggtta aaaatgagct gatttaacaa aaatttaacg cgaattttaa	5460
	caaaatatta acgcttacia tttagggtgg acttttcggg gaaatgtgcg cggaaccctt	5520
50	atttgtttat ttttctaaat acattcaa atgtatccgc tcatgagaca ataaccctga	5580
	taaatgcttc aataatattg aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc	5640
	cttattccct tttttgcggc attttgccct cctgtttttg ctcaaccaga aacgctgggtg	5700
55	aaagtaaaag atgctgaaga tcagttgggt gcacgagtggt gttacatcga actggatctc	5760
	aacagcggta agatccttga gagttttcgc ccgaagaac gttttccaat gatgagcact	5820

ES 2 709 216 T3

5	tttaaagttc tgctatgtgg cgcggtatta tcccgtattg acgccgggca agagcaactc	5880
	ggtcgccgca tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag	5940
	catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat	6000
10	aacactgcmg ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt	6060
	ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa	6120
	gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac aacgttgcmg	6180
15	aaactattaa ctggcgaaat acttactcta gcttcccggc aacaattaat agactggatg	6240
	gaggcgata aagttgcagg accacttctg cgctcgccc ttccggctgg ctggtttatt	6300
	gctgataaat ctggagccgg tgagcgtggg tctcgcggtg tcattgcagc actggggcca	6360
20	gatggtaagc cctcccgtat cgtagttagt tacacgacgg ggagtcaggc aactatggat	6420
	gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg gtaactgtca	6480
	gaccaagttt actcatatat acttttagatt gatttaaaac ttcattttta atttaaaagg	6540
25	atctaggtga agatcctttt tgataatctc atgacaaaaa tcccttaacg tgagttttcg	6600
	ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt	6660
	ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccacgc taccagcggg ggtttgtttg	6720
30	ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata	6780
	ccaaataactg tccttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca	6840
35	ccgcctacat acctcgctct gctaatacctg ttaccagtgg ctgctgccag tggcgataag	6900
	tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc	6960
	tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga	7020
40	tacctacagc gtgagctatg agaaagcgcc acgcttcccg aaggggagaaa ggcggacagg	7080
	tatccggtaa gcggcagggg cggaacagga gagcgacga gggagcttcc agggggaaac	7140
	gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg	7200
45	tgatgctcgt cagggggggc gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg	7260
	ttcctggcct tttgctggcc ttttgctcac atgttctttc ctgctttatc ccctgattct	7320
	gtggataacc gtattaccgc ctttgagtga gctgataccg ctgcgccgag ccgaacgacc	7380
50	gagcgcagcg agtcagtga cgaggaagcg gaagagcgcc caatacgcaa accgcctctc	7440
	cccgcgcggt ggccgattca ttaatgcagc tggcacgaca ggtttcccga ctggaaagcg	7500
55	ggcagtgaac gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta	7560
	cactttatgc ttccggctcg tatgttgtgt ggaattgtga gcggataaca atttcacaca	7620
60	ggaaacagct atgacatga ggcgcgccgg attc	7654
65		

ES 2 709 216 T3

<210> 36
 <211> 947
 5 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> ArtTall-Alw
 10 <400> 36

 Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp
 1 5 10 15

 Asp Asp Asp Lys Pro Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Pro
 20 25 30

 Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln
 35 40 45

 Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln
 50 55 60

 His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val
 65 70 75 80

 Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr
 85 90 95

 Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val
 100 105 110

 Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu
 115 120 125

 Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Ser Gly Leu Asp
 130 135 140

 Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val
 145 150 155 160

 Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn
 165 170 175

 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
 180 185 190

 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
 195 200 205

 His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly

ES 2 709 216 T3

	210	215	220	
5	Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys 225 230 235 240			
10	Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn 245 250 255			
15	Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val 260 265 270			
20	Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala 275 280 285			
25	Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu 290 295 300			
30	Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala 305 310 315 320			
35	Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg 325 330 335			
40	Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val 340 345 350			
45	Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val 355 360 365			
50	Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln 370 375 380			
55	Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu 385 390 395 400			
60	Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr 405 410 415			
65	Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala 420 425 430			
	Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly 435 440 445			
	Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys 450 455 460			

ES 2 709 216 T3

	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	465	470	475	480
5	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly		485	490	495
10	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys		500	505	510
15	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn		515	520	525
20	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val		530	535	540
25	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala		545	550	555
30	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu		565	570	575
35	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala		580	585	590
40	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Arg	Pro	Ala	Leu	Glu	Ser	Ile	Val	Ala		595	600	605
45	Gln	Leu	Ser	Arg	Pro	Asp	Pro	Ala	Leu	Ala	Arg	Ser	Ala	Leu	Thr	Asn		610	615	620
50	Asp	His	Leu	Val	Ala	Leu	Ala	Cys	Leu	Gly	Gly	Arg	Pro	Ala	Leu	Asp		625	630	635
55	Ala	Val	Lys	Lys	Gly	Leu	Pro	His	Ala	Pro	Ala	Leu	Ile	Lys	Arg	Thr		645	650	655
60	Asn	Arg	Arg	Ile	Pro	Glu	Gly	Ser	Asp	Arg	Leu	Asn	Lys	Glu	Thr	Asn		660	665	670
65	Ile	Leu	Leu	Val	Glu	Gln	Leu	Glu	Glu	Thr	Leu	Asn	Arg	Asn	Arg	Ile		675	680	685
	Leu	Phe	Glu	Lys	Asn	Ser	Ser	Ile	Ala	Gln	Ala	Pro	Ile	Gly	Glu	Ile		690	695	700
	Lys	Asn	Tyr	Arg	Tyr	His	Leu	Glu	Glu	Leu	Leu	Phe	Glu	Asn	Asn	Glu		705	710	715
																				720

ES 2 709 216 T3

5	Lys Lys Phe Ala Glu Asn Gln Lys Asn Glu Trp Asp Glu Ile Leu Ala	725	730	735
10	Tyr Met Asp Leu Leu Ile Ser Pro Lys Pro Ile Ser Ile Glu Ile Ala	740	745	750
15	Asp Lys Glu Ile Ser Ile Pro Ser Gly Glu Arg Pro Ala Tyr Phe Glu	755	760	765
20	Trp Val Leu Trp Arg Ala Phe Leu Ala Leu Asn His Leu Ile Ile Glu	770	775	780
25	Pro Gln Gln Cys Arg Arg Phe Lys Val Asp Gln Asp Phe Lys Pro Ile	785	790	795
30	His Asn Ala Pro Gly Gly Gly Ala Asp Val Ile Phe Glu Tyr Glu Asn	805	810	815
35	Phe Lys Ile Leu Gly Glu Val Thr Leu Thr Ser Asn Ser Arg Gln Glu	820	825	830
40	Ala Ala Glu Gly Glu Pro Val Arg Arg His Ile Ala Val Glu Thr Val	835	840	845
45	Asn Thr Pro Asp Lys Asp Val Tyr Gly Leu Phe Leu Ala Leu Thr Ile	850	855	860
50	Asp Thr Asn Thr Ala Glu Thr Phe Arg His Gly Ala Trp Tyr His Gln	865	870	875
55	Glu Glu Leu Met Asp Val Lys Ile Leu Pro Leu Thr Leu Glu Ser Phe	885	890	895
60	Lys Lys Tyr Leu Glu Ser Leu Arg Lys Lys Asn Gln Val Glu Thr Gly	900	905	910
65	Ile Phe Asp Leu Lys Lys Met Met Asp Glu Ser Leu Lys Leu Arg Glu	915	920	925
	Thr Leu Thr Ala Pro Gln Trp Lys Asn Glu Ile Thr Asn Lys Phe Ala	930	935	940
	Arg Pro Ile	945		

ES 2 709 216 T3

<210> 37
 <211> 869
 <212> PRT
 5 <213> Secuencia artificial

 <220>
 <223> ArtTall-CLEDORF

 10 <400> 37

 Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp
 1 5 10 15

 15 Asp Asp Asp Lys Pro Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Pro
 20 25 30

 20 Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln
 35 40 45

 25 Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln
 50 55 60

 His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val
 65 70 75 80

 30 Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr
 85 90 95

 35 Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val
 100 105 110

 Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu
 115 120 125

 40 Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Ser Gly Leu Asp
 130 135 140

 45 Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val
 145 150 155 160

 50 Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn
 165 170 175

 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
 180 185 190

 55 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
 195 200 205

 60 His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
 210 215 220

 65

ES 2 709 216 T3

	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	
	225					230					235					240	
5																	
	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	
					245					250					255		
10																	
	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	
				260					265					270			
15																	
	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	
			275					280					285				
20																	
	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	
	290						295					300					
25																	
	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	
	305					310					315					320	
30																	
	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	
					325					330					335		
35																	
	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	
				340					345					350			
40																	
	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	
			355					360					365				
45																	
	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	
	370						375					380					
50																	
	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	
	385					390					395					400	
55																	
	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	
					405					410					415		
60																	
	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	
				420					425					430			
65																	
	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	
			435					440					445				
70																	
	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly	Lys	
	450						455					460					
75																	
	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	
	465					470					475					480	

ES 2 709 216 T3

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser His Asp Gly
 485 490 495
 5
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
 500 505 510
 10
 Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
 515 520 525
 15
 Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
 530 535 540
 20
 Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
 545 550 555 560
 25
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
 565 570 575
 30
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
 580 585 590
 35
 Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala
 595 600 605
 40
 Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Arg Ser Ala Leu Thr Asn
 610 615 620
 45
 Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp
 625 630 635 640
 50
 Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr
 645 650 655
 55
 Asn Arg Arg Ile Pro Glu Gly Ser Asp Arg Leu Asn Lys Leu Ala Lys
 660 665 670
 60
 Ser Ser Gln Ser Glu Thr Lys Glu Lys Leu Arg Glu Lys Leu Arg Asn
 675 680 685
 65
 Leu Pro His Glu Tyr Leu Ser Leu Val Asp Leu Ala Tyr Asp Ser Lys
 690 695 700
 70
 Gln Asn Arg Leu Phe Glu Met Lys Val Ile Glu Leu Leu Thr Glu Glu
 705 710 715 720
 75
 Cys Gly Phe Gln Gly Leu His Leu Gly Gly Ser Arg Arg Pro Asp Gly

ES 2 709 216 T3

	725								730						735	
5	Val	Leu	Tyr	Thr	Ala	Gly	Leu	Thr	Asp	Asn	Tyr	Gly	Ile	Ile	Leu	Asp
				740					745					750		
10	Thr	Lys	Ala	Tyr	Ser	Ser	Gly	Tyr	Ser	Leu	Pro	Ile	Ala	Gln	Ala	Asp
			755					760					765			
15	Glu	Met	Glu	Arg	Tyr	Val	Arg	Glu	Asn	Gln	Thr	Arg	Asp	Glu	Leu	Val
		770					775					780				
20	Asn	Pro	Asn	Gln	Trp	Trp	Glu	Asn	Phe	Glu	Asn	Gly	Leu	Gly	Thr	Phe
	785					790					795					800
25	Tyr	Phe	Leu	Phe	Val	Ala	Gly	His	Phe	Asn	Gly	Asn	Val	Gln	Ala	Gln
					805					810					815	
30	Leu	Glu	Arg	Ile	Ser	Arg	Asn	Thr	Gly	Val	Leu	Gly	Ala	Ala	Ala	Ser
				820					825					830		
35	Ile	Ser	Gln	Leu	Leu	Leu	Leu	Ala	Asp	Ala	Ile	Arg	Gly	Gly	Arg	Met
			835					840					845			
40	Asp	Arg	Glu	Arg	Leu	Arg	His	Leu	Met	Phe	Gln	Asn	Glu	Glu	Phe	Leu
		850					855					860				
45	Leu	Glu	Gln	Glu	Leu											
	865															
<210>	38															
<211>	867															
<212>	PRT															
<213>	Secuencia artificial															
<220>																
<223>	ArtTal1-Clo051															
<400>	38															
50	Met	Gly	Pro	Lys	Lys	Lys	Arg	Lys	Val	Ala	Ala	Ala	Asp	Tyr	Lys	Asp
	1				5					10					15	
55	Asp	Asp	Asp	Lys	Pro	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Val	Pro
				20					25					30		
60	Ala	Ser	Pro	Ala	Ala	Gln	Val	Asp	Leu	Arg	Thr	Leu	Gly	Tyr	Ser	Gln
			35					40					45			
65	Gln	Gln	Gln	Glu	Lys	Ile	Lys	Pro	Lys	Val	Arg	Ser	Thr	Val	Ala	Gln
	50						55					60				

ES 2 709 216 T3

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

ES 2 709 216 T3

	305		310		315		320									
5	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg
				325					330						335	
10	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val
				340					345					350		
15	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val
			355					360					365			
20	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln
	370						375					380				
25	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu
	385					390				395						400
30	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr
					405					410					415	
35	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala
				420					425					430		
40	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly
			435				440					445				
45	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly	Lys
	450					455						460				
50	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala
	465					470				475						480
55	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly
					485				490						495	
60	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys
			500						505				510			
65	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn
		515						520					525			
70	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val
	530						535					540				
75	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala
	545					550					555					560

ES 2 709 216 T3

5	Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu	565	570	575
10	Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala	580	585	590
15	Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala	595	600	605
20	Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Arg Ser Ala Leu Thr Asn	610	615	620
25	Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp	625	630	635
30	Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr	645	650	655
35	Asn Arg Arg Ile Pro Glu Gly Ser Asp Arg Leu Asn Glu Gly Ile Lys	660	665	670
40	Ser Asn Ile Ser Leu Leu Lys Asp Glu Leu Arg Gly Gln Ile Ser His	675	680	685
45	Ile Ser His Glu Tyr Leu Ser Leu Ile Asp Leu Ala Phe Asp Ser Lys	690	695	700
50	Gln Asn Arg Leu Phe Glu Met Lys Val Leu Glu Leu Leu Val Asn Glu	705	710	715
55	Tyr Gly Phe Lys Gly Arg His Leu Gly Gly Ser Arg Lys Pro Asp Gly	725	730	735
60	Ile Val Tyr Ser Thr Thr Leu Glu Asp Asn Phe Gly Ile Ile Val Asp	740	745	750
65	Thr Lys Ala Tyr Ser Glu Gly Tyr Ser Leu Pro Ile Ser Gln Ala Asp	755	760	765
	Glu Met Glu Arg Tyr Val Arg Glu Asn Ser Asn Arg Asp Glu Glu Val	770	775	780
	Asn Pro Asn Lys Trp Trp Glu Asn Phe Ser Glu Glu Val Lys Lys Tyr	785	790	795
	Tyr Phe Val Phe Ile Ser Gly Ser Phe Lys Gly Lys Phe Glu Glu Gln	805	810	815

ES 2 709 216 T3

5 Leu Arg Arg Leu Ser Met Thr Thr Gly Val Asn Gly Ser Ala Val Asn
 820 825 830
 Val Val Asn Leu Leu Leu Gly Ala Glu Lys Ile Arg Ser Gly Glu Met
 835 840 845
 10 Thr Ile Glu Glu Leu Glu Arg Ala Met Phe Asn Asn Ser Glu Phe Ile
 850 855 860
 15 Leu Lys Tyr
 865
 <210> 39
 <211> 918
 <212> PRT
 20 <213> Secuencia artificial
 <220>
 <223> ArtTall-Mly
 25 <400> 39
 Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp
 1 5 10 15
 30 Asp Asp Asp Lys Pro Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Pro
 20 25 30
 35 Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln
 35 40 45
 40 Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln
 50 55 60
 45 His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val
 65 70 75 80
 Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr
 85 90 95
 50 Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val
 100 105 110
 55 Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu
 115 120 125
 60 Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Ser Gly Leu Asp
 130 135 140

ES 2 709 216 T3

5	Thr	Gly	Gln	Leu	Leu	Lys	Ile	Ala	Lys	Arg	Gly	Gly	Val	Thr	Ala	Val	145	150	155	160
	Glu	Ala	Val	His	Ala	Trp	Arg	Asn	Ala	Leu	Thr	Gly	Ala	Pro	Leu	Asn	165	170	175	
10	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly	Lys	180	185	190	
15	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	195	200	205	
20	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	210	215	220	
	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	225	230	235	240
25	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	245	250	255	
30	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	260	265	270	
35	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	275	280	285	
40	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	290	295	300	
	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	305	310	315	320
45	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	325	330	335	
50	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	340	345	350	
55	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	355	360	365	
60	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	370	375	380	
65	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	385	390	395	400

ES 2 709 216 T3

	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	
					405					410					415		
5																	
	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	
				420					425					430			
10	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	
			435					440					445				
15	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly	Lys	
	450						455					460					
20	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	
	465					470					475					480	
	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	
					485					490					495		
25																	
	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	
				500					505					510			
30	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	
			515					520					525				
35	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	
	530						535					540					
40	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	
	545					550					555					560	
	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	
					565					570					575		
45																	
	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	
				580					585					590			
50	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Arg	Pro	Ala	Leu	Glu	Ser	Ile	Val	Ala	
			595					600					605				
55	Gln	Leu	Ser	Arg	Pro	Asp	Pro	Ala	Leu	Ala	Arg	Ser	Ala	Leu	Thr	Asn	
	610						615					620					
60	Asp	His	Leu	Val	Ala	Leu	Ala	Cys	Leu	Gly	Gly	Arg	Pro	Ala	Leu	Asp	
	625					630					635					640	
65	Ala	Val	Lys	Lys	Gly	Leu	Pro	His	Ala	Pro	Ala	Leu	Ile	Lys	Arg	Thr	
					645					650					655		

ES 2 709 216 T3

	Asn	Arg	Arg	Ile	Pro	Glu	Gly	Ser	Asp	Arg	Leu	Asn	Ile	Asn	Ser	Lys	
				660					665					670			
5	Ile	Lys	Gln	Leu	Asp	Asp	Ser	Ile	Asn	Val	Glu	Ser	Leu	Lys	Ile	Asp	
			675					680					685				
10	Asp	Ala	Lys	Asp	Leu	Leu	Asn	Asp	Leu	Glu	Ile	Gln	Arg	Lys	Ala	Lys	
		690						695				700					
15	Thr	Ile	Glu	Asp	Thr	Val	Asn	His	Leu	Lys	Leu	Arg	Ser	Asp	Ile	Glu	
	705					710					715					720	
20	Asp	Ile	Leu	Asp	Val	Phe	Ala	Lys	Ile	Lys	Lys	Arg	Asp	Val	Pro	Asp	
				725						730					735		
25	Val	Pro	Leu	Phe	Leu	Glu	Trp	Asn	Ile	Trp	Arg	Ala	Phe	Ala	Ala	Leu	
				740					745					750			
30	Asn	His	Thr	Gln	Ala	Ile	Glu	Gly	Asn	Phe	Ile	Val	Asp	Leu	Asp	Gly	
			755					760					765				
35	Met	Pro	Leu	Asn	Thr	Ala	Pro	Gly	Lys	Lys	Pro	Asp	Ile	Glu	Ile	Asn	
		770					775					780					
40	Tyr	Gly	Ser	Phe	Ser	Cys	Ile	Val	Glu	Val	Thr	Met	Ser	Ser	Gly	Glu	
	785					790					795					800	
45	Thr	Gln	Phe	Asn	Met	Glu	Gly	Ser	Ser	Val	Pro	Arg	His	Tyr	Gly	Asp	
				805						810					815		
50	Leu	Val	Arg	Lys	Val	Asp	His	Asp	Ala	Tyr	Cys	Ile	Phe	Ile	Ala	Pro	
				820					825					830			
55	Lys	Val	Ala	Pro	Gly	Thr	Lys	Ala	His	Phe	Phe	Asn	Leu	Asn	Arg	Leu	
			835					840					845				
60	Ser	Thr	Lys	His	Tyr	Gly	Gly	Lys	Thr	Lys	Ile	Ile	Pro	Met	Ser	Leu	
	850						855					860					
65	Asp	Asp	Phe	Ile	Cys	Phe	Leu	Gln	Val	Gly	Ile	Thr	His	Asn	Phe	Gln	
	865					870					875					880	
70	Asp	Ile	Asn	Lys	Leu	Lys	Asn	Trp	Leu	Asp	Asn	Leu	Ile	Asn	Phe	Asn	
				885						890					895		
75	Leu	Glu	Ser	Glu	Asp	Glu	Glu	Ile	Trp	Phe	Glu	Glu	Ile	Ile	Ser	Lys	

ES 2 709 216 T3

	900	905	910
5	Ile Ser Thr Trp Ala Ile		
	915		
	<210> 40		
10	<211> 881		
	<212> PRT		
	<213> Secuencia artificial		
	<220>		
15	<223> ArtTall-Pept071		
	<400> 40		
20	Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp		
	1 5 10 15		
25	Asp Asp Asp Lys Pro Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Pro		
	20 25 30		
30	Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln		
	35 40 45		
35	Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln		
	50 55 60		
40	His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val		
	65 70 75 80		
45	Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr		
	85 90 95		
50	Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val		
	100 105 110		
55	Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu		
	115 120 125		
60	Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Ser Gly Leu Asp		
	130 135 140		
65	Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val		
	145 150 155 160		
	Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn		
	165 170 175		
	Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys		
	180 185 190		

ES 2 709 216 T3

5	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	195	200	205
10	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	210	215	220
15	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	225	230	235
20	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	245	250	255
25	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	260	265	270
30	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	275	280	285
35	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	290	295	300
40	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	305	310	315
45	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	325	330	335
50	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	340	345	350
55	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	355	360	365
60	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	370	375	380
65	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	385	390	395
	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	405	410	415
	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	420	425	430
	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly			

ES 2 709 216 T3

	435						440						445					
5	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly	Lys		
	450						455					460						
10	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala		
	465					470					475					480		
15	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly		
					485					490					495			
20	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys		
				500					505					510				
25	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn		
			515					520					525					
30	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val		
	530						535					540						
35	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala		
	545					550					555					560		
40	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu		
					565					570					575			
45	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala		
				580					585					590				
50	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Arg	Pro	Ala	Leu	Glu	Ser	Ile	Val	Ala		
			595					600					605					
55	Gln	Leu	Ser	Arg	Pro	Asp	Pro	Ala	Leu	Ala	Arg	Ser	Ala	Leu	Thr	Asn		
		610					615					620						
60	Asp	His	Leu	Val	Ala	Leu	Ala	Cys	Leu	Gly	Gly	Arg	Pro	Ala	Leu	Asp		
	625					630					635					640		
65	Ala	Val	Lys	Lys	Gly	Leu	Pro	His	Ala	Pro	Ala	Leu	Ile	Lys	Arg	Thr		
					645					650					655			
70	Asn	Arg	Arg	Ile	Pro	Glu	Gly	Ser	Asp	Arg	Leu	Asn	Lys	Ile	Ser	Lys		
				660					665					670				
75	Thr	Asn	Val	Leu	Glu	Leu	Lys	Asp	Lys	Val	Arg	Asp	Lys	Leu	Lys	Tyr		
			675					680					685					

ES 2 709 216 T3

5 Val Asp His Arg Tyr Leu Ala Leu Ile Asp Leu Ala Tyr Asp Gly Thr
 690 695 700
 Ala Asn Arg Asp Phe Glu Ile Gln Thr Ile Asp Leu Leu Ile Asn Glu
 705 710 715 720
 10 Leu Lys Phe Lys Gly Val Arg Leu Gly Glu Ser Arg Lys Pro Asp Gly
 725 730 735
 15 Ile Ile Ser Tyr Asn Ile Asn Gly Val Ile Ile Asp Asn Lys Ala Tyr
 740 745 750
 Ser Thr Gly Tyr Asn Leu Pro Ile Asn Gln Ala Asp Glu Met Ile Arg
 755 760 765
 20 Tyr Ile Glu Glu Asn Gln Thr Arg Asp Glu Lys Ile Asn Ser Asn Lys
 770 775 780
 25 Trp Trp Glu Ser Phe Asp Asp Lys Val Lys Asp Phe Asn Tyr Leu Phe
 785 790 795 800
 30 Val Ser Ser Phe Phe Lys Gly Asn Phe Lys Asn Asn Leu Lys His Ile
 805 810 815
 Ala Asn Arg Thr Gly Val Ser Gly Gly Ala Ile Asn Val Glu Asn Leu
 820 825 830
 35 Leu Tyr Phe Ala Glu Glu Leu Lys Ala Gly Arg Leu Ser Tyr Val Asp
 835 840 845
 40 Ser Phe Lys Met Tyr Asp Asn Asp Glu Ile Tyr Val Gly Asp Phe Ser
 850 855 860
 45 Asp Tyr Ser Tyr Val Lys Phe Ala Ala Glu Glu Glu Gly Glu Tyr Leu
 865 870 875 880
 Thr
 50 <210> 41
 <211> 831
 <212> PRT
 <213> Secuencia artificial
 55 <220>
 <223> ArtTall-Sbf
 <400> 41
 60 Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp
 65

ES 2 709 216 T3

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

ES 2 709 216 T3

	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	
				260					265					270			
5																	
	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	
			275					280					285				
10	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	
		290					295					300					
15	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	
	305					310					315					320	
20	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	
					325					330					335		
	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	
				340					345					350			
25																	
	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	
			355					360					365				
30	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	
		370					375					380					
35	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	
	385					390					395					400	
40	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	
					405					410					415		
	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	
				420					425					430			
45																	
	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	
			435					440					445				
50	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly	Lys	
		450					455					460					
55	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	
	465					470					475					480	
	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	
					485					490					495		
60																	
	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	
				500					505					510			
65																	

ES 2 709 216 T3

5	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	
			515					520					525				
10	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	
		530					535					540					
15	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	
	545					550					555					560	
20	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	
					565					570					575		
25	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	
				580					585					590			
30	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Arg	Pro	Ala	Leu	Glu	Ser	Ile	Val	Ala	
			595					600					605				
35	Gln	Leu	Ser	Arg	Pro	Asp	Pro	Ala	Leu	Ala	Arg	Ser	Ala	Leu	Thr	Asn	
		610					615					620					
40	Asp	His	Leu	Val	Ala	Leu	Ala	Cys	Leu	Gly	Gly	Arg	Pro	Ala	Leu	Asp	
	625					630					635					640	
45	Ala	Val	Lys	Lys	Gly	Leu	Pro	His	Ala	Pro	Ala	Leu	Ile	Lys	Arg	Thr	
					645					650					655		
50	Asn	Arg	Arg	Ile	Pro	Glu	Gly	Ser	Asp	Arg	Leu	Asn	Ile	Ser	Val	Asp	
				660					665					670			
55	Leu	Pro	Gly	Gly	Glu	Glu	Phe	Leu	Leu	Ser	Pro	Ala	Gly	Gln	Asn	Pro	
			675					680					685				
60	Leu	Leu	Lys	Lys	Met	Val	Glu	Glu	Phe	Val	Pro	Arg	Phe	Ala	Pro	Arg	
		690					695					700					
65	Ser	Thr	Val	Leu	Tyr	Leu	Gly	Asp	Thr	Arg	Gly	Lys	His	Ser	Leu	Phe	
	705					710					715					720	
70	Glu	Arg	Glu	Ile	Phe	Glu	Glu	Val	Leu	Gly	Leu	Thr	Phe	Asp	Pro	His	
					725					730					735		
75	Gly	Arg	Met	Pro	Asp	Leu	Ile	Leu	His	Asp	Glu	Val	Arg	Gly	Trp	Leu	
				740					745					750			
80	Phe	Leu	Met	Glu	Ala	Val	Lys	Ser	Lys	Gly	Pro	Phe	Asp	Glu	Glu	Arg	
			755					760					765				

ES 2 709 216 T3

5 His Arg Ser Leu Gln Glu Leu Phe Val Thr Pro Ser Ala Gly Leu Ile
 770 775 780
 Phe Val Asn Cys Phe Glu Asn Arg Glu Ser Met Arg Gln Trp Leu Pro
 785 790 795 800
 10 Glu Leu Ala Trp Glu Thr Glu Ala Trp Val Ala Glu Asp Pro Asp His
 805 810 815
 15 Leu Ile His Leu Asn Gly Ser Arg Phe Leu Gly Pro Tyr Glu Arg
 820 825 830
 <210> 42
 <211> 831
 <212> PRT
 20 <213> Secuencia artificial
 <220>
 <223> ArtTal1-Sdal
 25 <400> 42
 Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp
 1 5 10 15
 30 Asp Asp Asp Lys Pro Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Pro
 20 25 30
 35 Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln
 35 40 45
 40 Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln
 50 55 60
 45 His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val
 65 70 75 80
 Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr
 85 90 95
 50 Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val
 100 105 110
 55 Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu
 115 120 125
 60 Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Ser Gly Leu Asp
 130 135 140
 65

ES 2 709 216 T3

5	Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val 145 150 155 160
10	Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn 165 170 175
15	Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys 180 185 190
20	Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala 195 200 205
25	His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly 210 215 220
30	Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys 225 230 235 240
35	Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn 245 250 255
40	Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val 260 265 270
45	Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala 275 280 285
50	Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu 290 295 300
55	Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala 305 310 315 320
60	Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg 325 330 335
65	Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val 340 345 350
	Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val 355 360 365
	Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln 370 375 380
	Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu 385 390 395 400

ES 2 709 216 T3

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
 405 410 415
 5
 Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala
 420 425 430
 10
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
 435 440 445
 15
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
 450 455 460
 20
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
 465 470 475 480
 25
 His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser His Asp Gly
 485 490 495
 30
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
 500 505 510
 35
 Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
 515 520 525
 40
 Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
 530 535 540
 45
 Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
 545 550 555 560
 50
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
 565 570 575
 55
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
 580 585 590
 60
 Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala
 595 600 605
 65
 Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Arg Ser Ala Leu Thr Asn
 610 615 620
 70
 Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp
 625 630 635 640
 75
 Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr

ES 2 709 216 T3

	645	650	655
5	Asn Arg Arg Ile Pro Glu Gly Ser Asp Arg Leu Asn Ile Ser Val Asp 660 665 670		
10	Leu Ala Asp Gly Asp Glu Phe Leu Leu Ser Pro Ala Gly Gln Asn Pro 675 680 685		
15	Leu Leu Lys Lys Met Val Glu Glu Phe Met Pro Arg Phe Ala Pro Gly 690 695 700		
20	Ala Lys Val Leu Tyr Ile Gly Asp Trp Arg Gly Lys His Thr Arg Phe 705 710 715 720		
25	Glu Lys Arg Ile Phe Glu Glu Thr Leu Gly Leu Thr Phe Asp Pro His 725 730 735		
30	Gly Arg Met Pro Asp Leu Val Leu His Asp Lys Val Arg Lys Trp Leu 740 745 750		
35	Phe Leu Met Glu Ala Val Lys Ser Lys Gly Pro Phe Asp Glu Glu Arg 755 760 765		
40	His Arg Thr Leu Arg Glu Leu Phe Ala Thr Pro Val Ala Gly Leu Val 770 775 780		
45	Phe Val Asn Cys Phe Glu Asn Arg Glu Ala Met Arg Gln Trp Leu Pro 785 790 795 800		
50	Glu Leu Ala Trp Glu Thr Glu Ala Trp Val Ala Asp Asp Pro Asp His 805 810 815		
55	Leu Ile His Leu Asn Gly Ser Arg Phe Leu Gly Pro Tyr Glu Arg 820 825 830		
60	Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp 1 5 10 15		
65	Asp Asp Asp Lys Pro Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Pro 20 25 30		

<210> 43
 <211> 876
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> ArtTal1-Stsl

<400> 43

ES 2 709 216 T3

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

ES 2 709 216 T3

[illegible]

ES 2 709 216 T3

5	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	
	530						535					540					
10	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	
	545					550					555					560	
15	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	
					565					570					575		
20	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	
				580					585					590			
25	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Arg	Pro	Ala	Leu	Glu	Ser	Ile	Val	Ala	
			595					600					605				
30	Gln	Leu	Ser	Arg	Pro	Asp	Pro	Ala	Leu	Ala	Arg	Ser	Ala	Leu	Thr	Asn	
	610						615					620					
35	Asp	His	Leu	Val	Ala	Leu	Ala	Cys	Leu	Gly	Gly	Arg	Pro	Ala	Leu	Asp	
	625					630					635					640	
40	Ala	Val	Lys	Lys	Gly	Leu	Pro	His	Ala	Pro	Ala	Leu	Ile	Lys	Arg	Thr	
					645					650					655		
45	Asn	Arg	Arg	Ile	Pro	Glu	Gly	Ser	Asp	Arg	Leu	Asn	Asp	Val	Val	Leu	
				660					665					670			
50	Glu	Lys	Ser	Asp	Ile	Glu	Lys	Phe	Lys	Asn	Gln	Leu	Arg	Thr	Glu	Leu	
			675					680					685				
55	Thr	Asn	Ile	Asp	His	Ser	Tyr	Leu	Lys	Gly	Ile	Asp	Ile	Ala	Ser	Lys	
	690						695					700					
60	Lys	Lys	Thr	Ser	Asn	Val	Glu	Asn	Thr	Glu	Phe	Glu	Ala	Ile	Ser	Thr	
	705					710					715					720	
65	Lys	Ile	Phe	Thr	Asp	Glu	Leu	Gly	Phe	Ser	Gly	Lys	His	Leu	Gly	Gly	
					725					730					735		
70	Ser	Asn	Lys	Pro	Asp	Gly	Leu	Leu	Trp	Asp	Asp	Asp	Cys	Ala	Ile	Ile	
				740					745					750			
75	Leu	Asp	Ser	Lys	Ala	Tyr	Ser	Glu	Gly	Phe	Pro	Leu	Thr	Ala	Ser	His	
			755					760					765				
80	Thr	Asp	Ala	Met	Gly	Arg	Tyr	Leu	Arg	Gln	Phe	Thr	Glu	Arg	Lys	Glu	
	770						775					780					

