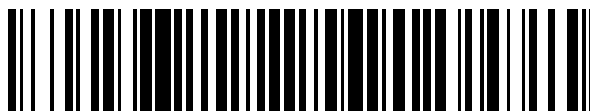


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 753 201**

51 Int. Cl.:

C12N 15/67

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **05.10.2016 PCT/EP2016/073814**

87 Fecha y número de publicación internacional: **13.04.2017 WO17060314**

96 Fecha de presentación y número de la solicitud europea: **05.10.2016 E 16778784 (5)**

97 Fecha y número de publicación de la concesión europea: **25.09.2019 EP 3337902**

54 Título: **Secuencias UTR 3' para la estabilización del ARN**

30 Prioridad:

07.10.2015 WO PCT/EP2015/073180

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

07.04.2020

73 Titular/es:

**BIONTECH RNA PHARMACEUTICALS GMBH
(50.0%)**

An der Goldgrube 12

55131 Mainz, DE y

**TRON - TRANSLATIONALE ONKOLOGIE AN DER
UNIVERSITÄTSMEDIZIN DER JOHANNES
GUTENBERG- UNIVERSITÄT MAINZ (50.0%)**

72 Inventor/es:

ORLANDINI VON NIESSEN, ALEXANDRA;

FESSER, STEPHANIE;

VALLAZZA, BRITTA;

BEISSERT, TIM;

KUHN, ANDREAS;

SAHIN, UGUR y

POLEGANOV, MARCO ALEXANDER

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 753 201 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

Secuencias UTR 3' para la estabilización del ARN

5 El uso del ARN ofrece una alternativa atractiva respecto al ADN para sortear los riesgos potenciales de seguridad relacionados con el uso terapéutico del ADN. *El ARN transcrito In vitro* (IVT-ARN) es de particular interés en los enfoques terapéuticos. Las ventajas de un uso terapéutico del ARN incluyen la expresión transitoria y un carácter no transformante. El ARN no necesita ingresar al núcleo para expresarse y, además, no puede integrarse en el genoma del huésped, de esta manera elimina el riesgo de oncogénesis. Cuando se usa para la vacunación, la inyección del ARN puede inducir
10 respuesta inmunitaria celular y humoral *in vivo*. Sin embargo, el uso del ARN para aplicaciones clínicas es muy restringido, especialmente por la corta vida media del ARN.

Los vectores IVT pueden usarse de manera estandarizada como molde para la transcripción *in vitro*. Tales vectores IVT pueden tener la siguiente estructura: un promotor 5' de la ARN polimerasa que permite la transcripción del ARN, seguido de un gen de interés que está flanqueado 3' y/o 5' por regiones no traducidas (UTR), y un casete de poliadenilo 3' que contiene nucleótidos de A. Antes de la transcripción *in vitro* el plásmido circular se linealiza en dirección 3' del casete de poliadenilo mediante enzimas de restricción de tipo II (la secuencia de reconocimiento corresponde al sitio de escisión). Por lo tanto, el casete de poliadenilo corresponde a la secuencia de poli(A) posterior en el transcrito.

20 Las células dendríticas inmaduras humanas (hiDC) se usan ampliamente para desarrollar y mejorar las inmunoterapias para el tratamiento del cáncer. Las hiDC cargadas con el ARNm transcrito *in vitro* (IVT), que codifica un antígeno tumoral específico (TA), pueden inducir una respuesta antitumoral eficaz. Sin embargo, un requisito previo para una respuesta inmunitaria eficaz, mediante el uso de vacunas contra el cáncer basadas en ARN, es la alta estabilidad y la eficiencia de la traducción del ARN. Ambas pueden mejorarse mediante modificaciones estructurales del 5'-CAP, la cola 3' poli(A), así
25 como también las regiones no traducidas (UTR) en 5' y 3'. Los elementos de secuencia dentro de las UTR afectan la eficiencia de la traducción (principalmente 5'-UTR) y la estabilidad del ARN (principalmente UTR-3').

En trabajos anteriores demostramos que dos copias consecutivas de la UTR-3' de beta-globina humana (ahora llamada 2hBg; además, anteriormente 2βgUTR) contribuyen a una mayor estabilidad de la transcripción y eficiencia de la traducción (Holtkamp (2006) Blood 108:4009-4017). Sin embargo, la presencia de dos copias idénticas de la secuencia UTR-3' de la beta-globina humana en el ADN plasmídico, que finalmente se usa como molde para la transcripción del ARN *in vitro* conlleva el riesgo de recombinación durante su propagación en *E. coli*. De manera similar, cualquier enfoque de clonación, especialmente mediante el uso de amplificación basada en PCR, es muy difícil. Lo mismo es cierto para la amplificación basada en PCR de la región que codifica el ARN de 2hBg en el extremo 3' para usarse como molde para la transcripción *in vitro*, porque aquí se han observado errores de hibridación, que conducen a la omisión de una copia de UTR-3' de la beta-globina humana. Para evitar estos problemas, buscamos identificar secuencias novedosas que tengan un efecto estabilizador en el ARNm transcrito *in vitro* al menos similar a, idealmente incluso mejor que, la secuencia de 2hBg.

40 Xia y otros (International Journal of Molecular Medicine, 2013, 31(5):1081-1086) se refiere al gen "Potenciador de la Separación del Amino Terminal" (AES).

El objeto de la presente invención fue proporcionar ARN con aumento de la estabilidad y/o eficiencia de la traducción y los medios para obtener tal ARN. Debería ser posible obtener mayores grados de expresión mediante el uso de dicho ARN en la terapia.

Este objeto se logra de acuerdo con la invención mediante el contenido de las reivindicaciones.

La presente invención se refiere a la estabilización del ARN, en particular el ARNm, y un aumento en la traducción del ARNm. La presente invención se refiere particularmente a una modificación del ARN, en particular del ARN transcrito *in vitro*, que resulta en un aumento de la estabilidad del transcrito y/o de la eficiencia de la traducción.

De acuerdo con la invención, se demostró que determinadas secuencias en la región no traducida en 3' (UTR) de una molécula de ARN mejoran la estabilidad y la eficiencia de la traducción.

55 Mediante el uso del ARN modificado de acuerdo con la invención en la transfección de células dendríticas (DC), será posible, por ejemplo, aumentar la densidad de complejos de péptido/MHC específicos de antígeno sobre las células transfectadas y su capacidad para estimular y expandir las células T CD4⁺ y CD8⁺ específicas para un antígeno. Por lo tanto, la invención, en una modalidad, se refiere a una estrategia para optimizar las vacunas de ARN para transfectar DC o vacunas de DC transfectadas con ARN mediante el uso de ARN que se ha modificado por las modificaciones del ARN descritas de acuerdo con la invención.

De acuerdo con la invención, la modificación, y de esta manera la estabilización y/o el aumento en la eficiencia de la traducción, del ARN se logra, preferentemente, mediante la modificación genética de los vectores de expresión que, preferentemente, sirven como molde para la transcripción del ARN *in vitro*. Estos vectores de expresión permiten la
65

transcripción del ARN con una región no traducida en 3' descrita de acuerdo con la invención, y, preferentemente, entre la secuencia codificante de un péptido o proteína (marco abierto de lectura) y la secuencia de poli(A).

Estos vectores pueden permitir, además, la transcripción del ARN con una secuencia de poli(A) que, preferentemente, tiene un extremo abierto en dicho ARN, es decir, no hay nucleótidos distintos de los nucleótidos A que flanquean dicha secuencia de poli(A) en su extremo 3'. Puede lograrse una secuencia de poli(A) de extremo abierto en el ARN mediante la introducción de un sitio de escisión de restricción de tipo IIS en un vector de expresión que permite transcribir el ARN bajo el control de un promotor de la ARN polimerasa 5' y que contiene un casete de poliadenilo, en donde la secuencia de reconocimiento se ubica en dirección 3' del casete de poliadenilo, mientras que el sitio de escisión se ubica en dirección 5' y, por lo tanto, dentro del casete de poliadenilo. La escisión por restricción en el sitio de escisión por restricción de tipo IIS permite que un plásmido se linealice dentro del casete de poliadenilo. El plásmido linealizado puede usarse después como molde para la transcripción *in vitro*, el transcrito resultante termina en una secuencia de poli(A) expuesta. Además, una interrupción opcional del casete de poliadenilo 3' por una secuencia de nucleótidos aleatoria, con una distribución igual de los 4 nucleótidos (enlazador), aumenta la estabilidad del casete de poliadenilo 3' en *E. coli*.

Resumen de la invención

En un aspecto, la invención se refiere a una molécula de ácido nucleico como se definió en la reivindicación 1 adjunta.

En una modalidad, la secuencia de ácido nucleico (c-8) comprende una combinación de (c-4) y (c-6). En una modalidad, (c-4) se ubica 5' respecto a (c-6). En una modalidad, la combinación de (c-4) y (c-6) comprende la secuencia de ácido nucleico de la SEQ ID NO: 174, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico.

En una modalidad, la molécula de ácido nucleico de la invención comprende además (d) una secuencia de ácido nucleico que, cuando se transcribe bajo el control del promotor (a), codifica para una secuencia de ácido nucleico que es una secuencia de poliadenilo que opcionalmente comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A. En una modalidad, dicha secuencia de poliadenilo comprende al menos 20 nucleótidos A, preferentemente al menos 40, al menos 80, al menos 100 o al menos 120 nucleótidos A, preferentemente nucleótidos A consecutivos. En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A es una secuencia, preferentemente, una secuencia arbitraria, de 2 o más nucleótidos consecutivos, en donde el primer y el último nucleótido de dicha secuencia de 2 o más nucleótidos consecutivos es un nucleótido distinto de un nucleótido A. En una modalidad, dicha secuencia de ácido nucleico (d) es una secuencia de ácido nucleico que, cuando se transcribe bajo el control del promotor (a), codifica para una secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A y que exhiben una mayor estabilidad tras la propagación de dicha molécula de ácido nucleico en *Escherichia coli* en comparación con una molécula de ácido nucleico que comprende en lugar de dicha secuencia de ácido nucleico (d) una secuencia de ácido nucleico (d)' que, cuando se transcribe bajo el control del promotor (a), codifica para una secuencia de poliadenilo de la misma longitud que dicha secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de nucleótidos A. En una modalidad, dicha secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A, comprende al menos 80 nucleótidos, preferentemente, al menos 90 o 100 nucleótidos. En una modalidad, dicha secuencia de ácido nucleico, que es una secuencia de poliadenilo, que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de nucleótidos A comprende al menos 90 nucleótidos, preferentemente al menos 100 nucleótidos, preferentemente al menos 110 nucleótidos. En una modalidad, dicha secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A comprende aproximadamente 120 nucleótidos. En modalidades particulares, dicha secuencia de ácido nucleico que es una secuencia de poliadenilo, que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de nucleótidos A, comprende hasta 200, preferentemente hasta 150, y, en particular, hasta 130 nucleótidos. En una modalidad, al menos 90 %, preferentemente al menos 92 %, preferentemente al menos 95 %, 97 % o 98 % de los nucleótidos de dicha secuencia de ácido nucleico que es una secuencia de poliadenilo que opcionalmente comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A son nucleótidos A en dicha secuencia de poliadenilo (sin incluir los nucleótidos A en dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A).

En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A se ubica dentro de una región desde la posición 21 a la posición 80, preferentemente desde la posición 21 a la posición 60, con mayor preferencia desde la posición 31 a la posición 50 de dicha secuencia de poliadenilo.

En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A está precedida por al menos 20 residuos A, preferentemente al menos 30, 40 o 50 residuos A en dicha

secuencia de poliadenilo. En modalidades particulares, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de nucleótidos A está precedida por hasta 80 residuos A, preferentemente hasta 70 o 60 residuos A en dicha secuencia de poliadenilo.

5 En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A es seguida por al menos 20 residuos A, preferentemente al menos 30, 40, 50, 60 o 70 residuos A en dicha secuencia de poliadenilo. En modalidades particulares, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A es seguida por hasta 100 residuos A, preferentemente hasta 80 residuos A en dicha secuencia de poliadenilo.

10 En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A está precedida por 20 a 50, preferentemente 30 a 40 residuos A en dicha secuencia de poliadenilo y es seguida por 30 a 80, preferentemente 40 a 70 residuos A en dicha secuencia de poliadenilo.

15 En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A tiene una longitud de al menos 3, al menos 4, al menos 5, al menos 6, al menos 8, preferentemente al menos 10, con mayor preferencia al menos 15 nucleótidos.

20 En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de nucleótidos A tiene una longitud de no más de 50, preferentemente no más de 30, con mayor preferencia no más de 20 nucleótidos.

25 En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A no comprende más de 3, preferentemente no más de 2, preferentemente no hay residuos A consecutivos.

En una modalidad, las secuencias de ácido nucleico (b), (c) y (d) bajo el control del promotor (a) pueden transcribirse para dar un transcrito común. En una modalidad, las secuencias de ácido nucleico transcritas a partir de las secuencias de ácido nucleico (c) y opcionalmente (d) son activas para aumentar la eficiencia de traducción y/o la estabilidad de la secuencia de ácido nucleico transcrita a partir de la secuencia de ácido nucleico transcribible (b).

30 En una modalidad, en el transcrito dicha secuencia de ácido nucleico, que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A, se ubica en el extremo 3'.

35 En una modalidad, la molécula de ácido nucleico de la invención es una molécula de ADN. En una modalidad, dicha molécula de ácido nucleico es un vector de expresión o plásmido tal como un vector IVT.

En una modalidad, la molécula de ácido nucleico de la invención es una molécula circular cerrada o una molécula lineal.

40 En una modalidad, la secuencia de ácido nucleico transcribible comprende una secuencia de ácido nucleico que codifica para un péptido o proteína y la secuencia de ácido nucleico para introducir una secuencia de ácido nucleico transcribible es un sitio de clonación múltiple.

45 En una modalidad, la molécula de ácido nucleico de la invención comprende, además, uno o más miembros seleccionados del grupo que consiste en: (i) un gen reportero; (ii) un marcador de selección; y (iii) un origen de replicación.

En una modalidad, la molécula de ácido nucleico de la invención es adecuada, en particular después de la linealización, para la transcripción del ARN *in vitro*, en particular el ARNm.

50 Antes de la transcripción *in vitro*, los vectores circulares de IVT generalmente se linealizan en dirección 3' del casete de poliadenilo mediante enzimas de restricción de tipo II (la secuencia de reconocimiento corresponde al sitio de escisión). Por lo tanto, el casete de poliadenilo corresponde a la secuencia de poli(A) posterior en el transcrito. Como un resultado de este procedimiento, algunos nucleótidos permanecen como parte del sitio de escisión enzimática después de la linealización y extienden o enmascaran la secuencia de poli(A) en el extremo 3'. Sin embargo, se descubrió que el ARN que tiene una secuencia poli(A) de extremo abierto se traduce de manera más eficiente que el ARN que tiene una secuencia poli(A) con un terminal enmascarado.

55 En consecuencia, las moléculas de ácido nucleico de la invención cuando se usan como vectores de expresión permiten, preferentemente, la transcripción del ARN con una secuencia de poli(A) que, preferentemente, tiene un extremo abierto en dicho ARN, es decir, no hay nucleótidos distintos de los nucleótidos A que flanquean dicha secuencia de poli(A) en su extremo 3'. Una secuencia de poli(A) de extremo abierto en el ARN puede lograrse mediante la introducción de un sitio de escisión de restricción de tipo IIS en un vector de expresión que permite transcribir el ARN bajo el control de un promotor 5' de la ARN polimerasa y que contiene un casete de poliadenilo, en donde la secuencia de reconocimiento se ubica en dirección 3' del casete de poliadenilo, mientras que el sitio de escisión se ubica en dirección 5' y, por lo tanto, dentro del casete de poliadenilo. La escisión por restricción en el sitio de escisión por restricción de tipo IIS permite que

un plásmido se linealice dentro del casete de poliadenilo. El plásmido linealizado puede usarse después como molde para la transcripción *in vitro*, el transcrito resultante termina en una secuencia de poli(A) expuesta.

En consecuencia, en una modalidad, se prefiere que la molécula de ácido nucleico de la invención pueda escindir, preferentemente enzimáticamente o de otra manera bioquímica, dentro de la secuencia de ácido nucleico (d) de tal manera que dicha escisión resulte en una molécula de ácido nucleico que comprende, en la dirección de transcripción 5' → 3', el promotor (a), las secuencias de ácido nucleico (b) y (c), y al menos una parte de la secuencia de ácido nucleico (d), en donde la al menos una parte de la secuencia de ácido nucleico (d), cuando se transcribe bajo el control del promotor (a), codifica para dicha secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A y en donde en el transcrito el nucleótido terminal 3' es un nucleótido A de dicha secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A.

Preferentemente, después de la escisión, la molécula de ácido nucleico, en el extremo de la cadena que sirve como molde para la secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A, tiene un nucleótido T que es parte de la secuencia de ácido nucleico que sirve como molde para la secuencia de ácido nucleico que es una secuencia de poliadenilo que opcionalmente comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A.

La molécula de ácido nucleico de la invención es, preferentemente, una molécula circular cerrada antes de la escisión y una molécula lineal después de la escisión.

Preferentemente, la escisión se realiza con la ayuda de un sitio de escisión de restricción que es, preferentemente, un sitio de escisión de restricción para una endonucleasa de restricción de tipo IIS.

En una modalidad, la secuencia de reconocimiento para la endonucleasa de restricción de tipo IIS se ubica en 5-26 pares de bases, preferentemente 24-26 pares de bases, en dirección 3' del extremo 3' de la secuencia de ácido nucleico (d).

En una modalidad, una molécula de ácido nucleico de acuerdo con la invención está en una conformación circular cerrada y, preferentemente, adecuada para la transcripción del ARN *in vitro*, en particular el ARNm, en particular después de la linealización.

En aspectos adicionales, la invención se refiere al ARN que se obtiene mediante la transcripción, preferentemente, la transcripción *in vitro*, como se definió en la reivindicación 8 adjunta.

Por lo tanto, la invención en un aspecto se refiere al ARN como se definió en la reivindicación 9 adjunta.

En una modalidad, el ARN comprende además (d) una secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A. En una modalidad, dicha secuencia de ácido nucleico (d) se ubica en el extremo 3' de dicho ARN.

En una modalidad, las secuencias de ácido nucleico (c) y, opcionalmente, (d) son activas para aumentar la eficiencia de traducción y/o la estabilidad de la secuencia de ácido nucleico que codifica para un péptido o proteína.

En una modalidad, el ARN comprende además (e) un casquete 5'.

Las modalidades de la región no traducida en 3' y la secuencia de ácido nucleico, que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A, son como se describió anteriormente para las moléculas de ácido nucleico de la invención.

En un aspecto adicional, la invención se refiere a un método para obtener ARN como se definió en la reivindicación 12 adjunta.

En un aspecto adicional, la invención se refiere a un método para obtener un péptido o proteína como se definió en la reivindicación 13 adjunta.

En una modalidad, el método para obtener el ARN o el método para obtener un péptido o proteína comprende además, antes de la transcripción de la molécula de ácido nucleico, la escisión de la molécula de ácido nucleico.

En un aspecto adicional, la invención se refiere a un método para obtener ARN como se definió en la reivindicación 15 adjunta.

En una modalidad, el método comprende además el acoplamiento de una secuencia de ácido nucleico (c) que, cuando se transcribe, codifica para una secuencia de ácido nucleico que es una secuencia de poliadenilo que opcionalmente comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de nucleótidos A, en el extremo 3' de la secuencia de ácido nucleico (b).

En una modalidad, las secuencias de ácido nucleico (a), (b), y (c) pueden transcribirse para proporcionar un transcrito común en el que las secuencias de ácido nucleico transcritas a partir de las secuencias de ácido nucleico (b) y, opcionalmente, (c) son activas para aumentar la eficiencia de la traducción y/o la estabilidad de la secuencia de ácido nucleico transcrita a partir de la secuencia de ácido nucleico transcribible (a).

Las modalidades de la región no traducida en 3' y la secuencia de ácido nucleico, que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A, son como se describió anteriormente para las moléculas de ácido nucleico de la invención.

En un aspecto adicional, la invención se refiere a un método para obtener un péptido o proteína como se definió en la reivindicación 17 adjunta.

En una modalidad de cualquiera de los métodos de la invención, la transcripción se realiza *in vitro*.

En una modalidad, el método para obtener el ARN o el método para obtener un péptido o proteína comprende además, antes de la transcripción de la molécula de ácido nucleico, la escisión de la molécula de ácido nucleico.

En una modalidad, la escisión está dentro de la secuencia de ácido nucleico que, cuando se transcribe, codifica para una secuencia de ácido nucleico que es una secuencia de poliadenilo que opcionalmente comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A, de manera que la transcripción del ácido nucleico que se obtiene de esta manera genera un transcrito que tiene en su extremo terminal 3' dicha secuencia de ácido nucleico que es una secuencia de poliadenilo que opcionalmente comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de nucleótidos A, en donde el nucleótido terminal 3' de dicho transcrito es un nucleótido A de la secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A.

En todos los aspectos de los métodos de acuerdo con la invención, la escisión se realiza, preferentemente, con la ayuda de un sitio de escisión de restricción que es, preferentemente, un sitio de escisión de restricción para una endonucleasa de restricción de tipo IIS.

En una modalidad, la secuencia de reconocimiento para la endonucleasa de restricción de tipo IIS es de 5-26 pares de bases, preferentemente, 24-26 pares de bases, en dirección 3' del extremo 3' de la secuencia de ácido nucleico que, cuando se transcribe, codifica para una secuencia de ácido nucleico que es una secuencia de poliadenilo que comprende opcionalmente dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A.

La invención se refiere, además, al ARN que se obtiene mediante los métodos de acuerdo con la invención para obtener ARN.

La invención puede utilizarse, por ejemplo, para aumentar la expresión de proteínas recombinantes en la transcripción y expresión celular. Más específicamente, es posible, cuando se producen proteínas recombinantes, usar los vectores de expresión de la invención para la transcripción de ácidos nucleicos recombinantes y la expresión de proteínas recombinantes en sistemas basados en células. Esto incluye, por ejemplo, la preparación de anticuerpos recombinantes, hormonas, citocinas, enzimas y similares. Esto permite, entre otros, disminuir los costos de producción.

Además, es posible usar las moléculas de ácido nucleico de la invención para las aplicaciones de terapia génica. En consecuencia, una molécula de ácido nucleico de la invención puede ser un vector de terapia génica y usarse para la expresión de un transgén. Para este fin, puede usarse cualquier sistema de vectores basado en ácido nucleico (ADN/ARN) (por ejemplo, plásmidos, adenovirus, vectores de poxvirus, vectores de virus de la influenza, vectores de alfavirus y similares). Las células pueden transfectarse *in vitro* con estos vectores, por ejemplo, en linfocitos o células dendríticas, o de otra manera *in vivo* mediante la administración directa.

El ARN de la invención (por ejemplo, que se obtiene mediante el uso de una molécula de ácido nucleico descrita en la presente descripción como un molde de transcripción) puede usarse, por ejemplo, para la expresión transitoria de genes, con posibles campos de aplicación que son vacunas basadas en ARN que se transfectan *in vitro* en células o se

administran directamente *en vivo*, expresión transitoria *in vitro* de proteínas recombinantes funcionales, por ejemplo, para iniciar los procesos de diferenciación en las células o para estudiar las funciones de las proteínas, y la expresión transitoria *in vivo*, de proteínas recombinantes funcionales tales como la eritropoyetina, hormonas, inhibidores de la coagulación, etcétera, en particular como productos farmacéuticos.

El ARN de la invención puede usarse en particular para transfectar células presentadoras de antígeno y, por lo tanto, como una herramienta para suministrar el antígeno que se presentará y para cargar las células presentadoras de antígeno, con dicho antígeno para ser presentado que corresponde al péptido o proteína expresado a partir de dicho ARN o derivado del mismo, en particular por medio de procesamiento intracelular tal como escisión, es decir, el antígeno a presentarse es, por ejemplo, un fragmento del péptido o de la proteína que se expresa a partir del ARN. Tales células presentadoras de antígeno pueden usarse para estimular las células T, en particular las células T CD4⁺ y/o CD8⁺.

En consecuencia, en un aspecto adicional, la invención se refiere a un ARN de la invención para usar en un método para transfectar una célula huésped. En una modalidad, la célula huésped es una célula presentadora de antígeno, en particular una célula dendrítica, un monocito o un macrófago.

En un aspecto adicional, la invención se refiere a un ARN de la invención para usar en la terapia, en particular para usar en un método de vacunación.

En un aspecto adicional, la invención se refiere a una composición farmacéutica tal como una composición de vacuna que comprende el ARN de la invención.

En un aspecto adicional, la invención se refiere al ARN de la invención para los usos descritos en la presente descripción.

Descripción detallada de la invención

Aunque la presente invención se describe en detalle más abajo, debe entenderse que esta invención no se limita a las metodologías, protocolos y reactivos particulares descritos en la presente descripción, ya que estos pueden variar. Debe entenderse, también, que la terminología utilizada en la presente descripción es para el propósito de describir modalidades particulares solamente, y no pretende limitar el alcance de la presente invención la cual se limitará solamente por las reivindicaciones anexas. A menos que se defina de otra manera, todos los términos técnicos y científicos que se usan en la presente descripción tienen el mismo significado que el que entiende comúnmente un experto en la técnica.

A continuación, se describirán los elementos de la presente invención. Estos elementos se enumeran con modalidades específicas, sin embargo, debe entenderse que pueden combinarse de cualquier manera y en cualquier número para crear modalidades adicionales. Los ejemplos descritos de diversas maneras y las modalidades preferidas no deben interpretarse como que limitan la presente invención a solamente las modalidades descritas explícitamente. Debe entenderse que esta descripción sustenta y abarca modalidades que combinan las modalidades descritas explícitamente con cualquier cantidad de elementos descritos y/o preferidos. Además, cualquiera de las permutaciones y combinaciones de todos los elementos descritos en esta solicitud deben considerarse descritas por la descripción de la presente solicitud a menos que el contexto lo indique de cualquier otra manera. Por ejemplo, si en una modalidad preferida una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de nucleótidos A está precedida por al menos 20 residuos A en dicha secuencia de poliadenilo y si en otra modalidad preferida una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A es seguida por al menos 20 residuos A en dicha secuencia de poliadenilo, es una modalidad preferida contemplada que una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos a los nucleótidos A sea precedida y seguida por al menos 20 residuos A en dicha secuencia de poliadenilo.

Preferentemente, los términos que se usan en la presente descripción se definen como se describe en "A multilingual glossary of biotechnological terms: (IUPAC Recommendations)", H.G.W. Leuenberger, B. Nagel, y H. Kölbl, Eds., Helvetica Chimica Acta, CH-4010 Basel, Suiza, (1995).

La práctica de la presente invención empleará, a menos que se indique de otra manera, métodos convencionales de química, bioquímica, biología celular, inmunología y técnicas de ADN recombinante que se explican en la literatura del campo (consultar, por ejemplo, Molecular Cloning: A Laboratory Manual, 2da Edición, J. Sambrook y otros eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989).

A lo largo de esta descripción y de las reivindicaciones que siguen, a menos que el contexto lo requiera de cualquier otra manera, la palabra "comprender", y variaciones tales como "comprende" y "que comprende", se entenderán que implican la inclusión de un miembro, entero o etapa o grupo de miembros, enteros o etapas indicadas pero no la exclusión de cualquier otro miembro, entero o etapa o grupo de miembros, enteros o etapas. Los términos "un" y "una" y "el/la" y referentes similares usados en el contexto para describir la invención (especialmente en el contexto de las reivindicaciones) se deben interpretar para cubrir tanto el singular como el plural, a menos que se indique de cualquier otra forma en la presente descripción o claramente se contradiga por el contexto. La enumeración de los intervalos de valores en la presente descripción pretende servir simplemente como un método abreviado para referirse individualmente a cada valor separado que se encuentra dentro del intervalo. A menos que se indique de otra manera en la presente

descripción, cada valor individual se incorpora en la descripción como si se enumerara individualmente en la presente descripción. Todos los métodos que se describen en la presente descripción pueden llevarse a cabo en cualquier orden adecuado a menos que se indique de otra manera en la presente descripción o que el contexto lo contradiga claramente de otra manera. El uso de cualquiera y de todos los ejemplos, o lenguaje ilustrativo (por ejemplo, "tal como") que se proporciona en la presente descripción simplemente tiene el propósito de ilustrar mejor la invención y no representa una limitación en el alcance de la invención que se reivindica de otra manera. El lenguaje en la descripción no deberá interpretarse como indicación de cualquier elemento no reivindicado esencial para la práctica de la invención.

Se citan varios documentos a lo largo del texto de esta descripción. Nada debe interpretarse en la presente descripción como una admisión de que la invención no tiene derecho a preceder dicha descripción en virtud de una invención anterior.

La presente invención describe moléculas de ácido nucleico tales como plásmidos de ADN útiles como vectores de expresión de ARN que comprenden secuencias de ácido nucleico que codifican regiones no traducidas en 3' (UTR) modificadas en el ARN que tienen un efecto estabilizador sobre el ARN y/o aumentan la eficiencia de la traducción del ARN.

El término "secuencia de ácido nucleico que, cuando se transcribe, codifica para una región no traducida en 3' en el transcrito" se refiere a una secuencia de ácido nucleico que contiene una cadena molde que codifica para dicha región no traducida en 3'. Preferentemente, dicha secuencia de ácido nucleico comprende una cadena codificante que comprende la misma secuencia de ácido nucleico que dicha región no traducida en 3' del transcrito de ARN producido (aunque con timina sustituida por uracilo). Por lo tanto, de acuerdo con la invención, una "secuencia de ácido nucleico que, cuando se transcribe, codifica para una región no traducida en 3' en el transcrito", en una modalidad, comprende una cadena codificante que comprende una región no traducida en 3' como se especificó en la presente descripción (aunque con timina sustituida por uracilo).

El término "FCGRT" se refiere al fragmento Fc de IgG, receptor, transportador, alfa e incluye el gen FCGRT. Este gen codifica un receptor que se une a la región Fc de la inmunoglobulina G monomérica. La proteína codificada transfiere los anticuerpos de inmunoglobulina G de la madre al feto a través de la placenta. Esta proteína además se une a la inmunoglobulina G para proteger el anticuerpo de la degradación.

El término "secuencia de ácido nucleico de la región no traducida en 3' de FCGRT, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico" se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 1 a 50 del listado de secuencias o un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico. En una modalidad, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 1 a 50. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente, que consiste en la secuencia de ácido nucleico de la SEQ ID NO: 27 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de la SEQ ID NO: 27.

El término "LSP1" se refiere a la Proteína 1 Específica de Linfocitos e incluye el gen LSP1. Este gen codifica una proteína de unión a actina F intracelular. La proteína se expresa en linfocitos, neutrófilos, macrófagos y en el endotelio, y puede regular la motilidad de los neutrófilos, la adhesión a las proteínas de la matriz de fibrinógeno, y la migración transendotelial.

El término "secuencia de ácido nucleico de la región no traducida en 3' de LSP1, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico" se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 51 a 72 del listado de secuencias o un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico. En una modalidad, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 51 a 72. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente, que consiste en la secuencia de ácido nucleico de la SEQ ID NO: 52 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de la SEQ ID NO: 52.

El término "CCL22" se refiere al Ligando de Quimiocina 22 (motivo C-C) e incluye el gen CCL22. El producto de este gen se une al receptor de quimiocinas CCR4. Las quimiocinas pueden desempeñar un papel en el tráfico de los linfocitos T activados a los sitios inflamatorios y otros aspectos de la fisiología de los linfocitos T activados.

El término "secuencia de ácido nucleico de la región no traducida en 3' de CCL22, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico" se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 73 a 85 del listado de secuencias o un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido

- nucleico. En una modalidad, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 73 a 85. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente, que consiste en la secuencia de ácido nucleico de la SEQ ID NO: 79 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de la SEQ ID NO: 79.
- El término "AES" se refiere al Potenciador de la Separación del Amino Terminal e incluye el gen AES. La proteína codificada por este gen pertenece a la familia de proteínas groucho/TLE, puede funcionar como un homooligómero o como un heterooligómero con otros miembros de la familia para reprimir de manera dominante la expresión de otros genes miembros de la familia.
- El término "secuencia de ácido nucleico de la región no traducida en 3' de AES, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico" se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 86 a 89 del listado de secuencias o un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico. En una modalidad, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 86 a 89. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente, que consiste en la secuencia de ácido nucleico de la SEQ ID NO: 86 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de la SEQ ID NO: 86. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en la secuencia de ácido nucleico de las posiciones 1 a 68, posiciones 1 a 102, posiciones 35 a 102, posiciones 35 a 136, o posiciones 68 a 136 de la SEQ ID NO: 86 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de las posiciones 1 a 68, posiciones 1 a 102, posiciones 35 a 102, posiciones 35 a 136, o posiciones 68 a 136 de la SEQ ID NO: 86.
- El término "PLD3" se refiere al Miembro 3 de la Familia de Fosfolipasa D e incluye el gen PLD3. Este gen codifica un miembro de la familia de enzimas fosfolipasa D (PLD) que catalizan la hidrólisis de los fosfolípidos de membrana. La proteína codificada es una proteína de membrana de tipo II de un solo paso y contiene dos dominios de fosfodiesterasa PLD. Esta proteína influye en el procesamiento de la proteína precursora beta amiloide. Las mutaciones en este gen se asocian con el riesgo de enfermedad de Alzheimer.
- El término "secuencia de ácido nucleico de la región no traducida en 3' de PLD3, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico" se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 90 a 104 del listado de secuencias o un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico. En una modalidad, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 90 a 104. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente, que consiste en la secuencia de ácido nucleico de la SEQ ID NO: 96 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de la SEQ ID NO: 96.
- El término "MT_RNR1" se refiere al ARN 12S codificado mitocondrialmente e incluye el gen MT_RNR1. Este gen de ARN pertenece a la clase Mt_rRNA. Las enfermedades asociadas con MT-RNR1 incluyen cardiomiopatía restrictiva y neuropatía auditiva. Entre sus vías relacionadas se encuentran la biogénesis de los ribosomas en eucariotas y la fidelidad de la traducción de CFTR (mutaciones de clase I).
- El término "secuencia de ácido nucleico de la región no traducida en 3' de MT_RNR1, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico" se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 105 a 121 del listado de secuencias o un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico. En una modalidad, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 105 a 121. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente, que consiste en la secuencia de ácido nucleico de la SEQ ID NO: 115 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de la SEQ ID NO: 115. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende,

preferentemente que consiste en la secuencia de ácido nucleico de las posiciones 1 a 71, posiciones 1 a 107, posiciones 37 a 107, posiciones 37 a 142, o posiciones 71 a 142 de la SEQ ID NO: 115 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de las posiciones 1 a 71, posiciones 1 a 107, posiciones 37 a 107, posiciones 37 a 142, o posiciones 71 a 142 de la SEQ ID NO: 115.

El término "HLA-DRB4" se refiere al complejo principal de histocompatibilidad, clase II, DR Beta 4 e incluye el gen HLA-DRB4. HLA-DRB4 pertenece a los parálogos de la cadena beta de HLA clase II. Esta molécula de clase II es un heterodímero que consiste en una cadena alfa (DRA) y una cadena beta (DRB), ambas ancladas en la membrana. Desempeña un papel central en el sistema inmunitario mediante la presentación de los péptidos derivados de las proteínas extracelulares. Las moléculas de clase II se expresan en células presentadoras de antígeno (APC: linfocitos B, células dendríticas, macrófagos).

El término "secuencia de ácido nucleico de la región no traducida en 3' de HLA-DRB4, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico" se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 122 a 143 del listado de secuencias o un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico. En una modalidad, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 122 a 143. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente, que consiste en la secuencia de ácido nucleico de la SEQ ID NO: 126 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de la SEQ ID NO: 126.

El término "cualquier combinación de dos o más de las secuencias, fragmentos y/o variantes de ácido nucleico" con respecto a las secuencias de ácido nucleico de las regiones no traducidas en 3' de determinados genes, fragmentos de estos, o variantes de dichas secuencias o fragmentos de ácido nucleico significa que 2 o más, 3 o más, o 4 o más y preferentemente hasta 6 o hasta 5 de dichas secuencias, fragmentos y/o variantes de ácido nucleico están alineados cabeza a cola, opcionalmente separados por enlazadores. En una modalidad, la combinación de dos o más secuencias, fragmentos y/o variantes de ácido nucleico comprende dos o más secuencias, fragmentos y/o variantes de ácido nucleico diferentes y/o dos o más idénticas. En una modalidad, la combinación de dos o más secuencias, fragmentos y/o variantes de ácido nucleico comprende dos o más secuencias, fragmentos y/o variantes de ácido nucleico diferentes de la región no traducida en 3' de los mismos y/o diferentes genes.

En una modalidad, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 144 a 220, preferentemente las SEQ ID NOs: 174 y 208 a 220. En una modalidad, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente que consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 144 a 220, preferentemente las SEQ ID NOs: 174 y 208 a 220 o un fragmento de esta, o una variante de dicha secuencia de ácido nucleico o fragmento. En una modalidad particularmente preferida, el término se refiere a una secuencia de ácido nucleico que comprende, preferentemente, que consiste en la secuencia de ácido nucleico de la SEQ ID NO: 174 o que comprende, preferentemente que consiste en una secuencia de ácido nucleico que es al menos 90 %, preferentemente al menos 95 %, con mayor preferencia al menos 98 % idéntica a la secuencia de ácido nucleico de la SEQ ID NO: 174.

El término "enlazador" de acuerdo con la invención se refiere a una secuencia de ácido nucleico añadida entre dos secuencias de ácido nucleico para conectar dichas dos secuencias de ácido nucleico. No existe una limitación particular con respecto a la secuencia del enlazador.

De acuerdo con la invención, una molécula de ácido nucleico o una secuencia de ácido nucleico se refiere a un ácido nucleico que es, preferentemente, ácido desoxirribonucleico (ADN) o ácido ribonucleico (ARN). De acuerdo con la invención, los ácidos nucleicos comprenden ADN genómico, ADNc, ARNm, moléculas producidas recombinantemente y sintetizadas químicamente. De acuerdo con la invención, un ácido nucleico puede estar en la forma de una molécula monocatenaria o bicatenaria, y lineal o circular cerrada covalentemente.

En el contexto de la presente invención, el término "ARN" se refiere a una molécula que comprende residuos de ribonucleótidos y preferentemente está compuesta totalmente o esencialmente de residuos de ribonucleótidos. El término "ribonucleótido" se refiere a un nucleótido con un grupo hidroxilo en la posición 2' de un grupo β -D-ribofuranosa. El término "ARN" comprende ARN bicatenario, ARN monocatenario, ARN aislado tal como ARN purificado parcialmente o completamente, ARN esencialmente puro, ARN sintético, y ARN producido de manera recombinante tal como ARN modificado que difiere del ARN de origen natural por la adición, delección, sustitución y/o alteración de uno o más nucleótidos. Dichas alteraciones pueden incluir la adición de material no nucleotídico, tal como al(los) extremo(s) de un ARN o internamente, por ejemplo, a uno o más nucleótidos del ARN. Los nucleótidos en las moléculas de ARN pueden comprender también nucleótidos no estándar, tales como nucleótidos que no son de origen natural o nucleótidos o

desoxinucleótidos sintetizados químicamente. Estos ARN alterados pueden referirse como análogos, o particularmente análogos de los ARN de origen natural. De acuerdo con la invención, el ARN incluye ARNm.

El término "ARNm" significa "ARN mensajero" y se refiere a un transcrito que se genera mediante el uso de un molde de ADN y codifica para un péptido o proteína. Típicamente, el ARNm comprende una 5'-UTR, una región codificante de proteínas, una UTR-3' y una secuencia de poli(A). El ARNm puede generarse mediante la transcripción *in vitro* a partir de un molde de ADN. La metodología de la transcripción *in vitro* se conoce por el experto. Por ejemplo, existe una variedad de kits de transcripción *in vitro* disponibles comercialmente. De acuerdo con la invención, el ARNm puede modificarse mediante modificaciones de estabilización y protección adicionales, además de las modificaciones de acuerdo con la invención.

En una modalidad de la presente invención, el ARN es ARN autorreplicante, tal como ARN monocatenario autorreplicante. En una modalidad, el ARN autorreplicante es ARN monocatenario de sentido positivo. En una modalidad, el ARN autorreplicante es ARN viral o ARN derivado de ARN viral. En una modalidad, el ARN autorreplicante es un ARN genómico alfaviral o se deriva de un ARN genómico alfaviral. En una modalidad, el ARN autorreplicante es un vector de expresión génica viral. En una modalidad, el virus es el virus del bosque Semliki. En una modalidad, el ARN autorreplicante contiene uno o más transgenes. En una modalidad, si el ARN es ARN viral o derivado de ARN viral, los transgenes pueden reemplazar parcialmente o completamente las secuencias virales tal como las secuencias virales que codifican proteínas estructurales. En una modalidad, el ARN autorreplicante es el ARN transcrito *in vitro*.

El término "casquete 5'" se refiere a una estructura de casquete que se encuentra en el extremo 5' de una molécula de ARNm y consiste, generalmente, en un nucleótido de guanosina que se conecta al ARNm a través de un enlace trifosfato inusual 5' a 5'. En una modalidad, esta guanosina se metila en la posición 7. El término "casquete 5' convencional" se refiere a un ARN de origen natural con casquete en 5', preferentemente, a la metilguanósina con casquete en 7 (m⁷G). En el contexto de la presente invención, el término "casquete 5'" incluye un análogo del casquete en 5' que se asemeja a la estructura del casquete de ARN y se modifica para que posea la capacidad de estabilizar el ARN si se une a este, preferentemente, *in vivo* y/o en una célula. Proporcionar un ARN con un casquete 5' o análogo del casquete 5' puede lograrse mediante la transcripción *in vitro* de un molde de ADN en presencia de dicho casquete 5' o análogo del casquete 5', en donde dicho casquete 5' se incorpora cotranscripcionalmente en la cadena de ARN generada, o el ARN puede generarse, por ejemplo, mediante la transcripción *in vitro*, y el casquete 5' puede generarse postranscripcionalmente mediante el uso de enzimas que añaden el casquete, por ejemplo, las enzimas del virus vaccinia que añaden el casquete.

El término "ácido nucleico" de acuerdo con la invención comprende, además, una derivatización química de un ácido nucleico en una base de nucleótidos, en el azúcar o en el fosfato, y ácidos nucleicos que contienen nucleótidos no naturales y análogos de nucleótidos.

