



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 346 645**

51 Int. Cl.:
C12N 15/11 (2006.01)
C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **03707900 .1**
96 Fecha de presentación : **12.03.2003**
97 Número de publicación de la solicitud: **1490489**
97 Fecha de publicación de la solicitud: **29.12.2004**

54 Título: **Procedimientos y medios de supervisión y modulación del silenciamiento génico.**

30 Prioridad: **14.03.2002 US 363852 P**

45 Fecha de publicación de la mención BOPI:
19.10.2010

45 Fecha de la publicación del folleto de la patente:
19.10.2010

73 Titular/es: **Commonwealth Scientific and Industrial
Research Organisation
Limestone Avenue
Campbell, ACT 2612, AU**

72 Inventor/es: **Waterhouse, Peter;
Wesley, Susan y
Helliwell, Chris**

74 Agente: **Carpintero López, Mario**

ES 2 346 645 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

Procedimientos y medios de supervisión y modulación del silenciamiento génico.

5 **Campo de la invención**

La presente invención se refiere a procedimientos para alterar la expresión de genes en organismos eucariotas, tales como plantas, pero también en animales tales como nemátodos, insectos y artrópodos, mamíferos incluyendo humanos, o levaduras, hongos o mohos, usando ARNbc capaz de alterar la expresión de genes diana, o genes que codifican dicho ARNbc. También se proporcionan organismos eucariotas no humanos que comprenden dicho ARNbc o genes que lo codifican.

En un primer aspecto, la invención proporciona procedimientos y medios para supervisar el silenciamiento de un gen diana en una célula eucariota midiendo el grado de silenciamiento de un segundo gen, en los que el silenciamiento del gen diana y del segundo gen se obtiene a través de la acción de una única molécula de ARNbc inicial proporcionada a la célula eucariota.

En un segundo aspecto, la invención proporciona procedimientos y medios para modular el grado de silenciamiento de un gen diana en una célula eucariota, incluyendo en la molécula única inicial de ARNbc proporcionada a la célula eucariota además del ARNbc diana que induce el silenciamiento del gen diana, una cantidad de secuencias de ARNbc no relacionadas con el ARNbc diana, en una proporción que refleja la modulación deseada del grado de silenciamiento del gen diana.

25 **Técnica antecedente**

Hasta recientemente, se conocían dos procedimientos predominantes para la modulación de la expresión génica en organismos eucariotas, que se conocen en la técnica como regulación negativa “antisentido” o regulación negativa “con sentido”.

Trabajos recientes han demostrado que la eficacia silenciadora podría mejorarse mucho tanto a nivel cuantitativo como cualitativo usando construcciones quiméricas que codifican ARN capaz de formar ARN bicatenario mediante la formación de pares de bases entre las secuencias de nucleótidos de ARN antisentido y con sentido respectivamente complementarias y homólogas a las secuencias diana. Tales ARN bicatenarios (ARNbc) también se conocen como ARN en horquilla (ARNh) o ARN interferente (ARNi).

La regulación negativa de la expresión de secuencias de nucleótidos diana usando ARNbc se ha descrito como un procedimiento de análisis conveniente para el esclarecimiento de la función de ácidos nucleicos cuyo papel se desconoce (véase, por ejemplo, el documento WO99/53050 o WO00/01846). El silenciamiento génico mediado por ARNbc también puede usarse para modular la expresión de uno o más genes diana en un organismo para obtener un organismo modificado con un fenotipo o rasgo deseado (documento WO99/53050).

Sería útil para las aplicaciones anteriormente mencionadas tener una forma de supervisar fácilmente la regulación negativa cuantitativa y cualitativa de los ácidos nucleicos diana. Esto es especialmente cierto cuando la determinación de la regulación negativa de la expresión del gen diana a través del análisis del fenotipo causado por la regulación negativa es costosa en mano de obra, coste y tiempo, requiere el uso de procedimientos de análisis especializados, etc. Cuando el silenciamiento génico mediado por ARNbc se usa para identificación/validación de la función de genes desconocidos, en los que por definición el fenotipo a buscar no se conoce, puede ser fundamental tener un procedimiento alternativo para supervisar la eficacia del silenciamiento.

Además, sería conveniente poder modular el grado de silenciamiento de un gen o genes diana particulares en un organismo, por ejemplo generando una población de organismos que exhiben un amplio espectro en los grados individuales de silenciamiento génico mediado por ARNbc e identificando los organismos con el grado de silenciamiento deseado.

Fire y col., 1998, describen interferencia genética específica por introducción experimental de ARN bicatenario en *Caenorhabditis elegans*.

El documento WO 99/32619 proporciona un procedimiento para introducir un ARN en una célula viva para inhibir la expresión génica de un gen diana en esa célula. El procedimiento puede ponerse en práctica *ex vivo* o *in vivo*. El ARN tiene una región con estructura bicatenaria. La inhibición es específica de secuencia puesto que las secuencias de nucleótidos de la región doble del ARN y/o una porción del gen diana son idénticas.

Waterhouse y col. 1998, describen que la resistencia a virus y el silenciamiento génico en plantas pueden inducirse mediante expresión simultánea de ARN con sentido y antisentido. El ARN con sentido y antisentido puede localizarse en un transcrito que tenga autocomplementariedad.

Hamilton y col. 1998, describe que un transgén con ADN repetido, es decir, copias invertidas de su región no traducida 5', causa supresión post-transcripcional, de alta frecuencia, de la expresión de ACC-oxidasa en tomate.

ES 2 346 645 T3

El documento WO 98/53083 describe construcciones y procedimientos para mejorar la inhibición de un gen diana dentro de un organismo que implican insertar en el vector de silenciamiento génico una secuencia de repetición invertida de toda o parte de una región polinucleotídica dentro del vector.

5 El documento WO 99/53050 proporciona procedimientos y medios para reducir la expresión fenotípica de un ácido nucleico de interés en células eucariotas, especialmente en células vegetales, introduciendo genes quiméricos que codifican moléculas de ARN con sentido y antisentido dirigidas hacia el ácido nucleico diana. Estas moléculas son capaces de formar una región ARN bicatenaria mediante la formación de pares de bases entre las regiones con las secuencias de nucleótidos con sentido y antisentido o mediante la introducción de las propias moléculas de ARN.
10 Preferentemente, las moléculas de ARN comprenden simultáneamente secuencias de nucleótidos tanto con sentido como antisentido.

El documento WO 99/49029 se refiere, en general, a un procedimiento para modificar la expresión génica y a genes sintéticos para modificar la expresión de genes endógenos en una célula, tejido u órgano de un organismo transgénico,
15 en particular en un animal o planta transgénica. También se proporcionan genes sintéticos y construcciones genéticas capaces de formar un ARNbc que son capaces de reprimir, retrasar o reducir de otra manera la expresión de un gen endógeno o un gen diana de un organismo cuando se introducen en él.

El documento WO 99/61631 se refiere a procedimientos para alterar la expresión de un gen diana en una planta
20 usando fragmentos de ARN con sentido y antisentido del gen. Los fragmentos de ARN con sentido y antisentido son capaces de emparejarse y formar una molécula de ARN bicatenario, alterando así la expresión del gen. La presente invención también se refiere a plantas, su progenie y semillas de las mismas obtenidas usando estos procedimientos.

El documento WO 00/01846 proporciona un procedimiento para identificar ADN responsable de conferir un fenotipo particular en una célula, comprendiendo dicho procedimiento a) construir una genoteca de ADNc o genómica del ADN de la célula en un vector adecuado en una orientación relativa a un promotor o promotores capaces de iniciar la transcripción del ADNc o ADN a ARN bicatenario (bc) tras la unión de un factor de transcripción apropiado al promotor o promotores; b) introducir la genoteca en una o más células que comprenden el factor de transcripción, y c) identificar y aislar un fenotipo particular de una célula que comprende la genoteca e identificar el fragmento de ADN
30 o ADNc de la genoteca responsable de conferir el fenotipo. Usando esta técnica, también es posible asignar función a una secuencia de ADN conocida mediante a) la identificación de homólogos de la secuencia de ADN en una célula, b) el aislamiento del homólogo o los homólogos del ADN pertinente o un fragmento del mismo de la célula, c) la clonación del homólogo o fragmento del mismo en un vector apropiado en una orientación relativa a un promotor adecuado capaz de iniciar la transcripción de ARNbc a partir de dicho homólogo de ADN o fragmento tras la unión de un factor de transcripción apropiado al promotor y d) introducir el vector en la célula de la etapa a) que comprende el factor de transcripción.
35

El documento WO 00/44914 también describe una composición y procedimientos para la atenuación *in vivo* e *in vitro* de la expresión génica usando ARN bicatenario, especialmente en pez cebra.
40

El documento WO 00/49035 describe un procedimiento para silenciar la expresión de un gen endógeno en una célula, comportando el procedimiento la sobreexpresión en la célula de una molécula de ácido nucleico del gen endógeno y una molécula antisentido que incluye una molécula de ácido nucleico complementaria a la molécula de ácido nucleico del gen endógeno, en el que la sobreexpresión de la molécula de ácido nucleico del gen endógeno y la molécula antisentido en la célula silencia la expresión del gen endógeno.
45

Smith y col., 2000, así como el documento WO 99/53050, describieron que el ARNbc que contiene intrones aumentaba más la eficacia del silenciamiento. El ARN en horquilla que contiene intrones también es conocido a menudo como ARNhi.
50

Como ilustran las referencias anteriormente mencionadas, el silenciamiento génico mediado por ARNbc es un fenómeno que ocurre en una amplia serie de organismos eucariotas, incluyendo plantas, levaduras u hongos, insectos, artrópodos y animales vertebrados, incluyendo mamíferos.

55 El documento WO 93/23551 describe un procedimiento para la inhibición de dos o más genes diana que comprende introducir en la planta un gen de control único que tiene regiones de ADN definidas homólogas a cada uno de los genes diana y un promotor operativo en plantas adaptado a transcribir a partir de tales regiones definidas ARN antisentido o con sentido que inhibe la expresión de cada uno de los genes diana.

60 El documento WO 99/49029 describe un procedimiento para fijar simultáneamente como objetivo la expresión de varios genes diana que se coexpresan en una célula específica, por ejemplo usando una molécula de ácido nucleico dispersada o molécula de ácido nucleico extraña que comprende secuencias de nucleótidos que son básicamente idénticas a cada uno de dichos genes diana coexpresados.

65 El documento WO 02/10365 proporciona un procedimiento para la supresión génica en eucariotas mediante la transformación con una construcción recombinante que contiene un promotor, al menos una secuencia de nucleótidos antisentido y/o con sentido para el gen o los genes que se van a suprimir, en el que se inhibe el transporte del núcleo al citoplasma de los productos de transcripción de la construcción. En una realización, el transporte del núcleo

ES 2 346 645 T3

5 al citoplasma se inhibe en ausencia de una UTR 3' normal. La construcción puede incluir opcionalmente al menos una ribozima que se autoescinde. La construcción puede también incluir opcionalmente secuencias con sentido y/o antisentido de múltiples genes que se van a regular negativamente de manera simultánea usando un único promotor. También se describen vectores, plantas, animales, semillas, gametos y embriones que contienen la construcción recombinante.

Elbashir y col., 2002, METHODS Vol. 26, páginas 199-213 analiza la estrategia de diseño para ARNpi y varios protocolos de silenciamiento, que incluyen inactivación génica doble en células HeLa SS6 de lamina A/C y NuMA.

10 Wesley y col. 2001 (The Plant Journal 27 (6), 581-590) describe los vectores de silenciamiento pHANNIAL y pHELLSGATE usados en la presente solicitud.

15 Sin embargo, ninguna de las referencias de la técnica anterior mencionadas anteriormente ha abordado el problema de supervisar el grado de silenciamiento de un gen diana con un fenotipo que es difícil o prácticamente imposible medir. La técnica anterior también continúa siendo defectuosa para proporcionar un procedimiento para modular o ajustar el grado de silenciamiento de un gen o genes diana específicos que usan ARNbc.

20 Éstos y otros problemas se han resuelto como se describe en lo sucesivo en el presente documento en las distintas realizaciones y reivindicaciones.

Sumario de la invención

25 La invención proporciona un procedimiento para supervisar la reducción de la expresión de un gen diana en una célula de un organismo eucariota, tal como una planta, animal, levadura, hongo o moho, que comprende las etapas de

a) proporcionar a la célula de dicho organismo eucariota un ARNbc que comprende una primera región, una segunda región, una tercera región y una cuarta región en el que

30 i) la primera región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tienen al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de dicho gen diana;

35 ii) la segunda región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tienen al menos un 94% de identidad de secuencia con una secuencia de nucleótidos complementaria a 19 nucleótidos consecutivos de dicha región de nucleótidos con sentido del gen diana;

iii) la primera región y la segunda región son capaces de formar una región de ARN bicatenario;

40 iv) la tercera región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de un segundo gen, tal como un gen endógeno o un transgén, integrado de forma estable en el genoma de la célula eucariota y que es distinto del gen diana;

45 v) la cuarta región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tienen al menos un 94% de identidad de secuencia con el complemento de los 19 nucleótidos consecutivos de la región de nucleótidos con sentido del segundo gen; y

vi) la tercera y cuarta región son capaces de formar una región de ARN bicatenario; y

50 b) supervisar la reducción de la expresión del gen diana analizando la reducción en la expresión de dicho segundo gen,

55 en el que el organismo eucariota es una planta, una levadura, un hongo, un moho, o un animal a condición de que cuando dicho organismo es un animal, la célula no sea una célula en un cuerpo humano o animal.

El ARNbc puede transcribirse a partir de un gen quimérico comprendido en células del organismo eucariota, comprendiendo el gen quimérico:

60 c) una región promotora que funciona en la célula eucariota; unida operativamente a

d) una región de ADN que cuando se transcribe produce la molécula de ARNbc; y

65 e) una región de terminación de transcripción y poliadenilación que funciona en las células del organismo eucariota.

ES 2 346 645 T3

También es un objeto de la invención proporcionar una molécula de ARN como se ha descrito así como el uso de dicha molécula de ARN para medir la reducción de la expresión de un gen diana.

5 Otro objeto de la invención es proporcionar una molécula de ADN para medir la reducción de la expresión de un gen diana en una célula eucariota, que comprende

- a) una región promotora que funciona en la célula eucariota; unida operativamente a
- 10 b) una región de ADN que cuando se transcribe produce una molécula de ARNbc, en la que el ARNbc comprende una primera, segunda, tercera y cuarta región;
 - 15 i) comprendiendo la primera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido del gen diana;
 - 20 ii) comprendiendo la segunda región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con una secuencia de nucleótidos complementaria a 19 nucleótidos consecutivos de la región de nucleótidos con sentido del gen diana;
 - 25 iii) siendo capaces la primera región y dicha segunda región de formar una región de ARN bicatenaria;
 - iv) comprendiendo la tercera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de un segundo gen presente en la célula eucariota y que es distinto del gen diana;
 - 30 v) comprendiendo la cuarta región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con el complemento de los 19 nucleótidos consecutivos de la región de nucleótidos con sentido del segundo gen;
 - 35 vi) siendo capaces la tercera y cuarta región de formar una región de ARN bicatenaria; en la que las regiones de ARN bicatenarias formadas entre la primera y la segunda región y la región de ARN bicatenaria formada entre la tercera y cuarta región son aproximadamente iguales en tamaño; y
- c) una región de terminación de transcripción y poliadenilación que funciona en células del organismo eucariota, en la que dicho segundo gen es un gen endógeno del organismo eucariota o un transgén integrado de forma estable en el genoma de células del organismo eucariota, así como el uso de dicha molécula de ADN para medir la expresión de un gen diana mediante la medición de la reducción en expresión de un segundo gen.