ES 2 709 216 T3

5 Glu Ile Lys Pro Thr Trp Trp Asp Ile Ala Pro Glu His Leu Asp Asn
785 790 795 800

10 Thr Tyr Phe Ala Tyr Val Ser Gly Ser Phe Ser Gly Asn Tyr Lys Glu
805 810 815

15 Gln Leu Gln Lys Phe Arg Gln Asp Thr Asn His Leu Gly Gly Ala Leu
820 825 830

20 Glu Phe Val Lys Leu Leu Leu Leu Ala Asn Asn Tyr Lys Thr Gln Lys
835 840 845

25 Met Ser Lys Lys Glu Val Lys Lys Ser Ile Leu Asp Tyr Asn Ile Ser
850 855 860

30 Tyr Glu Glu Tyr Ala Pro Leu Leu Ala Glu Ile Glu
865 870 875

<210> 44
<211> 864
<212> PRT
<213> Secuencia artificial

35 Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp
1 5 10 15

40 Asp Asp Asp Lys Pro Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Pro
20 25 30

45 Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln
35 40 45

50 Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln
50 55 60

55 His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val
65 70 75 80

60 Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr
85 90 95

65 Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val
100 105 110

ES 2 709 216 T3

5 Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu
115 120 125

10 Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Ser Gly Leu Asp
130 135 140

15 Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val
145 150 155 160

20 Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn
165 170 175

25 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
180 185 190

30 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
195 200 205

35 His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
210 215 220

40 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
225 230 235 240

45 Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
245 250 255

50 Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
260 265 270

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

60 Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
290 295 300

65 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
305 310 315 320

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
325 330 335

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
340 345 350

Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val
355 360 365

ES 2 709 216 T3

5	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	
	370						375					380					
10	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	
	385					390					395					400	
15	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	
					405					410					415		
20	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	
				420					425					430			
25	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	
			435					440					445				
30	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly	Lys	
	450						455					460					
35	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	
	465					470					475					480	
40	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	
					485					490					495		
45	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	
				500					505					510			
50	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	
			515					520					525				
55	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	
	530						535					540					
60	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	
	545					550					555					560	
65	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	
				565					570						575		
70	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	
			580						585					590			
75	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Arg	Pro	Ala	Leu	Glu	Ser	Ile	Val	Ala	
			595					600					605				
80	Gln	Leu	Ser	Arg	Pro	Asp	Pro	Ala	Leu	Ala	Arg	Ser	Ala	Leu	Thr	Asn	
	610						615					620					

ES 2 709 216 T3

5	Asp 625	His	Leu	Val	Ala	Leu 630	Ala	Cys	Leu	Gly	Gly 635	Arg	Pro	Ala	Leu	Asp 640
10	Ala	Val	Lys	Lys	Gly 645	Leu	Pro	His	Ala	Pro 650	Ala	Leu	Ile	Lys	Arg	Thr 655
15	Asn	Arg	Arg	Ile 660	Pro	Glu	Gly	Ser	Asp 665	Arg	Leu	Asn	Gln	Leu	Val	Lys 670
20	Ser	Glu	Leu	Glu	Glu	Lys	Lys	Ser 680	Glu	Leu	Arg	His	Lys 685	Leu	Lys	Tyr
25	Val	Pro 690	His	Glu	Tyr	Ile	Glu	Leu	Ile	Glu	Ile	Ala 700	Arg	Asn	Ser	Thr
30	Gln 705	Asp	Arg	Ile	Leu	Glu	Met	Lys	Val	Met	Glu	Phe	Phe	Met	Lys	Val 720
35	Tyr	Gly	Tyr	Arg	Gly 725	Lys	His	Leu	Gly	Gly 730	Ser	Arg	Lys	Pro	Asp	Gly 735
40	Ala	Ile	Tyr	Thr	Val	Gly	Ser	Pro	Ile 745	Asp	Tyr	Gly	Val	Ile	Val	Asp 750
45	Thr	Lys	Ala	Tyr	Ser	Gly	Gly	Tyr	Asn 760	Leu	Pro	Ile	Gly 765	Gln	Ala	Asp
50	Glu	Met	Gln	Arg	Tyr	Val	Glu	Glu	Asn	Gln	Thr	Arg	Asn	Lys	His	Ile 770
55	Asn	Pro	Asn	Glu	Trp	Trp	Lys	Val	Tyr	Pro	Ser	Ser	Val	Thr	Glu	Phe 800
60	Lys	Phe	Leu	Phe	Val	Ser	Gly	His	Phe	Lys	Gly	Asn	Tyr	Lys	Ala	Gln 815
65	Leu	Thr	Arg	Leu	Asn	His	Ile	Thr	Asn	Cys	Asn	Gly	Ala	Val	Leu	Ser 830
70	Val	Glu	Glu	Leu	Leu	Ile	Gly	Gly	Glu	Met	Ile	Lys	Ala	Gly	Thr	Leu 845
75	Thr	Leu	Glu	Glu	Val	Arg	Arg	Lys	Phe	Asn	Asn	Gly	Glu	Ile	Asn	Phe 860

ES 2 709 216 T3

<210> 45
<211> 7374
<212> ADN
<213> Secuencia artificial

5

<220>
<223> ArtTall-Reportero
<400> 45

10	cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgccatt	60
	gacgtcaata atgacgtatg ttcccatagt aacgccataa gggactttcc attgacgtca	120
	atgggtggag tattttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc	180
15	aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta	240
	catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	300
	catggtgatg cgggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	360
20	atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	420
	ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt	480
25	acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg	540
	ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct	600
	agaggatccg gtactcgacg aactgcaga gacctacttc actaacaacc ggtatggtcg	660
30	cgagtagctt ggcaactggc gtcgttttac aacgtcgtga ctgggaaaac cctggcgcta	720
	cccaacttaa tcgccttgca gcacatcccc ctttcgccag ctggcgtaat agcgaagagg	780
	cccgccaccga tcgcccttcc caacagttgc gcagcctgaa tggcgaatgg cgctttgcct	840
35	ggtttccggc accagaagcg gtgccggaaa gctggctgga gtgcgatctt cctgaggccg	900
	atactgtcgt cgtcccctca aactggcaga tgcacggtta cgatgcgccc atctacacca	960
40	acgtgacctt tcccattacg gtcaatccgc cgtttgttcc cacggagaat ccgacgggtt	1020
	gttactcgct cacatttaat gttgatgaaa gctggctata aaaccggtac agttcggcca	1080
	ccatggctgt attctgggac gttttcacac tcttctaacg tcccagaata ctcgagtagc	1140
45	ttggcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt	1200
	aatcgccttg cagcacatcc ccctttcgcc agctggcgta atagcgaaga ggcccgacc	1260
	gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggcgctttgc ctggtttccg	1320
50	gcaccagaag cgggtgccgga aagctggctg gagtgcgatc ttcttgaggc cgatactgtc	1380
	gtcgtcccct caaactggca gatgcacggt tacgatgcgc ccatctacac caacgtgacc	1440
	tatccatta cgggtcaatcc gccgtttgtt cccacggaga atccgacggg ttgttactcg	1500
55	ctcacattta atgttgatga aagctggcta caggaaggcc agacgcgaat tatttttgat	1560
	ggcgtaact cggcgtttca tctgtggtgc aacgggcgct gggtcggtta cggccaggac	1620
60	agtcgtttgc cgtctgaatt tgacctgacg gcatttttac gcgccggaga aaaccgcctc	1680

65

ES 2 709 216 T3

	gcggtgatgg	tgctgcgctg	gagtgacggc	agttatctgg	aagatcagga	tatgtggcgg	1740
	atgagcggca	ttttccgtga	cgtctcgttg	ctgcataaac	cgactacaca	aatcagcgat	1800
5	ttccatggtg	ccactcgttt	taatgatgat	ttcagcccg	ctgtactgga	ggctgaagtt	1860
	cagatgtgcg	gcgagttg	tgactaccta	cgggtaacag	tttctttatg	gcagggtgaa	1920
	acgcaggtcg	ccagcggcac	cgcgcctttc	ggcggtgaaa	ttatcgatga	gcgtggtggt	1980
10	tatgccgatc	gcgtcacact	acgtctgaac	gtcgaaaacc	cgaaactgtg	gagcggcgaa	2040
	atcccgaatc	tctatcgtgc	gggtggtgaa	ctgcacaccg	ccgacggcac	gctgattgaa	2100
	gcagaagcct	gcgatgtcgg	tttccgcgag	gtgcggattg	aaaatggtct	gctgctgctg	2160
15	aacggcaagc	cggttctgat	tcgaggcggt	aaccgtcacg	agcatcatcc	tctgcatggg	2220
	caggatcatg	atgagcagac	gatggtgcag	gatatcctgc	tgatgaagca	gaacaacttt	2280
	aacgccgtgc	gctgttcgca	ttatccgaac	catccgctgt	ggtacacgct	gtgcgaccgc	2340
20	tacggcctgt	atgtggtgga	tgaagccaat	attgaaaccc	acggcatggg	gccaatgaat	2400
	cgtctgaccg	atgatccg	ctggctaccg	gcgatgagcg	aacgcgtaac	gcgaatggg	2460
25	cagcgcgatc	gtaatcacc	gagtgatgat	atctggtcgc	tggggaatga	atcaggccac	2520
	ggcgctaata	acgacgcgct	gtatcgctgg	atcaaactct	tcgatccttc	ccgcccggg	2580
	cagtatgaag	gcggcggagc	cgacaccacg	gccaccgata	ttatgtgccc	gatgtacgcg	2640
30	cgcgtggatg	aagaccagcc	cttcccggct	gtgccgaaat	gggccatcaa	aaaatggctt	2700
	tcgctacctg	gagagacgcg	cccgtgatc	ctttgcgaat	acgcccacgc	gatgggtaac	2760
	agtcttggcg	gtttcgctaa	atactggcag	gcgtttcgtc	agtatccccg	tttacagggc	2820
35	ggcttcgtct	gggactgggt	ggatcagtcg	ctgattaaat	atgatgaaaa	cggcaaccgc	2880
	tggctcgctt	acggcgggtga	ttttggcgat	acgccgaacg	atcgccagtt	ctgtatgaac	2940
	ggtctggtct	ttgccgaccg	cacgccgcat	ccagcgtgga	cgggaagcaa	acaccagcag	3000
40	cagtttttcc	agttccgttt	atccgggcaa	accatcgaag	tgaccagcga	atacctgttc	3060
	cgtcatagcg	ataacgagct	cctgcactgg	atggtggcgc	tggatggtaa	gccgctggca	3120
45	agcggatgaag	tgccctctgga	tgctcgtcca	caaggtaaac	agttgattga	actgcctgaa	3180
	ctaccgcagc	cggagagcgc	cgggcaactc	tggctcacag	tacgcgtagt	gcaaccgaac	3240
	gcgaccgcat	ggtcagaagc	cgggcacatc	agcgcctggc	agcagtggcg	tctggcggaa	3300
50	aacctcagtg	tgacgctccc	cgcgcgtccc	cacgccatcc	cgcactctgac	caccagcgaa	3360
	atggattttt	gcactcagct	gggtaataag	cgttggcaat	ttaaccgcca	gtcaggcttt	3420
	ctttcacaga	tgtggattgg	cgataaaaaa	caactgctga	cgcgcgtg	cgatcagttc	3480
55	acccgtgcac	cgctggataa	cgacattggc	gtaagtgaag	cgaccgcgat	tgaccctaac	3540
60							
65							

ES 2 709 216 T3

	gcctgggtcg aacgctggaa ggcggcgggc cattaccagg ccgaagcagc gttgttgacg	3600
	tgcacggcag atacacttgc tgatgcggtg ctgattacga ccgctcacgc gtggcagcat	3660
5	caggggaaaa ccttatttat cagccggaaa acctaccgga ttgatggtag tggcctaatg	3720
	gcgattaccg ttgatgttga agtggcgagc gatacacgcg atccggcgcg gattggcctg	3780
	aactgccagc tggcgagcgt agcagagcgg gtaaaactggc tcggattagg gccgcaagaa	3840
10	aactatccc accgccttac tgccgcctgt tttgaccgct gggatctgcc attgtcagac	3900
	atgtatacc cgtacgtctt cccgagcgaa aacggctctgc gctgcgggac gcgcgaattg	3960
	aattatggcc cacaccagtg gcgcggcgac ttccagttca acatcagccg ctacagtcaa	4020
15	cagcaactga tggaaaccag ccatcgccat ctgctgcacg ccgaagaagg cacatggctg	4080
	aatatcgacg gtttccatat ggggattggt ggcgacgact cctggagccc gtcagtatcg	4140
	gcggaattac agctgagcgc cggtcgctac cattaccagt tggctctggtg tcaaaaataa	4200
20	taataaccgg gcaggccatg tctgccgta tttgcgtaa ggaaatccat tatgtactat	4260
	ttaaaaaaca caaacttttg gatgttcggt ttattctttt tcttttactt ttttatcatg	4320
	ggagcctact tcccgttttt cccgatttgg ctacatgaca tcaaccatat cagcaaaagt	4380
25	gatacgggta ttatttttgc cgtatattct ctgttctcgc tattattcca accgctgttt	4440
	ggtctgcctt ctgacaaaact cggcctcgac tctaggcggc cgcggggatc cagacatgat	4500
	aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa aatgctttat	4560
30	ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca ataaacaagt	4620
	taacaacaac aattgcattc attttatgtt tcagggttcag ggggaggtgt gggaggtttt	4680
35	ttcggatcct ctagagtoga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct	4740
	gtttcctgtg tgaaattggt atccgctcac aattccacac aacatacgag ccggaagcat	4800
	aaagtgtaaa gcctggggtg cctaagtagt gagctaactc acattaattg cgttgcgctc	4860
40	actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg	4920
	cgcggggaga ggcggtttgc gtattggcg ctcttcgct tctcgcctca ctgactcgt	4980
	gcgctcggtc gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg taatacgggt	5040
45	atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc	5100
	caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggtccgcc cccctgacga	5160
	gcacacaaaa aatcgacgct caagtacagag gtggcgaaac ccgacaggac tataaagata	5220
50	ccaggcggtt ccccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac	5280
	cggataacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg	5340
	taggtatctc agttcggtgt aggtcggtcg ctccaagctg ggetgtgtgc acgaaccccc	5400
55	cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtgaa	5460
60		
65		

ES 2 709 216 T3

	acacgactta tgcgcaactgg cagcagccac tggtaacagc attagcagag cgaggatgt	5520
	aggcgggtgct acagagttct tgaagtgtg gcctaactac ggctacacta gaaggacagt	5580
5	atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg	5640
	atccggcaaa caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac	5700
	gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacgggt ctgacgctca	5760
10	gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac	5820
	ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac	5880
	ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt	5940
15	tcggttcaccc atagttgcct gactccccgt cgtgtagata actacgatac gggagggtt	6000
	accatctggc ccagtgctg caatgatacc gcgagacca cgctcaccgg ctccagattt	6060
	atcagcaata aaccagccag ccggaaggcg cgagcgaga agtggtcctg caactttatc	6120
20	cgctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa	6180
	tagtttgccg aacgttggtt ccattgctac aggcacgtg gtgtcacgct cgtcgtttgg	6240
25	tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatgtt	6300
	gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc	6360
	agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt	6420
30	aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg	6480
	gcgaccgagt tgctcttgcc cggcgtaaat acgggataat accgcgccac atagcagaac	6540
	tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc	6600
35	gctgttgaga tccagttcga tgtaaccac tctgtcaccc aactgatctt cagcatcttt	6660
	tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaagg	6720
	aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag	6780
40	catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa	6840
	acaaataggg gttccgcgca catttcccg aaaagtgcc cctgacgtct aagaaacat	6900
45	tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtctcgcgcg	6960
	tttcggtgat gacggtgaaa acctctgaca catgcagctc ccggagacgg tcacagcttg	7020
	tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg gtgttggcgg	7080
50	gtgtcggggc tggcttaact atgcggcatc agagcagatt gtactgagag tgcaccatat	7140
	gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcacaggg gccattcgcc	7200
	attcaggctg cgcaactgtt gggaaggcg atcgggtcgg gcctcttcgc tattacgcca	7260
55	gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag ggttttccca	7320
	gtcacgacgt tgtaaaacga cggccagtga attcgagctt gcacgcctgc aggt	7374
60		
65		

ES 2 709 216 T3

<210> 46
 <211> 7374
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> TalRabl-Reportero

<400> 46

10	cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgccatt	60
	gacgtcaata atgacgtatg ttcccatagt aacgccataa gggactttcc attgacgtca	120
15	atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc	180
	aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta	240
	catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	300
20	catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	360
	atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	420
	ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt	480
25	acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg	540
	ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct	600
30	agaggatccg gtactcgacg aactgcaga gacctacttc actaacaacc ggtatggtcg	660
	cgagtagctt ggcactggcc gtcgttttac aacgtcgtga ctgggaaaac cctggcggtta	720
	cccaacttaa tcgccttgca gcacatcccc ctttcgccag ctggcgtaat agcgaagagg	780
35	cccgcaccga tcgcccttcc caacagttgc gcagcctgaa tggcgaatgg cgctttgcct	840
	ggtttccggc accagaagcg gtgccggaaa gctggctgga gtgcgatctt cctgaggccg	900
	atactgtcgt cgtccctca aactggcaga tgcacgggta cgatgcgccc atctacacca	960
40	acgtgaccta tccattacg gtcaatccgc cgtttgttcc cacggagaat ccgacggggt	1020
	gttactcgct cacatttaat gttgatgaaa gctggctata aaaccggtac agttcggcca	1080
45	ccatggctcg gtgcaccaa acttttcaca ctcttctaag ttttggtgca cacgagtagc	1140
	ttggcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt	1200
	aatcgccttg cagcacatcc ccctttcgcc agctggcgta atagcgaaga ggcccgacc	1260
50	gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggcgctttgc ctggtttccg	1320
	gcaccagaag cggtgccgga aagctggctg gagtgcgatc ttcctgaggc cgatactgtc	1380
	gtcgtcccct caaactggca gatgcacggt tacgatgcgc ccatctacac caacgtgacc	1440
55	tatcccatta cggtaaatcc gccgtttgtt cccacggaga atccgacggg ttgttactcg	1500
	ctcacattta atgttgatga aagctggcta caggaaggcc agacgcgaat tatttttgat	1560

ES 2 709 216 T3

	ggcgttaact	cggcgtttca	tctgtggtgc	aacgggcgct	gggtcgggta	cggccaggac	1620
	agtcgtttgc	cgtctgaatt	tgacctgagc	gcattttttac	gcgccggaga	aaaccgcctc	1680
5	gcggtgatgg	tgctgcgctg	gagtgcgggc	agttatctgg	aagatcagga	tatgtggcgg	1740
	atgagcggca	ttttccgtga	cgtctcggtt	ctgcataaac	cgactacaca	aatcagcgat	1800
	ttccatgttg	ccactcgctt	taatgatgat	ttcagccgcg	ctgtactgga	ggctgaagtt	1860
10	cagatgtgcg	gcgagttgcg	tgactacctt	cgggtaaacg	tttcttttat	gcagggtgaa	1920
	acgcaggctg	ccagcggcac	cgcgcctttc	ggcggtgaaa	ttatcgatga	gcgtggtggt	1980
	tatgccgatc	gcgtcacact	acgtctgaac	gtcgaaaacc	cgaactgtg	gagcgcgcaa	2040
15	atcccgaatc	tctatcgtgc	ggtggttgaa	ctgcacaccg	ccgacggcac	gctgattgaa	2100
	gcagaagcct	gcgatgtcgg	tttccgcgag	gtgcggattg	aaaatggtct	gctgctgctg	2160
	aacggcaagc	cgttgctgat	tcgagggcgt	aaccgtcacg	agcatcatcc	tctgcatggt	2220
20	caggatcatg	atgagcagac	gatggtgcag	gatatcctgc	tgatgaagca	gaacaacttt	2280
	aacgccgtgc	gctgtttcga	ttatccgaac	catccgctgt	ggtacacgct	gtgcgaccgc	2340
25	tacggcctgt	atgtggtgga	tgaagccaat	attgaaaccc	acggcatggt	gccaatgaat	2400
	cgtctgaccg	atgatccgcg	ctggctaccg	gcgatgagcg	aacgcgtaac	gcgaatggtg	2460
	cagcgcgatc	gtaatcacc	gagtgtgatc	atctggtcgc	tggggaatga	atcaggccac	2520
30	ggcgctaata	acgacgcgct	gtatcgctgg	atcaaactct	tcgatccttc	ccgcccgggtg	2580
	cagtatgaag	gcggcggagc	cgacaccacg	gccaccgata	ttatttgccc	gatgtacgcg	2640
	cgcgtggatg	aagaccagcc	cttcccggct	gtgccgaaat	ggtccatcaa	aaaatggctt	2700
35	tcgctacctg	gagagacgcg	cccgtgatc	ctttgcgaat	acgcccacgc	gatgggtaac	2760
	agtcttggcg	gtttcgctaa	atactggcag	gcgtttcgtc	agtatccccg	tttacagggc	2820
	ggcttcgtct	gggactgggt	ggatcagtcg	ctgattaaat	atgatgaaaa	cggcaacccg	2880
40	tggtcggctt	acggcgggtga	ttttggcgat	acgccgaacg	atcgccagtt	ctgtatgaac	2940
	ggtctggtct	ttgccgaccg	cacgccgcat	ccagcgtga	cggaaagcaa	acaccagcag	3000
45	cagtttttcc	agttccgttt	atccgggcaa	accatcgaag	tgaccagcga	atacctgttc	3060
	cgtcatagcg	ataacgagct	cctgcactgg	atggtggcgc	tggtatggtg	gccgctggca	3120
	agcgggtgaag	tgccctctgga	tgctcgctcca	caaggtaaac	agttgattga	actgcctgaa	3180
50	ctaccgcagc	cggagagcgc	cgggcaactc	tggctcacag	tacgcgtagt	gcaaccgaac	3240
	gcgaccgcat	ggtcagaagc	cgggcacatc	agcgcctggc	agcagtggcg	tctggcggaa	3300
	aacctcagtg	tgacgctccc	cgcgcgtccc	cacgccatcc	cgcactctgac	caccagcgaa	3360
55	atggattttt	gcactcgagct	gggtaataag	cgttggcaat	ttaaccgcca	gtcaggcttt	3420
60							
65							

ES 2 709 216 T3

	ctttcacaga	tgtggattgg	cgataaaaaa	caactgctga	cgccgctgcg	cgatcagttc	3480
	acccgtgcac	cgctggataa	cgacattggc	gtaagtgaag	cgacccgcat	tgaccctaac	3540
5	gcctgggtcg	aacgctggaa	ggcggcgggc	cattaccagg	ccgaagcagc	gttggtgcag	3600
	tgcacggcag	atacacttgc	tgatgcggtg	ctgattacga	ccgctcacgc	gtggcagcat	3660
	caggggaaaa	ccttatttat	cagccggaaa	acctaccgga	ttgatggtag	tggtcaaattg	3720
10	gcgattaccg	ttgatgttga	agtggcgagc	gatacaccgc	atccggcgcg	gattggcctg	3780
	aactgccagc	tggcgaggt	agcagagcgg	gtaaactggc	tgggattagg	gccgcaagaa	3840
	aactatcccg	accgccttac	tgccgcctgt	tttgaccgct	gggatctgcc	attgtcagac	3900
15	atgtataccc	cgtagctctt	cccgcgcgaa	aacgggtctgc	gctgcgggac	gcgcgaattg	3960
	aattatggcc	cacaccagtg	gcgcggcgac	ttccagttca	acatcagccg	ctacagtcaa	4020
20	cagcaactga	tggaaaccag	ccatcgccat	ctgctgcacg	cggaagaagg	cacatggctg	4080
	aatatcgacg	gtttccatat	ggggattggg	ggcgacgact	cctggagccc	gtcagtatcg	4140
	gcggaattac	agctgagcgc	cggtcgctac	cattaccagt	tgggtctggtg	tcaaaaataa	4200
25	taataaccgg	gcaggccatg	tctgcccgta	tttcgcgtaa	ggaaatccat	tatgtactat	4260
	ttaaaaaaca	caaacttttg	gatgttcggg	ttattctttt	tcttttactt	ttttatcatg	4320
	ggagcctact	tcccgttttt	cccgatattg	ctacatgaca	tcaaccatat	cagcaaaagt	4380
30	gatacgggta	ttatttttgc	cgctatttct	ctgttctcgc	tattattcca	accgctgttt	4440
	ggtctgcttt	ctgacaaact	cggcctcgac	tctaggcggc	cgcggggac	cagacatgat	4500
	aagatacatt	gatgagtttg	gacaaaccac	aactagaatg	cagtgaaaaa	aatgctttat	4560
35	ttgtgaaatt	tgtgatgcta	ttgctttatt	tgttaaccatt	ataagctgca	ataaacaagt	4620
	taacaacaac	aattgcattc	attttatggt	tcaggttcag	ggggaggtgt	gggaggtttt	4680
40	ttcggatcct	ctagagtcga	cctgcaggca	tgcaagcttg	gcgtaatcat	ggcatagct	4740
	gtttcctgtg	tgaattgtt	atccgctcac	aattccacac	aacatacgag	ccggaagcat	4800
	aaagtgtaaa	gcctgggggtg	cctaagtgtg	gagctaactc	acattaattg	cgttgcgctc	4860
45	actgcccgtc	ttccagtcgg	gaaacctgtc	gtgccagctg	cattaatgaa	tcggccaacg	4920
	cgcggggaga	ggcggtttgc	gtattgggcg	ctcttccgct	tcctcgctca	ctgactcgct	4980
	gcgctcggtc	gttcggctgc	ggcgagcggg	atcagctcac	tcaaaggcgg	taatacgggt	5040
50	atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	5100
	caggaaccgt	aaaaaggccg	cggtgctggc	gtttttccat	aggctccgcc	cccctgacga	5160
	gcatcacaaa	aatcgacgct	caagtcagag	gtggcgaaac	ccgacaggac	tataaagata	5220
55	ccaggcggtt	ccccctggaa	gctccctcgt	gcgctctcct	gttccgaccc	tgcgcgttac	5280
	cggataacctg	tccgcctttc	tcccttcggg	aagcgtggcg	ctttctcata	gtcacgctg	5340
60							
65							

ES 2 709 216 T3

	taggtatctc agttcgggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc	5400
	cggtcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggttaag	5460
5	acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatatgt	5520
	aggcgggtgct acagagttct tgaagtgggtg gcctaactac ggctacacta gaaggacagt	5580
10	atcttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg	5640
	atccggcaaa caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac	5700
	gcgagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca	5760
15	gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa g gatcttcac	5820
	ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac	5880
	ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt	5940
20	tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt	6000
	accatctggc ccagtgctg caatgatacc gcgagacca cgctcaccgg ctccagattt	6060
25	atcagcaata aaccagccag ccggaagggc cgagcgaga agtggtcctg caactttatc	6120
	cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa	6180
	tagtttgccg aacgttggtt ccattgctac aggcacgtg gtgtcacgct cgtcgtttgg	6240
30	tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatggt	6300
	gtgcaaaaa gcggttagct ccttcgggtc tccgatcgtt gtcagaagta agttggccgc	6360
	agtggtatca ctcatgggtt tggcagcact gcataattct cttactgtca tgccatccgt	6420
35	aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg	6480
	gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac	6540
40	tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc	6600
	gctgttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt cagcatcttt	6660
	tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaatgccg caaaaaagg	6720
45	aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag	6780
	catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa	6840
	acaaataggg gttccgcgca catttccccg aaaagtcca cctgacgtct aagaaacat	6900
50	tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtctcgcgcg	6960
	tttcggtgat gacggtgaaa acctctgaca catgcagctc ccggagacgg tcacagcttg	7020
55	tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg gtgttgccgg	7080
	gtgtcggggc tggcttaact atcgggcacg agagcagatt gtactgagag tgcaccatat	7140
	gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcacaggg gccattcgcc	7200

ES 2 709 216 T3

	attcaggctg cgcaactgtt gggaaggcg atcgggtgcgg gcctcttcgc tattacgcca	7260
	gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag ggttttccca	7320
5	gtcacgacgt tgtaaaacga cggccagtga attcgagctt gcatgcctgc aggt	7374
	<210> 47	
10	<211> 7377	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
15	<223> TalRab2-Reportero	
	<400> 47	
20	cgttacataa cttacggtaa atggcccggc tggctgaccg cccaacgacc cccgcccatt	60
	gacgtcaata atgacgtatg ttcccatagt aacgccaaata gggactttcc attgacgtca	120
	atgggtggag tattttacggt aaactgccc a ttggcagta catcaagtgt atcatatgcc	180
25	aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta	240
	catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	300
	catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	360
30	atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	420
	ggactttcca aaatgtcgta acaactccgc cccattgacg caaatggcg gtaggcgtgt	480
	acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg	540
35	ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct	600
	agaggatccg gtactcgacg aactgcaga gacctacttc actaacaacc ggtatggtcg	660
40	cgagtagctt ggcactggcc gtcgttttac aacgtcgtga ctgggaaaac cctggcgcta	720
	cccaacttaa tcgccttgca gcacatcccc ctttcgccag ctggcgtaat agcgaagagg	780
	ccgcaccga tcgcccttcc caacagttgc gcagcctgaa tggcgaatgg cgctttgcct	840
45	ggtttccggc accagaagcg gtgccggaaa gctggctgga gtgcgatctt cctgaggccg	900
	atactgtcgt cgtcccctca aactggcaga tgcacggtta cgatgcgccc atctacacca	960
	acgtgaccta tccattacg gtcaatccgc cgtttgttcc cacggagaat ccgacgggtt	1020
50	gttactcgct cacatttaat gttgatgaaa gctggctata aaaccggtac agttcggcca	1080
	ccatggtcga tgggtggccg gtagttttca cactcttctc actaccgggc caccacgagt	1140
	agcttggcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa	1200
55	cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggccgc	1260
	accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgctt tgcctggttt	1320
60	ccggcaccag aagcgggtgc ggaaagctgg ctggagtgcg atcttcctga ggccgatact	1380
	gtcgtcgtcc cctcaaactg gcagatgcac ggttacgatg cgcccatcta caccaacgtg	1440
65		

ES 2 709 216 T3

	acctatccca ttacggtcaa tccgcggttt gttcccacgg agaatccgac gggttgttac	1500
	tcgctcacat ttaatgttga tgaaagctgg ctacaggaag gccagacgcg aattatTTTT	1560
5	gatggcggtta actcggcggtt tcatctgtgg tgcaacgggc gctgggtcgg ttacggccag	1620
	gacagtcggt tgccgtctga atttgacctg agcgcatTTT tacgcgccgg agaaaaccgc	1680
10	ctcgcggtga tgggtctgcg ctggagtgcg ggcagttatc tggaagatca ggatatgtgg	1740
	cggatgagcg gcattttccg tgacgtctcg ttgctgcata aaccgactac acaaatcagc	1800
	gatttccatg ttgccactcg ctttaatatg gatttcagcc gcgctgtact ggaggctgaa	1860
15	gttcagatgt gcggcgagtt gcgtgactac ctacgggtaa cagtttcttt atggcaggggt	1920
	gaaacgcagg tcgccagcgg caccgcgcct ttcggcggtg aaattatcga tgagcgtggt	1980
	ggttatgccg atcgcgtcac actacgtctg aacgtcgaaa acccgaaact gtggagcgcc	2040
20	gaaatcccg atctctatcg tgcggtggtt gaactgcaca ccgccgacgg cacgctgatt	2100
	gaagcagaag cctgcgatgt cggtttccgc gaggtgcgga ttgaaaatgg tctgctgctg	2160
25	ctgaacggca agccgttgct gattcgaggc gttaaccgtc acgagcatca tcctctgcat	2220
	ggtcaggtca tggatgagca gacgatggtg caggatatcc tgctgatgaa gcagaacaac	2280
	tttaacgcg tcgctgttgc gcattatccg aaccatccgc tgtggtacac gctgtgcgac	2340
30	cgctacggcc tgtatgtggt ggatgaagcc aatattgaaa cccacggcat ggtgccaatg	2400
	aatcgtctga ccgatgatcc gcgctggcta ccggcgatga gcgaacgcgt aacgcgaatg	2460
	gtgcagcgcg atcgtaatca cccgagtgtg atcatctggt cgctggggaa tgaatcaggc	2520
35	cacggcgcta atcacgacgc gctgtatcgc tggatcaa atctgtcgatcc ttcccgcccc	2580
	gtgcagtatg aaggcgggcg agccgacacc acggccaccg atattatttg cccgatgtac	2640
40	gcgcgcgtgg atgaagacca gcccttcccc gctgtgccga aatgggtccat caaaaaatgg	2700
	ctttcgctac ctggagagac gcgcccgctg atcctttgcg aatacgcca cgcgatgggt	2760
	aacagtcttg gcggtttcgc taaatactgg caggcgtttc gtcagtatcc ccgtttacag	2820
45	ggcggcttcg tctgggactg ggtggatcag tcgctgatta aatatgatga aaacggcaac	2880
	ccgtggtcgg cttacggcgg tgattttggc gatacgccga acgatcgcca gttctgtatg	2940
	aacggtcttg tctttgccga ccgcacgccg catccagcgc tgacggaagc aaaacaccag	3000
50	cagcagtttt tccagttccg tttatccggg caaaccatcg aagtgaccag cgaatacctg	3060
	ttccgtcata gcgataacga gctcctgcac tggatggtgg cgctggatgg taagccgctg	3120
	gcaagcggtg aagtgcctct ggatgtcgct ccacaaggta aacagttgat tgaactgcct	3180
55	gaactaccgc agccggagag cgccgggcaa ctctggctca cagtacgcgt agtgaaccg	3240
	aacgcgaccg catggtcaga agccgggcac atcagcgcct ggcagcagtg gcgtctggcg	3300