Los términos "fragmento" o "fragmento de una secuencia de ácido nucleico" se refieren a una parte de una secuencia de ácido nucleico, es decir, una secuencia que representa la secuencia de ácido nucleico acortada en los extremos 5' y/o 3'. Preferentemente, un fragmento cuando reemplaza dicha secuencia de ácido nucleico en una molécula de ARN retiene la estabilidad del ARN y/o la eficiencia de la traducción. Preferentemente, un fragmento de una secuencia de ácido nucleico comprende al menos 80 %, preferentemente al menos 90 %, 95 %, 96 %, 97 %, 98 %, o 99 % de los residuos de nucleótidos de dicha secuencia de ácido nucleico.

El término "variante" con respecto a, por ejemplo, secuencias de ácido nucleico o aminoácidos, de acuerdo con la invención, incluye cualquiera de las variantes, en particular, mutantes, variantes de empalme, conformaciones, isoformas, variantes alélicas, variantes de especies y homólogos de especies, en particular aquellos que son de origen natural. Una variante alélica se refiere a una alteración en la secuencia normal de un gen, cuya relevancia frecuentemente no está clara. La secuenciación génica completa a menudo identifica numerosas variantes alélicas para un gen determinado. Un homólogo de especie es una secuencia de ácido nucleico o aminoácido con una especie de origen diferente de una secuencia determinada de un ácido nucleico o aminoácido.

De acuerdo con la invención, las variantes de ácido nucleico incluyen delecciones, adiciones, mutaciones y/o inserciones de nucleótidos, únicas o múltiples en comparación con el ácido nucleico de referencia. Las delecciones incluyen la eliminación de uno o más nucleótidos del ácido nucleico de referencia. Las variantes de adición comprenden fusiones terminales 5' y/o 3' de uno o más nucleótidos, tales como 1, 2, 3, 5, 10, 20, 30, 50, o más nucleótidos. Las mutaciones pueden incluir, pero no se limitan a sustituciones, en donde se elimina al menos un nucleótido en la secuencia y se inserta otro nucleótido en su lugar (tal como transversiones y transiciones), sitios abásicos, sitios de uniones cruzadas, y bases alteradas o modificadas químicamente. Las inserciones incluyen la adición de al menos un nucleótido en el ácido nucleico de referencia.

Con respecto a las moléculas de ácido nucleico, el término "variante" incluye secuencias de ácidos nucleicos degenerados, en donde un ácido nucleico degenerado de acuerdo con la invención es un ácido nucleico que difiere de un ácido nucleico de referencia en la secuencia de codones debido a la degeneración del código genético.

Preferentemente, el grado de identidad entre una secuencia de ácido nucleico dada y una secuencia de ácido nucleico que es una variante de dicha secuencia de ácido nucleico dada será al menos 70 %, preferentemente al menos 75 %, o al menos 80 %.

preferentemente al menos 80 %, con mayor preferencia al menos 85 %, incluso con mayor preferencia al menos 90 % o con la máxima preferencia al menos 95 %, 96 %, 97 %, 98 % o 99 %. El grado de identidad es dado, preferentemente, para una región de al menos aproximadamente 30, al menos aproximadamente 50, al menos aproximadamente 70, al menos aproximadamente 90, al menos aproximadamente 100, al menos aproximadamente 150, al menos aproximadamente 200, al menos aproximadamente 250, al menos aproximadamente 300, o al menos aproximadamente 400 nucleótidos. En modalidades preferidas, el grado de identidad está dado por la longitud completa de la secuencia del ácido nucleico de referencia.

"Similitud de secuencia" indica el porcentaje de aminoácidos que son idénticos o que representan sustituciones conservadoras de aminoácidos. "Identidad de secuencia" entre dos polipéptidos o secuencias de ácido nucleico indica el porcentaje de aminoácidos o nucleótidos que son idénticos entre las secuencias.

El término "% idéntico" pretende referirse, en particular, a un porcentaje de nucleótidos que son idénticos en un alineamiento óptimo entre dos secuencias a comparar, dicho porcentaje es puramente estadístico, y las diferencias entre las dos secuencias pueden distribuirse aleatoriamente en toda la longitud de la secuencia y la secuencia a compararse puede comprender adiciones o deleciones en comparación con la secuencia de referencia, para obtener un alineamiento óptimo entre dos secuencias. Las comparaciones de dos secuencias generalmente se realizan mediante la comparación de dichas secuencias, después de un alineamiento óptimo, con respecto a un segmento o "ventana de comparación", para identificar regiones locales de secuencias correspondientes. El alineamiento óptimo de las secuencias para una comparación puede realizarse manualmente o con la ayuda del algoritmo de homología local de Smith y Waterman, 1981, *Ads App. Math.* 2, 482, con la ayuda del algoritmo de homología local de Neddleman y Wunsch, 1970, *J. Mol. Biol.* 48, 443, y con la ayuda del algoritmo de búsqueda de similitud de Pearson y Lipman, 1988, *Proc. Natl Acad. Sci. USA* 85, 2444, o con la ayuda de programas informáticos que usan dichos algoritmos (GAP, BESTFIT, FASTA, BLAST P, BLAST N y TFASTA en el paquete de programas informáticos de Wisconsin Genetics, Genetics Computer Group, 575 Science Drive, Madison, Wis.).

El porcentaje de identidad se obtiene mediante la determinación del número de posiciones idénticas en las que se corresponden las secuencias que se comparan, se divide este número entre el número de posiciones comparadas, y se multiplica este resultado por 100.

Por ejemplo, puede usarse el programa BLAST "BLAST 2 sequences" que está disponible en el sitio web <http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>.

Un ácido nucleico es "capaz de hibridarse" o "se hibrida" con otro ácido nucleico si las dos secuencias son complementarias entre sí. Un ácido nucleico es "complementario" a otro ácido nucleico si las dos secuencias son capaces de formar un dúplex estable entre sí. De acuerdo con la invención, la hibridación se realiza, preferentemente, en condiciones que permiten la hibridación específica entre polinucleótidos (condiciones rigurosas). Las condiciones rigurosas se describen, por ejemplo, en *Molecular Cloning: A Laboratory Manual*, J. Sambrook y otros, Editores, 2da Edición, Cold Spring Harbor Laboratory press, Cold Spring Harbor, Nueva York, 1989 o *Current Protocols in Molecular Biology*, F.M. Ausubel y otros, Editores, John Wiley & Sons, Inc., Nueva York y se refieren, por ejemplo, a una hibridación a 65 °C en un tampón de hibridación (3,5 x SSC, Ficoll al 0,02 %, polivinilpirrolidona al 0,02 %, albúmina de suero bovino al 0,02 %, NaH₂PO₄ 2,5 mM (pH 7), SDS al 0,5 %, EDTA 2 mM). SSC es cloruro de sodio 0,15 M/citrato de sodio 0,15 M, pH 7. Después de la hibridación, se lava la membrana a la que se ha transferido el ADN, por ejemplo, en 2 x SSC a temperatura ambiente y después en 0,1-0,5 x SSC/0,1 x SDS a temperaturas de hasta 68 °C.

Un porcentaje de complementariedad indica el porcentaje de residuos contiguos en una molécula de ácido nucleico que pueden formar enlaces de hidrógeno (por ejemplo, apareamiento de bases de Watson-Crick) con una segunda secuencia de ácido nucleico (por ejemplo, 5, 6, 7, 8, 9, 10 de 10 que son 50 %, 60 %, 70 %, 80 %, 90 % y 100 % complementarias). "Perfectamente complementario" o "completamente complementario" significa que todos los residuos contiguos de una secuencia de ácido nucleico se unirán por enlaces de hidrógeno con la misma cantidad de residuos contiguos en una segunda secuencia de ácido nucleico. Preferentemente, el grado de complementariedad de acuerdo con la invención es al menos 70 %, preferentemente, al menos 75 %, preferentemente, al menos 80 %, con mayor preferencia, al menos 85 %, aún con mayor preferencia, al menos 90 % o con la máxima preferencia, al menos 95 %, 96 %, 97 %, 98 % o 99 %. Con la máxima preferencia, el grado de complementariedad de acuerdo con la invención es 100 %.

El término "derivado" comprende la formación de un derivado químico de un ácido nucleico en una base nucleotídica, en el azúcar o en el fosfato. El término "derivado" comprende, además, ácidos nucleicos que contienen nucleótidos y análogos de nucleótidos que no son de origen natural. Preferentemente, la formación de un derivado de un ácido nucleico aumenta su estabilidad.

Los fragmentos o variantes de secuencias de ácido nucleico específicas o secuencias de ácido nucleico que tienen un grado particular de identidad con secuencias de ácido nucleico específicas tienen, preferentemente, al menos una propiedad funcional de dichas secuencias específicas y, preferentemente, son funcionalmente equivalentes a dichas secuencias específicas, por ejemplo, secuencias de ácido nucleico que presentan propiedades idénticas o similares a las de las secuencias de ácido nucleico específicas.

Una propiedad importante es retener o mejorar la estabilidad de una molécula de ARN y/o la eficiencia de la traducción e incluye en particular la capacidad de aumentar, en un enlace funcional a un ácido nucleico que puede transcribirse en ARN (secuencia de ácido nucleico transcribible) o una secuencia de ácido nucleico codificante para un péptido o proteína, la estabilidad y/o la eficiencia de la traducción del ARN producido a partir de este ácido nucleico o de la secuencia de ácido nucleico codificante para un péptido o proteína en la molécula completa de ARN.

En una modalidad, si una secuencia específica de ácido nucleico es activa a fines de aumentar la eficiencia de la traducción y/o la estabilidad de otra secuencia de ácido nucleico, un fragmento o variante de la secuencia de ácido nucleico específica o una secuencia de ácido nucleico que tiene un grado particular de identidad con la secuencia de ácido nucleico específica, además, es activa a fines de aumentar la eficiencia de la traducción y/o la estabilidad de la otra secuencia de ácido nucleico (cuando esta sustituye la secuencia de ácido nucleico específica). Un fragmento o variante de la secuencia de ácido nucleico específica o una secuencia de ácido nucleico que tiene un grado particular de identidad con la secuencia de ácido nucleico específica puede ser tan activa o más activa que la secuencia de ácido nucleico específica o la actividad de un fragmento o variante de la secuencia de ácido nucleico específica o de una secuencia de ácido nucleico que tiene un grado particular de identidad con la secuencia de ácido nucleico específica puede ser al menos 50 %, al menos 60 %, al menos 70 %, al menos 80 % o al menos 90 % de la actividad de la secuencia de ácido nucleico específica.

De acuerdo con la invención, "enlace funcional" o "unido funcionalmente" se refiere a una conexión dentro de una relación funcional. Un ácido nucleico está "unido funcionalmente" si este se relaciona funcionalmente con otra secuencia de ácido nucleico. Por ejemplo, un promotor está unido funcionalmente a una secuencia codificante si influye en la transcripción de dicha secuencia codificante. Los ácidos nucleicos unidos funcionalmente son típicamente adyacentes entre sí, cuando es apropiado, separados por secuencias de ácido nucleico adicionales, y, en modalidades particulares, se transcriben mediante la ARN polimerasa para proporcionar una molécula de ARN única (transcrito común). Preferentemente, una secuencia que es una variante con respecto a una secuencia específica, cuando sustituye la secuencia específica en una molécula de ARN retiene la estabilidad del ARN y/o la eficiencia de la traducción.

De acuerdo con la invención, una "secuencia de ácido nucleico que se deriva de una secuencia de ácido nucleico" se refiere a un ácido nucleico que es una variante del ácido nucleico del que se deriva.

El término "extremo 3' de un ácido nucleico" se refiere de acuerdo con la invención a ese extremo que tiene un grupo hidroxilo libre. En una representación esquemática de ácidos nucleicos bicatenarios, en particular ADN, el extremo 3' siempre está en el lado derecho. El término "extremo 5' de un ácido nucleico" se refiere de acuerdo con la invención a ese extremo que tiene un grupo fosfato libre. En una representación esquemática de ácidos nucleicos de doble cadena, en particular ADN, el extremo 5' siempre está en el lado izquierdo.

extremo 5' 5'--P-NNNNNNN-OH-3' extremo 3'

3'-HO-NNNNNNN-P--5'

En modalidades particulares, un ácido nucleico está unido funcionalmente de acuerdo con la invención a secuencias de control de la expresión que pueden ser homólogas o heterólogas con respecto al ácido nucleico.

Una secuencia de ácido nucleico transcribible, en particular una secuencia de ácido nucleico codificante para un péptido o proteína, y una secuencia de control de la expresión están unidas "funcionalmente" entre sí, si están unidas covalentemente entre sí de tal manera que la transcripción o expresión de la secuencia de ácido nucleico transcribible y en particular codificante está bajo el control o bajo la influencia de la secuencia de control de la expresión. Si la secuencia de ácido nucleico debe traducirse en un péptido o proteína funcional, la inducción de una secuencia de control de la expresión unida funcionalmente a la secuencia codificante resulta en la transcripción de dicha secuencia codificante, sin causar un corrimiento del marco en la secuencia codificante o la secuencia codificante es incapaz de traducirse en el péptido o proteína deseada.

El término "secuencia de control de la expresión" comprende, de acuerdo con la invención, promotores, secuencias de unión a ribosomas y otros elementos de control que controlan la transcripción de un gen o la traducción del ARN derivado. En modalidades particulares de la invención, pueden regularse las secuencias de control de la expresión. La estructura exacta de las secuencias de control de la expresión puede variar en dependencia de la especie o tipo de célula, pero generalmente incluye las secuencias no transcritas en 5' y las secuencias no traducidas en 5' y 3' involucradas en el inicio de la transcripción y la traducción, respectivamente, tales como caja TATA, secuencia de casquete, secuencia CAAT y similares. Más específicamente, las secuencias de control de la expresión no transcritas en 5' incluyen una región promotora que incluye una secuencia promotora para el control de la transcripción del gen unido funcionalmente. Las secuencias de control de la expresión pueden incluir, además, las secuencias potenciadoras o secuencias activadoras en dirección 5'.

Las secuencias de ácido nucleico especificadas en la presente descripción, en particular las secuencias de ácido nucleico transcribibles y codificantes, pueden combinarse con cualquier secuencia de control de la expresión, en particular promotores, que pueden ser homólogos o heterólogos a dichas secuencias de ácido nucleico, con el término "homólogo"

que se refiere al hecho de que una secuencia de ácido nucleico además está unida funcionalmente de manera natural a la secuencia de control de la expresión, y el término "heterólogo" se refiere al hecho de que una secuencia de ácido nucleico no está unida funcionalmente de manera natural a la secuencia de control de la expresión.

5 El término "promotor" o "región promotora" se refiere a una secuencia de ADN en dirección 5' (5') de la secuencia codificante de un gen, que controla la expresión de dicha secuencia codificante porque proporciona un sitio de reconocimiento y de unión para la ARN polimerasa. La "región promotora" puede incluir sitios de reconocimiento o de unión adicionales para otros factores involucrados en la regulación de la transcripción de dicho gen. Un promotor puede controlar la transcripción de un gen procariótico o eucariótico. Un promotor puede ser "inducible" e iniciar la transcripción en respuesta a un inductor, o puede ser "constitutivo" si la transcripción no está controlada por un agente inductor. Un promotor inducible se expresa solo en un grado muy pequeño o nada en absoluto, si un inductor está ausente. En presencia del inductor, el gen se activa o aumenta el nivel de la transcripción. Esto está mediado, en general, por la unión de un factor de transcripción específico.

15 Los ejemplos de promotores preferidos de acuerdo con la invención son los promotores para la polimerasa SP6, T3 o T7.

De acuerdo con la invención, el término "expresión" se usa en su significado más general y comprende la producción de ARN o de ARN y proteína. Comprende también la expresión parcial de ácidos nucleicos. Además, la expresión puede ser transitoria o estable. Con respecto al ARN, el término "expresión" o "traducción" se refiere al proceso en los ribosomas de una célula mediante el cual una cadena de ARN mensajero dirige el ensamblaje de una secuencia de aminoácidos para producir un péptido o proteína.

25 El término "secuencias de ácido nucleico que pueden transcribirse para proporcionar un transcrito común" significa que dichas secuencias de ácido nucleico están unidas funcionalmente entre sí de tal manera que, cuando sea apropiado después de la linealización, tal como por la escisión, mediante enzima de restricción, de la molécula de ácido nucleico que comprende dichas secuencias de ácido nucleico, en particular de una molécula de ácido nucleico circular cerrada, la transcripción bajo el control de un promotor resulta en una molécula de ARN que comprende los transcritos de dichas secuencias de ácido nucleico unidos covalentemente entre sí, cuando sea apropiado, separados por secuencias ubicadas de por medio.

30 En el contexto de la presente invención, el término "transcripción" se refiere a un proceso, en donde el código genético en una secuencia de ADN se transcribe a ARN. Posteriormente, el ARN puede traducirse en proteína. De acuerdo con la presente invención, el término "transcripción" comprende la "transcripción *in vitro*", en donde el término "*transcripción in vitro*" se refiere a un proceso en donde el ARN, en particular el ARNm, se sintetiza *in vitro* en un sistema libre de células. Preferentemente, los vectores de clonación se aplican para la generación de transcritos. Estos vectores de clonación se designan generalmente como vectores de transcripción y se abarcan de acuerdo con la presente invención por el término "vector". De acuerdo con la presente invención, el ARN preferentemente es ARN transcrito *in vitro* (ARN-IVT) y puede obtenerse mediante la transcripción *in vitro* de un molde de ADN apropiado. El promotor para controlar la transcripción puede ser cualquier promotor para cualquier ARN polimerasa. Un molde de ADN para la transcripción *in vitro* puede obtenerse mediante la clonación de un ácido nucleico, en particular ADNc, y su introducción en un vector apropiado para la transcripción *in vitro*. El ADNc puede obtenerse mediante transcripción inversa del ARN.

45 El término "secuencia de ácido nucleico transcrita a partir de una secuencia de ácido nucleico" se refiere al ARN, cuando sea apropiado como parte de una molécula completa de ARN, que es un producto de la transcripción de la última secuencia de ácido nucleico.

50 El término "secuencia de ácido nucleico que es activa para aumentar la eficiencia y/o estabilidad de la traducción de una secuencia de ácido nucleico" significa que la primera secuencia de ácido nucleico es capaz de modificar, en un transcrito común con la segunda secuencia de ácido nucleico, la eficiencia de la traducción y/o la estabilidad de dicha segunda secuencia de ácido nucleico de tal manera que dicha eficiencia de la traducción y/o estabilidad se aumenta en comparación con la eficiencia de la traducción y/o estabilidad de dicha segunda secuencia de ácido nucleico sin dicha primera secuencia de ácido nucleico. En este contexto, el término "eficiencia de la traducción" se refiere a la cantidad de producto de traducción proporcionado por una molécula de ARN dentro de un período de tiempo particular y el término "estabilidad" se refiere a la vida media de una molécula de ARN.

55 La modificación, y de esta manera la estabilización y/o el aumento en la eficiencia de la traducción, del ARN puede lograrse de acuerdo con la invención mediante la modificación genética de la expresión de las moléculas de ácido nucleico de la invención cuando se usan como vectores de expresión de tal manera que permiten la transcripción del ARN con regiones no traducidas en 3' como se describe en la presente descripción en su extremo 3', y preferentemente entre la secuencia codificante para un péptido o proteína (marco abierto de lectura) y la secuencia de poli(A)

60 El término "región no traducida en 3'" se refiere a una región que se ubica en el extremo 3' de un gen, en dirección 3' del codón de terminación de una región codificante de proteína, y que se transcribe pero no se traduce en una secuencia de aminoácidos, o a la región correspondiente en una molécula de ARN.

65

De acuerdo con la invención, se considera que una primera región polinucleotídica se ubica en dirección 3' de una segunda región polinucleotídica, si el extremo 5' de dicha primera región polinucleotídica es la parte de dicha primera región polinucleotídica más cercana al extremo 3' de dicha segunda región polinucleotídica.

5 La región no traducida en 3' se extiende, típicamente, desde el codón de terminación para un producto de la traducción hasta la secuencia de poli(A) que generalmente se une después del proceso de transcripción. Las regiones no traducidas en 3' del ARNm de mamífero, típicamente, tienen una región de homología conocida como la secuencia de hexanucleótidos AAUAAA. Presumiblemente, esta secuencia es la señal de unión de poli(A) y se ubica frecuentemente de 10 a 30 bases en dirección 5' del sitio de unión de poli(A).

10 Las regiones no traducidas en 3' pueden contener una o más repeticiones invertidas que pueden plegarse para dar estructuras de bucle y tallo que actúan como barreras para las exoribonucleasas o interactúan con proteínas que se conocen que aumentan la estabilidad del ARN (por ejemplo, proteínas de unión al ARN).

15 Las regiones no traducidas en 5' y/o en 3' pueden, de acuerdo con la invención, unirse funcionalmente a un ácido nucleico transcribible y en particular codificante, para que estas regiones se asocien con el ácido nucleico de tal manera que aumente la estabilidad y/o la eficiencia de la traducción del ARN transcrito a partir de dicho ácido nucleico transcribible.

20 Las regiones no traducidas en 3' de los ARNm de inmunoglobulinas son relativamente cortas (menos de aproximadamente 300 nucleótidos), mientras que las regiones no traducidas en 3' de otros genes son relativamente largas. Por ejemplo, la región no traducida en 3' del tPA es aproximadamente 800 nucleótidos de longitud, la del factor VIII es aproximadamente 1800 nucleótidos de longitud y la de la eritropoyetina es aproximadamente 560 nucleótidos de longitud.

25 De acuerdo con la invención puede determinarse, si una región no traducida en 3' o una secuencia de ácido nucleico derivada de esta aumenta la estabilidad y/o la eficiencia de la traducción del ARN, mediante la incorporación de la región no traducida en 3' o la secuencia de ácido nucleico derivada de esta en la región no traducida en 3' de un gen y medir si dicha incorporación aumenta la cantidad de proteína sintetizada.

30 Lo anterior se aplica en consecuencia al caso en el que, de acuerdo con la invención, un ácido nucleico comprende dos o más regiones no traducidas en 3' que, preferentemente, se acoplan secuencialmente con o sin un enlazador intermedio, preferentemente en una "relación de cabeza a cola" (es decir, las regiones no traducidas en 3' tienen la misma orientación, preferentemente, la orientación de origen natural en un ácido nucleico).

35 De acuerdo con la invención, el término "gen" se refiere a una secuencia de ácido nucleico particular que es responsable de producir uno o más productos celulares y/o lograr una o más funciones intercelulares o intracelulares. Más específicamente, dicho término se refiere a una sección de ADN que comprende un ácido nucleico codificante para una proteína específica o una molécula de ARN funcional o estructural.

40 La poliadenilación es la adición de una secuencia o cola de poli(A) a un transcrito de ARN primario. La secuencia de poli(A) consiste en múltiples monofosfatos de adenosina. En otras palabras, es un tramo de ARN que solo tiene bases de adenina. En eucariotas, la poliadenilación es parte del proceso que produce ARN mensajero maduro (ARNm) para la traducción. Este, por lo tanto, forma parte del proceso más amplio de expresión génica. El proceso de poliadenilación comienza cuando la transcripción de un gen finaliza, o termina. El segmento más 3' del pre-ARNm recién creado se escinde primero mediante un conjunto de proteínas; después estas proteínas sintetizan la secuencia de poli(A) en el extremo 3' del ARN. La secuencia de poli(A) es importante para la exportación nuclear, la traducción, y la estabilidad del ARNm. La secuencia se acorta con el tiempo y, cuando es lo suficientemente corta, el ARNm se degrada enzimáticamente.

45 Los términos "secuencia de poliadenilo", "secuencia de poli(A)" o "cola de poli(A)" se refieren a una secuencia de residuos de adenilo que se ubica, típicamente, en el extremo 3' de una molécula de ARN. La invención proporciona que tal secuencia se una durante la transcripción del ARN por medio de un molde de ADN sobre la base de residuos repetidos de timidilo en la cadena complementaria a la cadena codificante, mientras que dicha secuencia normalmente no está codificada en el ADN pero se une al extremo 3' libre del ARN mediante una ARN polimerasa independiente del molde después de la transcripción, en el núcleo. De acuerdo con la invención, en una modalidad, una secuencia de poli(A) tiene al menos 20, preferentemente al menos 40, preferentemente al menos 80, preferentemente al menos 100 y
50 preferentemente hasta 500, preferentemente hasta 400, preferentemente hasta 300, preferentemente hasta 200, y en particular hasta 150, nucleótidos A, preferentemente nucleótidos A consecutivos, y en particular aproximadamente 120 nucleótidos A. El término "nucleótidos A" o "A" se refiere a residuos de adenilo.

55 En una modalidad preferida, una molécula de ácido nucleico de acuerdo con la invención es un vector. El término "vector" se usa en la presente en su significado más general y comprende cualquier vehículo intermediario para un ácido nucleico que, por ejemplo, permite que dicho ácido nucleico se introduzca en células huésped procariontas y/o eucariotas y, cuando sea apropiado, se integre en un genoma. Preferentemente, tales vectores se replican y/o se expresan en la célula. Los vectores comprenden plásmidos, fagémidos o genomas virales. El término "plásmido", como se usa en la presente descripción, se refiere generalmente a una construcción de material genético extracromosómico, generalmente un ADN
60 dúplex circular, que puede replicarse independientemente del ADN cromosómico.

Los ácidos nucleicos descritos en la presente descripción pueden ser moléculas recombinantes y/o aisladas.

- Una "molécula aislada" como se usa en la presente descripción, pretende referirse a una molécula que es sustancialmente libre de otras moléculas tal como otro material celular. El término "ácido nucleico aislado" significa de acuerdo con la invención que el ácido nucleico se (i) amplificó *in vitro*, por ejemplo mediante reacción en cadena de la polimerasa (PCR), (ii) se produjo por vía recombinante mediante clonación, (iii) se purificó, por ejemplo, mediante escisión y fraccionamiento electroforético en gel, o (iv) se sintetizó, por ejemplo, mediante síntesis química. Un ácido nucleico aislado es un ácido nucleico disponible para su manipulación mediante las técnicas del ADN recombinante.
- El término "recombinante" en el contexto de la presente invención significa "hecho a través de ingeniería genética". Preferentemente, un "objeto recombinante" tal como una célula recombinante en el contexto de la presente invención no es de origen natural.
- El término "de origen natural", como se usa en la presente descripción, se refiere al hecho de que un objeto puede encontrarse en la naturaleza. Por ejemplo, es de origen natural un péptido o ácido nucleico que está presente en un organismo (lo que incluye los virus) y que puede aislarse a partir de una fuente en la naturaleza y que no se ha modificado intencionalmente por el hombre en el laboratorio.
- De acuerdo con la invención, el término "célula huésped" se refiere a cualquier célula que puede transformarse o transfectarse con un ácido nucleico exógeno. El término "célula huésped" comprende, de acuerdo con la invención, células procariotas (por ejemplo, *E. Coli*) o eucariotas (por ejemplo, células de levadura y células de insectos). Se prefieren particularmente células de mamífero, tales como células de humanos, ratones, hamsters, cerdos, cabras y primates. Las células pueden derivarse de una multiplicidad de tipos de tejidos y comprenden células primarias y líneas celulares. Los ejemplos específicos incluyen queratinocitos, leucocitos de sangre periférica, células madre de médula ósea y células madre embrionarias. En otras modalidades, la célula huésped es una célula presentadora de antígeno, en particular una célula dendrítica, un monocito o un macrófago. Un ácido nucleico puede estar presente en la célula huésped en una sola copia o en varias copias y, en una modalidad se expresa en la célula huésped.
- E. Coli* es una bacteria gramnegativa, facultativamente anaeróbica, en forma de bastón del género *Escherichia* que se encuentra comúnmente en el intestino inferior de los organismos de sangre caliente. La bacteria puede cultivarse fácilmente y económicamente en un entorno de laboratorio, y se ha investigado intensamente durante más de 60 años. *E. coli* es el organismo modelo procariota más estudiado, y una especie importante en los campos de la biotecnología y la microbiología, donde sirvió como el organismo huésped para la mayoría del trabajo con el ADN recombinante. Las cepas de *E. coli* de acuerdo con la invención incluyen: AG1, AB1157, B2155, BL21, BNN93, BNN97, BW26434, C600, CSH50, D1210, DB3.1, DH1, DH5α, DH10B, DH12S, DM1, *E. cloni(r)*, *E.coli* K12 ER2738, ER2566, ER2267, HB101, IJ1126, IJ1127, JM83, JM101, JM103, JM105, JM106, JM107, JM108, JM109, JM110, JM2.300, LE392, Mach1, MC1061, MC4100, MFDpir, MG1655, OmniMAX2, RR1, RV308, SOLR, SS320, STBL2, STBL3, STBL4, SURE, SURE2, TG1, TOP10, Top10F', W3110, WM3064, XL1-Blue, XL2-Blue, XL1-Red y XL10-Gold.
- De acuerdo con la presente invención, el término "péptido" comprende oligo- y polipéptidos y se refiere a las sustancias que comprenden dos o más, preferentemente 3 o más, preferentemente 4 o más, preferentemente 6 o más, preferentemente 8 o más, preferentemente 10 o más, preferentemente 13 o más, preferentemente 16 o más, preferentemente 20 o más, y hasta, preferentemente 50, preferentemente 100 o preferentemente 150, aminoácidos consecutivos unidos entre sí a través de enlaces peptídicos. El término "proteína" se refiere a péptidos grandes, preferentemente, péptidos que tienen al menos 151 aminoácidos, pero los términos "péptido" y "proteína" se usan en la presente descripción generalmente como sinónimos.
- Los términos "péptido" y "proteína" comprenden de acuerdo con la invención sustancias que no solo contienen componentes de aminoácidos sino además componentes no aminoácidos tales como estructuras de azúcares y fosfatos, y además comprenden sustancias que contienen enlaces tales como enlaces éster, tioéter o disulfuro.
- De acuerdo con la presente invención, un ácido nucleico tal como el ARN puede codificar un péptido o proteína. En consecuencia, una secuencia de ácido nucleico transcribible o un transcrito de esta pueden contener un marco abierto de lectura (ORF) que codifica un péptido o proteína. Dicho nucleico puede expresar el péptido o proteína codificado. Por ejemplo, dicho ácido nucleico puede ser un ácido nucleico que codifica y expresa un antígeno o un péptido o proteína farmacéuticamente activo tal como un compuesto inmunológicamente activo (que preferentemente no es un antígeno).
- De acuerdo con la invención, el término "ácido nucleico que codifica un péptido o proteína" significa que el ácido nucleico, si está presente en el entorno apropiado, preferentemente dentro de una célula, puede dirigir el ensamblaje de aminoácidos para producir el péptido o proteína durante el proceso de traducción. Preferentemente, el ARN de acuerdo con la invención es capaz de interactuar con la maquinaria de traducción celular y permitir la traducción del péptido o proteína.
- De acuerdo con la invención, en una modalidad, el ARN comprende o consiste en ARN farmacéuticamente activo. Un "ARN farmacéuticamente activo" puede ser ARN que codifica un péptido o proteína farmacéuticamente activo.

- Un "péptido o proteína farmacéuticamente activo" tiene un efecto positivo o ventajoso sobre la afección o el estado de enfermedad de un sujeto cuando se administra al sujeto en una cantidad con eficacia terapéutica. Preferentemente, un péptido o proteína farmacéuticamente activo tiene propiedades curativas o paliativas y puede administrarse para mejorar, aliviar, mitigar, revertir, retrasar el inicio o disminuir la gravedad de uno o más síntomas de una enfermedad o trastorno.
- Un péptido o proteína farmacéuticamente activo puede tener propiedades profilácticas y puede usarse para retrasar el comienzo de una enfermedad o para disminuir la gravedad de tal enfermedad o afección patológica. El término "péptido o proteína farmacéuticamente activo" incluye proteínas o polipéptidos completos, y además puede referirse a fragmentos farmacéuticamente activos de estos. Además, puede incluir análogos farmacéuticamente activos de un péptido o proteína. El término "péptido o proteína farmacéuticamente activo" incluye péptidos y proteínas que son antígenos, es decir, el péptido o proteína provoca una respuesta inmunitaria en un sujeto, que puede ser terapéutica o parcial o totalmente protectora.
- Los ejemplos de proteínas farmacéuticamente activas incluyen, pero no se limitan a, citocinas y proteínas del sistema inmunitario tales como compuestos inmunológicamente activos (por ejemplo, interleucinas, factor estimulante de colonias (CSF), factor estimulante de colonias de granulocitos (G-CSF), factor estimulantes de colonias de granulocitos y macrófagos (GM-CSF), eritropoyetina, factor de necrosis tumoral (TNF), interferones, integrinas, adresinas, selectinas, receptores de direccionamiento a tejidos, receptores de células T, inmunoglobulinas, antígenos solubles del complejo principal de histocompatibilidad, antígenos inmunológicamente activos tales como antígenos bacterianos, parasitarios o virales, alérgenos, autoantígenos, anticuerpos), hormonas (insulina, hormona tiroidea, catecolaminas, gonadotrofinas, hormonas tróficas, prolactina, oxitocina, dopamina, somatotropina bovina, leptinas y similares), hormonas de crecimiento (por ejemplo, hormona de crecimiento humana), factores de crecimiento (por ejemplo, factor de crecimiento epidérmico, factor de crecimiento nervioso, factor de crecimiento similar a la insulina y similares), receptores de factor de crecimiento, enzimas (activador tisular del plasminógeno, estreptoquinasa, enzimas de la síntesis o degradación del colesterol, enzimas esteroideogénicas, quinasas, fosfodiesterasas, metilasas, desmetilasas, deshidrogenasas, celulasas, proteasas, lipasas, fosfolipasas, aromatasas, citocromos, adenilato o guanilato ciclasas, neuramidasa y similares) receptores (receptores de hormonas esteroideas, receptores de péptidos), proteínas de unión (proteínas de unión a la hormona del crecimiento o a factores de crecimiento y similares), factores de transcripción y traducción, proteínas supresoras del crecimiento tumoral (por ejemplo, proteínas que inhiben la angiogénesis), proteínas estructurales (tales como colágeno, fibroína, fibrinógeno, elastina, tubulina, actina, y miosina), proteínas de la sangre (trombina, albúmina sérica, Factor VII, Factor VIII, insulina, Factor IX, Factor X, activador tisular del plasminógeno, proteína C, factor von Willebrand, antitrombina III, glucocerebrosidasa, eritropoyetina factor estimulante de colonias de granulocitos (G-CSF) o Factor VIII modificado, anticoagulantes y similares.
- En una modalidad, la proteína farmacéuticamente activa de acuerdo con la invención es una citocina que está implicada en la regulación de la homeostasis linfóide, preferentemente una citocina que está implicada en, y preferentemente, induce o mejora el desarrollo, activación, expansión, diferenciación y/o supervivencia de las células T. En una modalidad, la citocina es una interleucina. En una modalidad, la proteína farmacéuticamente activa de acuerdo con la invención es una interleucina seleccionada del grupo que consiste en IL-2, IL-7, IL-12, IL-15 e IL-21.
- El término "compuesto inmunológicamente activo" se refiere a cualquier compuesto que altera una respuesta inmunitaria, preferentemente mediante la inducción y/o supresión de la maduración de las células inmunitarias, mediante la inducción y/o supresión de la biosíntesis de citocinas, y/o mediante la alteración de la inmunidad humoral a través de la estimulación de la producción de anticuerpos por las células B. Los compuestos inmunológicamente activos poseen una actividad inmunestimulante potente que incluye, pero no se limita a, actividad antiviral y antitumoral, y además pueden regular negativamente otros aspectos de la respuesta inmunitaria, por ejemplo, cambiar la respuesta inmunitaria de una respuesta inmunitaria TH2, que es útil para tratar una amplia gama de enfermedades mediadas por TH2. Los compuestos inmunológicamente activos pueden ser útiles como adyuvantes de vacunas.
- Si, de acuerdo con la presente invención, es conveniente inducir o mejorar una respuesta inmunitaria mediante el uso de ARN como se describe en la presente descripción, la respuesta inmunitaria puede activarse o potenciarse mediante el ARN. Por ejemplo, las proteínas o péptidos codificados por los ARN o productos de procesamiento de estos pueden presentarse mediante las proteínas del complejo principal de histocompatibilidad (MHC) expresadas en las células presentadoras de antígeno. El complejo MHC péptido puede ser reconocido por las células inmunitarias, tal como las células T, lo que conducen a su activación.
- En una modalidad, el ARN que codifica para un antígeno, tal como un antígeno asociado a la enfermedad, se administra a un mamífero, en particular si es conveniente tratar a un mamífero que tiene una enfermedad que involucra el antígeno. El ARN es tomado por las células presentadoras de antígeno del mamífero (monocitos, macrófagos, células dendríticas u otras células). Se forma un producto antigénico de la traducción del ARN y el producto se presenta en la superficie de las células para su reconocimiento por las células T. En una modalidad, el antígeno se presenta sobre la superficie celular para su reconocimiento por las células T CAR modificadas genéticamente dirigidas al antígeno. En una modalidad, el antígeno o un producto producido mediante procesamiento opcional de este se presenta en la superficie celular en el contexto de las moléculas del MHC para el reconocimiento por las células T a través de su receptor de células T.
- Alternativamente, la presente invención prevé modalidades en donde el ARN que expresa un antígeno se introduce en células presentadoras de antígeno *ex vivo*, por ejemplo, células presentadoras de antígeno tomadas de un paciente, y las

células presentadoras de antígeno, opcionalmente propagadas clonalmente *ex vivo*, se trasplantan nuevamente en el mismo paciente. Las células transfectadas pueden reintroducirse en el paciente mediante el uso de cualquier medio conocido en la técnica, preferentemente en forma estéril mediante administración intravenosa, intracavitaria, intraperitoneal o intratumoral.

Los métodos de la invención pueden implicar una célula presentadora de antígeno para expresar el ARN que codifica el antígeno. Para este fin, los métodos de la invención pueden implicar la introducción del ARN que codifica los antígenos en células presentadoras de antígeno tales como las células dendríticas. Para la transfección de las células presentadoras de antígeno tales como células dendríticas, puede usarse una composición farmacéutica que comprende el ARN que codifica el antígeno. Puede administrarse a un paciente un vehículo de suministro que direcciona el ARN a una célula dendrítica o a otra célula presentadora de antígeno, lo que resulta en la transfección que ocurre *in vivo*.

De acuerdo con la invención, se prefiere usar formulaciones del ARN codificante de un antígeno las que suministran el ARN con alta selectividad a las células presentadoras de antígeno tales como las células dendríticas (DC) en el bazo después de la administración sistémica. Por ejemplo, formulaciones nanoparticuladas de ARN con un tamaño de partícula definido en donde la carga neta de las partículas es cercana a cero o negativa, tal como los lipoplejos de ARN, cargados negativamente o electro neutros, y liposomas, por ejemplo, los lipoplejos que comprenden DOTMA y DOPE o DOTMA y colesterol, conducen a la expresión sustancial del ARN en las DC de bazo después de la administración sistémica. Se determinó una fuerte expresión en las células diana (bazo) mientras que la expresión en otros órganos fue baja.

Como se usa en la presente descripción, el término "nanopartícula" se refiere a cualquier partícula que tiene un diámetro que la haga adecuada para la administración sistémica, en particular parenteral, de, en particular, ácidos nucleicos, típicamente, un diámetro de menos que 1000 nanómetros (nm). En algunas modalidades, una nanopartícula tiene un diámetro de menos que 600 nm. En algunas modalidades, una nanopartícula tiene un diámetro de menos que 400 nm.

Como se usa en la presente descripción, el término "formulación nanoparticulada" o los términos similares se refieren a cualquier sustancia que contiene al menos una nanopartícula. En algunas modalidades, una composición nanoparticulada es una colección uniforme de nanopartículas. En algunas modalidades, las composiciones nanoparticuladas son dispersiones o emulsiones. En general, una dispersión o emulsión se forma cuando se combinan al menos dos materiales inmiscibles.

El término "lipoplejo" o "lipoplejo de ácido nucleico", en particular "lipoplejo de ARN", se refiere a un complejo de lípidos y ácidos nucleicos, en particular ARN. Los lipoplejos se forman espontáneamente cuando los liposomas catiónicos, que además a menudo incluyen un lípido "auxiliar" neutro, se mezclan con ácidos nucleicos.

Si la presente invención se refiere a una carga tal como una carga positiva, carga negativa o carga neutra o un compuesto catiónico, compuesto negativo o compuesto neutro, esto significa generalmente que la carga mencionada está presente a un pH seleccionado, tal como un pH fisiológico. Por ejemplo, el término "lípido catiónico" significa un lípido que tiene una carga neta positiva a un pH seleccionado, tal como un pH fisiológico. El término "lípido neutro" significa un lípido que no tiene carga neta positiva o negativa y puede presentarse en la forma de un ion anfótero no cargado o neutro a un pH seleccionado, tal como un pH fisiológico. Por "pH fisiológico" se entiende, en la presente descripción, un pH de aproximadamente 7,5.

Los vehículos nanoparticulados tales como los vehículos lipídicos contemplados para usar en la presente invención incluyen cualquier sustancia o vehículo con el que pueda asociarse un ácido nucleico, tal como el ARN, por ejemplo mediante la formación de complejos con el ácido nucleico o mediante la formación de vesículas en las que el ácido nucleico está encerrado o encapsulado. Esto puede resultar en una mayor estabilidad del ácido nucleico en comparación con el ácido nucleico desnudo. En particular, la estabilidad del ácido nucleico en la sangre puede aumentarse.

Los lípidos catiónicos, los polímeros catiónicos y otras sustancias con cargas positivas pueden formar complejos con los ácidos nucleicos cargados negativamente. Estas moléculas catiónicas pueden usarse para formar complejos de ácidos nucleicos, de esta manera forman, por ejemplo, los llamados lipoplejos o poliplexos, respectivamente, y se ha demostrado que estos complejos suministran ácidos nucleicos a las células.

Las preparaciones de ácido nucleico nanoparticuladas para usar en la presente invención pueden obtenerse mediante diversos protocolos y a partir de diversos compuestos complejantes de ácido nucleico. Los lípidos, polímeros, oligómeros o anfífilos son agentes complejantes típicos. En una modalidad, el compuesto complejante comprende al menos un agente seleccionado del grupo que consiste en protamina, polietilenimina, una poli-L-lisina, una poli-L-arginina o una histona.

De acuerdo con la invención, la protamina es útil como agente vehículo catiónico. El término "protamina" se refiere a cualquiera de varias proteínas fuertemente básicas de peso molecular relativamente bajo que son ricas en arginina y se encuentran asociadas especialmente con el ADN en lugar de histonas somáticas en las células espermáticas de varios animales (como los peces). En particular, el término "protamina" se refiere a proteínas que se encuentran en los espermatozoides de los peces que son fuertemente básicas, solubles en agua, no se coagulan por calor, y producen principalmente arginina después de la hidrólisis. En forma purificada, se usan en una formulación de insulina de acción prolongada y para neutralizar los efectos anticoagulantes de la heparina.