40 La invención proporciona también organismos eucariotas no humanos que comprenden una molécula de ARN o una molécula de ADN como se describe en el presente documento.

La invención también proporciona un procedimiento para identificar, dentro de una población de organismos eucariotas con silenciamiento génico mediado por ARNbc, aquellos organismos con el grado deseado de silenciamiento de un gen diana, que comprende:

- a) proporcionar a las células eucariotas de los organismos eucariotas un ARNbc que comprende una primera región, una segunda región, una tercera región y una cuarta región
 - 50 i) comprendiendo la primera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido del gen diana;
 - 55 ii) comprendiendo la segunda región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con una secuencia de nucleótidos complementaria a 19 nucleótidos consecutivos de la región de nucleótidos con sentido del gen diana;
 - 60 iii) siendo capaces la primera región y la segunda región de formar una región de ARN bicatenaria;
 - iv) comprendiendo la tercera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de un segundo gen presente en las células eucariotas y que es distinto del gen diana;
 - 65 v) comprendiendo la cuarta región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con el complemento de los 19 nucleótidos consecutivos de la región de nucleótidos con sentido del segundo gen;

ES 2 346 645 T3

vi) siendo capaces la tercera y cuarta región de formar una región de ARN bicatenaria; y

- b) identificar el organismo con el grado deseado de silenciamiento del gen diana, seleccionando aquellos organismos con el grado deseado de silenciamiento del segundo gen, en el que el organismo eucariota es una planta, una levadura, un hongo, un mohó, o un animal, a condición de que cuando dicho organismo es un animal, la célula no sea una célula en un cuerpo humano o animal.

Breve descripción de las figuras

La figura 1 muestra semillas de plantas T1 transformadas con la construcción FLC/CHS, que emite fluorescencia con luz UV, comparadas con semillas de tipo silvestre C24 (C24), un tipo silvestre C24 que contiene un gen GUS quimérico bajo control de un promotor CaMV35S (GUS) y una línea silenciada CHS homocigota (CHS). Los números indican los días hasta el florecimiento de las líneas celulares de plantas transgénicas.

La figura 2 muestra un análisis de transferencia de Northern de ARN preparado a partir de líneas vegetales transgénicas, que comprende genes que codifican ARNbc que tienen como objetivo tanto FLC como CHS, expuesto a una sonda FLC (panel superior) y una sonda CHS (panel inferior) (Fig. 2A). La Fig. 2B es una representación gráfica de la cantidad de ARNm detectada a partir de la expresión de FLC trazada como función de la cantidad de ARNm detectada de la expresión de CHS.

Descripción de realizaciones detalladas

La invención se basa en la observación inesperada de que cuando una célula eucariota, tal como una célula vegetal, comprende un ARN bicatenario (ARNbc) en el que el ARNbc comprende simultáneamente regiones complementarias antisentido y con sentido para al menos dos genes diana, existe una correspondencia entre el grado de silenciamiento de expresión de todos los genes fijados como objetivo. La correspondencia además depende del tamaño relativo de las regiones antisentido y con sentido diseñadas para reducir la expresión de los diferentes genes diana.

Esta correspondencia puede usarse oportunamente para supervisar el silenciamiento de genes cuya expresión da lugar a un fenotipo que no es fácil de supervisar. Esto puede conseguirse uniendo las regiones complementarias con sentido y antisentido adecuadas para reducir o silenciar la expresión de dicho gen, a regiones complementarias con sentido y antisentido adecuadas para reducir o silenciar la expresión de un gen cuya expresión da como resultado un fenotipo que puede supervisarse de una manera fácil.

Así pues, en una realización de la invención, se proporciona un procedimiento para supervisar la reducción o silenciamiento de la expresión de un gen diana en una célula eucariota, que comprende las etapas de

- Proporcionar a la célula eucariota un ARNbc que comprende una primera región, una segunda región, una tercera región y una cuarta región, en las que
- la primera región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tienen al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de la región de nucleótidos con sentido del gen diana;
- la segunda región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tienen al menos un 94% de identidad de secuencia con una secuencia de nucleótidos complementaria a 19 nucleótidos consecutivos de la región de nucleótidos con sentido del gen diana;
- la primera región y la segunda región son capaces de formar una región de ARN bicatenario, preferentemente sobre la extensión completa de la primera y segunda región, y al menos sobre la extensión de los 19 nucleótidos mencionados;
- la tercera región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de la región de nucleótidos con sentido de un segundo gen presente en la célula eucariota y que es diferente del gen diana;
- la cuarta región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tienen al menos un 94% de identidad de secuencia con el complemento de 19 nucleótidos consecutivos de la región de nucleótidos con sentido del mismo segundo gen;
- la tercera y cuarta región son capaces de formar una región de ARN bicatenario, preferentemente sobre la extensión completa de la primera y segunda región, y al menos sobre la extensión de los 19 nucleótidos mencionados;
- y
- monitorizar la expresión del grado de silenciamiento del gen diana, analizando la reducción en expresión del segundo gen.

ES 2 346 645 T3

Preferentemente, la primera y segunda región son capaces de formar una región de ARN bicatenario al mismo tiempo que cuando la tercera y cuarta región están formando un ARN bicatenario.

5 El segundo gen diferente del gen diana se usa como gen informador para supervisar el grado de silenciamiento del gen diana y siempre que se hace referencia en el presente documento a un gen informador, se entiende que esta expresión se usa para referirse a un segundo gen presente en la célula eucariota y que es diferente del gen diana.

10 Para los propósitos de la presente invención, es igual que el segundo gen sea un gen endógeno, presente normalmente en la célula eucariota, o un transgén, que se ha introducido en la célula eucariota por intervención humana en algún momento de su historia, preferentemente integrado en el genoma de la célula eucariota de manera estable.

Preferentemente, el gen informador debería tener un fenotipo cuyo análisis es más fácil que el análisis del gen o genes diana, en términos de requisitos de costes, especialización, tiempo, mano de obra, aparatos usados, etc.

15 También preferentemente, el silenciamiento del gen informador no debería tener una influencia negativa en la viabilidad de la célula huésped o el organismo huésped, aunque la ausencia de influencia negativa puede depender de condiciones particulares.

20 Oportunamente, el silenciamiento de la expresión del gen informador puede dar como resultado un fenotipo visible, aunque la visibilidad del fenotipo puede, de nuevo, ser condicional.

25 Los ejemplos de genes informadores adecuados para la aplicación de los procedimientos de la invención en células vegetales y plantas incluyen, pero no se limitan a gen de la Chalcona sintasa (CHS, en el que la regulación negativa de la expresión da como resultado una acumulación de compuestos fluorescentes UV en el color de la cubierta de la semilla); el gen de la fitoeno desaturasa (PDS, en el que la regulación negativa de la expresión del gen da como resultado fotoblanqueo), locus de flor C (FLC, en el que la regulación negativa de la expresión del gen da como resultado florecimiento temprano), gen de insensibilidad a etileno 2 (EIN2, en el que la regulación negativa del gen da como resultado plantas que son insensibles al etileno, y crecerán en medios que contienen ácido 1-aminociclopropano-1-carboxílico (AAC)), genes marcadores visuales tales como los genes del color de la cubierta de las semillas (por ejemplo gen R en maíz), los genes GUS o GFP expresables en plantas, fitocromo B y similares.

30 Los ejemplos de genes informadores adecuados para la aplicación de los procedimientos de la invención en células animales y animales incluyen, pero no se limitan a genes GUS o GFP unidos operativamente a regiones de expresión adecuadas para células animales pero también genes como *unc* en *Caenorhabditis elegans*, cuyo silenciamiento causa un patrón de retorcimiento característico en nemátodos y similares.

35 Los ejemplos de los genes informadores adecuados para la aplicación de los procedimientos de la invención en células fúngicas incluyen pero no se limitan a genes GUS o GFP unidos operativamente a regiones de expresión adecuadas para células fúngicas, pero también genes cuyo silenciamiento causa crecimiento auxotrófico (tales como, por ejemplo *trpC*), un fenotipo que puede explorarse fácilmente en medios mínimos. No hace falta decir que en estos casos, una copia maestra de la genoteca de las células fúngicas silenciadas debe mantenerse en condiciones que permitan el crecimiento, por ejemplo en presencia del compuesto nutriente requerido.

40 Preferentemente, el gen informador debe expresarse al menos en condiciones en las que se expresa el gen diana.

45 Los 19 nucleótidos que son al menos un 94% idénticos o complementarios a la secuencia de nucleótidos de 19 nucleótidos consecutivos del gen diana y/o el gen informador pueden estar comprendidos dentro de una molécula de ARN mayor. Dicha molécula de ARN tendrá una secuencia de nucleótidos que varía entre tan poco como 19 pb y una longitud igual al tamaño del gen diana de grado de variación global de identidad de secuencia.

50 Para el objeto de la presente invención, la "identidad de secuencia" de dos secuencias de nucleótidos o aminoácidos relacionadas, expresada como porcentaje, se refiere al número de posiciones en las dos secuencias alineadas óptimamente que tienen restos idénticos (x100) dividido por el número de posiciones comparadas. Un hueco, es decir, una posición en un alineamiento en el que un resto está presente en una secuencia pero no en la otra se considera una posición con restos no idénticos. El alineamiento de las dos secuencias se realiza por el algoritmo de Needleman y Wunsch (Needleman y Wunsch 1970). El alineamiento de secuencias asistido por ordenador anterior, puede realizarse oportunamente usando un programa de software convencional tal como el GAP, que es parte del Wisconsin Package Versión 10.1 (Genetics Computer Group, Madison, Wisconsin, Estados Unidos) usando la matriz de puntuación por defecto con una penalización de creación de huecos de 50 y una penalización de extensión de huecos de 3. Las secuencias se indican como "esencialmente similares" cuando dicha secuencia tiene una identidad de secuencia de al menos aproximadamente un 75%, particularmente al menos aproximadamente un 80%, más particularmente al menos aproximadamente un 85%, bastante particularmente aproximadamente un 90%, especialmente aproximadamente un 95%, más especialmente aproximadamente un 100%, y bastante especialmente son idénticas. Está claro que cuando las secuencias de ARN van a ser esencialmente similares o tienen un cierto grado de identidad de secuencia con secuencias de ADN, la timina (T) en la secuencia de ADN se considera igual al uracilo (U) en la secuencia de ARN. Así pues, cuando se indica en esta solicitud que una secuencia de 19 nucleótidos consecutivos tiene una identidad de secuencia del 94% con una secuencia de 19 nucleótidos, esto quiere decir que al menos 18 de los 19 nucleótidos de la primera secuencia son idénticos a 18 de los 19 nucleótidos de la segunda secuencia.

ES 2 346 645 T3

De esta manera, las regiones de nucleótidos con sentido o antisentido mencionadas pueden ser de aproximadamente 50 nt, 100 nt, 200 nt, 300 nt, 500 nt, 1000 nt, 2000 nt, o incluso aproximadamente 5000 nt de longitud o mayores, teniendo cada una una identidad de secuencia global respectivamente de aproximadamente un 40%, 50%, 60%, 70%, 80%, 90% o 100%. Cuanto más larga es la secuencia menos riguroso será el requisito para la identidad de secuencia global.

La primera, segunda, tercera y cuarta regiones pueden separarse, cada una, mediante una región espaciadora que tiene una secuencia de nucleótidos que no está relacionada con la secuencia de nucleótidos del gen diana o el gen informador.

Para el objeto de la invención, aunque las regiones se denominan consecutivamente, el orden de las regiones de ARNbc en la molécula de ARNbc no es importante. En otras palabras, no importa si, por ejemplo, la primera o segunda región o, como alternativa, la tercera o cuarta región están localizadas en el extremo 5' o 3' de la molécula de ARN.

Como se usa en el presente documento, debe interpretarse que “que comprende” especifica la presencia de las características, números enteros, etapas o componentes a los que se hace referencia, pero no excluye la presencia o adición de una o más características, números enteros, etapas o componentes, o grupos de los mismos. Así pues, por ejemplo un ácido nucleico o proteína que comprende una secuencia de nucleótidos o aminoácidos, puede comprender más nucleótidos o aminoácidos que los que de hecho se han citado, es decir, puede estar incluido en un ácido nucleico o proteína mayor. Un gen quimérico que comprende una región de ADN que está definida funcional o estructuralmente, puede comprender regiones adicionales de ADN.

Por lo tanto, una molécula de ARNbc que comprende una primera, segunda, tercera y cuarta región, como se define en el presente documento, puede incluir adicionalmente, por ejemplo, una quinta y sexta región que tengan una región de nucleótidos de 19 nucleótidos con al menos un 94% de identidad o complementariedad de secuencia con, por ejemplo, un gen diana, que puede ser el mismo gen diana que el primer gen diana o puede ser uno diferente.

En una realización de la invención, la molécula de ARNbc puede además comprender una o más regiones que tienen al menos un 94% de identidad de secuencia con regiones de 19 nucleótidos consecutivos de los nucleótidos con sentido de los genes diana, diferentes de los 19 nucleótidos consecutivos como se definen en la primera región, y una o más regiones que tienen al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos del complemento de los nucleótidos con sentido del gen diana, diferentes de los 19 nucleótidos consecutivos como se definen en la segunda región, en los que estas regiones adicionales pueden formar pares de bases entre ellos. De forma similar, el ARNbc puede también comprender adicionalmente una o más regiones que tienen al menos un 94% de identidad de secuencia con regiones de 19 nucleótidos consecutivos de un gen informador diferentes de los 19 nucleótidos de la tercera región y una o más regiones que tienen al menos un 94% de identidad de secuencia con las regiones complementarias de 19 nucleótidos consecutivos de los genes informadores, en las que estas regiones adicionales son capaces de formar pares de bases entre ellas. De nuevo, no se requiere un orden particular de las regiones, y estas regiones pueden dispersarse entre ellas. Así pues, por ejemplo, las regiones de ARNbc que se dirigen a silenciar el gen diana pueden alternarse con regiones de ARNbc dirigidas a silenciar el gen informador siempre que, por supuesto, aún sea posible la formación de pares de bases entre regiones de ARN complementarias.