60

65

ES 2 709 216 T3

	gaaaacctca	gtgtgacgct	ccccgccg	tcccacgcca	tcccgcacct	gaccaccagc	3360
	gaaatggatt	tttgcacga	gctgggtaat	aagcggttgc	aatttaaccg	ccagtcaggc	3420
5	tttctttcac	agatgtggat	tggcgataaa	aaacaactgc	tgacgccgct	gcgcgatcag	3480
	ttcacccgtg	caccgctgga	taacgacatt	ggcgtaagt	aagcgacccg	cattgaccct	3540
	aacgcctggg	tcgaacgctg	gaaggcgccg	ggccattacc	agggcgaagc	agcgttggtg	3600
10	cagtgcacgg	cagatacact	tgctgatg	gtgctgatta	cgaccgctca	cgcgtaggag	3660
	catcagggga	aaaccttatt	tatcagccgg	aaaacctacc	ggattgatgg	tagtggtcaa	3720
15	atggcgatta	ccgttgatgt	tgaagtggcg	agcgatacac	cgcatccggc	gcggattggc	3780
	ctgaactgcc	agctggcgca	ggtagcagag	cgggtaaact	ggctcggatt	agggccgcaa	3840
	gaaaactatc	ccgaccgcct	tactgccgcc	tgttttgacc	gctgggatct	gccattgtca	3900
20	gacatgtata	ccccgtacgt	cttcccagac	gaaaacggtc	tgcgctgcgg	gacgcgcgaa	3960
	ttgaattatg	gcccacacca	gtggcgccgg	gacttccagt	tcaacatcag	ccgctacagt	4020
	caacagcaac	tgatggaaac	cagccatcgc	catctgctgc	acgcggaaga	aggcacatgg	4080
25	ctgaatatcg	acggtttcca	tatggggatt	ggtagcgacg	actcctggag	cccgtcagta	4140
	tcggcggaat	tacagctgag	cgccggctgc	taccattacc	agttggctctg	gtgtcaaaaa	4200
30	taataataac	cgggcaggcc	atgtctgccc	gtatttcg	taaggaaatc	cattatgtac	4260
	tatttaaaaa	acacaaactt	ttggatgttc	ggtttattct	ttttctttta	cttttttacc	4320
	atgggagcct	acttcccggt	tttcccgatt	tggctacatg	acatcaacca	tatcagcaaa	4380
35	agtgatacgg	gtattatatt	tgcgctatt	tctctgttct	cgctattatt	ccaaccgctg	4440
	tttggctctg	tttctgacaa	actcgccctc	gactctaggg	ggccgcgggg	atccagacat	4500
	gataagatac	attgatgagt	ttggacaaac	cacaactaga	atgcagtga	aaaaatgctt	4560
40	tattttgtgaa	attttgtgat	ctattgcttt	atttgaacc	attataagct	gcaataaaca	4620
	agttaacaac	aacaattgca	ttcattttat	gtttcagggt	cagggggagg	tgtgggaggt	4680
45	tttttcggat	cctctagagt	cgacctgcag	gcatgcaagc	ttggcgtaat	catggtcata	4740
	gctgtttcct	gtgtgaaatt	gttatccgct	cacaattcca	cacaacatac	gagccggaag	4800
	cataaagtgt	aaagcctggg	gtgcctaatt	agtgcgctaa	ctcacattaa	ttgcgttg	4860
50	ctcactgccc	gctttccagt	cgggaaacct	gtcgtgccag	ctgcattaat	gaatcgcca	4920
	acgcgcgggg	agaggcggtt	tgcgtattgg	gcgctcttcc	gcttcctcgc	tcactgactc	4980
	gctgcgctcg	gtcgttcggc	tgcggcgagc	ggtagcagct	cactcaaagg	cggtaatacg	5040
55	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaaag	gccagcaaaa	5100
	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	5160
60	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	gactataaag	5220

ES 2 709 216 T3

	ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	5280
	taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	5340
5	ctgtagggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgcacgaacc	5400
	ccccgttcag	cccgaccgct	gcgccttata	cggtaactat	cgtcttgagt	ccaacccggt	5460
10	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	5520
	tgtaggcggt	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	5580
	agtatttggg	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	5640
15	ttgatccggc	aaacaaacca	ccgctggtag	cgggtggtttt	tttgtttgca	agcagcagat	5700
	tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	5760
	tcagtggaac	gaaaactcac	gttaagggat	tttggtcatg	agattatcaa	aaaggatctt	5820
20	cacctagatc	cttttaaatt	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	5880
	aacttgggtct	gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	5940
25	atttcgttca	tccatagttg	cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	6000
	cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	6060
	tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	6120
30	atccgcctcc	atccagtcta	ttaattgttg	cggggaagct	agagtaagta	gttcgccagt	6180
	taatagtttg	cgcaacgttg	ttgccattgc	tacaggcatc	gtggtgtcac	gctcgtcggt	6240
35	tggtatggct	tcattcagct	ccggttccca	acgatcaagg	cgagttacat	gatcccccat	6300
	gttgtgcaaa	aaagcggtta	gctccttcgg	tcctccgata	gttgtcagaa	gtaagttggc	6360
	cgcagtgtta	tcactcatgg	ttatggcagc	actgcataat	tctcttactg	tcatgccatc	6420
40	cgtaagatgc	ttttctgtga	ctggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	6480
	gcggcgaccg	agttgctctt	gcccggcgtc	aatacgggat	aataccgcgc	cacatagcag	6540
	aactttaaaa	gtgctcatca	ttggaacacg	ttcttcgggg	cgaaaactct	caaggatctt	6600
45	accgctgttg	agatccagtt	cgatgtaacc	cactcgtgca	cccaactgat	cttcagcatc	6660
	ttttactttc	accagcgttt	ctgggtgagc	aaaaacagga	aggcaaaatg	ccgcaaaaaa	6720
50	gggaataagg	gcgacacgga	aatgttgaat	actcatactc	ttcctttttc	aatattattg	6780
	aagcatttat	caggggttatt	gtctcatgag	cggatacata	tttgaatgta	tttagaaaaa	6840
	taaacaaata	ggggttccgc	gcacatttcc	ccgaaaagtg	ccacctgacg	tctaagaaac	6900
55	cattattatc	atgacattaa	cctataaaaa	taggcgtatc	acgaggccct	ttcgtctcgc	6960
	gcgtttcggg	gatgacgggtg	aaaacctctg	acacatgcag	ctcccgga	cggtcacagc	7020
60	ttgtctgtaa	gcggatgccg	ggagcagaca	agcccgtcag	ggcgcgtcag	cgggtgttgg	7080

ES 2 709 216 T3

	cgggtgtcgg ggctggctta actatgcggc atcagagcag attgtactga gagtgcacca	7140
	tatgcgggtgt gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggcgccattc	7200
5	gccattcagg ctgcgcaact gttgggaagg gcgatcggtg cgggcctctt cgctattacg	7260
	ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	7320
	ccagtcacga cgttgtaaaa cgacggccag tgaattcgag cttgcatgcc tgcaggt	7377
10	<210> 48 <211> 7383 <212> ADN <213> Secuencia artificial	
15	<220> <223> ArtTal1/TalRab2-Reportero	
20	<400> 48	
	cgttacataa cttacggtaa atggcccgcc tggttgaccg cccaacgacc cccgccatt	60
	gacgtcaata atgacgtatg ttcccatagt aacgccataa gggactttcc attgacgtca	120
25	atgggtggag tatttacggg aaactgccca cttggcagta catcaagtgt atcatatgcc	180
	aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta	240
	catgacctta tgggactttc ctacttgga gtacatctac gtattagtca tcgctattac	300
30	catgggtgatg cggtttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	360
	atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	420
35	ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt	480
	acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg	540
	ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct	600
40	agaggatccg gtactcgagg aactgcaga gacctacttc actaacaacc ggtatggctg	660
	cgagtagctt ggcactggcc gtgcgttttac aacgtcgtga ctgggaaaac cctggcggtta	720
45	cccaacttaa tcgccttgca gcacatcccc ctttcgccag ctggcgtaat agcgaagagg	780
	cccgaccga tcgccttcc caacagttgc gcagcctgaa tggcgaaatg cgctttgcct	840
	ggtttccggc accagaagcg gtgccgaaa gctggctgga gtgcgatctt cctgaggccg	900
50	atactgtcgt cgtcccctca aactggcaga tgcacgggta cgatgcgccc atctacacca	960
	acgtgacctt tcccattacg gtcaatccgc cgtttgttcc cacggagaat ccgacggggt	1020
	gttactcgct cacatttaat gttgatgaaa gctggctata aaaccgggtac agttcggcca	1080
55	ccatggtcgt attctgggac gtttttcaca ctcttctaaa ctaccgggac accacgggtc	1140
	gcgagtagct tggcactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcggt	1200
60	acccaactta atgccttgc agcacatccc ctttcgccca gctggcgtaa tagcgaagag	1260
	gcccgcaccg atgccttcc ccaacagttg gcgagcctga atggcgaaatg gcgctttgcc	1320

ES 2 709 216 T3

	tggtttccgg caccagaagc ggtgccggaa agctggctgg agtgcgatct tccctgaggcc	1380
	gatactgtcg tcgtcccctc aaactggcag atgcacggtt acgatgcgcc catctacacc	1440
5	aacgtgacct atcccattac ggtcaatccg ccgtttgttc ccacggagaa tccgacgggt	1500
	tgttactcgc tcacatttaa tgttgatgaa agctggctac aggaaggcca gacgcgaatt	1560
10	atTTTTgatg gcgttaactc ggcgtttcat ctgtggtgca acgggcgctg ggtcggttac	1620
	ggccaggaca gtcgtttgcc gtctgaattt gacctgagcg cttttttacg cgccggagaa	1680
	aaccgcctcg cggatgatgg gctgcgctgg agtgacggca gttatctgga agatcaggat	1740
15	atgtggcgga tgagcggcat tttccgtgac gtctcgttgc tgcataaacc gactacacaa	1800
	atcagcgatt tccatgttgc cactcgcttt aatgatgatt tcagccgcgc tgtactggag	1860
	gctgaagttc agatgtgcgg cgagttgcgt gactacctac gggtaacagt ttctttatgg	1920
20	cagggtgaaa cgcaggtcgc cagcggcacc ggccttttcg gcggtgaaat tatcgatgag	1980
	cgtggtggtt atgccgatcg cgtcacacta cgtctgaacg tcgaaaaccc gaaactgtgg	2040
25	agcgcgaaa tcccgaatct ctatcgtcgc gtggttgaaac tgcacaccgc cgacggcacg	2100
	ctgattgaag cagaagcctg cgatgtcggg ttccgcgagg tgcggattga aaatggtctg	2160
	ctgctgctga acggcaagcc gttgctgatt cgaggcggtta accgtcacga gcatcatcct	2220
30	ctgcatggtc aggtcatgga tgagcagacg atggtgcagg atatcctgct gatgaagcag	2280
	aacaacttta acgccgtcgc ctgttcgcat tatccgaacc atccgctgtg gtacacgctg	2340
35	tgcgaccgct acggcctgta tgtggtggat gaagccaata ttgaaacca cggcatggtg	2400
	ccaatgaatc gtctgaccga tgatccgcgc tggctaccgg cgatgagcga acgcgtaacg	2460
	cgaatggtgc agcgcgatcg taatcaccgg agtgtgatca tctggtcgct ggggaatgaa	2520
40	tcagggcacg gcgctaatac cgacgcgctg tatcgctgga tcaaactctgt cgatccttcc	2580
	cgcccgtgtc agtatgaagg cggcggagcc gacaccacgg ccaccgatat tttttgcccg	2640
	atgtacgcgc gcgtggatga agaccagccc ttcccggtg tgcgaaatg gtccatcaaa	2700
45	aaatggcttt cgctacctgg agagacgcgc ccgctgatcc tttgcgaata cgcccacgcg	2760
	atgggtaaca gtcttgccgg tttcgctaaa tactggcagg cgtttcgtca gtatccccgt	2820
50	ttacagggcg gcttcgtctg ggactgggtg gatcagtcgc tgattaaata tgatgaaaac	2880
	ggcaaccggt ggtcggctta cggcggatgat tttggcgata cgccgaacga tcgccagttc	2940
	tgtatgaacg gtctggtctt tgccgaccgc acgccgcac cagcgtgac ggaagcaaaa	3000
55	caccagcagc agtttttcca gttccgttta tccgggcaaa ccatcgaagt gaccagcgaa	3060
	tacctgttcc gtcatagcga taacgagctc ctgcactgga tgggtggcgct ggatggtaag	3120
60	ccgctggcaa gcggtgaagt gcctctggat gtcgctccac aaggtaaaca gttgattgaa	3180

ES 2 709 216 T3

	ctgcctgaac taccgcagcc ggagagcgcc gggcaactct ggctcacagt acgcgtagtg	3240
	caaccgaacg cgaccgcatg gtcagaagcc gggcacatca gcgcctggca gcagtggcgt	3300
5	ctggcggaag acctcagtgt gacgctcccc gccgcgtccc acgccatccc gcatctgacc	3360
	accagcgaaa tggatttttg catcgagctg ggtaataaagc gttggcaatt taaccgccag	3420
10	tcaggctttc tttcacagat gtggattggc gataaaaaac aactgctgac gccgctgcgc	3480
	gatcagttca cccgtgcacc gctggataac gacattggcg taagtgaagc gacccgcatt	3540
	gaccctaacg cctgggtcga acgctggaag gggcggggcc attaccaggc cgaagcagcg	3600
15	ttgttgcaagt gcacggcaga tacacttgct gatgcggtgc tgattacgac cgctcacgcg	3660
	tggcagcatc aggggaaaac cttattttatc agccggaaaa cctaccggat tgatggtagt	3720
	ggtcaaattg cgattaccgt tgatgttgaa gtggcgagcg atacaccgca tccggcgcg	3780
20	attggcctga actgccagct ggcgcaggta gcagagcggg taaactggct cggattaggg	3840
	ccgcaagaaa actatcccga ccgccttact gccgcctgtt ttgaccgctg ggatctgcc	3900
25	ttgtcagaca tgtatacccc gtacgtcttc ccgagcgaag acggtctgcg ctgcgggacg	3960
	cgcgaaattga attatggccc acaccagtgg cgcggcgact tccagttcaa catcagccgc	4020
	tacagtcaac agcaactgat ggaaaccagc catcgccatc tgctgcacgc ggaagaaggc	4080
30	acatggctga atatcgacgg tttccatatg gggattggtg gcgacgactc ctggagcccg	4140
	tcagtatcgg cggaattaca gctgagcgcc ggtcgctacc attaccagtt ggtctggtgt	4200
	caaaaaataat aataaccggg caggccatgt ctgcccgat ttcgcgtaag gaaatccatt	4260
35	atgtactatt taaaaaacac aaacttttgg atgttcggtt tattcttttt cttttacttt	4320
	tttatcatgg gagcctactt cccgtttttc ccgatttggc tacatgacat caaccatctc	4380
	agcaaaagtg atacgggtat tatttttgcc gctatttctc tgttctcgct attattccaa	4440
40	ccgctgtttg gtctgctttc tgacaaaactc ggccctcgact ctaggcggcc gcggggatcc	4500
	agacatgata agatacattg atgagtttgg acaaaccaca actagaatgc agtgaaaaaa	4560
45	atgctttatt tgtgaaattt gtgatgctat tgcttttatt gtaaccatta taagctgcaa	4620
	taaacaagtt aacaacaaca attgcattca ttttatgttt caggttcagg gggaggtgtg	4680
	ggaggttttt tcggtacctc tagagtcgac ctgcaggcat gcaagcttgg cgtaatcatg	4740
50	gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc	4800
	cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg agctaactca cattaattgc	4860
	gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc attaatgaat	4920
55	cggccaaacg gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac	4980
	tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcgg	5040
60	aatacgggtta tccacagaat caggggataa gcgaggaaag aacatgtgag caaaaggcca	5100

ES 2 709 216 T3

	gcaaaaggcc	aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	5160
	ccctgacgag	catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	5220
5	ataaagatac	caggcgtttc	cccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	5280
	gccgcttacc	ggatacctgt	ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	5340
10	ctcacgctgt	aggtatctca	gttcgggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	5400
	cgaaccccc	gttcagcccc	accgctgcgc	cttatccggt	aactatcgtc	ttgagtccaa	5460
	cccggtaaga	cacgacttat	cgccactggc	agcagccact	ggtaacagga	ttagcagagc	5520
15	gaggatatgta	ggcggtgcta	cagagttctt	gaagtgggtg	cctaactacg	gctacactag	5580
	aaggacagta	tttggatatct	gcgctctgct	gaagccagtt	accttcggaa	aaagagttgg	5640
20	tagctcttga	tccggcaaac	aaaccaccgc	tggtagcggg	ggtttttttg	tttgcaagca	5700
	gcagattacg	cgcagaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	5760
	tgacgctcag	tggaaacgaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	5820
25	gatcttcacc	tagatccttt	taaattaaaa	atgaagtttt	aatcaatct	aaagtatata	5880
	tgagtaaact	tgggtctgaca	gttaccaatg	cttaatcagt	gaggcaccta	tctcagcgat	5940
30	ctgtctattt	cgttcatcca	tagttgcctg	actccccgtc	gtgtagataa	ctacgatacg	6000
	ggagggctta	ccatctggcc	ccagtgtgtc	aatgataccg	cgagaccac	gctcaccggc	6060
	tccagattta	tcagcaataa	accagccagc	cgggaagggc	gagcgcagaa	gtggtcctgc	6120
35	aactttatcc	gcctccatcc	agtctattaa	ttggtgccgg	gaagctagag	taagtagttc	6180
	gccagttaat	agtttgcgca	acgttggtgc	cattgctaca	ggcatcgtgg	tgtcacgctc	6240
40	gtcgtttggt	atggcttcat	tcagctccgg	ttcccaacga	tcaaggcgag	ttacatgatc	6300
	ccccatgttg	tgcaaaaaag	cggttagctc	cttcgggtcct	ccgatcgttg	tcagaagtaa	6360
	gttggccgca	gtgttatcac	tcattggttat	ggcagcactg	cataattctc	ttactgtcat	6420
45	gccatccgta	agatgctttt	ctgtgactgg	tgagtactca	accaagtcac	tctgagaata	6480
	gtgtatgcgg	cgaccgagtt	gctcttgccc	ggcgtcaata	cgggataata	ccgcgccaca	6540
50	tagcagaact	ttaaaagtgc	tcattcattgg	aaaacgttct	tcggggcgaa	aactctcaag	6600
	gatcttaccg	ctgttgagat	ccagttcgat	gtaaccact	cgtgcaccca	actgatcttc	6660
	agcatctttt	actttcacca	gcgtttctgg	gtgagcaaaa	acaggaaggc	aaaatgccgc	6720
55	aaaaaaggga	ataagggcga	cacggaaatg	ttgaatactc	atactcttcc	tttttcaata	6780
	ttattgaagc	atttatcagg	gttattgtct	catgagcgga	tacatatattg	aatgtattta	6840
60	gaaaaataaa	caaatagggg	ttccgcgcac	atttccccga	aaagtgccac	ctgacgtcta	6900
	agaaaccatt	attatcatga	cattaaccta	taaaaatagg	cgtatcacga	ggccctttcg	6960

65

ES 2 709 216 T3

	tctcgcgcgt ttcggtgatg acggtgaaaa cctctgacac atgcagctcc cggagacggt	7020
	cacagcttgt ctgtaagcgg atgccgggag cagacaagcc cgtcagggcg cgtcagcggg	7080
5	tggtggcggg tgcggggct ggcttaacta tgcggcatca gagcagattg tactgagagt	7140
	gcaccatatg cgggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg	7200
10	ccattcgcca ttcagcgtgc gcaactgttg ggaagggcga tcggtgcggg cctcttcgct	7260
	attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg taacgccagg	7320
	gttttcccag tcacgacgtt gtaaacgac ggccagtga ttcgagcttg catgcctgca	7380
15	ggt	7383
	<210> 49	
	<211> 5566	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> pCMV-hLuc	
25	<400> 49	
	ggtaccgagc tcttacgcgt gctagcccg gctcaggag cttggcccat tgcatacgtt	60
30	gtatccatat cataatatgt acatttatat tggctcatgt ccaacattac cgccatgttg	120
	acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcataagccc	180
	atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccac	240
35	cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac	300
	tttccattga cgtcaatggg tggagtattt acgctaaact gccacttgg cagtacatca	360
	agtgtatcat atgccaaagta cgcctctat tgacgtcaat gacggtaaata ggccgcctg	420
40	gcattatgcc cagtacatga ccttatggga ctttcctact tggcagtaca tctacgtatt	480
	agtcacgcgt attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg	540
45	gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga gtttggtttg	600
	gcacccaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaata	660
	gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctcgtttag tgaaccgtca	720
50	gatgcctgg agacgccatc cacgctgttt tgacctccat agaagacacc gggaccgatc	780
	cagcctccgc ggccccgaat tagcttggca ttccggtact gttggtaaag ccaccatgga	840
55	agacgccaaa aacataaaga aaggcccggc gccattctat ccgctggaag atggaaccgc	900
	tggagagcaa ctgcataagg ctatgaagag atacgccctg gttcctggaa caattgcttt	960
	tacagatgca catatcgagg tggacatcac ttacgtgag tacttcgaaa tgtccgttcg	1020
60	gttggcagaa gctatgaaac gatatgggct gaatacaaat cacagaatcg tcgtatgcag	1080
	tgaaaactct cttcaattct ttatgccggt gttgggcgcg ttatttatcg gagttgcagt	1140
65		

ES 2 709 216 T3

	tgcgcccgcg aacgacattt ataatgaacg tgaattgctc aacagtatgg gcatttcgca	1200
	gcctaccgtg gtgttcgttt ccaaaaaggg gttgcaaaaa attttgaacg tgcaaaaaaa	1260
5	gctcccaatc atccaaaaaa ttattatcat ggattctaaa acggattacc agggatttca	1320
	gtcgatgtac acgttcgtca catctcatct acctcccggt tttaatgaat acgattttgt	1380
10	gccagagtcc ttcgataggg acaagacaat tgcactgata atgaactcct ctggatctac	1440
	tggtctgcct aaaggtgtcg ctctgcctca tagaactgcc tgcgtgagat tctcgcgcatgc	1500
	cagagatcct atttttggca atcaaatcat tccggatact gcgattttta gtgttggtcc	1560
15	attccatcac ggttttggaa tgtttactac actcggatat ttgatattgt gatttcagat	1620
	cgtcttaatg tatagatttg aagaagagct gtttctgagg agccttcagg attacaagat	1680
	tcaaagtgcg ctgctggtgc caaccctatt ctcttctctc gccaaaagca ctctgattga	1740
20	caaatacgat ttatctaatt tacacgaaat tgcttctggt ggcgctcccc tctctaagga	1800
	agtcggggaa gcggttgcca agaggttcca tctgccaggt atcaggcaag gatatgggct	1860
25	cactgagact acatcagcta ttctgattac acccgagggg gatgataaac cgggcgcggt	1920
	cggtaaagtt gttccatttt ttgaagcgaa ggttgtggat ctggataacc ggaaaacgct	1980
	gggcgttaat caaagaggcg aactgtgtgt gagaggctct atgattatgt ccggttatgt	2040
30	aaacaatccg gaagcgacca acgccttgat tgacaaggat ggatggctac attctggaga	2100
	catagcttac tgggacgaag acgaacactt cttcatcggt gaccgcctga agtctctgat	2160
35	taagtacaaa ggctatcagg tggctcccgc tgaattggaa tccatcttgc tccaacaccc	2220
	caacatcttc gacgcagggt tgcgaggctt tcccgcgat gacgcgggtg aacttcccgc	2280
	cgccgttggt gttttggagc acggaaagac gatgacggaa aaagagatcg tggattacgt	2340
40	cgccagtc aa gtaacaaccg cgaaaaagtt gcgcggagga gttgtgtttg tggacgaagt	2400
	accgaaaggt cttaccggaa aactcgacgc aagaaaaatc agagagatcc tcataaaggc	2460
	caagaagggc ggaaagatcg ccgtgtaatt ctagagtcgg gccggccggc cgcttcgagc	2520
45	agacatgata agatacattg atgagtttgg acaaaccaca actagaatgc agtgaaaaaa	2580
	atgctttatt tgtgaaattt gtgatgctat tgctttattt gtaaccatta taagctgcaa	2640
50	taaacaagtt aacaacaaca attgcattca ttttatgttt caggttcagg gggagggtgtg	2700
	ggagggtttt taaagcaagt aaaacctcta caaatgtggt aaaatcgata aggatccgtc	2760
	gaccgatgcc cttgagagcc ttcaaccag tcagctcctt ccggtgggcg cggggcatga	2820
55	ctatcgtcgc cgcacttatg actgtcttct ttatcatgca actcgtagga cagggtgccg	2880
	cagcgtctt ccgcttctc gtcactgac tcgctgcgct cggtcgttcg gctgcggcga	2940
60	gcggtatcag ctcaactcaa gccggtaata cggttatcca cagaatcagg ggataacgca	3000

ES 2 709 216 T3

	ggaaagaaca	tgtgagcaaa	aggccagcaa	aagccagga	accgtaaaaa	ggccgcgttg	3060
	ctggcggtttt	tccataggct	ccgccccct	gacgagcatc	acaaaaatcg	acgctcaagt	3120
5	cagaggtggc	gaaacccgac	aggactataa	agataccagg	cgtttcccc	tggaagctcc	3180
	ctcgtgcgct	ctcctgttcc	gaccctgcgc	cttaccggat	acctgtccgc	ctttctccct	3240
10	tcgggaagcg	tgccgctttc	tcatagctca	cgctgtagg	atctcagttc	ggtgtaggtc	3300
	gttcgctcca	agctgggctg	tgtgcacgaa	cccccggtc	agcccgaccg	ctgcgcctta	3360
	tccggtaact	atcgtcttga	gtccaaccgc	gtaagacacg	acttatcgcc	actggcagca	3420
15	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	3480
	tggtggccta	actacggcta	cactagaaga	acagtatttg	gtatctgcgc	tctgctgaag	3540
	ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaaac	caccgctggt	3600
20	agcggtggtt	tttttggttg	caagcagcag	attacgcgca	gaaaaaagg	atctcaagaa	3660
	gatcctttga	tcttttctac	ggggtctgac	gctcagtgga	acgaaaactc	acgttaagg	3720
	attttggtca	tgagattatc	aaaaaggatc	ttcacctaga	tcctttttaa	ttaaaaatga	3780
25	agttttaaat	caatctaaag	tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	3840
	atcagtgagg	cacctatctc	agcgatctgt	ctatttcggt	catccatagt	tgcctgactc	3900
30	cccgtcgtgt	agataactac	gatacgggag	ggcttaccat	ctggccccag	tgctgcaatg	3960
	ataccgcgag	accacgcctc	accggctcca	gatttatcag	caataaacca	gccagccgga	4020
	agggccgagc	gcagaagtgg	tcctgcaact	ttatccgcct	ccatccagtc	tattaattgt	4080
35	tgccgggaag	ctagagtaag	tagttcgcca	gttaatagtt	tgcgcaacgt	tgttgccatt	4140
	gctacaggca	tcgtgggtgc	acgctcgtcg	tttggtatgg	cttcattcag	ctccggttcc	4200
	caacgatcaa	ggcgagttac	atgatcccc	atgttgtgca	aaaaagcgg	tagctccttc	4260
40	ggtcctccga	tcgttgtcag	aagtaagttg	gccgcagtg	tatcactcat	ggttatggca	4320
	gcactgcata	attctcttac	tgtcatgcca	tcgtaagat	gcttttctgt	gactggtgag	4380
	tactcaacca	agtcattctg	agaatagtgt	atgcggcgac	cgagttgctc	ttgccggcg	4440
45	tcaatacggg	ataataccgc	gccacatagc	agaactttaa	aagtgtcat	cattgaaaa	4500
	cgttcttcgg	ggcgaaaact	ctcaaggatc	ttaccgctgt	tgagatccag	ttcgatgtaa	4560
50	cccactcgtg	cacccaactg	atcttcagca	tcttttactt	tcaccagcgt	ttctgggtga	4620
	gcaaaaacag	gaaggcaaaa	tgccgcaaaa	aagggaataa	ggcgacacg	gaaatgttga	4680
	atactcatac	tcttcctttt	tcaatattat	tgaagcattt	atcagggtta	ttgtctcatg	4740
55	agcggataca	tatttgaatg	tatttagaaa	aataaaciaa	taggggttcc	gcgcacattt	4800
	ccccgaaaag	tgccacctga	cgcgccctgt	agcggcgcat	taagcgcggc	gggtgtggtg	4860
60	gttacgcgca	gcgtgaccgc	tacacttgcc	agcgccctag	cgcccgctcc	tttcgctttc	4920

ES 2 709 216 T3

	ttcccttcct ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc	4980
	cctttagggg tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggg	5040
5	gatgggtcac gtagtgggcc atcgccctga tagacggttt ttcgcccttt gacgttgag	5100
	tccacgttct ttaatagtgg actcttggtc caaactggaa caaactcaa ccctatctcg	5160
10	gtctattctt ttgatttata agggattttg ccgatttcgg cctattgggt aaaaaatgag	5220
	ctgatttaac aaaaatttaa cgcgaatttt aacaaaatat taacgcttac aatttgccat	5280
	tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg tcggggcctc ttcgctatta	5340
15	cgccagccca agctaccatg ataagtaagt aatattaagg tacgggaggt acttgagcgc	5400
	gccgaataa aatatcttta ttttcattac atctgtgtgt tgggtttttg tgtgaatcga	5460
20	tagtactaac atacgctctc catcaaaaca aaacgaaaca aaacaaacta gcaaaatagg	5520
	ctgtccccag tgcaagtgca ggtgccagaa catttctcta tcgata	5566

25 <210> 50
 <211> 2961
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> pBS
 <400> 50

35	gtaaaacgac ggccagttag cgcgcgtaat acgactcact atagggcgaa ttggagctcc	60
	accgcggtgg cgcccgctct agaactagtg gatcccccg gctgcaggaa ttcgatatca	120
40	agcttatcga taccgtcgac ctcgaggggg ggcccggtac ccagcttttg ttcccttag	180
	tgaggggttaa ttgcgcgctt ggcgtaatca tgggtcatagc tgtttcctgt gtgaaattgt	240
	tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctggggg	300
45	gcctaattgag tgagctaact cacattaatt gcgttcgct cactgccgcg tttccagtcg	360
	ggaaacctgt cgtgccagct gcattaatga atcgccaac gcgcggggag aggcggtttg	420
	cgtattgggc gctcttcgcg ttctcgcgc actgactcgc tcgctcgggt cgttcggctg	480
50	cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga atcaggggat	540
	aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc	600
55	gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia aaatcgacgc	660
	tcaagtcaaga ggtggcgaaa cccgacagga ctataaagat accaggcggt tccccctgga	720
	agctccctcg tcgctctctc tgttcgacc ctgccgctta ccggatacct gtccgccttt	780
60	ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcgggtg	840
	taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc	900