De acuerdo con la invención, el término "protamina" como se usa en la presente descripción significa que comprende cualquier secuencia de aminoácidos de protamina que se obtiene o se deriva a partir de fuentes naturales o biológicas que incluyen fragmentos de esta y formas multiméricas de dicha secuencia de aminoácidos o fragmento de esta. Además, el término abarca polipéptidos (sintetizados) que son artificiales y diseñados específicamente para fines específicos y no pueden aislarse de fuentes naturales o biológicas.

La protamina usada de acuerdo con la presente invención puede ser protamina sulfatada o hidrocloruro de protamina. En una modalidad preferida, la fuente de protamina usada para la producción de las nanopartículas descritas en la presente descripción es la protamina 5000 que contiene protamina a más de 10 mg/mL (5000 unidades neutralizantes de heparina por mL) en una solución de sales isotónica.

Los liposomas son vesículas lipídicas microscópicas que a menudo tienen una o más bicapas de un lípido formador de vesículas, tal como un fosfolípido, y son capaces de encapsular un fármaco. En el contexto de la presente invención pueden emplearse diferentes tipos de liposomas, que incluyen, sin limitarse a estos, vesículas multilamelares (MLV), vesículas unilamelares pequeñas (SUV), vesículas unilamelares grandes (LUV), liposomas estabilizados estéricamente (SSL), vesículas multivesiculares (MV), y grandes vesículas multivesiculares (LMV), así como también otras formas de bicapa conocidas en la técnica. El tamaño y la lamellaridad del liposoma dependerán de la forma de preparación y la selección del tipo de vesículas a usarse dependerá del modo de administración preferido. Existen numerosas otras formas de organización supramolecular en las que los lípidos pueden estar presentes en un medio acuoso, que comprenden fases lamelares, fases hexagonales y fases hexagonales inversas, fases cúbicas, micelas, micelas inversas compuestas de monocapas. Estas fases pueden obtenerse, además, en combinación con ADN o ARN, y la interacción con ARN y ADN puede afectar sustancialmente el estado de la fase. Las fases descritas pueden estar presentes en las formulaciones nanoparticuladas del ácido nucleico de la presente invención.

Para la formación de lipoplejos de ácido nucleico a partir de ácido nucleico y liposomas, puede usarse cualquier método adecuado para formar liposomas siempre que proporcione los lipoplejos de ácido nucleico previstos. Los liposomas pueden formarse mediante el uso de métodos estándar tales como el método de evaporación inversa (REV), el método de inyección de etanol, el método de deshidratación-rehidratación (DRV), sonicación u otros métodos adecuados.

Después de la formación de liposomas, los liposomas pueden dimensionarse para obtener una población de liposomas que tienen un intervalo de tamaños sustancialmente homogéneo.

Los lípidos formadores de bicapa tienen, típicamente, dos cadenas hidrocarbonadas, particularmente cadenas de acilo, y un grupo cabeza, ya sea polar o no polar. Los lípidos formadores de bicapa están compuestos de lípidos de origen natural o de origen sintético, que incluyen los fosfolípidos, tales como fosfatidilcolina, fosfatidiletanolamina, ácido fosfatídico, fosfatidilinositol, y esfingomielina, donde las dos cadenas hidrocarbonadas son, típicamente, entre aproximadamente 14-22 átomos de carbono de longitud, y tienen diversos grados de insaturación. Otros lípidos adecuados para usar en la composición de la presente invención incluyen glicolípidos y esteroides tales como colesterol y sus diversos análogos, que además pueden usarse en los liposomas.

Los lípidos catiónicos, típicamente, tienen una porción lipofílica, tal como un esteroil, una cadena de acilo o diacilo, y tienen una carga neta general positiva. El grupo cabeza del lípido, típicamente, lleva la carga positiva. El lípido catiónico tiene, preferentemente, una carga positiva de 1 a 10 valencias, con mayor preferencia una carga positiva de 1 a 3 valencias, y con mayor preferencia una carga positiva de 1 valencia. Los ejemplos de lípidos catiónicos incluyen, pero no se limitan a 1,2-di-O-octadecenil-3-trimetilamonio propano (DOTMA); dimetildioctadecilamonio (DDAB); 1,2-dioleoil-3-trimetilamonio-propano (DOTAP); 1,2-dioleoil-3-dimetilamonio-propano (DODAP); 1,2-diaciloxi-3-dimetilamonio propanos; 1,2-dialquiloxi-3-dimetilamonio propanos; cloruro de dioctadecil-dimetilamonio (DODAC), 1,2-dimiristoiloxipropil-1,3-dimetilhidroxietil-amonio (DMRIE) y trifluoroacetato de 2,3-dioleoiloxi-N- [[2(espermina carboxamida)etil]-N,N-dimetil-1-propanamio (DOSPA). Los preferidos son DOTMA, DOTAP, DODAC, y DOSPA. El más preferido es DOTMA.

Además, las nanopartículas descritas en la presente descripción incluyen, además, preferentemente un lípido neutro en vista de la estabilidad estructural, y similares. El lípido neutro puede seleccionarse apropiadamente en vista de la eficiencia del suministro del complejo ácido nucleico-lípido. Los ejemplos de lípidos neutros incluyen, pero no se limitan a, 1,2-di-(9Z-octadecenoi)-sn-glicero-3-fosfoetanolamina (DOPE), 1,2-dioleoil-sn-glicero-3-fosfocolina (DOPC), diacilfosfatidilcolina, diacilfosfatidil etanolamina, ceramida, esfingomielina, cefalina, esteroil, y cerebrósido. El preferido es DOPE y/o DOPC. El más preferido es DOPE. En el caso en que un liposoma catiónico incluye tanto un lípido catiónico como un lípido neutro, la relación molar del lípido catiónico con respecto al lípido neutro puede determinarse apropiadamente en vista de la estabilidad del liposoma y similares.

De acuerdo con una modalidad, las nanopartículas descritas en la presente descripción pueden comprender fosfolípidos. Los fosfolípidos pueden ser un glicerofosfolípido. Los ejemplos de glicerofosfolípido incluyen, sin limitación, tres tipos de lípidos: (i) fosfolípidos zwitteriónicos, que incluyen, por ejemplo, fosfatidilcolina (PC), fosfatidilcolina de yema de huevo, PC derivada de soja en forma natural, parcialmente hidrogenada o totalmente hidrogenada, dimiristoil fosfatidilcolina (DMPC) esfingomielina (SM); (ii) fosfolípidos cargados negativamente: que incluyen, por ejemplo, fosfatidilserina (PS), fosfatidilinositol (PI), ácido fosfatídico (PA), fosfatidilglicerol (PG) dipalmitoil PG, dimiristoil fosfatidilglicerol (DMPG); derivados sintéticos en los que el conjugado produce un fosfolípido zwitteriónico cargado negativamente, tal es el caso

de metoxi-polietileno, glicol-diésteroil fosfatidiletanolamina (mPEG-DSPE); y (iii) fosfolípidos catiónicos, que incluyen, por ejemplo, fosfatidilcolina o esfingomielina de los cuales el fosfomonoéster fue O-metilado para formar los lípidos catiónicos.

La asociación del ácido nucleico con el portador lípido puede ocurrir, por ejemplo, mediante el llenado de espacios intersticiales del portador con el ácido nucleico, de manera que el portador atrapa físicamente al ácido nucleico, o mediante enlace covalente, iónico, o de hidrógeno, o por medio de adsorción mediante enlaces no específicos. Cualquiera que sea el modo de asociación, el ácido nucleico debe retener sus propiedades terapéuticas, es decir, las que codifican el antígeno.

El término "enfermedad" se refiere a una afección anormal que afecta el cuerpo de un individuo. Una enfermedad a menudo se interpreta como una afección médica que se asocia con signos y síntomas específicos. Una enfermedad puede causarse por factores originalmente de una fuente externa, tal como enfermedad infecciosa, o puede causarse por disfunciones internas, tal como enfermedades autoinmunitarias.

De acuerdo con la invención, el término "enfermedad" además se refiere a enfermedades cancerosas. Los términos "enfermedad cancerosa" o "cáncer" (término médico: neoplasia maligna) se refieren a una clase de enfermedades en las que un grupo de células muestra un crecimiento incontrolado (división más allá de los límites normales), invasión (intrusión y destrucción de tejidos adyacentes) y, a veces, metástasis (propagación a otras localizaciones en el cuerpo a través de la linfa o la sangre). Estas tres propiedades malignas de los cánceres los diferencian de los tumores benignos, que son autolimitados y no invaden ni hacen metástasis. La mayoría de los cánceres forman un tumor, es decir, una inflamación o lesión formada por un crecimiento anormal de células (llamadas células neoplásicas o células tumorales), pero algunos, como la leucemia, no lo hacen. Los ejemplos de cánceres incluyen, pero no se limitan a, carcinoma, linfoma, blastoma, sarcoma, glioma y leucemia. Más particularmente, los ejemplos de tales cánceres incluyen cáncer óseo, cáncer sanguíneo, cáncer de pulmón, cáncer de hígado, cáncer de páncreas, cáncer de piel, cáncer de la cabeza o cuello, melanoma maligno cutáneo o intraocular, cáncer uterino, cáncer de ovarios, cáncer rectal, cáncer de la región anal, cáncer de estómago, cáncer de colon, cáncer de mama, cáncer de próstata, cáncer uterino, carcinoma de los órganos sexuales y reproductivos, Enfermedad de Hodgkin, cáncer del esófago, cáncer del intestino delgado, cáncer del sistema endocrino, cáncer de la glándula tiroides, cáncer de la glándula paratiroides, cáncer de la glándula adrenal, sarcoma de tejidos blandos, cáncer de la vejiga, cáncer de los riñones, carcinoma celular renal, carcinoma de la pelvis renal, neoplasias del sistema nervioso central (CNS), cáncer neuroectodérmico, tumores del eje espinal, glioma, meningioma, y adenoma pituitario. El término "cáncer" de acuerdo con la invención también comprende, las metástasis del cáncer.

El término "enfermedad infecciosa" se refiere a cualquier enfermedad que puede transmitirse de individuo a individuo o de organismo a organismo, y es causada por un agente microbiano (por ejemplo, resfriado común). Los ejemplos de enfermedades infecciosas incluyen enfermedades infecciosas virales, tales como SIDA (HIV), hepatitis A, B o C, herpes, herpes zoster (varicela), sarampión alemán (virus de la rubéola), fiebre amarilla, dengue, etcétera, flavivirus, virus de la gripe, enfermedades infecciosas hemorrágicas (virus de Marburg o Ébola), y síndrome respiratorio agudo severo (SRAS), enfermedades infecciosas bacterianas, tales como la enfermedad del legionario (*Legionella*), enfermedades de transmisión sexual (por ejemplo clamidia o gonorrea), úlcera gástrica (*Helicobacter*), cólera (*Vibrio*), tuberculosis, difteria, infecciones por *E.coli*, *Staphylococci*, *Salmonella* o *Streptococci* (tetanus); infecciones por protozoarios patógenos tal como malaria, enfermedad del sueño, leishmaniasis; toxoplasmosis, es decir, infecciones por *Plasmodium*, *Trypanosoma*, *Leishmania* y *Toxoplasma*; o infecciones fúngicas, que son causadas, por ejemplo, por *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Coccidioides immitis*, *Blastomyces dermatitidis* o *Candida albicans*.

El término "enfermedad autoinmunitaria" se refiere a cualquier enfermedad en la cual el cuerpo produce una respuesta inmunogénica (es decir, sistema inmunitario) a algún componente de sus propios tejidos. En otras palabras, el sistema inmunitario pierde su capacidad de reconocer algunos tejidos o sistemas dentro del cuerpo como propios y direcciona y ataca estos como si fueran extraños. Las enfermedades autoinmunitarias pueden clasificarse en aquellas en las que predominantemente se afecta un órgano (por ejemplo, anemia hemolítica y tiroiditis autoinmunitaria), y aquellas en las que el proceso de la enfermedad autoinmunitaria se difunde a través de muchos tejidos (por ejemplo, lupus eritematoso sistémico). Por ejemplo, se cree que la esclerosis múltiple es causada por las células T que atacan las vainas que rodean las fibras nerviosas del cerebro y la médula espinal. Esto resulta en pérdida de la coordinación, debilidad, y visión borrosa. Las enfermedades autoinmunitarias son conocidas en la técnica e incluyen, por ejemplo, tiroiditis de Hashimoto, enfermedad de Grave, lupus, esclerosis múltiple, artritis reumatoide, anemia hemolítica, tiroiditis autoinmunitaria, lupus eritematoso sistémico, enfermedad celíaca, enfermedad de Crohn, colitis, diabetes, esclerodermia, psoriasis y similares.

De acuerdo con la invención, puede estimularse una respuesta inmunitaria mediante la introducción en un sujeto de un ARNm adecuado que codifica para un antígeno o un fragmento de este, por ejemplo, un antígeno asociado a la enfermedad.

El término "antígeno" se refiere a un agente que comprende un epítipo contra el que se genera una respuesta inmunitaria. El término "antígeno" incluye en particular proteínas, péptidos, polisacáridos, ácidos nucleicos, especialmente ARN y ADN, y nucleótidos. El término "antígeno" también incluye agentes, que se convierten en antigénicos - y sensibilizantes - solo a través de la transformación (por ejemplo de manera intermedia en la molécula o por terminación con la proteína del cuerpo). Un antígeno es, preferentemente, presentado por las células del sistema inmunitario, tales como las células presentadoras de antígeno, como las células dendríticas o los macrófagos. Adicionalmente, un antígeno o un producto procesado de éste, preferentemente, se reconoce por un receptor de células T o B, o por una molécula de

inmunoglobulina, tal como un anticuerpo. En una modalidad preferida, el antígeno es un antígeno asociado a la enfermedad, tal como un antígeno asociado a tumores, un antígeno viral, o un antígeno bacteriano.

El término "antígeno asociado a la enfermedad" se usa en su sentido más amplio para referirse a cualquier antígeno asociado con una enfermedad. Un antígeno asociado a la enfermedad es una molécula que contiene epítomos que estimularán el sistema inmunitario del huésped para producir una respuesta inmunitaria celular específica al antígeno y/o una respuesta humoral de anticuerpos contra la enfermedad. Por lo tanto, el antígeno asociado a la enfermedad puede usarse con fines terapéuticos. Los antígenos asociados a la enfermedad se asocian, preferentemente, con infección por microbios, típicamente, antígenos microbianos, o asociados con cáncer, típicamente, tumores.

El término "enfermedad que involucra un antígeno" se refiere a cualquier enfermedad que involucre un antígeno, por ejemplo, una enfermedad que se caracteriza por la presencia y/o la expresión de un antígeno. La enfermedad que involucra un antígeno puede ser una enfermedad infecciosa, una enfermedad autoinmunitaria, o una enfermedad cancerosa o simplemente cáncer. Como se mencionó anteriormente, el antígeno puede ser un antígeno asociado a la enfermedad, tal como un antígeno asociado a tumor, un antígeno viral, o un antígeno bacteriano.

En una modalidad, un antígeno asociado a la enfermedad es un antígeno asociado a tumor. En esta modalidad, la presente invención puede ser útil en el tratamiento del cáncer o de la metástasis del cáncer. Preferentemente, el órgano o tejido enfermo se caracteriza por células enfermas tales como células cancerosas que expresan un antígeno asociado a la enfermedad y/o se caracterizan por la asociación de un antígeno asociado a la enfermedad con su superficie. La inmunización con antígenos intactos o sustancialmente intactos asociados a tumores o fragmentos de estos, tales como péptidos o ácidos nucleicos de MHC clase I y clase II, en particular el ARNm, que codifica tal antígeno o fragmento hace posible producir una respuesta de tipo MHC clase I y/o una de clase II y, por lo tanto, estimular células T tales como linfocitos T citotóxicos CD8+, que son capaces de destruir las células cancerosas, y/o células T CD4+. Tal inmunización puede provocar, además, una respuesta inmunitaria humoral (respuesta de células B) que resulta en la producción de anticuerpos contra el antígeno asociado a tumor. Además, las células presentadoras de antígenos (APC) tales como las células dendríticas (DC) pueden cargarse con péptidos presentados por MHC de clase I mediante transfección con ácidos nucleicos que codifican antígenos tumorales *in vitro* y administrarse a un paciente. En una modalidad, el término "antígeno asociado a tumor" se refiere a un constituyente de las células cancerosas que puede derivarse del citoplasma, la superficie celular y el núcleo celular. En particular, se refiere a aquellos antígenos que se producen, preferentemente en gran cantidad, intracelularmente o como antígenos de superficie en las células tumorales. Los ejemplos de antígenos tumorales incluyen HER2, EGFR, VEGF, antígeno CAMPATH1, CD22, CA-125, HLA-DR, linfoma de Hodgkin o mucina-1, pero no se limitan a estos.

De acuerdo con la presente invención, un antígeno asociado a tumor comprende, preferentemente, cualquier antígeno que es característico para tumores o cánceres, así como también para células tumorales o cancerosas con respecto al tipo y/o nivel de expresión. En una modalidad, el término "antígeno asociado a tumor" se refiere a proteínas que en condiciones normales, es decir, en un sujeto saludable, se expresan específicamente en un número limitado de tejidos y/u órganos o en etapas específicas del desarrollo, por ejemplo, el antígeno asociado a tumor puede expresarse en condiciones normales específicamente en tejido de estómago, preferentemente, en la mucosa gástrica, en órganos reproductores, por ejemplo, en testículos, en tejido trofoblástico, por ejemplo, en la placenta, o en células de la línea germinal, y se expresan o se expresan de manera aberrante en uno o más tejidos tumorales o cancerosos. En este contexto, "un número limitado" significa, preferentemente, no más de 3, con mayor preferencia no más de 2 o 1. Los antígenos asociados a tumores en el contexto de la presente invención incluyen, por ejemplo, antígenos de diferenciación, preferentemente, antígenos de diferenciación de tipo celular específico, es decir, proteínas que en condiciones normales se expresan específicamente en un cierto tipo de célula, en cierta etapa de diferenciación, antígenos de cáncer/testículos, es decir, proteínas que en condiciones normales se expresan específicamente en testículos y algunas veces en la placenta, y antígenos específicos de línea germinal. En el contexto de la presente invención, el antígeno asociado a tumor, preferentemente, no se expresa, o solo rara vez, en tejidos normales o está mutado en células tumorales. Preferentemente, el antígeno asociado a tumor o la expresión aberrante del antígeno asociado a tumor identifica las células cancerosas. En el contexto de la presente invención, el antígeno asociado a tumor que se expresa por una célula cancerosa en un sujeto, por ejemplo, un paciente que padece una enfermedad cancerosa, es preferentemente una proteína propia en dicho sujeto. En modalidades preferidas, el antígeno asociado a tumor en el contexto de la presente invención se expresa en condiciones normales específicamente en un tejido u órgano que no es esencial, es decir, tejidos u órganos que cuando son dañados por el sistema inmunitario no conducen a la muerte del sujeto, o en órganos o estructuras del cuerpo que no son, o son difícilmente asequibles por el sistema inmunitario. Preferentemente, un antígeno asociado a tumor se presenta en el contexto de las moléculas de MHC por una célula cancerosa en la que se expresa.

Los ejemplos de antígenos de diferenciación que cumplen idealmente los criterios para antígenos asociados a tumor como se contempla por la presente invención como estructuras diana en la inmunoterapia tumoral, en particular, en la vacunación tumoral son las proteínas de superficie celular de la familia Claudina, tales como CLDN6 y CLDN18.2. Estos antígenos de diferenciación se expresan en tumores de diversos orígenes, y son particularmente adecuados como estructuras diana en relación con la inmunoterapia del cáncer mediada por anticuerpos, debido a su expresión selectiva (sin expresión en un tejido normal de toxicidad relevante) y a su localización en la membrana plasmática.

- Otros ejemplos de antígenos que pueden ser útiles en la presente invención son p53, ART-4, BAGE, beta-catenina/m, Bcr-abL CAMEL, CAP-1, CASP-8, CDC27/m, CDK4/m, CEA, CLAUDIN-12, c-MYC, CT, Cyp-B, DAM, ELF2M, ETV6-AML1, G250, GAGE, GnT-V, Gap100, HAGE, HER-2/neu, HPV-E7, HPV-E6, HAST-2, hTERT (o hTRT), LAGE, LDLR/FUT, MAGE-A, preferentemente, MAGE-A1, MAGE-A2, MAGE-A3, MAGE-A4, MAGE-A5, MAGE-A6, MAGE-A7, MAGE-A8, MAGE-A9, MAGE-A10, MAGE-A11, o MAGE-A12, MAGE-B, MAGE-C, MART-1/Melan-A, MC1R, Myosin/m, MUC1, MUM-1, -2, -3, NA88-A, NF1, NY-ESO-1, NY-BR-1, p190 menor BCR-abL, Pm1/RARa, PRAME, proteinasa 3, PSA, PSM, RAGE, RU1 o RU2, SAGE, SART-1 o SART-3, SCGB3A2, SCP1, SCP2, SCP3, SSX, SURVIVIN, TEL/AML1, TPI/m, TRP-1, TRP-2, TRP-2/INT2, TPTE y WT, preferentemente, WT-1.
- El término "antígeno viral" se refiere a cualquier componente viral que tiene propiedades antigénicas, es decir, que es capaz de provocar una respuesta inmunitaria en un individuo. El antígeno viral puede ser una ribonucleoproteína viral o una proteína de envoltura.
- El término "antígeno bacteriano" se refiere a cualquier componente bacteriano que tiene propiedades antigénicas, es decir, que es capaz de provocar una respuesta inmunitaria en un individuo. El antígeno bacteriano puede derivarse de la pared celular o de la membrana del citoplasma de la bacteria.
- El "procesamiento del antígeno" se refiere a la degradación de un antígeno en productos de procesamiento, que son fragmentos de dicho antígeno (por ejemplo, la degradación de una proteína en péptidos) y la asociación de uno o más de estos fragmentos (por ejemplo, mediante unión) con las moléculas MHC para la presentación por células, preferentemente células presentadoras de antígeno, a células T específicas.
- El término "respuesta inmunitaria", como se usa en la presente descripción, se refiere a una reacción del sistema inmunitario tal como a organismos inmunogénicos, tales como bacterias o virus, células o sustancias. El término "respuesta inmunitaria" incluye la respuesta inmunitaria innata y la respuesta inmunitaria adaptativa. Preferentemente, la respuesta inmunitaria está relacionada con una activación de células inmunitarias, una inducción de la biosíntesis de citocinas y/o la producción de anticuerpos. Se prefiere que la respuesta inmunitaria comprenda las etapas de activación de las células presentadoras de antígeno, tales como células dendríticas y/o macrófagos, la presentación de un antígeno o fragmento de este por dichas células presentadoras de antígeno y la activación de células T citotóxicas debido a esta presentación.
- El término "tratar" o "tratamiento" se refiere a cualquier tratamiento que mejora el estado de salud y/o prolonga (aumenta) la esperanza de vida de un individuo. Dicho tratamiento puede eliminar la enfermedad en un individuo, detener o enlentecer el desarrollo de una enfermedad en un individuo, inhibir o enlentecer el desarrollo de una enfermedad en un individuo, disminuir la frecuencia o gravedad de los síntomas en un individuo, y/o disminuir la recurrencia en un individuo quien actualmente tiene o quien previamente tuvo una enfermedad.
- En particular, el término "tratamiento de una enfermedad" incluye curar, acortar la duración, mejorar, retrasar o inhibir la progresión o el empeoramiento de una enfermedad o los síntomas de ésta.
- El término "inmunoterapia" se refiere a un tratamiento que involucra, preferentemente, una reacción inmunitaria específica y/o una(s) función inmunitaria efectora.
- El término "inmunización" o "vacunación" describe el proceso de tratamiento de un sujeto por razones terapéuticas o profilácticas.
- El término "sujeto" o "individuo", como se usa en la presente descripción, se refiere, preferentemente, a mamíferos. Por ejemplo, los mamíferos en el contexto de la presente invención son humanos, primates no humanos, animales domesticados tales como perros, gatos, ovejas, vacuno, cabras, cerdos, caballos, etcétera, animales de laboratorio tales como ratones, ratas, conejos, conejillos de indias, etcétera, así como también animales en cautiverio, tal como animales de los zoológicos. En una modalidad preferida, el sujeto es un humano.
- El término "célula presentadora de antígeno" (APC) se refiere a una célula de una variedad de células capaces de exponer, adquirir, y/o presentar al menos un antígeno o fragmento antigénico sobre (o en) su superficie celular. Las células presentadoras de antígeno pueden distinguirse en células presentadoras de antígeno profesionales y células presentadoras de antígeno no profesionales.
- El término "células presentadoras de antígeno profesionales" se refiere a las células presentadoras de antígeno que expresan constitutivamente las moléculas del complejo principal de histocompatibilidad clase II (MHC clase II) requeridas para la interacción con las células T vírgenes. Si una célula T interactúa con el complejo de la molécula MHC de clase II en la membrana de la célula presentadora de antígeno, la célula presentadora de antígeno produce una molécula coestimuladora que induce la activación de la célula T. Las células presentadoras de antígeno profesionales comprenden células dendríticas y macrófagos.
- El término "células presentadoras de antígenos no profesionales" se refiere a células presentadoras de antígenos que no expresan constitutivamente las moléculas de MHC de clase II, sino que son estimuladas por determinadas citocinas tales

como el interferón gamma. Las células presentadoras de antígeno no profesionales ilustrativas incluyen fibroblastos, células epiteliales tímicas, células epiteliales tiroideas, células gliales, células beta pancreáticas o células endoteliales vasculares.

5 El término "complejo principal de histocompatibilidad" y la abreviatura "MHC" incluyen las moléculas MHC de clase I y MHC de clase II y se refiere a un complejo de genes que se producen en todos los vertebrados. Las proteínas o moléculas del MHC son importantes para la señalización entre linfocitos y células presentadoras de antígeno o células enfermas en reacciones inmunitarias, en donde las proteínas o moléculas del MHC se unen a péptidos y los presentan para su reconocimiento por los receptores de células T. Las proteínas codificadas por el MHC se expresan en la superficie de las células y muestran tanto antígenos propios (fragmentos de péptidos de la propia célula) como antígenos no propios (por ejemplo, fragmentos de microorganismos invasores) a una célula T.

De acuerdo con la invención el término "receptor de antígeno quimérico (CAR)" es sinónimo de los términos "receptor de célula T quimérico" y "receptor de células T artificial".

15 Estos términos se refieren a receptores ingenierizados, que confieren una especificidad arbitraria tal como la especificidad de un anticuerpo monoclonal en una célula inmunitaria efectora tal como una célula T. De esta forma, pueden generarse un gran número de células T específicas de cáncer para la transferencia adoptiva de células. Por lo tanto, un CAR puede estar presente en las células T, por ejemplo, en lugar de, o además de, el propio receptor de células T de las células T. Tales células T no requieren necesariamente el procesamiento y la presentación de un antígeno para el reconocimiento de la célula diana si no que pueden reconocer preferentemente con especificidad cualquier antígeno presente en una célula diana. Preferentemente, dicho CAR se expresa en la superficie de las células. Para el propósito de la presente invención las células T que comprenden un CAR se comprenden por el término "célula T" como se usa en la presente descripción.

25 De acuerdo con la invención, el término "CAR" (o "receptor de antígeno quimérico") se refiere a un receptor artificial que comprende una sola molécula o un complejo de moléculas que reconoce, es decir, que se unen a, una estructura diana (por ejemplo, un antígeno) sobre una célula diana tal como una célula cancerosa (por ejemplo, mediante la unión de un dominio de unión a antígeno con un antígeno expresado en la superficie de la célula diana) y puede conferir especificidad a una célula efectora inmunitaria tal como una célula T que expresa dicho CAR en la superficie celular. Preferentemente, el reconocimiento de la estructura diana mediante un CAR resulta en la activación de una célula efectora inmunitaria que expresa dicho CAR. Un CAR puede comprender una o más unidades de proteína, dichas unidades de proteína comprenden uno o más dominios como se describió en la presente descripción. El término "CAR" no incluye los receptores de células T.

35 En una modalidad, un fragmento variable de cadena sencilla (scFv) derivado de un anticuerpo monoclonal se fusiona a CD3-zeta transmembranal y de endodominio. Tales moléculas resultan en la transmisión de una señal zeta en respuesta al reconocimiento por el scFv de su diana antigénica en una célula diana y la muerte de la célula diana que expresa el antígeno diana. Los dominios de reconocimiento de antígenos que pueden usarse además, incluyen entre otros las cadenas sencillas alfa y beta del receptor de célula T (TCR). De hecho, casi cualquier cosa que se una a una diana determinada con alta afinidad puede usarse como un dominio de reconocimiento de antígenos.

40 Seguido del reconocimiento de los antígenos, los receptores se agrupan y se transmite una señal a la célula. A este respecto, un "dominio de señalización de células T" es un dominio, preferentemente un endodominio, que transmite una señal de activación a la célula T después que se une el antígeno. El componente de endodominio más utilizado es CD3-zeta.

45 La terapia de transferencia adoptiva de células con células T ingenierizadas con CAR que expresan receptores de antígeno quiméricos es un terapéutico anticancerígeno prometedor debido a que las células T modificadas con CAR pueden ingenierizarse para hacer diana virtualmente cualquier antígeno tumoral. Por ejemplo, las células T de los pacientes pueden modificarse genéticamente (genéticamente modificadas) para expresar CAR dirigidos específicamente hacia los antígenos en las células tumorales de los pacientes, después se inyectan nuevamente al paciente.

50 De acuerdo con la invención, un CAR puede sustituir la función de un receptor de célula T y, en particular, puede conferir reactividad tal como la actividad citolítica a una célula tal como una célula T. Sin embargo, en contraste con la unión del receptor de células T a un complejo antígeno péptido-MHC, un CAR puede unirse a un antígeno, en particular cuando se expresa en la superficie celular.

55 De acuerdo con la invención, los CAR pueden comprender generalmente tres dominios. El primer dominio es el dominio de unión que reconoce y se une al antígeno. El segundo dominio es el dominio de coestimulación. El dominio de coestimulación sirve para potenciar la proliferación y supervivencia de los linfocitos citotóxicos después de la unión del CAR a una porción diana. La identidad del dominio de coestimulación está limitada solo porque este tiene la capacidad de potenciar la proliferación celular y la supervivencia después de la unión de la porción diana mediante el CAR. Los dominios de coestimulación adecuados incluyen CD28, CD137 (4-1BB), un miembro de la familia de receptores del factor de necrosis tumoral (TNF), CD134 (OX40), un miembro de la superfamilia de receptores TNFR y CD278 (ICOS), una molécula coestimuladora de la superfamilia CD28 expresada

en células T activadas. La persona experta comprenderá que las variantes de secuencia de estos dominios de coestimulación observados pueden usarse sin afectar negativamente a la invención, donde las variantes tienen la misma actividad o actividad similar que el dominio en el que se modelan. Tales variantes tendrán al menos aproximadamente 80 % de identidad de secuencia con la secuencia de aminoácidos del dominio del que se derivan. En algunas modalidades de la invención, las construcciones CAR comprenden dos dominios de coestimulación. Si bien las combinaciones particulares incluyen todas las variaciones posibles de los cuatro dominios señalados, los ejemplos específicos incluyen CD28+CD137 (4-1BB) y CD28+CD134 (OX40).

El tercer dominio es el dominio de señalización de activación (o dominio de señalización de células T). El dominio de señalización de activación sirve para activar linfocitos citotóxicos después de la unión del CAR al antígeno. La identidad del dominio de señalización de activación se limita solamente en que este tiene la capacidad de inducir la activación del linfocito citotóxico seleccionado después de la unión del antígeno por el CAR. Los dominios de señalización de activación adecuados incluyen la cadena CD3 [zeta] de células T y el receptor Fc [gamma]. El experto entenderá que las variantes de secuencia de estos dominios de señalización de activación observados pueden usarse sin afectar negativamente a la invención, donde las variantes tienen la misma actividad o actividad similar que el dominio en el que se modelan. Tales variantes tendrán al menos aproximadamente 80 % de identidad de secuencia con la secuencia de aminoácidos del dominio del que se derivan.

Los CAR pueden comprender los tres dominios, juntos en forma de una proteína de fusión. Tales proteínas de fusión generalmente comprenderán un dominio de unión, uno o más dominios de coestimulación, y un dominio de señalización de activación, unido en una dirección N-terminal a C-terminal. Sin embargo, los CAR no se limitan a esta disposición y otras disposiciones son aceptables e incluyen un dominio de unión, un dominio de señalización de activación, y uno o más dominios de coestimulación. Se debe entender que debido a que el dominio de unión debe ser libre para unirse al antígeno, la colocación del dominio de unión en la proteína de fusión generalmente será tal que se logre la exposición de la región sobre el exterior de la célula. De la misma manera, debido a que los dominios de coestimulación y señalización de activación sirven para inducir la actividad y la proliferación de los linfocitos citotóxicos, la proteína de fusión generalmente presentará estos dos dominios en el interior de la célula. Los CAR pueden incluir elementos adicionales, tales como un péptido señal para garantizar la exportación adecuada de la proteína de fusión a la superficie de las células, un dominio transmembrana para garantizar que la proteína de fusión se mantenga como una proteína integral de membrana y un dominio de bisagra (o región espaciadora) que imparte flexibilidad al dominio de unión y permite una unión fuerte al antígeno.

Las células usadas junto con el sistema CAR de la presente invención son preferentemente células T, en particular linfocitos citotóxicos, preferentemente seleccionados de células T citotóxicas, células asesinas naturales (NK), y células asesinas activadas por linfocinas (LAK). Después de la activación, cada uno de estos linfocitos citotóxicos desencadena la destrucción de las células diana. Por ejemplo, las células T citotóxicas desencadenan la destrucción de las células diana por uno o ambos de los siguientes medios. Primero, después de la activación, las células T liberan citotoxinas tales como perforina, granzimas, y granzulina. La perforina y la granzulina crean poros en la célula diana, y las granzimas entran en la célula y desencadenan una cascada de caspasas en el citoplasma que induce la apoptosis (muerte celular programada) de la célula. Segundo, la apoptosis puede inducirse mediante la interacción de los ligando Fas-Fas entre las células T y las células diana. Los linfocitos citotóxicos serán, preferentemente, células autólogas, aunque pueden usarse células heterólogas o células alogénicas.

Pueden usarse una diversidad de métodos para introducir construcciones CAR en las células T, que incluyen la transfección de ADN no viral, los sistemas basados en transposones y los sistemas basados en virus. La transfección de ADN no viral tiene un bajo riesgo de mutagénesis insercional. Los sistemas basados en transposones pueden integrar los transgenes de manera más eficiente que los plásmidos que no contienen un elemento de integración. Los sistemas basados en virus incluyen el uso de retrovirus y vectores lentivirales. Los retrovirus y son relativamente fáciles de producir, transducen de manera eficiente y permanente las células T, y preliminarmente han demostrado su seguridad desde el punto de vista de la integración en las células T humanas primarias. Los vectores lentivirales además transducen de manera eficiente y permanente las células T, pero su fabricación es más costosa. Además, son potencialmente más seguros que los sistemas basados en retrovirus.

El ARN descrito en la presente descripción (por ejemplo, que se obtiene mediante el uso de una molécula de ácido nucleico descrita en la presente descripción como un molde de transcripción) es útil además en la reprogramación o la desdiferenciación de células somáticas en células similares a las células madre, es decir, células que tienen características de células madre, *in vitro* o *in vivo*. Esto puede implicar la expresión transitoria de factores de reprogramación *in vitro* o *in vivo* para iniciar los procesos de reprogramación o desdiferenciación en las células. Por lo tanto, en una modalidad, el péptido o la proteína codificada por un ácido nucleico tal como el ARN descrito en la presente descripción es un factor que permite la reprogramación de células somáticas en células que tienen características de células madre. Las células similares a células madre pueden proporcionarse de acuerdo con la invención sin generar embriones o fetos. La desdiferenciación de las células somáticas en células que tienen características de células madre, en particular la pluripotencia, puede realizarse mediante la introducción del ARN codificante para factores que inducen la desdiferenciación de las células somáticas en las células somáticas (además denominados factores de transcripción de reprogramación (rTF)) y mediante el cultivo de las células somáticas que permiten que las células se desdiferencien. Después de desdiferenciarse, las células podrían inducirse a rediferenciarse en el mismo o diferente tipo de célula somática tal como neuronal, hematopoyética, muscular, epitelial, y otros tipos de células. Por lo tanto, tales células

similares a células madres tienen aplicaciones médicas para el tratamiento de enfermedades degenerativas mediante "terapia celular" y pueden usarse en estrategias terapéuticas novedosas en el tratamiento de trastornos cardíacos, neurológicos, endocrinos, vasculares, retinianos, dermatológicos, musculoesqueléticos, y otras enfermedades.

5 En consecuencia, la invención se refiere, además, a un método para proporcionar células que tienen características de células madre que comprende las etapas de (i) proporcionar una población celular que comprende células somáticas, (ii) introducir el ARN de la invención capaz de expresar uno o más factores que permiten la reprogramación de las células somáticas a células que tienen características de células madre en las células somáticas, y (iii) permitir el desarrollo de células que tienen características de células madre. En una modalidad, el método comprende, además, introducir en las
10 células somáticas miARN que potencia la reprogramación de las células somáticas a células que tienen características de células madre.

En una modalidad, el uno o más factores comprenden OCT4 y SOX2. El uno o más factores pueden comprender, además, KLF4 y/o c-MYC y/o NANOG y/o LIN28. En una modalidad, el uno o más factores comprenden OCT4, SOX2, KLF4 y c-MYC y pueden comprender, además, LIN28 y opcionalmente NANOG. En una modalidad, el uno o más factores
15 comprenden OCT4, SOX2, NANOG y LIN28.

En una modalidad, el método comprende, además, la etapa de cultivar las células somáticas en presencia de al menos un inhibidor de histona desacetilasa, en donde el al menos un inhibidor de histona desacetilasa comprende, preferentemente, ácido valproico, butirato de sodio, trichostatina A y/o scriptaid.
20

En una modalidad, la etapa (iii) comprende cultivar las células somáticas en condiciones de cultivo de células madre embrionarias.

25 En una modalidad, las características de las células madre comprenden una morfología de células madre embrionarias.

En una modalidad, las células que tienen características de células madre tienen cariotipos normales, expresan actividad de telomerasa, expresan marcadores de superficie celular que son característicos para células madre embrionarias y/o expresan genes que son característicos para células madre embrionarias.
30

En una modalidad, las células que tienen características de células madre exhiben un estado pluripotente.

En una modalidad, las células que tienen características de células madre tienen el potencial de desarrollo para diferenciarse en derivados avanzados de todas las tres capas germinales primarias.
35

En una modalidad, las células somáticas son fibroblastos tales como fibroblastos de pulmón, fibroblastos de prepucio o fibroblastos dérmicos. Preferentemente, las células somáticas son células humanas.

En una modalidad, el ARN se introduce en las células somáticas mediante electroporación o lipofección. En una modalidad, el ARN se introduce en las células somáticas repetitivamente.
40

En una modalidad, la introducción de ARN capaz de expresar determinados factores, como se describió en la presente descripción, en células somáticas resulta en la expresión de dichos factores durante un período prolongado de tiempo, preferentemente, durante al menos 10 días, preferentemente, durante al menos 11 días y con mayor preferencia durante al menos 12 días. Para lograr tal expresión a largo plazo, el ARN se introduce preferentemente periódicamente (es decir, repetitivamente) en las células más de una vez, preferentemente, mediante el uso de electroporación. Preferentemente, el ARN se introduce en las células al menos dos veces, con mayor preferencia al menos 3 veces, con mayor preferencia al menos 4 veces, incluso con mayor preferencia al menos 5 veces hasta, preferentemente, 6 veces, con mayor preferencia hasta 7 veces o incluso hasta 8, 9 o 10 veces, preferentemente durante un período de tiempo de al menos 10 días, preferentemente durante al menos 11 días y con mayor preferencia durante al menos 12 días para asegurar la expresión de uno o más factores durante un período de tiempo prolongado. Preferentemente, los períodos de tiempo transcurridos entre las introducciones repetidas del ARN son de 24 horas a 120 horas, preferentemente de 48 horas a 96 horas. En una modalidad, los períodos de tiempo que transcurren entre las introducciones repetidas del ARN no son más de 72 horas, preferentemente no más de 48 horas o 36 horas. En una modalidad, antes de la siguiente electroporación, se permite que las células se recuperen de la electroporación previa. En cualquier caso, las condiciones deben seleccionarse de manera que los factores se expresen en las células en cantidades y durante períodos de tiempo que sustenten el proceso de reprogramación.
45
50
55

Una "célula madre" es una célula con la capacidad de autorrenovarse, permanecer indiferenciada, y diferenciarse. Una célula madre puede dividirse sin límite, durante al menos el tiempo de vida del animal en el que reside naturalmente. Una célula madre no está terminalmente diferenciada; no está en la etapa final de una vía de diferenciación. Cuando una célula madre se divide, cada célula hija puede permanecer como una célula madre o embarcarse en un curso que conduce a la diferenciación terminal.
60

Las células madre totipotentes son células que tienen propiedades de diferenciación totipotencial y que son capaces de desarrollarse en un organismo completo. Esta propiedad es poseída por las células hasta la etapa de 8 células después
65

de la fertilización del ovocito por los espermatozoides. Cuando estas células se aíslan y se trasplantan al útero, pueden convertirse en un organismo completo.