Oportunamente, el ARNbc como se ha descrito puede introducirse en la célula huésped mediante introducción y posible integración de un gen quimérico, cuya transcripción produce dicho ARNbc. Así pues, la invención también se dirige a proporcionar dicho gen quimérico que comprende

- un promotor o una región de promotor que es capaz de expresarse en células del organismo eucariota de interés; unida operativamente a una región de ADN que cuando se transcribe produce una molécula de ARNbc que comprende una primera región, una segunda región, una tercera región y una cuarta región, en la que
- la primera región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tienen al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido del gen diana;
- la segunda región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con el complemento de 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido del gen diana;
- la primera región y segunda región son capaces de formar una región de ARN bicatenario, preferentemente sobre toda la extensión de la primera y segunda región, y al menos entre los 19 nucleótidos mencionados de la primera y segunda región;
- la tercera región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tienen al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido de un segundo gen presente en la célula eucariota y que es diferente del gen diana;

ES 2 346 645 T3

- la cuarta región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tienen al menos un 94% de identidad de secuencia con el complemento de 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido del mismo gen;

5 - la tercera y cuarta región son capaces de formar una región de ARN bicatenario, preferentemente sobre toda la extensión de la primera y segunda región, y al menos sobre la extensión de los 19 nucleótidos mencionados;

y en la que las regiones de ARN bicatenario formadas entre la primera y la segunda región y la región de ARN bicatenario formada entre la tercera y cuarta región son aproximadamente iguales en tamaño; y

10

- una región de terminación de la transcripción y de poliadenilación adecuada para la célula eucariota elegida.

Como se usa en el presente documento, el término “promotor” indica cualquier ADN que reconoce y al que se une (directa o indirectamente) una ARN polimerasa dependiente de ADN durante la iniciación de la transcripción. Un promotor incluye el sitio de iniciación de la transcripción y los sitios de unión para factores de iniciación de la transcripción y ARN polimerasa, y puede comprender otros diversos sitios (por ejemplo, potenciadores), a los que pueden unirse las proteínas reguladoras de la expresión génica.

15

La expresión “región reguladora”, como se usa en el presente documento, se refiere a cualquier ADN que esté involucrado en dirigir la transcripción y supervisar (es decir, regular) el ritmo y nivel de transcripción de una secuencia de ADN dada, tal como un ADN que codifica una proteína o polipéptido. Por ejemplo, una región reguladora 5’ (o “región promotora”) es una secuencia de ADN situada aguas arriba (es decir, 5’) de una secuencia codificante y que comprende el promotor y la secuencia líder no traducida 5’. Una región reguladora 3’ es una secuencia de ADN localizada aguas abajo (es decir, 3’) de la secuencia codificante y que comprende señales adecuadas de terminación (y/o regulación) de la transcripción, incluyendo una o más señales de poliadenilación.

20

25

En una realización de la invención, el promotor es un promotor constitutivo. En otra realización de la invención, la actividad promotora se potencia por estímulos externos o internos (promotor inducible), tales como pero no limitados a hormonas, compuestos químicos, impulsos mecánicos, condiciones de estrés abiótico o biótico. La actividad del promotor también puede regularse de una forma temporal o espacial (promotores específicos de tejido; promotores regulados por el desarrollo).

30

En una realización particular de la invención, el promotor es un promotor expresable en plantas. Como se usa en el presente documento, la expresión “promotor expresable en plantas” se refiere a una secuencia de ADN que es capaz de supervisar (iniciar) la transcripción en una célula vegetal.

35

Esto incluye cualquier promotor de origen vegetal, pero también cualquier promotor de origen no vegetal que es capaz de dirigir la transcripción en una célula vegetal, es decir, ciertos promotores de origen viral o bacteriano tales como el CaMV35S (Hapster y col., 1988), el promotor del virus del trébol subterráneo N° 4 o N° 7 (documento WO9606932), promotores del gen ADN-T pero también promotores específicos de tejido o específicos de órgano que incluyen, pero no se limitan a promotores específicos de semilla (por ejemplo, documento WO89/03887), promotores específicos de primordios de órganos (An y col., 1996), promotores específicos de tallo (Keller y col., 1988), promotores específicos de hoja (Hudspeth y col., 1989), promotores específicos de mesófilo (tales como los promotores de la Rubisco inducibles por luz), promotores específicos de raíz (Keller y col., 1989), promotores específicos de tubérculo (Keil y col., 1989), promotores específicos de tejido vascular (Peleman y col., 1989), promotores selectivos de estambres (documentos WO 89/10396, WO 92/13956), promotores específicos de zona de dehiscencia (documento WO 97/13865) y similares.

40

45

En otra realización particular de la invención, el promotor es un promotor expresable en hongos. Como se usa en el presente documento, la expresión “promotor expresable en hongos” se refiere a una secuencia de ADN que es capaz de supervisar (iniciar) la transcripción en una célula fúngica, tal como, pero sin limitarse al promotor del gen *trpC* de *A. nidulans*, o el promotor de GAL4 de *S. cerevisiae* que es capaz de supervisar (iniciar) la transcripción en una célula fúngica.

55

En otra realización particular más de la invención, el promotor es un promotor expresable en animales. Como se usa en el presente documento, la expresión “promotor expresable en animales” se refiere a una secuencia de ADN que es capaz de supervisar (iniciar) la transcripción en una célula animal e incluye pero no se limita a promotores tardíos y tempranos de SV40, promotores CMV-IE de citomegalovirus, promotor RSV-LTR, promotor SCSV, promotor SCBV y similares.

60

Las moléculas de ARNbc útiles para la invención pueden producirse también por transcripción *in vitro*. Para este objetivo, el promotor de los genes quiméricos de acuerdo con la invención puede ser un promotor reconocido por una ARN polimerasa de subunidad sencilla de bacteriófago, tal como los promotores reconocidos por la ARN polimerasa de subunidad sencilla de bacteriófago tales como las ARN polimerasas derivadas de los fagos de *E. coli* T7, T3, ϕ I, ϕ II, W31, H, Y, A1, 122, cro, C21, C22 y C2; fago gh-1 de *Pseudomonas putida*; fago SP6 de *Salmonella typhimurium*; fago IV de *Serratia marcescens*; fago VIII de *Citrobacter* y fago N° 11 de *Klebsiella* [Hausmann, Current Topics in Microbiology and Immunology, 75: 77-109 (1976); Korsten y col., J. Gen Virol. 43: 57-73 (1975); Dunn y col., Nature

65

ES 2 346 645 T3

New Biology, 230: 94-96 (1971); Towle y col., J. Biol. Chem. 250: 1723-1733 (1975); Butler y Chamberlin, J. Biol. Chem., 257: 5772-5778 (1982)]. Son ejemplos de tales promotores un promotor específico de la ARN polimerasa de T3 y un promotor específico de la ARN polimerasa de T7 respectivamente. Un promotor de T3 que va usarse como un primer promotor en el CIG puede ser cualquier promotor de los genes de T3 que se describen en McGraw y col., Nucl. Acid Res. 13: 6753-6766 (1985). Como alternativa, un promotor de T3 puede ser un promotor de T7 que está modificado en las posiciones de nucleótidos -10, -11 y -12 para que se reconozcan por la ARN polimerasa de T3 [(Klement y col., J. Mol. Biol. 215, 21-29 (1990)]. Un promotor de T3 preferido es el promotor que tiene la secuencia “consenso” de un promotor de T3, como se describe en la Patente de Estados Unidos 5.037.745. Un promotor de T7 que puede usarse de acuerdo con la invención en combinación con la ARN polimerasa de T7 comprende un promotor de uno de los genes de T7 como se describe en Dunn y Studier, J. Mol. Biol. 166: 477-535 (1983). Un promotor de T7 preferido es el promotor que tiene la secuencia “consenso” para un promotor de T7, como se describe en Dunn y Studier (anteriormente).

El ARNbc puede producirse en grandes cantidades poniendo en contacto el ADN del vector aceptor con la ARN polimerasa de subunidad sencilla de bacteriófago apropiada en condiciones bien conocidas para el experto en la materia. El ARNbc así producido puede usarse después para introducirse en células propensas al silenciamiento génico, tales como células vegetales, células fúngicas o células animales. El ARNbc puede introducirse en células animales mediante liposomas u otros agentes de transfección (por ejemplo, reactivo de transfección Clonfection o el kit de transfección de mamíferos CalPhos de ClonTech) y puede usarse para procedimientos de tratamiento de animales, incluyendo humanos, mediante el silenciamiento de los genes diana apropiados. El ARNbc puede introducirse en la célula de varias formas diversas. Por ejemplo, el ARNbc puede administrarse por microinyección, bombardeo con partículas cubiertas por el ARNbc, inmersión de la célula o los organismos en una solución del ARNbc, electroporación de las membranas celulares en presencia del ARNbc, liberación mediada por liposomas del ARNbc y transfección mediada por compuestos químicos tales como el fosfato de calcio, infección viral, transformación y similares. El ARNbc puede introducirse junto con componentes que aumentan la captación de ARN por la célula, estabilizan las cadenas híbridadas, o aumentan de otra manera la inhibición del gen diana. En el caso de un animal completo, el ARNbc se introduce oportunamente por inyección o perfusión en una cavidad o espacio intersticial de un organismo, o sistémicamente por administración oral, tópica, parenteral (incluyendo administración subcutánea, intramuscular o intravenosa), vaginal, rectal, intranasal, oftálmica o intraperitoneal. El ARNbc puede administrarse también a través de un dispositivo de liberación prolongada implantable.

Los genes quiméricos de acuerdo con la invención capaces de producir un ARNbc también pueden equiparse con cualquier promotor procarionota adecuado para la expresión de ARNbc en un huésped procarionota específico. El huésped procarionota puede usarse como fuente de ARNbc, por ejemplo, alimentando con él a un animal, tal como un nematodo o un insecto, en el que se prevé el silenciamiento del gen diana y se controla por reducción de la expresión del gen informador. En este caso, estará claro que el gen diana y los genes informadores deberían ser genes presentes en las células del organismo eucariota diana y no del organismo huésped procarionota. El ARNbc de acuerdo con la invención o los genes quiméricos capaces de producir dichas moléculas de ARNbc pueden producirse así en un organismo huésped. El ARNbc puede administrarse a otro organismo diana (por ejemplo a través de la alimentación, administración oral, como una molécula de ADN o ARN desnuda o encapsulada en un liposoma, en una partícula de virus o una partícula de virus atenuado, o en una partícula inerte, etc.) y efectuar una reducción de la expresión del gen en el gen o genes diana y gen o genes informadores en otro organismo. En este caso, estará claro que el gen diana y los genes informadores deberían ser genes presentes en las células del organismo (eucariota) diana y no del organismo huésped.

Las regiones adecuadas de terminación de transcripción y poliadenilación incluyen, pero no se limitan a la señal de poliadenilación de SV40, la señal de poliadenilación de HSV TK, el terminador del gen de la nopalina sintasa de *Agrobacterium tumefaciens*, el terminador del transcrito de CaMV 35S, terminadores del virus del enanismo del trébol subterráneo, el terminador del gen *trpC* de *Aspergillus nidulans* y similares.

La invención también pretende proporcionar las moléculas de ARNbc que pueden obtenerse por transcripción de estos genes quiméricos y que son útiles para los procedimientos de acuerdo con la invención.

Es otro objeto de la invención proporcionar células eucariotas y organismos no humanos eucariotas que contienen las moléculas de ARNbc de la invención o contienen los genes quiméricos capaces de producir las moléculas ARNbc de la invención. En una realización preferida, los genes quiméricos están integrados de forma estable en el genoma de las células del organismo eucariota.

En otra realización preferida, los genes quiméricos pueden proporcionarse en una molécula de ADN capaz de replicarse de forma autónoma en las células del organismo eucariota, tal como, por ejemplo, vectores virales. El gen quimérico o el ARNbc también pueden proporcionarse transitoriamente a las células del organismo eucariota.

Las moléculas de ARNbc de acuerdo con este primer aspecto de la invención también pueden usarse para identificar, dentro de una población de organismos con silenciamiento génico mediado por ARNbc, los que tienen el grado deseado de silenciamiento de un gen o genes diana. Para este objetivo se genera una población de organismos que contienen ARNbc, en la que el ARNbc comprende una primera región, una segunda región, una tercera región y una cuarta región, comprendiendo la primera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido del gen diana. La segunda región comprende una secuencia de nucleótidos de al menos 19 nucleótidos

ES 2 346 645 T3

consecutivos que tiene al menos un 94% de identidad de secuencia con el complemento de 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido del gen diana. La tercera región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido de un segundo gen presente en la célula eucariota y que es diferente del gen diana. La cuarta región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con el complemento de 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido del mismo segundo gen. La primera y segunda y la tercera y cuarta región son capaces de formar una región de ARN bicatenario, preferentemente sobre toda la extensión de la primera y segunda o tercera y cuarta regiones, y al menos entre los tramos mencionados de 19 nucleótidos, y en los que las regiones de ARN bicatenario formadas entre la primera y segunda región y la región de ARN bicatenario formada entre la tercera y cuarta región son aproximadamente iguales en tamaño. La población de organismos que contienen ARNbc se analiza después para la regulación negativa de la expresión del gen informador y se seleccionan y aíslan los organismos que contienen ARNbc en los que la regulación negativa de la expresión del gen informador corresponde al grado deseado de regulación negativa.

La correspondencia en el grado de silenciamiento de dos o más genes inducido por la presencia de ARNbc que comprende las regiones de ARN con sentido y antisentido complementarias bicatenarias para los dos o más genes diana, también puede usarse para modular el grado de silenciamiento de esos genes diana. La correspondencia en el grado de silenciamiento de dos o más genes depende del tamaño relativo de las regiones de ARN con sentido y antisentido complementarias bicatenarias. Por consiguiente, el grado de silenciamiento de un gen particular puede modularse mediante la inclusión en la misma molécula de ARNbc que comprende regiones de ARN complementarias con sentido y antisentido, diseñada para silenciar la expresión de un gen o genes diana particulares, un exceso de secuencias complementarias no relacionadas, también capaces de formar un ARN bicatenario mediante formación de pares de bases.

Sin restringir la invención a un particular modo de actuar, se piensa que la enzima de células eucariotas que es responsable de generar moléculas de ARN pequeñas de aproximadamente 21 nt a partir del ARNbc (tal como DICER en *Drosophila*) puede saturarse mediante la inclusión en el ARNbc de un exceso de secuencias de ARNbc (es decir, moléculas de ARN complementarias) cuya secuencia no está relacionada con la secuencia de nucleótidos del gen o genes diana que van a silenciarse.

Así pues, en una realización de la invención, se proporciona un procedimiento para modular la reducción de la expresión de un gen diana en un organismo eucariota, que comprende las etapas de

- proporcionar a la célula eucariota un ARNbc que comprende una primera región, una segunda región, una tercera región y una cuarta región, en el que
- la primera región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido del gen diana;
- la segunda región comprende una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con el complemento de 19 nucleótidos consecutivos de la secuencia de nucleótidos con sentido del gen diana;
- la primera región y la segunda región son capaces de formar una región de ARN bicatenario, preferentemente sobre toda la extensión de la primera y segunda región, y al menos entre los tramos mencionados de 19 nucleótidos;
- la tercera y cuarta región comprenden secuencias de nucleótidos complementarias no relacionadas con la secuencia de nucleótidos del gen diana y son capaces de formar un ARN bicatenario. El tamaño del ARN bicatenario que puede formarse por formación de pares de bases entre la tercera y cuarta región, es preferentemente igual o mayor que el tamaño del ARN bicatenario que puede formarse por formación de pares de bases entre la primera y la segunda región. El tamaño de la tercera/cuarta región puede también ser más pequeño que el tamaño de la primera/segunda región, dependiendo del grado deseado de modulación del silenciamiento.