65

ES 2 709 216 T3

	gccttatccg gtaactatcg tcttgagtcc aaccggttaa gacacgactt atcgccactg	960
	gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc	1020
5	ttgaagtggg ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg	1080
	ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaacacc	1140
	gctggtagcg gtgggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct	1200
10	caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaaacga aaactcacgt	1260
	taagggattt tgggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa	1320
	aaatgaagtt ttaaatacaat cttaaagtata tatgagtaaa cttgggtctga cagttaccaa	1380
15	tgcttaataca gtgaggcacc tatctcagcg atctgtctat ttcgttcac ctagttgcc	1440
	tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg cccagtgct	1500
	gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca	1560
20	gccggaaggc cgcgagcgcag aagtggctct gcaactttat ccgcctccat ccagtctatt	1620
	aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttggt	1680
	gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc	1740
25	ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa agcgggttagc	1800
	tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt	1860
	atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact	1920
30	ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc	1980
	ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt	2040
35	ggaaaaagctt cttcggggcg aaaaactctca aggatcttac cgctgttgag atccagttcg	2100
	atgtaaccca ctcggtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct	2160
	gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa	2220
40	tggtgaatac tcatactctt cttttttcaa tattattgaa gcatttatca gggttattgt	2280
	ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc	2340
	acatttcccc gaaaagtgcc acctaaattg taagcggttaa tattttgtta aaattcgcgt	2400
45	taaatttttg ttaaatacagc tcatttttta accaataggc cgaaatcggc aaaatccctt	2460
	ataaatcaaa agaataagacc gagatagggt tgagtgttgt tccagtttg aacaagagtc	2520
	cactattaaa gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg	2580
50	gccactacg tgaaccatca ccctaataca gttttttggg gtcgaggtgc cgtaaagcac	2640
	taaatcggaa ccctaagggg agccccgat ttagagcttg acggggaaag ccggcgaaacg	2700
	tggcgagaaa ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg gcaagtgtag	2760
55	cggctcacgt gcgcgtaacc accacacccg ccgcgcttaa tgcgcgcta cagggcgct	2820
	cccattcgcc attcaggctg cgcaactgtt gggaagggcg atcgggtgcg gcctcttcgc	2880
	tattacgccg gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag	2940
60	ggttttccca gtcacgacgt t	2961

ES 2 709 216 T3

<210> 51
 <211> 7164
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> pCMVbeta

<400> 51

5									
10		gaattc	gagc	ttgcat	gcct	gcaggt	cggt	acataa	actta
		cggt	aaatgg	ccgcct	ggc	60			
		tgaccg	ccca	acgacccc	cg	cccatt	gacg	tcaata	aatga
		cgta	tgttcc	catagta	acg	120			
15		ccaata	ggga	ctttcc	attg	acgtca	aatgg	gtggag	tatt
		tacggt	aaac	tgcccact	tg	180			
		gcagta	catc	aagtgt	atca	tatgcca	agt	acgcccc	cta
		ttgacgt	caa	tgacgg	taaa	240			
20		tggccc	gcct	ggcatt	atgc	ccagtaca	tg	accttat	ggg
		actttc	ctac	ttggcag	tac	300			
		atctac	gtat	tagtcat	cg	tattacca	tg	gtgatgc	ggt
		tttggc	agta	catcaat	ggg	360			
		cgtggat	agc	ggtttg	actc	acgggg	at	ccaagt	ctcc
		accccatt	ga	cgtcaat	ggg	420			
25		agtttg	tttt	ggcacca	aaa	tcaacg	ggg	ac	tttccaa
		aat	gtcgt	aacaa	ctccg	cccca	480		
		ttgacg	caaa	tggg	cggtag	gcgtgt	acg	tgggag	gtct
		atataag	cag	agctc	gttta	540			
30		gtgaacc	gct	agatcg	cctg	gagacg	ccat	ccacg	ctgtt
		ttgac	ctcca	tagaag	acac	600			
		cgggacc	gat	ccagc	ctccg	gactcta	gag	gatccg	gtac
		tcgagga	act	gaaaa	accag	660			
		aaagtta	act	ggtaag	ttta	gtcttt	ttgt	ctttt	atttc
		aggtccc	gga	tccggt	ggtg	720			
35		gtgcaa	aatca	aagaact	gct	cctcag	tgga	tgttgc	cttt
		acttcta	ggc	ctgtac	ggaa	780			
		gtgtta	cttc	tgtct	taaaa	gctgcg	gaat	tgtac	ccgcg
		gccgca	attc	ccgggg	atcg	840			
40		aaagag	cctg	ctaaag	caaa	aaaga	agtca	ccatgt	cggt
		tacttt	gacc	aacaag	aacg	900			
		tgatttt	ctgt	tgccgg	tctg	ggaggc	attg	gtctgg	acac
		cagca	aggag	ctgct	caagc	960			
		gcgat	cccgt	cgtttt	tacaa	cgctgt	gact	gggaaa	accc
		tggcgt	tacc	caactta	atc	1020			
45		gccttg	cagc	acatcccc	ct	ttcgcc	agct	ggcgta	aatag
		cgaag	aggcc	cgcacc	gatc	1080			
		gccctt	ccca	acagtt	gcgc	agcctg	aatg	gcgaat	ggcg
		ctttgc	ctgg	tttccg	gcac	1140			
50		cagaag	cgggt	gccgg	aaagc	tggctg	gagt	gcgat	cttcc
		tgagg	ccgat	actgt	cgctg	1200			
		tcccct	caaa	ctggc	agatg	cacggt	tacg	atgcg	cccat
		ctacac	caaac	gtaac	ctatc	1260			
55		ccattac	ggt	caatcc	gccg	tttgtt	ccca	cggaga	aatcc
		gacggg	ttgt	tactcg	ctca	1320			
		cattta	aatgt	tgatga	aagc	tggctac	agg	ccagac	gcga
		attatt	tttgat	ggcg	1380				

ES 2 709 216 T3

	ttaactcggc gtttcatctg tggcgcaacg ggcgctgggt cggttacggc caggacagtc	1440
5	gtttgccgtc tgaatttgac ctgagcgcat ttttacgcgc cggagaaaac cgcctcgcgg	1500
	tgatggtgct gcgttggagt gacggcagtt atctggaaga tcaggatatg tggcggatga	1560
	gcggcatttt ccgtgacgtc tcgttgctgc ataaaccgac tacacaaatc agcgatttcc	1620
10	atgttgccac tcgctttaat gatgatttca gccgcgctgt actggaggct gaagttcaga	1680
	tgtgcggcga gttgcgtgac tacctacggg taacagtttc tttatggcag ggtgaaacgc	1740
	aggtcgccag cggcaccgcg cctttcggcg gtgaaattat cgatgagcgt ggtggttatg	1800
15	ccgatcgcgt cactacgt ctgaacgtcg aaaacccgaa actgtggagc gccgaaatcc	1860
	cgaatctcta tcgtgcggtg gttgaactgc acaccgccga cggcacgctg attgaagcag	1920
20	aagcctgcga tgcggtttc cgcgaggtgc ggattgaaaa tggctcgtc ctgctgaacg	1980
	gcaagccgtt gctgattcga ggcgttaacc gtcacgagca tcatcctctg catggtcagg	2040
	tcatggatga gcagacgatg gtgcaggata tcctgctgat gaagcagaac aactttaacg	2100
25	ccgtgcgctg ttcgcattat ccgaaccatc cgctgtggta cacgctgtgc gaccgctacg	2160
	gcctgtatgt ggtggatgaa gccaatattg aaaccacgg catggtgcc aatgaatcgtc	2220
	tgaccgatga tccgcgctgg ctaccggcga tgagcgaacg cgtaacgcga atggtgcagc	2280
30	gcgatcgtaa tcaccogagt gtgatcatct ggtcgctggg gaatgaatca ggccacggcg	2340
	ctaatacga cgcgctgtat cgctggatca aatctgtcga tccttcccgc ccggtgcagt	2400
35	atgaaggcgg cggagccgac accacggcca ccgatattat ttgccgatg tacgcgcgcg	2460
	tggatgaaga ccagcccttc ccggtgtgc cgaaatggc catcaaaaaa tggctttcgc	2520
	tacctggaga gacgcgccg ctgatccttt gcgaatacgc ccacgcgatg ggtaacagtc	2580
40	ttggcgggtt cgctaaatac tggcagcgct ttcgtcagta tcccgttta caggcgcgct	2640
	tcgtctggga ctgggtggat cagtcgctga ttaaatatga tgaaaacggc aaccctgggt	2700
	cggcttacgg cggtgatttt ggcgatacgc cgaacgatcg ccagttctgt atgaacggtc	2760
45	tggctcttgc cgaccgcacg ccgcataccg cgctgacgga agcaaaacac cagcagcagt	2820
	ttttccagtt ccgtttatcc gggcaaacca tcgaagtac cagcgaatac ctgttccgtc	2880
50	atagcgataa cgagctcctg cactggatgg tggcgctgga tggtaagccg ctggcaagcg	2940
	gtgaagtgcc tctggatgtc gtcacacaag gtaaacagtt gattgaactg cctgaactac	3000
	cgcagccgga gagcgccggg caactctggc tcacagtacg cgtagtgcaa ccgaacgcga	3060
55	ccgatgggtc agaagccggg cacatcagcg cctggcagca gtggcgctctg gcggaaaacc	3120
	tcagtgtgac gctccccgcc gcgtcccacg ccatcccga tctgaccacc agcgaaatgg	3180
60	atttttgcat cgagctgggt aataagcgtt ggcaatttaa ccgccagtca ggctttcttt	3240
	cacagatgtg gattggcgat aaaaaacaac tgctgacgcc gctgcgcgat cagttcaccc	3300

65

ES 2 709 216 T3

	gtgcaccgct	ggataacgac	attggcgtaa	gtgaagcgac	ccgcattgac	cctaacgcct	3360
5	gggtcgaacg	ctggaaggcg	gcggggccatt	accaggccga	agcagcggtg	ttgcagtgca	3420
	cggcagatac	acttgctgat	gcggtgctga	ttacgaccgc	tcacgcgtgg	cagcatcagg	3480
	ggaaaacctt	atttatcagc	cggaaaacct	accggattga	tggtagtggg	caaattggcg	3540
10	ttaccgttga	tgttgaagtg	gcgagcgata	caccgcatcc	ggcgcggatt	ggcctgaact	3600
	gccagctggc	gcaggtagca	gagcgggtaa	actggctcgg	attagggccg	caagaaaact	3660
15	atcccagaccg	ccttactgcc	gcctgttttg	accgctggga	tctgccattg	tcagacatgt	3720
	ataccccgta	cgtcttcccg	agcgaaaacg	gtctgcgctg	cgggacgcgc	gaattgaatt	3780
	atggcccaca	ccagtggcgc	ggcgacttcc	agttcaacat	cagccgctac	agtcaacagc	3840
20	aactgatgga	aaccagccat	cgccatctgc	tgcacgcgga	agaaggcaca	tggtctgaata	3900
	tcgacggttt	ccatatgggg	attggtggcg	acgactcctg	gagcccgtca	gtatcggcgg	3960
	aattacagct	gagcgcgggt	cgctaccatt	accagttggg	ctgggtgtcaa	aaataataat	4020
25	aaccgggag	gccatgtctg	cccgtatttc	gcgtaaggaa	atccattatg	tactatttaa	4080
	aaaaacacaaa	cttttggatg	ttcggtttat	tctttttctt	ttactttttt	atcatgggag	4140
30	cctacttccc	gtttttcccg	atttggctac	atgacatcaa	ccatatcagc	aaaagtgata	4200
	cgggtattat	ttttgccgct	atttctctgt	tctcgctatt	attccaaccg	ctgtttggtc	4260
	tgttttctga	caaactcggc	ctcgactcta	ggcggccgcg	gggatccaga	catgataaga	4320
35	tacattgatg	agtttggaca	aaccacaact	agaatgcagt	gaaaaaaatg	ctttatttgt	4380
	gaaatttgtg	atgctattgc	tttatttcta	accattataa	gctgcaataa	acaagttaac	4440
40	aacaacaatt	gcattcatth	tatgtttcag	gttcaggggg	aggtgtggga	ggtttttttcg	4500
	gatcctctag	agtcgacctg	caggcatgca	agcttggcgt	aatcatggtc	atagctgttt	4560
	cctgtgtgaa	attgttatcc	gctcacaatt	ccacacaaca	tacgagccgg	aagcataaag	4620
45	tgtaaagcct	ggggtgccta	atgagtgagc	taactcacat	taattgcgtt	gcgctcactg	4680
	cccgcctttcc	agtcgggaaa	cctgtcgtgc	cagctgcatt	aatgaatcgg	ccaacgcgcg	4740
	gggagaggcg	gtttgcgtat	tgggcgctct	tccgcttcct	cgctcactga	ctcgctgcgc	4800
50	tcggtcgttc	ggctgcggcg	agcggatatca	gctcactcaa	aggcggtaat	acggttatcc	4860
	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	aaaggccagg	4920
55	aaccgtaaaa	aggccgcgtt	gctggcggtt	ttccataggc	tccgcccccc	tgacgagcat	4980
	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaiaaccga	caggactata	aagataccag	5040
	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgaccctgcc	gcttaccgga	5100
60	tacctgtccg	cctttctccc	ttcggaagc	gtggcgcttt	ctcatagctc	acgctgtagg	5160

ES 2 709 216 T3

	tatctcagtt	cggtgtaggt	cgttcgctcc	aagctgggct	gtgtgcacga	acccccggt	5220
5	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	agccaaccc	ggtaagacac	5280
	gacttatcgc	cactggcagc	agccactggg	aacaggatta	gcagagcgag	gtatgtaggc	5340
	gggtctacag	agttcttgaa	gtggtggcct	aactacggct	acactagaag	gacagtattt	5400
10	ggatatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	ctcttgatcc	5460
	ggcaaaacaaa	ccaccgctgg	tagcgggtgg	ttttttgttt	gcaagcagca	gattacgcgc	5520
	agaaaaaag	gatctcaaga	agatcctttg	atcttttcta	cggggtctga	cgctcagtgg	5580
15	aacgaaaact	cacgttaagg	gattttggtc	atgagattat	caaaaaggat	cttcacctag	5640
	atccttttaa	attaaaaatg	aagttttaaa	tcaatctaaa	gtatatatga	gtaaacttgg	5700
	tctgacagtt	accaatgctt	aatcagttag	gcacctatct	cagcgatctg	tctatttcgt	5760
20	tcatccatag	ttgcctgact	ccccgtcgtg	tagataacta	cgatacggga	gggcttacca	5820
	tctggcccca	gtgctgcaat	gataccgcga	gaccacgct	caccggctcc	agatttatca	5880
25	gcaataaacc	agccagccgg	aagggccgag	cgcagaagtg	gtcctgcaac	tttatccgcc	5940
	tccatccagt	ctattaattg	ttgccgggaa	gctagagtaa	gtagttcgcc	agttaatagt	6000
	ttgcgcaacg	ttgttgccat	tgctacaggc	atcgtggtgt	cacgctcgtc	gtttggtatg	6060
30	gcttcattca	gctccggttc	ccaacgatca	aggcgagtta	catgatcccc	catgttgctg	6120
	aaaaaagcgg	ttagctcctt	cggtcctccg	atcgttgtca	gaagtaagtt	ggccgcagtg	6180
	ttatcactca	tggttatggc	agcactgcat	aattctctta	ctgtcatgcc	atccgtaaga	6240
35	tgcttttctg	tgactgggtga	gtactcaacc	aagtcattct	gagaatagtg	tatgcggcga	6300
	ccgagttgct	cttgcccggc	gtcaatacgg	gataatacgg	cgccacatag	cagaacttta	6360
	aaagtgctca	tatttgaaa	acgttcttcg	gggcgaaaac	tctcaaggat	cttaccgctg	6420
40	ttgagatcca	gttcgatgta	acccactcgt	gcacccaact	gatcttcagc	atcttttact	6480
	ttcaccagcg	tttctgggtg	agcaaaaaca	ggaaggcaaa	atgccgcaaa	aaagggaata	6540
	agggcgacac	ggaaatgttg	aatactcata	ctcttccttt	ttcaatatta	ttgaagcatt	6600
45	tatcagggtt	attgtctcat	gagcggatac	atatttgaat	gtatttagaa	aaataaacia	6660
	ataggggttc	cgcgcacatt	tccccgaaaa	gtgccacctg	acgtctaaga	aaccattatt	6720
50	atcatgacat	taacctataa	aataggcgt	atcacgaggc	cctttcgtct	cgcgcgtttc	6780
	gggtgatgacg	gtgaaaacct	ctgacacatg	cagctcccgg	agacggtcac	agcttgctg	6840
	taagcggatg	ccgggagcag	acaagcccgt	cagggcgctg	cagcgggtgt	tggcgggtgt	6900
55	cggggctggc	ttactatgc	ggcatcagag	cagattgtac	tgagagtgca	ccatatgcgg	6960
	tgtgaaatac	cgcacagatg	cgtaaggaga	aaataccgca	tcaggcgcca	ttcgccattc	7020
	aggctgcgca	actgttgga	agggcgatcg	gtgcgggcct	cttcgctatt	acgccagctg	7080
60	gcgaaagggg	gatgtgctgc	aaggcgatta	agttgggtaa	cgccagggtt	ttcccagtca	7140
	cgacgttgta	aaacgacggc	cagt				7164

65

ES 2 709 216 T3

<210> 52
<211> 7867
<212> ADN

<213> Secuencia artificial

<220>
<223> pCAG-TalRab2-Clo051

<400> 52

5	ggcgcgccgg attcgacatt gattattgac tagttattaa tagtaatcaa ttacgggggtc	60
15	attagttcat agcccatata tggagttccg cgttacataa cttacggtaa atggcccgcc	120
	tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg ttcccatagt	180
20	aacgccaata gggactttcc attgacgtca atgggtggag tatttacggt aaactgcccc	240
	cttggcagta catcaagtgt atcatatgcc aagtacgccc cctattgacg tcaatgacgg	300
	taaatggccc gcctggcatt atgccagta catgacctta tgggactttc ctacttggca	360
25	gtacatctac gtattagtca tcgctattac catggtcgag gtgagcccca cgttctgctt	420
	cactctcccc atctcccccc cctccccacc cccaattttg tattttattta ttttttaatt	480
	atthttgtgca gcgatggggg cggggggggg gggggggcgc gcgccaggcg gggcggggcg	540
30	gggcgagggg cggggcgggg cgaggcgagg aggtgcggcg gcagccaatc agagcggcg	600
	gctccgaaag tttcctttta tggcgaggcg gcggcgggcg cggccctata aaaagcgaag	660
35	cgcgcggcgg gcgggagtcg ctgcgcgctg ccttcgcccc gtgccccgct ccgcgcgcgc	720
	ctcgcgcgcg ccgccccggc tctgactgac cgcgttactc ccacaggtga gcgggcggga	780
	cggcccttct cctccgggct gtaattagcg cttgggtttaa tgacggcttg tttcttttct	840
40	gtggctgctg gaaagccttg aggggctccg ggagggccct ttgtgcgggg gggagcggt	900
	cggggggtgc gtgctgtgtg gtgtgcgtgg ggagcgccgc gtgcggctcc gcgctgccc	960
45	gcggctgtga gcgctgcggg cgcggcgcgg ggctttgtgc gctccgcagt gtgcgcgagg	1020
	ggagcgcggc cgggggcggg gccccgcggt gcgggggggg ctgcgagggg aacaaaggct	1080
	gcgtgcgggg tgtgtgcgtg ggggggtgag cagggggtgt gggcgctcg gtcgggctgc	1140
50	aacccccctt gcacccccct ccccgagttg ctgagcacgg cccggcttcg ggtgcggggc	1200
	tccgtacggg gcgtggcgcg gggctcgccc tgccggggcg ggggtggcgg caggtggggg	1260
	tgccggggcg ggcggggccg cctcggggcg gggagggctc gggggagggg cgcggcgggc	1320
55	cccggagcgc cggcggtgt cgaggcgcg cgagccgcag ccattgcctt ttatggtaat	1380
	cgtgcgagag ggcgcaggga cttcctttgt cccaaatctg tgccggagccg aaatctggga	1440
60		
65		

ES 2 709 216 T3

5	ggcgccgccc	cacccctct	agcgggccc	ggcggaagc	gtcgccccc	ggcaggaag	1500
	aaatgggccc	ggagggcctt	cgtgcgtcgc	cgcgcgccc	tcccttctc	cctctccagc	1560
	ctcggggctg	tccgcgggg	gacggctgcc	ttcgggggg	acggggcagg	gcgggggttc	1620
10	gcttctggcg	tgtgaccggc	ggctctagag	cctctgctaa	ccatgttcat	gccttcttct	1680
	ttttcctaca	gacctaata	taataatacg	actcactata	ggggccgcca	ccatgggacc	1740
	taagaaaaag	aggaaggtgg	cggccgctga	ctacaaggat	gacgacgata	aaccaggtgg	1800
15	cggaggtagt	ggcgaggtg	gggtaccgc	cagtccagca	gccaggtgg	atctgagaac	1860
	cctcggctac	agccagcagc	agcaggagaa	gatcaaacca	aaggtgcggt	ccaccgtcgc	1920
	tcagcaccat	gaagcactgg	tggggcacgg	tttcacacac	gcccattattg	tggctctgtc	1980
20	tcagcatccc	gctgcactcg	ggactgtggc	cgtcaaata	caggacatga	tcgcgctct	2040
	gcctgaggca	accacgaag	ccattgtggg	cgtcggaaag	cagtggagcg	gtgccagagc	2100
	actcgaagca	ctcctcaccg	tcgccgggga	actgcggggt	ccaccactcc	agtcgggact	2160
25	ggacactgga	cagctgctga	agatcgctaa	acgcggcgga	gtgacagctg	tggagctgt	2220
	gcacgcttgg	aggaatgctc	tgacaggagc	cccactgaat	ctgacacccc	agcaggtggt	2280
	ggccattgct	agcaacaatg	ggggcaagca	ggctctggag	acagtgcagc	gcctgctgcc	2340
30	tgtgctgtgc	caggctcacg	gactgactcc	acagcaggtg	gtggccatcg	cttccaacaa	2400
	tggagggaaa	caggctctgg	aaacagtgca	gaggctgctg	cccgtgctgt	gccaggctca	2460
	tggactgaca	cctcagcagg	tcgtcgccat	tgttcttaac	ggcgagggga	agcaggctct	2520
35	ggagactgtg	cagagactgc	tgccagtgtc	gtgccaggcc	catggactga	cccctcagca	2580
	ggtcgtggct	atcgctagta	acaatggcgg	aaaacaggct	ctggaaactg	tcagcgggct	2640
	gctccccgtg	ctgtgccagg	cccacggcct	cactccacag	caggctcgtc	ctatgcctc	2700
40	taataacggg	ggcaagcagg	ctctggagac	agtacagcgc	ctgttaccgc	tgctgtgcca	2760
	ggcacacggc	ctcacacctc	agcaggctcg	ggcaatcgct	tcccatgacg	gagggaacaa	2820
	ggctctggaa	acggctcaga	ggctgctccc	cgtgctgtgc	caagctcacg	gcctcaccgc	2880
45	tcagcaggtg	gtcgtattg	cttctcatga	tggcggaag	caggctctgg	agaccgtgca	2940
	gagactgctc	cctgtgctgt	gccaagccca	cggcctgact	ccacagcagg	tcgtggccat	3000
	cgctagtcac	gacgggggca	aacaggctct	ggaaacagta	cagcggctgt	taccctgtgt	3060
50	gtgccaagcc	catggcctca	cacctcagca	agtcgtcgtc	atcgctagca	acaatggagg	3120
	gaagcaggct	ctggagacgg	tgacagcctc	gctcccagtg	ctgtgccaag	ctcatggcct	3180
	cacccctcag	caagtcgtcg	caattgcttc	caataacggc	ggaaaacagg	ctctggaaac	3240
55	cgtccagagg	ctgctgcccg	tgctgtgcca	agcacatggc	ttaactccac	agcaagtggg	3300
	ggccattgct	tctaattggg	gcggaaagca	ggccctggag	acagtccaga	gactgttgcc	3360
60							
65							

ES 2 709 216 T3

	cgtgctgtgc caagcgcatg gactgacacc tgaacaggtc gtcgctatcg ctagtaatat	3420
5	tgggggcaaa caggccctgg aaacagtgc ggggctgctt cccgtgctgt gccaggcgca	3480
	tggactcaca cccagcagg tcgtcgcaat cgcctctaataacggaggga agcaggccct	3540
10	ggaaaccgtg cagagactgt tacctgtgct gtgccaggca catggtctga caccacagca	3600
	ggtggtcgca attgctagca atggcggagg gaagcaggcc ctggagactg tccagagact	3660
	gctaccctgt ctgtgccaaag cgcacggcct gacccacag caggctcgtcg ctattgcttc	3720
15	taatggcgga gggcggcctg ctctggagag cattgtggct cagctgtcca ggcccgatcc	3780
	tgccctggct agatccgcac tactaaca tcatctggct gctctcgctt gcctcggtgg	3840
20	acggcccgt ctggacgcag tcaaaaagg tctccccat gctccgcac tgatcaagag	3900
	aaccaacagg agaattcctg agggatccga tcgtttaaac gaaggcatca aaagcaacat	3960
	ctccctcctg aaagacgaac tccgggggca gattagccac attagtcacg aatacctctc	4020
25	cctcatcgac ctggctttcg atagcaagca gaacaggctc tttgagatga aagtgtgga	4080
	actgctcgtc aatgagtacg ggttcaagg tgcacacctc ggcggatcta ggaaaccaga	4140
	cggcatcgtg tatagtacca cactggaaga caactttggg atcattgtgg ataccaaggc	4200
30	atactctgag ggttatagtc tgcccatttc acaggccgac gagatggaac ggtacgtgcg	4260
	cgagaactca aatagagatg aggaagtcaa ccctaacaag tgggtgggaga acttctctga	4320
35	ggaagtgaag aaatactact tcgtctttat cagcgggtcc ttcaagggtta aatttgagga	4380
	acagctcagg agactgagca tgactaccgg cgtgaatggc agcgccgtca acgtggtcaa	4440
	tctgctcctg ggcgtgaaa agattcggag cggagagatg accatcgaag agctggagag	4500
40	ggcaatgttt aataatagcg agtttatcct gaaatactga acgcgtaaat gattgcagat	4560
	ccactagttc tagaattcca gctgagcgcc ggtcgctacc attaccagtt ggtctggtgt	4620
45	caaaaataat aataaccggg caggggggat ctgcatggat ctttgtgaag gaaccttact	4680
	tctgtggtgt gacataattg gacaaactac ctacagagat ttaaagctct aaggtaaata	4740
	taaaattttt aagtgtataa tgtgttaaac tactgattct aattgtttgt gtattttaga	4800
50	ttccaacctt tggaactgat gaatgggagc agtgggtgaa tgccagatcc agacatgata	4860
	agatacattg atgagtttg acaaaccaca actagaatgc agtgaaaaaa atgctttatt	4920
55	tgtgaaatgt gtgatgctat tgctttatgt gtaaccatta taagctgcaa taaacaagtt	4980
	aacaacaaca attgcattca ttttatgttt cagggttcagg gggagggtgt ggagggtttt	5040
	taaagcaagt aaaacctcta caaatgtggt atggctgatt atgatctgcg gccgccactg	5100
60	gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg ttaccaact taatgcctt	5160
	gcagcacatc cccctttcgc cagctggcgt aatagcgaag aggccgcac cgatcgccct	5220

65

ES 2 709 216 T3

	tcccaacagt	tgcgagcct	gaatggcgaa	tggaacgcgc	cctgtagcgg	cgcattaagc	5280
5	gcggcgggtg	tggtgggttac	gcgcagcgtg	accgctacac	ttgccagcgc	cctagcgccc	5340
	gctcctttcg	ctttcttccc	ttcctttctc	gccacgttcg	cgggctttcc	cgtcaagct	5400
	ctaaatcggg	ggctcccttt	agggttcoga	tttagtgctt	tacggcacct	cgaccccaaa	5460
10	aaacttgatt	agggtgatgg	ttcacgtagt	gggccatcgc	cctgatagac	ggtttttcgc	5520
	cctttgacgt	tggagtccac	gttctttaat	agtggactct	tgttccaaac	tggaacaaca	5580
15	ctcaacccta	tctcgggtcta	ttcttttgat	ttataaggga	ttttgccgat	ttcggcctat	5640
	tggttaaaaa	atgagctgat	ttaacaaaaa	tttaacgcga	attttaacaa	aatattaacg	5700
	cttacaattt	agggtggcact	tttcggggaa	atgtgcgcgg	aaccctatt	tgtttatttt	5760
20	tctaaataca	ttcaaataatg	tatccgctca	tgagacaata	accctgataa	atgcttcaat	5820
	aatattgaaa	aaggaagagt	atgagtattc	aacatttccg	tgtcgccctt	attccctttt	5880
	ttgcggcatt	ttgccttcct	gtttttgctc	accagaaaac	gctgggtgaaa	gtaaaagatg	5940
25	ctgaagatca	gttgggtgca	cgagtgggtt	acatcgaact	ggatctcaac	agcggtaaga	6000
	tccttgagag	ttttcgcccc	gaagaacgtt	ttccaatgat	gagcactttt	aaagttctgc	6060
30	tatgtggcgc	ggtattatcc	cgtattgacg	cggggcaaga	gcaactcggg	cggcgcatat	6120
	actattctca	gaatgacttg	gttgagtact	caccagtcac	agaaaagcat	cttacggatg	6180
	gcatgacagt	aagagaatta	tgacgtgctg	ccataaccat	gagtataaac	actgcggcca	6240
35	acttacttct	gacaacgatc	ggaggaccga	aggagctaac	cgtttttttg	cacaacatgg	6300
	gggatcatgt	aactcgcctt	gatcgttggg	aaccggagct	gaatgaagcc	ataccaaacg	6360
	acgagcgtga	caccacgatg	cctgtagcaa	tggaacaac	gttgcgcaaa	ctattaactg	6420
40	gcgaactact	tactctagct	tcccggcaac	aattaataga	ctggatggag	gcggataaag	6480
	ttgcaggacc	acttctgcgc	tggcccttc	cggctggctg	gtttattgct	gataaatctg	6540
45	gagccggtga	gcgtgggtct	cgcggtatca	ttgcagcact	ggggccagat	ggtaagccct	6600
	cccgtatcgt	agttatctac	acgacgggga	gtcaggcaac	tatggatgaa	cgaaatagac	6660
	agatcgctga	gatagggtgcc	tactgatta	agcattggta	actgtcagac	caagtttact	6720
50	catatatact	ttagattgat	ttaaaacttc	atttttaatt	taaaaggatc	taggtgaaga	6780
	tcctttttga	taatctcatg	acccaaatcc	cttaacgtga	gttttcgttc	cactgagcgt	6840
	cagaccccg	agaaaagatc	aaaggatcct	cttgagatcc	tttttttctg	cgcgtaatct	6900
55	gctgcttgca	aacaaaaaaa	ccaccgctac	cagcgggtgt	ttgtttgccg	gatcaagagc	6960
	taccaactct	ttttccgaag	gtaactggct	tcagcagagc	gcagatacca	aatactgtcc	7020
60	ttctagtgtg	gccgtagtta	ggccaccact	tcaagaactc	tgtagcaccg	cctacatacc	7080
	tcgctctgct	aatcctgtta	ccagtggctg	ctgccagtgg	cgataagtcg	tgtcttaccg	7140
65							

ES 2 709 216 T3

5 gggttgactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt 7200
 5 cgtgcacaca gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg 7260
 agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggttaagcg 7320
 10 gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatatcttt 7380
 atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag 7440
 gggggcggag cctatggaaa aacgccagca acgcggcctt ttacgggttc ctggcctttt 7500
 15 gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta 7560
 ttaccgcctt tgagtgaagt gataccgctc gccgcagccg aacgaccgag cgcagcgagt 7620
 cagtgaagca ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc 7680
 20 cgattcatta atgcagctgg cagcagaggt ttcccgaact gaaagcgggc agtgagcgca 7740
 acgcaattaa tgtgagttag ctactcatt aggcacccca ggctttacac tttatgcttc 7800
 cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg 7860
 25 accatga 7867

<210> 53
 <211> 935
 30 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> TalRab2-Clo051
 35 <400> 53

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

65

ES 2 709 216 T3

[illegible]

ES 2 709 216 T3

5	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	355	360	365	
	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	370	375	380	
10	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	385	390	395	400
15	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	405	410	415	
20	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	420	425	430	
	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	435	440	445	
25	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	450	455	460	
30	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	465	470	475	480
35	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	485	490	495	
	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	500	505	510	
40	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	515	520	525	
45	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	530	535	540	
	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	545	550	555	560
50	Ser	Asn	Ile	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	565	570	575	
55	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	580	585	590	
60	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	595	600	605	
65																				