- 5 Las células madre pluripotentes son células capaces de desarrollarse en diversas células y tejidos derivados de las capas ectodérmica, mesodérmica y endodérmica. Las células madre pluripotentes que se derivan de la masa celular interna ubicada en el interior del blastocisto, que se generan 4-5 días después de la fertilización se denominan "células madre embrionarias" y pueden diferenciarse en células de diversos tejidos, pero no pueden formar nuevos organismos vivos.
- 10 Las células madre multipotentes son células madre que se diferencian normalmente solo en tipos de células específicas de su tejido y órgano de origen. Las células madre multipotentes están involucradas no solo en el crecimiento y el desarrollo de diversos tejidos y órganos durante los períodos fetal, neonatal y adulto, sino además en el mantenimiento de la homeostasis del tejido adulto y la función de inducir la regeneración después del daño tisular. Las células multipotentes específicas de tejido se denominan colectivamente "células madre adultas".
- 15 Una "célula madre embrionaria" o "ESC" es una célula madre que está presente o se aísla de un embrión. Puede ser pluripotente, que tiene la capacidad de diferenciarse en todas y cada una de las células presentes en el organismo, o multipotente, con la capacidad de diferenciarse en más de un tipo de célula.
- 20 Como se usa en la presente descripción, "embrión" se refiere a un animal en las primeras etapas de su desarrollo. Estas etapas se caracterizan por la implantación y la gastrulación, donde se definen y establecen las tres capas germinales y mediante la diferenciación de las capas germinales en los órganos y sistemas de órganos respectivos. Las tres capas germinales son el endodermo, el ectodermo y el mesodermo.
- 25 Un "blastocisto" es un embrión en una etapa temprana del desarrollo en el que el óvulo fertilizado experimenta la escisión, y se forma o se ha formado una capa esférica de células que rodean una cavidad llena de líquido. Esta capa esférica de células es el trofoectodermo. Dentro del trofoectodermo hay un grupo de células denominado masa celular interna (ICM). El trofoectodermo es el precursor de la placenta, y la ICM es la precursora del embrión.
- 30 Una célula madre adulta, llamada además célula madre somática, es una célula madre que se encuentra en un adulto. Una célula madre adulta se encuentra en un tejido diferenciado, puede autorrenovarse, y puede diferenciarse, con algunas limitaciones, para producir tipos de células especializadas de su tejido de origen. Los ejemplos incluyen células madre mesenquimales, células madre hematopoyéticas, y células madre neurales.
- 35 Una "célula diferenciada" es una célula madura que ha experimentado cambios progresivos en el desarrollo a una forma o función más especializada. La diferenciación celular es el proceso que experimenta una célula a medida que madura a un tipo de célula claramente especializado. Las células diferenciadas tienen características distintas, realizan funciones específicas, y tienen menos probabilidades de dividirse que sus contrapartes menos diferenciadas.
- 40 Una célula "indiferenciada", por ejemplo, una célula inmadura, embrionaria, o primitiva, típicamente tiene una apariencia inespecífica, puede realizar múltiples actividades no específicas, y puede desempeñarse mal, si es que lo hace, en funciones realizadas típicamente por células diferenciadas.
- 45 "Célula somática" se refiere a todas y cada una de las células diferenciadas y no incluye células madre, células germinales, o gametos. Preferentemente, "célula somática" como se usa en la presente descripción se refiere a una célula diferenciada terminalmente.
- 50 Como se usa en la presente descripción, "comprometido" se refiere a las células que se consideran comprometidas permanentemente con una función específica. Las células comprometidas además se denominan "células diferenciadas terminalmente".
- 55 Como se usa en la presente descripción, "diferenciación" se refiere a la adaptación de las células para una forma o función particular. En las células, la diferenciación conduce a una célula más comprometida.
- 60 Como se usa en la presente descripción, "desdiferenciación" se refiere a la pérdida de especialización en forma o función. En las células, la desdiferenciación conduce a una célula menos comprometida.
- 65 Como se usa en la presente descripción, "reprogramación" se refiere al restablecimiento del programa genético de una célula. Una célula reprogramada exhibe, preferentemente, pluripotencia.
- 60 Los términos "desdiferenciada" y "reprogramada" o términos similares se usan indistintamente en la presente descripción para designar células derivadas de células somáticas que tienen características de células madre. Sin embargo, dichos términos no pretenden limitar el tema descrito en la presente descripción por consideraciones mecánicas o funcionales.
- 65 El término "ARN que induce el desarrollo de características de células madre" o "ARN capaz de expresar uno o más factores que permiten la reprogramación de las células somáticas a células que tienen características de células madre" se refiere al ARN que cuando se introduce en una célula somática induce a la célula a desdiferenciarse.

Como se usa en la presente descripción, "célula germinal" se refiere a una célula reproductiva tal como un espermatozoido o un ovocito, o una célula que se convertirá en una célula reproductiva.

Como se usa en la presente descripción, "pluripotente" se refiere a células que pueden dar lugar a cualquier tipo de célula, excepto las células de la placenta u otras células de soporte del útero.

Los términos tales como "célula que tiene características de células madre", "célula que tiene propiedades de células madre" o "célula similar a células madre" se usan en la presente descripción para designar células que, aunque derivan de células somáticas diferenciadas que no son células madre, exhiben una o más características típicas de células madre, en particular de células madre embrionarias. Tales características incluyen una morfología de células madre embrionarias, tal como colonias compactas, alta relación de núcleo a citoplasma y nucléolos prominentes, cariotipos normales, expresión de actividad de telomerasa, expresión de marcadores de superficie celular que son característicos de células madre embrionarias, y/o la expresión de genes que son característicos de las células madre embrionarias. Los marcadores de superficie celular que son característicos de las células madre embrionarias se seleccionan, por ejemplo, del grupo que consiste en antígeno embrionario específico de etapa 3 (SSEA-3), SSEA-4, antígeno relacionado con tumores 1-60 (TRA-1-60), TRA-1-81, y TRA-2-49/6E. Los genes que son característicos de las células madre embrionarias se seleccionan, por ejemplo, del grupo que consiste en OCT4 endógeno, NANOG endógeno, factor de crecimiento y diferenciación 3 (GDF3), expresión reducida 1 (REX1), factor de crecimiento de fibroblastos 4 (FGF4), gen específico de células embrionarias 1 (ESG1), asociado al desarrollo pluripotencial 2 (DPPA2), DPPA4, y telomerasa transcriptasa inversa (TERT). En una modalidad, la una o más características típicas de las células madre incluyen la pluripotencia.

En una modalidad de la invención, las características de las células madre comprenden una morfología de células madre embrionarias, en donde dicha morfología de células madre embrionarias comprende, preferentemente, criterios morfológicos seleccionados del grupo que consiste en colonias compactas, alta relación de núcleo a citoplasma y nucléolos prominentes. En determinadas modalidades, las células que tienen características de células madre tienen cariotipos normales, expresan actividad de telomerasa, expresan marcadores de superficie celular que son característicos de células madre embrionarias y/o expresan genes que son característicos de células madre embrionarias. Los marcadores de superficie celular que son característicos de las células madre embrionarias pueden seleccionarse del grupo que consiste en antígeno embrionario específico de etapa 3 (SSEA-3), SSEA-4, antígeno relacionado con tumores 1-60 (TRA-1-60), TRA-1-81, y TRA-2-49/6E y los genes que son característicos de las células madre embrionarias pueden seleccionarse del grupo que consiste en OCT4 endógeno, NANOG endógeno, factor de crecimiento y diferenciación 3 (GDF3), expresión reducida 1 (REX1), factor de crecimiento de fibroblastos 4 (FGF4), gen específico de células embrionarias 1 (ESG1), asociado al desarrollo pluripotencial 2 (DPPA2), DPPA4 y telomerasa transcriptasa inversa (TERT).

Preferentemente, las células que tienen características de células madre son células somáticas desdiferenciadas y/o reprogramadas. Preferentemente, las células que tienen características de células madre exhiben las características esenciales de las células madre embrionarias tal como un estado pluripotente. Preferentemente, las células que tienen características de células madre tienen el potencial de desarrollo para diferenciarse en derivados avanzados de todas las tres capas germinales primarias. En una modalidad, la capa germinal primaria es el endodermo y el derivado avanzado es el tejido epitelial similar al intestino. En una modalidad adicional, la capa germinal primaria es el mesodermo y el derivado avanzado es el músculo estriado y/o el cartílago. Aún en una modalidad adicional, la capa germinal primaria es el ectodermo y el derivado avanzado es el tejido neural y/o el tejido epidérmico. En una modalidad preferida, las células que tienen características de células madre tienen el potencial de desarrollo para diferenciarse en células neuronales y/o células cardíacas.

En una modalidad, las células somáticas son células madre embrionarias derivadas de células somáticas con un fenotipo mesenquimatoso. En una modalidad preferida, las células somáticas son fibroblastos tales como fibroblastos fetales o fibroblastos postnatales o queratinocitos, preferentemente, queratinocitos derivados del folículo piloso. En modalidades adicionales, los fibroblastos son fibroblastos de pulmón, fibroblastos de prepucio o fibroblastos dérmicos. En modalidades particulares, los fibroblastos son fibroblastos como los depositados en la Colección Americana de Tipos de Cultivos (ATCC) bajo el núm. de Catálogo CCL-186, como los depositados en la Colección Americana de Tipos de Cultivos (ATCC) bajo el núm. de Catálogo CRL-2097 o como los depositados en la Colección Americana de Tipos de Cultivos (ATCC) bajo el núm. de Catálogo CRL-2522, o como los distribuidos por System Biosciences bajo el núm. de catálogo. PC501A-HFF. En una modalidad, los fibroblastos son fibroblastos dérmicos humanos adultos. Preferentemente, las células somáticas son células humanas. De acuerdo con la presente invención, las células somáticas pueden modificarse genéticamente.

El término "factor" de acuerdo con la invención cuando se usa junto con la expresión de este por ARN incluye proteínas y péptidos, así como también derivados y variantes de estos. Por ejemplo, el término "factor" comprende OCT4, SOX2, NANOG, LIN28, KLF4 y c-MYC.

Los factores pueden ser de cualquier especie animal; por ejemplo, mamíferos y roedores. Los ejemplos de mamíferos incluyen, pero no se limitan a humanos y primates no humanos. Los primates incluyen, pero no se limitan a humanos, chimpancés, babuinos, monos cynomolgus, y cualquiera de los otros monos del Nuevo o Viejo Mundo. Los roedores incluyen, pero no se limitan a ratón, rata, conejillo de indias, hámster y jerbo.

De acuerdo con la presente invención, uno o más factores capaces de permitir la reprogramación de células somáticas a células que tienen características de células madre comprenden un conjunto de factores seleccionados del grupo que consiste en (i) OCT4 y SOX2, (ii) OCT4, SOX2 y uno o ambos de NANOG y LIN28, (iii) OCT4, SOX2 y uno o ambos de KLF4 y c-MYC. En una modalidad, dicho uno o más factores capaces de expresarse por el ARN comprenden OCT4, SOX2, NANOG y LIN28 u OCT4, SOX2, KLF4 y c-MYC. Preferentemente, el ARN se introduce en dichas células somáticas mediante electroporación o microinyección. Preferentemente, la invención comprende, además, permitir el desarrollo de células que tienen características de células madre, por ejemplo, mediante el cultivo de la célula somática en condiciones de cultivo de células madre embrionarias, preferentemente, en condiciones adecuadas para mantener células madre pluripotentes en un estado indiferenciado.

OCT4 es un factor de transcripción de los factores de transcripción eucariota POU y un indicador de pluripotencia de las células madre embrionarias. Es una proteína de unión a Octómero expresada por vía materna. Se ha observado que está presente en los ovocitos, la masa celular interna de los blastocitos y además en la célula germinal primordial. El gen *POU5F1* codifica la proteína OCT4. Los sinónimos del nombre del gen incluyen *OCT3*, *OCT4*, *OTF3* y *MGC22487*. La presencia de OCT4 en concentraciones específicas es necesaria para que las células madre embrionarias permanezcan indiferenciadas. Preferentemente, la "proteína OCT4" o simplemente "OCT4" se refiere a OCT4 humano.

Sox2 es un miembro de la familia de genes Sox (caja HMG relacionada con SRY) que codifica los factores de transcripción con un único dominio de unión al ADN de HMG. Se ha encontrado que SOX2 controla las células progenitoras neurales mediante la inhibición de su capacidad para diferenciarse. La represión del factor resulta en la delaminación de la zona ventricular, que es seguida por una salida del ciclo celular. Estas células, además, comienzan a perder su carácter progenitor a través de la pérdida de marcadores progenitores y de diferenciación neuronal temprana. Preferentemente, la "proteína SOX2" o simplemente "SOX2" se refiere a SOX2 humano.

NANOG es un gen de homeodominio de tipo NK-2, y se ha propuesto que desempeña un papel clave en el mantenimiento de la pluripotencia de las células madre presumiblemente mediante la regulación de la expresión de genes críticos para la renovación y diferenciación de células madre embrionarias. NANOG se comporta como un activador de la transcripción con dos dominios de activación inusualmente fuertes integrados en su extremo terminal C. La disminución de la expresión de NANOG induce la diferenciación de las células madre embrionarias. Preferentemente, la "proteína NANOG" o simplemente "NANOG" se refiere a NANOG humano.

LIN28 es una proteína citoplasmática conservada con un apareamiento inusual de motivos de unión a ARN: un dominio de choque frío y un par de dedos de zinc CCHC de tipo retroviral. En los mamíferos, es abundante en diversos tipos de células indiferenciadas. En células pluripotentes de mamíferos, se observó LIN28 en complejos sensibles a RNasa con la Proteína de Unión a poli(A), y en fracciones polisomales de gradientes de sacarosa, lo que sugiere que está asociado con la traducción de los ARNm. Preferentemente, la "proteína LIN28" o simplemente "LIN28" se refiere a LIN28 humano.

El factor similar a Krueppel (KLF4) es un factor de transcripción de dedos de zinc, que se expresa fuertemente en las células epiteliales postmitóticas de diferentes tejidos, por ejemplo, el colon, el estómago y la piel. KLF4 es esencial para la diferenciación terminal de estas células y se involucra en la regulación del ciclo celular. Preferentemente, la "proteína KLF4" o simplemente "KLF4" se refiere a KLF4 humano.

MYC (cMYC) es un protooncogén, que se sobreexpresa en una amplia gama de cánceres humanos. Cuando es mutado específicamente, o se sobreexpresa, aumenta la proliferación celular y funciona como un oncogén. El gen MYC codifica para un factor de transcripción que regula la expresión del 15 % de todos los genes a través de la unión en las secuencias de Caja Potenciadora (Cajas-E) y el reclutamiento de las histonas acetiltransferasas (HAT). MYC pertenece a la familia de factores de transcripción MYC, que además incluye los genes N-MYC y L-MYC. Los factores de transcripción de la familia MYC contienen el dominio bHLH/LZ (cremallera básica de leucina Hélice-Bucle-Hélice). Preferentemente, la "proteína cMYC" o simplemente "cMYC" se refiere a cMYC humano.

Una referencia, en la presente descripción a factores específicos tales como OCT4, SOX2, NANOG, LIN28, KLF4 o c-MYC debe comprenderse a fin de que además se incluyan todas las variantes de estos factores. En particular, debe comprenderse a fin de que además se incluyan todas las variantes de empalme, variantes modificadas postraduccionalmente, conformaciones, isoformas y especies homólogas de estos factores que las células expresan de manera natural.

El término "miARN" (microARN) se relaciona con los ARN no codificantes de 21-23 nucleótidos de longitud que se encuentran en las células eucariotas que, mediante la inducción de la degradación y/o la prevención de la traducción de los ARNm diana, modulan una gran cantidad de funciones celulares, que incluyen las relacionadas con la autorenovación/diferenciación de las ESC y progresión del ciclo celular. Los miARN son reguladores postranscripcionales que se unen a secuencias complementarias en los transcritos de ARN mensajero diana (ARNm), que generalmente resulta en la represión de la traducción o en la degradación de la diana y en el silenciamiento génico. Se ha encontrado que los miARN en la combinación correcta son capaces de inducir la reprogramación celular directa de células somáticas a células que tienen características de células madre in vitro. Por ejemplo, se ha observado que el conglomerado de miARN 302-367 potencia la reprogramación de células somáticas.

Preferentemente, la etapa de permitir el desarrollo de células que tienen características de células madre comprende cultivar las células somáticas en condiciones de cultivo de células madre embrionarias, preferentemente condiciones adecuadas para mantener células madre pluripotentes en un estado indiferenciado.

5 Preferentemente, para permitir el desarrollo de células que tienen características de células madre, las células se cultivan en presencia de uno o más inhibidores de la ADN metiltransferasa y/o uno o más inhibidores de la histona desacetilasa. Los compuestos preferidos se seleccionan del grupo que consiste en 5'-azacitidina (5'-azaC), ácido suberoilánilida hidroxámico (SAHA), dexametasona, tricostatina A (TSA), butirato de sodio (NaBu), Scriptaid y ácido valproico (VPA). Preferentemente, las células se cultivan en presencia de ácido valproico (VPA), preferentemente, en una concentración
10 de entre 0,5 y 10 mM, con mayor preferencia entre 1 y 5 mM, con la máxima preferencia en una concentración de aproximadamente 2 mM.

Los métodos de la presente invención pueden usarse para efectuar la desdiferenciación de cualquier tipo de célula somática. Las células que pueden usarse incluyen células que pueden ser desdiferenciadas o reprogramadas por los
15 métodos de la presente invención, particularmente células que están total o parcialmente diferenciadas, con mayor preferencia completamente diferenciadas. Preferentemente, la célula somática es una célula diploide derivada de organismos multi-celulares pre-embrionarios, embrionarios, fetales y postnatales. Los ejemplos de células, que pueden usarse incluyen pero no se limitan a fibroblastos, tales como fibroblastos fetales y neonatales o fibroblastos adultos, queratinocitos, en particular queratinocitos primarios, con mayor preferencia queratinocitos derivados de pelo, células
20 adiposas, células epiteliales, células epidérmicas, condrocitos, células del cúmulus, células neurales, células gliales, astrocitos, células cardíacas, células esofágicas, células musculares, melanocitos, células hematopoyéticas, osteocitos, macrófagos, monocitos y células mononucleares.

Las células con las que los métodos de la invención pueden usarse pueden ser de cualquier especie animal; por ejemplo, mamíferos y roedores. Los ejemplos de células de mamífero que pueden ser desdiferenciadas y re-diferenciadas por la
25 presente invención incluyen, pero no se limitan a células de primates humanos y no humanos. Las células de primates con la que la invención puede realizarse incluyen pero no se limitan a las células de seres humanos, chimpancés, babuinos, monos cynomolgus, y cualquiera de los otros monos del Nuevo o Viejo Mundo. Las células de roedores con la que la invención puede realizarse incluyen, pero no se limitan a células de ratón, rata, conejillo de indias, hámster y jerbo.
30

Las células desdiferenciadas preparadas de acuerdo con la presente invención se espera que muestren muchos de los mismos requisitos que las células madre pluripotentes y puedan expandirse y mantenerse en condiciones usadas para las células madre embrionarias, por ejemplo, medio celular ES o cualquier medio que sustente el crecimiento de las células embrionarias. Las células madre embrionarias retienen su pluripotencia *in vitro* cuando se mantienen sobre fibroblastos fetales inactivados tales como fibroblastos embrionarios de ratón o fibroblastos humanos irradiados (por ejemplo, fibroblastos de prepucio humano, fibroblastos de piel humana, fibroblastos de endometrio humano, fibroblastos de oviducto humano) en cultivo. En una modalidad, las células alimentadoras humanas pueden ser células alimentadoras autólogas derivadas del mismo cultivo de células reprogramadas por diferenciación directa.
35

Además, las células madre embrionarias humanas pueden propagarse con éxito sobre Matrigel en un medio condicionado con fibroblastos fetales de ratón. Las células madre humanas pueden crecer en cultivo durante un período de tiempo prolongado y permanecer indiferenciadas en condiciones de cultivo específicas.
40

En determinadas modalidades, las condiciones de cultivo celular pueden incluir poner en contacto las células con factores que pueden inhibir la diferenciación o de cualquier otra manera potenciar la desdiferenciación de las células, por ejemplo, prevenir la diferenciación de células en células no-ES, trofoectodermo u otros tipos de células.
45

Las células desdiferenciadas preparadas de acuerdo con la presente invención pueden evaluarse mediante los métodos que incluyen supervisar los cambios en el fenotipo de las células y caracterizar su expresión génica y proteica. La expresión génica puede determinarse mediante RT-PCR, y los productos de traducción pueden determinarse mediante inmunocitoquímica e transferencia Western. Particularmente, las células desdiferenciadas pueden caracterizarse para determinar el patrón de expresión génica y si las células reprogramadas presentan un patrón de expresión génica similar al patrón de expresión esperado de las células de control pluripotentes indiferenciadas, tal como las células madre embrionarias, mediante el uso de procedimientos bien conocidos en la técnica, que incluyen la transcriptómica.
50

La expresión de los siguientes genes de células desdiferenciadas puede evaluarse al respecto: OCT4, NANOG, factor de crecimiento y diferenciación 3 (GDF3), expresión reducida 1 (REX1), factor de crecimiento de fibroblastos 4 (FGF4), gen específico de la célula embrionaria 1 (ESG1), asociado al desarrollo pluripotencial 2 (DPPA2), DPPA4, telomerasa transcriptasa inversa (TERT), antígeno embrionario 3 (SSEA-3), SSEA-4, antígeno asociado a tumor-1-60 (TRA-1-60), TRA-1-81, y TRA-2-49/6E.
55
60

Las células madre embrionarias o indiferenciadas con las que pueden compararse las células reprogramadas pueden ser de la misma especie que las células somáticas diferenciadas. Alternativamente, las células madre indiferenciadas o embrionarias con las que pueden compararse las células reprogramadas pueden ser de una especie diferente como las células somáticas diferenciadas.
65

En algunas modalidades, existe una similitud en el patrón de expresión génica entre una célula reprogramada y una célula indiferenciada, por ejemplo, célula madre embrionaria, si determinados genes expresados específicamente en una célula indiferenciada se expresan, además, en la célula reprogramada. Por ejemplo, determinados genes, por ejemplo, telomerasa, que son típicamente indetectables en las células somáticas diferenciadas pueden usarse para controlar la extensión de la reprogramación. Del mismo modo, para determinados genes, la ausencia de expresión puede usarse para evaluar la extensión de la reprogramación.

La capacidad de auto-renovación, marcada por la inducción de la actividad de la telomerasa, es otra de las características de las células madre que pueden supervisarse en las células desdiferenciadas.

El análisis del cariotipo puede realizarse por medio de las extensiones de los cromosomas a partir de las células mitóticas, cariotipado espectral, ensayos de la longitud del telómero, hibridación genómica total, u otros procedimientos bien conocidos en la técnica.

Mediante el uso de la presente invención, el ARN codificante de los factores apropiados se incorpora en una o más células somáticas, por ejemplo, mediante electroporación. Después de la incorporación, las células se cultivan preferentemente mediante el uso de condiciones que sustentan el mantenimiento de las células desdiferenciadas (es decir, condiciones de cultivo de las células madre). Las células desdiferenciadas pueden expandirse después e inducir que se re-diferencien en diferentes tipos de células somáticas que se necesitan para la terapia celular. Las células desdiferenciadas que se obtienen de acuerdo con la presente invención pueden inducirse a diferenciarse, *in vitro* o *in vivo*, en uno o más tipos de células somáticas deseadas.

Preferentemente, las células desdiferenciadas que se obtienen de acuerdo con la presente invención pueden dar lugar a células de cualquiera de las tres capas germinales embrionarias, es decir, endodermo, mesodermo, y ectodermo. Por ejemplo, las células desdiferenciadas pueden diferenciarse en músculo esquelético, esqueleto, dermis de la piel, tejido conectivo, sistema urogenital, corazón, sangre (células linfáticas), y bazo (mesodermo); estómago, colon, hígado, páncreas, vejiga urinaria; revestimiento de la uretra, partes epiteliales de la tráquea, pulmones, faringe, tiroides, paratiroides, intestino (endodermo); o del sistema nervioso central, retina y cristalino, cráneo y sensorial, ganglios y nervios, células de pigmento, tejido conectivo de la cabeza, epidermis, pelo, glándulas mamarias (ectodermo). Las células desdiferenciadas que se obtienen de acuerdo con la presente invención pueden re-diferenciarse *in vitro* o *in vivo* mediante el uso de procedimientos conocidos en la técnica.

En una modalidad de la presente invención, las células reprogramadas que resultan de los métodos de esta invención se usan para producir una progenie diferenciada. Por lo tanto, en un aspecto, la presente invención proporciona un método para producir células diferenciadas, que comprende: (i) obtener células reprogramadas mediante el uso de los métodos de esta invención; e (ii) inducir la diferenciación de las células reprogramadas para producir células diferenciadas. La etapa (ii) puede realizarse *in vivo* o *in vitro*. Además, la diferenciación puede inducirse a través de la presencia de factores de diferenciación apropiados, que pueden o bien añadirse o están presentes *in situ*, por ejemplo, en un cuerpo, un órgano o tejido en el que se introdujeron las células reprogramadas. Las células diferenciadas pueden usarse para derivar las células, tejidos y/u órganos que se usan ventajosamente en el área de trasplante de células, tejidos y/u órganos. Si se desea, las modificaciones genéticas pueden introducirse, por ejemplo, en células somáticas antes de la reprogramación. Las células diferenciadas de la presente invención preferentemente no poseen la pluripotencia de una célula madre embrionaria, o una célula germinal embrionaria, y son, en esencia, células específicas de tejido parcial o totalmente diferenciadas.

Una ventaja de los métodos de la presente invención es que las células reprogramadas que se obtienen mediante la presente invención pueden diferenciarse sin previa selección o purificación o establecimiento de una línea celular. En consecuencia, en determinadas modalidades, una población heterogénea de células que comprenden células reprogramadas se diferencian en un tipo celular deseado. En una modalidad, una mezcla de células que se obtienen a partir de los métodos de la presente invención se expone a uno o más factores de diferenciación y se cultivan *in vitro*.

Los métodos de diferenciación de las células reprogramadas que se obtienen por los métodos descritos en la presente descripción pueden comprender una etapa de permeabilización de la célula reprogramada. Por ejemplo, las células generadas mediante las técnicas de reprogramación descritas en la presente descripción, o alternativamente una mezcla heterogénea de células que comprenden células reprogramadas, pueden permeabilizarse antes de la exposición a uno o más factores de diferenciación o extracto celular u otra preparación que comprende factores de diferenciación.

Por ejemplo, las células diferenciadas pueden obtenerse mediante el cultivo de las células reprogramadas indiferenciadas en presencia de al menos un factor de diferenciación y seleccionarse de células diferenciadas a partir del cultivo. La selección de células diferenciadas puede basarse en el fenotipo, tal como la expresión de ciertos marcadores celulares presentes en las células diferenciadas, o mediante ensayos funcionales (por ejemplo, la capacidad de realizar una o más funciones de un tipo de célula diferenciada en particular).

En otra modalidad, las células reprogramadas de acuerdo con la presente invención se modifican genéticamente mediante la adición, delección, o modificación de su(s) secuencia(s) de ADN.

Las células reprogramadas o desdiferenciadas preparadas de acuerdo con la presente invención o las células derivadas de las células reprogramadas o desdiferenciadas son útiles en la investigación y en la terapia. Las células pluripotentes reprogramadas pueden diferenciarse en cualquiera de las células en el cuerpo, que incluyen, sin limitación, células de la piel, cartilago, músculo esquelético óseo, músculo cardíaco, renales, hepáticas, sangre y formadoras de la sangre, precursoras vasculares y endotelio vascular, beta pancreáticas, neuronas, glías, retinales, neuronales, intestinales, pulmonares, y hepáticas.

Las células reprogramadas son útiles para la terapia regenerativa/reparadora y pueden trasplantarse a un paciente que lo necesite. En una modalidad, las células son autólogas con el paciente.

Las células reprogramadas proporcionadas de acuerdo con la presente invención pueden usarse, por ejemplo, en las estrategias terapéuticas en el tratamiento de trastornos cardíacos, neurológicos, endocrinos, vasculares, retineanos, dermatológicos, muscular-esqueléticos, y otras enfermedades.

Por ejemplo, y sin pretender ser una limitación, las células reprogramadas de la presente invención pueden usarse para reponer las células en animales cuyas células naturales se agotaron debido a la edad o a la terapia de ablación tales como radioterapia y quimioterapia del cáncer. En otro ejemplo no limitante, las células reprogramadas de la presente invención son útiles en la regeneración de órganos y en la reparación de tejidos. En una modalidad de la presente invención, las células reprogramadas pueden usarse para revitalizar el tejido muscular dañado que incluye los músculos distróficos y los músculos dañados por eventos isquémicos, tales como infartos de miocardio. En otra modalidad de la presente invención, las células reprogramadas descritas en la presente descripción pueden usarse para mejorar la cicatrización en los animales, que incluyen los humanos, después de una lesión traumática o de la cirugía. En esta modalidad, las células reprogramadas de la presente invención se administran sistémicamente, tal como por vía intravenosa, y migran al sitio del tejido recién traumatizado reclutadas por las citoquinas circulantes que se secretan por las células dañadas. En otra modalidad de la presente invención, las células reprogramadas pueden administrarse localmente a un sitio de tratamiento que lo necesita o en reparación o regeneración.

En una modalidad de la invención, los ácidos nucleicos tal como el ARN se administran a un paciente mediante métodos ex vivo, es decir, mediante la extracción de células de un paciente, la modificación genética de dichas células y la reintroducción de las células modificadas en el paciente. La transfección y los métodos de transducción se conocen por el experto en la técnica.

El término "transfección" se refiere a la introducción de ácidos nucleicos, en particular ARN, en una célula. Para los propósitos de la presente invención, el término "transfección" también incluye la introducción de un ácido nucleico en una célula o la captación de un ácido nucleico por tal célula, en donde la célula puede estar presente en un sujeto, por ejemplo, un paciente. Por lo tanto, de acuerdo con la presente invención, una célula para la transfección de un ácido nucleico descrito en la presente descripción puede estar presente *in vitro* o *in vivo*, por ejemplo, la célula puede formar parte de un órgano, un tejido y/o un organismo de un paciente. De acuerdo con la invención, la transfección puede ser transitoria o estable. Para algunas aplicaciones de la transfección, es suficiente si el material genético transfectado se expresa solo transientemente. Como el ácido nucleico que se introduce en el proceso de transfección no se integra usualmente en el genoma nuclear, el ácido nucleico extraño se diluirá a través de la mitosis o se degradará. Las células que permiten la amplificación episomal de ácido nucleicos reducen grandemente la tasa de la dilución. Si se desea que el ácido nucleico transfectado permanezca realmente en el genoma de la célula y sus células hijas, debe ocurrir una transfección estable. El ARN puede transfectarse en las células para expresar transientemente su proteína codificada.

De acuerdo con la presente invención, puede usarse cualquier técnica útil para introducir, es decir, transferir o transfectar ácidos nucleicos en las células. Preferentemente, el ARN se transfecta en células mediante técnicas estándar. Tales técnicas incluyen electroporación, lipofección y microinyección. En una modalidad particularmente preferida de la presente invención, el ARN se introduce en las células mediante electroporación. La electroporación o electropermeabilización se refiere a un aumento significativo en la conductividad eléctrica y la permeabilidad de la membrana plasmática celular causada por un campo eléctrico aplicado externamente. Por lo general, se usa en biología molecular como una vía para introducir alguna sustancia en una célula. De acuerdo con la invención, se prefiere que la introducción de ácido nucleico codificante para una proteína o péptido en las células resulte en la expresión de dicha proteína o péptido.

De acuerdo con la invención, los ácidos nucleicos pueden dirigirse a células particulares. En tales modalidades, un vehículo usado para la administración de un ácido nucleico a una célula (por ejemplo, un retrovirus o un liposoma) puede tener una molécula de direccionamiento. Por ejemplo, una molécula, tal como un anticuerpo específico para una proteína de la superficie de la membrana en la célula diana, o un ligando para un receptor en la célula diana puede incorporarse en, o unirse al vehículo de ácido nucleico. Si se desea la administración de un ácido nucleico mediante liposomas, en la formulación del liposoma pueden incorporarse proteínas que se unen a una proteína de la superficie de la membrana asociada con la endocitosis para facilitar el direccionamiento y/o la absorción. Tales proteínas incluyen proteínas de la cápsida o fragmentos de estas que son específicas para un tipo particular de células, anticuerpos para proteínas que se internalizan, y proteínas que se dirigen a un sitio intracelular, y similares.

El término "reportero" se refiere a una molécula, típicamente un péptido o proteína, que está codificada por un gen reportero y se mide en un ensayo reportero. Los sistemas convencionales usualmente emplean un reportero enzimático y miden la actividad de dicho reportero.

5 El término "sitio de clonación múltiple" se refiere a una región de ácido nucleico que contiene sitios de enzimas de restricción, cualquiera de los cuales puede usarse para la escisión de, por ejemplo, un vector y la inserción de un ácido nucleico.

10 De acuerdo con la invención, dos elementos tales como nucleótidos o aminoácidos son consecutivos, si están directamente adyacentes entre sí, sin ninguna interrupción. Por ejemplo, una secuencia de x nucleótidos consecutivos N se refiere a la secuencia (N)_x.

15 La "endonucleasa de restricción" o "enzima de restricción" se refiere a una clase de enzimas que escinden los enlaces fosfodiéster en ambas cadenas de una molécula de ADN dentro de secuencias de bases específicas. Estas reconocen sitios de unión específicos, denominados secuencias de reconocimiento, en una molécula de ADN de doble cadena. Los sitios en los que dichos enlaces fosfodiéster en el ADN son escindidos por dichas enzimas se denominan como sitios de escisión. En el caso de las enzimas tipo IIS, el sitio de escisión se ubica a una distancia definida del sitio de unión al ADN. De acuerdo con la invención, el término "endonucleasa de restricción" comprende, por ejemplo, las enzimas SapI, EciI, BpiI, AarI, AclI, BaeI, BbvCI, PpiI y PstI, BsrDI, BtsI, EarI, BmrI, BsaI, BsmBI, FaeI, BbsI, BciVI, BfuAI, BspMI, BseRI, EciI, BtgZI, BpuEI, BsgI, MmeI, CspCI, BaeI, BsaMI, Mva1269I, PctI, Bse3DI, BseMI, Bst6I, Eam1104I, Ksp632I, BfiI, Bso31I, BspTNI, Eco31I, Esp3I, BfuI, Acc36I, AarI, Eco57I, Eco57MI, GsuI, AclI, Hin4I, PpiI, y PstI.

25 El término "estabilidad" del ARN se refiere a la "vida media" del ARN. La "vida media" se refiere al período de tiempo que se necesita para eliminar la mitad de la actividad, cantidad, o número de moléculas. En el contexto de la presente invención, la vida media de un ARN es indicativa de la estabilidad de dicho ARN.

30 Los ácidos nucleicos tales como el ARN descrito en la presente descripción, en particular cuando se usan para los tratamientos descritos en la presente descripción, pueden presentarse en forma de una composición farmacéutica o kit que comprende el ácido nucleico y, opcionalmente, uno o más vehículos, diluyentes y/o excipientes aceptables farmacéuticamente.

Las composiciones farmacéuticas son, preferentemente, estériles y contienen una cantidad eficaz del ácido nucleico.

35 Las composiciones farmacéuticas se proporcionan usualmente en una forma de dosificación uniforme y pueden prepararse de una manera conocida en la técnica. La composición farmacéutica puede estar, por ejemplo, en forma de una solución o suspensión.

40 La composición farmacéutica puede comprender sales, sustancias tamponantes, conservantes, portadores, diluyentes y/o excipientes los que son, preferentemente, aceptables farmacéuticamente. El término "aceptable farmacéuticamente" se refiere a la no toxicidad de un material que no interfiere con la acción del componente activo de la composición farmacéutica.

45 Las sales que no son aceptables farmacéuticamente pueden usarse para preparar sales farmacéuticamente aceptables y se incluyen en la invención. Las sales aceptables farmacéuticamente de este tipo comprenden, de manera no limitante, aquellas preparadas a partir de los ácidos siguientes: ácidos clorhídrico, bromhídrico, sulfúrico, nítrico, fosfórico, maleico, acético, salicílico, cítrico, fórmico, malónico, succínico y similares. Las sales aceptables farmacéuticamente pueden prepararse además como sales de metales alcalinos o sales de metales alcalinotérreos, tales como sales de sodio, sales de potasio o sales de calcio.

50 Las sustancias tamponantes adecuadas para usar en la composición farmacéutica incluyen ácido acético en una sal, ácido cítrico en una sal, ácido bórico en una sal y ácido fosfórico en una sal.

Los conservantes adecuados para usar en la composición farmacéutica incluyen cloruro de benzalconio, clorobutanol, parabeno y timerosal.

55 El término "vehículo" se refiere a un componente orgánico o inorgánico, de una naturaleza natural o no natural (sintética), con el que el componente activo se combina para facilitar, potenciar o permitir la aplicación. De acuerdo con la invención, el término "vehículo" incluye además uno o más agentes de relleno sólidos o líquidos, diluyentes o sustancias de encapsulación, que son adecuadas para la administración a un paciente.

60 Las sustancias vehículos posibles para la administración parenteral son, por ejemplo, agua estéril, solución de glucosa, Ringer, lactato de Ringer, solución de cloruro de sodio estéril, polialquilenglicoles, naftalenos hidrogenados y, en particular, polímeros biocompatibles de lactida, copolímeros de lactida/glicolida o copolímeros de polioxietileno/polioxipropileno.

65 El término "excipiente" cuando se usa en la presente descripción pretende indicar todas las sustancias que pueden presentarse en una composición farmacéutica y que no son ingredientes activos tales como, por ejemplo, vehículos,

aglutinantes, lubricantes, espesantes, agentes activos de superficie, conservantes, emulsionantes, tampones, agentes saborizantes, o colorantes.

Las composiciones farmacéuticas descritas en la presente descripción pueden administrarse por cualquier vía convencional, tal como por administración parenteral que incluye la inyección o la infusión. La administración es preferentemente parenteral, por ejemplo, intravenosa, intraarterial, subcutánea, en el nódulo linfático, intradérmica o intramuscular.

Las composiciones adecuadas para la administración parenteral comprenden usualmente una preparación acuosa o no acuosa estéril del compuesto activo, que es, preferentemente, isotónica con respecto a la sangre del receptor. Los ejemplos de portadores y solventes compatibles son la solución de Ringer y la solución de cloruro de sodio isotónica. Adicionalmente, se usan aceites fijados, usualmente estériles, como solución o medio de suspensión.

Los agentes y composiciones descritos en la presente descripción se administran, preferentemente, en cantidades eficaces. Una "cantidad eficaz" se refiere a la cantidad que logra una reacción deseada o un efecto deseado solo o junto con dosis adicionales. En el caso del tratamiento de una enfermedad particular o de una afección particular, la reacción deseada se relaciona, preferentemente, con la inhibición del curso de la enfermedad. Esto comprende ralentizar el progreso de la enfermedad y, en particular, interrumpir o revertir el progreso de la enfermedad. La reacción deseada en un tratamiento de una enfermedad o de una afección puede ser también un retraso en la aparición o una prevención de la aparición de dicha enfermedad o de dicha afección.

Una cantidad eficaz de un agente o composición descrito en la presente descripción dependerá de la afección a tratar, la gravedad de la enfermedad, los parámetros individuales del paciente, que incluyen la edad, la condición fisiológica, la talla y el peso, la duración del tratamiento, el tipo de una terapia de acompañamiento (si se presenta), la vía específica de administración y factores similares. En consecuencia, las dosis administradas de los agentes descritos en la presente descripción pueden depender de varios de estos parámetros. En el caso que una reacción en un paciente sea insuficiente con una dosis inicial pueden usarse dosis más altas (o dosis eficazmente mayores alcanzadas por una vía de administración diferente, más localizada).

La presente invención se describe en detalle mediante las siguientes figuras y ejemplos que deben interpretarse solamente a modo de ilustración y no por medio de limitación. Sobre la base de la descripción y los ejemplos, otras modalidades son accesibles para el trabajador experto y están igualmente dentro del alcance de la invención.

Figuras

Figura 1: Resumen del proceso de selección *in vivo*

Para preparar la biblioteca de partida, se cultivaron células dendríticas inmaduras humanas en presencia de actinomicina D, un inhibidor de la transcripción, durante cinco horas para preseleccionar los ARN estables. El ARNm celular restante se extrajo y se purificó mediante el uso del kit Poly(A)Purist (Ambion) y luego se fragmentó con Nuclease P1 (Roche). Para esto, 10 µg de ARN se incubaron durante 45 minutos con 0,3 U NP-1 en 8 µL 50 mM de tampón de NaAc (pH 5,5) en un volumen de reacción total de 24 µL. Después de la purificación con columnas RNeasy (Qiagen), los fragmentos estaban listos para retrotranscribirse en ADNc. Se realizó la síntesis de la primera y segunda cadena mediante el uso y el seguimiento del protocolo del kit de síntesis de la 1era cadena de ADNc RevertAid Premium (Fermentas) y un cebador hexámero con una secuencia de cebador definida y un sitio de restricción NotI. Para rellenar los voladizos en 5' y eliminar los voladizos en 3', el ADNc se incubó a continuación con 12,5 U de ADN polimerasa T4 durante 5 minutos a 15 °C. La reacción se terminó mediante la adición de 5 µL EDTA 0,5 mM, pH 8,0 y el ADNc se purificó mediante el uso de columnas NucleoBond (Macherey-Nagel). La digestión de la biblioteca de ADNc con NotI (NEB) produjo fragmentos con un extremo romo y pegajoso. Los fragmentos se seleccionaron adicionalmente por tamaño a través de la preparación de gel para asegurar la eliminación de todos los fragmentos más pequeños de 150 pb. Para la clonación de la biblioteca, el vector se digirió con EcoRV y NotI, como se muestra en el panel A, dejando un extremo romo y pegajoso, respectivamente. En la siguiente etapa, la biblioteca se ligó en el vector mediante el uso de la ADN ligasa T4 (Fermentas). La mezcla de ligación se usó directamente como molde para la PCR como se proporciona en la Tabla 4 mediante el uso de la ADN polimerasa Phusion™ de alta fidelidad de inicio en caliente (Finnzymes). Después de la purificación, se usó el producto de PCR como molde para la transcripción de T7 como se muestra en la Tabla 5. La incubación se realizó a 37° C. Después de 30 min se añadieron a la reacción 0.75 µL de GTP 100 mM. La reacción se detuvo después de 2.5 h mediante la adición de TURBO DNasa (2U/µL, Ambion) y se incubó durante otros 15 min a 37 °C. La reacción finalmente se limpió mediante columnas RNeasy (Qiagen). La biblioteca de ARN podría entonces usarse para el procedimiento de selección que comienza con la electroporación del ARN en hiDC como se describió anteriormente (Kuhn y otros, 2010). Después del cultivo para la selección, la extracción y purificación del ARN se realizó mediante el uso de columnas RNeasy (Qiagen) y siguiendo las instrucciones del fabricante. El ARN se usó seguidamente como molde para la síntesis de ADNc mediante el uso de la transcriptasa inversa Superscript II (Invitrogen) y siguiendo las instrucciones del fabricante y un cebador dT18. Posteriormente se usó el ADNc como molde para la PCR como se describió anteriormente. Finalmente, los productos de la PCR podrían usarse como molde para la transcripción-T7 (ver arriba) para comenzar la siguiente ronda de selección (panel B). Los controles de calidad de las muestras de ADN/ADNc y ARN se realizaron mediante gel de agarosa y bioanalizador AGILENT 2100, respectivamente.