Preferentemente, la primera y segunda región son capaces de formar una región de ARN bicatenario a la vez que la tercera y cuarta región están formando un ARN bicatenario. Además, preferentemente, el gen diana es un gen endógeno de la célula eucariota o un transgén integrado de manera estable en el genoma de las células eucariotas.

Como se usa en el presente documento, se dice que una secuencia está “no relacionada” con un gen diana, cuando la identidad de secuencia global entre la secuencia no relacionada y la secuencia diana es de menos del 50%, particularmente menos del 45%, más particularmente menos del 35%. La secuencia no relacionada preferentemente no debe tener una secuencia de nucleótidos que tenga al menos un 94% de identidad de secuencia con un tramo de 19 nucleótidos consecutivos del gen diana o del complemento del mismo.

ES 2 346 645 T3

5 La variación natural en regulación negativa de la expresión de un gen diana que ocurre entre líneas diferentes de un organismo eucariota que comprende la misma molécula de ARNbc se desplazará hacia el extremo inferior del espectro del silenciamiento génico. Esto se debe a la inclusión de secuencias de nucleótidos de ARNbc extra que no están relacionadas con el gen diana, que se unen operativamente al ARNbc formado por la primera y segunda región de ARN.

10 La relación de tamaño del ARN bicatenario que puede formarse por formación de pares de bases entre la tercera y la cuarta región y el tamaño del ARN bicatenario que puede formarse por formación de pares de bases entre la primera y la segunda región puede variar de 0,1 a 20, preferentemente de 1 a 10. Cuanto mayor sea la relación mencionada más organismos de una población de organismos eucariotas que comprenden ese ARNbc tendrán un menor grado de silenciamiento génico (es decir, un mayor nivel de expresión del gen diana).

15 Los 19 nucleótidos de la primera o segunda regiones de ARN que son idénticas o complementarias en al menos un 94% a una secuencia de nucleótidos de 19 nucleótidos consecutivos del gen diana pueden estar comprendidos dentro de una molécula de ARN mayor. Tal molécula de ARN más larga tendrá una secuencia de nucleótidos que varía entre tan poco como 19 pb y una longitud igual a la del tamaño del gen diana con un grado global variable de identidad de secuencia.

20 Las mencionadas regiones de nucleótidos con sentido o antisentido pueden así ser de aproximadamente 50 nt, 100 nt, 200 nt, 300 nt, 500 nt, 1000 nt, 2000 nt o incluso aproximadamente 5000 nt o mayores en longitud, teniendo cada una una identidad de secuencia global de respectivamente aproximadamente un 40%, 50%, 60%, 70%, 80%, 90% o 100%. Cuanto más larga es la secuencia, menos riguroso es el requisito de identidad de secuencia global.

25 La primera, segunda, tercera y cuarta regiones pueden separarse, cada una, por una región espaciadora que tiene una secuencia de nucleótidos que no está relacionada con la secuencia de nucleótidos del gen diana o informador.

30 Oportunamente, las moléculas de ARNbc pueden introducirse en las células u organismos eucariotas por transcripción a partir de un gen quimérico que comprende un promotor o una región promotora capaz de transcribir una región de ADN, que cuando se transcribe produce el ARNbc de acuerdo con este aspecto de la invención. Se entiende que diferentes realizaciones que conciernen promotores adecuados, etc., mencionadas en relación con el primer aspecto de la invención también pueden aplicarse aquí.

35 La invención también pretende proporcionar ARNbc o genes quiméricos capaces de producir tales moléculas de ARNbc de acuerdo con el segundo aspecto de la invención, así como organismos eucariotas no humanos que comprenden tal ARNbc o genes quiméricos.

40 La parte de repeticiones invertidas de los genes quiméricos de la invención puede construirse convencionalmente por clonación recombinatoria, usando los medios y procedimientos descritos en las solicitudes de Patente de Estados Unidos 60/244067 y 60/333743 o el documento WO02/059294. Los vectores de acuerdo con las solicitudes de Patente de Estados Unidos 60/244067 o 60/333743 pueden modificarse para incluir una tercera y cuarta regiones capaces de formar un gen de ARN bicatenario, en el que la tercera y cuarta regiones son capaces de regular negativamente la expresión de un gen marcador tal como EIN2, PHYB, FLC y CHS.

45 Los genes quiméricos que codifican ARNbc de acuerdo con la invención comprenden preferentemente un intrón, preferentemente en la región entre la segunda y tercera región, para aumentar la eficacia (como se describe en el documento WO99/53050). Como se usa en el presente documento, un "intrón" o secuencia intermedia se usa para referirse a una región de ADN dentro de una región transcrita mayor, que se transcribe en el núcleo para producir una región de ARN que es parte de un ARN mayor, sin embargo, dicho ARN que corresponde a la secuencia intrón se retira del ARN mayor cuando se transfiere al citoplasma. El ARN correspondiente también se conoce como un intrón o secuencia intermedia. Las secuencias intrón están flanqueadas por sitios de corte, y pueden hacerse intrones sintéticos uniendo sitios de corte apropiados a cualquier secuencia que tenga un punto de ramificación apropiado. Los ejemplos de intrones incluyen el intrón *pdk2*, el intrón catalasa de ricino, el intrón Delta 12 desaturasa de algodón, el intrón Delta 12 desaturasa de *Arabidopsis*, el intrón de ubiquitina de maíz o el intrón de SV40.

55 En una realización de la invención, combinando el primer y segundo aspecto de la invención, una molécula de ARNbc puede proporcionarse a una célula eucariota para supervisar la regulación negativa de la expresión de un gen diana mediante el análisis de la regulación negativa de un segundo gen o gen informador, como se describe en el presente documento. En tal realización, la relación entre el tamaño del ARNbc que tiene secuencia para el gen diana o su complemento y el tamaño del ARNbc que tiene secuencia para el segundo gen o su complemento está entre 2 y 60 10. Se espera que de esta forma, mediante la identificación de los organismos en los que la expresión del segundo gen se regula negativamente de manera eficaz, se identificará una población de organismos en la que la expresión del gen diana se regula negativamente de forma severa.

65 Los procedimientos y medios descritos en el presente documento pueden aplicarse a cualquier organismo eucariota en el que tiene lugar el silenciamiento génico, que incluye pero no se limita a plantas (tales como maíz, algodón, *Arabidopsis*, arroz, verduras, soja, tabaco, árboles etc.), animales invertebrados (tales como insectos, marisco, moluscos, crustáceos tales como cangrejos, langostas y gambas), animales vertebrados (peces, aves, mamíferos, humanos), levaduras y hongos entre otros.

ES 2 346 645 T3

Los siguientes ejemplos no limitantes describen un procedimiento y medios para supervisar el silenciamiento mediado por ARNbc de la expresión de un gen diana a través del análisis de un segundo gen, además de procedimientos y medios para modular el grado de silenciamiento génico mediado por ARNbc en células eucariotas.

5 A menos que se indique otra cosa en los Ejemplos, todas las técnicas de ADN recombinante se llevan a cabo de acuerdo con protocolos convencionales como se describe en Sambrook y col. (1989) *Molecular Cloning: A Laboratory Manual*, Segunda Edición, Cold Spring Harbor Laboratory Press, NY y en los Volúmenes 1 y 2 de Ausubel y col. (1994) *Current Protocols in Molecular Biology*, Current Protocols, Estados Unidos. Se describen materiales y procedimientos convencionales para el trabajo molecular en plantas en *Plant Molecular Biology Labfax* (1993) por R.D.D. Croy, conjuntamente publicado por BIOS Scientific Publications Ltd (UK) y Blackwell Scientific Publications, UK. Otras referencias para técnicas de biología molecular convencionales incluyen Sambrook y Russell (2001) *Molecular Cloning: A Laboratory Manual*, Tercera Edición, Cold Spring Harbor Laboratory Press, NY, Volúmenes I y II de Brown (1998) *Molecular Biology LabFax*, Segunda Edición, Academic Press (UK). Pueden encontrarse materiales y procedimientos convencionales para las reacciones en cadena de la polimerasa en Dieffenbach y Dveksler (1995) *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, y en McPherson y col., (2000) *PCR-Basics: From Background to Bench*, Primera Edición, Springer Verlag, Alemania.

A lo largo de la descripción y los ejemplos, se hace referencia a las siguientes secuencias:

20 SEC ID N° 1: cebador oligonucleotídico para reacción PCR del fragmento CHS/FLC.

SEC ID N° 2: cebador oligonucleotídico para reacción PCR del fragmento CHS/FLC.

25 SEC ID N° 3: cebador oligonucleotídico para reacción PCR del fragmento CHS/FLC.

SEC ID N° 4: cebador oligonucleotídico para reacción PCR del fragmento CHS/FLC.

SEC ID N° 5: sitio de recombinación attB1.

30 SEC ID N° 6: sitio de recombinación attB2.

SEC ID N° 7: pHELLSGATE 8.

35 Ejemplos

Ejemplo 1

Una construcción de ARNhi que tiene como objetivo dos genes da silenciamiento estequiométrico

40 *Construcción de una construcción de ARNhi que tiene como objetivo dos genes*

Se construyó un gen quimérico que codifica una construcción de ARNhi que tiene como objetivo dos genes diferentes que comprende 300 nt del gen del Locus de Flor C (*FLC*) (N° de Acceso del Genbank AF116527, bases 620-920) seguidos de 300 nt del gen de la Chalcona sintasa (*CHS*) (N° de Acceso del Genbank Y18603, bases 147-532) en orientación con sentido, y una secuencia de nucleótidos complementaria a los mencionados 300 nt de CHS seguida de la secuencia de nucleótidos complementaria a los mencionados 300 nt de FLC.

50 El inserto CHS/FLC se generó amplificando un fragmento de 300 pb de FLC con los cebadores oligonucleotídicos que tienen las siguientes secuencias:

- 5' CTCGAGTCTAGAGGAGCAGAAGCTGAGATGGAG 3' (SEC ID N° 1) (cebador 138)

55 - 5' CTGCAGGAATAAGGTACAAAGTTCATC 3' (SEC ID N° 2) (cebador 136) y clonando el producto resultante en pGEM T-easy (Promega, Madison, WI).

El fragmento de CHS se amplificó con los cebadores oligonucleotídicos que tienen las siguientes secuencias:

60 - 5' CTGCAGGCACTGCTAACCCCTGAGAAAC 3' (SEC ID N° 3) (cebador 133)

- 5' GGTACCTTGACTTGGGCTGGCCCCACT 3' (SEC ID N° 4) (cebador 143) y también se clonaron en un pGEM T-easy.

65 Un fragmento de 300 pb se escindió de pGEM T-easy CHS que contenía el fragmento de CHS y se insertó en un sitio PstI de pGEM T-easy FLC. El inserto quimérico CHS/FLC resultante se amplificó usando cebadores con las mismas secuencias específicas de gen que 138 y 143 pero modificadas para comprender también la secuencia de

ES 2 346 645 T3

nucleótidos que codifica los sitios de recombinación attB1 (GGGGACAAGTTTGTACAAAAAAGCAGGCT; SEC ID N° 5) y attB2 (GGGACCACTTTGTACAAGAAAGCTGGGT; SEC ID N° 6), respectivamente en sus extremos 5' con Taq ADN polimerasa F1 (Fisher Biotec, Subiaco, WA, Australia) usando el protocolo del fabricante.

5 Los productos de PCR se precipitaron mediante la adición de 3 volúmenes de TE y dos volúmenes de PEG 3000 al 30% (p/v), MgCl₂ 30 mM y centrifugación a 13000 g durante 15 minutos. La reacción de recombinación de productos de PCR con pDONR201 (Invitrogen, Groningen, Países Bajos) se llevó a cabo en un volumen total de 10 μ l con 2 μ l de tampón BP clonasa (Invitrogen), 1-2 μ l de producto de PCR, 150 ng de vector plasmídico y 2 μ l de BP clonasa (Invitrogen). La reacción se incubó a temperatura ambiente (25°C) durante entre 1 h y una noche. Después de la
10 incubación, se añadió 1 μ l de proteinasa K (2 μ g/ μ l; Invitrogen) y se incubó durante 10 minutos a 37°C. Se usaron 1-2 μ l de la mezcla para transformar *E. coli* DH5 α y las colonias se seleccionaron con los antibióticos apropiados. Los clones se comprobaron por digestión de minipreparaciones de ADN o por PCR.

Las reacciones de recombinación de clones de pDONR201 con pHellsgate 8 se llevaron a cabo en un volumen total
15 de 10 μ l con 2 μ l de tampón LR clonasa (Invitrogen), 2 μ l de clon de pDONR201 (aproximadamente 150 ng), 300 ng de pHellsgate 8 y 2 μ l de LR clonasa (Invitrogen). La reacción se incubó durante una noche a temperatura ambiente y la proteinasa se trató y se usó para transformar DH5 α como para la reacción de BP clonasa.

pHELLSGATE 8 es un vector adecuado para la clonación recombinante de tal manera que se genera una repetición
20 invertida del ADN inserto de interés. La construcción de pHELLSGATE 8, y su uso para la clonación de recombinación se han descrito en la solicitud PCT PCT/AU02/00073, US 60/264.067 o la solicitud de prioridad de Estados Unidos US 60/333.743. pHELLSGATE 8 comprende los siguientes elementos de ADN que codifican

- 25 • una secuencia de extremo derecho de la región de ADN-T de *Agrobacterium tumefaciens*;
- una región promotora de CaMV 35S;
- un sitio de recombinación attR1;
- 30 • un gen marcador de selección ccdB de *E. coli*;
- un sitio de recombinación attR2;
- 35 • una secuencia intrón de *pdk2* (intrón 2 de piruvato ortofosfato diquinasa de *Flavineria trinerva*);
- un sitio de recombinación de attR2;
- un gen marcador de selección ccdB;
- 40 • un sitio de recombinación attR1;
- una región terminadora del gen de la octopina sintasa de *Agrobacterium tumefaciens* (ocs);
- un gen *npII* quimérico flanqueado por un promotor del gen de la nopalina sintasa (*nos*) y un terminador nos;
- 45 • una región de ADN-T de extremo izquierdo de *Agrobacterium tumefaciens*;
- un origen de replicación para *E. coli* y un gen de resistencia a estreptomycinina.