ES 2 709 216 T3

5	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	610	615	620	
10	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	625	630	635	640
15	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	645	650	655	
20	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Arg	Pro	Ala	Leu	Glu	660	665	670	
25	Ser	Ile	Val	Ala	Gln	Leu	Ser	Arg	Pro	Asp	Pro	Ala	Leu	Ala	Arg	Ser	675	680	685	
30	Ala	Leu	Thr	Asn	Asp	His	Leu	Val	Ala	Leu	Ala	Cys	Leu	Gly	Gly	Arg	690	695	700	
35	Pro	Ala	Leu	Asp	Ala	Val	Lys	Lys	Gly	Leu	Pro	His	Ala	Pro	Ala	Leu	705	710	715	720
40	Ile	Lys	Arg	Thr	Asn	Arg	Arg	Ile	Pro	Glu	Gly	Ser	Asp	Arg	Leu	Asn	725	730	735	
45	Glu	Gly	Ile	Lys	Ser	Asn	Ile	Ser	Leu	Leu	Lys	Asp	Glu	Leu	Arg	Gly	740	745	750	
50	Gln	Ile	Ser	His	Ile	Ser	His	Glu	Tyr	Leu	Ser	Leu	Ile	Asp	Leu	Ala	755	760	765	
55	Phe	Asp	Ser	Lys	Gln	Asn	Arg	Leu	Phe	Glu	Met	Lys	Val	Leu	Glu	Leu	770	775	780	
60	Leu	Val	Asn	Glu	Tyr	Gly	Phe	Lys	Gly	Arg	His	Leu	Gly	Gly	Ser	Arg	785	790	795	800
65	Lys	Pro	Asp	Gly	Ile	Val	Tyr	Ser	Thr	Thr	Leu	Glu	Asp	Asn	Phe	Gly	805	810	815	
	Ile	Ile	Val	Asp	Thr	Lys	Ala	Tyr	Ser	Glu	Gly	Tyr	Ser	Leu	Pro	Ile	820	825	830	
	Ser	Gln	Ala	Asp	Glu	Met	Glu	Arg	Tyr	Val	Arg	Glu	Asn	Ser	Asn	Arg	835	840	845	
	Asp	Glu	Glu	Val	Asn	Pro	Asn	Lys	Trp	Trp	Glu	Asn	Phe	Ser	Glu	Glu	850	855	860	

ES 2 709 216 T3

5 Val Lys Lys Tyr Tyr Phe Val Phe Ile Ser Gly Ser Phe Lys Gly Lys
865 870 875 880

10 Phe Glu Glu Gln Leu Arg Arg Leu Ser Met Thr Thr Gly Val Asn Gly
885 890 895

15 Ser Ala Val Asn Val Val Asn Leu Leu Leu Gly Ala Glu Lys Ile Arg
900 905 910

20 Ser Gly Glu Met Thr Ile Glu Glu Leu Glu Arg Ala Met Phe Asn Asn
915 920 925

25 Ser Glu Phe Ile Leu Lys Tyr
930 935

<210> 54
<211> 7867
<212> ADN
<213> Secuencia artificial

<220>
<223> pCAG-RabChtTal1-Clo051

<400> 54

ggcgcgccgg attcgacatt gattattgac tagttattaa tagtaatcaa ttacgggggtc 60

35 attagttcat agcccatata tggagttccg cgttacataa cttacggtaa atggcccgcc 120

tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg ttcccatagt 180

40 aacgccaata gggactttcc attgacgtca atgggtggag tattttacggt aaactgccca 240

cttggcagta catcaagtgt atcatatgcc aagtagccc cctattgacg tcaatgacgg 300

taaatggccc gcctggcatt atgcccagta catgacctta tgggactttc ctacttggca 360

45 gtacatctac gtattagtca tcgctattac catggtcgag gtgagcccca cgttctgctt 420

cactctcccc atctcccccc cctccccacc cccaattttg tattttattta ttttttaatt 480

attttgtgca gcgatggggg cggggggggg gggggggcgc gcgccaggcg gggcggggcg 540

50 gggcgagggg cggggcgggg cgaggcggag aggtgcggcg gcagccaatc agagcggcgc 600

gctccgaaag ttccctttta tggcgaggcg gcggcggcgg cggccctata aaaagcgaag 660

cgcgcgggcg gcgggagtcg ctgcgcgctg ccttcgcccc gtgcccgcgt ccgccgcgcg 720

55 ctgcgcgcgc ccgccccggc tctgactgac cgcgttactc ccacaggtga gcgggcggga 780

cggcccttct cctccgggct gtaattagcg cttggtttaa tgacggcttg tttcttttct 840

60 gtggctgcgt gaaagccttg aggggctccg ggagggccct ttgtgcgggg gggagcggct 900

cgggggggtgc gtgcgtgtgt gtgtgcgtgg ggagcgcgcg gtgcggctcc gcgctgcccg 960

65

ES 2 709 216 T3

5	gcggctgtga gcgctgcggg cgcggcgcg	ggctttgtgc gctccgcagt gtgcgcgagg	1020
	ggagcgcggc cgggggcggt gccccgcggt	gcgggggggg ctgcgagggg aacaaaggct	1080
	gcgtgcgggg tgtgtgcgtg ggggggtgag	caggggggtgt gggcgcgctcg gtcgggctgc	1140
10	aacccccct gcacccccct ccccgagttg	ctgagcacgg cccggcttcg ggtgcggggc	1200
	tccgtacggg gcgtggcgcg gggctcgccg	tgccggggcg ggggtggcgg caggtggggg	1260
	tgccggggcg ggcggggccg cctcgggccc	gggagggctc gggggagggg cgcggcgccc	1320
15	cccgagcgc cggcggtgtg cgaggcgcg	cgagccgcag ccattgcctt ttatggtaat	1380
	cgtgcgagag ggcgcaggga cttcctttgt	cccaaactcg tgcggagccg aaatctggga	1440
	ggcgccgccg caccacctct agcgggcgcg	ggcggaagcg gtgcggcgcc ggcaggaagg	1500
20	aaatgggcgg ggagggcctt cgtgcgtcgc	cgcgcgcgcc tccccctctc cctctccagc	1560
	ctcggggctg tccgcggggg gacggctgcc	ttcggggggg acggggcagg gcggggttcg	1620
	gcttctggcg tgtgaccggc ggctctagag	cctctgctaa ccatgttcat gccttcttct	1680
25	ttttcctaca gatccttaat taataatacg	actcactata ggggccgcca ccatgggacc	1740
	taagaaaaag aggaaggtgg cggccgctga	ctacaaggat gacgacgata aaccaggtgg	1800
	cggaggtagt ggcggaggtg gggtaaccgc	cagtccagca gccaggtgg atctgagaac	1860
30	cctcggctac agccagcagc agcaggagaa	gatcaaacca aaggtgcggg ccaccgtcgc	1920
	tcagcaccat gaagcactgg tggggcacgg	tttcacacac gcccatattg tggctctgtc	1980
	tcagcatccc gctgcactcg ggactgtggc	cgtcaaatat caggacatga tcgccgtct	2040
35	gcctgaggca acccacgaag ccattgtggg	cgtcggaaag cagtggagcg gtgccagagc	2100
	actogaagca ctctcacccg tcgccgggga	actgcggggg ccaccactcc agtccggact	2160
	ggacactgga cagctgctga agatcgctaa	acgcggcgga gtgacagctg tggaaagctgt	2220
40	gcacgcttg aggaatgctc tgacaggagc	cccactgaat cttacaccog aacaggtggg	2280
	ggccatcgct agtaacattg ggggcaaaca	ggctctggaa acagtacagc ggctgttacc	2340
	tgtgctgtgc caggctcatg gcctcacacc	tcagcaggtc gtcgcaatcg cctccaatgg	2400
45	cggaggggaag caggccctgg aaacggtgca	gagactgtta ccagtgtgtg gccaggccca	2460
	tggcctaaca cccagcagg tgggtggcat	cgccagccac gacggcgcca agcaggccct	2520
	ggaaaccgtg cagaggctgc tgcctgtgct	gtgccaggct catggcctga cacctgagca	2580
50	ggtcgtcgcc atcgccagca acatcggcgg	caagcaggcc ctggaaaccg tgcagaggct	2640
	gctgccagtg ctgtgccagg cccatggctt	aacacccgaa cagggtgtgg ccatcgcttc	2700
	taatattggg ggcaagcagg ccctggaaac	agtccagaga ctgttgcttg tgctgtgcca	2760
55	ggctcatggc ttgacacctc agcaggtcgt	cgctatcgcc tctaataagg ggggcaagca	2820
	ggctctggag acagtacagc gcctgttacc	agtgtgtgtc caggcccacg ggctcacacc	2880

ES 2 709 216 T3

5	ccagcaggtg	gtggcaatcg	cttcccatga	cggagggaaa	caggctctgg	aaacgggtcca	2940
	gaggctgctc	cctgtgctgt	gccaggctca	cggctctaaca	ccccagcagg	tggtagccat	3000
	tgctagcaac	aatgggggca	agcaggctct	ggagacagtg	cagcgcttgc	tgctgtgct	3060
10	gtgccaggct	catggcctca	cacctcagca	ggctcgtcgc	atcgccagcc	acgacggcgg	3120
	caagcaggcc	ctggaaaccg	tgacagaggct	gctgccagtg	ctgtgccagg	cccatggcct	3180
	aacaccccag	cagggtggtg	caatcgcttc	caatggcgga	gggaagcagg	ccctggaaac	3240
15	ggtgcagaga	ctgttacctg	tgctgtgcca	ggctcatggc	ctgacacctg	agcaggctgt	3300
	cgctatcgct	agcaatatcg	gaggggaagca	ggctctggaa	actgtccagc	gcctgctccc	3360
	agtgtgtgtc	caggcccatg	gcttaacacc	ccagcaggtg	gtggcaattg	ctagcaatgg	3420
20	cggaggggaag	caggccctgg	agactgtcca	gagactgcta	cctgtgctgt	gccaggctca	3480
	tggtctgaca	cctcagcagg	tcgtcgctat	cgctctaat	aaggggggca	agcaggctct	3540
	ggagacagta	cagcgctgtg	taccagtgtc	gtgccaggcc	cacgggctca	caccccagca	3600
25	ggtggtggcc	atcgccagca	acggcggcgg	caagcaggcc	ctggaaaccg	tgacagaggct	3660
	gctgcctgtg	ctgtgccagg	ctcacggcct	gacccacag	caggctcgtc	ctattgtctc	3720
	taatggcgga	ggcgggcctg	ctctggagag	cattgtggct	cagctgtcca	ggcccgatcc	3780
30	tgccctggct	agatccgcac	tactaaca	tcactctggtc	gctctcgtt	gcctcggtag	3840
	acggcccgtc	ctggacgcag	tcaaaaagg	tctccccc	gctccgcac	tgatcaagag	3900
	aaccaacagg	agaattcctg	agggatccga	tcgttttaac	gaaggcatca	aaagcaacat	3960
35	ctccctcctg	aaagacgaac	tccgggggca	gattagccac	attagtcacg	aatacctctc	4020
	cctcatcgac	ctggctttcg	atagcaagca	gaacaggctc	tttgagatga	aagtgtgga	4080
	actgtcgtc	aatgagtacg	ggttcaagg	tcgacacctc	ggcggatcta	ggaaccaga	4140
40	cggcatcgtg	tatagtacca	cactggaaga	caactttggg	atcattgtgg	ataccaaggc	4200
	atactctgag	ggttatagtc	tgcccatttc	acaggccgac	gagatggaac	ggtacgtg	4260
	cgagaactca	aatagagatg	aggaagtcaa	ccctaacaag	tggtgggaga	acttctctga	4320
45	ggaagtgaag	aaatactact	tcgtctttat	cagcgggtcc	ttcaagggtg	aatttgagga	4380
	acagctcagg	agactgagca	tgactaccgg	cgtgaatggc	agcgccgtca	acgtggtcaa	4440
	tctgtcctg	ggcgctgaaa	agattcggag	cggagagatg	accatcgaag	agctggagag	4500
50	ggcaatgttt	aataatagcg	agtttatcct	gaaatactga	acgcgtaaat	gattgcagat	4560
	ccactagttc	tagaattcca	gctgagcgcc	ggctcgtacc	attaccagtt	ggtctggtgt	4620
	caaaaataat	aataaccggg	caggggggat	ctgcatggat	ctttgtgaag	gaaccttact	4680
55	tctgtggtgt	gacataattg	gacaaactac	ctacagagat	ttaaagctct	aaggtaaata	4740
60							
65							

ES 2 709 216 T3

5	taaaatTTTT aagtgtataa tgtgttaaAC tactgattct aattgtttgt gtattttaga	4800
	ttccaaccta tggaactgat gaatgggagc agtgggtggaa tgccagatcc agacatgata	4860
	agatacattg atgagtttgg acaaaccaca actagaatgc agtgaaaaaa atgctttatt	4920
10	tgtgaaatTT gtgatgctat tgctttatTT gtaaccatta taagctgcaa taaacaagtt	4980
	aacaacaaca attgcattca ttttatgttt cagggttcagg gggaggtgtg ggaggttttt	5040
	taaagcaagt aaaacctcta caaatgtggt atggctgatt atgatctgcg gccgccactg	5100
15	gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg ttaccCAAact taatcgCctt	5160
	gcagcacatc cccctttcgc cagctggcgt aatagcgaag aggcccgcac cgatcgccct	5220
	tccaacagtt tgcgcagcct gaatggcgaa tggaacgcgc cctgtagcgg cgcattaagc	5280
20	gcggcgggtg tgggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgcgc	5340
	gctcctttcg ctttcttccc ttccctttctc gccacgttcg ccggctttcc ccgtcaagct	5400
25	ctaaatcggg ggctcccttt agggttccga tttagtgtt tacggcacct cgaccccaaa	5460
	aaacttgatt aggggtgatg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc	5520
	cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca	5580
30	ctcaacccta tctcggtcta ttcttttgat ttataaggga ttttgccgat ttcggcctat	5640
	tgggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attttaacaa aatattaacg	5700
	cttacaatTT aggtggcact tttcggggaa atgtgcgcgg aaccctatt tgtttatTTT	5760
35	tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat	5820
	aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt	5880
40	ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg	5940
	ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga	6000
	tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc	6060
45	tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggc cgccgcatac	6120
	actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg	6180
	gcatgacagt aagagaatta tgcagtgtcg ccataaccat gaggataaac actgcggcca	6240
50	acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg	6300
	gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg	6360
55	acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg	6420
	gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcggataaag	6480
	ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg	6540
60	gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct	6600
	cccgatatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac	6660
65		

ES 2 709 216 T3

	agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac caagtttact	6720
5	catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga	6780
	tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt	6840
10	cagaccccgt agaaaagatc aaaggatcctt cttgagatcc tttttttctg cgcgtaatct	6900
	gctgcttgca aacaaaaaaa ccaccgctac cagcgggtggt ttgtttgccg gatcaagagc	6960
	taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc	7020
15	ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc	7080
	tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtgc tgtcttaccg	7140
	ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt	7200
20	cgtgcacaca gcccgacttg gagcgaacga cctacaccga actgagatac ctacagcgtg	7260
	agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggttaagcg	7320
	gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatatcttt	7380
25	atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag	7440
	gggggctggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt	7500
30	gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta	7560
	ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt	7620
	cagtgagcga ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc	7680
35	cgattcatta atgcagctgg cacgacaggt ttcccactg gaaagcgggc agtgagcgca	7740
	acgcaattaa tgtgagttag ctactcatt aggcacccca ggctttacac tttatgcttc	7800
	cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg	7860
40	accatga	7867

<210> 55
 <211> 935
 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> RabChtTal1-Clo051
 50
 <400> 55

55	Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp	1 5 10 15
60	Asp Asp Asp Lys Pro Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Pro	20 25 30
65	Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln	

ES 2 709 216 T3

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

ES 2 709 216 T3

5 Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
290 295 300

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
305 310 315 320

10 Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
325 330 335

15 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
340 345 350

20 Val Ala Ile Ala Ser Asn Lys Gly Gly Lys Gln Ala Leu Glu Thr Val
355 360 365

25 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
370 375 380

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
385 390 395 400

30 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
405 410 415

35 Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala
420 425 430

40 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
435 440 445

45 Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
450 455 460

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
465 470 475 480

50 His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
485 490 495

55 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
500 505 510

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn
515 520 525

60 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
530 535 540

65

ES 2 709 216 T3

5	Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala 545 550 555 560
10	Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu 565 570 575
15	Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala 580 585 590
20	Ile Ala Ser Asn Lys Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg 595 600 605
25	Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val 610 615 620
30	Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val 625 630 635 640
35	Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln 645 650 655
40	Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu 660 665 670
45	Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Arg Ser 675 680 685
50	Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg 690 695 700
55	Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu 705 710 715 720
60	Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Gly Ser Asp Arg Leu Asn 725 730 735
65	Glu Gly Ile Lys Ser Asn Ile Ser Leu Leu Lys Asp Glu Leu Arg Gly 740 745 750
	Gln Ile Ser His Ile Ser His Glu Tyr Leu Ser Leu Ile Asp Leu Ala 755 760 765
	Phe Asp Ser Lys Gln Asn Arg Leu Phe Glu Met Lys Val Leu Glu Leu 770 775 780
	Leu Val Asn Glu Tyr Gly Phe Lys Gly Arg His Leu Gly Gly Ser Arg 785 790 795 800

ES 2 709 216 T3

5		Lys Pro Asp Gly Ile Val Tyr Ser Thr Thr Leu Glu Asp Asn Phe Gly	805	810	815
10		Ile Ile Val Asp Thr Lys Ala Tyr Ser Glu Gly Tyr Ser Leu Pro Ile	820	825	830
15		Ser Gln Ala Asp Glu Met Glu Arg Tyr Val Arg Glu Asn Ser Asn Arg	835	840	845
20	v	Asp Glu Glu Val Asn Pro Asn Lys Trp Trp Glu Asn Phe Ser Glu Glu	850	855	860
25		Val Lys Lys Tyr Tyr Phe Val Phe Ile Ser Gly Ser Phe Lys Gly Lys	865	870	875
30		Phe Glu Glu Gln Leu Arg Arg Leu Ser Met Thr Thr Gly Val Asn Gly	885	890	895
35		Ser Ala Val Asn Val Val Asn Leu Leu Leu Gly Ala Glu Lys Ile Arg	900	905	910
40		Ser Gly Glu Met Thr Ile Glu Glu Leu Glu Arg Ala Met Phe Asn Asn	915	920	925
45		Ser Glu Phe Ile Leu Lys Tyr	930	935	
50		<210> 56			
55		<211> 7867			
60		<212> ADN			
65		<213> Secuencia artificial			
70		<220>			
75		<223> pCAG-RabChtTal2-Clo051			
80		<400> 56			
85		ggcgcgccgg attcgacatt gattattgac tagttattaa tagtaatcaa ttacgggggtc			60
90		attagttcat agcccatata tggagttccg cgttacataa cttacggtaa atggcccgcc			120
95		tggctgaccg cccaacgacc cccgccatt gacgtcaata atgacgtatg ttcccatagt			180
100		aacgccaata gggactttcc attgacgtca atgggtggag tatttacggt aaactgcca			240
105		cttggcagta catcaagtgt atcatatgcc aagtacgccc cctattgacg tcaatgacgg			300
110		taaatggccc gcctggcatt atgcccagta catgacctta tgggactttc ctacttggca			360
115		gtacatctac gtattagtca tcgctattac catggtcgag gtgagcccca cgttctgctt			420
120		cactctcccc atctcccccc cctccccacc cccaattttg tatttattta ttttttaatt			480

ES 2 709 216 T3

5	at t t t t g t g c a	g c g a t g g g g g	c g g g g g g g g	g g g g g g g c g	g c g c c a g g c g	g g g c g g g c g	540
	g g g c g a g g g g	c g g g g c g g g	c g a g g c g g a g	a g g t g c g g c g	g c a g c c a a t c	a g a g c g g c g c	600
	g c t c c g a a a g	t t t c c t t t t a	t g g c g a g g c g	g c g g c g g c g g	c g g c c c t a t a	a a a a g c g a a g	660
10	c g c g c g g c g g	g c g g g a g t c g	c t g c g c g c t g	c c t t c g c c c c	g t g c c c c g c t	c c g c c g c c g c	720
	c t c g c g c c g c	c c g c c c c g g c	t c t g a c t g a c	c g c g t t a c t c	c c a c a g g t g a	g c g g g c g g g a	780
	c g g c c c t t c t	c c t c c g g g c t	g t a a t t a g c g	c t t g g t t t a a	t g a c g g c t t g	t t t c t t t t c t	840
15	g t g g c t g c g t	g a a a g c c t t g	a g g g g c t c c g	g g a g g g c c c t	t t g t g c g g g g	g g g a g c g g c t	900
	c g g g g g g t g c	g t g c g t g t g t	g t g t g c g t g g	g g a g c g c c g c	g t g c g g c t c c	g c g c t g c c c g	960
	g c g g c t g t g a	g c g c t g c g g g	c g c g g c g c g g	g g c t t t g t g c	g c t c c g c a g t	g t g c g c g a g g	1020
20	g g a g c g c g g c	c g g g g g c g g t	g c c c c g c g g t	g c g g g g g g g g	c t g c g a g g g g	a a c a a a g g c t	1080
	g c g t g c g g g g	t g t g t g c g t g	g g g g g g t g a g	c a g g g g g t g t	g g g c g c g t c g	g t c g g g c t g c	1140
	a a c c c c c c c t	g c a c c c c c c t	c c c c g a g t t g	c t g a g c a c g g	c c c g g c t t c g	g g t g c g g g g c	1200
25	t c c g t a c g g g	g c g t g g c g c g	g g g c t c g c c g	t g c c g g g c g g	g g g g t g g c g g	c a g g t g g g g g	1260
	t g c c g g g c g g	g g c g g g g c c g	c c t c g g g c c g	g g g a g g g c t c	g g g g g a g g g g	c g c g c g g g c c	1320
	c c c g g a g c g c	c g g c g g c t g t	c g a g g c g c g g	c g a g c c g c a g	c c a t t g c c t t	t t a t g g t a a t	1380
30	c g t g c g a g a g	g g c g c a g g g a	c t t c c t t t g t	c c c a a a t c t g	t g c g g a g c c g	a a a t c t g g g a	1440
	g g c g c c g c c g	c a c c c c c t c t	a g c g g g c g c g	g g g c g a a g c g	g t g c g g c g c c	g g c a g g a a g g	1500
	a a a t g g g c g g	g g a g g g c c t t	c g t g c g t c g c	c g c g c c g c c g	t c c c c t t c t c	c c t c t c c a g c	1560
35	c t c g g g g c t g	t c c g c g g g g g	g a c g g c t g c c	t t c g g g g g g g	a c g g g g c a g g	g c g g g g t t c g	1620
	g c t t c t g g c g	t g t g a c c g g c	g g c t c t a g a g	c c t c t g c t a a	c c a t g t t c a t	g c c t t c t t c t	1680
	t t t t c c t a c a	g a t c c t t a a t	t a a t a a t a c g	a c t c a c t a t a	g g g g c c g c c a	c c a t g g g a c c	1740
40	t a a g a a a a a g	a g g a a g g t g g	c g g c c g c t g a	c t a c a a g g a t	g a c g a c g a t a	a a c c a g g t g g	1800
	c g g a g g t a g t	g g c g g a g g t g	g g g t a c c c g c	c a g t c c a g c a	g c c c a g g t g g	a t c t g a g a a c	1860
	c c t c g g c t a c	a g c c a g c a g c	a g c a g g a g a a	g a t c a a a c c a	a a g g t g c g g t	c c a c c g t c g c	1920
45	t c a g c a c c a t	g a a g c a c t g g	t g g g g c a c g g	t t t c a c a c a c	g c c c a t a t t g	t g g c t c t g t c	1980
	t c a g c a t c c c	g c t g c a c t c g	g g a c t g t g g c	c g t c a a a t a t	c a g g a c a t g a	t c g c c g c t c t	2040
	g c c t g a g g c a	a c c c a c g a a g	c c a t t g t g g g	c g t c g g a a a g	c a g t g g a g c g	g t g c c a g a g c	2100
50	a c t c g a a g c a	c t c c t c a c c g	t c g c c g g g g a	a c t g c g g g g t	c c a c c a c t c c	a g t c c g g a c t	2160
	g g a c a c t g g a	c a g c t g c t g a	a g a t c g c t a a	a c g c g g c g g a	g t g a c a g c t g	t g g a a g c t g t	2220
	g c a c g c t t g g	a g g a a t g c t c	t g a c a g g a g c	c c c a c t g a a t	c t t a c a c c c c	a g c a g g t g g t	2280
55	g g c c a t t g c t	a g c a a c a a t g	g g g g c a a g c a	g g c t c t g g a g	a c a g t g c a g c	g c c t g c t g c c	2340
	t g t g c t g t g c	c a g g c t c a t g	g c c t c a c a c c	t c a g c a g g t c	g t c g c c a t t g	c t t c t a a c a a	2400
60							
65							

ES 2 709 216 T3

5	tggaggggaag	caggctctgg	agactgtgca	gagactgctg	ccagtgtgtg	gccaggccca	2460
	tggcctaaca	ccccagcagg	tggtggccat	cgccagccac	gacggcggca	agcaggccct	2520
	ggaaaccgtg	cagaggctgc	tgctgtgtct	gtgccaggct	catggcctga	cacctcagca	2580
10	ggtcgtcgcc	atcgccagcc	acgacggcgg	caagcaggcc	ctggaaaccg	tgcagaggct	2640
	gctgccagt	ctgtgccagg	cccatggctt	aacaccccag	caggtggtgg	ccatcgctag	2700
	tcatgacggg	ggcaaacagg	ctctggaaac	agtacagcgg	ctgttacctg	tgtgtgtcca	2760
15	ggctcatggc	ttgacacctc	agcaggctgt	cgctatcgcc	tctaataagg	ggggcaagca	2820
	ggctctggag	acagtacagc	gcctgttacc	agtgtgtgtc	caggcccacg	ggctcacacc	2880
	ccagcagggt	gtggcaattg	cttccaataa	gggcggaaaa	caggctctgg	aaaccgtcca	2940
20	gaggctgctg	cctgtgtgtg	gccaggctca	cggtctaaca	ccccagcagg	tggtggccat	3000
	cgcttccaac	ggagggggca	aacaggctct	ggaaacagt	cagaggctgc	tgctgtgtct	3060
25	gtgccaggct	catggcctca	cacctgagca	ggtcgtcgcc	atcgccagca	acatcggcgg	3120
	caagcaggcc	ctggaaaccg	tgcagaggct	gctgccagt	ctgtgccagg	cccatggcct	3180
	aacaccccag	caggtggtgg	caattgcttc	caataagggc	ggaaaacagg	ctctggaaac	3240
30	cgtccagagg	ctgtgcctg	tgtgtgtcca	ggctcatggc	ctgacacctc	agcaggctgt	3300
	cgcaatcgcc	tccaatggcg	gaggggaagca	ggccctggaa	acggtgcaga	gactgttacc	3360
	agtgtgtgtc	caggcccattg	gcttaacacc	ccagcagggt	gtggcaatcg	cctctaataa	3420
35	gggaggggaag	caggccctgg	aaaccgtgca	gagactgtta	cctgtgtgtg	gccaggctca	3480
	tggcttgaca	cctcagcagg	tcgtcgctat	cgctagtcac	gatggcggaa	aacaggctct	3540
40	ggaaactgtg	cagcggctgc	tcccagtgtc	gtgccaggcc	cacgggctca	caccccagca	3600
	ggtggtggcc	atcgccagca	acaagggcgg	caagcaggcc	ctggaaaccg	tgcagaggct	3660
	gctgcctgtg	ctgtgccagg	ctcacggcct	gaccccacag	caggtcgtcg	ctattgcttc	3720
45	taatggcgga	gggcggcctg	ctctggagag	cattgtggct	cagctgtcca	ggcccgatcc	3780
	tgccctggct	agatccgcac	tcaactaacga	tcatctggct	gctctcgctt	gcctcgggtg	3840
50	acggcccgcct	ctggacgcag	tcaaaaagg	tctcccccat	gctcccgcac	tgatcaagag	3900
	aaccaacagg	agaattcctg	agggatccga	tcgttttaaac	gaaggcatca	aaagcaacat	3960
	ctccctcctg	aaagacgaac	tccgggggca	gattagccac	attagtcacg	aatacctctc	4020
55	cctcatcgac	ctggctttcg	atagcaagca	gaacaggctc	tttgagatga	aagtgtgtga	4080
	actgctcgtc	aatgagtacg	ggttcaagg	tcgacacctc	ggcggatcta	ggaaaccaga	4140
	cggcatcgtg	tatagtacca	caactggaaga	caactttggg	atcattgtgg	ataccaaggc	4200
60	atactctgag	ggttatagtc	tgcccatttc	acaggccgac	gagatggaac	ggtacgtgcg	4260

65

ES 2 709 216 T3

5	cgagaactca aatagagatg aggaagtcaa ccctaacaag tgggtgggaga acttctctga	4320
	ggaagtgaag aaatactact tcgtctttat cagcgggtcc ttcaagggtta aatttgagga	4380
	acagctcagg agactgagca tgactaccgg cgtgaatggc agcgccgtca acgtgggtcaa	4440
10	tctgctcctg ggcgctgaaa agattcggag cggagagatg accatcgaag agctggagag	4500
	ggcaatgttt aataatagcg agtttatcct gaaatactga acgcgtaaat gattgcagat	4560
	ccactagttc tagaattcca gctgagcgcc ggtcgctacc attaccagtt ggtctggtgt	4620
15	caaaaataat aataaccggg caggggggat ctgcatggat ctttgtgaag gaaccttact	4680
	tctgtggtgt gacataattg gacaaactac ctacagagat ttaaagctct aaggtaaata	4740
	taaaattttt aagtgtataa tgtgttaaac tactgattct aattgtttgt gtattttaga	4800
20	ttccaacctta tggaactgat gaatgggagc agtgggtggaa tgccagatcc agacatgata	4860
	agatacattg atgagtttgg acaaaccaca actagaatgc agtgaaaaaa atgctttatt	4920
	tgtgaaattt gtgatgctat tgctttatct gtaaccatta taagctgcaa taaacaagtt	4980
25	aacaacaaca attgcattca ttttatgttt caggttcagg gggaggtgtg ggaggttttt	5040
	taaagcaagt aaaacctcta caaatgtggt atggctgatt atgatctgcg gccgccactg	5100
	gccgtcgttt tacaacgtcg tgactgggaa aacctggcg ttaccaact taatcgccct	5160
30	gcagcacatc cccctttcgc cagctggcgt aatagcgaag aggcccgcac cgatcgccct	5220
	tcccaacagt tgcgcagcct gaatggcgaa tggaacgcgc cctgtagcgg cgcattaagc	5280
	gcggcgggtg tgggtggttac gcgcagcgtg accgctacac ttgccagcgc ctagcgccc	5340
35	gctcctttcg ctttcttccc ttcttttctc gccacgttcg ccggctttcc ccgtcaagct	5400
	ctaaatcggg ggctcccttt agggttccga tttagtgtct tacggcacct cgaccccaaa	5460
	aaacttgatt aggtgatggg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc	5520
40	cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca	5580
	ctcaacccta tctcgggtcta ttcttttgat ttataaggga ttttgccgat ttcggcctat	5640
	tgggtaaaaa atgagctgat ttaacaaaaa tttaacgcga attttaacaa aatattaacg	5700
45	cttacaattt aggtggcact tttcggggaa atgtgcgcgg aaccctatt tgtttatattt	5760
	tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat	5820
	aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgcctt attccctttt	5880
50	ttgcggcatt ttgccttctt gtttttgctc acccagaaac gctggtgaaa gtaaaagatg	5940
	ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggttaaga	6000
	tccttgagag ttttcgcccc gaagaacggt ttccaatgat gagcactttt aaagttctgc	6060
55	tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggg cgccgcatac	6120
	actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg	6180
60		
65		

ES 2 709 216 T3

5	gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggccca	6240
	acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg	6300
	gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg	6360
10	acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg	6420
	gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcgataaaag	6480
	ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg	6540
15	gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct	6600
	cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac	6660
20	agatcgctga gatagggtgc tcaactgatta agcattggta actgtcagac caagtttact	6720
	catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga	6780
	tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt	6840
25	cagaccccggt agaaaagatc aaaggatcct cttgagatcc tttttttctg cgcgtaatct	6900
	gctgcttgca aacaaaaaaaa ccaccgctac cagcgggtggg ttgtttgccg gatcaagagc	6960
30	taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc	7020
	ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc	7080
	tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg	7140
35	ggttggaactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt	7200
	cgtgcacaca gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg	7260
	agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaaagc	7320
40	gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatatctt	7380
	atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag	7440
45	gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt	7500
	gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta	7560
	ttaccgcctt tgagtgaact gataccgctc gccgcagccg aacgaccgag cgcagcgagt	7620
50	cagtgaacga ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc	7680
	cgattcatta atgcagctgg cagcagaggt ttcccactg gaaagcgggc agtgagcgca	7740
55	acgcaattaa tgtgagttag ctcaactcatt aggcacccca ggctttacac tttatgcttc	7800
	cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg	7860
60	accatga	7867
65		

ES 2 709 216 T3

<210> 57
 <211> 935
 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> RabChtTal2-Clo051
 <400> 57