Figura 2: Esquema de la apariencia de la muestra dentro del vector luc2CPmut

(A) Un elemento único o dos elementos (en dirección 5' y 3') se clonaron como UTR-3' en el vector como se indica. Además se presentan las muestras de control NEG (control negativo sin inserción de una UTR 3'), hBg y 2hBg. Preparación de ARN para rondas de selección. (B) Para la electroporación en hiDC se usó el vector como molde para la PCR mediante el uso de cebadores alargados que comprenden el promotor T7 y la cola poli(A). El producto de la PCR se usó posteriormente como molde para la síntesis de T7 *in vitro* para producir el respectivo IVT-RNA.

Figura 3: Efecto de las secuencias seleccionadas sobre la estabilidad de los ARN que codifican luc2Cpmut

Los resultados muestran la actividad de la luciferasa, la vida media y la proteína total, a lo largo del tiempo, de los ARN que contienen las secuencias seleccionadas como UTR 3' en comparación con nuestro estándar de oro 2hBg después de la electroporación en células dendríticas inmaduras humanas (el NEG es como se definió en la Figura 2). El panel superior proporciona los cursos de tiempo de tres ARN ilustrativos con las UTR 3' como se indicó. En el panel inferior izquierdo, se muestra la vida media de los ARN con la UTR 3' respectiva como se indicó en relación con un ARN con 2hBg. De manera similar, la expresión relativa de proteína total en comparación con un ARN con 2hBg se proporciona en el panel inferior derecho.

Figura 4: Actividad de luciferasa representativa mediante el uso de luc2mut como gen reportero y UTR-3' recién seleccionadas. Después de la electroporación de ARN con UTR 3' como se indicó, en células dendríticas inmaduras humanas, se midió la actividad de luciferasa durante 72 h.

Figura 5: Resultados representativos de la electroporación con ARN-IVT en fibroblastos

panel izquierdo: vector basado en luc2CPmut. panel derecho: vector basado en luc2mut

Figura 6: Resultados representativos de la electroporación con ARN-IVT en células T

El panel más a la izquierda proporciona la expresión relativa de proteína total de un ARN con FI UTR 3' en comparación con un ARN con 2hBg, en células T CD4+ y CD8+. De manera similar, la eficiencia de la traducción relativa y la vida media del ARNm de un ARN con FI UTR 3' en comparación con un ARN con 2hBg en células T CD4+ y CD8+ se proporciona en el panel central y en el panel más a la derecha, respectivamente.

Figura 7: Arquitectura del ARN e integridad para evaluar los ARN con nucleótidos modificados

A: Los ARN usados en los ensayos de luciferasa se construyeron como se representa aquí. Como casquete 5' se usó β -S-ARCA (D2). Como 5'UTR se usó la 5'UTR de alfa globina humana, que incluye una secuencia de Kozak. Después del gen de luciferasa de luciérnaga, se clonaron las dos 3'UTR para compararse. Como cola poliA, se usó una secuencia A30L70.

B: Antes de la transfección, se verificó la integridad de los ARN en un Bioanalizador 2100 (Agilent). Todos los ARN tenían una integridad suficientemente alta y comparable además y, por lo tanto, podían usarse en los experimentos.

Figura 8: Efecto de FI UTR 3' sobre la estabilidad y funcionalidad del ARN *in vivo*

La luciferasa y el ARNm de gp70 que contenían FI UTR 3' o 2hBg 3'UTR se formularon con F12 y se administraron i.v. en ratones BALB/c. Después de la administración del ARNm de Luciferasa, la expresión se controló después de 6 h y 24 h; El ARNm de gp70 se administró en el día 0 y el día 6 y la activación inmunitaria se analizó en el día 10 mediante tinción de CD8 y gp70 tet+.

A) Muestra los niveles de expresión de luciferasa a las 6h y 24h después de la inyección de ARNm no modificado y ARNm modificado m1Y que contiene FI 3'UTR o la 2hBg 3'UTR. Tanto el ARNm de Luciferasa no modificado como el m1Y modificado que contiene la FI 3'UTR muestran niveles de expresión comparables como el ARNm correspondiente que contiene la 2hBg 3'UTR.

B) Muestra el porcentaje de células T CD8 específicas de gp70 en respuesta al ARNm de gp70 que contiene la FI 3'UTR o la 2hBg 3'UTR. Las dos 3'UTR se desempeñan igualmente bien en la inducción de inmunidad específica de antígeno después de dos inmunizaciones, con un aumento significativo de las células T CD8 específicas de antígeno en el bazo de aquellos ratones que recibieron ARNm gp70 que contiene la FI 3'UTR.

Estadísticas: ANOVA de una vía seguido de la prueba de Tukey, * p <0.05.

Figura 9: Efecto de las UTR estabilizadoras sobre la estabilidad del ARN autorreplicante

La Luciferasa desestabilizada (Luc2CP) se clonó inmediatamente en dirección 5' del elemento de secuencia conservada en 3' de un ARN autorreplicante (replicón) derivado del virus Semliki Forest no citotóxico. El replicón de ARN se preparó

mediante la transcripción *in vitro* de un plásmido linealizado correspondiente y se electroporó en las células. La expresión de luciferasa se midió mediante la adición de sustrato luminiscente durante 96 a 120 h. (A) Curso temporal de la expresión de luciferasa en un experimento representativo con células BHK21. (B) Curso temporal de la expresión de luciferasa en un experimento representativo con fibroblastos de prepucio humano (HFF). Para reducir la citotoxicidad de los interferones tipo I liberados, se cotransfectó el ARNm del virus Vaccinia B18R en cada muestra. Para inhibir la activación de la proteína quinasa R y aumentar el nivel general de traducción, se cotransfectó el ARNm del virus Vaccinia E3 en cada muestra.

Figura 10: tramos de homología en el elemento FI.

Se predijo que los tramos de secuencia subrayados se aparearían entre sí. Para la construcción "mutación 8nt", se mutó el primer elemento por aaagggu para interrumpir las interacciones con el segundo elemento.

Figura 11: Artefactos en la IVT basado en los moldes de PCR mediante el uso de 2hBgUTR.

A: Representación esquemática de la generación del molde para IVT mediante PCR. El cebador 5' se hibrida en dirección 5' del promotor T7, el cebador 3' contiene una cola de poliA de 120 nt y se hibrida con el poliA codificado por el plásmido y parte de la 3'UTR. En el caso de 2hBgUTR, puede producirse un error de hibridación mediante la hibridación a la primera repetición. B: Productos de PCR de un plásmido que contiene el 2hBgUTR. La flecha roja representa el producto secundario, que representa un truncamiento de 1hBg. C: Por lo tanto, el ARN transcrito de un producto de PCR además presenta un subproducto acortado (flecha). D: Productos de PCR de un plásmido que contiene el elemento FI como 3'UTR. No hay productos secundarios visibles. E: El ARNm resultante tiene la alta integridad esperada, sin picos laterales adicionales.

Figura 12: Representación esquemática de los elementos UTR truncados y la vida media de las construcciones de ARNm correspondientes.

El panel superior de la figura A muestra una representación esquemática de los elementos UTR truncados en referencia a las posiciones de ácido nucleico de la secuencia de longitud completa del elemento F de la SEQ ID NO: 86 cubierto por esas variantes truncadas.

El panel inferior de la figura A muestra la vida media relativa del ARNm que comprende el UTR truncado en referencia al ARNm que comprende la secuencia de longitud completa del elemento F de la SEQ ID NO: 86. Los ARNm codificantes de un reportero de luciferasa se electroporaron en hiDC y se siguió su expresión en el tiempo mediante las mediciones de luciferasa para determinar la vida media relativa del ARN.

El panel superior de la Figura B muestra una representación esquemática de los elementos UTR truncados en referencia a las posiciones de ácido nucleico de la secuencia de longitud completa del elemento I de la SEQ ID NO: 115 cubierta por esas variantes truncadas.

El panel inferior de la Figura B muestra la vida media relativa del ARNm que comprende el UTR truncado en referencia al ARNm que comprende la secuencia de longitud completa del elemento I de la SEQ ID NO: 115. Los ARNm codificantes de un reportero de luciferasa se electroporaron en hiDC y se siguió su expresión en el tiempo mediante las mediciones de luciferasa para determinar la vida media relativa del ARN.

Figura 13: Vida media relativa y expresión de proteínas a partir de construcciones de ARNm que comprenden elementos F, I o FI hacia UTR aleatorias.

La Figura 13 muestra la vida media relativa y la expresión de proteínas a partir de construcciones de ARNm que comprenden elementos F, I o FI hacia UTR aleatorias. Para esta longitud completa, los elementos individuales F e I, así como también la combinación de FI, se compararon hacia una UTR 3' aleatoria (257nt de longitud). Todos los elementos se clonaron en construcciones codificantes para luciferasa, se transcribieron a ARNm *in vitro*, se electroporaron en hiDC, se midió la expresión de luciferasa a lo largo del tiempo, y se calcularon las vidas medias relativas y la expresión de proteínas totales.

Figura 14: Elementos UTR para reprogramación celular.

La Figura 14A muestra la línea de tiempo para la reprogramación de fibroblastos de prepucio humano primarios. Se sembraron 40,000 células en una placa de 12 pocillos y se lipofectaron durante tres (3x) o cuatro (4x) días consecutivos con mezclas de ARNm que estaban compuestas de 0.33 µg de ARN sin modificar transcrito *in vitro* (IVT) que codifica la reprogramación TF OCT4, SOX2, KLF4, cMYC, NANOG y LIN28 (OSKMNL) (1:1:1:1:1:1) con 0.08 µg de cada B18R, E3 y K3 (EKB) y 0.17 µg de una mezcla de miARN compuesta de miARN 302a-d y 367 (1:1:1:1:1:1). De esta manera, las construcciones de ARN solo diferían en su 3'UTR que consiste en una repetición en tándem de la 3'UTR de la β-globina humana (2hBg), el elemento F-I (FI) o el elemento I-F (IF). Desde el día 9 en adelante, se observó la formación de colonias y el análisis de colonias se realizó en el d11.

La Figura 14B muestra una tinción de fosfatasa alcalina (AP) de las colonias establecidas y la Figura 14C muestra un gráfico de barras correspondiente que representa el recuento numérico de las colonias positivas de AP.

La Figura 14D muestra la morfología de las colonias de células iPS resultantes mediante el uso de los ARN que contienen la FI-UTR. Era como células semejantes a hES, con células pequeñas muy compactas en colonias distintas y bordes bien definidos.

La Figura 14E muestra las colonias preparadas como en D teñidas positivas para AP con cuatro y diez aumentos.

La Figura 14F muestra colonias preparadas como en D en una tinción vital del marcador de superficie celular TRA-1-60, de las hES.

La Figura 14G muestra la expresión de ARNm de los marcadores de las hES, OCT4 (endógeno), NANOG (endógeno), LIN28 (endógeno), TERT y REX1 evaluados mediante sedimentación de las colonias, aislamiento del ARN total y cuantificación mediante qRT-PCR.

EJEMPLOS

10 Ejemplo 1: Identificación de elementos de secuencia que estabilizan los ARNm

Para identificar elementos de secuencia novedosos que estabilizan los ARNm, hemos desarrollado un proceso de selección *in vivo* mediante el uso de hiDC como entorno selectivo para el ARN transcrito *in vitro*. La biblioteca del ARN de partida se construyó mediante el uso de secuencias de ARNm de origen natural derivadas de hiDC. Antes del aislamiento del ARN, las células se cultivaron durante 5 horas en presencia del inhibidor de la transcripción Actinomicina D (ActD) para preseleccionar los ARN estables. Después se extrajo el ARNm restante y se redujo a fragmentos de 200-800 nucleótidos, se transcribió inversamente y se clonó como UTR-3' en un vector que portaba una secuencia de hAg 5'UTR y un gen reportero, que se eligió como base del proceso de selección. El molde de ADN usado para la posterior transcripción del ARNm de la biblioteca se amplificó mediante PCR, durante la cual se introdujo un promotor T7 a través del 5' y una cola PoliA A60 a través del cebador 3'. El ARNm transcrito se introdujo después en el proceso de selección *in vivo*, que se compone de varias rondas de transcripción *in vitro* de la biblioteca, electroporación de los ARN correspondientes en hiDC, y extracción y amplificación de secuencias estables después de puntos de tiempo definidos. La amplificación de las secuencias seleccionadas se realizó mediante PCR con cebadores específicos, después de la síntesis de ADNc. Los productos de PCR resultantes se usaron posteriormente como moldes para la nueva biblioteca del ARNm. Esto se realizó durante seis rondas, con extracción de los ARN restantes después de 24 horas en la ronda 1, 48 horas en las rondas 2 y 3, 72 horas en las rondas 4 y 5, y finalmente 96 horas, así como también una y dos semanas en la ronda 6 (después de la electroporación, las células se dividieron en tres partes y después se cosecharon individualmente en los puntos de tiempo dados).

El monitoreo del proceso de selección después de las rondas 1 a 5 demostró un aumento significativo de la vida media promedio del conjunto de ARN correspondiente, lo que es indicativo de un enriquecimiento de los elementos UTR-3' estabilizadores (Tabla 1). Sin embargo, el aumento de la estabilidad fue menos pronunciado con rondas más altas. Por lo tanto, el proceso de selección se detuvo después de una sexta ronda final, en la que se extrajo el ARN de las células después de 96 horas, una semana, y dos semanas. Para caracterizar las secuencias seleccionadas, se secuenciaron más de 350 clones individuales, 108 de la ronda 5, 88 de la ronda 6/96 horas, 110 de la ronda 6/1 semana, y 96 de la ronda 6/2 semanas. Todas las secuencias se compararon entre sí, así como también y con el BLAST para identificar su origen genómico. Aquí, se analizó especialmente si las secuencias se derivaban de UTR 5' o 3' endógenas o de la región codificante. Finalmente, su nivel de expresión en las hiDC se descargó de NextBio (Illumina). En total, pudieron identificarse siete grupos, (i) para los que se encontraron secuencias múltiples, (ii) los que se originaron a partir de las UTR-3' de ARN endógenos o de un ARN no codificante endógeno, y (iii) los que se expresaron claramente en las hiDC (Tabla 2). Estos se derivan de los siguientes genes: Fragmento Fc de IgG, receptor, transportador, alfa (B, FCGRT, NM_001136019), Proteína específica de linfocitos 1 (D, LSP1, NM_002339), Ligando de quimiocina 22 (E, CCL22, NM_002990), Potenciador de la separación del amino terminal (F, AES, NM_198969), Miembro 3 de la familia de fosfolipasa D (G, PLD3, NM_001031696), ARN 12S codificado mitocondrialmente (I, MT_RNR1, NC_012920), Complejo principal de histocompatibilidad clase II DR beta 4 (J, HLA-DRB4, NM_021983). Tenga en cuenta que, por simplicidad, las letras mayúsculas B a I dadas entre paréntesis se usan a continuación como abreviaturas para estos elementos. Es importante destacar que, en todos los casos, los clones para una secuencia difieren en sus extremos exactos 5'- y 3'-, lo que demuestra que provienen de diferentes clones iniciales y no se enriquecen artificialmente durante el proceso (ver los apéndices para obtener una lista completa de todas las secuencias identificadas en la selección).

50 Ejemplo 2: Caracterización de elementos de secuencia individuales identificados

Para la caracterización de los elementos de secuencia identificados, se eligió un candidato representativo de cada grupo (las secuencias detalladas están marcadas en el apéndice). Después esta secuencia se clonó como UTR-3' en un vector con un gen reportero de luciferasa, cuyo nivel de expresión puede analizarse en el tiempo después de la transferencia a las células. Se ha demostrado previamente que a partir de la observación del patrón de expresión para la proteína, puede inferirse con precisión la estabilidad relativa y la eficiencia de la traducción del ARN (Kuhn 2010 Gene Ther.). El reportero específico usado en este experimento, luc2CPmut, es una forma desestabilizada de luciferasa (Promega). Esto permite detectar incluso pequeños cambios en la estabilidad del ARN. Después se comparó el ARN transcrito *in vitro* procedente de estos vectores con nuestro ARNm estándar de oro, es decir, el que contiene el 2hBg UTR-3', con respecto a la estabilidad del ARN y a la eficiencia de la traducción. Como muestras de control se usaron un ARN transcrito *in vitro* sin una UTR-3' (es decir, que solo contenía secuencias usadas para clonar los insertos) y uno con un solo elemento de Beta-globina (IhB).

Comenzando con los vectores que contienen UTR, la región a transcribirse se amplificó por PCR mediante el uso de un cebador 5' que contiene el promotor T7 y un cebador 3' con una cola poli(A) de 60 nucleótidos. La purificación de los

fragmentos de PCR se realizó mediante el uso de AGENCOURT AMPURE XP (Beckman Coulter). Se añadieron 0,6 volúmenes de perlas a cada reacción de PCR y se mezclaron. Después de una incubación de 15 minutos a RT, los productos de PCR unidos a las perlas se separaron mediante soporte magnético del exceso de cebadores, nucleótidos, sales y enzimas. Las perlas se lavaron dos veces durante 30 segundos con etanol al 80 % para eliminar aún más los contaminantes. Los productos de PCR deseados finalmente se eluyeron dos veces con 30 μ L ddH₂O y se usaron como moldes para la transcripción in vitro de los ARN correspondientes. Para las *transcripciones in vitro* se usaron ARN polimerasa T7 (Fermentas), el tampón de la reacción respectivo y NTP 6 mM. Para la formación eficiente del casquete del ARN, la concentración de GTP se redujo a 1,5 mM y se añadieron 6 mM de β -S-ARCA(D2) a la reacción y se incubaron durante 2,5 h a 37 °C. El ARN se purificó mediante perlas magnéticas carboxiladas (Invitrogen) y la concentración y calidad del ARN se evaluaron mediante espectrofotometría y análisis en un Bionanalyzer 2100 (Agilent).

De acuerdo con sus identificaciones en el enfoque de detección, todas las secuencias nuevas mostraron características muy similares en comparación con 2hBg con respecto a la estabilidad del ARN con el grupo I (mtRNRI) como el mejor (Figura 3; Tabla 3). Es importante destacar que, cada elemento individual confirió estabilización al ARN en comparación con el ARN sin una UTR-3' e incluso en comparación con el ARN solo con una copia única del elemento Beta-globina. La eficiencia de la traducción no se afectó significativamente, como se observó por la correlación directa entre la estabilidad del ARN y la proteína total expresada en el tiempo.

Ejemplo 3: Combinación de elementos de secuencia individuales.

En un experimento adicional, las secuencias individuales de cada grupo se combinaron entre sí de una manera bien ponderada (Figura 2). La racionalidad detrás de esto fue nuestra observación previa de que la combinación de dos UTR-3' tenía un efecto adicional sobre la estabilidad y la eficiencia de la traducción del ARN (Holtkamp y otros 2006). La estabilidad y la eficiencia de la traducción del ARN se calcularon en R mediante interpolación de los valores de luciferasa medidos con una spline, a partir de la cual la pendiente ascendente más pronunciada se definió como la eficiencia de la traducción y la vida media de la señal como estabilidad. La integral de la spline interpolada se interpretó como la expresión de proteína total. En total, se clonaron 64 combinaciones, es decir, todas las combinaciones posibles de las siete secuencias recientemente identificadas y de la beta-globina UTR-3' humana (Tabla 6). Como se describió anteriormente, el ARN se preparó a partir de estos moldes de ADN, y después se electroporó en hiDC. Como controles, además se incluyeron los ARN que contenían los elementos individuales. Para la mayoría de los siete elementos nuevos se observó que al menos una combinación con otro elemento proporciona un ARN con una mayor estabilidad que solo con el elemento individual (Tabla 7 a Tabla 13). Curiosamente, en la mayoría de los casos, la combinación con el elemento I (mtRNRI) aumentó la vida media del ARN. Aquí, la estabilidad del ARN fue generalmente incluso mayor en comparación con un ARN con 2hBg UTR-3' (Tabla 7 a Tabla 13). Casi todas las combinaciones tuvieron un efecto positivo en la eficiencia de la traducción del ARN. En total, los efectos combinados sobre la estabilidad del ARN y la eficiencia de la traducción resultan en un aumento de la expresión total de proteína de hasta 1,74 veces. Por lo tanto, podríamos identificar elementos individuales (con longitudes inferiores a 233 nucleótidos), así como también combinaciones de dos elementos diferentes que dan lugar al ARN con un aumento en la estabilidad y/o eficiencia de la traducción, pero al mismo tiempo se evitarían los problemas de tener dos copias idénticas de un elemento como se describió anteriormente para 2hBg.

Para verificar los resultados obtenidos con la forma desestabilizada de luciferasa, se repitieron los experimentos anteriores con ARN que portaban la luciferasa estándar (Promega), y las siguientes UTR-3' seleccionadas: mtRNRI (I), mtRNRI-AES (IF), AES-mtRNRI (FI), mtRNRI-hBg (IhBg) y hBg-mtRNRI (IhBg). Como se muestra en la Figura 4 y Tabla 14, pueden obtenerse resultados equivalentes como se observó anteriormente, mediante la verificación de que los elementos nuevos, individualmente o en combinación, aumentan la estabilidad del ARNm y/o la eficiencia de la traducción de manera similar al elemento 2hBg.

Ejemplo 4: Análisis de ARNm que portan elementos de secuencia seleccionados en otros tipos de células

Las UTR-3' mtRNRI y AES recién seleccionadas se evaluaron además en diferentes tipos de células y líneas celulares para ver si hay una especificidad de hiDC. Las secuencias se evaluaron en fibroblastos humanos (HFF), mioblastos murinos (C2C12) (Figura 5) y células T (Figura 6) para evaluar si además se estabilizan en estas células.

Las células HFF y C2C12 se cosecharon y se prepararon para la electroporación. A continuación se electroporaron 2,0 μ g de IVT-RNA junto con 1,0 μ g del ARN codificante para GFP que contiene las 3'UTR indicadas. Después de la electroporación, las células se dividieron. Para medir la actividad luciferasa se distribuyeron 5000 células por pocillo en una placa de 96 pocillos por triplicado en un total de 7 puntos de tiempo (2, 4, 8, 24, 48 y 72 h). Se sembraron 2E+05 células por pocillo en placas de 6 pocillos para cosechar para FACS después de 24 h (señal GFP). Esto permitió monitorear las eficiencias de la transfección. Estas diferían entre 72 y 90 % y podrían incluirse en el cálculo de la vida media. Los resultados obtenidos con HFF y C2C12 así como también con las células T confirmaron los resultados obtenidos previamente con hiDC. La combinación de I con F fue en particular de 2 a 3 veces mejor en la vida media en comparación con 2hBg. Además, FI mostró una eficiencia de la traducción 3 veces mejor en las células C2C12 y una producción de proteínas 2 veces mejor en el tiempo, en comparación con nuestro estándar de oro. Estos resultados mostraron que I y F no son específicos de hiDC, sino que, además, mejoran la estabilidad del ARNm y la eficiencia de la traducción en otras células.

Ejemplo 5: El FI 3'UTR aumenta la expresión del ARNm modificado

Para algunas aplicaciones, que incluyen la terapia de reemplazo de proteínas, se prefieren los ARNm con nucleótidos modificados en lugar de los no modificados debido a su inmunogenicidad disminuida (Kariko y otros, 2008). Sin embargo, las modificaciones de base pueden tener un efecto sobre la estabilidad de un ARNm ya sea por influir directamente en la interacción con una proteína de unión al ARN correspondiente, o por alterar la formación de la estructura secundaria del ARN. En consecuencia, las UTR 3' seleccionadas podrían comportarse de manera diferente en el contexto de los ARNm modificados. Por lo tanto, comparamos la combinación de F e I con 2hBgUTR en el contexto de ARNm modificado m1Y en células hiDC, HFF, células T CD8+ y CD4+ y en MEF murinos, C2C12 y bmDC. Se usó Luciferasa como reportero (ver Figura 7A para el diseño de la construcción). Para la generación de los ARNm modificados, en la reacción de IVT, U se sustituyó completamente por m1Y. En todos los experimentos, el ARN no modificado se incluyó como un control. Las integridades de los ARNm obtenidos no se afectaron por el intercambio de UTP por m1YTP (Figura 7B). Las células se electroporaron mediante el uso de las condiciones descritas en la Tabla 15, y los niveles de luciferasa se midieron a las 3, 6, 12, 24, 48, 72 y 96 h.

La electroporación del ARNm no modificado de luciferasa podría reproducir los efectos vistos anteriormente: En todos los tipos de células, el elemento FI fue igual o superior al control 2hBg para transmitir la estabilidad del ARN (Tabla 16A). Mientras que en las DC murinas y en las células T humanas, las vidas medias del ARNm eran comparables entre las dos 3'UTR, el elemento FI aumentó las vidas medias de ARNm hasta 1,69 veces en las células HFF. La cantidad de proteína total aumentó en todas las líneas celulares, más prominentemente en las células HFF (2,45 veces).

Con el ARNm modificado, el elemento FI además condujo a un aumento en la vida media de ARNm en comparación con 2hBg en hiDC, la cantidad de proteína total aumentó más del doble (Tabla 16B). Además, los resultados en otros tipos de células son similares a los obtenidos con el ARNm no modificado: El elemento FI fue superior a 2hBg en todos los experimentos que involucraron las células HFF, MEF y C2C12, y comparable en células T y DC murinas (Tabla 16B). Por lo tanto, la modificación del U no altera la capacidad del elemento FI para estabilizar el ARNm.

Ejemplo 6: El FI 3'UTR aumenta la expresión del ARNm independientemente del método de transfección

Hasta ahora, todos los experimentos se realizaron con la electroporación como método de transfección. Con la electroporación, el ARNm suministrado llega directamente al citoplasma, bajo la evasión de una ruta de captación endosómica, la que se sigue después de la transfección mediante lipofección. Para ver si el elemento FI además funciona en estas condiciones, las células se lipofectaron con los mismos ARNm de Luciferasa, que contenían FI y 2hBg que se usaron en experimentos anteriores, mediante el uso de RNAiMAX como reactivo de transfección. Además después de la lipofección, el elemento FI aumentó la expresión de Luciferasa, aunque el aumento fue menos pronunciado en comparación con los experimentos donde el ARN se suministró mediante electroporación (Tabla 16C). Por lo tanto, el método de transfección no tiene un impacto en el efecto estabilizador del elemento FI.

Ejemplo 7: El ARNm que contiene FI 3'UTR y 2hBgUTR conducen a una expresión de proteínas comparable y a la activación inmunitaria *in vivo*.

Para evaluar la expresión de proteínas a partir del ARNm que contiene el FI 3'UTR *in vivo*, los mismos ARNm de Luciferasa que contienen FI y 2hBg que se usaron en experimentos anteriores se formularon con F12 y se administraron i.v. en ratones BALB/c. Como se muestra en la Figura 8, la expresión de luciferasa fue comparable para ambas 3'UTR. La respuesta inmunitaria específica al antígeno se indujo, además, en un grado comparable con el efecto del ARNm que contiene FI 3'UTR que es ligeramente más fuerte en el bazo.

Ejemplo 8: IF UTR conduce a una mayor estabilidad del ARN autorreplicante *in vitro*.

Los ARN autorreplicantes transcritos *in vitro* (replicón de ARN) derivado de genomas alfavirales son potentes vectores de vacuna. El replicón de ARN codifica en los dos primeros tercios el complejo enzimático necesario para la replicación citoplasmática (replicasa) del replicón de ARN. Esta replicasa reconoce una estructura interna del ARN que actúa como promotor subgenómico para la síntesis dependiente de replicasa de los ARN subgenómicos. Los transgenes o antígenos para la vacunación están codificados en este ARN subgenómico que es significativamente más corto que el replicón completo. En general, tanto el ARN genómico (es decir, el replicón de ARN de longitud completa) como el ARN subgenómico se asemejan al ARNm celular. Ambos están flanqueados por las UTR, ambos están protegidos por casquetes y son poliadenilados. Las enzimas responsables de la formación del casquete y de la poliadenilación están contenidas en el complejo de la enzima replicasa. Los elementos de secuencia conservada (CSE) dentro de las UTR, que se superponen con el ORF de la replicasa en el caso del 5'CSE, son necesarios para la unión de la replicasa y actúan como promotores de la síntesis de la cadena negativa (3'CSE) o la síntesis de la cadena positiva (5'CSE).

Para evaluar si las UTR estabilizadoras novedosas, identificadas y validadas para no replicar el ARNm transcrito *in vitro*, proporcionan una mayor estabilidad y, por lo tanto, una mayor expresión transgénica, del ARN replicón, clonamos las secuencias respectivas en vectores molde de ARN replicón. Como el 3'CSE necesita ubicarse inmediatamente adyacente a la cola poli-A, insertamos los UTR novedosos inmediatamente en dirección 5' del 3'CSE de un replicón que codifica la luciferasa desestabilizada (Luc2CP). El replicón de ARN se sintetizó mediante la transcripción *in vitro* de plásmidos molde

linealizados similares al ARNm de IVT. El replicón de ARN se introdujo en las células (BHK21 y HFF) mediante electroporación, y se evaluó la expresión de luciferasa. Como se muestra en la Figura 9, todas las UTR insertadas aumentaron la traducción de Luc2CP en ambas líneas celulares usadas. Curiosamente, la combinación UTR "IF" resultó en un aumento notable de la traducción.

Ejemplo 9: Los intercambios de nucleótidos de hasta el 90 % de homología no tienen impacto en las propiedades estabilizadoras del elemento FI

Debido al procedimiento de selección que se aplicó para identificar los elementos estabilizadores UTR novedosos, se obtuvieron secuencias en un determinado intervalo de tamaño. La identificación de las mismas secuencias con extremos prolongados en 5' y 3' proporcionó una primera indicación de la longitud mínima requerida. Sin embargo, la región mínima requerida para que cada elemento ejerza su efecto estabilizador podría ser aún más corta. Adicionalmente, ligeras variaciones de las secuencias podrían seguir siendo funcionales, es decir, la identidad de cualquier nucleótido individual podría no ser de suma importancia para las propiedades estabilizantes del elemento FI. Para ver en qué grado los elementos son robustos frente a los intercambios de nucleótidos, se evaluaron las secuencias UTR 3' con 97,5 %, 95,0 %, 92,5 % y 90,0 % de homología con el elemento FI original para determinar la expresión de proteína total y la vida media del ARNm en hiDC. Los nucleótidos que se cambiaron se eligieron aleatoriamente a lo largo de toda la secuencia (secuencias 208-211, modificaciones aleatorias). Los ARNm de luciferasa con estos elementos modificados como 3'UTR se transcribieron *in vitro*, se electroporaron en hiDC y su expresión se siguió en el tiempo mediante las mediciones de luciferasa después de 3, 6, 24, 48 y 72 h. Los ARNm de luciferasa con el elemento FI modificado produjeron la misma cantidad global de proteína y tuvieron aproximadamente la misma vida media (Tabla 17).

Además de las sustituciones aleatorias con grados crecientes como se describió anteriormente, se generó otro conjunto de elementos FI modificados mediante la introducción racional de sustituciones de nucleótidos que probablemente alteren la estructura secundaria del elemento FI. Para múltiples secuencias UTR 3' naturales se conoce que su estructura secundaria es importante porque proporciona los sitios de unión para proteínas reguladoras, que influyen en la estabilidad del ARNm (Address y otros, 1997; Putland y otros, 2002; Crucs y otros, 2000; Adams y otros, 2003). Dos secuencias de 8nt que son perfectamente complementarias entre sí están presentes en el elemento FI, una en el elemento F y la otra en el elemento I (Figura 10). El apareamiento de bases de estas dos regiones puede verse, además, en la mayoría de las predicciones múltiples. mFold (Zuker, 2003) es un programa de computadora que permite predicciones de estructura secundaria de secuencias de entrada. Para verificar la importancia de este elemento de estructura secundaria específico, la secuencia se cambió de una manera que elimina el emparejamiento de bases (secuencia 212, mutación 8nt). Además de estas secuencias complementarias bastante largas, las predicciones mfold para FI 3'UTR se seleccionaron para elementos de estructura presentes en la mayoría de los pliegues de salida, que por lo tanto deberían tener una alta probabilidad de formación *in vivo*. Los nucleótidos involucrados en el apareamiento de bases de estos pliegues se cambiaron a 97,5 %, 95,0 %, 92,5 % y 90,0 % de homología con las secuencias FI originales mediante intercambio con sus parejas de apareamiento de bases, de esta manera se retiene la estructura secundaria de la secuencia (secuencias 217-220, modificaciones que retienen la estructura). Además, las mismas secuencias se intercambiaron en una sola cadena de la parte bicatenaria, de esta manera se destruye deliberadamente la estructura secundaria. En estos casos, la identidad de la secuencia original fue 98,75%, 97,50%, 96,25% y 95,00%, respectivamente (secuencias 213-216, modificaciones desestabilizadoras de la estructura).

Los ARN de luciferasa con los elementos UTR 3' modificados descritos se transcribieron *in vitro*, se electroporaron en hiDC y su expresión se siguió en el tiempo mediante las mediciones de luciferasa después de 3, 6, 24, 48 y 72 h. Con ninguna estrategia de modificación pudo observarse un impacto significativo en la vida media del ARNm. Por lo tanto, las propiedades estabilizadoras del elemento FI parecen ser robustas frente a los cambios en su secuencia de nucleótidos o estructura secundaria al menos hasta 10,0 % de nucleótidos variados. Además, no pudo observarse una disminución en la cantidad de proteína total después de la modificación de la secuencia de FI (Tabla. 18 A y B).

Ejemplo 10: El uso del elemento FI en lugar de 2hBg evita errores de hibridación en la amplificación basada en PCR de la región codificante del ARN

Como se ha demostrado, el elemento FI es igual o superior al 2hBg 3'UTR con respecto a la estabilidad del ARNm y a la eficiencia de la traducción. Otra ventaja del elemento FI es su secuencia no repetitiva, mientras que las dos copias de hBg 3'UTR pueden causar problemas en algunos casos.

Esto es más obvio cuando el molde de ADN para la transcripción de ARN se amplifica mediante PCR. En tales casos, se adiciona la cola poliA de longitud completa con el oligo cebador 3' que se une en el propio extremo 3' de la UTR 3' (Figura 11 A). En el caso de 2hBgUTR, emergen productos secundarios truncados durante la PCR, que después de la secuenciación resultó que consisten en ARNm con solo la repetición de 1hBg en la UTR (Figura 11 B). Después de la transcripción, el truncamiento es, además, visible en el ARNm (Figura 11C). Este fenómeno ocurre en la mayoría de las reacciones de PCR con construcciones que contienen el elemento 2hBgUTR y no puede eliminarse completamente mediante esfuerzos de optimización que incluyen temperatura de hibridación del cebador, composición del tampón, secuencia del cebador o polimerasas alternativas. Incluso después de la inserción de una secuencia de enlace única entre la UTR 3' y la cola de poliA, el problema persiste. Es importante destacar que la fuerza del pico lateral se correlacionó con el rendimiento de la reacción de PCR, lo que indica errores de hibridación de los fragmentos de PCR cortos truncados,

que aumentan con cada ciclo de PCR, como causa probable del problema. Por lo tanto, no pudieron identificarse condiciones satisfactorias para los moldes de ADN codificantes para los ARN con el 2hBg UTR-3'.

Por el contrario, la PCR de los moldes de ADN con el elemento FI no produjo ningún producto secundario truncado (Figura 11 D), y además el ARNm resultante no mostró un pico adicional en el perfil del Bioanalizador (Figura 11 E). Por lo tanto, el elemento FI constituye una mejora considerable como una 3'UTR en comparación con el 2hBgUTR con respecto a la integridad del molde de PCR y a la calidad del ARN correspondiente.

Ejemplo 11: Propiedades estabilizadoras de ARN de los subfragmentos de los elementos F e I

Debido al procedimiento de selección que se aplicó para identificar los elementos estabilizadores UTR novedosos, se obtuvieron secuencias en un determinado intervalo de tamaño. La identificación de las mismas secuencias con extremos prolongados en 5' y 3' proporcionó una primera indicación de la longitud mínima requerida. Sin embargo, la región mínima requerida para que cada elemento ejerza su efecto estabilizador podría ser incluso más corta.

Con este fin, tanto para el elemento F como para el I, se diseñaron cinco construcciones reporteras de Luciferasa, cada una de las cuales contiene una UTR acortada que cubre un fragmento diferente del elemento original acortado en el extremo 5' y/o 3' (ver Figura 12 paneles superiores A y B, respectivamente). Estas construcciones reporteras se transcribieron *in vitro*, se electroporaron en hiDC y su expresión se siguió en el tiempo mediante las mediciones de luciferasa a las 3, 6, 24, 48 y 72 h después de la electroporación. Las curvas de expresión resultantes se analizaron para la vida media relativa del ARN con el ARN que contiene la longitud completa establecido en 1 (ver Figura 12 paneles inferiores A y B, respectivamente).

Para el elemento F, no pudo observarse una vida media del ARNm significativamente disminuida para ninguna subsecuencia evaluada, lo que indica una participación redundante y no cooperativa de varias subsecuencias a lo largo del elemento F en su papel estabilizador. Se podría obtener un resultado similar para el elemento I, aunque aquí pudo observarse una ligera caída en el rendimiento cuando solo se usó la región central (nt37-107) como 3'UTR.

Para poner estos resultados en perspectiva, los elementos F e I individuales de longitud completa, así como también la combinación FI, se compararon con una UTR 3' seleccionada al azar de la biblioteca inicial (257nt de longitud). Esto se obtuvo mediante la clonación de un conjunto de ADN inicial y mediante la selección aleatoria de un clon único. Como se describió anteriormente, los ARN codificantes de luciferasa con las secuencias UTR respectivas se electroporaron en hiDC, se midió la expresión de luciferasa en el tiempo, y se calcularon las vidas medias relativas y la expresión de proteína total. En comparación con los elementos F, I y FI, el ARN con la UTR 3' seleccionada aleatoriamente es significativamente menos estable (Figura 13, panel superior). El efecto de las UTR seleccionadas es aún más pronunciado para la expresión de proteína total (Figura 13, panel inferior). Esto indica claramente que el efecto de los fragmentos de los elementos F e I como se describió anteriormente son específicos para las secuencias seleccionadas y no causado simplemente por la presencia de una secuencia UTR 3'. Esto está en línea con el aumento observado en la estabilidad del ARN del conjunto durante la selección (ver arriba).

Ejemplo 12: Uso de elementos UTR estabilizadores para la reprogramación de células madre

Se sembraron 40,000 células en una placa de 12 pocillos y se lipofectaron durante tres (3x) o cuatro (4x) días consecutivos con mezclas de ARNm que estaban compuestas de 0,33 μ g de ARN sin modificar transcrito *in vitro* (IVT) que codifica la reprogramación TF OCT4, SOX2, KLF4, cMYC, NANOG y LIN28 (OSKMNL) (1:1:1:1:1:1) con 0,08 μ g de cada B18R, E3 y K3 (EKB) y 0,1,7 μ g de una mezcla de miARN compuesta de miARN 302a-d y 367 (1:1:1:1:1:1). De esta manera, las construcciones de ARN solo diferían en su 3'UTR que consiste en una repetición en tándem de la 3'UTR de la β -globina humana (2hBg), el elemento F-I (FI) o el elemento I-F (IF). Las células se cultivaron en medio de células madre embrionarias humanas (hES) y las lipofecciones se realizaron mediante el uso de RNAiMAX de acuerdo con las instrucciones del fabricante. Desde el día 9 en adelante, se observó la formación de colonias y el análisis de colonias se realizó en el d11 (ver la Figura 14A para una descripción general de la línea de tiempo). Las colonias establecidas se tiñeron para fosfatasa alcalina (AP) el día 11 mediante el uso de un kit de tinción AP. Para una visión general, las tinciones representativas se muestran en la Figura 14B. Se hizo obvio que la incorporación del elemento FI resulta en cantidades más altas de colonias AP positivas (oscuras). Se contaron las colonias teñidas para AP y se confirmaron los resultados de la visión general: En comparación con el 2hBg-UTR usado anteriormente, la sustitución con el FI-UTR conduce a un exceso de colonias de 3-4 veces cuando las células se lipofectaron 3 veces. La sustitución con el IF-UTR resulta en un exceso de 2 veces. Con cuatro transfecciones, estos efectos son menos pronunciados. Aquí no se observó ninguna mejora con el IF-UTR. Por un lado, el proceso parece estar saturado con cuatro transfecciones, mientras que por el otro lado el recuento de colonias está sesgado en cierta medida debido al crecimiento excesivo de colonias (ver Figura 14C). La morfología de la colonia de las colonias de células iPS resultantes mediante el uso de ARN que contenían el FI-UTR era similar a una célula hES, con células pequeñas muy compactas en colonias distintas y bordes bien definidos (figura 14D). Estas colonias podrían teñirse positivas para AP (Figura 14E) y el marcador de superficie celular de las hES TRA-1-60 (Figura 14F). La tinción vital TRA-1-60 se realizó con el anticuerpo Stain-Alive TRA-1-60 (Stemgent) de acuerdo con las instrucciones del fabricante. Se muestran imágenes representativas de colonias. Para evaluar aún más la pluripotencia de las colonias, las células se sedimentaron, se aisló el ARN total y la expresión del ARNm de los marcadores de hES OCT4 (endógeno), NANOG (endógeno), LIN28 (endógeno), TERT y REX1 se cuantificó mediante qRT-PCR. La expresión

de ARNm se normalizó con respecto a la de HPRT y se muestra como las veces de inducción en comparación con los niveles de los transcritos de las células de entrada. El análisis de colonias después de 3 lipofecciones se muestra en la Figura 14G. Todos los marcadores analizados se expresaron altamente en comparación con las células de entrada, lo que indica la pluripotencia de las células reprogramadas. La superioridad del ARNm sintético que contiene FI se confirmó mediante una mayor expresión del marcador endógeno en comparación con la reprogramación con los ARNm que contienen 2hBg e IF.

Estos resultados muestran, que la sustitución de la 2hBg-UTR con la FI-UTR resulta en una tecnología de reprogramación basada en el ARN más rápida y eficiente. Esto probablemente se basa en la expresión más larga y más alta de los factores de transcripción de reprogramación que resultan de la sustitución con el elemento FI. La orientación del elemento FI parece de esta manera indispensable, ya que el beneficio no se observó con las construcciones IF. La reprogramación exitosa de las células mediante el ARNm que contienen FI se confirmó por la morfología similar a las células hES, la actividad AP y la expresión de los marcadores de la superficie celular de las hES y de los marcadores endógenos de las colonias de las células iPS resultantes.

Tablas

Tabla 1 Vida media del ARNm en horas (h) calculada a partir de datos de experimentos de transcriptasa inversa-PCR (RT-PCR) en tiempo real para monitorear el progreso de la selección. Los ARNm se cuantificaron 8, 24 y 48 horas después de la electroporación. En el experimento I (izquierda), cada muestra se analizó solo una vez. En consecuencia, no se proporciona una desviación estándar.