50 La secuencia completa de pHELLSGATE 8 se representa en SEC ID N° 7.

Tras la clonación recombinante del CHS/FLC, el inserto amplificado por PCR como se describe anteriormente, se
55 genera un gen quimérico que comprende

- una región promotora de CaMV35S;
- una secuencia de nucleótidos que corresponde a las bases 620-920 de *FLC*;
- 60 • una secuencia de nucleótidos que corresponde a las bases 147-532 de *CHS*;
- un intrón *pdk2*
- una secuencia de nucleótidos complementaria a las bases 147-532 de *CHS*;
- 65 • una secuencia de nucleótidos complementaria a las bases 620-920 de *FLC*;
- una región terminadora de ocs.

ES 2 346 645 T3

Este gen quimérico se localiza entre las secuencias de los extremos de ADN-T de *Agrobacterium*, junto con el gen marcador seleccionable expresable en plantas. El vector puede así introducirse en *Agrobacterium* que comprende las funciones auxiliares requeridas, y usarse para transformar células vegetales.

5 *Transformación de plantas*

La transformación de ecotipo *Arabidopsis* C24 se realizó mediante el procedimiento de inmersión floral (Clough y Bent, 1998). Las plantas se seleccionaron en medio MS solidificado con agar suplementado con timentina a 100 mg/l y kanamicina a 50 mg/l.

10

Procedimientos de análisis fenotípico para plantas en las que los genes FLC o CHS están silenciados

Las plantas con ARNhi *FLC* T1 se valoraron transfiriéndolas a placas MS y contando los días hasta la floración de las flores o de las hojas en roseta comparadas con las plantas de tipo silvestre C24 y las líneas mutantes *flc*. Las líneas vegetales en las que la expresión del gen *FLC* se reduce o elimina florecen antes que las del ecotipo C24 de tipo silvestre. Un bloque de biosíntesis de antocianina por reducción de la expresión del gen de la chalcona sintasa (*CHS*) lleva a una acumulación de malonil-CoA, que aumenta la fluorescencia de la semilla observada bajo luz UV.

15

Todas las plantas T1 transformadas con la construcción hi *CHS/FLC* florecieron antes que el tipo silvestre C24 (36 días) y más tarde que la línea *flc-13* marcada con transposón (17 días; Sheldon y col, 1999). Las semillas de cuatro plantas hi *CHS/FLC* que florecieron a tiempos diferentes se ensayaron con respecto a su fluorescencia (Figura 1) comparadas con el tipo silvestre C24, una línea 35S::GUS y una línea *CHS* silenciada por ARNhi homocigota (Wesley y col, 2001). Las concentraciones relativas de los diferentes intermedios en la ruta de la chalcona sintasa también se determinaron por análisis de HPLC.

25

La línea *CHS* ARNhi muestra una fuerte fluorescencia que refleja la acumulación de malonil-CoA en esta línea mientras que el tipo silvestre C24 y las líneas 35S::GUS no muestran fluorescencia bajo luz UV. Las cuatro líneas hi *CHS/FLC* examinadas mostraron algo de fluorescencia, mostrando las líneas de floración más temprana la mayor fluorescencia.

30

Ejemplo 2

35 *Modulación del grado de silenciamiento del gen CHS por dilución con secuencias extra no diana en el brazo del ARNhi*

Para ensayar la influencia de la inclusión de nucleótidos extra, no relacionados con las secuencias de ADN diana, en las moléculas de ARNhi o ARNbc sobre el grado de silenciamiento de un gen diana, se generaron varios genes quiméricos que comprendían cada uno 100 pb de la secuencia de nucleótidos de *CHS* (que corresponde al N° de Acceso del Genbank Y18603) en orientación tanto con sentido como antisentido (conocidas como *CHS*_{antisentido} y *CHS*_{sentido} respectivamente).

40

Progresivamente, más secuencias de nucleótidos (desde 100 pb hasta 900 pb; conocidas como Extra-NNN-nt_{sentido} o Extra-NNN-nt_{antisentido}) no relacionadas con el gen *CHS* se insertaron tanto en orientación con sentido como antisentido. Las secuencias de nucleótidos no relacionadas se obtuvieron del gen *FLC* (que corresponde al N° de Acceso del Genbank Nr AF116527).

45

Las diferentes construcciones de nucleótidos *CHS/Extra* se amplificaron por PCR con cebadores ampliados attB1 y attB2 y se introdujeron en pHELLSGATE 8 como se describe en el Ejemplo 1. Los plásmidos resultantes contienen así los siguientes genes de ARNbc quiméricos:

50

1. <CaMV35S-*CHS*_{sentido}-Extra_100_nt_{sentido}-intrónpdk2-Extra_100_nt_{antisentido}-*CHS*_{antisentido}-terminador ocs> o
2. <CaMV35S-*CHS*_{sentido}-Extra_300_nt_{sentido}-intrónpdk2-Extra_300_nt_{antisentido}-*CHS*_{antisentido}-terminador ocs> o
3. <CaMV35S-*CHS*_{sentido}-Extra_500_nt_{sentido}-intrónpdk2-Extra_500_nt_{antisentido}-*CHS*_{antisentido}-terminador ocs> o
4. <CaMV35S-*CHS*_{sentido}-Extra_900_nt_{sentido}-intrónpdk2-Extra_900_nt_{antisentido}-*CHS*_{antisentido}-terminador ocs>

55

60

Las líneas de planta *Arabidopsis* transformadas que comprenden uno de los cuatro genes quiméricos mencionados anteriormente se generaron como se describe en el Ejemplo 1, y la fluorescencia bajo luz UV para supervisar el silenciamiento del gen *CHS* se valoró como se describe en el Ejemplo 1 para semillas de plantas de cada línea. Las líneas transgénicas independientes se clasificaron de acuerdo con el grado de fluorescencia. Los resultados se resumen en la Tabla 1.

65

ES 2 346 645 T3

TABLA 1

Número de líneas transgénicas independientes que muestran distintas clases de fluorescencia

	Nivel de fluorescencia			
	-	+	++	+++
100 nt CHS+ 100 nt extra	12	0	20	0
100 nt CHS+ 300 nt extra	11	18	11	1
100 nt CHS+ 500 nt extra	5	7	13	0
100 nt CHS+ 900 nt extra	18	8	3	0

A partir de estos resultados está claro que cuanto mayor es la relación de las secuencias de nucleótidos extra frente a las secuencias diana en la construcción de ARNhi, mayor es la proporción de líneas de plantas transgénicas obtenidas con un nivel bajo de fluorescencia, es decir, con un grado más bajo de silenciamiento de la expresión del gen diana CHS.

La inclusión de secuencias extra no relacionadas en ARNhi o ARNbc desplaza eficazmente la distribución del silenciamiento en la población de líneas transgénicas hacia el extremo inferior del espectro del silenciamiento génico.

Ejemplo 3

Análisis Northern de líneas de plantas transgénicas que comprenden un gen quimérico que codifica simultáneamente regiones de ARNbc dirigidas a FLC y CHS

El ARN se preparó a partir de diferentes líneas de planta *Arabidopsis* que comprenden dos construcciones en horquilla FLCCHS. Una construcción fue un vector Hellsgate12 modificado (véase el documento WO02/059294) con una horquilla FLC próxima al intrón con un fragmento de CHS insertado en los sitios attR. La otra construcción fue un Hellsgate12 modificado con una horquilla CHS próxima al intrón con un fragmento de FLC insertado en los sitios attR. Se procesaron y se sometieron a transferencia dos geles idénticos y se sondearon con sondas de ARN antisentido de FLC o CHS. Las manchas de transferencia se trataron con ARNasa para asegurar que las señales fueran específicas. El auto-radiograma resultante se representa en la Figura 2A. Las estimaciones de la cantidad de hibridación para las diferentes sondas y líneas se resumen en la Tabla 2 y está representadas gráficamente en la Figura 2B. Parece haber una alta relación directa lineal entre la cantidad de señales de FLC y CHS en cada línea, lo que indica que cuando un gen se silencia el otro gen también se silencia en el mismo grado.

(Tabla pasa a página siguiente)

ES 2 346 645 T3

TABLA 2

Estimaciones de señales de hibridación en distintas plantas usando sondas antisentido de FLC o CHS

5

10

15

20

25

30

35

40

45

Planta	ARNm de FLC	ARNm de CHS
HG12 2FLC + 2CHS N° 1	13867	15154
HG 12 2FLC + 2CHS N° 2	21962	26467
HG 12 2FLC + 2CHS N° 3	40434	40096
HG12 2FLC + 2CHS N° 4	0	0
HG 12 2FLC + 2CHS N° 5	41122	38987
HG12 2FLC + 2CHS N° 6	16609	23091
HG12 2FLC + 2CHS N° 7	18417	3357
HG 12 2FLC + 2CHS N° 8	16463	17400
HG 12 2FLC + 2CHS N° 9	47210	59680
HG12 2FLC + 2CHS N° 10	10792	18879
HG12 2CHS + 2FLC N° 1	3399	9694
HG12 2CHS + 2FLC N° 2	3692	14317
HG12 2CHS + 2FLC N° 3	1725	51004
HG12 2CHS + 2FLC N° 4	4620	16067
HG12 2CHS + 2FLC N° 5	21826	38643
HG12 2CHS + 2FLC N° 7	13288	34043
HG12 2CHS + 2FLC N° 8	6716	5949
HG12 2CHS + 2FLC N° 9	3027	8036
HG12 2CHS + 2FLC N° 10	1021	4677
C24 tipo silvestre	23376	36279

Referencias

An y col., 1996 *The Plant Cell* 8, 15-30

Butler y Chamberlin, 1982 *J. Biol. Chem.*, 257: 5772-5778

Dunn y Studier 1983 *J. Mol. Biol.* 166: 477-535

Dunn y col., 1971 *Nature New Biology*, 230: 94-96

Fire y col., 1998 *Nature* 391, 806-811

Hamilton y col. 1998 *Plant J.* 15: 737-746

Hapster y col., 1988 *Mol. Gen. Genet.* 212, 182-190

Hausmann, 1976 *Current Topics in Microbiology and Immunology*, 75: 77-109

Hudspeth y col., 1989 *Plant Mol Biol* 12: 579-589

Keil y col., 1989 *EMBO J.* 8: 1323-1330

Keller y col., 1988 *EMBO J.* 7: 3625-3633

ES 2 346 645 T3

Keller y col., 1989 *Genes Devel.* 3: 1639-1646

Klement y col., 1990 *J. Mol. Biol.* 215, 21-29

5 **Korsten** y col., 1975 *J. Gen Virol.* 43: 57-73

McGraw y col, 1985 *Nucl. Acid Res.* 13: 6753-6766

Needleman y **Wunsch** 1970

10 **Peleman** y col., 1989 *Gene* 84: 359-369

Smith y col., 2000 *Nature* 407: 319-320

15 **Towle** y col., *J. Biol. Chem.* 250: 1723-1733 (1975)

Waterhouse y col. 1998 *Proc. Natl. Acad. Sci. USA* 95: 13959-13964.

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Un procedimiento de supervisión de la reducción de la expresión de un gen diana en una célula de un organismo eucariota, comprendiendo dicho procedimiento las etapas de

- a) proporcionar a la célula de dicho organismo eucariota un ARN^{bc} que comprende una primera región, una segunda región, una tercera región y una cuarta región
 - i) comprendiendo dicha primera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de dicho gen diana;
 - ii) comprendiendo dicha segunda región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con una secuencia de nucleótidos complementaria a 19 nucleótidos consecutivos de dicha región de nucleótidos con sentido del gen diana;
 - iii) siendo capaces dicha primera región y dicha segunda región de formar una región de ARN bicatenario;
 - iv) comprendiendo dicha tercera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de un segundo gen presente en dicha célula eucariota y que es diferente de dicho gen diana;
 - v) comprendiendo dicha cuarta región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con el complemento de dichos 19 nucleótidos consecutivos de dicha región de nucleótidos con sentido de dicho segundo gen;
 - vi) siendo capaces dicha tercera y cuarta región de formar una región de ARN bicatenario; y
- b) supervisar la reducción de la expresión de dicho gen diana mediante el análisis de la reducción en expresión de dicho segundo gen,

en el que el organismo eucariota es una planta, una levadura, un hongo, un mohó, o un animal, a condición de que cuando dicho organismo es un animal, la célula no sea una célula en un cuerpo humano o animal.

2. El procedimiento de la reivindicación 1, en el que dicho organismo eucariota se selecciona del grupo constituido por algodón, patata, maíz, trigo, arroz, caña de azúcar, colza, *Arabidopsis*, remolacha azucarera, tabaco y soja.

3. El procedimiento de la reivindicación 1 o reivindicación 2, en el que dicho organismo eucariota se selecciona del grupo constituido por insectos, marisco, moluscos, crustáceos, cangrejos, langostas, gambas, peces, aves, mamíferos y seres humanos.

4. El procedimiento de una cualquiera de las reivindicaciones 1 a 3, en el que dicho segundo gen es un gen endógeno presente en dicha célula de dicho organismo eucariota.

5. El procedimiento de una cualquiera de las reivindicaciones 1 a 3, en el que dicho segundo gen es un transgén integrado de forma estable en el genoma de dicha célula de dicho organismo eucariota.

6. El procedimiento de la reivindicación 1 o reivindicación 2, en el que dicho organismo es una planta y dicho segundo gen se selecciona del grupo que consiste en PDS, EIN2, FLC y PhyB.

7. El procedimiento de una cualquiera de las reivindicaciones 1 a 3, en el que dicha célula de dicho organismo eucariota comprende un gen GUS o GFP expresado de forma funcional y dicho segundo gen es un gen GUS o GFP.

8. El procedimiento de una cualquiera de las reivindicaciones 1 a 7, en el que las regiones de ARN bicatenario formadas entre la primera y la segunda región y la región de ARN bicatenaria formada entre la tercera y cuarta región son aproximadamente iguales en tamaño.

9. El procedimiento de una cualquiera de las reivindicaciones 1 a 7, en el que dicha primera y segunda región y dicha tercera y cuarta región tienen una longitud de aproximadamente 300 nt.

10. El procedimiento de una cualquiera de las reivindicaciones 1 a 9, en el que dicho ARN^{bc} se transcribe a partir de un gen quimérico comprendido dentro de las células de dicho organismo eucariota, comprendiendo dicho gen quimérico:

ES 2 346 645 T3

- a) una región promotora que funciona en dicha célula eucariota; unida operativamente a
- b) una región de ADN que al transcribirse produce dicha molécula de ARNbc; y
- 5 c) una región de terminación de la transcripción y de poliadenilación que actúa en dicho organismo eucariota.

10 11. El procedimiento de acuerdo con la reivindicación 10, en el que dicho gen quimérico está integrado de forma estable en el genoma de células de dicho organismo eucariota.

12. Una molécula de ARN como se describe en una cualquiera de las reivindicaciones 1 a 11 en la que las regiones de ARN bicatenario formadas entre la primera y la segunda región y la región de ARN bicatenario formada entre la tercera y cuarta región son aproximadamente iguales en tamaño.

15 13. Uso de una molécula de ARN de acuerdo con la reivindicación 12 para medir la reducción de expresión de un gen diana en una célula de un organismo eucariota, en el que dicho organismo eucariota es una planta, una levadura, un hongo, un moho o un animal, a condición de que cuando dicho organismo es un animal la célula no sea una célula de un cuerpo humano o animal.