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

ES 2 709 216 T3

5	Gly 225	Lys	Gln	Ala	Leu	Glu 230	Thr	Val	Gln	Arg	Leu 235	Leu	Pro	Val	Leu	Cys 240
10	Gln	Ala	His	Gly	Leu 245	Thr	Pro	Gln	Gln	Val 250	Val	Ala	Ile	Ala	Ser 255	His
15	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val 265	Gln	Arg	Leu	Leu	Pro	Val
20	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala
25	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val 300	Gln	Arg	Leu	Leu
30	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala 320
35	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala 330	Leu	Glu	Thr	Val	Gln	Arg 335
40	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val
45	Val	Ala	Ile	Ala	Ser	Asn	Lys	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val
50	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln
55	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Lys	Gly	Gly	Lys	Gln	Ala	Leu	Glu 400
60	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr 415
65	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala
	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly
	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly	Lys
	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala

ES 2 709 216 T3

5	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Lys	Gly	
					485					490					495		
	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	
				500					505					510			
10	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	
			515					520					525				
15	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	
		530					535					540					
20	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	Ile	Ala	
	545					550					555					560	
25	Ser	Asn	Lys	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	
				565						570					575		
	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	Val	Ala	
				580					585					590			
30	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	
			595					600						605			
35	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	Gln	Val	
		610					615					620					
40	Val	Ala	Ile	Ala	Ser	Asn	Lys	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	
	625					630					635					640	
	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Gln	
					645					650					655		
45	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Arg	Pro	Ala	Leu	Glu	
				660					665					670			
50	Ser	Ile	Val	Ala	Gln	Leu	Ser	Arg	Pro	Asp	Pro	Ala	Leu	Ala	Arg	Ser	
			675					680					685				
55	Ala	Leu	Thr	Asn	Asp	His	Leu	Val	Ala	Leu	Ala	Cys	Leu	Gly	Gly	Arg	
		690					695					700					
	Pro	Ala	Leu	Asp	Ala	Val	Lys	Lys	Gly	Leu	Pro	His	Ala	Pro	Ala	Leu	
		705				710					715					720	
60	Ile	Lys	Arg	Thr	Asn	Arg	Arg	Ile	Pro	Glu	Gly	Ser	Asp	Arg	Leu	Asn	
					725					730					735		
65																	

ES 2 709 216 T3

5 Glu Gly Ile Lys Ser Asn Ile Ser Leu Leu Lys Asp Glu Leu Arg Gly
 740 745 750

 Gln Ile Ser His Ile Ser His Glu Tyr Leu Ser Leu Ile Asp Leu Ala
 755 760 765

10 Phe Asp Ser Lys Gln Asn Arg Leu Phe Glu Met Lys Val Leu Glu Leu
 770 775 780

15 Leu Val Asn Glu Tyr Gly Phe Lys Gly Arg His Leu Gly Gly Ser Arg
 785 790 795 800

20 Lys Pro Asp Gly Ile Val Tyr Ser Thr Thr Leu Glu Asp Asn Phe Gly
 805 810 815

 Ile Ile Val Asp Thr Lys Ala Tyr Ser Glu Gly Tyr Ser Leu Pro Ile
 820 825 830

25 Ser Gln Ala Asp Glu Met Glu Arg Tyr Val Arg Glu Asn Ser Asn Arg
 835 840 845

30 Asp Glu Glu Val Asn Pro Asn Lys Trp Trp Glu Asn Phe Ser Glu Glu
 850 855 860

35 Val Lys Lys Tyr Tyr Phe Val Phe Ile Ser Gly Ser Phe Lys Gly Lys
 865 870 875 880

 Phe Glu Glu Gln Leu Arg Arg Leu Ser Met Thr Thr Gly Val Asn Gly
 885 890 895

40 Ser Ala Val Asn Val Val Asn Leu Leu Leu Gly Ala Glu Lys Ile Arg
 900 905 910

45 Ser Gly Glu Met Thr Ile Glu Glu Leu Glu Arg Ala Met Phe Asn Asn
 915 920 925

50 Ser Glu Phe Ile Leu Lys Tyr
 930 935

55 <210> 58
 <211> 6607
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> pRab38-chtTAL

60 <400> 58

65

ES 2 709 216 T3

5	caccgcatta ccctgggcgt tgaaaccgaa gaagacctgg atttgaaata ggcgttttct	60
	ttacatttct aaagtgggac tcctcacttg taaaaggaaa aataatgata cttttaagac	120
	ttccaggatg actaaatggg gtgtatgaga agatttataa acatctgccg ctacttacia	180
10	tgataagacc acttgtgtgt tgttcagctt ggagaattta ggataggagt ggaggctgaa	240
	agaaaagtaa gcccttagca tttcctctca ggtggcctct actttaggtc attaacagtt	300
	gaataggcgc taagagatag cattaccact ttatagaagc ccaggcaaaa ggagattaaa	360
15	gggtttgcct aaattctttc aactctaagg gccagagaag acctaagtct actgctttgc	420
	tgtttctcaa ggtctcccca actttacaac actgtgtggg tggcaacagg gcttaatagc	480
	ctcagaagac ctgggtatct ttcgacactc agttctctcc ccggcagaac gtggaaaaca	540
20	aaatccacat aagtttgtgt catggacggg aggcgagaga aaaatctctg tgaaaggagt	600
	aaagcactgt gcaaatacca gcttgacagg cagtagcact ggggtcccg gtccttttagc	660
25	ttccagtccc aggagttgct cttgtctcct cccactctgg agtcgcgaga gtaggaagga	720
	ggattaaacc cgggggagga gttccgcacc agtcctctat cctgcgccag cagccttagc	780
	ctaagcggcc acatagagct ccggtctccg tcggtgcccc gccccggctg tgcttcccag	840
30	agcaagctcc aggtccgca agaccccgcg gcctccagga tgcagacacc tcacaaggag	900
	cacctgtaca agctgctggg gatcggcgac ctggtagtgg gcaagaccag cattattaaa	960
	cggtagctgc atcaaaattt ctctctcat tatcgagcca ccattgggtg ggacttcgcg	1020
35	ctgaagggtg tccactggga ccagagacg gtggtgctct tgcagctctg ggacattgct	1080
	ggtgagcgat cagagcagcg cgcaacgggt gagggaggag tgagccagtg aggagttcgg	1140
	gggtgaagggt tcggggagtg gaaaatgact ttccagtccg ttccagtccc gggacccttg	1200
40	agtgcacatc agcaggagat ccggatcgcc tgggcgctcc actcttgga agtttggtt	1260
	aatggcttgg aaacctgatt tcaaagaaat ggaagtgttt tcttttcttt ctttctttt	1320
45	tttttttttt ttttttctt ttgctgttgt ttctgttga gtcgtcccc ctctacctgt	1380
	aacttctaga taacttcgct ggctctcact ggctgtgaga aagcgaacca ctttctctg	1440
	ggattcttgg gtgcagagaa ggctgtcgcc tggactcaca aggagattgt agtcgcattc	1500
50	ttgtttcatt ctagtctttt tctggacaca ggtagcccg acttgcccc gagtatctca	1560
	cgtggctttc atccttcgtg ttttagagggg aagcccctag gaaatttaag aaggagcagg	1620
	attatcttag gaatttagtt tctttcaaat ctcaactacta tcatctcctt gcttattggc	1680
55	ctcttcagtc agaaaaattt gagatgctaa atttgtatac atctagaacg aactatctct	1740
	tctcactcca ctcccctctt ccccatctct cttccgtctc cctccatcct tggtatctc	1800
	ttcttcactt tccatttcaa acaggagact gtgtatgttt tttaggaaaa cattaaaaaa	1860
60	aaaaccacaa aaacaaaaac aaaacggaga cagggtcccg tcatgtaact ctgctaact	1920

65

ES 2 709 216 T3

5	atatcaagct gaccttgacc tcatagagac ccacttgccct ctgcctccct agaggcaagg	1980
	gtcgggggta tggatgatgtt aatgtcgttt gctttaagat tccttgattt gatcttggtg	2040
	tatTTTTTga gaaatctaaa gtatgaaatc agagtttgac taacagcttc taccagctcc	2100
10	tagccacaat aaagactgag gcaggctata gttagtgtctc aatactgggt cctacctggc	2160
	tgcttgtaac ctgggcatgc ctagcattct agatgctaac tcaccaaagc agtagcattt	2220
	taagctgcaa atggctaggc agcgacagct caagaatctt cttgctttgg agttttaaac	2280
15	tccaatgaga ttttccatga tccctttcaa ataaccctac ttaatctctc ttcatagccc	2340
	acagtaccaa gaagcctttg ataagctctg gattgaaaag aagcagttct ttttcaaaag	2400
	atgtgctcat ttgaactagt gcatttccct ggaacacatt tgccaggact tgagatgggc	2460
20	actaagaagg aaaattcctc aaaggacatg tacagtcttg agatgcattc gcttctgtag	2520
	ccatgagctt gctgggtcttg agataagggtt agttgggtga gctaggttca tgggttgagg	2580
25	tctttggcag ttctagagaa gcatgagcta ttagagactt ggagattgca tcaagtagag	2640
	ccttttgagc ttttccactgt gtacctgggc cctctgtcgc tgcacgtttt agtgtctgaa	2700
	atgtctttca gctgtagcag ttttctcggg accccagttt aaaatagctt actgtttaaa	2760
30	agatgtagct gtagctagca ttattgaact agcataatta tagtctaaat agcattatgt	2820
	cttcagcctt gttatatgtt ggtgagtttt agtttctctt tctaaacggg aagaacagaa	2880
	agatgtaatg attctgagct tccagagtga gacacctcta gagagaaata ccttcttctg	2940
35	aagactaccg tgtgattaca gataaattct gatatctttg tttagctttt gatatctata	3000
	aacagggagt gtattttatc tctccaaatg agagaagaat aaacaataat gcaaggtaaa	3060
40	ggcaatagtg ctacactcta ggagttacca ctctttgtac atttatttat aaatactaag	3120
	caagaggaac atgccataca tacactgact aagtcctaac aagtggcagt tcttatatca	3180
	cacatttatc ttgccctcaa atgccagtcc agcatcagtt tagtctcatg catttggcag	3240
45	cataaaggcag tttgagttcc acacttgctc tcagaagcaa tttaactccc acacttggga	3300
	atcctttcct aagccacagt ttcagaccaa agttttggtg aaggctataa tcacagaagt	3360
	ctgcacaagt agggagtctg aaggatctga gctccattca gcagtcagag cggcatccaa	3420
50	ccccaggta atgctcagct cactttgata acttcaagct caaaggccct gaactgctga	3480
	gttgagggtt gaaagatgtt tgggtaaaag caaggtaatt ggcggatagg atggttgtaa	3540
	cgtaattgtt tcaagttgta ttagagacct ctgggttcta aggggatatg aaatccaacc	3600
55	tccactctcc actgagattc aagttaggtt aagtatgcct ttgagtaccc tcaagtcaca	3660
	gcatgccact ctctttttct taactctaata atgtatctat aaagaacggg tagtagtcaa	3720
60	ctgagtcgac ggtatcgata agcttgatcc agcttttggt cccttttagtg agggttaatt	3780

65

ES 2 709 216 T3

5	gcgcgccttgg	cgtaatcatg	gtcatagctg	tttcctgtgt	gaaattgtta	tccgctcaca	3840
	attccacaca	acatacgagc	cggaagcata	aagtgtaaa	cctgggggtgc	ctaattgagt	3900
	agctaactca	cattaattgc	gttgcgctca	ctgcccgctt	tccagtcggg	aaacctgtcg	3960
10	tgccagctgc	attaatgaat	cggccaaacgc	gcggggagag	gcggtttgcg	tattgggcgc	4020
	tcttccgctt	cctcgctcac	tgactcgctg	cgctcggtcg	ttcggtcgcg	gcgagcggta	4080
	tcagctcact	caaaggcgg	aatacggtta	tccacagaat	caggggataa	cgcaggaaag	4140
15	aacatgtgag	caaaaggcca	gcaaaaggcc	aggaaccgta	aaaaggccgc	gttgctggcg	4200
	ttttccata	ggctccgccc	ccctgacgag	catcacaaaa	atcgacgctc	aagtcagagg	4260
	tggcgaaacc	cgacaggact	ataaagatac	caggcgtttc	cccctggaag	ctccctcgtg	4320
20	cgctctcctg	ttccgaccct	gccgcttacc	ggatacctgt	ccgcctttct	cccttcggga	4380
	agcgtggcgc	tttctcatag	ctcacgctgt	aggtatctca	gttcgggtgt	ggtcgttcgc	4440
25	tccaagctgg	gctgtgtgca	cgaaccccc	gttcagcccg	accgctgcgc	cttatccggt	4500
	aactatcgtc	ttgagtccea	cccggtaaga	cacgacttat	cgccactggc	agcagccact	4560
	ggtaacagga	ttagcagagc	gaggtatgta	ggcggtgcta	cagagttctt	gaagtggtag	4620
30	cctaactacg	gctacactag	aaggacagta	tttggtatct	gcgctctgct	gaagccagtt	4680
	accttcggaa	aaagagttgg	tagctcttga	tccggcaaac	aaaccaccgc	tggtagcggg	4740
	ggtttttttg	tttgcaagca	gcagattacg	cgcagaaaaa	aaggatctca	agaagatcct	4800
35	ttgatctttt	ctacggggtc	tgacgctcag	tggaaacgaa	actcacgtta	agggattttg	4860
	gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	taaattaa	atgaagtttt	4920
	aaatcaatct	aaagtatata	tgagtaaact	tgggtctgaca	gttaccaatg	cttaatcagt	4980
40	gaggcaccta	tctcagcgat	ctgtctattt	cgttcaccca	tagttgcctg	actccccgtc	5040
	gtgtagataa	ctacgatacg	ggagggctta	ccatctggcc	ccagtgcctg	aatgataccg	5100
45	cgagaccac	gtcacccggc	tccagattta	tcagcaataa	accagccagc	cggaaagggcc	5160
	gagcgcgaaa	gtggctcctg	aactttatcc	gcctccatcc	agtctattaa	ttgttgccgg	5220
	gaagctagag	taagtagttc	gccagttaat	agtttgcgca	acgttggtgc	cattgctaca	5280
50	ggcatcgtgg	gtgcacgctc	gtcgtttggg	atgggttcac	tcagctccgg	ttcccaacga	5340
	tcaaggcgag	ttacatgatc	ccccatgttg	tgcaaaaaag	cggttagctc	cttcggtcct	5400
	ccgatcgttg	tcagaagtaa	gttggccgca	gtgttatcac	tcattggttat	ggcagcaact	5460
55	cataattctc	ttactgtcat	gccatccgta	agatgctttt	ctgtgactgg	tgagtactca	5520
	accaagtcat	tctgagaata	gtgtatgcgg	cgaccgagtt	gctcttgccc	ggcgtcaata	5580
	cgggataata	ccgcgccaca	tagcagaact	ttaaaagtgc	tcattcattgg	aaaacgttct	5640
60	tcggggcgaa	aactctcaag	gatcttaccg	ctgttgagat	ccagttcgat	gtaaccact	5700

ES 2 709 216 T3

	cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttcttg gtgagcaaaa	5760
	acaggaaggc aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc	5820
5	atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga	5880
	tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga	5940
10	aaagtgccac ctaaattgta agcgttaata ttttgtaaa attcgcgtta aatttttgtt	6000
	aaatcagctc attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag	6060
	aatagaccga gataggggtg agtggtgttc cagtttgga caagagtcca ctattaaaga	6120
15	acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg	6180
	aaccatcacc ctaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc	6240
20	ctaaaggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg	6300
	aagggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc	6360
	gcgtaaccac cacacccgcc gcgcttaatg cgccgctaca gggcgcgctc cattcgccat	6420
25	tcaggctgcg caactgttgg gaaggcgat cggtgcgggc ctcttcgcta ttacgccagc	6480
	tggcgaaagg gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt	6540
30	cacgacgttg taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat	6600
	tggagct	6607
35	<210> 59 <211> 202 <212> ADN <213> Mus musculus	
40	<400> 59	
	atgcagacac ctcaaaagga gcacctgtac aagctgctgg tgatcggcga cctgggtgtg	60
	ggcaagacca gcattatcaa gcgctatgtg caccaaaact tctcctcgca ctaccgggcc	120
45	accattggtg tggacttcgc gctgaagggtg ctccactggg acccagagac ggtggtgcgc	180
	ttgcagctct gggacattgc tg	202
50	<210> 60 <211> 8218 <212> ADN <213> Secuencia artificial	
55	<220> <223> pRab38-chtTAL-neo	
	<400> 60	
60	caccgcatta ccctgggcgt tgaaccgaa gaagacctgg atttgaaata ggcgttttct	60
	ttacatttct aaagtgggac tcctcacttg taaaaggaaa aataatgata cttttaagac	120
65	ttccaggatg actaaatggg gtgtatgaga agatttataa acatctgccg ctacttacaa	180

ES 2 709 216 T3

	tgataagacc	acttgtgtgt	tgttcagctt	ggagaattta	ggataggagt	ggaggctgaa	240
5	agaaaagtaa	gcccttagca	tttcctctca	ggtggcctct	acttttagtc	attaacagtt	300
	gaataggcgc	taagagatag	cattaccact	ttatagaagc	ccaggcaaaa	ggagattaaa	360
	gggtttgcct	aaattctttc	aactctaagg	gccagagaag	acctaagtct	actgctttgc	420
10	tgtttctcaa	ggtctcccca	actttacaac	actgtgtggg	tggcaacagg	gcttaatagc	480
	ctcagaagac	ctgggtatth	ttcgacactc	agttctctcc	ccggcagAAC	gtggaaaaca	540
	aaatccacat	aagtttgtgt	catggacggg	aggcgagaga	aaaatctctg	tgaaggaggt	600
15	aaagcactgt	gcaaatacca	gcttgacagg	cagtagcact	ggggtcccgg	gtcctttagc	660
	ttccagtccc	aggagttgct	cttgtctcct	cccactctgg	agtccgcaga	gtaggaagga	720
20	ggattaaacc	cgggggagga	gttccgcacc	agctccctat	cctgcgccag	cacgcctagc	780
	ctaagcgcgc	acatagagct	ccggtctccg	tcggtgccca	gccccggctg	tgcttcccag	840
	agcaagctcc	aggctccgca	agaccgcggg	gcctccagga	tgcagacacc	tcacaaggag	900
25	cacctgtaca	agctgctggt	gatcggcgac	ctggtagtgg	gcaagaccag	cattattaaa	960
	cggtagctgc	atcaaaatac	cgggtagggg	aggcgctttt	cccaaggcag	tctggagcat	1020
	gcgcttttagc	agccccgctg	ggcacttggc	gctacacaag	tggcctctgg	cctcgcacac	1080
30	attccacatc	caccggtagg	cgccaaccgg	ctccgttctt	tgggtggccc	ttcgcgccac	1140
	cttctactcc	tcccctagtc	aggaagttcc	ccccgcgcc	gcagctcgcg	tcgtgcagga	1200
35	cgtgacaaat	ggaagtagca	cgtctcacta	gtctcgtgca	gatggacagc	accgctgagc	1260
	aatggaagcg	ggtaggcctt	tggggcagcg	gccaatagca	gctttgctcc	ttcgctttct	1320
	gggctcagag	gctgggaagg	ggtgggtccg	ggggcgggct	caggggcggg	ctcaggggcg	1380
40	gggcgggcgc	ccgaaggtcc	tccggaggcc	cggcattctg	cacgcttcaa	aagcgcacgt	1440
	ctgccgcgct	gttctcctct	tcctcatctc	cgggcctttc	gacctgcagc	caatatggga	1500
	tcggccattg	aacaagatgg	attgcacgca	ggttctccgg	ccgcttgggt	ggagaggcta	1560
45	ttcggctatg	actgggcaca	acagacaatc	ggctgctctg	atgccgccgt	gttccggctg	1620
	tcagcgcagg	ggcgcccggt	tctttttgtc	aagaccgacc	tgtccggtgc	cctgaatgaa	1680
50	ctgcaggacg	aggcagcgcg	gctatcgtgg	ctggccacga	cgggcgttcc	ttgcgcagct	1740
	gtgctcgacg	ttgtcactga	agcgggaagg	gactggctgc	tattgggcga	agtgccgggg	1800
	caggatctcc	tgtcatctca	ccttgctcct	gccgagaaag	tatccatcat	ggctgatgca	1860
55	atgcggcggc	tgcatacgct	tgatccggct	acctgcccac	tcgaccacca	agcgaaacac	1920
	cgcacgcagc	gagcacgtac	tcggatggaa	gccggtcttg	tcgatcagga	tgatctggac	1980
60	gaagagcatc	aggggctcgc	gccagccgaa	ctgttcgccca	ggctcaaggc	gcgcagcccc	2040
65							

ES 2 709 216 T3

	gacggcgatg atctcgtcgt gacccatggc gatgcctgct tgccgaatat catggtggaa	2100
5	aatggccgct tttctggatt catcgactgt ggccggctgg gtgtggcgga ccgctatcag	2160
	gacatagcgt tggctacccg tgatattgct gaagagcttg gcggcgaatg ggctgaccgc	2220
	ttcctcgtgc tttacggtat cgcgcgtccc gattcgcagc gcatcgccct ctatcgccct	2280
10	cttgacgagt tcttctgagg ggatcaattc tctagagctc gctgatcagc ctcgactgtg	2340
	ccttctagtt gccagccatc tgttgtttgc ccctcccccg tgccttcctt gaccctggaa	2400
	ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt	2460
15	aggtgtcatt ctattctggg ggggtggggtg gggcaggaca gcaaggggga ggattgggaa	2520
	gacaatagca ggcattgctgg ggatgcgggtg ggctctatgg cttctgagggc ggaaagaacc	2580
	agctgggggt tctcctctca ttatcgagcc accattgggtg tggacttcgc gctgaagggtg	2640
20	ctccactggg acccagagac ggtggtgcgc ttgcagctct gggacattgc tggtagcgga	2700
	tcagagcagc gcgcaacggg tgagggtgga gtgagccagt gaggagtctg ggggtgaagg	2760
25	ttcggggagt ggaaaatgac ttttcagtcg gttccagtcc cgggaccctt gagtgcatac	2820
	aagcaggaga tccgcatcgc ctgggcgctc cactcttgga aagtttggct taatggcttg	2880
	gaaacctgat ttcaaagaaa tggaagtgtt ttcttttctt tctttccttt tttttttttt	2940
30	ttttttttct tttgctgttg tttctgttgg agtcgtcccc actctacctg taacttctag	3000
	ataacttcgc tggctctcac tggctgtgag aaagcgaacc actttctcct gggattcctg	3060
	ggtgcagaga aggctgtcgc ctggactcac aaggagattg tagtcgcatt cttgtttcat	3120
35	tctagtcctt ttctggacac aggtagccgc gacttgcccc agagtatctc acgtggcttt	3180
	catccttcgt gtttagaggg gaagccccta ggaaatttaa gaaggagcag gattatctta	3240
	ggaatttagt ttctttcaaa tctcactact atcatctcct tgcttattgg cctcttcagt	3300
40	cagaaaaatt tgagatgcta aatttgata catctagaac gaactatctc ttctcactcc	3360
	actcccctct tccccatctc tcttcogtct ccctccatcc ttggctatct cttcttcact	3420
45	ttccatttca aacaggagac tgtgtatgtt ttttaggaaa acattaaaaa aaaaaccaca	3480
	aaaacaaaaa caaacgggag acagggtccc gtcattgaac tctgctaacc tatatcaagc	3540
	tgaccttgac ctcatagaga ccacttgcc tctgcctccc tagaggcaag ggtcgggggt	3600
50	atggtgatgt taatgtcgtt tgctttaaga ttccctgatt tgatcttggg gtattttttg	3660
	agaaatctaa agtatgaaat cagagtttga ctaacagctt ctaccagctc ctagccacaa	3720
	taaagactga ggcaggctat agttagtgtc caatactggg tcctacctgg ctgcttgtaa	3780
55	cctgggcatg cctagcattc tagatgctaa ctcaccaaag cagtagcatt ttaagctgca	3840
	aatggctagg cagcgacagc tcaagaatct tcttgctttg gagtttttaa ctccaatgag	3900
60	attttccatg atccctttca aataacccta cttaatctct cttcatagcc cacagtacca	3960

65

ES 2 709 216 T3

	agaagccttt gataagctct ggattgaaaa gaagcagttc tttttcaaaa gatgtgctca	4020
5	tttgaactag tgcatttccc tggaaacact ttgccaggac ttgagatggg cactaagaag	4080
	gaaaattcct caaaggacat gtacagtctt gagatgcatt cgcttctgta gccatgagct	4140
	tgctggctct gagataaggt tagttgggtg agctaggttc atgggttggg gtctttggca	4200
10	gttctagaga agcatgagct attagagact tggagattgc atcaagtaga gccttttgag	4260
	cttttctactg tgtacctggg ccctctgtcg ctgcacgttt tagtgtctga aatgtctttc	4320
	agctgtagca gttttctcgg gacccagtt taaaatagct tactgtttaa aagatgtagc	4380
15	tgtagctagc attattgaac tagcataatt atagtctaaa tagcattatg tcttcagcct	4440
	tggttatatgt tggtaggttt tagtttcctc ttctaaacgg gaagaacaga aagatgtaat	4500
20	gattctgagc ttccagagtg agacacctct agagagaaat accttcttct gaagactacc	4560
	gtgtgattac agataaattc tgatatcttt gtttagcttt tgatatctat aaacagggag	4620
	tgtattttat ctctccaaat gagagaagaa taaacaataa tgcaaggtaa aggcaatagt	4680
25	gctacactct aggagttacc actctttgta catttattta taaatactaa gcaagaggaa	4740
	catgccatac atacactgac taagtcctaa caagtggcag ttcttatatc acacatttat	4800
30	cttgccctca aatgccagtc cagcatcagt ttagtctcat gcatttggca gcataaggca	4860
	gtttgagttc cacacttgct ctccagaagca atttaactcc cacacttggg aatcctttcc	4920
	taagccacag ttccagacca aagttttggt gaaggctata atcacagaag tctgcacaag	4980
35	tagggagtct gaaggatctg agctccattc agcagtcaga gcggcatcca accccaaggt	5040
	aatgctcagc tcactttgat aacttcaagc tcaaaggccc tgaactgctg agttggaggt	5100
	tgaaagatgt ttgggtaaaa gcaaggtaat tggcggatag gatggttgta acgtaattgt	5160
40	ttcaagttgt attagagacc tctgggttct aaggggatat gaaatccaac ctccactctc	5220
	cactgagatt caagttaggt taagtatgcc tttgagtacc ctcaagtcac agcatgccac	5280
45	tctccttttc ttaactctaa tatgtatcta taaagaacgg gtagtagtca actgagtcga	5340
	cggtatcgat aagcttgatc cagcttttgt tccctttagt gagggttaat tgcgcgcttg	5400
	gcgtaatcat ggtcatagct gtttcctgtg tgaaattggt atccgctcac aattccacac	5460
50	aacatacgag ccggaagcat aaagtgtaaa gcctggggtg cctaagagt gagctaactc	5520
	acattaattg cggtgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgccagctg	5580
55	cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct	5640
	tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcgg atcagctcac	5700
	tcaaagggcg taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga	5760
60	gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cggtgctggc gtttttccat	5820

65

ES 2 709 216 T3

	aggctccgcc	cccctgacga	gcatacaaaa	aatcgacgct	caagtcagag	gtggcgaaac	5880
5	ccgacaggac	tataaagata	ccaggcggtt	ccccctggaa	gctccctcgt	gcgctctcct	5940
	gttccgaccc	tgccgcttac	cggatacctg	tccgcctttc	tcccttcggg	aagcgtggcg	6000
	ctttctcata	gctcacgctg	taggtatctc	agttcgggtg	aggtcgttcg	ctccaagctg	6060
10	ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgcg	ccttatccgg	taactatcgt	6120
	cttgagtcca	acccggtaag	acacgactta	tgcgcactgg	cagcagccac	tggtaacagg	6180
	attagcagag	cgagggtatg	aggcgggtgct	acagagttct	tgaagtgggtg	gcctaactac	6240
15	ggctacacta	gaaggacagt	atgttggtatc	tgcgctctgc	tgaagccagt	taccttcgga	6300
	aaaagagttg	gtagctcttg	atccggcaaa	caaaccaccg	ctggtagcgg	tggttttttt	6360
20	gtttgcaagc	agcagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	6420
	tctacggggg	ctgacgctca	gtggaacgaa	aactcacggt	aagggatttt	ggcatgaga	6480
	ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	6540
25	taaagtatat	atgagtaaac	ttgggtctgac	agttaccaat	gcttaatcag	tgaggcacct	6600
	atctcagcga	tctgtctatt	tcgttcaccc	atagttgcct	gactccccgt	cgtgtagata	6660
	actacgatac	gggagggcct	accatctggc	cccagtgtcg	caatgatacc	gcgagaccca	6720
30	cgctcaccgg	ctccagattt	atcagcaata	aaccagccag	ccggaagggc	cgagcgcaga	6780
	agtggctcctg	caactttatc	cgccctcatc	cagtctatta	attggtgccg	ggaagctaga	6840
35	gtaagtagtt	cgccagttaa	tagtttgccg	aacggtgttg	ccattgctac	aggcatcgtg	6900
	gtgtcacgct	cgtcgtttgg	tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	6960
	gttacatgat	cccccatggt	gtgcaaaaaa	gcggttagct	ccttcggtcc	tccgatcgtt	7020
40	gtcagaagta	agttggccgc	agtgttatca	ctcatgggta	tggcagcact	gcataattct	7080
	cttactgtca	tgccatccgt	aagatgcttt	tctgtgactg	gtgagtactc	aaccaagtca	7140
	ttctgagaat	agtgtatgcg	gcgaccgagt	tgtctttgcc	cggcgtcaat	acgggataat	7200
45	accgcgccac	atagcagaac	tttaaaagtg	ctcatcattg	gaaaacgttc	ttcggggcga	7260
	aaactctcaa	ggatcttacc	gctgttgaga	tccagttcga	tgtaacccac	tcgtgcaccc	7320
50	aactgatctt	cagcatcttt	tactttcacc	agcgtttctg	ggtgagcaaa	aacaggaagg	7380
	caaaatgccg	caaaaaaggg	aataagggcg	acacggaaat	gttgaatact	catactcttc	7440
	ctttttcaat	attattgaag	catttatcag	ggttattgtc	tcatgagcgg	atacatattt	7500
55	gaatgtattt	agaaaaataa	acaaataggg	gttccgcgca	catttccccg	aaaagtgcc	7560
	cctaaattgt	aagcgttaat	attttggtta	aattcgcgtt	aaatttttgt	taaatcagct	7620
	cattttttta	ccaataggcc	gaaatcggca	aaatccctta	taaatcaaaa	gaatagaccg	7680
60	agataggggt	gagtgttggt	ccagtttgga	acaagagtcc	actattaaag	aacgtggact	7740

65

ES 2 709 216 T3

	ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg ccactacgt gaaccatcac	7800
5	cctaatacaag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac cctaaaggga	7860
	gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga	7920
	aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg cgcgtaacca	7980
10	ccacacccgc cgcgcttaat gcgcgcgtac agggcgcgctc ccattcgcca ttcaggctgc	8040
	gcaactgttg ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag	8100
15	ggggatgtgc tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt	8160
	gtaaaacgac ggccagttag cgcgcgtaat acgactcact atagggcgaa ttggagct	8218
	<210> 61	
	<211> 9989	
20	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> pCMV-Rab-Reportero (higro)	
25	<400> 61	
	gaattcgagc ttgcatgcct gcaggtcggt acataactta cggtaaatgg cccgcctggc	60
30	tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc catagtaacg	120
	ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg	180
35	gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa	240
	tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac	300
	atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg	360
40	cgtggatagc ggtttgactc acggggatgt ccaagtctcc acccatttga cgtcaatggg	420
	agtttggtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgccccca	480
45	ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag agctcgttta	540
	gtgaaccgtc agatcgcctg gagacgcat ccacgctggt ttgacctcca tagaagacac	600
	cgggaccgat ccagcctccg gactctagag gatccggtac tcgaggacac tgagagacc	660
50	tacttcacta acaaccggtg tggtcgccag tagcttggca ctggccgtcg ttttacaacg	720
	tcgtgactgg gaaaaccctg gcgttaccca acttaatcgc cttgcagcac atcccccttt	780
	cgccagctgg cgtaatatgc aagaggcccc caccgatcgc ccttcccaac agttgcgcag	840
55	cctgaatggc gaatggcgct ttgcctgggt tccggcacca gaagcgggtg cggaaagctg	900
	gctggagtgc gatcttcctg aggcggatag tgcgtcgtc ccctcaaact ggcagatgca	960
60	cggttacgat gcgcccattc acaccaacgt gacctatccc attacggtca atccgccgtt	1020
	tgttcccacg gagaatccga cgggttggtt ctgcgtcaca tttaatgttg atgaaagctg	1080