Muestra	Vida media del ARNm	Muestra	Vida media del ARNm
2hBg	7,5 h	2hBg	13,5 ± 0,2 h
lib	4,5 h	Rn4	13,9 ± 0,7 h
Rn1	4,9 h	Rn5	16,5 ± 0,7 h
Rn2	6,7 h		
Rn3	7,5 h		

Tabla 2 Descripción general de los 7 grupos principales con la región de unión (BR) dentro de la UTR-3' de la secuencia BLASTed. Se muestran la abreviatura del grupo, el número de clones identificados para el grupo (núm.), el origen genómico con la abreviatura respectiva (Abr.), el código NCBI y la posición dentro de la secuencia con respecto a la región codificante. De acuerdo con NextBio, todas las secuencias se regulan positivamente en hiDC.

Grupo	núm.	Resultado del BLAST con la secuencia representativa de cada grupo. Homo Sapiens	Abr.	Código NCBI	BR
B	50	Fragmento Fc de IgG. receptor. transportador. alfa. ARNm (clon de ADNc)	FCGRT	NM_001136019	UTR-3'
D	22	Proteína 1 específica de linfocitos. ARNm	LSP1	NM_002339	UTR-3'
E	13	Ligando 22 de quimiocina (motivo C-X-C). ARNm	CCL22	NM_002990	UTR-3'
F	4	Potenciador de la Separación del Amino Terminal. ARNm	AES	NM_198969	UTR-3'
G	15	Familia de fosfolipasa D. miembro 3. ARNm	PLD3	NM_001031696	CDS+UTR-3'
I	17	ARN 12S codificado mitocondrialmente	MT-RNR1	NC_012920	ncRNA
J	22	Complejo principal de histocompatibilidad, clase II. DR beta 4. ARNm	HLA-DRB4	NM_021983	UTR-3'

Tabla 3 Valores calculados en relación con nuestro estándar de oro 2hBg para la vida media y la proteína total a lo largo del tiempo. Se muestran el nombre del grupo y el gen respectivo.

Gen	RefSeq	Con relación a 2hBg	
		Vida media	Proteína total en el tiempo
Fragmento Fc de IgG, receptor, transportador, alfa	NM_001136019	0,89 ± 0,15	0,96 ± 0,15
Proteína 1 específica de linfocitos	NM_002339	0,80 ± 0,21	0,75 ± 0,03
Ligando de quimiocina 22	NM_002990	0,82 ± 0,16	0,66 ± 0,12
Potenciador de la Separación del Amino Terminal	NM_198969	0,90 ± 0,06	0,95 ± 0,01
Miembro 3 de la familia de Fosfolipasa D	NM_001031696	0,79 ± 0,21	0,66 ± 0,13
ARN 12S codificado mitocondrialmente	NC_012920	1,15 ± 0,09	0,94 ± 0,08
Complejo principal de histocompatibilidad clase II DR beta 4	NM_021983	0,89 ± 0,08	0,89 ± 0,09

Tabla 4. Condiciones de PCR para la amplificación de la biblioteca y las etapas de selección posterior.

Tiempo	Temperatura	Etapas
1 min 30 s	98 °C	Desnaturalización inicial
20 s	98 °C	Desnaturalización
30 s	65 °C	Hibridación
45 s	72 °C	Extensión
5 min	72 °C	Extensión final
∞	4 °C	Mantenimiento



Tabla 5 Reacción de transcripción IVT-T7.

	Conc./Vol.	Conc. Final
ddH2O		Ad 50 µL
Casquete D1	Variable	6,0 mM
ATP/CTP/UTP	100 mM	7,5 mM
GTP	100 mM	1,5 mM
Tampón T7	10x	1x
Producto de la PCR	Variable	0,05 µg/µL
enzima T7 mix HC	10x	1x

Tabla 6. Combinaciones clonadas y comparadas con nuestro estándar de oro 2hBg (esquina inferior derecha). Se resaltan los elementos únicos clonados dos veces.

	I	G	B	D	J	E	F	hBg
I	II	GI	BI	DI	JI	EI	FI	hBgI
G	IG	GG	BG	DG	JG	EG	FG	hBgG
B	IB	GB	BB	DB	JB	EB	FB	hBgB
D	ID	GD	BD	DD	JD	ED	FD	hBgD
J	IJ	GJ	BJ	DJ	JJ	EJ	FJ	hBgJ
E	IE	GE	BE	DE	JE	EE	FE	hBgE
F	IF	GF	BF	DF	JF	EF	FF	hBgF
hBg	lhBg	ghBg	bhBg	dhBg	jhBg	ehBg	fhBg	2hBg

Tabla 7 Resultado de FCGRT (grupo B) clonado como elemento único o combinado en dirección 5' con una de las otras secuencias del grupo como elemento en dirección 3'. Los valores en negrita son > 1,0. Los valores son con relación a 2hBg

Muestra	Vida media	Con relación a 2hBg	
		Eficiencia de la traducción	Proteína total en el tiempo
B	0,840	1,320	1,300
BB	0,580	1,530	0,900
BI	0,920	1,750	1,410
BG	0,780	2,300	1,430
BD	0,730	1,970	1,220
BJ	0,710	1,910	1,190
BE	0,720	1,500	1,030
BF	0,760	1,720	1,220
BhBg	0,970	2,200	1,740
hBgB	0,640	1,750	1,030
2hBg	1,000	1,000	1,000

Tabla 8 Resultado de LSP1 (grupo D) clonado como elemento único o combinado en dirección 5' con una de las otras secuencias del grupo como elemento en dirección 3'. Los valores en negrita son > 1,0. Los valores son con relación a 2hBg.

	Muestra	Con relación a 2hBg		
		Vida media	Eficiencia de la traducción	Proteína total en el tiempo
5	D	0,770	0,860	1,250
	DD	0,680	1,130	1,000
	DI	0,960	1,440	1,270
	DG	0,700	1,530	1,110
10	DB	0,640	0,900	0,760
	DJ	0,640	1,040	0,890
	DE	0,690	1,000	0,970
15	DF	0,750	1,080	1,000
	DhBg	0,840	1,120	1,020
	hBgD	0,820	1,490	1,160
	2hBg	1,000	1,000	1,000

Tabla 9 Resultado de CCL22 (grupo E) clonado como elemento único o combinado en dirección 5' con una de las otras secuencias del grupo como elemento en dirección 3'. Los valores en negrita son > 1,0. Los valores son con relación a 2hBg.

	Muestra	Con relación a 2hBg		
		Vida media	Eficiencia de la traducción	Proteína total en el tiempo
25	E	0,760	0,970	0,940
	EE	0,600	0,950	0,670
	EI	0,890	1,120	0,960
	EG	0,680	1,590	0,940
30	EB	0,570	1,470	0,850
	ED	0,650	1,350	0,950
	EJ	0,600	1,230	0,760
	EF	0,760	1,100	0,860
35	EhBg	0,690	1,190	0,780
	hBgE	0,880	1,630	1,050
	2hBg	1,000	1,000	1,000

Tabla 10 Resultado de AES (grupo F) clonado como elemento único o combinado en dirección 5' con una de las otras secuencias del grupo como elemento en dirección 3'. Los valores en negrita son > 1,0. Los valores son con relación a 2hBg.

	Muestra	Con relación a 2hBg		
		Vida media	Eficiencia de la traducción	Proteína total en el tiempo
50	F	0,500	1,760	0,970
	FF	0,910	1,770	1,410
	FI	1,100	1,490	1,290
	FG	0,850	1,680	0,980
55	FB	0,720	1,360	0,860
	FD	0,490	1,350	0,620
	FJ	0,780	1,720	1,090
	FE	0,730	1,660	1,080
60	FhBg	1,050	1,900	1,530
	hBgF	0,940	2,250	1,500
	2hBg	1,000	1,000	1,000

Tabla 11 Resultado de PLD3 (grupo G) clonado como elemento único o combinado en dirección 5' con una de las otras secuencias del grupo como elemento en dirección 3'. Los valores en negrita son > 1,0. Los valores son con relación a 2hBg.

5	Muestra	Con relación a 2hBg		
		Vida media	Eficiencia de la traducción	Proteína total en el tiempo
10	G	0,740	1,260	1,110
	GG	0,480	1,080	0,690
	GI	0,990	1,010	1,000
	GB	0,520	0,970	0,620
	GD	0,630	1,170	0,780
15	GJ	0,520	0,940	0,640
	GE	0,500	0,730	0,550
	GF	0,620	0,790	0,680
20	GhBg	0,740	0,990	0,860
	hBgG	0,720	1,160	0,910
	2hBg	1,000	1,000	1,000

Tabla 12 Resultado de mtRNRI (grupo I) clonado como elemento único o combinado en dirección 5' con una de las otras secuencias del grupo como elemento en dirección 3'. Los valores en negrita son. Los valores son con relación a 2hBg.

25	Muestra	Con relación a 2hBg		
		Vida media	Eficiencia de la traducción	Proteína total en el tiempo
30	I	1,080	1,020	1,440
	II	1,170	0,830	1,030
	IG	1,040	1,250	1,310
	IB	1,100	1,200	1,180
35	ID	1,190	1,580	1,510
	IJ	1,080	1,430	1,330
	IE	1,060	1,000	1,070
40	IF	1,220	1,130	1,290
	IhBg	1,230	1,110	1,210
	hBgl	1,210	1,420	1,270
	2hBg	1,000	1,000	1,000

Tabla 13 Resultado de HLA-DRB4 (grupo J) clonado como elemento único o combinado en dirección 5' con una de las otras secuencias del grupo como elemento en dirección 3'. Los valores en negrita son > 1,0. Los valores son con relación a 2hBg.

50	Muestra	Con relación a 2hBg		
		Vida Media	Eficiencia de la traducción	Proteína total en el tiempo
55	J	0,790	0,930	0,920
	JJ	0,490	0,960	0,540
	JI	0,880	1,110	0,900
	JG	0,420	1,280	0,630
	JB	0,480	1,000	0,520
60	JD	0,500	1,370	0,830
	JE	0,420	0,950	0,520
	JF	0,570	1,190	0,800
	JhBg	0,730	1,100	0,800
65	hBgJ	0,770	1,530	1,080
	2hBg	1,000	1,000	1,000

Tabla 14 Resultados representativos mediante el uso de luc2mut como gen reportero y las UTR-3' recién seleccionadas después de la electroporación en hiDC. La actividad de luciferasa se midió durante 96 h. Los valores son con relación a 2hBg.

Muestra	Con relación a 2hBg		
	Vida media	Eficiencia de la traducción	Proteína total en el tiempo
noUTR	0,300	0,694	0,139
hBg	0,360	1,216	0,437
I	0,800	1,132	0,936
IF	1,110	1,050	1,133
FI	1,020	0,818	0,847
lhBg	0,880	0,860	0,792
hBgl	0,840	0,776	0,681
2hBg	1,000	1,000	1,000

Tabla 15: Condiciones de Electroporación

La tabla resume los detalles del protocolo de electroporación para todos los tipos de células usados. La cantidad de células indicadas bajo el recuento celular se mezcló con la cantidad de ARN indicada en μg o pmol tanto en las cubetas de electroporación como en las placas de electroporación de 96 pocillos (como se indica en el formato) en medio X-VIVO15 (Lonza). La electroporación se realizó mediante la aplicación de un pulso con la longitud y el voltaje designado enumerado bajo V. Posteriormente, la suspensión celular se diluyó en medio de crecimiento y se distribuyó en 96 pocillos con la densidad enumerada bajo células/puntos de tiempo.

	Formato	tamaño de la brecha	recuento de células	ARN [μg]	ARN [pmol]	células/puntos de tiempo	V	pulso
hiDCs	cubeta	4mm	1,00E+06		10	5,00E+04	300	1×12ms
HFF	96-pocillos	4mm	7,00E+04	2		1,00E+04	200	1×24ms
CD8+	cubeta	4mm	2,50E+06		10	1,67E+05	500	1×3ms
CD4+	cubeta	4mm	2,50E+06		10	1,67E+05	500	1×3ms
MEF	96-pocillos	4mm	7,00E+04	2		1,00E+04	200	5×6ms/ 400ms
C2C12	96-pocillos	4mm	7,00E+04	2		1,00E+04	240	5×5ms/ 400ms
bmDCs	cubeta	4mm	1,00E+06		10	5,00E+04	400	1×5ms

Tabla 16

Vida media y proteína total del elemento FI con relación a 2hBgUTR que contiene ARNm no modificado y modificado, después de la electroporación, y ARN no modificado después de la lipofección. Los plásmidos que codifican el gen de luciferasa de luciérnaga que contiene ya sea FI o 2hBg como 3'UTR se linealizaron en dirección 3' del poli(dA:dT) con una enzima de restricción de clase IIS, de esta manera se genera un molde sin nucleótido adicional más allá del poli(dA:dT). El ADN plasmídico linealizado se purificó mediante el uso de perlas magnéticas carboxiladas (Invitrogen), se cuantificó por espectrofotometría y se sometió a transcripciones *in vitro*. Para las transcripciones *in vitro* de fabricación casera se usaron las ARN polimerasas T7 suplementadas con inhibidores de RNasa y pirofosfatasa con NTP 7,5 mM en un tampón de Hepes 125mM pH 8,35, MgOAc2 34mM, DTT 10mM y espermidina 2mM. Para una eficiente formación del casquete del ARN se añadió β -S-ARCA(D2) 6 mM a la reacción y la concentración inicial de GTP se redujo a 1,5 mM, que se ajustó a 7,5 mM en un proceso de alimentación por lotes durante 2,5 h a 37 °C. El ARN se purificó mediante perlas magnéticas carboxiladas (Invitrogen) y la concentración y calidad del ARN se evaluaron mediante espectrofotometría y análisis en un Bioanalizador 2100 (Agilent).

A) Muestra que la vida media de los ARNm no modificados que contienen el elemento FI son mayores o comparables a los que contienen el 2hBg 3'UTR en varias líneas celulares humanas y murinas. La cantidad de fibroblastos humanos (HFF), células T CD8+ y CD4+, fibroblastos embrionarios murinos (MEF), células de mioblastoma (C2C12) y DC murinas como se enumeran en la Tabla 15 se mezclaron con la cantidad respectiva de ARN (Tabla 15) en medio X-VIVO15 (Lonza) y se sometieron a electroporación. El número indicado de células se sembró en placas de 96 pocillos en 100 μ L de medio de crecimiento apropiado con aditivos. A las 2, 6, 24, 48, 72 y 96 horas después de la siembra, se determinó la actividad de luciferasa de luciérnaga mediante la adición de Luciferina (Promega) en un lector de fluorescencia (TECAN).

B) Muestra que la vida media de los ARNm modificados m1Y que contienen el elemento FI son más altas o comparables a las que contienen el 2hBg 3'UTR en diferentes líneas celulares humanas y murinas. La cantidad de células dendríticas inmaduras humanas (iDC), fibroblastos (HFF), células T CD8+ y CD4+, fibroblastos embrionarios murinos (MEF), células de mioblastoma (C2C12) y DC murinas como se enumeran en la Tabla 15 se mezclaron con la cantidad respectiva de ARN modificado m1Y (Tabla 15) en medio X-VIVO15 (Lonza) y se sometieron a electroporación. El número indicado de células se sembró en placas de 96 pocillos en 100 μ L de medio de crecimiento apropiado con aditivos. A las 2, 6, 24, 48, 72 y 96 horas después de la siembra, se determinó la actividad de luciferasa de luciérnaga mediante la adición de Luciferina (Promega) en un lector de fluorescencia (TECAN).

C) Muestra que la vida media de los ARNm no modificados que contienen el elemento FI son mayores o comparables a los que contienen el 2hBg 3'UTR en diferentes líneas celulares, además, cuando el ARN se transfectó mediante lipofección. Se incubaron 50 ng de ARN durante 15-30 minutos con 0,2 μ L RNAiMAX y se administró a las células 1E04, HFF, MEF o C2C12 en 96 pocillos. Los niveles de luciferasa se midieron a las 3, 6, 12, 24, 48, 72 y 96 h mediante la adición de Luciferina (Promega) en un lector de fluorescencia (TECAN).

	A		B		C	
	ARNm no modificado		ARNm modificado m1Y		Lipofección	
	con relación a 2hBg		con relación a 2hBg		con relación a 2hBg	
	Vida media proteína total		Vida media proteína total		Vida media proteína total	
	hiDCs		1,29		2,24	
	C2C12		1,58		1,09	
	HFF		1,83		1,14	
	MEF		1,18		1,11	
	CD4+		1,02		1,46	
	CD8+		1,05		1,33	
	bmDC		1,09		1,34	

Tabla 17: 10 μ g del ARN que codifica para la luciferasa de luciérnaga que contiene ya sea el elemento FI o variaciones del elemento FI con la homología designada a la secuencia FI original como 3'UTR se electroporaron en hiDC en un formato de 96 pocillos. La expresión de luciferasa se siguió en el tiempo a las 3, 6, 24, 48, y 72 h, y a partir de la curva de expresión resultante se calculó la vida media de ARNm y la cantidad de proteína total traducida a partir del ARN.

	con relación a la secuencia FI		
	Vida media		proteína total
	% de homología		
	97,5	1,0+/-0,1	1,3+/-0,2
	95,0	1,0+/-0,0	1,2+/-0,2
	92,5	1,1+/-0,1	1,4+/-0,1
	90,0	0,9+/-0,1	1,1+/-0,2

Tabla 18: 10µg del ARN que codifica para la luciferasa de luciérnaga que contiene ya sea el elemento FI o variaciones del elemento FI que contiene la estructura que retiene o destruye las mutaciones y con la homología designada a la secuencia FI original como 3'UTR se electroporaron en hiDC en un formato de 96 pocillos. La expresión de luciferasa se siguió en el tiempo a las 3, 6, 24, 48, y 72 h, y a partir de la curva de expresión resultante se calculó la vida media de ARNm y la cantidad de proteína total.

5

10

15

20

A			B		
Modificaciones que retienen la estructura			Modificaciones que desestabilizan la estructura		
con relación a la secuencia FI			con relación a la secuencia FI		
Vida media	proteína total		Vida media	proteína total	
% de homología			% de homología		
97,5	1,2+/-0,1	1,6+/-0,3	98,75	1,1+/-0,1	1,5+/-0,1
95,0	1,1+/-0,1	1,7+/-0,3	97,50	1,1+/-0,1	1,4+/-0,1
92,5	1,1+/-0,1	1,5+/-0,3	96,25	1,0+/-0,1	1,5+/-0,1
90,0	1,1+/-0,2	1,4+/-0,1	95,00	1,0+/-0,0	1,1+/-0,2
			Mutación 8nt	0,9 +/- 0,0	1,3 +/- 0,4

Las secuencias descritas en la presente descripción son las siguientes:

Grupo B

25

>Rn5-2pl-A4_For2

30

CAUCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUCCUGGAAUCUGACCAUUCGUUGUC
 UGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUG
 AGACUGACUCCCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUG
 CCGCUGCUGAUCCAUGCCGGUGUGACC

35

>Rn5-2pl-A3_For2

40

GCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUCCUGGAAUCUGACCAUUCGUUGUCUGCUAU
 GCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUG
 ACUCCCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUG
 CUGAUCCAUGCCGGUGUGACC

45

>Rn5C5_For2

50

UUCUGCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUCCUAGAAUCUGACCAUUCGUUG
 UCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCA
 UGAGACUGACUCCCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCU
 UGCCGCUGCUGAUCCAUGCCGGUGAGACC

55

>Rn5E6_For2

UGCUGCUGCUGCUGCUGCGGGUCUUCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCC
 UCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCCA
 CUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCA
 UUGCCGGCGGACA

>Rn6-1WoC3_For2

GCUGCUGCUGCGGGUCUUCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAA
 GACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCCAUGCUCU
 GCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUUGCCGG
 CGUACC

>Rn6-1WoB12_For2

CUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUG
 CUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCCAUGCUCUGCCUGCCUCUCCCCAC
 UGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUUGCCGGCGGACC

>Rn6-1WoB1_For2

UCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUUCUGGAAUCUGACCAUUCGUUGUCUG
 CUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAG
 ACUGACUUCCCAUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCC
 GCUGCUGAUCCAUUGCCGGUGGGACC

>Rn6-1WoF3_For2

CUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCCAUGCUCUGCCUGC
 CUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUUGCCGGUAGAAC
 C

>Rn6-1Wo_H11_b

UCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUUCUGGAAUCUGACCAUUCGUUGUCUG
 CUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAG
 ACUGACUUCCCAUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCC
 GCUGCUGAUCCAUUGCCGGUGGGACC

>Rn6-2WoG8_b

GCUGCUGCUGCUGCUGCUGCGGGUCUUCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGU
 CCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUCC
 CACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUC
 CAUUGCCGGUGUGACC

>Rn5-2pl-B3_For2

UCUGGCCUCACUGAGUCUGAAGAGCUGUUAACUACCAUGGCCAGUCCUCCCUGAGUCUGA
CCAUCUCCAUCUCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUCCUGGAAUCUGACCAU
UCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCG
GGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCC
CCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn5_F5_b

UCUGGCCUCACUGAGUCUGAAGAGCUGUUAACUACCAUGGCCAGUCCUCCCUGAGUCUGA
CCAUCUCCAUCUCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUCCUGGAAUCUGACCAU
UCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCG
GGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCC
CCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn5B8_For2

CUACCAUGGCCAGUCCUCCCUGAGUCUGACCAUCUCCAUCUCCUGCUGCUGCUGCUGCUGC
UGCUGCGGGUCUCCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGA
CUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGC
CUGCCUUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUG
UGACC

>Rn6-1WoH9_For2

GUCCUCCCUGAGUCUGACCAUCUCCAUCUCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCU
UCCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGC
UGCUIUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCC
ACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUCUGACA

>Rn6-2WoC11_For2

GUCCUCCCUGAGUCUGACCAUCUCCAUCUCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCU
UCCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGC
UGCUIUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCCCCA
CUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn5_C3_b

CCAUCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUUCUGGAAUCUGACCAUUCGUUGU
 CUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUGCUACUGCCCGGGCCCAUG
 AGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUG
 CCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-2WoH5_For2

GCCAGUCCUCCUGAGUCUGACCAUCUCCAUCCUGCUGCUGCUGCUGCUGCUGCGG
 GUCUUCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGC
 CUGCUGAUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCU
 CCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-96hE12_For2

UGCCUUCGGUCUCCUGCUGCUCUGGCCUCACUGAGUCUGAAGAGCUGUUAACUACCAUG
 GCCAGUCCUCCUGAGUCUGACCAUCUCCAUCCUGCUGCUGCUGCUGCUGCUGCGG
 GUCUUCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGC
 CUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCU
 CCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-96h-2pl-E9_F

GGCCAGUCCUCCUGAGUCUGACCAUCUCCAUCCUGCUGCUGCUGCUGCUGCUGCG
 GGUCUUCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUG
 CCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUC
 UCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-96h-2pl-H10_

GGCCAGUCCUCCUGAGUCUGACCAUCUCCAUCCUGCUGCUGCUGCUGCUGCUGCG
 GGUCUUCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUG
 CCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUC
 UCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-1WoB11_For2

UGACCAUCUCCAUCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUUCUGGAAUCUGAC
 CAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGC
 CCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACA
 GCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-1WoF7_For2

CCAGUCCUCCUGAGUCUGACCAUCUCCAUCCUGCUGCUGCUGCUGCUGCUGCUGCGGG
 UCUUCCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCC
 UGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCCUUGCUCUGCCUGCCUCUC

CCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUUGCCUGUGUGACCA

>Rn6-1WoA7_For2

UGACCAUCUCCAUCCUGCUGCUGCUGCUGCUGCUGCGGGUCUUCUGGAAUCUGACCAUUCG
 UUGUCUGCUAUGCCCGUCCUACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGC
 CCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCG
 CCUUGUCGCUGCUGAUCCAUUGCCGGUGUGACAC

>Rn6-2WoD11_b

GACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCU
 GCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUUGCCGG
 UGUGACCC

>Rn6-2WoG3_For2

CUCCUGAGUCUGACCAUCUCCAUCCUGCUGCUGCUGCUGCUGCUGCGGGUCUUC
 UGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGC
 UUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACU
 GCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUUGCCGGUGUGACC

>Rn6-2WoC2_For2

UCCAUCCUGCUGCUGCUGCUGCUGCUGCUGCGGGUCUUCUGGAAUCUGACCAUUCGUU
 GUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCC
 AUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCC
 UUGCCGCUGCUGAUCCAUUGCCGGUGUGACC

>Rn6-1WoD6_For2

UGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
 GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
 GCUGAUCCAUUGCCGGUGUGACC

>Rn6-1WoD10_For2

UGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
GCUGAUCCAUGCCGGUGUGACC

>Rn6-2WoG5_For2

GCGGGUCUCCUGGAAUCUGAACAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGA
CUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGC
CUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUUGCUGAUCCAUGCCGGUGUGAC
C

>Rn6-96h-2pl-G8_F

GUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGG
CCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCC
GCCUUGCCGCUUGCUGAUCCAUGCCGGUGUGACC

>Rn6-1WoE7_For2

CAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGC
UCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUUGCUGAUCCAUGC
CGGUGUGACC

>Rn6-1Wo_A12_b

CCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGACCCAUGAGACUGACUUC
CACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUUGCUGAU
CAUUGCCGGUGUGACUGC

>Rn6-1WoG11_For2

CUUCCAUCCUGCUGCUGCUGCUGCUGCUGCGGGUCUCCUGGAAUCUGACCAUUCGU
UGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCC
CAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGC
CUUGCCGCUUGCUGAUCCAUGCCGGUGUGACCCC

>Rn6-1WoH5_For2

CUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUCCC
ACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUUGCUGAUCC
AUUGCCGGUGUGACC

>Rn6-1WoH4_For2

AAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCU
CUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCC
GGUGUGACC

>Rn6-2WoB4_For2

CUCCAGCUCGCUUCCAUUUGCUUGCAGAAGUUCUCGCGUGGUCACGAAGCUUGCGCUCC
UUGGAGGCCUCAGCAACAGCAUCAUAAGCUGAGCUUCCAGCUCUUCCUGAGCUUCUCA
GCUCUCCGCAUUUCCUGCCGCAUGGCGUCCACCUUCUGCGUGGCCACCUCCAUCUCCUCC
UCCUUGUCUCGCGAGCUGCCGGGACACCUUCUGCGCUAAGAUGGGAUACGGCAUUGAGGGA
UCAUGUGUAAGGAUCCGAUCUGCUUCUGGCCUCACUGAGUCUGAAGAGCUGUUAACUAC
CAUGGCCAGUCCUCCUGAGUCUGACCAUCUCCAUCUGCUGCUGCUGCUGCUGCUGCU
GCGGGUCUCCUGGAAUCUGACCAUUCGUUGUCUGCUAUGUCCGUCCUCACCAAGACUGA
CUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGC
CUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGAC
C

>Rn6-96h-2pl-A5_F

CUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGC
CUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUG
UGACC

>Rn6-1WoC8_For2

CCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUG
CUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUG

ACGGUGUGACC

>Rn5D1_For2

UAACUACCAUGGCCAGUCCUCCUGAGUCUGACCAUCUCCAUCUGCUGCUGCUGCUGC
UGCUGCGGGUCUCCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCGUCCUCACCAAGA
CUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGC
CUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGUUGAUCCAUGCCGGUG
UGACC

>Rn6-2WoG10_For2

CCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUG
CUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGUUGAUCCAUG
UCGGUGUGACC

>Rn6-1Wo_E4_b

CCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUG
CUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUG
CCGGUGUGACC

>Rn6-2WoF3_For2

CUGCUAUGCCUGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAU
GAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCAUGCACUGGCACAGCCCCGCCU
GCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-96h-2pl-B10

CUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAU
GAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCU
GCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-96h-2pl-C10

GCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGA
GACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGC
CGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-1WoB6_For2

UCUUCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCC
UGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUC
CCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-96h-2pl-D6_F

CUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUG
CUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCAC
UGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>Rn6-96h-2pl-E6_F

CUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUG
CUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCAC

UGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUUGCCGGUGUGACC

>Rn6-2WoF10_For2

GGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCU
UUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUG
CACUGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUUGCCGGUGUGACC

>Rn6-1WoG9_For2

CCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUG
CUCUGCCUGCCUCUCCCCACUGCACUGGCAUAGCCCCGCCUUGCCGCUGCUGAUCCAUUG
CCGGUGUGACC

>Rn6-96hC12_For2

CUUCCUGGAAUCUGACCAUUCGUUGUCUGCUAUGCCCGUCCUCACCAAGACUGACUGCCU
GCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCC
CCACUGCACUGGCAUAGCCCCGCCUUGCCGCUGCUGAUCCAUUCGGUGUGACC

Grupo D

>Rn6-1WoF2_For2

CAGACACCCGCCCCCGGCCUGGCUAAGAAUUGCUUCCUGUUGCCAGCAUGACCUACC
CUCGCCUCUUGAUGCCAUCGCGUGCCACCUCUUGCUUCCUGGACCCUUGAGCCUCUC
UGCCCUUCCACUCUCUGACC

>Rn6-2WoD8_For2

UUCAGCCAGACACCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUG
ACCUACCCUCGCCUCUUGAUGCCAUCGCGUGCCACCUCUUGCUUCCUGGACCCUUA
GCCUCUCUGCCCUUCCACUCUCUGACCCC

>Rn6-1WoD5_For2

CUCGCUUCCUGGGUCUGCAGGUCCAGCCGGCUGGCACCCUCCAUGUACCCAGGGGAGAUU
 CCAGCCAGACACCCGCCCCCGGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGAC
 CUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAGC
 CUCUCUGCCCUUCCACUCUCUGACCACCGCCC

>Rn5-2pl-D3_Fo2

UCCAGCCAGACACCCGCCCCCGGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGA
 CCUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAG
 CUUCUCUGCCCUUCCACUCUCUGG

>Rn6-2WoA8_Fo2

CGCUUCCUGGGUCUGCAGGUCCAGCCGGCUGGCACCCUCCACGUACCCAGGGGAGAUUCC
 AGCCAGACACCCGCCCCCGGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCU
 ACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAGCCU
 CUCUGCCCUUCCACUCUCUGACCACCG

>Rn6-2WoD7_Fo2

CAUGUACCCAGGGGAGAUUCCAGCCAGACACCCGCCCCCGGGCCUGGCUAAGAAGUUGC
 UUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUU
 UGCUCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCG

>Rn6-2WoB8_Fo2

CUCGCUUCCUGGGUCUGCAGGUCCAGCCGGCUGGCACCCUCCAUGUACCCAGGGGAGAUU
 CCAGCCAGACACCCGCCCCCGGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGAC
 CUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAGC
 CUCUCUGCCCCCCCCGAU

>Rn6-96h-2pl-H6_F

CCCAGCUCCCUAGGCGUCCCAUCUCGCUUCCUGGGUCUGCAGGUCCAGCCGGCUGGCACC
 CUCCAUGUACCCAGGGGAGAUUCCAGCCAGACACCCGCCCCCGGGCCUGGCUAAGAAGU
 UGCUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUC
 UUUUGCUCCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCUUUGACCCCCAUCUUA

>Rn6-96h-2pl-F10

GGCCACCGGGCAUGGGAAGUAUGAGAAGGUGCUUGUGGAAGGGGGCCCGGCUCCCUAGGC
 GUCCCAUCUCGCUUCCUGGGUCUGCAGGUCCAGCCGGCUGGCACCCUCCAUGUACCCAGG
 GAGAUUCCAGCCAGACACCCGCCCCCGGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAG
 CAUGACCUACCCUCGCCUCUUGAUGCCAUCGCGUGCCACCUCUUUGCUCCUGGACCC
 UUUAGCCUCUCUGCCCUUCCACUCUCUGACCCC

>Rn5H_For2

UGUACCCAGGGGAGAUUCCAGCCAGACACCCGCCCCCGGGCCUGGCUAAGAAGUUGCUU
 CCUGUUGCCAGCAUGACCUACCCUCGCCUCUUGAUGCCAUCGCGUGCCACCUCUUUG
 CUCCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCUCUGACCACCACCCCC

>Rn5G7_For2

CCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUUGAU
 GCCAUCCGCGUGCCACCUCUUUGCUCCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCU
 CUGACCACAGCCCC

>Rn6-1WoG5_For2

CCCGGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUUG
 AUGCCAUCGCGUGCCACCUCUUUGCUCCUGGACCCUUUAGCCUCUCUGCCCUUCCACU
 CUCUGACCACCGCCCCCGCC

>Rn6-1WoA8_For2

CCGGCUGGCACCCUCCAUGUACCCAGGGGAGAUUCCAGCCAGACACCCGCCCCCGGGCC
 UGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUUGAUGCCAUC
 GCUGCCACCUCUUUGCUCCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCUCUGACCA
 CCGCCCC

>Rn6-96h_D3_b

GCCGGCUGGCACCCUCCAUGUACCCAGGGGAGAUUCCAGCCAGACACCCGCCCCCGGGCC
 CUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCUACCCUAGCCUCUUUGAUGCCAUC
 CGCUGCCACCUCUUUGCUCCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCUCUGAC
 CACCGCCCC

>Rn6-96hC11_For2

UCCAGCCAGACACCCGCCCCCGGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGA
 CCUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUGCUCCUGGACCCUUUAG
 CCUCUCUGCCCUUCCACUCUCUGACCACCACCCC

>Rn5H1_Fo2

GCCAGACACCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCUA
CCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAGCCUC
UCUGCCCUUCCACUCUCUGACCCCCC

>Rn6-1WoG2_Fo2

UCCAGCCAGACACCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGA
CCUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAG
CCUCUCUGCCCUUCCACUCUCUGACCCCCC

>Rn6-1WoG7_Fo2

CGGCUCCCUAGGCGUCCCAUCUCGCUUCCUGGGUCUGCAGGUCCAGCCGGCUGGCACCCU
CCAUGUACCCAGGGGAGAUUCCAGCCAGACACCCGCCCCCGGCCUGGCUAAGAAGUUG
CUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUU
UUGCUCCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCUCUGACCACUGCCCC

>Rn6-96hB11_Fo2

UGCAGGUCCAGCCGGCUGGCACCCUCCAUGUACCCAGGGGAGAUUCCAGCCAGACACCCA
CCCCCGGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUU
GAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAGCCUCUCUGCCCUUCCAC
UCUCUGACCACUACCCC

>Rn6-2WoF8_Fo2

UUCAGCCAGACACCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUG
ACCUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUA
GCCUCUCUGCCCUUCCACUCUCUGACCACUGCCCC

>Rn6-96h_A9_b

CCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCUACCCUCGCC
UCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAGCCUCUCUGCCCU
UCCACUCUCUGACC

>Rn6-1WoH3_Fo2

CAGCCAGACACCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACC
UACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAGCC
UCUCUGCCCUUCCACUCUCUGAACACC

Grupo E

5 >Rn6-2WoE2_For2

GAGCCUACUCUGAUGACCGUGGCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUC

CUGCCAUUAUAGCUGCUCUCCCGCCAGAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUC

10 CUGUGGCUGUCACCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCCUGACCCUC

GAACCCAUCCUA

15 >Rn6-1WoD3_For2

GAGCCUACUCUGAUGACCGUGGCCUUGGCUCCUCCAGGAAGGCUCAGGCGCCCUACCUC

CUGCCAUUAUAGCUGCUCUCCCGCCAGAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUC

20 CAUCCUGUGGCUGUCACCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCCUGAC

CCCUCUAACCC

25 >Rn6-2WoG7_For2

GCCUACUCUGAUGACCGUGGCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCU

GCCAUUAUAGCUGCUCUCCCGCCAGAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCA

30 UCCUGUGGCUGUCACCCUUGGUCCACUGCCAUCUCCCCC

>Rn6-2WoH2_For2

35 GAGCCUACUCUGAUGACCGUGGCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUC

CUGCCAUUAUAGCUGCUCUCCCGCCAGAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUC

40 CAUCCUGUGGCUGUCACCCUUGGUCACCUCCGUGCUGUACUGCCAUCUCCCCCUGAC

CCC

>Rn6-2WoC1_For2

45 GAAGAGCCUACUCUGAUGACCGUGGCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACC

UCCUGCCAUAUAGCUGCUCUCCCGCCAGAGCCUGUGCCAACUCUCUGCAUUCUCCUGAU

CUCCAUCUCCUGUGGCUGUCAUCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCCU

50 GACCCUCUAACCCAUCUCCUGCCUCCUCCUGCAGUCAGAGGGUCCUGUCCCAACCA

>Rn6-1Wo_C12_b

55 UGUGGCCUUGGCUCCUCCAGGAAGGCUAAGGAGCCCUACCUCUCCUGCCAUAUAGCUGCU

CCCCGCCAGAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCUCCUGUGGCUGUCA

60 CCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCCUGACCC

>Rn6-1WoE12_For2

65

GCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUUAUAGCUGCUCUCCC
GCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCUCCUGUGGCUGUCACCCU
UGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCC

>Rn6-2WoF5_For2

AGAGCCUACUCUGAUGACCGUGGCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUC
CCUGCCAUUAUAGCUGCUCUCCCGCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCU
CCAUCUCCUGUGGCUGUCACCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCCUGA
CCCCUCUAAACCAUCCUCUGCCUCCUCCUCCUGCAGUCAGAGGGUCCUGUCCCAUCAGCG
AUUCCCCUGCUUAAACCCUCCAUGACUCCCCACUGCCCUAAGCUGAGGUCAGUCUCCCA
AGCCUGACAU

>Rn5-2pl-H3_For2

UAUAGCUGCUCUCCCGCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCUCCU
GUGGCUGUCACCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCCUGACCCUCUA
ACCAUCCUCUGCCUCCUCCUCCUGCAGUCAGAGGGUCCUGUCCCAUCAGCGAUUCCCU
GCUUAAACCCUCCAUGACAGCCC

>Rn6-2WoA3_For2

UCUGCAUUCUCCUGAUCUCCAUCUCCUGUGGCUGUCACCCUUGGUCACCUCCGUGCUGUCAC
UGCCAUCUCCCCCUGACCCUCUAACCAUCCUCUGCCUCCUCCUCCUGCAGUCAGAGGG
UCCUGUCCCAUCAGCGAUUCCCUUGCUUAAAGCCUCCAUGACUCCCC

>Rn6-96hF12_For2

CUCCUGCCAUUAUAGCUGCUCUCCCGCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGA
UCUCCAUCUCCUGUGGCUGUCACCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCC
UGACCCUCUAACCAUCCUCUGCCUCCUCCUCCUGCAGUCAGAGGGUCCUGUCCCAUCA
GCGAUUCCCUUGCUUAAACCCUCCAUGACUCCCCAA

>Rn6-96hE11_For2

GCCUACUCUGAUGACCGUGGCCUUGGGUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCU
GCCAUUAUAGCUGCUCUCCCGCCAGAAGUCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCA
UCCUGUGGCUGUCACCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCCUGACCC
CUCUAAACCAUCCUCUGCCUCCUCCUCCUGCAGUCAGAGGGUCCUGUCCCAUCAGCGAUU
CCCCUGCUUAAACCCUCCAUGACUCCCCUCU

>Rn6-96h-2pl-A11

CUACCUCUCCUGCCAUAUAAGCUGCUCCCCGCCAGAAGCCUGUGCCAACUCUCUGCAUUC
 CUGAUCUCCAUCCCUGUGGCUGUCACCCUUGGUCACCUCGUGCUGUCACUGCCAUCUCC
 CCCCUGACCCC

Grupo F

>Rn6-1WoB5_For2

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCCGGGUCCCAGGUAUGCUCCACCUCACCUGCCCCACUCACCACCUCUGCU
 AGUUCAGACACCUC

>Rn6-2WoE11_a

CCGGCCCUUCCCCCGUUUGAACAUUGUGUAACCGACAGUCUGCCUGGGCCACAGCCCUCU
 CACCUCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGCACCCCGAG
 UCUCUCCCCGACCCCGGGUCCCAGGUAUGCUCCACCUCACCUGCCCCACUCACCACCUC
 UGCUAGUUCAGACACCCCGCG

>Rn6-96h_E3_b

CCUUCUCCCCGUUUUGAACAUUGUGUAACCGACAGUCUGCCUGGGCCACAGCCCUCUACCC
 UGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGCACCCCGAGUCUCC
 CCGACCCCGGGUCCCAGGUAUGCUCCACCUCACCUGCCCCACUCACCACCUCUGCUA
 GUUCCAGACACCUCAC

>Rn6-96h-2pl-B6_F

UCUGCCUGGGCCACAGCCCUCUACCCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCU
 UUCCCGUCCUGGGCACCCCGAGUCUCCCCGACCCCGGGUCCCAGGUAUGCUCCACCUC
 CACCUGCCCCACUCACCACCUCUGCUAGUUCAGACACCUCACGCCACCUGGUCCUCU
 CCAUCGCCCACAAAAGGGGGGGCACGAGGGACGAGCUUAGCUGAGCUGGGAGGAGCAGG
 GUGAGGGUGGGCGACCCAGGAUUCUCCCCACCCC

Grupo G

>Rn5_D5_b

5 UGACACCUCAGCUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUG
GGCAGGCCAAGGCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUC
CCCGCACCCCCGCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCC
CACCUCUACCUCACCCCA

10 >Rn5B2_For2

15 CUCAGCUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGG
CCAAGGCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCA
CCCCCGCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCCACCUC
UACCUCACUCCC

20 >Rn5G3_For2

25 UCUGAGGCCCCGAUCCAGUGGGCAGGCCAAGGCCUGCUGGGCCCCCGCGGACCCAGGUGCUC
CUGGGUACGGUCCCUGUCCCCGCACCCCCGCUUCUGUCUGCCCCAUUGUGGCUCCUCAG
GCUCUCUCCCCUGCUCUCCCACCUCUACCUCACCCCC

30 >Rn6-96hF11_For2

35 GGCCCCGAUCCAGUGGGCAGGCCAAGGCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGG
UCACGGUCCCUGUCCCCGCACCCCCGCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCU
CUCCCCUGCUCUCCCACCUCUACCUCACCCCC

>Rn6-96h-2pl-D8_F

40 GCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCC
GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCCACCUCUACCU
CCGCCCCC

45 >Rn5C4_For2

50 CGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCCGCUUCUGUCUGCCC
CAUUGUGGCUCUUAAGGCUCUCUCCCCUGCUCUCCCACCUCUUAACCUCACCCCCUAC

>Rn6-2WoD3_For2

55 CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGGCCAAG
GCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCC
GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCCACCUCUACCU
60 CCACCCCCAC

>Rn6-96h-2pl-C6_F

65

5 CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGGCCAAG
GCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCC
GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCCACCUCUACCU
CCACCCCCAAC

10 >Rn6-96h-2pl-C7_F

15 CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGGCCAAG
GCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCC
GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCCACCUCUACCU
CCACCCCCAAC