20 14. Uso de una molécula de ARN de acuerdo con la reivindicación 12, para reducir la expresión de dicho gen diana y dicho segundo gen en una célula de un organismo eucariota, en el que dicho organismo eucariota es una planta, una levadura, un hongo, un moho o un animal, a condición de que cuando dicho organismo es un animal, la célula no sea una célula en un cuerpo humano o animal.

25 15. Una molécula de ADN para medir la reducción de la expresión de un gen diana en una célula de un organismo eucariota, que comprende

- a) una región promotora que actúa en dicha célula eucariota; unida operativamente a
- 30 b) una región de ADN que cuando se transcribe produce una molécula de ARNbc, comprendiendo dicho ARNbc una primera, segunda, tercera y cuarta región;
 - 35 i) comprendiendo dicha primera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de dicho gen diana;
 - 40 ii) comprendiendo dicha segunda región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con una secuencia de nucleótidos complementaria a 19 nucleótidos consecutivos de dicha región de nucleótidos con sentido del gen diana;
 - 45 iii) siendo capaces dicha primera región y dicha segunda región de formar una región de ARN bicatenario;
 - 50 iv) comprendiendo dicha tercera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de un segundo gen presente en dicha célula eucariota y que es diferente de dicho gen diana;
 - 55 v) comprendiendo dicha cuarta región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con el complemento de dichos 19 nucleótidos consecutivos de dicha región de nucleótidos con sentido de dicho segundo gen;
 - vi) siendo capaces dicha tercera y cuarta región de formar una región de ARN bicatenario;
- en la que las regiones de ARN bicatenario formadas entre la primera y la segunda región y la región de ARN bicatenario formada entre la tercera y cuarta región son aproximadamente iguales en tamaño; y
- 60 c) una región de terminación de la transcripción y de poliadenilación que actúa en dicho organismo eucariota, en la que dicho segundo gen es un gen endógeno de dicho organismo eucariota o un transgén integrado de manera estable en el genoma de células de dicho organismo eucariota.

65 16. Uso de una molécula de ADN de acuerdo con la reivindicación 15 para medir la expresión de dicho gen diana mediante la medición de la reducción de la expresión de dicho segundo gen en una célula de un organismo eucariota, en el que dicho organismo eucariota es una planta, una levadura, un hongo, un moho o un animal, a condición de que cuando dicho organismo es un animal, la célula no es una célula en un cuerpo animal o humano.

ES 2 346 645 T3

17. Uso de una molécula de ADN de acuerdo con la reivindicación 15 para reducir la expresión de dicho gen diana y dicho segundo gen, en una célula de un organismo eucariota, en el que dicho organismo eucariota es una planta, una levadura, un hongo, un moho o un animal, a condición de que cuando dicho organismo es un animal, la célula no sea una célula en un cuerpo humano o animal.

5

18. Un organismo eucariota no humano que comprende una molécula de ARN de acuerdo con la reivindicación 12.

10

19. Un organismo eucariota no humano que comprende una molécula de ADN de acuerdo con la reivindicación 15.

20. El organismo eucariota no humano de acuerdo con las reivindicaciones 18 ó 19, que se selecciona de una planta, un animal, una levadura, un hongo o un moho.

15

21. Un procedimiento para identificar, dentro de una población de organismos eucariotas con silenciamiento génico mediado por ARNbc, los organismos con el grado deseado de silenciamiento de un gen diana, que comprende:

20

a) proporcionar células de dichos organismos eucariotas con un ARNbc que comprende una primera región, una segunda región, una tercera región y una cuarta región

25

i) comprendiendo dicha primera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de dicho gen diana;

30

ii) comprendiendo dicha segunda región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con una secuencia de nucleótidos complementaria a 19 nucleótidos consecutivos de dicha región de nucleótidos con sentido del gen diana;

35

iii) siendo capaces dicha primera región y dicha segunda región de formar una región de ARN bicatenario;

40

iv) comprendiendo dicha tercera región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con 19 nucleótidos consecutivos de una región de nucleótidos con sentido de un segundo gen presente en dicha célula eucariota y que es diferente de dicho gen diana;

45

v) comprendiendo dicha cuarta región una secuencia de nucleótidos de al menos 19 nucleótidos consecutivos que tiene al menos un 94% de identidad de secuencia con el complemento de dichos 19 nucleótidos consecutivos de dicha región de nucleótidos con sentido de dicho segundo gen;

vi) siendo capaces dicha tercera y cuarta región de formar una región de ARN bicatenario; y

50

b) identificar dicho organismo con dicho grado deseado de silenciamiento de dicho gen diana, seleccionando dichos organismos con el grado deseado de silenciamiento de dicho segundo gen,

en el que dicho organismo eucariota es una planta, una levadura, un hongo, un moho o un animal, a condición de que cuando dicho organismo es un animal, la célula no sea una célula en un cuerpo humano o animal.

55

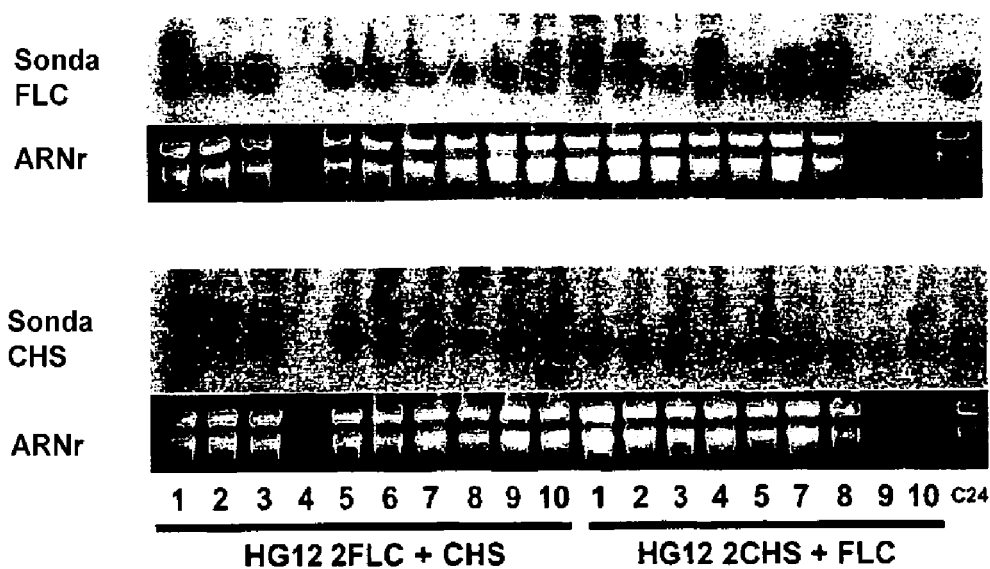
22. El procedimiento de acuerdo con la reivindicación 21, en el que las regiones de ARN bicatenario formadas entre la primera y la segunda región y la región de ARN bicatenario formada entre la tercera y cuarta región son aproximadamente iguales en tamaño.

60

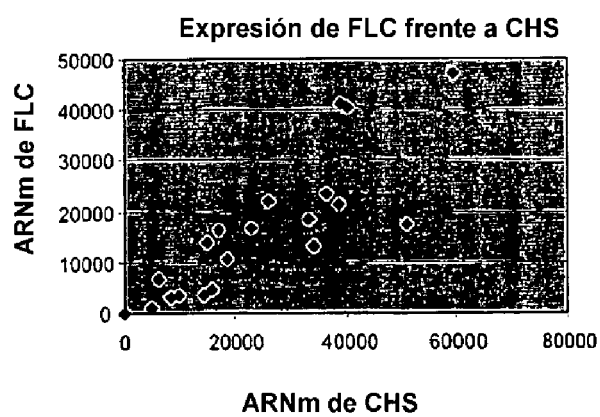
65



Figura 1



A.



B.

Figura 2

ES 2 346 645 T3

LISTA DE SECUENCIAS

- <110> Commonwealth Scientific and Industrial Research Organization
5 <120> Procedimientos y medios para control y modulación del silenciamiento génico
<130> COLINA-PCT
<150> US 60363852
<151> 14-03-2003
10 <160> 7
<170> PatentIn versión 3.0
- <210> 1
15 <211> 33
<212> ADN
<213> Artificial
- 20 <220>
<223> cebador oligonucleotídico para amplificación por PCR de fragmento CHS/FLC
- <400> 1
25 ctcgagtcta gaggagcaga agctgagatg gag 33
- <210> 2
30 <211> 27
<212> ADN
<213> Artificial
- 35 <220>
<223> cebador oligonucleotídico para amplificación por PCR de fragmento CHS/FLC
- <400> 2
40 ctgcaggaat aaggtacaaa gttc 27
- <210> 3
45 <211> 27
<212> ADN
<213> Artificial
- <220>
50 <223> cebador oligonucleotídico para amplificación por PCR de fragmento CHS/FLC
- <400> 3
55 ctgcaggcac tgctaaccct gagaaac 27
- <210> 4
60 <211> 27
<212> ADN
<213> Artificial
- <220>
65 <223> cebador oligonucleotídico para amplificación por PCR de fragmento CHS/FLC

ES 2 346 645 T3

<400> 4

ggtacctga ctgggctgg cccact 27

5 <210> 5

<211> 29

<212> ADN

10 <213> Artificial

<220>

<223> sitio de recombinación de attB1

15 <400> 5

ggggacaagt ttgtacaaa aagcaggct 29

20 <210> 6

<211> 27

<212> ADN

<213> Artificial

25

<220>

<223> sitio de recombinación de attB2

30 <400> 6

ggaccactt gtacaagaaa gctgggt 27

<210> 7

35 <211> 17476

<212> ADN

<213> Artificial

40 <220>

<223> plásmido pHELLSGATE 8

45

50

55

60

65

ES 2 346 645 T3

<400> 7

	ggccgcacta	gtgatatccc	gcccgcactg	cgcccgagg	catgagcgt	cgccccaat	60
5	tcccccata	gtgagtcgta	ttacaattca	ctggccgtcg	ttttacaacg	tcgtgactgg	120
	gaaaaccctg	gcgttaacca	acttaatcgc	cttgacagcac	atcccccttt	cgccagctgg	180
	cgtaatagcg	aagaggcccg	caccgatcgc	ccttcccaac	agttgagcag	cctgaatggc	240
	gaatggaaat	tgtaaacggt	aatgggtttc	tggagttaa	tgagctaagc	acatacgtca	300
10	gaaaccatta	ttgcgcgttc	aaaagtgcgc	taaggtcact	atcagctagc	aaatatttct	360
	tgtaaaaaat	gctccactga	cgttccataa	attccccctcg	gtatccaatt	agagtctcat	420
	attcactctc	aatccaaata	atctgcaatg	gcaattacct	tatecgcaac	ttctttacct	480
15	atctccgccc	ggatccgggc	aggttctcgc	gccgcttggg	tggagaggct	attcggctat	540
	gactgggcac	aacagacaat	cggtctctct	gatgccgccc	tgttccggct	gtcagcgcag	600
	gggcccgcg	ttctttttgt	caagaccgac	ctgtccggtg	ccctgaatga	actgcaggac	660
	gaggcagcgc	ggctatcgtg	gctggccacg	acgggcttcc	cttgagcagc	tgtgctcgac	720
20	gttgtcactg	aagcgggaag	ggactggctg	ctattgggcg	aagtgccggg	gcaggatctc	780
	ctgtcatctc	accttgctcc	tgccgagaaa	gtatccatca	tggctgatgc	aatgcccggg	840
	ctgcatacgc	ttgatccggc	tacctgccc	ttcgaccacc	aagcgaaca	tcgcatcgag	900
25	cgagcagcgt	ctcggatgga	agccggctct	gtcagtcagg	atgatctgga	cgaagagcat	960
	caggggctcg	cgccagccga	actgttccgc	aggtcaagg	cgccatgcc	cgacggcgag	1020
	gatctcgtcg	tgaccatgg	cgatgcctgc	ttgccgaata	tcatggtgga	aaatggccgc	1080
	ttttctggat	tcacgactg	tggccggctg	gggtgtggcg	accgctatca	ggacatagcg	1140
30	ttggctacct	gtgatattgc	tgaagagctt	ggcggcgaat	gggctgaccg	cttctcgtg	1200
	ctttacggta	tcgcccctcc	cgattcgcag	cgcatcgctt	tctatcgctt	tcttgacgag	1260
	ttcttctgag	cgggactctg	gggttcgaaa	tgaccgacca	agcagcgcgc	aacctgccat	1320
35	cacgagattt	cgattccacc	gccgccttct	atgaaagggt	gggcttcgga	atcgttttcc	1380
	gggacgcccg	ctggatgata	ctccagcgcg	gggatctcat	gctggagttc	ttcgcccacc	1440
	ccgatccaac	acttacgttt	gcaacgtcca	agagcaaata	gaccaacgac	gccggaagg	1500
40	tgccgcagcg	tgtggattgc	gtctcaattc	tctcttgag	gaatgcaatg	atgaatatga	1560
	tactgactat	gaaactttga	gggaatactg	cctagcaccg	tcacctcata	acgtgcatca	1620
	tgcatgccct	gacaacatgg	aacatcgcta	tttttctgaa	gaattatgct	cgttggagga	1680
	tgtcgcccga	attgcagcta	ttgccaacat	cgaactacc	ctcacgcatg	cattcatcaa	1740
45	tattattcoat	gcggggaaag	gcaagattaa	tccaactggc	aaatcatcca	gcgtgattgg	1800
	taacttcagt	tcagcagact	tgattcgttt	tgggtctacc	cacgttttca	ataaggacga	1860
	gatggtggag	taaagaagga	gtgcgtcgaa	gcagatcggt	caaacatttg	gcaataaagt	1920
50	ttcttaagat	tgaatcctgt	tgccggctct	gcgatgatta	tcatataaatt	tctgttgaat	1980
	tacgttaagc	atgtaataat	taacatgtaa	tgcatgacgt	tatttatgag	atgggttttt	2040
	atgattagag	tcccgcgaat	atacatttaa	tacgcgatag	aaaacaaaat	atagcgcgca	2100
55	aactaggata	aattatcgcg	cgccggtgta	tctatgttac	tagatcgaat	taattccagg	2160