ES 2 709 216 T3

	gctataaaac	cggtacagtt	cggccacccat	ggtcgtatca	agcgctatgt	gcacccaaaac	1140
5	ttctcctcgc	actaccgggc	caccattggg	cgagtagctt	ggcactggcc	gtcgtttttac	1200
	aacgtcgtga	ctgggaaaac	cctggcggtta	cccaacttaa	tcgccttgca	gcacatcccc	1260
	ctttcgccag	ctggcgtaat	agcgaagagg	cccgccaccga	tcgcccttcc	caacagttgc	1320
10	gcagcctgaa	tggcgaatgg	cgctttgcct	ggtttccggc	accagaagcg	gtgccggaaa	1380
	gctggctgga	gtgcgatcct	cctgaggccg	atactgtcgt	cgtccctca	aactggcaga	1440
	tgcacgggtta	cgatgcgcc	atctacacca	acgtgacct	tccattacg	gtcaatccgc	1500
15	cgtttgttcc	cacggagaat	ccgacgggtt	gttactcgct	cacatttaat	gttgatgaaa	1560
	gctggctaca	ggaaggccag	acgcgaatta	ttttgatgg	cgtaaactcg	gcgtttcatc	1620
20	tgtggtgcaa	cgggcgctgg	gtcggttacg	gccaggacag	tcgtttgccg	tctgaatttg	1680
	acctgagcgc	atttttacgc	gccggagaaa	accgcctcgc	ggatgatggg	ctgcgctgga	1740
	gtgacggcag	ttatctggaa	gatcaggata	tgtggcggt	gagcggcatt	ttcogtgacg	1800
25	tctcgttgct	gcataaaccg	actacacaaa	tcagcgattt	ccatggtgcc	actcgtttta	1860
	atgatgattt	cagccgcgct	gtactggagg	ctgaagtcca	gatgtgcggc	gagttgcgtg	1920
	actacctacg	ggtaacagtt	tctttatggc	agggtgaaac	gcaggtcgcc	agcggcaccg	1980
30	cgcccttcgg	cggtgaaatt	atcgatgagc	gtggtggtta	tgcgatcgcc	gtcacactac	2040
	gtctgaacgt	cgaaaacccg	aaactgtgga	gcgccgaaat	cccgaatctc	tatcgtgcgg	2100
35	tgggtgaact	gcacaccgcc	gacggcacgc	tgattgaagc	agaagcctgc	gatgtcggtt	2160
	tccgcgaggt	gcggattgaa	aatggtctgc	tgctgctgaa	cggaagccg	ttgctgattc	2220
	gaggcggtta	ccgtcacgag	catcatcctc	tgcattggtca	ggcatggat	gagcagacga	2280
40	tgggtgcagga	tatcctgctg	atgaagcaga	acaacttta	cgccgtgcgc	tgctcgatt	2340
	atccgaacca	tccgctgtgg	tacacgctgt	gcgaccgcta	cgccctgtat	gtggtggatg	2400
	aagccaatat	tgaaacccac	ggcatggtgc	caatgaatcg	tctgaccgat	gatccgcgct	2460
45	ggctaccggc	gatgagcgaa	cgcgtaacgc	gaatggtgca	gcgcgatcgt	aatcaccgga	2520
	gtgtgatcat	ctggtcgctg	gggaatgaat	caggccacgg	cgctaatcac	gacgcgctgt	2580
	atcgctggat	caaactctgc	gatccttccc	gcccggtgca	gtatgaaggc	ggcgagccg	2640
50	acaccacggc	caccgatatt	atttgcccga	tgtacgcgcg	cgtggatgaa	gaccagccct	2700
	tccgggctgt	gccgaaatgg	tccatcaaaa	aatggcttcc	gctacctgga	gagacgcgcc	2760
55	cgctgatcct	ttgcgaatac	gccacgcgca	tgggtaacag	tcttgccggg	ttcgctaaat	2820
	actggcaggc	gtttcgtcag	tatccccgtt	tacaggcgcg	cttcgtctgg	gactgggtgg	2880
	atcagtcgct	gattaaatat	gatgaaaacg	gcaaccgctg	gtcggttac	ggcggtgatt	2940
60	ttggcgatac	gccgaacgat	cgccagttct	gtatgaacgg	tctggtcttt	gccgaccgca	3000

ES 2 709 216 T3

	cgccgcatcc agcgctgacg gaagcaaaac accagcagca gtttttccag ttccgtttat	3060
5	ccggggcaaac catcgaagtg accagcgaat acctgttccg tcatagcgat aacgagctcc	3120
	tgcactggat ggtggcgctg gatggtaagc cgctggcaag cggatgaagtg cctctggatg	3180
	tcgctccaca aggtaaacag ttgattgaac tgcctgaact accgcagccg gagagcgcg	3240
10	ggcaactctg gctcacagta cgcgtagtgc aaccgaacgc gaccgcatgg tcagaagccg	3300
	ggcacatcag cgcctggcag cagtggcgtc tggcgaaaa cctcagtgtg acgctccccg	3360
15	ccgcgtccca cgccatcccg catctgacca ccagcgaaat ggatttttgc atcgagctgg	3420
	gtaataagcg ttggcaattt aaccgccagt caggctttct ttcacagatg tggattggcg	3480
	ataaaaaaca actgctgacg ccgctgcgcg atcagttcac ccgtgcaccg ctggataacg	3540
20	acattggcgt aagtgaagcg acccgcatg accctaacgc ctgggtcgaa cgctggaagg	3600
	cgccggggcca ttaccaggcc gaagcagcgt tgttgcatg caccggcagat acacttgctg	3660
	atgcggtgct gattacgacc gctcacgcgt ggcagcatca ggggaaaacc ttatttatca	3720
25	gccgaaaaac ctaccggatt gatggtagt gtcaaattggc gattaccgtt gatgttgaag	3780
	tggcgagcga tacaccgcat ccggcgcgga ttggcctgaa ctgccagctg gcgcaggtag	3840
30	cagagcgggt aaactggctc ggattagggc cgcaagaaaa ctatcccgac cgccttactg	3900
	ccgcctgttt tgaccgctgg gatctgccat tgtcagacat gtataccccg tacgtcttcc	3960
	cgagcgaaaa cggctcgcgc tgcgggacgc gcgaattgaa ttatggcca caccagtggc	4020
35	gcggcgactt ccagttcaac atcagccgct acagtcaaca gcaactgatg gaaaccagcc	4080
	atcgccatct gctgcacgcg gaagaaggca catggctgaa tatcgacggt ttccatatgg	4140
	ggattggtgg cgacgactcc tggagcccgt cagtatcggc ggaattccag ctgagcgcg	4200
40	gtcgctacca ttaccagttg gtctggtgtc aggggatccc ccgggctgca gccaatatgg	4260
	gatcgcccat tgaacaagat ggattgcacg cagggttctcc ggccgcttgg gtggagaggc	4320
45	tattcggcta tgactgggca caacagacaa tcggctgctc tgatgccgcc gtgttccggc	4380
	tgtcagcgca gggcgccccg gttctttttg tcaagaccga cctgtccggt gccctgaatg	4440
	aactgcagga cgaggcagcg cggctatcgt ggctggccac gacgggctt ccttgcgcag	4500
50	ctgtgctcga cgttgtcact gaagcgggaa gggactggct gctattgggc gaagtgccgg	4560
	ggcaggatct cctgtcatct caccttgctc ctgccgagaa agtatccatc atggctgatg	4620
55	caatgcggcg gctgcatacg cttgatccgg ctacctgcc attcgaccac caagcgaaac	4680
	atcgcatcga gcgagcacgt actcggatg aagccggtct tgtcgatcag gatgatctgg	4740
	acgaagagca tcaggggctc gcgccagccg aactgttcgc caggctcaag gcgcgcatgc	4800
60	ccgacggcga ggatctcgtc gtgacctatg gcgatgcctg cttgccgaat atcatggtgg	4860

65

ES 2 709 216 T3

	aaaatggccg cttttctgga ttcacgcact gtggccggct ggggtgtggcg gaccgctatc	4920
5	aggacatagc gttggctacc cgtgatattg ctgaagagct tggcggcgaa tgggctgacc	4980
	gcttcctcgt gctttacggt atcgccgctc ccgattcgca gcgcacgcgc ttctatcgcc	5040
	ttcttgacga gttcttctga ggggatcaat tctctagagc tcgctgatca gcctcgactg	5100
10	tgccttctag ttgccagcca tctgttggtt gccctcccc cgtgccttcc ttgacctgg	5160
	aagggtgccac tcccactgtc ctttcctaataaaaatgagga aattgcatcg cattgtctga	5220
	gtaggtgtca ttctattctg ggggggtgggg tggggcgagga cagcaagggg gaggattggg	5280
15	aagacaatag caggcatgct ggggatgcgg tgggctctat ggcttctgag acggaagaa	5340
	ccagctgggg ctcgatcctc tagagtcgac gtttgatctg atatcatcga tgaattctac	5400
20	cgggtagggg aggcgctttt cccaaggcag tctggagcat gcgctttagc agccccgctg	5460
	ggcacttggc gctacacaag tggcctctgg cctcgcacac attccacatc caccggtagg	5520
	cgccaaccgg ctccgttctt tgggtggcccc ttgcgcgcac cttctactcc tcccctagtc	5580
25	aggaagtcc cccccgccc gcagctcgcg tcgtgcagga cgtgacaaat ggaagtagca	5640
	cgtctcacta gtctcgtgca gatggacagc accgctgagc aatggaagcg ggtaggcctt	5700
	tggggcgagc gccaatagca gctttgctcc ttgcgtttct gggctcagag gctgggaagg	5760
30	ggtgggtccg ggggcgggct caggggcggg ctccagggcg gggcgggcgc ccgaaggctc	5820
	tccggaggcc cggcattctg cacgcttcaa aagcgcacgt ctgccgcgct gttctcctct	5880
35	tcctcatctc cgggcctttc gaccgatcca gccgccacca tgaaaaagcc tgaactcacc	5940
	gcgacgtctg tcgagaagtt tctgatcgaa aagtctgaca gcgtctccga cctgatgcag	6000
	ctctcggagg gcgaagaatc tcgtgctttc agcttcgatg taggagggcg tggatatgtc	6060
40	ctgcgggtaa atagctgcgc cgatgggttc taaaaagatc gttatgttta tcggcacttt	6120
	gcacgcggcc cgtcccgat tccggaagtg cttgacattg ggggaattcag cgagagcctg	6180
	acctattgca tctcccgccg tgcacagggt gtcacgttgc aagacctgcc tgaaacgaa	6240
45	ctgcccgcgt ttctgcagcc ggtcgcggag gccatggatg cgatcgctgc ggccgatctt	6300
	agccagacga gcgggttcgg cccattcgga ccgcaaggaa tcggtcaata cactacatgg	6360
	cgtgatttca tatgcgcgat tgctgatccc catgtgtatc actggcaaac tgtgatggac	6420
50	gacaccgtca gtgcgtccgt cgcgcaggct ctcgatgagc tgatgctttg ggccgaggac	6480
	tgccccgaag tccggcacct cgtgcacgcg gatttcggct ccaacaatgt cctgacggac	6540
55	aatggccgca taacagcggg cattgactgg agcgaggcga tgttcgggga ttcccaatac	6600
	gaggctcgcca acatcttctt ctggaggccg tggttggctt gtatggagca gcagacgcgc	6660
	tacttcgagc ggaggcatcc ggagcttgca ggatcgccgc ggctccgggc gtatatgttc	6720
60	cgcattgggtc ttgaccaact ctatcagagc ttgggtgacg gcaatttcga tgatgcagct	6780

65

ES 2 709 216 T3

	tgggcgcagg gtcgatgcga cgcaatcgtc cgatccggag ccgggactgt cgggcgtaca	6840
5	caaatcgccc gcagaagcgc ggccgtctgg accgatggct gtgtagaagt actcgccgat	6900
	agtggaaacc gacgccccag cactcgctccg agggcaaagg aatagtcgag aaattgatga	6960
	tctattaaac aataaagatg tccactaaaa tgggaagtttt tcctgtcata ctttgtaag	7020
10	aagggtgaga acagagtacc tacatcttga atggaaggat tggagctacg ggggtggggg	7080
	tggggtggga ttagataaat gcctgctctt tactgaaggc tctttactat tgctttatga	7140
15	taatgtttca tagttggata tcataattta aacaagcaaa accaaattaa gggccagctc	7200
	attcctccca ctcatgatct atagatcaaa catgcatgaa gtctctattc cgaagttcct	7260
	attctctaga aagtatagga acttcataaa acctgcaggc atgcaagcga tcgcggccgg	7320
20	ccaaggcccg cggggccact agttctagag cggccagctt ggcgtaatca tggcatagc	7380
	tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca	7440
25	taaagtgtaa agcctggggt gcctaagtag tgagctaact cacattaatt gcgttgcgct	7500
	cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac	7560
	gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttctctcgctc actgactcgc	7620
30	tgcgctcggt cgctcgctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgg	7680
	tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg	7740
	ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg	7800
35	agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat	7860
	accaggcggt tccccctgga agctccctcg tgcgctctcc tggtccgacc ctgccgtta	7920
40	ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct	7980
	gtaggtatct cagttcggtg taggtcgctt gctccaagct gggctgtgtg cacgaacccc	8040
	ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa	8100
45	gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg	8160
	taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact agaaggacag	8220
50	tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt	8280
	gatccggcaa acaaaccacc gctggtagcg gtgggttttt tgtttgcaag cagcagatta	8340
	cgcgagaaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc	8400
55	agtggaacga aaactcacgt taagggattt tggcatgag attatcaaaa aggatcttca	8460
	cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa	8520
60	cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat	8580
	ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct	8640

65

ES 2 709 216 T3

	taccatctgg	ccccagtgct	gcaatgatac	cgcgagaccc	acgctcacccg	gctccagatt	8700
5	tatcagcaat	aaaccagcca	gccggaaggg	ccgagcgcag	aagtggtcct	gcaactttat	8760
	ccgcctccat	ccagtctatt	aattgttgcc	gggaagctag	agtaagtagt	tcgccagtta	8820
10	atagtttgcg	caacgttggt	gccattgcta	caggcatcgt	ggtgtcacgc	tcgtcgtttg	8880
	gtatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	tcccccatgt	8940
	tgtgcaaaaa	agcggttagc	tccttcggtc	ctccgatcgt	tgtcagaagt	aagttggccg	9000
15	cagtgttatc	actcatgggt	atggcagcac	tgcataattc	tcttactgtc	atgccatccg	9060
	taagatgctt	ttctgtgact	ggtgagtact	caaccaagtc	attctgagaa	tagtgtatgc	9120
20	ggcgaccgag	ttgctcttgc	ccggcgtcaa	tacgggataa	taccgcgcca	catagcagaa	9180
	ctttaaaagt	gctcatcatt	ggaaaacgtt	cttcggggcg	aaaactctca	aggatcttac	9240
	cgctggtgag	atccagttcg	atgtaaccca	ctcgtgcacc	caactgatct	tcagcatctt	9300
25	ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaatgcc	gcaaaaaagg	9360
	gaataagggc	gacacggaaa	tgttgaatac	tcatactctt	cctttttcaa	tattattgaa	9420
30	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	9480
	aacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	acctgacgtc	taagaaacca	9540
	ttattatcat	gacattaacc	tataaaaata	ggcgtatcac	gaggcccttt	cgtctcgcgc	9600
35	gtttcggtga	tgacggtgaa	aacctctgac	acatgcagct	cccggagacg	gtcacagctt	9660
	gtctgtaagc	ggatgccggg	agcagacaag	cccgtcaggg	cgcgtcagcg	ggtgttgggc	9720
40	ggtgtcgggg	ctggcttaac	tatgcggcat	cagagcagat	tgtactgaga	gtgcaccata	9780
	tgcggtgtga	aataccgcac	agatgcgtaa	ggagaaaata	ccgcatcagg	cgccattcgc	9840
	cattcaggct	gcgcaactgt	tgggaagggc	gatcgggtgc	ggcctcttcg	ctattacgcc	9900
45	agctggcgaa	agggggatgt	gctgcaaggc	gattaagttg	ggtaacgcca	gggttttccc	9960
	agtcacgacg	ttgtaaaacg	acggccagt				9989

50

55

60

65

Reivindicaciones

1. Una molécula de ácido nucleico que codifica
5 (I) un polipéptido que tiene la actividad de una endonucleasa, que es
(a) una molécula de ácido nucleico que codifica un polipéptido que consiste en la secuencia de aminoácidos de la sec. con núm. de ident.:1;
10 (b) una molécula de ácido nucleico que consiste en la secuencia de nucleótidos de la sec. con núm. de ident.: 2;
(c) una molécula de ácido nucleico que codifica una endonucleasa, cuya secuencia de aminoácidos es al menos 70 % idéntica a la secuencia de aminoácidos de la sec. con núm. de ident.:1;
(d) una molécula de ácido nucleico que consiste en una secuencia de nucleótidos que es al menos 50 % idéntica a la secuencia de nucleótidos de la sec. con núm. de ident.: 2;
15 (e) una molécula de ácido nucleico que es degenerada con respecto a la molécula de ácido nucleico de (d); o
(f) una molécula de ácido nucleico correspondiente a la molécula de ácido nucleico de cualquiera de (a) hasta (e) en donde T se reemplaza por U;
(II) un fragmento del polipéptido de (I) que tiene la actividad de una endonucleasa.
- 20 2. La molécula de ácido nucleico de conformidad con la reivindicación 1, en donde en (I)(c) en dicha secuencia de aminoácidos que tiene al menos 70 % de identidad de secuencia con respecto a la sec. con núm. de ident.: 1 los residuos de aminoácidos P66, D67, D84 y/o K86 de la sec. con núm. de ident.: 1 no están modificados.
- 25 3. La molécula de ácido nucleico de conformidad con la reivindicación 1 o 2 que codifica, además, un dominio de unión al ADN.
4. La molécula de ácido nucleico de conformidad con la reivindicación 3, en donde el dominio de unión al ADN es un motivo efector TAL de una proteína efectora TAL.
- 30 5. Un vector que comprende la molécula de ácido nucleico de conformidad con cualquiera de las reivindicaciones 1 a 4.
6. Una célula huésped que comprende la molécula de ácido nucleico de conformidad con cualquiera de las reivindicaciones 1 a 4 o el vector de conformidad con la reivindicación 5.
- 35 7. Una proteína o proteína de fusión que tiene la actividad de una endonucleasa en donde dicha proteína o proteína de fusión está codificada por la molécula de ácido nucleico de conformidad con cualquiera de las reivindicaciones 1 a 4 o el vector de conformidad con la reivindicación 5.
- 40 8. Un método para la modificación *in vitro* de una secuencia diana en el genoma de una célula eucariótica, el método comprende la etapa:
(a) introducir en dicha célula la molécula de ácido nucleico de conformidad con la reivindicación 3 o 4, el vector de conformidad con la reivindicación 5 o la proteína o proteína de fusión de conformidad con la reivindicación 7.
- 45 9. El método de conformidad con la reivindicación 8, en donde la modificación de dicha secuencia diana es mediante recombinación homóloga con una secuencia de ácido nucleico donante, que comprende, además, la etapa:
(b) introducir una molécula de ácido nucleico en dicha célula, en donde dicha molécula de ácido nucleico comprende dicha secuencia de ácido nucleico donante, en donde dicha secuencia de ADN donante está flanqueada en el extremo 5' por un primer elemento flanqueante y en el extremo 3' por un segundo elemento flanqueante, en donde dicho primer y segundo elementos flanqueantes son diferentes y en donde cada uno de dichos primer y segundo elementos flanqueantes son homólogos a una secuencia de ADN continua sobre ambos lados de la doble hebra introducidas en (a) de conformidad con la reivindicación 8 dentro de dicha secuencia diana en el genoma de dicha célula eucariótica.
- 50 10. El método de conformidad con las reivindicaciones 8 o 9, en donde dicha célula se analiza para determinar la modificación exitosa de dicha secuencia diana en el genoma.
- 60 11. El método de conformidad con cualquiera de las reivindicaciones 8 a 10, en donde la célula se selecciona a partir del grupo que consiste en una célula de mamífero o vertebrado, una célula vegetal o una célula de hongo.
12. El método de conformidad con cualquiera de las reivindicaciones 8 a 11, en donde la célula es un ovocito no humano.

13. Un método para producir un vertebrado no humano o un mamífero no humano que porta una secuencia diana modificada en su genoma; el método que comprende producir una célula no humana de acuerdo con el método de conformidad con cualquiera de las reivindicaciones 9 a 12.
- 5 14. El método de conformidad con cualquiera de las reivindicaciones 8 a 13, en donde la célula se selecciona a partir del grupo que consiste en roedores, perros, félidos, primates, conejos, cerdos, vacas, pollos, pavos, faisanes, patos, gansos, codornices, avestruz, emúes, casuarios y pez cebra.
- 10 15. Un método para producir la proteína o proteína de fusión de conformidad con la reivindicación 7 que comprende las etapas: (a) cultivar la célula huésped de conformidad con la reivindicación 6 y (b) aislar la proteína o proteína de fusión que se produce.
- 15
- 20
- 25
- 30
- 35
- 40
- 45
- 50
- 55
- 60
- 65