20 >Rn6-96h-2pl-F8_F

25 CGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGGCCAAGGCCUGCU
GGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCCGCUUCUG
UCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCCACCUCUACCUCACCCA
CACC

30 >Rn6-96hH9_For2

35 UCCUGAGGGACUGGGACUCCCCUACAGCCAUGACCUUGACACCUCAGCUGACAGCGUGG
GCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGGCCAAGGCCUGCUGGGCC
CCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCCGCUUCUGUCUGC
CCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCCACCUCUACCUCACGCCAC

40 >Rn5_F10_b

45 CUGAGGGACUGGGACUCCCCUACAGCCAUGACCUUGACACCUCAGCUGACAGCGUGGGC
AACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGGCCAAGGCCUGCUGGGCCCC
CGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCCGCUUCUGUCUGCCC
CAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCCACCUCUACCUCACACCU

50 >Rn6-2WoF11_For2

55 UGCGGAGCCAGCUGGAGGCCAUUUUCCUGAGGGACUGGGACUCCCCUACAGCCAUGACC
UUGACACCUCAGCUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGU
GGGCAGGCCAAGGCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGU
CCCCGCACCCCCGCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUC
60 CCACCUCUACCUCACCC

>Rn6-1WoA9_For2

65

5 CUGGAGGCCAUUUUCCUGAGGGACUGGGACUCCCCUACAGCCAUGACCUUGACACCUCA
 GCUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCGAUCCAGUGGGCAGGCCAA
 10 GGCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUCACGGUCCCUUGUCCCCGCACCCC
 CGCUUCUGUCUGCCCCAUUGUGGCUCUCAGGCUCUCUCCCCUGCUCUCCACCCUACCU
 UCCCCCAC

>Rn6-1WoF9_For2

15 CUGGGACUCCCCUACAGCCAUGACCUUGACACCUAGCUGACAGCGUGGGCAACGCCUG
 CCGCCUGCUCUGAGGGCCAAUCCAGUGGGCAGGCCAAGGCCUGCUGGGCCCCCGCGGACC
 CAGGUGCUCUGGGUCACGGUCCCUUGUCCCCGCACCCCCGCUUCUGUCUGCCCCAUUGUGC
 20 CUCCUUAGGCUCUCUCCCCUGCUCUCCACCUUAACCUCCACCCCC

Grupo I

>Rn5_A7_b

25 GCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGUAAACAGCAGUGAUU
 AACUUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGGUCAAUUU
 30 CGUGCCAGCCACC

>Rn5_B6_b

35 CUUUUCUAUUAGCUCUUAGUAAGAUUACACAUGCAAGCAUCCCCGUUCCAGUGAGUUCACC
 CUCUAAAUCACCACGAUAAAAGGGACAAGCAUCAAGCACGCAGCAAUGCAGCUCAAAAC
 GCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAA
 40 AGUUUAACUAAGCUAUACUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCUACC

>Rn5D4_For2

45 GUUCCAGUGAGUUCACCCUCUAAAUCACCACGAUCAAAAGGGACAAGCAUCAAGCACGCA
 GCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAAC
 CUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGGUCAAUUUCGU
 50 GCCAGCCACC

>Rn5D2_For2

55 AAAGGGACAAGCAUCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCC
 CCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACU
 AACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACC

60 >Rn6-1Wo_D7_b

UCAAAGGGACAAGCAUCAAGCACGCAACAAUGCAGCUCAAAAACGCUUAGCCUAGCCAC
 ACCCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUA
 CACUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACC

>Rn6-96h-2pl-A9_F

UACACAUGCAAGCAUCCCCGUUCCAGUGAGUUCACCCUCUAAAUCACCACGAUCAAAAGG
 GACAAGCAUCAAGCACGCAGCAAUGCAGCUCAAAAACGCUUAGCCUAGCCACACCCCCAC
 GGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACC
 CCAGGGUUGGUCAAUUUCGUGCCAGCCACC

>Rn6-2WoH3_For2

CAUCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAAC
 AGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGU
 UGGUCAAUUUCGUGCCAACCACC

>Rn6-96hG11_For2

AAAGGGACAAGCAUCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCC
 CCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACU
 AACCCCAGGGUUGGUCAAUUUCGUGCCAACCACC

>Rn5E1_For2

CAAGCACGCAACAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
 AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
 UCAAUUUCGUGCCAACCACC

>Rn6-1WoA11_For2

CAUCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCAUGGGAAAC
 AGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGU
 UGGUCAAUUUCGUGCCAGCUCACC

>Rn6-2WoE7_For2

CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
 AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
 UCAAUUUCGUGCCAGCCACACC

>Rn6-96h-2pl-B5_F

CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
AGUGAUUAAACCUUUAGCAAUAAACGAAAGUUUUAACUAAGCUAUACUAACCCCAGGGUUGG
UCAAUUUCGUGCCAGCCACC

>Rn5H2_Fo2

CACGAUCAAAGGGACAAGCAUCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAG
CCACACCCCCACGGGAAACAGCAGUGAUUAAACCUUUAGCAAUAAACGAAAGUUUUAACUAA
GCUAUACUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACC

>Rn6-1WoF11_Fo2

UAAAUCACCACGAUCAAAGGGACAAGCAUCAAGCACGCAGCAAUGCAGCUCAAAACGCU
UAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAAACCUUUAGCAAUAAACGAAAGU
UUAACUAAGCUAUACUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACC

>Rn6-2WoB11_Fo2

AGCCUUCUUAUAGCUCUUAAGUAAGAUUACACAUGCAAGCAUCCCCGUUCCAGUGAGUUC
ACCCUCUAAAUCACCACGAUCAAAGGGACAAGCAUCAAGCACGCAGCAAUGCAGCUCAA
AACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAAACCUUUAGCAAUAAAC
GAAAGUUUAAACUAAGCUAUACUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACC

>Rn6-1WoA3_Fo2

GGGACAAGCAUCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCA
CGGGAAACAGCAGUGAUUAAACCUUUAGCAAUAAACGAAAGUUUAAACUAAGCUAUACUAAC
CCCAGGGUUGGUCAAUUUCGUGCCAGCCACC

>Rn6-1Wo_D2_b

GGGACAAGCAUCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCA
CGGGAAACAGCAGUGAUUAAACCUUUAGCAAUAAACGAAAGUUUAAACUAAGCUAUACUAAC
CCCAGGGUUGGUCAAUUUCGUGCCAGCCACC

Grupo J

>Rn5A1_Fo2

5 UUCUGCCCCAGCUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGA
GAGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCC
CUGUGGCUGCCUCAGCUCAUGCCUUUGGCCUGAAGUCCCAGCAUUGAUGGCAGCCCCUCA
10 UCUCUCCAAGUUUUGUGUCUCCCCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUC
CUUCUGUGCCACU

>Rn5B1_For2

15 UUCUGCCCCAGCUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGA
GAGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCC
CUGUGGCUGCCUCAGCUCAUGCCUUUGGCCUGAAGUCCCAGCAUUGAUGGCAGCCCCUCA
20 UCUCUCCAAGUUUUGUGUCUCCCCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUC
CUUCUGUGCCACU

>Rn5_A10_b

25 CCCCAGCUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGAC
CUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUG
30 GCUGCCUCAGCUCAUGCCUUUGGCCUGAAGUCCCAGCAUUGAUGGCAGCCCCUCAUCUUC
CAAGUUUUGUGUCUCCCCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCU
GUGCCACAAA

>Rn5_G1_b

35 CCCCAGCUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGAC
40 CUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUG
GCUGCCUCAGCUCAUGCCUUUGGCCUGAAGUCCCAGCAUUGAUGGCAGCCCCUCAUCUUC
CAAGUUUUGUGUCUCCCCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCU
45 GUGCCACAAA

>Rn6-1WoF5_For2

50 CUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCU
CCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCC
UCAGCUCAUGCCUUUGGCCUGAAGUCCCAGCAUUGAUGGCAGCCCCUCAUCUUCUCCAAGUU
55 UUGUGUCUCCCCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCC

>Rn6-2WoA5_For2

UGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAGGAGAGACCUUUCUCCG
 5 GACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCCUCA
 GCUCAUGCCUUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUUUUG
 UGCUCCCCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCA

>Rn6-2WoA7_Fo2

GUGAAGAUGACCACAUUCAAGGAAGAACCUUCUGCCCCAGCUUUGCAGGAUGAAACACUU
 15 CCCCCGCUUGGCUCUCCUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGG
 UUCAGCAGCUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUGGCCU
 20 GAAGUCCAGCAUUAUGGCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUACCUA
 UGCUUCCUGCCUCCCAUGCAUCUGUACUCCUGCUGUGCCA

>Rn6-2WoG2_Fo2

UCCACAAGAGAGACCUUUCUCCGGACCUGGCUGCUACUGGUUCAGCAGCUCUGCAGAAAA
 25 UGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCUGAAGUCCAGCAUUAUGGC
 AGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUACCUAUGCUUCCUGCCUCCCAUGCA
 30 UCUGUACUCCUGCUGUGCCACAAACAC

>Rn6-2WoH10_Fo2

UCCACAAGAGAGACCUUUCUCCGGACCUGGCUGCUACUGGUUCAGCAGCUCUGCAGAAAA
 35 UGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCUGAAGUCCAGCAUUAUGGC
 AGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUACCUAUGCUUCCUGCCUCCCAUGCA
 40 UCUGUACUCCUGCUGUGCCACAAACAC

>Rn6-96h-2pl-G7_F

GCUACUGGUUCAGCAGCUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACC
 45 UUUGGCCUGAAGUCCAGCAUUAUGGCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCU
 UUACCUAUGCUUCCUGCCUCCCAUGCAUCUGUACUCCUGCGU

>Rn5-2pl-B2_Fo2

AGAACCUUCUGCCCCAGCUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCC
 55 ACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAGCUCUGCAGAAAAUGU
 CCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCUGAAGUCCAGCAUUAUGGCAGC
 CCCUCAUCUCCAAGUUUUGUGCUCCCCUUUACCUAUGCUUCCUGCCUCCCAUGCAUCU
 60 GUACUCCUG

>Rn5-2pl-D1_Fo2

5 AGAACCUUCUGCCCCAGCUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCC
ACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAGCUCUGCAGAAAAUGU
CCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCUGAAGUCCCAGCAUUAUGGCAGC
CCCUCAUCUCCAAAGUUUUGUGCUCUUUUUACCUAAGCUUCCUGCCUCCCAUGCAUCU
10 GUACUCCUG

>Rn6-1WoA5_For2

15 UGAAGAUGACCACAUUCAAGGAAGAACCUUCUGCCCCAGCUUUGCAGGAUGAAACACUUC
CCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGUUCAGCAG
CUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCUGAAGUCCC
20 AGCAUUAUGGCAGCCCCUCAUCUCCAAAGUUUUGUGCUCUUUUUACCUAAGCUUCCU
GCCUCCCAUGCAUCUGUACUCCUG

>Rn6-1Wo_G10_b

25 UGAAGAUGACCACAUUCAAGGAAGAACCUUCUGCCCCAGCUUUGCAGGAUGAAACACUUC
CCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGGU
30 UCAGCAGCUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCUG
AAGUCCCAGCAUUAUGGCAGCCCCUCAUCUCCAAAGUUUUGUGCUCUUUUUACCUAU
35 GCUUCCUGCCUCCCAUGCAUCUGUACUCCC

>Rn6-2WoE4_For2

40 CCACAUUCAAGGAAGAACCUUCUGCCCCAGCUUUGCAGGAUGAAACACUCCCCGCUUGG
CUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAGCU
CUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCUGAAGUCCCAG
45 CAUUAUGGCAGCCCCUCAUCUCCAAAGUUUUGUGCUCUUUUUACCUAAGCUUCCUGC
CCCC

>Rn6-96hG12_For2

50 GUGAAGAUGACCACAUUCAAGGAAGAACCUUCUGCCCCAGCUUUGCAGGAUGAAACACUUC
CCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGG
55 UUCAGCAGCUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCU
GAAGUCCCAGCAUUAUGGCAGCCCCUCAUCUCCAAAGUUUUGUGCUCUUUUUACCUA
UGCUUCCUGCCCCCAU

>Rn6-96h-2pl-C12

AGAACCUUCUGCCCCAGCUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUUC
 ACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAGCUCUGCAGAAAAUGU
 CCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCUGAAGUCCCAGCAUUAUGGCAGC
 CCCUCAUCUCCAAGUUUUGUGCUCCCC

>Rn6-96h-2pl-A6_F

CUGAAGUGAAGAUGACCACAUUCAAGGAAGAACCUUCUGCCCCAGCUUUGCAGGAUGAAA
 CACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCU
 ACUGGUUCAGCAGCUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUU
 GGCCUGAAGUCCCAGCAUUAUGGCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUA
 CCUAAUGCUUCCUGCCUCCCAUGCAUCUGUACUCCU

>Rn6-96h-2pl-H5_F

CUGAAGUGAAGAUGACCACAUUCAAGGAAGAACCUUCUGCCCCAGCUUUGCAGGAUGAAA
 CACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCU
 ACUGGUUCAGCAGCUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUU
 GGCCUGAAGUCCCAGCAUUAUGGCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUA
 CCUAAUGCUUCCUGCCUCCCAUGCAUCUGUACUCCU

>Rn6-2WoG1_For2

AAGAUGACCACAUUCAAGGAAGAACCUUCUGCCCCAGCUUUGCAGGAUGAAACACUCCC
 CGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGGUUC
 AGCAGCUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCUGAA
 GUCCCAGCAUUAUGGCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUACCUA AUGC

UCCUGCCUCCCAUGCAUCUGUACUCCUGC

>Rn6-96h-2pl-D11

CCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGG
 UUCAGCAGCUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCU
 GAAGUCCCAGCAUUAUGGCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUACCUA
 UGCUUCCUGCCUCCCAUGCAUCUGUACUCCU

>Rn6-96h-2pl-F9_F

CCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGG
 UUCAGCAGCUCUGCAGAAAAUGUCCUCCCUUGUGGCUGCCUCAGCUCGUACCUUUGGCCU
 GAAGUCCCAGCAUUAUGGCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUACCUA
 UGCUUCCUGCCUCCCAUGCAUCUGUACUCCU

>hBg:

5 GAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCCCUAAGUCCAA
CUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAA
CAUUUAUUUUAUUGCUGCGUC

noUTR:
>

>BB

15 UGCCCUGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
20 GCUGAUCCAUUGCCGGUGUGACCUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUG
CUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCAC
UGGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUUGCCGGUGUGACC

>BD

25 UGCCCUGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
30 GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
GCUGAUCCAUUGCCGGUGUGACCUCUCCAGCCAGACACCCGCCCCCGGGCCUGGCUAAGA
AGUUGCUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUGAUGCCAUCGCGUGCCACC
35 UCCUUGUCCUGGACCCUUAAGCCUCUCUGCCCUUCCACUCUCUGACCCC

>BE

40 UGCCCUGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
GCUGAUCCAUUGCCGGUGUGACCUGCCUUGGCUCUCCAGGAAGGCUCAGGAGCCCUACCU
45 CCCUGCCAUAUAGCUGCUCCCCGCCAGAAGCCUGUGCCAACUCUCUGCAUUCCUGAUC
UCCAUCCUGUGGCUGUCACCCUUGGUCACCUCGUGCUGUCACUGCCAUCUCCCCCC

>BF

50 UGCCCUGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
55 GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
GCUGAUCCAUUGCCGGUGUGACCCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCUUUC
CCGUCCUGGGUACCCCGAGUCUCCCCGACCUCGGGUCCAGGUAUGCUCCACCUCAC
60 CUGCCCCACUCACCACCUCUGCUAGUCCAGACACCUC

>BG

65

5 UGCCCUGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
GCUGAUCCAUGCCGGUGUGACCCUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGG
CCCGAUCCAGUGGGCAGGCCAAGGCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUC
10 ACGGUCCCUGUCCCCGCACCCCCGCUUCUGUCUGCCCCAUUGUGGCUCUCAGGCUCUCU
CCCCUGCUCUCCCAACCUCUACCUCACCCCCAC

>BhBg

15 UGCCCUGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
GCUGAUCCAUGCCGGUGUGACCGAGAGCUCGCUUUCUUGCUGUCCAAUUUCUAUUAAAG
20 GUUCCUUUGUCCCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGGGCCUUGAGCA
UCUGGAUUCUGCCUAAUAAAAACAUAUUUUUCAUUGCUGCGUC

>BI

25 UGCCCUGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
GCUGAUCCAUGCCGGUGUGACCCAAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUA
30 GCCACACCCCCACGGGAAACAGCAGUGAUUAACCUUAGCAAUAAACGAAAGUUUAACUA
AGCUAUACUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

>BJ

40 UGCCCUGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACU
GACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCU
GCUGAUCCAUGCCGGUGUGACCCUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAU
45 UCUUCCACAAGAGAGACCUCUCCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGA
AAAUGUCCUCCCCUGUGGCUGCCUCAGCUCUUGCCUUUGGCCUGAAGUCCAGCAUUGAU
GGCAGCCCCUCAUCUCCAAGUUUUGUGCUCUUUUUACCUAACGCUUCCUGCCUCCCAU
50 GCAUCUGUACUCCUCC

>DB

55 UUCCAGCCAGACACCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUG
ACCUACCCUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCUGGACCCUUUA
GCCUCUCUGCCCUCUCCACUCUCUGACCCCUGCCGUCUCCACCAAGACUGACUGCCUGCU
60 GCUUUGCUACUGCCCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCA
CUGCACUGGCACAGCCCCGCCUUGCCGUGCUGAUCCAUGCCGGUGUGACC

65

>DD

5 U U C C A G C C A G A C A C C C G C C C C C G G C C C U G G C U A A G A A G U U G C U U C C U G U U G C C A G C A U G
A C C U A C C C U C G C C C U C U U G A U G C C A U C C G C U G C C A C C U C C U U U U G C U C C U G G A C C C U U A
G C C U C U C U G C C C U U C C A C U C U C U G A C C C C U U C C A G C C A G A C A C C C G C C C C C G G C C C U G G
10 C U A A G A A G U U G C U U C C U G U U G C C A G C A U G A C C U A C C C U C G C C C U C U U G A U G C C A U C C G C U
G C C A C C U C C U U U G C U C C U G G A C C C U U A G C C U C U C U G C C C U U C C A C U C U C U G A C C C C

>DE

15 U U C C A G C C A G A C A C C C G C C C C C G G C C C U G G C U A A G A A G U U G C U U C C U G U U G C C A G C A U G
A C C U A C C C U C G C C C U C U U G A U G C C A U C C G C U G C C A C C U C C U U U U G C U C C U G G A C C C U U A
20 G C C U C U C U G C C C U U C C A C U C U C U G A C C C C G C C U U G G C U C C U C C A G G A A G G C U C A G G A G C C
C U A C C U C C C U G C C A U U A U A G C U G C U C C C C G C C A G A A G C C U G U G C C A A C U C U C U G C A U U C C
C U G A U C U C C A U C C C U G U G G C U G U C A C C C U U G G U C A C C U C C G U G C U G U C A C U G C C A U C U C C
25 C C C C

>DF

30 U U C C A G C C A G A C A C C C G C C C C C G G C C C U G G C U A A G A A G U U G C U U C C U G U U G C C A G C A U G
A C C U A C C C U C G C C C U C U U G A U G C C A U C C G C U G C C A C C U C C U U U U G C U C C U G G A C C C U U A
G C C U C U C U G C C C U U C C A C U C U C U G A C C C C U G G U A C U G C A U G C A C G C A A U G C U A G C U G C C
35 C C U U C C C G U C C U G G G U A C C C C G A G U C U C C C C G A C C U C G G G U C C C A G G U A U G C U C C C A C
C U C C A C C U G C C C C A C U C A C C A C C U C U G C U A G U U C C A G A C A C C U C C

>DG

40 U U C C A G C C A G A C A C C C G C C C C C G G C C C U G G C U A A G A A G U U G C U U C C U G U U G C C A G C A U G
A C C U A C C C U C G C C C U C U U G A U G C C A U C C G C U G C C A C C U C C U U U U G C U C C U G G A C C C U U A
45 G C C U C U C U G C C C U U C C A C U C U C U G A C C C C U G A C A G C G U G G G C A A C G C C U G C C G C C U G C U
C U G A G G C C C G A U C C A G U G G G C A G G C C A A G G C C U G C U G G G C C C C G C G G A C C A G G U G C U C
U G G G U C A C G G U C C C U G U C C C C G A C C C C G C U U C U G U C U G C C C C A U U G U G G C U C C U C A G G
50 C U C U C U C C C C U G C U C U C C A C C U C U A C C U C C A C C C C C A C

>DhBg

55 U U C C A G C C A G A C A C C C G C C C C C G G C C C U G G C U A A G A A G U U G C U U C C U G U U G C C A G C A U G
A C C U A C C C U C G C C C U C U U G A U G C C A U C C G C U G C C A C C U C C U U U U G C U C C U G G A C C C U U A
G C C U C U C U G C C C U U C C A C U C U C U G A C C C C G A G A G C U C G C U U U C U G C U G U C C A A U U U C U A
U U A A A G G U U C C U U G U U C C C U A A G U C C A A C U A C U A A A C U G G G G G A U A U U A U G A A G G G C C U
60 U G A G C A U C U G G A U U C U G C C U A A U A A A A A C A U U U A U U U C A U U G C U G C G U C

>DI

65

UUC CAGCCAGACACCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUG
 ACCUACCCUCGCCUCUUGAUGCCAUCGCGUGCCACCUCUUGUUGCUUCCUGGACCCUUA

GCCUCUCUGCCCUUCCACUCUCUGACCCCCAAGCACGCGAGCAAUGCAGCUCAAAACGCUU
 AGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAACCUUAGCAAUAAACGAAAGUU
 UAACUAAGCUAUACUAACCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

>DJ

UUC CAGCCAGACACCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUG
 ACCUACCCUCGCCUCUUGAUGCCAUCGCGUGCCACCUCUUGUUGCUUCCUGGACCCUUA
 GCCUCUCUGCCCUUCCACUCUCUGACCCCCUUGCAGGAUGAAACACUCCCCGCUUGGC
 UCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAACUC
 UGCAGAAAAUGUCCUCCCCUGUGGCUGCCUCAGCUC AUGCCUUUGGCCUGAAGUCCAGC
 AUUGAUGGCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUACCUAACGCUUCCUGCC
 UCCCAUGCAUCUGUACUCCUCC

>EB

GCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUUAUAGCUGCUCCCC
 GCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCCUGUGGCUGUCACCCU
 UGGUCACCUCGUGCUGUCACUGCCAUCUCCCCUUGCCGUCUCCACCAAGACUGACUG
 CCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUCCACUGCUCUGCCUGCCUC
 UCCCCACUGCACUGGCACAGCCCGCCUUGCCGUGCUGAUCCAUUGCCGGUGUGACC

>ED

GCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUUAUAGCUGCUCCCC
 GCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCCUGUGGCUGUCACCCU
 UGGUCACCUCGUGCUGUCACUGCCAUCUCCCCCUUCCAGCCAGACACCCGCCCCCGG
 CCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUGAUGCCA
 UCCGCGUGCCACCUCUUGUUGCUUCCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCUCUGA
 CCCC

>EE

GCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUUAUAGCUGCUCCCC
 GCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCCUGUGGCUGUCACCCU
 UGGUCACCUCGUGCUGUCACUGCCAUCUCCCCCGCCUUGGCUCCUCCAGGAAGGCUCA
 GGAGCCCUACCUCUCCUGCCAUUAUAGCUGCUCCCCGCCAGAAGCCUGUGCCAACUCUCUG
 CAUUCUCCUGAUCUCCAUCCUGUGGCUGUCACCCUUGGUCACCUCGUGCUGUCACUGCC
 AUCUCCCCC

>EF

5 GCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUAUAGCUGCUCCCC
GCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCCUGUGGCUGUCACCCU
10 UGGUACACCUCCGUGCUGUCACUGCCAUCUCCCCCCCCUGGUACUGCAUGCACGCAAUGCUA
GCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUCCCCCGACCUCGGGUCCCAGGUAUGC
UCCCAACUCCACCUGCCCCACUCACCACCUCUGCUAGUCCAGACACCUCC

>EG

15 >GCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUAUAGCUGCUCCCC
CGCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCCUGUGGCUGUCACCC
20 UUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCCCCCUGACAGCGUGGGCAACGCCUGCC
GCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGGCCAAGGCCUGCUGGGCCCCCGCGGACCCA
GGUGCUCUGGGUCACGGUCCUGUCCCCGACCCCCGCUUCUGUCUGCCCCAUUGUGGGCU
25 CCUCAGGCUCUCUCCCCUGCUCUCCCAACUCUACCUCACCCCCAC

>EhBg

30 GCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUAUAGCUGCUCCCC
GCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCCUGUGGCUGUCACCCU
UGGUACACCUCCGUGCUGUCACUGCCAUCUCCCCCCCCGAGAGCUCGCUUUCUUGCUGUCCAA
35 UUUCUAUUAAGGUUCCUUUGUUCUUAAGUCCAACUACUAAACUGGGGGAUUAUUAUGAA
GGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAACAUAUUAUUUUAUUGCUGCGUC

>EI

40 GCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUAUAGCUGCUCCCC
GCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCCUGUGGCUGUCACCCU
45 UGGUACACCUCCGUGCUGUCACUGCCAUCUCCCCCCCCAAGCACGCAGCAAUGCAGCUCAAA
ACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACG
AAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGGUCAAUUUCUGUGCCAGCCACACC

>EJ

50 GCCUUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUAUAGCUGCUCCCC
55 GCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCUCCAUCCUGUGGCUGUCACCCU
UGGUACACCUCCGUGCUGUCACUGCCAUCUCCCCCCCCUUUGCAGGAUGAAACAUUCCCCG
CUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAG
60 CAACUCUGCAGAAAUGUCCUCCCCUGUGGCUGCCUCAGCUCUAGCCUUUGGCCUGAAGU
CCCAGCAUUGAUGGCAGCCCCUCAUCUCCCAAGUUUUGUGCUCUCCCUUUACCUAACGCUU
CCUGCCUCCCAUGCAUCUGUACUCCUCC

>FB

5 CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUGCUCCCACCUCACCUGCCCCACUCACCACCUCUGCU
 AGUUCCAGACACCUCUCCUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGC
 10 CCGGGCCCAUGAGACUGACUUCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACA
 GCCCCGCCUUGCCGUGCUGAUCCAUGCCGGUGUGACC

>FD

15 CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUGCUCCCACCUCACCUGCCCCACUCACCACCUCUGCU
 20 AGUUCCAGACACCUCUCCAGCCAGACACCCGCCCCCGGCCUAGGAAGUUGCU
 UCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUGAUGCCAUCGCGUGCCACCUCUUUU
 GCUCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCUCUGACCCC

>FE

25 CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUGCUCCCACCUCACCUGCCCCACUCACCACCUCUGCU
 30 AGUUCCAGACACCUCUCCUGGCUCCUCCAGGAAGGCUCAGGAGCCCUACCUCCUUGCC
 AUUAUAGCUGCUCCCCGCCAGAAGCCUGUGCCAACUCUCUGCAUCCUGAUCUCCAUC
 35 CUGUGGCUGUCACCCUUGGUCACCUCGUGCUGUCACUGCCAUCUCCCCC

>FF

40 CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUGCUCCCACCUCACCUGCCCCACUCACCACCUCUGCU
 AGUUCCAGACACCUCUCCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCU
 45 GGGUACCCCGAGUCUCCCCGACCUCGGGUCCCAGGUAUGCUCCCACCUCACCUGCCCC
 ACUCACCACCUCUGCUAGUCCAGACACCUCC

>FG

50 CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUGCUCCCACCUCACCUGCCCCACUCACCACCUCUGCU
 55 AGUUCCAGACACCUCUCCUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUC
 CAGUGGGCAGGCCAAGGCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCC
 CUGUCCCCGCACCCCCGCUUCUGUCUGCCCCAUUGUGGCUCUCCAGGCUCUCUCCCCUGC
 60 UCUCCACCUCUACCUCCACCCCCAC

>FhBg

65

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUCCCCGUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUGCUCACCACCUGCCCCACUCACCACCUCUGCU
 AGUUCAGACACCUCGAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCU
 UGUUCCCUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAU
 UCUGCCUAAUAAAAACAUTUAUUUUAUUGCUGCGUC

>FI

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUCCCCGUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUGCUCACCACCUGCCCCACUCACCACCUCUGCU
 AGUUCAGACACCUCCTCAAGCAGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACAC
 CCCCACGGGAAACAGCAGUGAUUAACCUUAGCAAUAAACGAAAGUUAACUAAGCUAUA
 CUAACCCACAGGGUUGGUCAAUUCGUGCCAGCCACACC

>FJ

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUCCCCGUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUGCUCACCACCUGCCCCACUCACCACCUCUGCU
 AGUUCAGACACCUCCTUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCA
 CAAGAGAGACCUCUUCUCCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAUGUC
 CUCCCCUGUGGCUGCCUCAGCUCUAGCCUUGGCCUGAAGUCCAGCAUUGAUGGCAGCC
 CCUCAUCUCCAAGUUUUGUGCUCUCCCUUUAACCUAACGCUUCCUGCCUCCCAUGCAUCUG
 UACUCCUCC

>GB

CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCGAUCCAGUGGGCAGGCCAAG
 GCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUCACGGUCCUGUCCCCGCACCCCC
 GCUUCUGUCUGCCCCAUUGUGGCUCUCAGGCUCUCUCCCCUGCUCUCCACCUCUACCU
 CCACCCCCACUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGC
 CCAUGAGACUGACUCCCCACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCG
 CCUUGCCGCUGCUGAUCCAUGCCGGUGUGACC

>GD

CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCGAUCCAGUGGGCAGGCCAAG
 GCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUCACGGUCCUGUCCCCGCACCCCC
 GCUUCUGUCUGCCCCAUUGUGGCUCUCAGGCUCUCUCCCCUGCUCUCCACCUCUACCU
 CCACCCCCACUUCAGCCAGACACCCGCCCCCGGCCUUGGCUAAGAAGUUGCUUCCUGU
 UGCCAGCAUGACCUCUCCUCGCCUCUUGAUGCCAUCCGCUGCCACCUCUUCUUGCUCU
 GGACCCUUAAGCCUCUCUGCCCUCCACUCUCUGACCCC

>GE

5 CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCGAUCCAGUGGGCAGGCCAAG
GCCUGCUGGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCC
10 GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCACCUCUACCU
CCACCCCCACGCCUUGGCUCUCCAGGAAGGCUCAGGAGCCCUACCUCCCUGCCAUAUA
GUGCUCUCCCCGCCAGAAGCCUGUGCCAACUCUCUGCAUCCCUGAUCUCCAUCCCUGUGG
15 CUGUCACCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCC

>GF

20 CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCGAUCCAGUGGGCAGGCCAAG
GCCUGCUGGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCC
GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCACCUCUACCU
25 CCACCCCCACCUGGUACUGCAUGCAGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUAC
CCCGAGUCUCCCCCGACCUCCGGGUGCCAGGUAUGCUCCACCUCACCUGCCCCACUCAC
CACCUCUGCUAGUUCAGACACCUC

>GG

30 CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCGAUCCAGUGGGCAGGCCAAG
GCCUGCUGGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCC
35 GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCACCUCUACCU
CCACCCCCACCUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCGAUCCAGUGG
GCAGGCCAAGGCCUGCUGGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCC
40 CCGCACCCCCGCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCC
ACCUCUACCUCACCCCCAC

>GhBg

45 CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCGAUCCAGUGGGCAGGCCAAG
GCCUGCUGGGGCCCCCGCGGACCCAGGUGCUCUGGGUACGGUCCCUGUCCCCGCACCCCC
50 GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCACCUCUACCU
CCACCCCCACGAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUC
55 CUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCC
UAAUAAAAACAUAUUAUUUCAUUGCUGCGUC

>GI

60

5 CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGGCCAAG
GCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUCACGGUCCUGUCCCCGCACCCCC
GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCACCUCUACCU
CCACCCCCACCAAGCACGCAGCAUUGCAGCUCAAAAACGCUUAGCCUAGCCACACCCCCAC
10 GGGAAACAGCAGUGAUUAACCUUAGCAAUAAACGAAAGUUUAAACUAAGCUAUACUAACC
CCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

>GJ

15 CUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCCGAUCCAGUGGGCAGGCCAAG
GCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUCACGGUCCUGUCCCCGCACCCCC
GCUUCUGUCUGCCCCAUUGUGGCUCCUCAGGCUCUCUCCCCUGCUCUCCACCUCUACCU
20 CCACCCCCACCUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGAG
AGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCC
UGUGGCUGCCUCAGCUCUAGCCUUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAU
25 CUUCCAAGUUUUGUGCUCUCCCUUUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCC
UCC

>hBgB

30 GAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCCCUAAGUCCAA
CUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAA
35 CAUUUAUUUUAUUGCUGCGUCUGCCCGUCCUCACCAAGACUGACUGCCUGCUGCUUUGC
UACUGCCCGGGCCCAUGAGACUGACUUCUCCACUGCUCUGCCUGCCUCUCCCCACUGCACU
GGCACAGCCCCGCCUUGCCGCUGCUGAUCCAUTGCCGGUGUGACC

>hBgD

40 GAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCCCUAAGUCCAA
CUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAA
45 CAUUUAUUUUAUUGCUGCGUCUCCAGCCAGACACCCGCCCCCGGCCUGGCUAAGAA
GUUGCUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUGAUGCCAUCGCGUGCCACCU
50 CCUUUUGCUCUCCUGGACCCUUAGCCUCUCUGCCCUUCCACUCUCUGACCCC

>hBgE

55 GAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCCCUAAGUCCAA
CUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAA
CAUUUAUUUUAUUGCUGCGUCGCCUUGGCUCUCCAGGAAGGCUCAGGAGCCCUACCUC
60 CCUGCCAUUAUAGCUGCUCUCCCGCCAGAAGCCUGUGCCAACUCUCUGCAUCCUGAUCU
CCAUCCUGUGGCUGUCACCCUUGGUACCUCCGUGCUGUCACUGCCAUCUCCCCC

>hBgF

5 GAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCCCUAAGUCCAA
CUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAA
CAUUUAUUUUCAUUGCUGCGUCCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUCC
10 CGUCCUGGGUACCCCGAGUCUCCCCGACCUCGGGUCCCAGGUAUGCUCACCACCUCACC
UGCCCCACUCACCACCUCUGCUAGUUCAGACACCUC

>hBgG

15 GAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCCCUAAGUCCAA
CUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAA
20 CAUUUAUUUUCAUUGCUGCGUCCUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGC
CCGAUCCAGUGGGCAGGCCAAGGCCUGCUGGGCCCCCGCGGACCCAGGUGCUCUGGGUCA
CGGUCCCUGUCCCCGACCCCGCUUCUGUCUGCCCCAUUGUGGCUCUCAGGCUCUCUC
25 CCCUGCUCUCCACCUCUACCUCACCCCCAC

>hBghBg

30 GAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCCCUAAGUCCAA
CUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAA
CAUUUAUUUUCAUUGCUGCGUCGAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGG
35 UCCUUGUUCCCUAAGUCCAAUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAU
CUGGAUUCUGCCUAAUAAAAAACAUUUUUUUCAUUGCUGCGUC

>hBgl

40 GAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCCCUAAGUCCAA
CUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAA
CAUUUAUUUUCAUUGCUGCGUCCAGCAGCAGCAAUGCAGCUAAAACGCUUAGCCUAG
45 CCACACCCCCACGGGAAACAGCAGUGAUUAACCUUAGCAAUAAACGAAAGUUUAACUAA
GCUAUACUAACCCAGGGUUGGUCAAUUCGUGCCAGCCACACC

>hBgJ

50 GAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUUCCCUAAGUCCAA
CUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAA
55 CAUUUAUUUUCAUUGCUGCGUCCUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAU
CUUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAA
AAUGUCCUCCCCUGUGGCUGCCUCAGCUCAUGCCUUGGCCUGAAGUCCAGCAUUGAUG
60 GCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUAACCUAACGCUUCCUGCCUCCCAUG
CAUCUGUACUCCUC

>IB

5 CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
UCAAUUUCGUGCCAGCCACACCUGCCGUCUACCAAGACUGACUGCCUGCUGCUUUGC
10 UACUGCCCCGGGCCCAUGAGACUGACUCCCCACUGCUCUGCCUGCCUCUCCCCACUGCACU
GGCACAGCCCCGCTUUGCCGCGUGCUGAUCCAUUGCCGGUGUGACC

>ID

15 CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
20 UCAAUUUCGUGCCAGCCACACCUUCCAGCCAGACACCCGCCCCCGGCCUGGCUAAGAA
GUUGCUUCCUGUUGCCAGCAUGACCUACCCUCGCCUCUUGAUGCCAUCGCGUGCCACCU
CCUUUUGCUCCUGGACCCUUUAGCCUCUCUGCCCUUCCACUCUCUGACCCC

>IE

25 CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
30 AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
UCAAUUUCGUGCCAGCCACACCGCCUUGGCUCUCCAGGAAGGCUCAGGAGCCCUACCTUC
CCUGCCAUUAUAGCUGCUCUCCCGCCAGAAGCCUGUGCCAACUCUCUGCAUUCUCCUGAUCU
35 CCAUCCUGUGGCUGUCACCCUUGGUCACCUCCGUGCUGUCACUGCCAUCUCCCCC

>IF

40 CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
UCAAUUUCGUGCCAGCCACACCCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUC
45 CGUCCUGGGUACCCCGAGUCUCCCCGACCUCGGGUCCAGGUAUGCUCCACCUCACC
UGCCCCACUCACCACCUCUGCUAGUUCAGACACCUCC

>IG

50 CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
55 UCAAUUUCGUGCCAGCCACACCCUGACAGCGUGGGCAACGCCUGCCGCCUGCUCUGAGGC
CCGAUCCAGUGGGCAGGCCAAGGCCUGCUGGGCCCCCGCGACCCAGGUGCUCUGGGUCA
CGGUCCCCUGUCCCCGACCCCCGCUUCUGUCUGCCCCAUUGUGGCUCUCCAGGCUCUCUC
60 CCCUGCUCUCCACCUCUACCUCCACCCCCAC

>lhBg

65

5 CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
UCAAUUUCGUGCCAGCCACACCGAGAGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGG
UUCUUUGUCCCCUAAGUCCAACUACUAAACUGGGGGGAUUAUGAAGGGCCUUGAGCAU
10 CUGGAUUCUGCCUAAUAAAAACAUUUAUUUCAUUGCUGCGUC

>II

15 CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
UCAAUUUCGUGCCAGCCACACCCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAG
20 CCACACCCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAA
GCUAUACUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

>IJ

25 CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC
AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGG
30 UCAAUUUCGUGCCAGCCACACCCUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAU
CUUCCACAAGAGAGACCUUUCUCCGGACCUGGUUGCACUGGUUCAGCAACUCUGCAGAA
AAUGUCCUCCCCUGUGGCUGCCUCAGCUCUAGCCUUUGGCCUGAAGUCCAGCAUUGAUG
35 GCAGCCCCUCAUCUCCAAGUUUUGUGCUCCCCUUUACCUAACGCUUCCUGCCUCCCAUG
CAUCUGUACUCCUCC

>JB

40 CUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCU
CCGGACCUGGUUGCACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCC
45 UCAGCUCAUGCCUUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUU
UUGUGCUCCCCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCUGCCCGU
50 CCUCACCAAGACUGACUGCCUGCUGCUUUGCUACUGCCCGGGCCCAUGAGACUGACUCC
CACUGCUCUGCCUGCCUCUCCCCACUGCACUGGCACAGCCCCGCCUUGCCGCUGCUGAUC
CAUUGCCGGUGUGACC

>JD

CUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUUCCACAAGAGAGACCUUUCU
 5 CCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCC
 UCAGCUC AUGCCUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUU
 UUGUGCUCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCUCCAGC
 10 CAGACACCCGCCCCCGGCCUGGCUAAGAAGUUGCUUCCUGUUGCCAGCAUGACCUACC
 CUCGCCUCUUUGAUGCCAUCGCGUGCCACCUCUUUUGCUCCUGGACCCUUUAGCCUCUC
 UGCCC UCCACUCUCUGACCCC

15 >JE

CUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUUCCACAAGAGAGACCUUUCU
 20 CCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCC
 UCAGCUC AUGCCUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUU
 UUGUGCUCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCGCCUUGG
 25 CUCCUCCAGGAAGGCUCAGGAGCCCUACCUCUCCUGCCAUAUAGCUGCUCCCGCCAGAA
 GCCUGUGCCAACUCUCUGCAUCCUUGAUCUCCAUCCUGUGGCUGUCACCCUUGGUCAC
 CUCCGUGCUGUCACUGCCAUCUCCCCC

30 >JF

CUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUUCCACAAGAGAGACCUUUCU
 35 CCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCC
 UCAGCUC AUGCCUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUU
 UUGUGCUCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCUGGUAC
 40 UGCAUGCACGCAAUGCUAGCUGCCCUUUCUCCUGUCCUGGGUACCCCGAGUCUCCCGAC
 CUCGGGUCCAGGUAUGCUCCACCUCUCCACCUGCCCCACUACCACCUCUGCUAGUCCA
 GACACCUCC

45 >JG

CUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUUCCACAAGAGAGACCUUUCU
 50 CCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCC
 UCAGCUC AUGCCUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUU
 UUGUGCUCUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCUGACAG
 55 CGUGGGCAACGCCUGCCGCCUGCUCUGAGGCCCGAUCCAGUGGGCAGGCCAAGGCCUGCU
 GGGCCCCCGCGACCCAGGUGCUCUGGGUACGGUCCUGUCCCGCACCCCCGCUUCUG
 UCUGCCCCAUUGUGGCUCUCUAGGCUCUCUCCUCCUGCUCUCCACCUCUACCUCACCCC
 CAC

60 >JhBg

65

CUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCU
 5 CCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCC
 UCAGCUC AUGCCUUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUU
 UUGUGCUCUUUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCGAGAGCU
 10 CGCUUUCUUGCUGUCCAAUUUCUAUUAAAAGGUUCCUUUGUCCCUAAGUCCAACUACUAA
 ACUGGGGGGAUUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAACAUUUAU
 UUUCAUUGCUGCGUC

15 >JI

CUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCU
 20 CCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCC
 UCAGCUC AUGCCUUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUU
 UUGUGCUCUUUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCCAAGCAC
 25 GCAGCAAUGCAGCUCAAAAAGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUU
 AACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCAGGGUUGGUCAAUUU
 CGUGCCAGCCACACC