60

65

ES 2 346 645 T3

	cggggaaggg	caatcagctg	ttgcccgtct	caactggtgaa	aagaaaaacc	accccagtac	2220
	attaaaaacg	tccgcaatgt	gttattaagt	tgtctaagcg	tcaatttggt	tacaccacaa	2280
5	tatatcctgc	caccagccag	ccaacagctc	cccgaccggc	agctcggcac	aaaatcacca	2340
	ctcgatacag	gcagcccac	agtcggggac	ggcgtcagcg	ggagagccgt	tgtaaggcgg	2400
	cagactttgc	tcattgttacc	gatgctattc	ggaagaacgg	caactaagct	gccgggtttg	2460
	aaacacggat	gatctcgcgg	agggtagcat	gttgattgta	acgatgacag	agcgttgctg	2520
10	cctgtgatca	aatatcatct	ccctcgcaga	gatccgaatt	atcagccttc	ttattcattt	2580
	ctcgcttaac	cgtgacaggc	tgctgatcct	gagaactatg	ccgacataat	aggaaatcgc	2640
	tggataaagc	cyctgaggaa	gctgagtggc	gctatttctt	tagaagtgaa	cgttgacgat	2700
15	gtcgcaggat	cttttccgct	gcataaccct	gcttcggggg	cattatagcg	attttttcgg	2760
	tatatccatc	ctttttcgca	cgatatacag	gattttgcca	aagggttcgt	gtagactttc	2820
	cttgggtgat	ccaacggcgt	cagccgggca	ggataggtga	agtaggccc	cccgcgagcg	2880
	ggtgttccct	cttcactgtc	ccttattcgc	acctggcggt	gctcaacggg	aatcctgctc	2940
20	tcgagggctg	gccggctacc	gccggcgtaa	cagatgaggg	caagcggatg	gctgatgaaa	3000
	ccaagccaac	caggggtgat	gctgccaact	tactgattta	gtgtatgatg	gtgtttttga	3060
	ggtgctccag	tggettctgt	ttctatcagc	tgccctcct	gttcagctac	tgacgggggtg	3120
25	gtgcgtaacg	gcaaaagcac	cgccggacat	cagcgcctac	tctgctctca	ctgccgtaaa	3180
	acatggcaac	tcagttcac	ttacaccgct	tctcaaccgg	gtacgcacca	gaaaatcatt	3240
	gatatggcca	tgaatggcgt	tggatgccgg	gcaacagccc	gcattatggg	cgttggcctc	3300
	aacacgattt	tcagtcactt	aaaaaactca	ggccgcagtc	ggtaacctcg	cgcatacagc	3360
30	cgggcagtga	cgctcatcgc	tcgcccggaaa	tggacgaaca	gtggggctat	gtcggggcta	3420
	aatcgcgcca	gcgctggctg	ttttacgcgt	atgacagtct	ccggaagacg	gttgttgcgc	3480
	acgtattcgg	tgaacgcact	atggcgacgc	tggggcgctt	tatgagcctg	ctgtcacctt	3540
35	ttgacgtggt	gatatggatg	acggatggct	ggccgcgtga	tgaatcccgc	ctgaagggaa	3600
	agctgcacgt	aatcagcaag	cgatatacgc	agcgaattga	gcggcataac	ctgaatctga	3660
	ggcagcacct	ggcacggctg	ggacggaagt	cgctgtcgtt	ctcaaaatcg	gtggagctgc	3720
	atgacaaaagt	catcgggcat	tatctgaaca	taaaacacta	tcaataagtt	ggagtcatta	3780
40	cccaaccagg	aagggcagcc	cacctatcaa	ggtgtaactg	cttccagacg	aacgaagagc	3840
	gattgaggaa	aagggcggcg	cggccggcat	gagcctgtcg	gcctacctgc	tggccgctcg	3900
	ccagggctac	aaaatcacgg	gcgtcgtgga	ctatgagcac	gtccgcgagc	tggcccgcac	3960
45	caatggcgac	ctgggcccgc	tgggcccctt	gctgaaactc	tggctcaccg	acgaccccg	4020
	cacggcgcg	ttcgggtgat	ccacgatcct	cgccctgctg	gcgaagatcg	aagagaagca	4080
	ggacgagctt	ggcaagggtca	tgatgggctg	ggtccgccc	agggcagagc	catgactttt	4140
	ttagccgcta	aaacggccgg	ggggtgcgcg	tgattgcaa	gcacgtccc	atgcgctcca	4200
50	tcaagaagag	cgacttcgcg	gagctggtat	tcgtgcaggg	caagattcgg	aataccaagt	4260
	acgagaagga	cggccagacg	gtctacggga	ccgacttcat	tgccgataag	gtggattatc	4320
	tggacaccaa	ggcaccaggc	gggtcaaate	aggaataagg	gcacattgcc	ccggcggtgag	4380
55	tcggggcaat	cccgcaagg	gggtgaatga	atcggacggt	tgaccggaag	gcatacaggc	4440
	aagaactgat	cgacgcgggg	ttttccgccc	aggatgccga	aaccatcgca	agccgcaccg	4500
	tcattgcgtg	gccccgcgaa	accttcagat	ccgtcggctc	gatggtccag	caagctacgg	4560
	ccaagatcga	gcgcgacagc	gtgcaactgg	ctccccctgc	cctgcccgcg	ccatcggccc	4620
60	ccgtggagcg	ttcgcgctgt	ctcgaacagg	aggcggcagg	tttggcgaag	tcgatgacca	4680
	tcgacacgcg	aggaactatg	acgaccaaga	agcgaaaaac	cgccggcgag	gacctggcaa	4740

65

ES 2 346 645 T3

	aacaggtcag	cgaggccaag	caggcccggt	tgctgaaaca	cacgaagcag	cagatcaagg	4800
	aatgcagct	ttccttggtc	gatattgctc	cgtggccgga	cacgatgcga	gcgatgccaa	4860
5	acgacacggc	ccgctctgcc	ctgttcacca	cgcgcaacaa	gaaaatcccg	cgcgaggcgc	4920
	tgcaaaacaa	ggtrattttc	cacgtcaaca	aggacgtgaa	gatcacctac	accggcgctc	4980
	agctgccccg	cgacgatgac	gaactgggtg	ggcagcaggt	gttggagtac	gcgaagcgca	5040
	cccctatcgg	cgagccgata	accttcacgt	tctacgagct	ttgccaggac	ctgggtgtgt	5100
10	cgatcaatgg	ccggtattac	acgaaggccg	aggaatgcct	gtcgcgccta	caggcgacgg	5160
	cgatgggctt	cacgtccgac	cgcgttgggc	acctggaatc	ggtgtcgcct	ctgcaccgct	5220
	tccgcgtcct	ggaccgtggc	aagaaaacgt	cccgttgcca	ggtcctgatc	gacgagggaa	5280
15	tcgtcgtgct	gtttgctggc	gaccactaca	cgaaattcat	atgggagaag	taccgcaagc	5340
	tgtcgcccgc	ggccccgacg	atgttcgact	atttcagctc	gcaccgggag	ccgtaccgcg	5400
	tcaagctgga	aaccttccgc	ctcatgtgcg	gatcggatc	cacccgcgtg	aagaagtggc	5460
	gcgagcaggt	cgccgaagcc	tgcgaagagt	tgccgaggcag	cgccctgggtg	gaacacgcct	5520
20	gggtcaatga	tgacctgggtg	cattgcaaac	gctagggcct	tgtggggtca	gttccggctg	5580
	ggggttcagc	agccagcgc	ttactggcat	ttcaggaaca	agcgggcaact	gctcgcgcga	5640
	cttgcttcgc	tcagtatcgc	tcgggacgca	cgccgcgctc	tacgaactgc	cgataaacag	5700
25	aggattaaaa	ttgacaattg	tgattaaggc	tcagattcga	cggttggag	cgccgcacgt	5760
	gcaggatttc	cgcgagatcc	gattgtcggc	cctgaagaaa	gctccagaga	tgttcgggtc	5820
	cgtttacgag	cacgaggaga	aaaagcccat	ggaggcgctc	gctgaacggg	tgcgagatgc	5880
	cgtggcatc	ggcgccctaca	tcgacggcga	gatcattggg	ctgtcgggtc	tcaaacagga	5940
30	ggacggcccc	aaggacgctc	acaaggcgca	tctgtccggc	gttttcgtgg	agcccgaaaca	6000
	gcgagggccga	ggggctcggc	gtatgctgct	gcgggcggtg	ccggcggtt	tattgctcgt	6060
	gatgatcgtc	cgacagatc	caacgggaat	ctggtggatg	cgcatcttca	tctcggcgc	6120
35	acttaatat	tcgctattct	ggagcttgtt	gtttattctg	gtctaccgcc	tgccgggcgg	6180
	ggtcgcggcg	acggtagggc	ctgtgcagcc	gctgatggtc	gtgttcattc	ctgccgctct	6240
	gctaggtagc	ccgatacgat	tgatggcggt	cctgggggct	atctgcggaa	ctgcggcgct	6300
40	ggcgcgtgtg	gtgttgacac	caaacgcagc	gctagatcct	gtcggcgctc	cagcgggcct	6360
	ggcggggggc	gtttccatgg	cgttcggaac	cgtgctgacc	cgcaagtggc	aacctcccgt	6420
	gcctctgctc	acctttaccg	cctggcaact	ggcggccgga	ggacttctgc	tcgttccagt	6480
	agcttttagtg	tttgatccgc	caatcccgat	gcctacagga	accaatgttc	tcggcctggc	6540
45	gtggctcggc	ctgatcggag	cgggtttaac	ctacttctt	tggttccggg	ggatctcgcg	6600
	actcgaacct	acagttgttt	ccttactggg	ctttctcagc	cgggatggcg	ctaagaagct	6660
	attgccgcgc	atcttcatat	gcggtgtgaa	ataccgcaca	gatgcgtaag	gagaaaatac	6720
	cgcacagggc	gctcttccgc	ttcctcgcct	actgactcgc	tcggctcggg	cgttcggctg	6780
50	cgccgagcgg	tatcagctca	ctcaaaggcg	gtaatacggg	tatccacaga	atcaggggat	6840
	aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	6900
	gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	6960
55	tcaagtcaga	ggtggcgaaa	cccgcagga	ctataaagat	accagggcgtt	ttcccctgga	7020
	agctcccctc	tgccctctcc	tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	7080
	ctcccttcgg	gaagcgtggc	gctttctcaa	tgetcagct	gtaggtatct	cagttcgggtg	7140
60	taggtcgttc	gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	7200
	gccttatccg	gtaactatcg	tcttgagctc	aaccgggtaa	gacacgactt	atcgcactcg	7260
	gcagcagcca	ctggtaacag	gattagcaga	gcgaggtatg	tagggcggtgc	tacagagttc	7320

65

ES 2 346 645 T3

ttgaagtggg ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg 7380
 ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 7440
 5 gctggtagcg gtgggttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatat 7500
 caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt 7560
 taagggattt tggatcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa 7620
 10 aatgaagtt ttaaataaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa 7680
 tgcttaataca gtgaggcacc tatctcagcg atctgtctat ttcgctcacc catagttgcc 7740
 tgactccccg tegtgtagat aactacgata cgggagggct taccatctgg cccagtgct 7800
 gcaatgatac cgcgagacc acgctcaccg gctccagatt tatcagcaat aaaccagcca 7860
 15 gccggaaggg ccgagcgcag aagtggctct gcaactttat ccgctccat ccagtcatt 7920
 aaacaagtgg cagcaacgga ttcgcaaac tgtcaegcct tttgtgcaa aagccgcgcc 7980
 aggtttgcga tccgctgtgc caggcgttag gcgtcatatg aagatttcgg tgatecctga 8040
 gcaggtggcg gaaacattgg atgctgagaa ccatttcatt gttcgtgaag tgttcgatgt 8100
 20 gcacctatcc gaccaaggct ttgaaactatc taccagaagt gtgagccct accggaagga 8160
 ttacatctcg gatgatgact ctgatgaaga ctctgcttgc tatggcgcat tcatcgacca 8220
 agagcttgtc ggaagattg aactcaactc aacatggaac gatctagcct ctatcgaaca 8280
 25 cattgttgtg tgcacacgc accgaggcaa aggagtcgcg cacagtctca tgaatttgc 8340
 gaaaaagtgg gcactaagca gacagctcct tggcatacga ttagagacac aaacgaacaa 8400
 tgtacctgcc tgcaatttgt acgcaaatg tggctttact ctggcgcca ttgacctgtt 8460
 cacgtataaa actagacctc aagtctcgaa cgaaacagcg atgtactggt actggttctc 8520
 30 gggagcacag gatgacgct aacaattcat tcaagccgac accgcttcgc ggcgcgctt 8580
 aattcaggag ttaaacaatca tgagggagc ggtgatcgcc gaagtatcga ctcaactatc 8640
 agaggtagtt ggcgtcatcg agcgcctct cgaaccgacg ttgctggccg tacatttgta 8700
 35 cggctccgca gtggatggcg gcctgaagcc acacagtgat attgatttgc tggttacggg 8760
 gaccgtaagg cttgatgaaa caacgcggcg agctttgatc aacgacctt tggaaacttc 8820
 ggcttcccc ggagagagcg agattctccg cgctgtagaa gtcaccattg ttgtgcacga 8880
 40 cgacatcatt ccgtggcggt atccagctaa gcgcgaactg caatttgag aatggcagcg 8940
 caatgacatt cttgcaggta tcttcagacc agccacgate gacattgatc tggctatctt 9000
 gctgacaaaa gcaagagaac atagcgttgc cttggtaggt ccagcggcgg aggaactctt 9060
 tgatccggtt cctgaacagg atctatttga ggcgctaaat gaaaccttaa cgctatggaa 9120
 45 ctgcgcgcc gactgggctg gcgatgagcg aaatgtagt cttacgttgt cccgcatttg 9180
 gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact gggcaatgga 9240
 gcgcctgcg gccacgtatc agcccgtcat acttgaagct aggcaggctt atcttgaca 9300
 agaagatcgc ttggcctcgc gcgcagatca gttggaagaa tttgttact acgtgaaagg 9360
 50 cgagatcacc aaggtagtgc gcaataatg tctaacaatt cgttcaagcc gacgcgctt 9420
 cgcggcgcgg cttaactcaa gcgttagaga gctggggaag actatgcgcg atctgttgaa 9480
 ggtggttcta agcctcgtac ttgcgatggc atcggggcag gcacttgctg acctgccaat 9540
 55 tgttttagtg gatgaagctc gtcttcccta tgactactcc ccatccaact acgacatttc 9600
 tccaagcaac tacgacaact ccataagcaa ttacgacaat agtccatcaa attacgacaa 9660
 ctctgagagc aactacgata atagttcatc caattacgac aatagtcgca acggaaatcg 9720
 60 taggcttata tatagcggaa atgggtctcg cactttcggc ggctactacg tcattgcaa 9780
 caatgggaca acgaaactct tttccacatc tggcaaaagg atgttctaca ccccaaaagg 9840
 ggggcgcggc gtctatggcg gcaaagatgg gagcttctgc ggggcattgg tegtataaa 9900