Figura 1

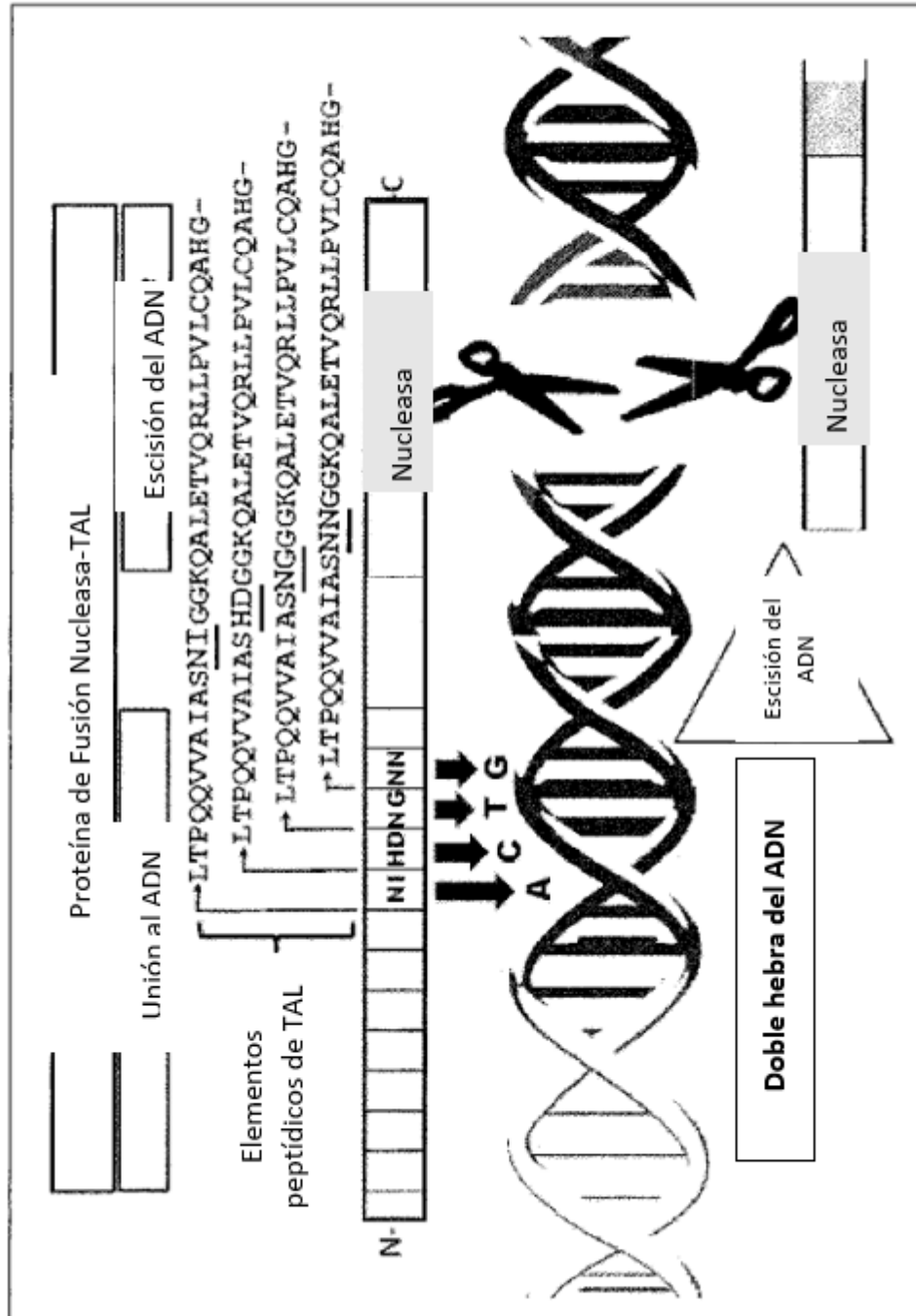


Figura 2

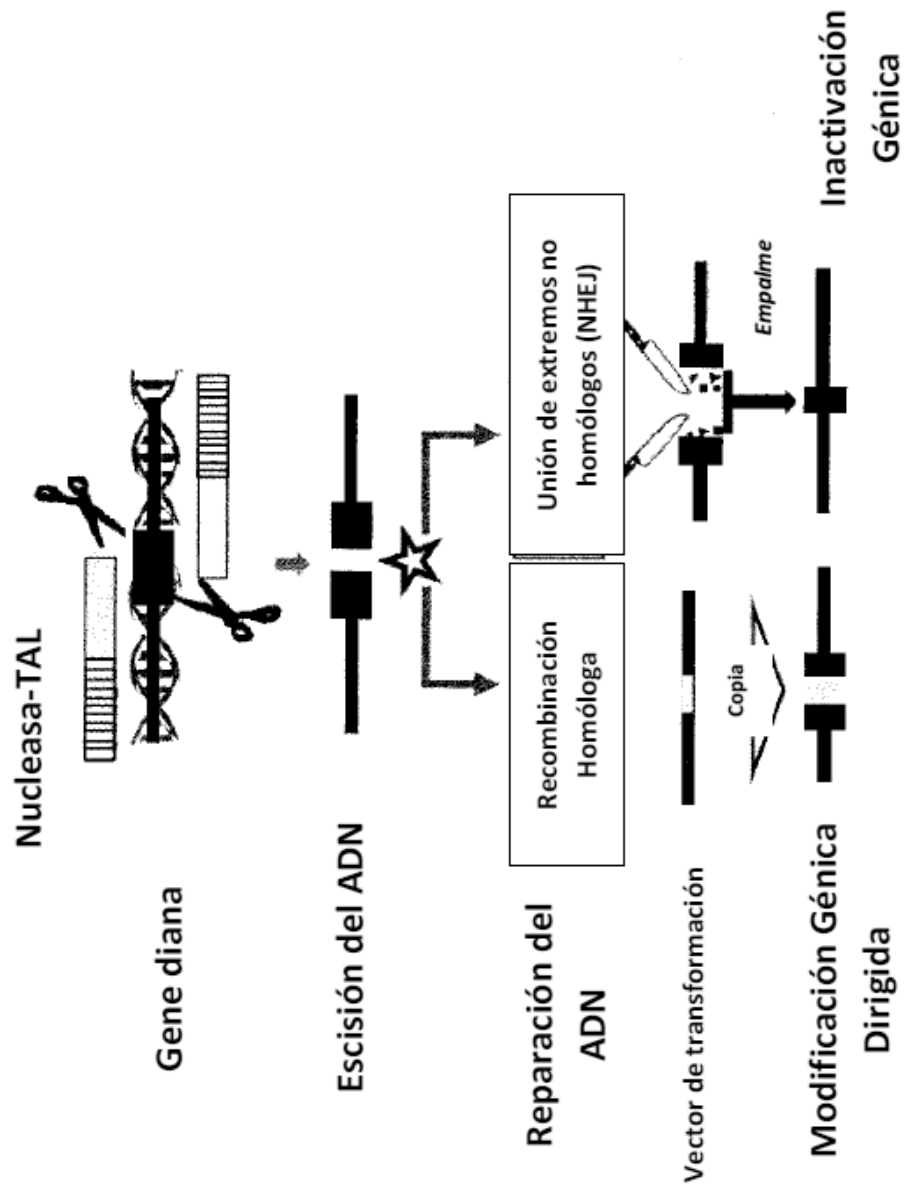


Figura 3

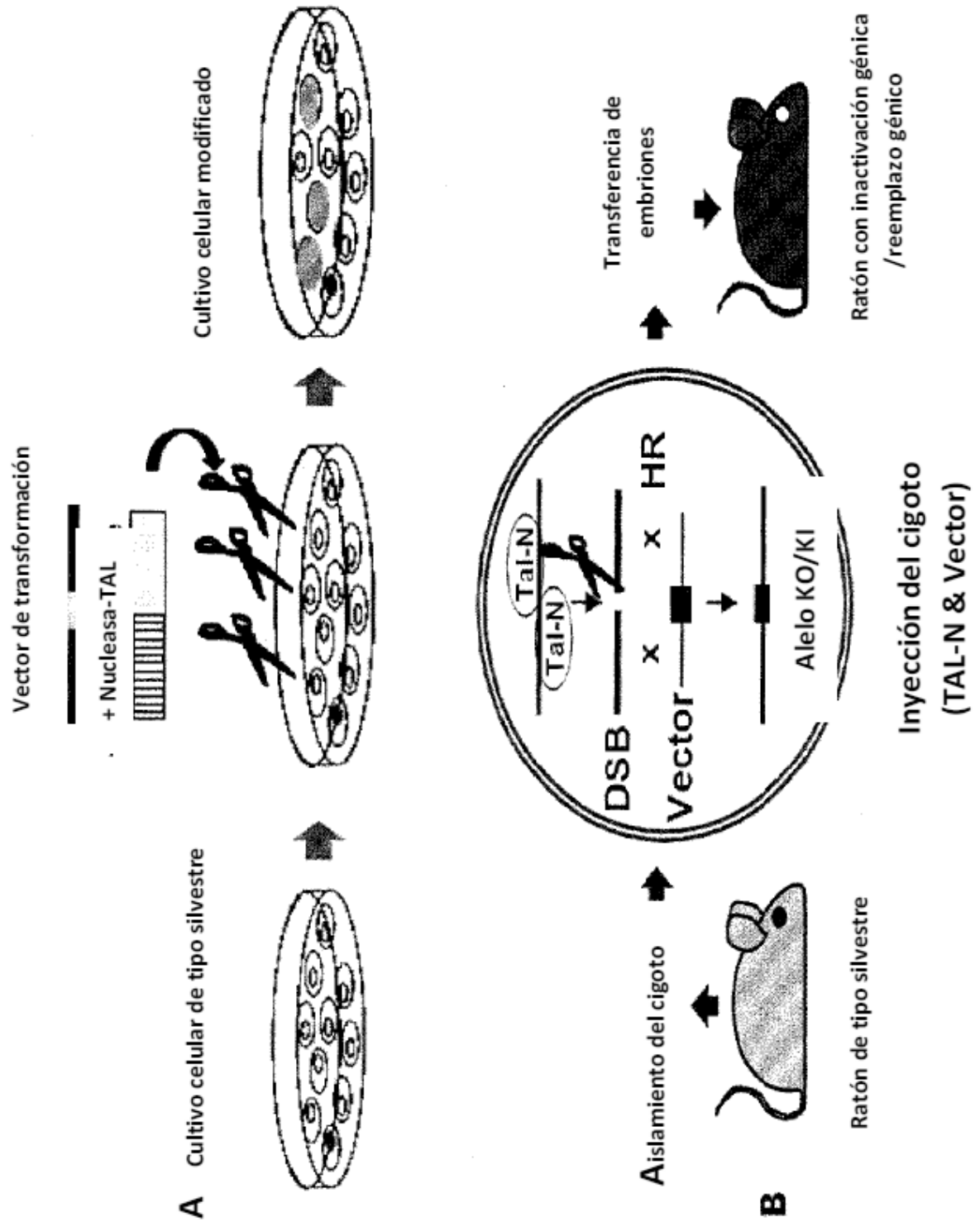


Figura 4

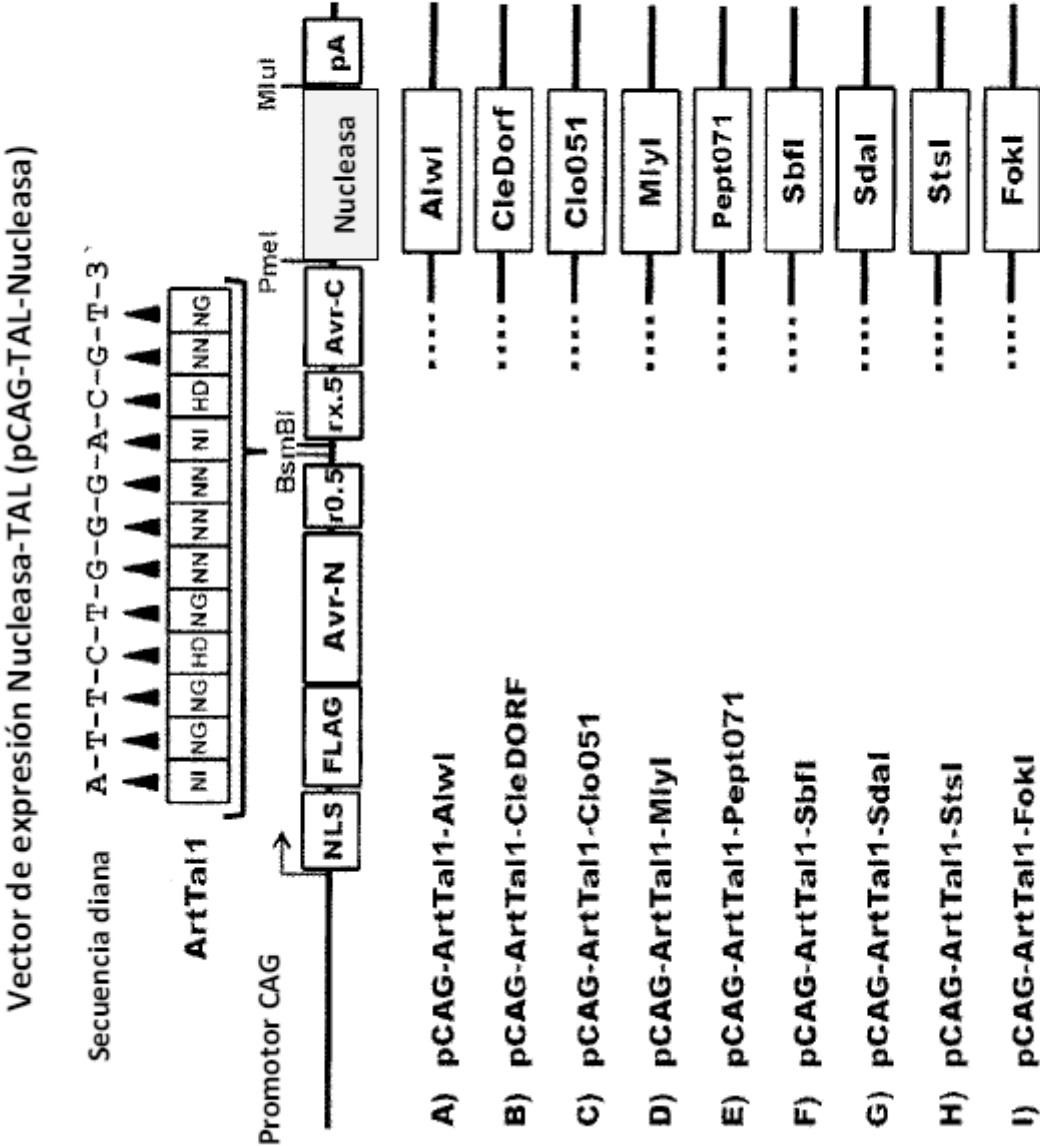


Figura 5

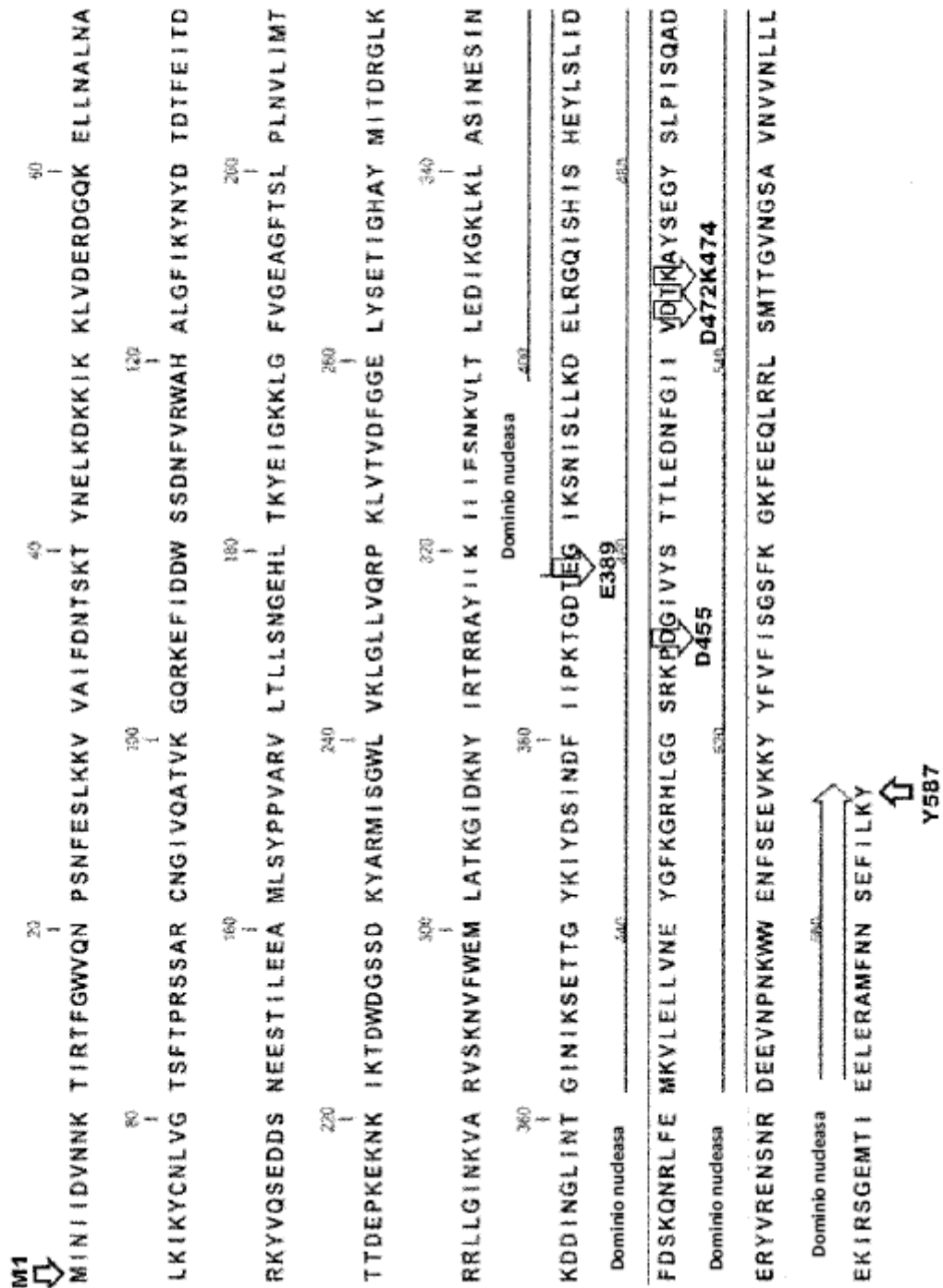


Figura 6

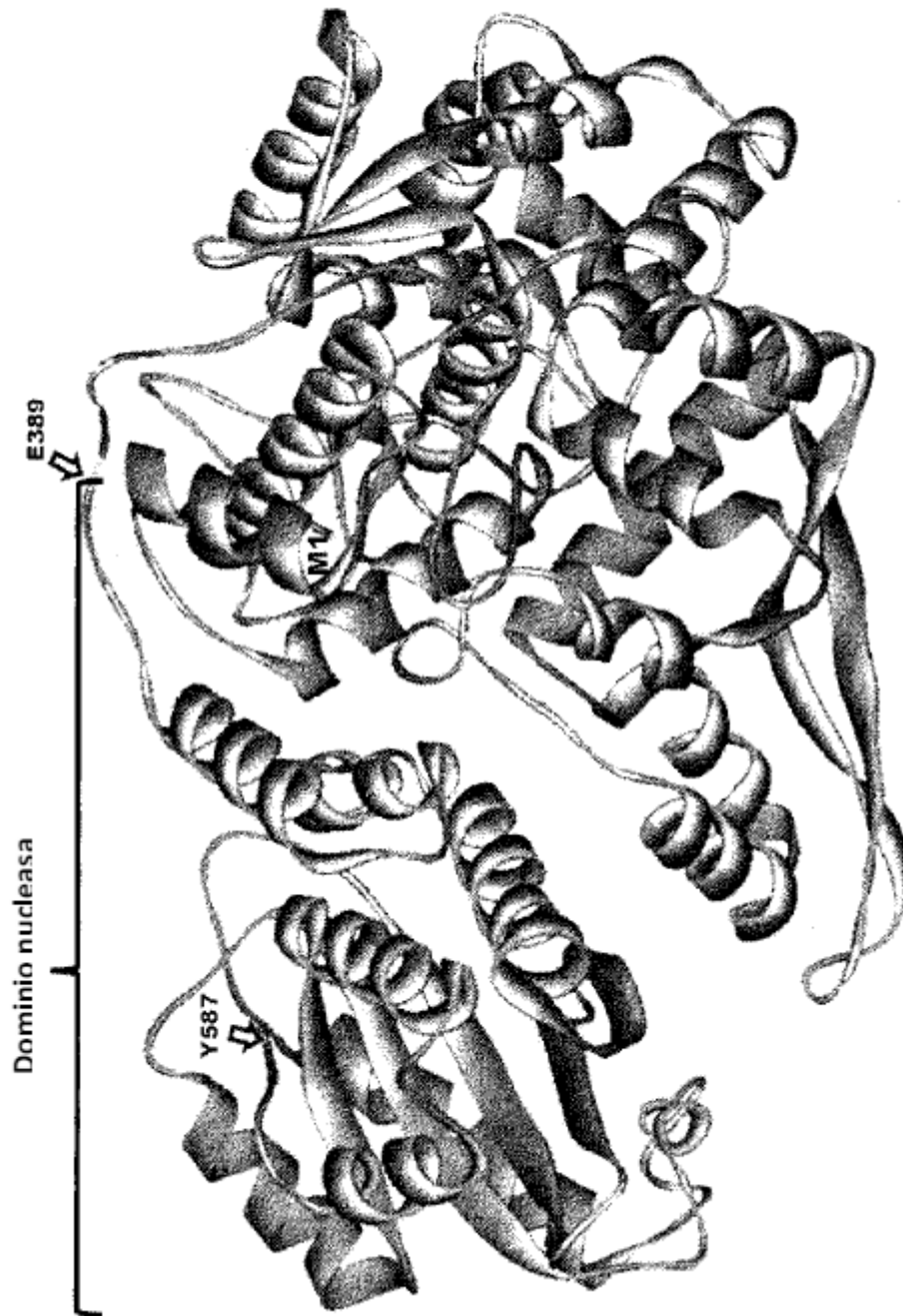


Figura 7

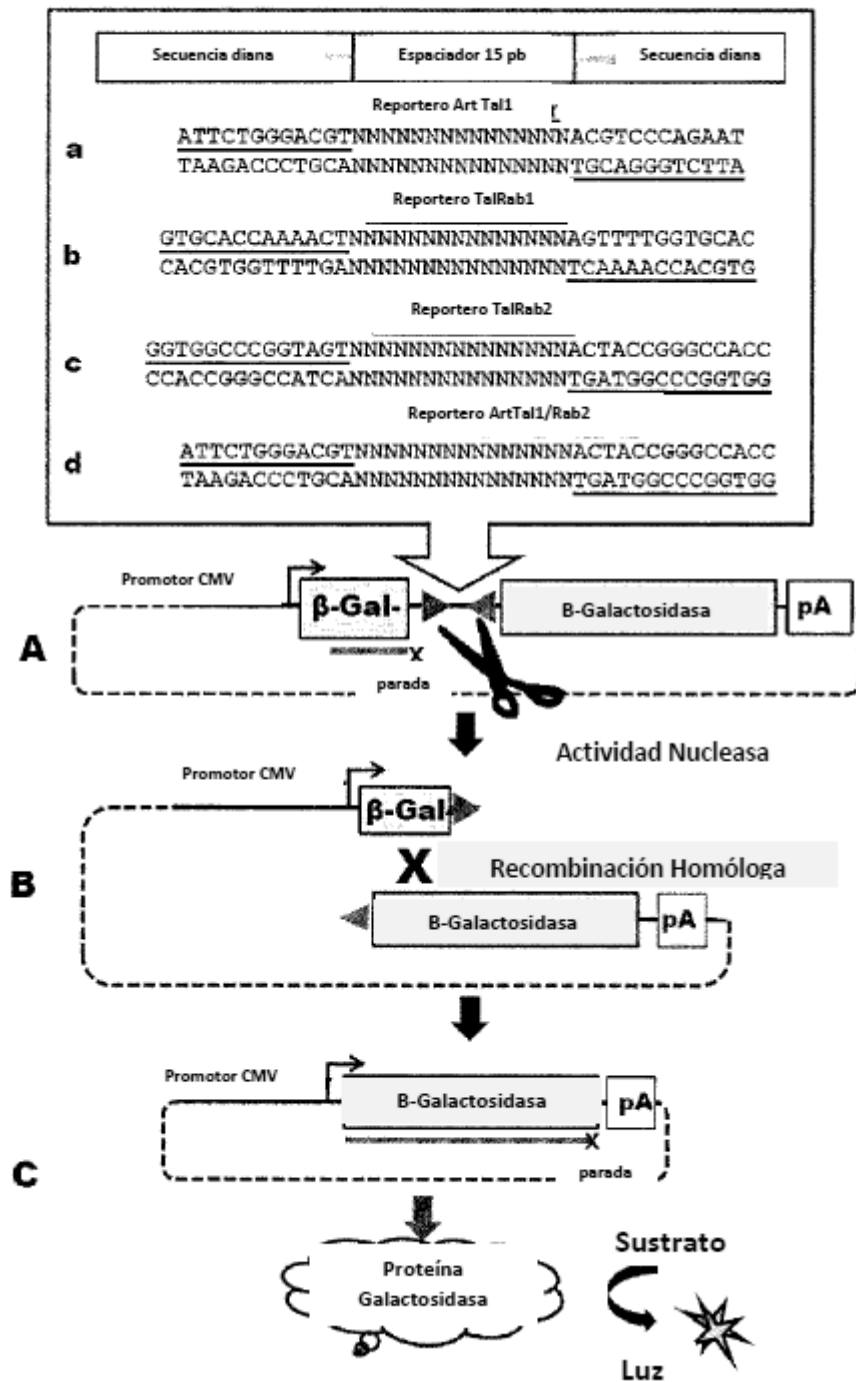


Figura 8

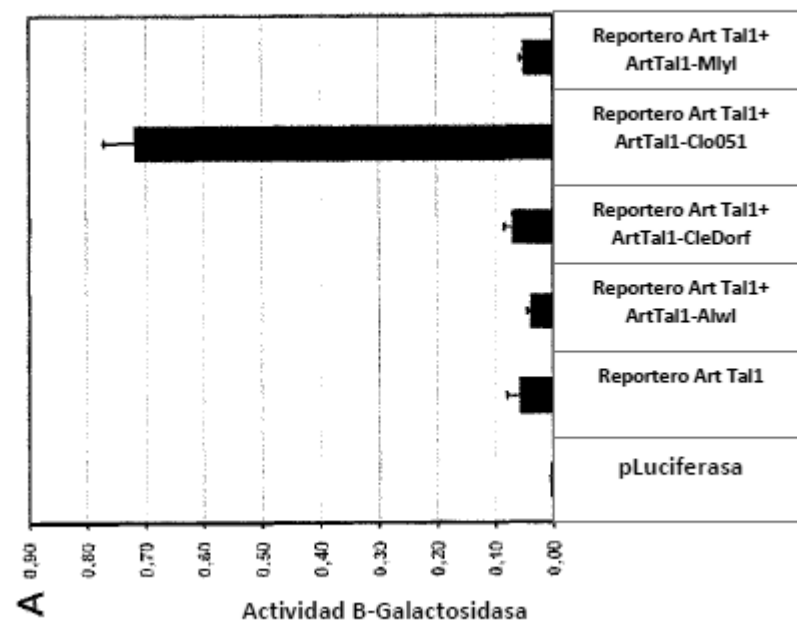
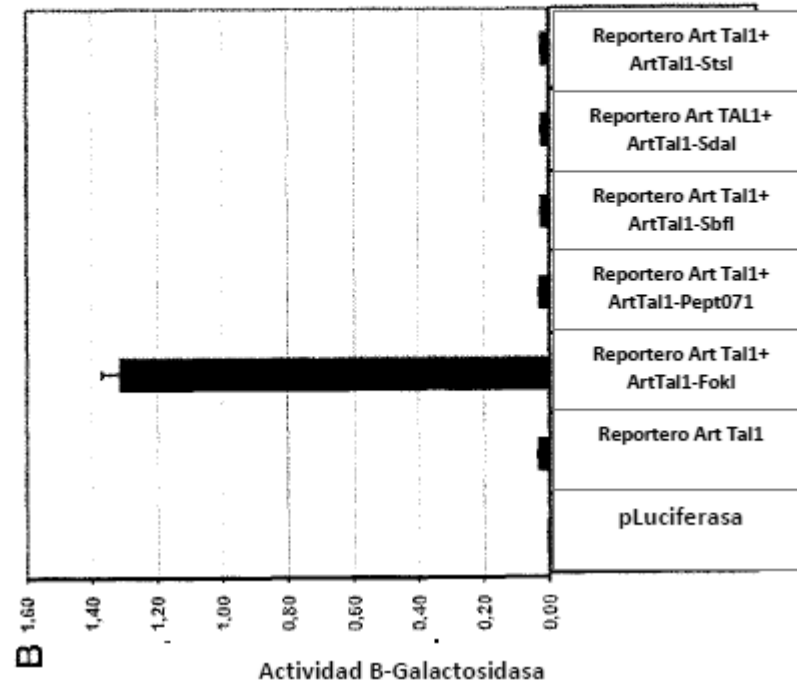


Figura 9

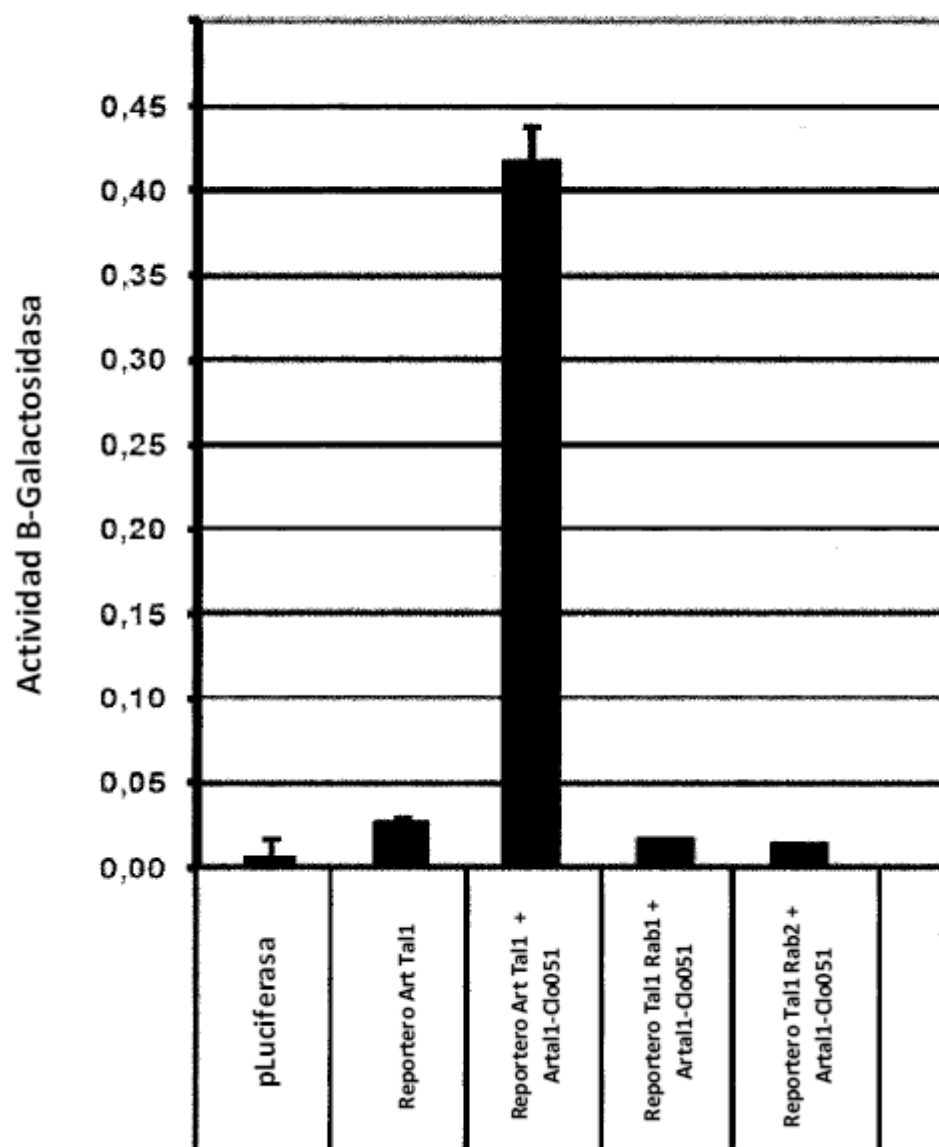


Figura 10

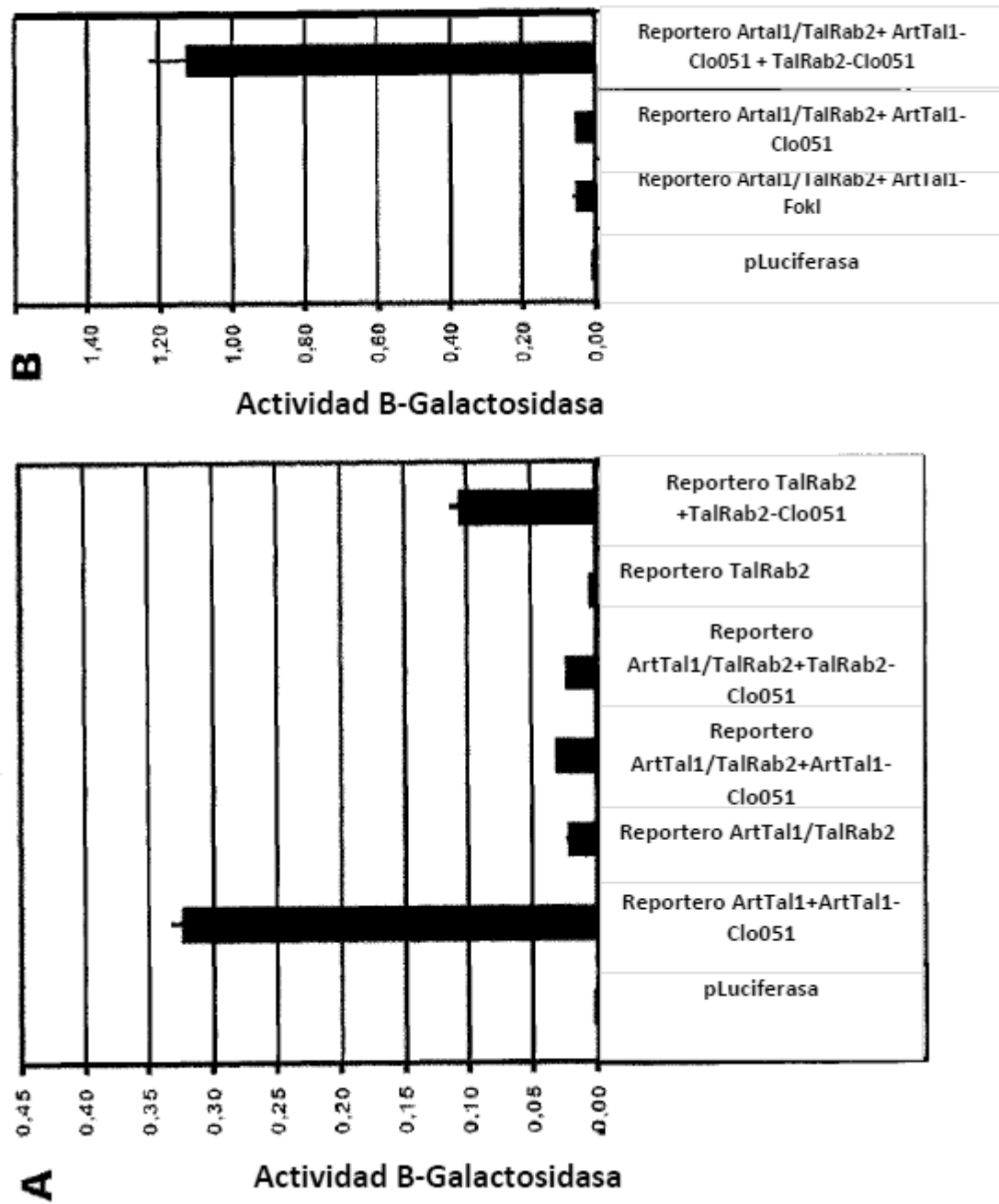


Figura 11

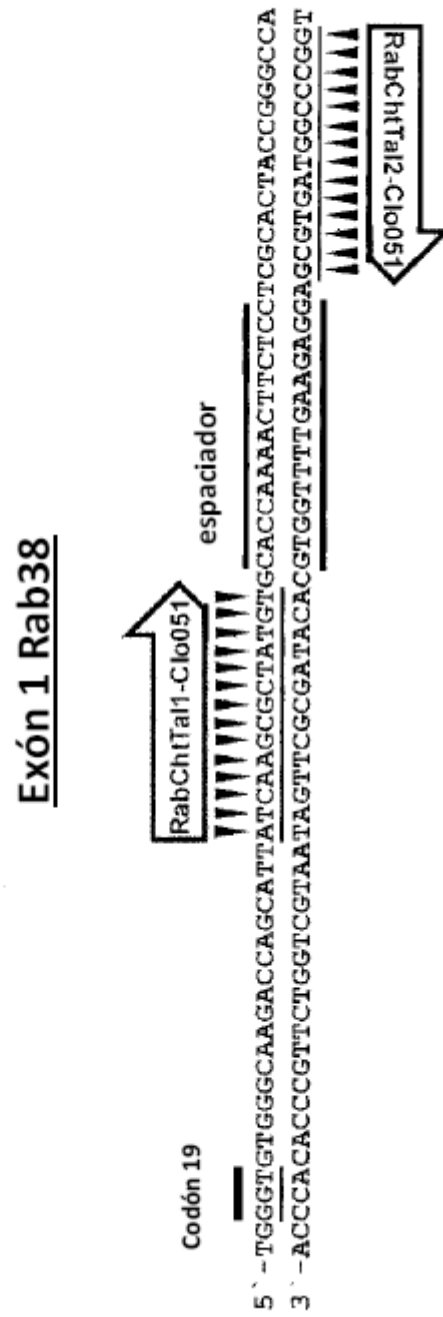


Figura 12

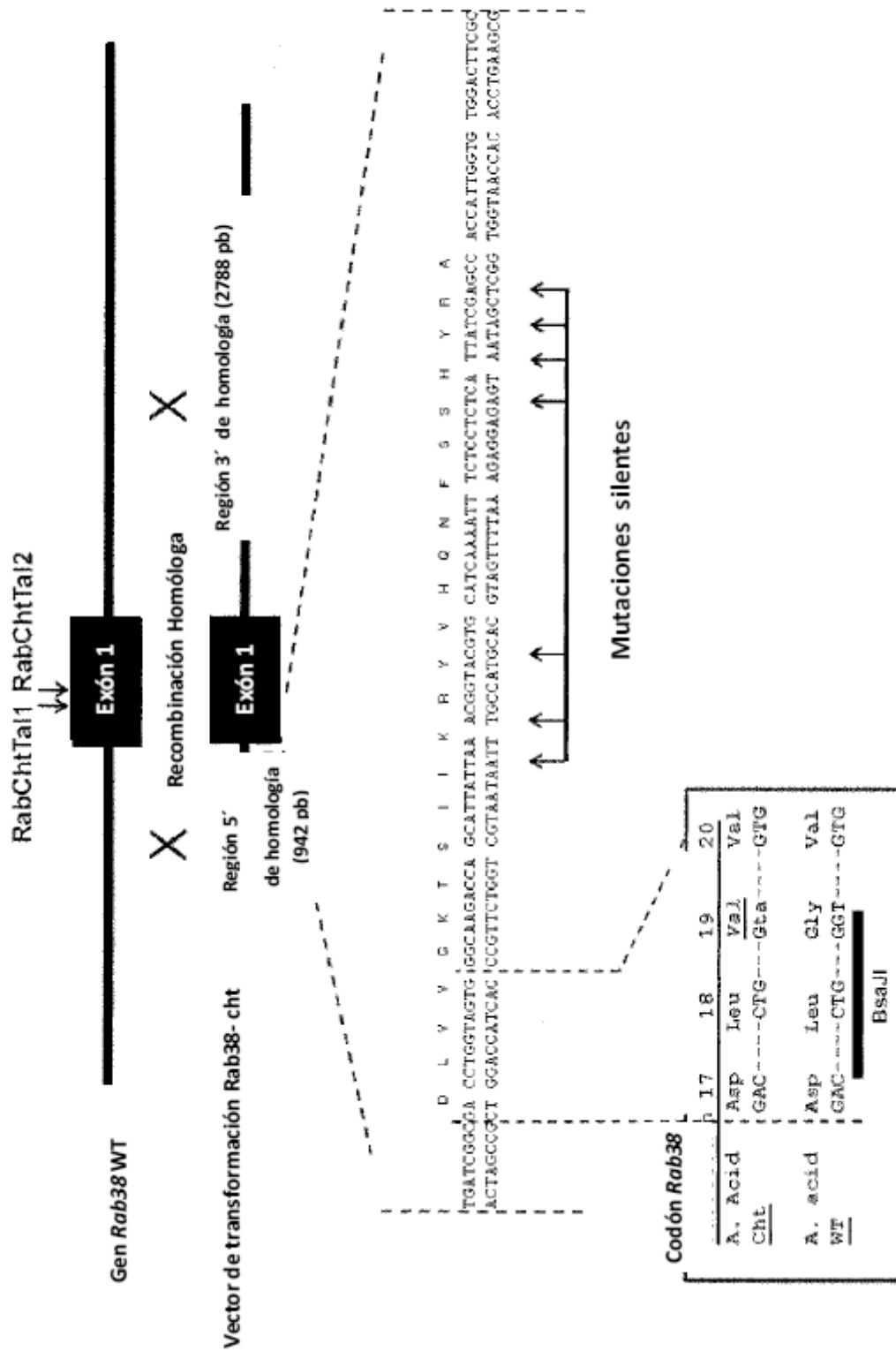


Figura 13

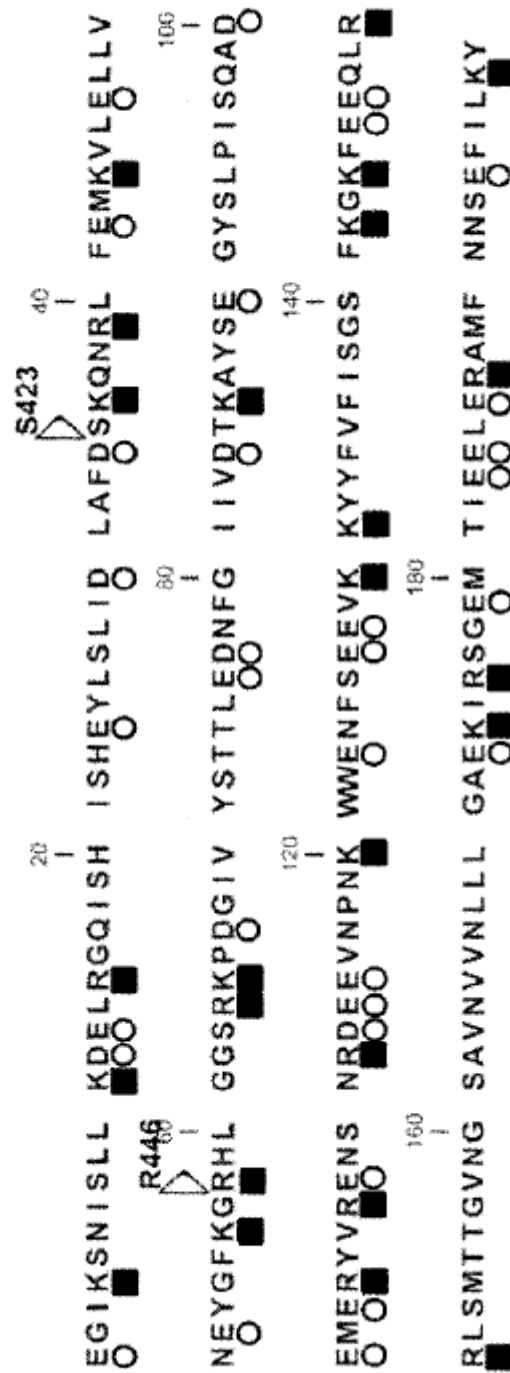


Figura 14