65

ES 2 346 645 T3

	tggccaattt	tcgcttgccc	tgacagataa	cggcctgaag	atcatgtatc	taagcaacta	9960
	gcctgctctc	taataaaatg	ttaggagctt	ggctgccatt	ttgggggtga	ggccgttcgc	10020
5	ggccgagggg	cgagcccctt	gggggatgg	gagcccgcg	ttagcgggcc	gggagggttc	10080
	gagaaggggg	ggcaccctcc	ttcggcgtgc	gcggtcacgc	gccagggcgc	agccctggtt	10140
	aaaaacaagg	tttataaata	ttggtttaa	agcaggttaa	aagacaggtt	agcgggtggc	10200
	gaaaaacggg	cggaaaccct	tgcaaatgct	ggattttctg	cctgtggaca	gcccccaaaa	10260
10	tgtcaatagg	tcgccccctc	atctgtcagc	actctgcccc	tcaagtgtca	aggatcgcgc	10320
	ccctcatctg	tcagtagtgc	cgcccccaa	gtgtcaatac	cgcagggcac	ttatccccag	10380
	gcttgtccac	atcatctgtg	ggaaactcgc	gtaaaatcag	gcgttttcgc	cgatttgcga	10440
15	ggctggccag	ctccacgtcg	ccggccgaaa	tcgagcctgc	ccctcatctg	tcaacgccgc	10500
	gccgggtgag	tcggcccttc	aagtgtcaac	gtccgccctt	catctgtcag	tgagggccaa	10560
	gttttcgcg	aggatccac	aacgccggcg	gccggccgcg	gtgtctcgca	cacggcttcg	10620
	acggcgtttc	tggcgcgttt	gcagggccat	agacggccgc	cagcccagcg	gcgagggcaa	10680
20	ccagcccggg	gagcgtcggg	aagggtcagc	atcttgcctg	gttcggatat	tttcgtggag	10740
	ttcccgccac	agaccgggat	tgaaggcgag	atccagcaac	tcgcgccaga	tcctcctgtg	10800
	acggaacttt	ggcgcgtgat	gactggccag	gacgtcggcc	gaaagagcga	caagcagatc	10860
25	acgattttcg	acagcgtcgg	atltgcgac	gaggattttt	cggcgcctgc	ctacgtccgc	10920
	gaccgcgttg	agggatcaag	ccacagcagc	ccactcgacc	ttctagccga	cccagacgag	10980
	ccaagggatc	tttttgggat	gctgctccgt	cgtcaggctt	tcgcagcttt	gggtggttga	11040
	acagaagtca	ttatcgtacg	gaatgccagc	actcccaggg	ggaaccctgt	ggttggcatg	11100
30	cacatacaaa	tggacgaacg	gataaacctt	ttcacgccct	tttaaatatc	cgttattcta	11160
	ataaacgctc	ttttctctta	ggtttaccgg	ccaatatatc	ctgtcaaaca	ctgatagttt	11220
	aaactgaagg	cgggaaacga	caatctgatc	atgagcggag	aattaaggga	gtcacgttat	11280
35	gacccccgcc	gatgacgcgg	gacaagccgt	tttacgtttg	gaactgacag	aaccgcaacg	11340
	attgaaggag	ccactcagcc	ccaatacgcg	aaccgcctct	ccccgcgcgt	tggccgattc	11400
	attaatgcag	ctggcacgac	aggtttcccg	actggaaagc	gggcagtgag	cgcaacgcaa	11460
40	ttaatgtgag	ttagctcact	cattaggcac	cccaggcttt	acactttatg	cttcggctc	11520
	gtatgtgtg	tggaaattgt	agcggataac	aatctcacac	aggaaacagc	tatgaccatg	11580
	attacgcaa	gctatttagg	tgacactata	gaatactcaa	gctatgcac	caacgcgttg	11640
	ggagctctcc	catatcgacc	tcgagcggc	cgctcgacga	attaattcca	atcccacaaa	11700
45	aatctgagct	taacagcaca	gttgcctctc	tcagagcaga	atcgggtatt	caacaccctc	11760
	atatcaacta	ctacgttgtg	tataacggtc	cacatgccgg	tatatacgat	gactggggtt	11820
	gtacaaaggc	ggcaacaaac	ggcgttcccg	gagttgcaca	caagaaattt	gccactatta	11880
50	cagaggcaag	agcagcagct	gacgcgtaca	caacaagtea	gcaaacagac	aggttgaact	11940
	tcatccccaa	aggagaagct	caactcaagc	ccaagagctt	tgctaaggcc	ctaacaagcc	12000
	caccaaaagc	aaaagcccac	tggctcacgc	taggaaccaa	aaggcccagc	agtgatccag	12060
	ccccaaaaga	gatctccttt	gcccgggaga	ttacaatgga	cgatttcctc	tatctttacg	12120
55	atctaggaag	gaagttcgaa	ggtgaagggt	acgacactat	gttcaccact	gataatgaga	12180
	aggttagcct	cttcaatttc	agaaagaatg	ctgaccacac	gatggttaga	gaggcctacg	12240
	cagcaggtct	catcaagacg	atctaccgca	gtaacaatct	ccaggagatc	aaataccttc	12300
60	ccaagaagg	taaagatgca	gtcaaaagat	tcaggactaa	ttgcatcaag	aacacagaga	12360
	aagacatatt	tctcaagatc	agaagtacta	ttccagtatg	gacgattcaa	ggcttgcttc	12420
	ataaaccaag	gcaagtaata	gagattggag	tctctaaaaa	ggtagttcct	actgaatcta	12480

65

ES 2 346 645 T3

	aggccatgca	tggagtctaa	gattcaaate	gaggatctaa	cagaactcgc	cgtgaagact	12540
	ggcgaacagt	tcatacagag	tcttttacga	ctcaatgaca	agaagaaaat	cttcgtaaac	12600
5	atggtggagc	acgacactct	ggtctactcc	aaaaatgtca	aagatacagt	ctcagaagac	12660
	caaagggcta	ttgagacttt	tcaacaaagg	ataatttcgg	gaaacctcct	cggattccat	12720
	tgccagcta	tctgtcactt	catcgaaagg	acagtagaaa	aggaagggtg	ctcctacaaa	12780
	tgccatcatt	gcgataaagg	aaaggctatc	attcaagatc	tctctgccga	cagtgggtccc	12840
10	aaagatggac	ccccacccac	gaggagcatc	gtggaaaaag	aagacgttcc	aaccacgtct	12900
	tcaaagcaag	tggattgatg	tgacatctcc	actgacgtaa	gggatgacgc	acaatcccac	12960
	tatccttcgc	aagacccttc	ctctatataa	ggaagttcat	ttcatttggg	gaggacacgc	13020
15	tcgagacaag	tttgtacaaa	aaagctgaac	gagaaacgta	aatgatata	aatatcaata	13080
	tattaaatta	gattttgcat	aaaaaacaga	ctacataata	ctgtaaaaca	caacatattc	13140
	agtcactatg	aatcaactac	ttagatggta	ttagtacctt	gtagtgcacc	gacagccttc	13200
	caaatgttct	tcgggtgatg	ctgccaaact	agtcgaccga	cagccttcca	aatgttcttc	13260
20	tcaaacggaa	tcgtcgtatc	cagcctactc	gctattgtcc	tcaatgcctt	attaaatcat	13320
	aaaaagaaat	aagaaaaaga	ggtgcgagcc	tcttttttgt	gtgacaaaat	aaaaacatct	13380
	acctattcat	atacgctagt	gtcatagtcc	tgaaaatcat	ctgcatcaag	aacaatttca	13440
25	caactcttat	acttttctct	tacaagtcgt	tcggcttcat	ctggatttcc	agcctctata	13500
	cttactaac	gtgataaagt	ttctgtaatt	tctactgtat	cgacctgcag	actggetgtg	13560
	tataagggag	cctgacattt	atattcccga	gaacatcagg	ttaatggcgt	ttttgatgtc	13620
	attttcggcg	tggctgagat	cagccacttc	ttccccgata	acggagaccg	gcacactggc	13680
30	catatcggtg	gtcatcatgc	gccagcttcc	atccccgata	tgaccaccgc	ggtaaagttc	13740
	acgggagact	ttatctgaca	gcagacgtgc	actggccagg	gggatcacca	tccgtcggcc	13800
	gggcgtgtca	ataatatcac	tctgtacatc	cacaaacaga	cgataacggc	tctctctttt	13860
35	ataggtgtaa	accttaaaact	gcatttcacc	agtccttgtt	ctcgtcagca	aaagagccgt	13920
	tcatttcaat	aaaccgggcg	acctcagcca	tcccttctctg	attttccgct	ttccagcgtt	13980
	cggcacgcag	acgacgggct	tcattctgca	tggttgtgct	taccagaccg	gagatattga	14040
40	catcatatat	gccttgagca	actgatagct	gtcgtgtgca	actgtcactg	taatacgcctg	14100
	cttcatagca	cacctctttt	tgacataact	cgggtagtgc	cgatcaactg	ctcattttctg	14160
	ccaaaagttg	gcccagggct	tcccggatc	aacagggaca	ccaggattta	tttattctgc	14220
	gaagtgatct	tccgtcacag	gtatttatcc	ggcgcaaagt	gcgtcgggtg	atgctgccaa	14280
45	cttagtcgac	tacaggtcac	taataccatc	taagtagttg	attcatagtg	actggatattg	14340
	ttgtgtttta	cagtattatg	tagtctgttt	tttatgcaaa	atctaattta	atatattgat	14400
	atltatatca	ttttacgttt	ctcgttcagc	tttcttctac	aaagtgttct	cgaggaattc	14460
	ggtaccccag	cttggtaagg	aaataattat	tttctttttt	ccttttagta	taaaatagtt	14520
50	aagtgatggt	aattagtatg	attataataa	tatagttggt	ataattgtga	aaaaataatt	14580
	tataaatata	ttgtttacat	aaacaacata	gtaatgtaaa	aaaatatgac	aagtgatgtg	14640
	taagacgaag	aagataaaaag	ttgagagtaa	gtatattatt	tttaatgaat	ttgatcgaac	14700
55	atgtaagatg	atatactagc	attaatattt	gttttaataca	taatagtaat	tctagctgggt	14760
	ttgatgaatt	aaatatcaat	gataaaatac	tatagtaaaa	ataagaataa	ataaattaaa	14820
	ataaatattt	tttatgatta	atagtttatt	atataattaa	atatctatac	cattactaaa	14880
60	tatttttagtt	taaaagttaa	taaatatttt	gttagaaatt	ccaatctgct	tgtaatttat	14940
	caataaacia	aatatataat	aacaagctaa	agtaacaaat	aatatcaaac	taatagaaac	15000
	agtaatctaa	tgtaacaaaa	cataatctaa	tgctaataata	acaaagcgca	agatctatca	15060

65

ES 2 346 645 T3

	ttttatatag	tattatthtc	aatcaacatt	cttattaatt	tctaaataat	acttgtagtt	15120
	ttattaactt	ctaaatggat	tgactattaa	ttaaataaat	tagtcgaaca	tgaataaaca	15180
5	aggtaacatg	atagatcatg	tcatttgtgt	atcattgatc	ttacatttgg	attgattaca	15240
	gttgggaagc	tgggttcgaa	atcgataagc	tgggatccctc	tagaccactt	tgtacaagaa	15300
	agctgaacga	gaaacgtaaa	atgatataaa	tatcaatata	ttaaattaga	ttttgcataa	15360
10	aaaacagact	acataatact	gtaaaacaca	acatatccag	tcactatgaa	tcaactactt	15420
	agatggat	agtgacctgt	agtcgactaa	gttggcagca	tcacccgacg	cactttgcgc	15480
	cgaataaata	cctgtgacgg	aagatcactt	cgcagaataa	ataaatcctg	gtgtccctgt	15540
	tgataaccggg	aagccctggg	ccaacttttg	gcgaaaatga	gacgttgatc	ggatttcaca	15600
15	actcttatac	ttttctctta	caagtcgttc	ggcttcatct	ggattttcag	cctctatact	15660
	tactaaacgt	gataaagttt	ctgtaatttc	tactgtatcg	acctgcagac	tggctgtgta	15720
	taagggagcc	tgacatttat	attccccaga	acatcaggtt	aatggcgttt	ttgatgtcat	15780
20	tttcgcggtg	gctgagatca	gccacttctt	ccccgataac	ggagaccggc	acactggcca	15840
	tatcgggtgt	catcatgccc	cagctttcat	ccccgatatg	caccaaccggg	taaagttcac	15900
	gggagacttt	atctgacagc	agacgtgcac	tggccagggg	gatcaccatc	cgtcgcccgg	15960
	gcgtgtcaat	aatatcactc	tgtacatcca	caaacagacg	ataacggctc	tctcttttat	16020
25	agggtgaaac	cttaaactgc	atttcaccag	tcctgtttct	cgtcagcaaa	agagccgttc	16080
	atttcaataa	accgggagac	ctcagccatc	ccttctgat	tttcgctttt	ccagcgttcg	16140
	gcacgcagac	gacgggcttc	attctgcatg	gttgtgctta	ccagaccgga	gatattgaca	16200
30	tcatatatgc	cttgagcaac	tgatagctgt	cgctgtcaac	tgteactgta	atacgtctgt	16260
	tcatagcaca	cctctttttg	acatacttct	gttcttgatg	cagatgattt	tcaggactat	16320
	gacactagcg	tatatgaata	ggtagatggt	tttattttgt	cacacaaaaa	agaggctcgc	16380
	acctcttttt	cttattttct	tttatgattt	aatacggcat	tgaggacaat	agcagtagg	16440
35	ctggatacga	cgattccggt	tgagaagaac	atctggaagg	ctgtcggctc	actaagttgg	16500
	cagcatcacc	cgaagaacat	ttggaaggct	gtcggtcgac	tacaggtcac	taataccatc	16560
	taagtagttg	atccatagtg	actggatag	ttgtgtttta	cagtattatg	tagtctgttt	16620
40	tttatgcaaa	atctaattta	atatattgat	atztatatca	ttttacgttt	ctcgttcagc	16680
	ttttttgtac	aaacttgtct	agagtcctgc	tttaatgaga	tatgcgagac	gcctatgatc	16740
	gcatgatatt	tgctttcaat	tctgttgtgc	acgttgtaaa	aaacctgagc	atgtgtagct	16800
	cagatcctta	ccgccgggtt	cggttcattc	taatgaatat	atcaccctgt	actatcgtat	16860
45	ttttatgaat	aatattctcc	gttcaattta	ctgattgtac	cctactactt	atagtacaa	16920
	tattaaaatg	aaaacaatat	attgtgctga	ataggtttat	agcgacatct	atgatagagc	16980
	gccacaataa	caaacaattg	cgttttatta	ttacaaatcc	aattttaaaa	aaagcggcag	17040
50	aaccgggtcaa	acctaaaaga	ctgattacat	aaatcttatt	caaatttcaa	aaggccccag	17100
	gggctagtat	ctacgcacaca	ccgagcggcg	aactaataac	gttcaactgaa	gggaactccg	17160
	gttccccgcc	ggcgcgcgatg	ggtgagatc	cttgaagttg	agtattggcc	gtccgctcta	17220
	ccgaaagtta	cgggcaccat	tcaaccgggt	ccagcaecggc	ggccgggtaa	ccgacttgct	17280
55	gccccgagaa	ttatgcagca	ttttttgggt	gtatgtgggc	cccaaatgaa	gtgcaggtea	17340
	aaccttgaca	gtgacgacaa	atcgttgggc	gggtccaggg	cgaattttgc	gacaacatgt	17400
	cgaggctcag	caggacctgc	aggcatgcaa	gctagcttac	tagtcatgca	tattctatag	17460
60	tgteacctaa	atctgc					17476