30 >JJ

CUUUGCAGGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCU
 35 CCGGACCUGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCC
 UCAGCUC AUGCCUUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUU
 UUGUGCUCUUUUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCCUUUGCA
 40 GGAUGAAACACUCCCCGCUUGGCUCUCAUUCUCCACAAGAGAGACCUUUCUCCGGACC
 UGGUUGCUACUGGUUCAGCAACUCUGCAGAAAAUGUCCUCCCCUGUGGCUGCCUCAGCUC
 45 AUGCCUUUGGCCUGAAGUCCAGCAUUGAUGGCAGCCCCUCAUCUCCAAGUUUUGUGCU
 CCCC UUUACCUAACGCUUCCUGCCUCCCAUGCAUCUGUACUCCUCC

> FI UTR 97,5 % de homología (modificaciones aleatorias)

50 CUGGUACUGCAUGGACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCAC
 CCCC GACCUCGGGUCCAGGUUUCGUCCACCUCACCUGCCCCACUACCACCUCUGCU
 AGUUC CAGACACCUC CCAAGCACGCAGCAAUGCAGTUCAAAACGCUUAGCAUAGCCACAC
 55 CCCCACGGGAAACAGCAGUGAUUAACCUUAGCAAUAUUCGAAUGUUUAACUAAGCUAUA
 CUAACCC CAGGGUUGGUCAAUUUCGUGCCAGCCACACC

60 > FI UTR 95 % de homología (modificaciones aleatorias)

CUCGUACUGCAUGGACGCAAUGCUAGCUGCCCCUUCUCCGUGCCUGGGUACCCCGAGUCAC
 CACCGACCUCGGGUCCCAGGUAUCGUCCACCUCACGUGCCCCACUCACCACCUCUGCU
 AGUUCAGACACCUCUCCCAAGCACGCAGCAAUGCAGCUAAAACGCUUAGCAUAGCCACAC
 CCCCACGGGAAACAGUAGUGAUUAACCUUUAGCAAUAAUCGAAUGUCUAACUAAGCUAUA
 CUAACCCACAGGGUUGAUCAAUACGUGCCAGCCACACC

> FI UTR 92,5 % de homología (modificaciones aleatorias)

CUCGUACUGCAUGGACGCAAUGCUAGCUGCCCCUUCUCCGUGCCUGGGUACCCCGAGUCAC
 CACCGACCUCGGGUCCCAGGUAUCGUCCACCUCACGUGCCCCACUCACCACCUCUUGCU
 AGUUCAGACACCUCUCCCAUGCACGCAGCAAUGCAGAUCAAACGCUUAGCAUAGCCACAC
 CCCCACGGGAAACAGUAGUGAUUAACCUUUAGCAAUAAUCGAAUGUCUAACUAAGCUAUA
 CUAACCCACAGGGUUGAUCAAUACGUGCCAGCCACACC

> FI UTR 90 % de homología (modificaciones aleatorias)

GUCGUACUGCAUGGACGCAAUGCUAGCAGCACCUCUUCUCCGUGCCUGGGUACCCCGAGUCAC
 CACCGACCUCGGGUCCCAGGUAUCGUCCACCUCACGUGCCCCACCCACCACCUCUUGCU
 AGUUCAGAGACCUCUCCCAUGCACGCAGCAAUGCAGAUCAAACGCUUAGCAUAGCCACAC
 CGCCACGGGAAACAGUAGUGAUCAAACCUUUAGCUAUAUAAUCGAAUGUCUAACUAAGCUAUA
 CUAACCACAGGGUUGAUCAAUACGUGCCAGCCAGACC

> Mutación FI 8nt

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCAAAGGGCUCUCCUGGGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUCGUCCACCUCACGUGCCCCACUCACCACCUCUGCU
 AGUUCAGACACCUCUCCCAAGCACGCAGCAAUGCAGCUAAAACGCUUAGCCUAGCCACAC
 CCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUA
 CUAACCCACAGGGUUGGUCAAUUUCGUGCCAGCCACACC

> FI UTR 98,75 % de homología (modificaciones que desestabilizan la estructura)

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUCUCCGUGGUCCGUACCCCGAGUCUC
 CCCCAGACCUCGGGUCCCAGGUAUCGUCCACCUCACGUGCCCCACUCACCACCUCUGCU
 AGUUCAGACACCUCUCCCAAGCACGCAGCAAUGCAGCUAAAACGCUUAGCCUAGCCACAC
 CCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUA
 CUAACCCACAGGGUUGGUCAAUUUCGUGCCAGCCACACC

> FI UTR 97,5 % de homología (modificaciones que desestabilizan la estructura)

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUGGACCGUACGGCGAGUCUC
 CCCCAGCCUCGGGUCCCAGGUAUGCUCACCUCACCUGCCCCACUCACCACCUCUGCU
 AGUUCCAGACACCUCCTCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACAC
 CCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUA
 CUAACCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

> FI UTR 96,25 % de homología (modificaciones que desestabilizan la estructura)

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUGCCGUGGACCGUACGGGCUGUCUC
 CCCCAGCCUCGGGUCCCAGGUAUGCUCACCUCACCUGCCCCACUCACCACCUCUGCU
 AGUUCCAGACACCUCCTCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACAC
 CCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUA
 CUAACCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

> FI UTR 95 % de homología (modificaciones que desestabilizan la estructura)

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUGGGCUGGACCGUACGGGCUGUCUC
 CCCCAGCCUCGGGUCCCAGGUAUGCUCACCUCACCUGCCCCACUCACCACCUCUGCU
 AGUUCCAGACACCUCCTCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACAC
 CCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUA
 CUAACCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

> FI UTR 97,5 % de homología (modificaciones que desestabilizan la estructura)

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUGGUCCGUACCCCGAGUCUC
 CCCCAGCCUCGGGUCCGACCUAUGCUCACCUCACCUGCCCCACUCACCACCUCUGCU
 AGUUCCAGACACCUCCTCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACAC
 CCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUA
 CUAACCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

> FI UTR 95 % (modificaciones que retienen la estructura)

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUGGACCGUACGGCGAGUCUC
 CCCCAGCCUCGCCUCGGUCCUAUGCUCACCUCACCUGCCCCACUCACCACCUCUGCU
 AGUUCCAGACACCUCCTCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACAC
 CCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUA
 CUAACCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

> FI UTR 92,5 % (modificaciones que retienen la estructura)

5 CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUGCCGUGGACCGUACGGGCUGUCUC
 CCCCAGACCAGCCCUCGGUCCUAUGCUCUCCACCUCUCCACCUGCCCCACUCACCACCUCUGCU
 AGUUCCAGACACCUCUCCCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACAC
 CCCCACGGCAAACAGCAGUGAUUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUA
 10 CUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

> FI UTR 90 % (modificaciones que retienen la estructura)

15 CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUGGGCUGGACCGUACGGGCUGUCUC
 CCCCAGACCAGCCCUCGGUCCUAUGCUCUCCACCUCUCCACCUGCCCCACUCACCACCUCUGCU
 AGUUCCAGACACCUCUCCCAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACAC
 20 CCCCAGCCCCAAACAGCAGUGAUUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUA
 CUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

25

Listado de secuencias

<110> BioNTech RNA Pharmaceuticals GmbH y otros

30

<120> SECUENCIAS UTR 3' PARA ESTABILIZACIÓN DEL ARN

<130> 674-146 PCT

35

<150> PCT/EP2015/073180

<151> 2015-10-07

<160> 221

40

<170> PatentIn versión 3.5

<210> 1

<211> 208

<212> RNA

45

<213> Homo sapiens

<400> 1

50

cauccugcug cugcugcugc ugcugcugcg ggucuuccug gaaucugacc auucguuguc 60

ugcuaugccc guccucacca agacugacug ccugcugcuu ugcuaugcc cgggcccaug 120

agacugacuu cccacugcuc ugccugccuc uccccacugc acuggcacag ccccgccuug 180

55

ccgcugcuga uccauugccg gugugacc 208

<210> 2

<211> 202

<212> ARN

60

<213> Homo sapiens

<400> 2

65

ES 2 753 201 T3

	gcugcugcugcug cugcugcugc ugccgggucuu ccuggaauuc gaccuuucgu ugucugcuau	60
	gcccguccuc accaagacug acugccugcu gcuuugcuac ugcccgggcc caugagacug	120
	acuuccacu gcucugccug ccucucccca cugcacuggc acagccccgc cuugccgcug	180
5	cugauccauu gccgguguga cc	202
	<210> 3	
	<211> 210	
10	<212> ARN	
	<213> Homo sapiens	
	<400> 3	
15	uuccugcugc ugucugcugcu gcugcugcug cgggucuucc uagaauucga ccauucguug	60
	ucugcuauug ccguccucac caagacugac ugccugcugc uuugcuacug cccggggcca	120
	ugagacugac uucccacugc ucugccugcc ucuccccacu gcacuggcac agccccgcu	180
20	ugccgcugcu gauccauugc cggugagacc	210
	<210> 4	
	<211> 193	
	<212> ARN	
25	<213> Homo sapiens	
	<400> 4	
	ugucugcugcu gcugcggguc uuccuggaau cugaccuuc guugucugcu augcccgucc	60
30	ucaccaagac ugacugccug cugcuuugcu acugcccggg cccaugagac ugacuuccca	120
	cugcucugcc ugccucucc cacugcacug gcacagcccc gccuugccgc ugcugaucca	180
35	uugccggcgg aca	193
	<210> 5	
	<211> 186	
	<212> ARN	
40	<213> Homo sapiens	
	<400> 5	
	gcugcugcgg gucuuccugg aaucugacca uucguugucu gcuaugcccg uccucaccaa	60
45	gacugacugc cugcugcuuu gcuacugccc gggcccauga gacugacuuc ccacugcucu	120
	gccugccucu cccacugca cuggcacagc cccgccuugc cgcugcugau ccuugccgg	180
	cguacc	186
50	<210> 6	
	<211> 170	
	<212> ARN	
	<213> Homo sapiens	
55	<400> 6	
	cuggaauucg accauucguu gucugcuauug cccguccuca ccaagacuga cugccugcug	60
60	cuuugcuacu gcccggggccc augagacuga cuucccacug cucugccugc cucuccccac	120
	ugcacuggca cagccccgcc uugccgcugc ugauccauug ccggcggacc	170
	<210> 7	
	<211> 206	
65	<212> ARN	

ES 2 753 201 T3

	<213> Homo sapiens	
	<400> 7	
5	uccugcugcu gcugcugcug cugcugcggg ucuuccugga aucugaccan ucguugucug	60
	cuaugcccgu ccucaccaag acugacugcc ugcugcuuug cuacugcccg ggcccaugag	120
	acugacuucc cacugcucug ccugccucuc cccacugcac uggcacagcc ccgccuugcc	180
10	gcugcugauc cauugccggu gggacc	206
	<210> 8	
	<211> 121	
15	<212> ARN	
	<213> Homo sapiens	
	<400> 8	
20	cugccugcug cuuugcuacu gcccggggcc augagacuga cuucccacug cucugccugc	60
	cucuccccac ugcacuggca cagccccgcc uugccgcugc ugauccauug ccgguagaac	120
	c	121
25	<210> 9	
	<211> 206	
	<212> ARN	
	<213> Homo sapiens	
30	<400> 9	
	uccugcugcu gcugcugcug cugcugcggg ucuuccugga aucugaccan uuguugucug	60
	cuaugcccgu ccucaccaag acugacugcc ugcugcuuug cuacugcccg ggcccaugag	120
35	acugacuucc cacugcucug ccugccucuc cccacugcac uggcacagcc ccgccuugcc	180
	gcugcugauc cauugccggu gggacc	206
40	<210> 10	
	<211> 196	
	<212> ARN	
	<213> Homo sapiens	
	<400> 10	
45	gcugcugcug cugcugcggg ucuuccugga aucugaccan ucguugucug cuaugcccgu	60
	ccucaccaag acugacugcc ugcugcuuug cuacugcccg ggcccaugag acugacuucc	120
	cacugcucug ccugccucuc cccacugcac uggcacagcc ccgccuugcc gcugcugauc	180
50	cauugccggu gugacc	196
	<210> 11	
	<211> 276	
55	<212> ARN	
	<213> Homo sapiens	
	<400> 11	

ES 2 753 201 T3

	ucuggccuca cugagucuga agagcuguua acuaccaugg ccaguccucc cugagucuga	60
5	ccaucuucca uccugcugcu gcugcugcug cugcugcggg ucuuccugga aucugaccan	120
	ucguugucug cuaugcccgu ccucaccaag acugacugcc ugcugcuuug cuacugcccg	180
	ggcccaugag acugacuucc cacugcucug ccugccucuc cccacugcac uggcacagcc	240
10	ccgccuugcc gcugcugauc cauugccggu gugacc	276
	<210> 12	
	<211> 276	
	<212> ARN	
15	<213> Homo sapiens	
	<400> 12	
	ucuggccuca cugagucuga agagcuguua acuaccaugg ccaguccucc cugagucuga	60
20	ccaucuucca uccugcugcu gcugcugcug cugcugcggg ucuuccugga aucugaccan	120
	ucguugucug cuaugcccgu ccucaccaag acugacugcc ugcugcuuug cuacugcccg	180
	ggcccaugag acugacuucc cacugcucug ccugccucuc cccacugcac uggcacagcc	240
25	ccgccuugcc gcugcugauc cauugccggu gugacc	276
	<210> 13	
	<211> 245	
30	<212> ARN	
	<213> Homo sapiens	
	<400> 13	
35	cuaccauggc caguccuccc ugagucugac caucuuccau ccugcugcug cugcugcugc	60
	ugcugcgggu cuuccugga ucugaccauu cguugucugc uaugcccgu cccaccaaga	120
	cugacugccu gcugcuuugc uacugcccgg gcccaugaga cugacuuccc acugcucugc	180
40	cugccuuucc ccacugcacu ggcacagccc cgccuugccg cugcugaucc auugccggug	240
	ugacc	245
	<210> 14	
45	<211> 233	
	<212> ARN	
	<213> Homo sapiens	
	<400> 14	
50	guccucccug agucugacca ucuuccaucc ugcugcugcu gcugcugcug cugcgggucu	60
	uccuggaaucc ugaccuuugc uugucugcua ugcccguccu caccaagacu gacugccugc	120
	ugcuuugcua cugcccgggc ccaugagacu gacuucccac ugcucugccu gccucucccc	180
55	acugcacugg cacagccccg ccuugccgcu gcugauccau ugccggucug aca	233
	<210> 15	
	<211> 232	
60	<212> ARN	
	<213> Homo sapiens	
	<400> 15	
65		

ES 2 753 201 T3

	guccuccug agucugacca ucuuccaucc ugcugcugcu gcugcugcug cugcgggucu	60
5	uccuggauc ugaccuuug uugucugcua ugccccuccu caccaagacu gacugccugc	120
	ugcuuugcua cugcccgggc ccaugagacu gacuuccac ugcucugccu gccuucccca	180
	cugcacuggc acagccccgc cuugccgcug cugauccau gccgguguga cc	232
10	<210> 16 <211> 208 <212> ARN <213> Homo sapiens	
15	<400> 16	
	ccauccugcu gcugcugcug cugcugcugc gggucuuccu ggaaucugac cauucguugu	60
	cugcuauGCC cguccacacc aagacugacu gccugcugcu ugcuaugcc cgggcccaug	120
20	agacugacuu cccacugcuc ugccugccuc ucccacugc acuggcacag ccccgccuug	180
	ccgcugcuga uccaugccg gugugacc	208
25	<210> 17 <211> 237 <212> ARN <213> Homo sapiens	
30	<400> 17	
	gccaguccuc ccugagucug accaucuucc auccugcugc ugcugcugcu gcugcugcgg	60
	gucuuccugg aaucugacca uucguugucu gcuauGCCg uccucaccaa gacugacugc	120
35	cugcugauuu gcuacugccc gggcccauga gacugacuuc ccacugcucu gccugccucu	180
	ccccacugca cuggcacagc cccgccuugc cgcugcugau ccauugccgg ugugacc	237
40	<210> 18 <211> 297 <212> ARN <213> Homo sapiens	
45	<400> 18	
	ugccuuccgu cuccugcugc uucuggccuc acugagucug aagagcuguu aacuaccaug	60
	gccaguccuc ccugagucug accaucuucc auccugcugc ugcugcugcu gcugcugcgg	120
50	gucuuccugg aaucugacca uucguugucu gcuauGCCg uccucaccaa gacugacugc	180
	cugcugcuuu gcuacugccc gggcccauga gacugacuuc ccacugcucu gccugccucu	240
	ccccacugca cuggcacagc cccgccuugc cgcugcugau ccauugccgg ugugacc	297
55	<210> 19 <211> 238 <212> ARN <213> Homo sapiens	
60	<400> 19	
65		

ES 2 753 201 T3

	ggccaguccu cccugagucu gaccaucuuc cauccugcug cugcugcugc ugcugcugcg	60
	ggucuuccug gaaucugacc auucguuguc ugcuaugccc guccucacca agacugacug	120
5	ccugcugcuu ugcuaugccc cgggcccuaug agacugacuu cccacugcuc ugccugccuc	180
	uccccacugc acuggcacag ccccgccuug ccgcugcuga uccauugccg gugugacc	238
10	<210> 20 <211> 238 <212> ARN <213> Homo sapiens	
	<400> 20	
15	ggccaguccu cccugagucu gaccaucuuc cauccugcug cugcugcugc ugcugcugcg	60
	ggucuuccug gaaucugacc auucguuguc ugcuaugccc guccucacca agacugacug	120
	ccugcugcuu ugcuaugccc cgggcccuaug agacugacuu cccacugcuc ugccugccuc	180
20	uccccacugc acuggcacag ccccgccuug ccgcugcuga uccauugccg gugugacc	238
25	<210> 21 <211> 219 <212> ARN <213> Homo sapiens	
	<400> 21	
30	ugaccaucu ccauccugcu gcugcugcug cugcugcugc gggucuuccu ggaucugac	60
	cauucguugu cugcuaugcc cguccucacc aagacugacu gccugcugcu uugcuaugc	120
	ccgggcccua gagacugacu ucccacugcu cugccugccu cuccccacug cacuggcaca	180
35	gccccgccuu gccgcugcug auccauugcc ggugugacc	219
40	<210> 22 <211> 237 <212> ARN <213> Homo sapiens	
	<400> 22	
45	ccaguccucc cugagucuga ccaucuucca uccugcugcu gcugcugcug cugcugcggg	60
	ucuuccugga aucugaccuu ucguugucug cuaugcccgu ccucaccaag acugacugcc	120
	ugcugcuuug cuacugcccg ggcccuaugag acugacuucc cucugcucug ccugccucuc	180
50	cccacugcac uggcacagcc cggccuugcc gcugcugauc cauugccugu gugacca	237
55	<210> 23 <211> 214 <212> ARN <213> Homo sapiens	
	<400> 23	
60	ugaccaucu ccauccugcu gcugcugcug cugcgggucu uccuggauc ugaccauucg	60
	uugucugcua ugcccguccu uaccaagacu gacugccugc ugcuuugcua cugcccgggc	120
	ccaugagacu gacuucccac ugcucugccu gccucucccc acugcacugg cacagccccg	180
65	ccuugucgcu gcugauccau ugccggugug acac	214

ES 2 753 201 T3

5	<210> 24		
	<211> 128		
	<212> ARN		
	<213> Homo sapiens		
	<400> 24		
10	gacugacugc cugcugcuuu gcuacugccc gggcccauga gacugacuuc ccacugcucu	60	
	gccugccucu cccacugca cuggcacagc cccgccuugc cgcugcugau ccuugccgg	120	
	ugugaccc	128	
15	<210> 25		
	<211> 230		
	<212> ARN		
	<213> Homo sapiens		
	<400> 25		
20	cucccugagu cugaccaucu uccauccugc ugcugcugcu gcugcugcug cgggucuucc	60	
	uggaaucuga ccuucguug ucugcuaugc ccguccucac caagacugac ugccugcugc	120	
	uuugcuacug cccggggcca ugagacugac uuccacugc ucugccugcc ucuccccacu	180	
25	gcacuggcac agccccgccu ugccgcugcu gauccauugc cggugugacc	230	
30	<210> 26		
	<211> 211		
	<212> ARN		
	<213> Homo sapiens		
	<400> 26		
35	uuccauccug cugcugcugc ugcugcugcu gcgggucuuc cuggaaucug accauucguu	60	
	gucugcuaug cccguccuca ccaagacuga cugccugcug cuuugcuacu gcccggggccc	120	
	augagacuga cuuccacug cucugccugc cucuccccac ugcacuggca cagccccgcc	180	
40	uugccgcugc ugauccauug cggugugac c	211	
45	<210> 27		
	<211> 143		
	<212> ARN		
	<213> Homo sapiens		
	<400> 27		
50	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60	
	gacuucccac ugcucugccu gccucucccc acugcacugg cacagccccg ccuugccgcu	120	
	gcugauccau ugccggugug acc	143	
55	<210> 28		
	<211> 143		
	<212> ARN		
	<213> Homo sapiens		
	<400> 28		
60			
65			

ES 2 753 201 T3

	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60
5	gacuucccac ugcucugccu gccucucucc acugcacugg cacagccccg ccuugccgcu	120
	gcugauccau ugccggugug acc	143
10	<210> 29 <211> 181 <212> ARN <213> Homo sapiens	
	<400> 29	
15	gcgggucuuu cuggaaucug aacauucguu gucugcuauug cccguccuca ccaagacuga	60
	cugccugcug cuuugcuacu gcccgggccc augagacuga cuucccacug cucugccugc	120
20	cucucuccac ugcacuggca cagccccgcc uugccgcugc ugauccauug ccggugugac	180
	c	181
25	<210> 30 <211> 154 <212> ARN <213> Homo sapiens	
	<400> 30	
30	guugucugcu augcccgucc ucaccaagac ugacugccug cugcuuugcu acugcccggg	60
	cccaugagac ugacuuccca cugcucugcc ugccucucc cacugcacug gcacagcccc	120
	gccuugccgc ugcugaucca uugccggugu gacc	154
35	<210> 31 <211> 130 <212> ARN <213> Homo sapiens	
40	<400> 31	
	caagacugac ugccugcugc uuugcuacug cccggggcca ugagacugac uucccacugc	60
45	ucugccugcc ucuccccacu gcacuggcac agccccgccu ugccgcugcu gauccauugc	120
	cggugugacc	130
50	<210> 32 <211> 138 <212> ARN <213> Homo sapiens	
	<400> 32	
55	ccucaccaag acugacugcc ugcugcuuug cuacugcccc gaccgaugag acugacuucc	60
	cacugcucug ccugccucuc cccacugcac uggcacagcc ccgccuugcc gcugcugauc	120
60	cauugccggu gugacugc	138
65	<210> 33 <211> 214 <212> ARN <213> Homo sapiens	
	<400> 33	

ES 2 753 201 T3

	cuuccauccu gcugcugcug gcugcugcugc ugcgggucuu ccuggaauccu gaccuauccu	60
5	ugucugcuau gcccguccuc accaagacug acugccugcu gcuuugcuac ugcccggggc	120
	caugagacug acuucccacu gcucugccug ccucucccca cugcacuggc acagccccgc	180
	cuugccgcug cugauccauu gccgguguga cccc	214
10	<210> 34 <211> 135 <212> ARN <213> Homo sapiens	
15	<400> 34	
	cucaccaaga cugacugccu gcugcuuugc uacugcccgg gcccaugaga cugacuuccc	60
	acugcucugc cugccucucc ccacugcacu ggcacagccc cgccuugccg cugcugaucc	120
20	auugccggug ugacc	135
	<210> 35 <211> 129 <212> ARN <213> Homo sapiens	
25	<400> 35	
30	aagacugacu gccugcugcu uugcuacugc ccggggccau gagacugacu ucccacugcu	60
	cugccugccu cuccccacug cacuggcaca gcccgcgccu gccgcugcug auccauugcc	120
	ggugugacc	129
35	<210> 36 <211> 541 <212> ARN <213> Homo sapiens	
40	<400> 36 cuccagcucg cuuccauuug cuugcagaag uucucgcugu gcucacgaag cuugcgcucc	60
	uuggaggccu cagcaacagc aucaucaagc ugagcuucca gcucuuccu gagcuucua	120
45	gcucuccgca uuuccugccg cauggcgucc accuucugcg uggccaccuc caucuccucc	180
	uccuugucuc gcagcugccg ggacaccuuc ugcgcuaaga ugggauacgg cauugagggga	240
	ucaaugugua aggauccgau cugcuucugc ccucacugag ucugaagagc uguuaacuac	300
50	cauggccagu ccuccugag ucugaccauc uuccauccug cugcugcugc ugucugcugcu	360
	gcgggucuu cuggaauccg accauucguu gucugcuau uccguccuca ccaagacuga	420
	cugccugcug cuuugcuacu gcccggggccc augagacuga cuucccacug cucugccugc	480
55	cucuccccac ugcacuggca cagccccgcc uugccgcugc ugauccauug ccggugugac	540
	c	541
60	<210> 37 <211> 125 <212> ARN <213> Homo sapiens	
65	<400> 37	

ES 2 753 201 T3

	cugacugccu gcugcuuugc uacugcccgg gcccaugaga cugacuuccc acugcucugc	60
	cugccucucc ccacugcacu ggcacagccc cgccuugccg cugcugaucc auugccggug	120
5	ugacc	125
	<210> 38	
	<211> 131	
	<212> ARN	
10	<213> Homo sapiens	
	<400> 38	
	ccaagacuga cugccugcug cuuugcuacu gcccgggccc augagacuga cuucccacug	60
15	cucugccugc cucucuccac ugcacuggca cagccccgcc uugccgcugc ugauccauug	120
	acggugugac c	131
20	<210> 39	
	<211> 245	
	<212> ARN	
	<213> Homo sapiens	
25	<400> 39	
	uaacuaccuau ggcaguccu cccugagucu gaccaucuuc cauccugcug cugcugcugc	60
	ugcugcgguu cuuccuggaa ucugaccuuu cguugucugc uaugcccguc cucaccaaga	120
30	cugacugccu gcugcuuugc uacugcccgg gcccaugaga cugacuuccc acugcucugc	180
	cugccucucc ccacugcacu ggcacagccc cgccuugccg cuguugaucc auugccggug	240
	ugacc	245
35	<210> 40	
	<211> 131	
	<212> ARN	
	<213> Homo sapiens	
40	<400> 40	
	ccaagacuga cugccugcug cuuugcuacu gcccgggccc augagacuga cuucccacug	60
	cucugccugc cucucuccac ugcacuggca cagccccgcc uugccgcugu ugauccauug	120
45	ucggugugac c	131
	<210> 41	
	<211> 131	
	<212> ARN	
50	<213> Homo sapiens	
	<400> 41	
	ccaagacuga cugccugcug cuuugcuacu gcccgggccc augagacuga cuucccacug	60
55	cucugccugc cucucuccac ugcacuggca cagccccgcc uugccgcugc ugauccauug	120
	ccggugugac c	131
60	<210> 42	
	<211> 149	
	<212> ARN	
	<213> Homo sapiens	
65	<400> 42	

ES 2 753 201 T3

	cugcuaugcc uguccucacc aagacugacu gccugcugcu uugcuacugc ccggggcccau	60
5	gagacugacu ucccacugcu cugccugccu cuccccaug cacuggcaca gcccgcgcuu	120
	gccgcugcug auccauugcc ggugugacc	149
10	<210> 43 <211> 149 <212> ARN <213> Homo sapiens	
	<400> 43	
15	cugcuaugcc cguccucacc aagacugacu gccugcugcu uugcuacugc ccggggcccau	60
	gagacugacu ucccacugcu cugccugccu cuccccacug cacuggcaca gcccgcgcuu	120
	gccgcugcug auccauugcc ggugugacc	149
20	<210> 44 <211> 147 <212> ARN <213> Homo sapiens	
25	<400> 44	
	gcuaugcccg uccucaccaa gacugacugc cugcugcuu gcuacugccc gggcccauga	60
30	gacugacuuc ccacugcucu gccugccucu ccccacugca cuggcacagc cccgcguugc	120
	cgcugcugau ccuugccgg ugugacc	147
35	<210> 45 <211> 176 <212> ARN <213> Homo sapiens	
	<400> 45	
40	ucuuccugga aucugaccau ucguugucug cuaugcccg cccaccaaag acugacugcc	60
	ugcugcuuug cuacugcccg gggcccaugag acugacuucc cacugcucug ccugccucuc	120
	cccacugcac uggcacagcc ccgcccugcc gcugcugauc cauugccggg gugacc	176
45	<210> 46 <211> 171 <212> ARN <213> Homo sapiens	
50	<400> 46	
	cuggaauug accauucguu gucugcuaug cccguccuca ccaagacuga cugccugcug	60
	cuuugcuacu gcccggggcc augagacuga cuucccacug cucugccugc cucuccccac	120
55	ugcaguggca cagccccgcc uugccgcugc ugauccauug ccggugugac c	171
60	<210> 47 <211> 171 <212> ARN <213> Homo sapiens	
	<400> 47	
65		

ES 2 753 201 T3

	cuggaaucug accauucguu gucugcuaug cccguccuca ccaagacuga cugccugcug	60
5	cuuugcuacu gcccggggccc augagacuga cuucccacug cucugccugc cucuccccac	120
	ugcacuggca cagccccgcc uugccgcugc ugauccauug ccggugugac c	171
10	<210> 48 <211> 169 <212> ARN <213> Homo sapiens	
	<400> 48	
15	ggaucugac cauucguugu cugcuaugcc cguccucacc aagacugacu gccugcugcu	60
	uugcuacugc ccggggcccau gagacugacu ucccacugcu cugccugccu cuccccacug	120
20	cacuggcaca gccccgccuu gccgcugcug auccauugcc ggugugacc	169
	<210> 49 <211> 131 <212> ARN <213> Homo sapiens	
25	<400> 49	
	ccaagacuga cugccugcug cuuugcuacu gcccggggccc augagacuga cuucccacug	60
30	cucugccugc cucuccccac ugcacuggca uagccccgcc uugccgcugc ugauccauug	120
	ccggugugac c	131
35	<210> 50 <211> 175 <212> ARN <213> Homo sapiens	
	<400> 50	
40	cuuccuggaa ucugaccauu cguugucugc uaugcccguc cucaccaaga cugacugccu	60
	gcugcuuugc uacugcccgg gcccaugaga cugacuuccc acugcucugc cugccucucc	120
	ccacugcacu ggcauagccc cgccuugccg cugcugaucc auuuccggug ugacc	175
45	<210> 51 <211> 140 <212> ARN <213> Homo sapiens	
50	<400> 51	
	cagacacccg cccccggcc cuggcuaaga auuugcuucc uguugccagc augaccuacc	60
	cucgccucuu ugaugccauc cgcugccacc uccuuuugcu ccuggacccu uuagccucuc	120
55	ugcccuucca cucucugacc	140
60	<210> 52 <211> 149 <212> ARN <213> Homo sapiens	
	<400> 52	
65		

ES 2 753 201 T3

	uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu ugccagcaug	60
	accuaccuc gccucuuuga ugccaucgc ugccaccucc uuuugcuccu ggacccuuua	120
	gccucucugc ccuuccacuc ucugacccc	149
5	<210> 53 <211> 212 <212> ARN <213> Homo sapiens	
10	<400> 53	
	cucgcuuuccu gggucugcag guccagccgg cuggcacccu ccauguacct aggggagauu	60
15	ccagccagac acccgccccc cggcccuggc uaagaaguug cuuccuguug ccagcaugac	120
	cuaccucgc cucuuugaug ccauccgcug ccaccuccuu uugcuccugg acccuuuagc	180
	cucucugccc uuccacucuc ugaccaccgc cc	212
20	<210> 54 <211> 144 <212> ARN <213> Homo sapiens	
25	<400> 54	
	uccagccaga caccgcccc cgggcccugg cuaagaagu gcuuccuguu gccagcauga	60
30	ccuaccucg ccucuuugau gccauccgc gccaccuccu uuugcuccug gacccuuuag	120
	cuucucugcc cuuccacucu cugg	144
	<210> 55 <211> 207 <212> ARN <213> Homo sapiens	
35	<400> 55	
	cgcuuuccugg gucugcaggu ccagccggcu ggcaccucc acguaccag gggagauucc	60
40	agccagacac ccgccccccg gccucggcua agaaguugcu uccuguugcc agcaugaccu	120
	accucgccc cuuugaugcc auccgcugcc accuccuuu gcuccuggac ccuuuagccu	180
45	cucugcccuu ccacucucug accaccg	207
	<210> 56 <211> 158 <212> ARN <213> Homo sapiens	
50	<400> 56	
	cauguaccta ggggagauuc cagccagaca cccgcccccc ggcccuggcu aagaaguugc	60
55	uuccuguugc cagcaugacc uaccucgccc ucuuugaugc cauccgcugc caccuccuuu	120
	ugcuccugga cccuuuagcc ucucugcccu uccacucg	158
60	<210> 57 <211> 197 <212> ARN <213> Homo sapiens	
65	<400> 57	

ES 2 753 201 T3

	cucgcuuccu gggucugcag guccagccgg cuggcacccu ccauguaccc aggggagauu	60
	ccagccagac acccgccccc cggcccuggc uaagaaguug cuuccuguug ccagcaugac	120
5	cuaccucgc cucuuugaug ccauccgcug ccaccuccuu uugcuccugg acccuuagc	180
	cucucugccc ccccgau	197
10	<210> 58 <211> 236 <212> ARN <213> Homo sapiens	
	<400> 58	
15	cccagcuccc uaggcguccc aucucguuc cugggucugc agguccagcc ggcuggcacc	60
	cuccauguac ccaggggaga uuccagccag acaccgccc cccggcccug gcuaagaagu	120
	ugcuuccugu ugccagcaug accuaccuc gccucuuuga ugccaucgc ugccaccuc	180
20	uuuuguccu ggaccuuua gccucucugc ccuuccacuc uuugaccccc aucua	236
25	<210> 59 <211> 273 <212> ARN <213> Homo sapiens	
	<400> 59	
30	ggccaccggg caugggaagu augagaaggu gcuugugaa gggggcccg cucuccuaggc	60
	gucccaucuc gcuuccuggg ucugcagguc cagccggcug gcaccucca uguaccag	120
	gagauuccag ccagacacc gcccccggc ccuggcuaa agguugcuuc cuguugccag	180
35	caugaccuac ccucgccuc uugaugccau ccgcugccac cuccuuuugc uccuggacc	240
	uuuagccuc cugcccuucc acucucugac ccc	273
40	<210> 60 <211> 171 <212> ARN <213> Homo sapiens	
	<400> 60	
45	uguaccag ggauguucca gccagacacc cggcccccg ccuggcuaa gaaguugcuu	60
	ccuguugcca gcaugaccua ccucgccuc uuugaugcca uccgcugcca ccuccuuug	120
50	cuccuggacc cuuuagccuc ucugcccuuc cacucucuga ccaccacccc c	171
55	<210> 61 <211> 134 <212> ARN <213> Homo sapiens	
	<400> 61	
	ccggcccugg cuaagaagu gcuuccuguu gccagcauga ccuaccucg ccuccuuugau	60
60	gccauccgc gccaccuccu uuugcuccug gaccuuuag ccucucugcc cuuccacucu	120
	cugaccacag cccc	134
65	<210> 62	

ES 2 753 201 T3

	<211> 140	
	<212> ARN	
	<213> Homo sapiens	
	<400> 62	
5	ccccggcccu ggcuagaag uugcuuccug uugccagcau gaccuaccu cgcucuuug	60
	augccaucg cugccaccuc cuuuugcucc uggaccuuu agccucucug cccuuccacu	120
	cucugaccac cgccccgcc	140
10	<210> 63	
	<211> 188	
	<212> ARN	
	<213> Homo sapiens	
15	<400> 63	
	ccggcuggca ccuuccaugu acccagggga gauuccagcc agacaccgc cccccggccc	60
20	uggcuaga guugcuuccu guugccagca ugaccuacc ucgccuuuu gaugccauc	120
	gcugccaccu ccuuuugcuc cuggaccuuu uagccucucu gcccuuccac ucucugacca	180
	ccgcccc	188
25	<210> 64	
	<211> 190	
	<212> ARN	
	<213> Homo sapiens	
30	<400> 64	
	gccggcuggc acccucaug uaccagggg agauuccagc cagacaccg cccccggcc	60
	cuggcuaga aguugcuucc uguugccagc augaccuacc cuagccucu ugaugccauc	120
35	cgugccacc uccuuuugc uccuggacc uuuagccucu cugcccuucc acucucugac	180
	caccgcccc	190
40	<210> 65	
	<211> 154	
	<212> ARN	
	<213> Homo sapiens	
	<400> 65	
45	uccagccaga caccgcccc cgggccugg cuaagaagu gcuuccugu gccagcauga	60
	ccuaccucg ccucuugau gccauccgc gccaccuccu uuugcuccug gaccuuuag	120
50	ccucucugcc cuuccacucu cugaccacca cccc	154
	<210> 66	
	<211> 146	
	<212> ARN	
	<213> Homo sapiens	
55	<400> 66	
	gccagacacc cgcccccg cccuggcuaa gaaguugcu ccuguugcca gcaugaccua	60
60	cccucgccuc uuugaugcca uccgucgcca ccuccuuuug cuccuggacc cuuuagccuc	120
	ucugcccuuc cacucucuga ccccc	146
65	<210> 67	

ES 2 753 201 T3

	<211> 150		
	<212> ARN		
	<213> Homo sapiens		
	<400> 67		
5	uccagccaga caccgcgcc cgggccugc cuaagaagu gcuuccugu gccagcauga	60	
	ccuaccucg ccucuugau gccaucgcu gccaccucc uuugcuccug gaccuuuag	120	
	ccucucugcc cuuccacucu cugaccccc	150	
10	<210> 68		
	<211> 233		
	<212> ARN		
	<213> Homo sapiens		
15	<400> 68		
	cggcuccua ggcguccau cucgcuuccu gggucugcag guccagccgg cuggcaccu	60	
20	ccauguacc aggggagau ccagccagac accgcgcgc cggccuggc uaagaaguug	120	
	cuuccuguug ccagcaugac cuaccucgc cucuuugaug ccauccgcug ccaccuccu	180	
	uugcuccug acccuuagc cucucugcc uuccacucuc ugaccacugc ccc	233	
25	<210> 69		
	<211> 197		
	<212> ARN		
	<213> Homo sapiens		
30	<400> 69		
	ugcaggucca gccggcugc accuccaug uaccagggg agauuccagc cagacacca	60	
35	cccccggcc cuggcuaaga aguugcuccu guugccagca ugaccuacc ucgccucuu	120	
	gaugccauc gcugccaccu ccuuuugcuc cuggacccu uagccucucu gcccuuccac	180	
	ucucugacca cuacc	197	
40	<210> 70		
	<211> 155		
	<212> ARN		
	<213> Homo sapiens		
	<400> 70		
45	uuccagccag acaccgcgc cccggccug gcuaagaagu ugcuccugu ugccagcaug	60	
	accuaccuc gccucuuga ugccaucgc ugccaccucc uuugcuccu ggaccuuua	120	
50	gccucucgc ccuuccacuc ucugaccacu gccc	155	
	<210> 71		
	<211> 134		
	<212> ARN		
55	<213> Homo sapiens		
	<400> 71		
	ccgcgcgcgc ggcgcugcu aagaaguugc uuccuguugc cagcaugacc uaccucgcc	60	
60	ucuugaugc cauccgcugc caccuccuu ugcuccugga ccuuuagcc ucucugccu	120	
	uccacucucu gacc	134	
65	<210> 72		

ES 2 753 201 T3

	<211> 147	
	<212> ARN	
	<213> Homo sapiens	
	<400> 72	
5	cagccagaca cccgcccccc ggcccuggcu aagaaguugc uuccuguugc cagcaugacc	60
	uaccucgcc ucuuugaugc cauccgcugc caccuccuuu ugcuccugga cccuuuagcc	120
	ucucugcccu uccacucucu gaacacc	147
10	<210> 73	
	<211> 192	
	<212> ARN	
	<213> Homo sapiens	
15	<400> 73	
	gagccuacuc ugaugaccgu ggccuuggcu ccuccaggaa ggcucaggag cccuaccucc	60
20	cugccauuau agcugcuccc cgccagaagc cugugccaac ucucugcauu cccugaucuc	120
	cuguggcugu caccuuggu caccuccgug cugucacugc caucucuccc cugaccccuc	180
	gaacccaucc ua	192
25	<210> 74	
	<211> 191	
	<212> ARN	
	<213> Homo sapiens	
30	<400> 74	
	gagccuacuc ugaugaccgu ggccuuggcu ccuccaggaa ggcucaggcg cccuaccucc	60
	cugccauuau agcugcuccc cgccagaagc cugugccaac ucucugcauu cccugaucuc	120
35	caucccugug gcugucaccc uuggucaccu ccgugcuguc acugccaucu cccccugac	180
	cccucuaacc c	191
40	<210> 75	
	<211> 162	
	<212> ARN	
	<213> Homo sapiens	
	<400> 75	
45	gccuacucug augaccgugg ccuuggcucc uccaggaagg cucaggagcc cuaccucccu	60
	gccauuauag cugcuccccg ccagaagccu gugccaacuc ucugcauucc cugaucucca	120
50	ucccuguggc ugucacccuu gguccacugc caucucuccc cc	162
	<210> 76	
	<211> 183	
	<212> ARN	
	<213> Homo sapiens	
55	<400> 76	
	gagccuacuc ugaugaccgu ggccuuggcu ccuccaggaa ggcucaggag cccuaccucc	60
60	cugccauuau agcugcuccc cgccagaagc cugugccaac ucucugcauu cccugaucuc	120
	caucccugug gcugucaccc uuggucaccu ccgugcuguu acugccaucu cccccugac	180
	ccc	183
65		

ES 2 753 201 T3

5	<210> 77 <211> 240 <212> ARN <213> Homo sapiens	
	<400> 77	
10	gaagagccua cucugaugac cguggccuug gcuccuccag gaaggcucag gagcccuacc	60
	uccugccau uauagcugcu ccccgccaga agccugugcc aacucucugc auucccugau	120
	cuccaucccu guggcuguca uccuugguca ccuccgugcu gucacugcca ucucuccccu	180
15	gacccucua acccauccuc ugccuccuc ccugcaguca gaggguccug uucccaacca	240
20	<210> 78 <211> 166 <212> ARN <213> Homo sapiens	
	<400> 78	
25	uguggccuug gcuccuccag gaaggcuaag gagcccuacc uccugccau uauagcugcu	60
	cccgccaga agccugugcc aacucucugc auucccugau cuccaucuccu guggcuguca	120
	ccuugguca ccuccgugcu gucacugcca ucucuccccu gacccc	166
30	<210> 79 <211> 155 <212> ARN <213> Homo sapiens	
35	<400> 79	
	gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuaua gcugcucccc	60
	gccagaagcc ugugccaacu cucugcauuc ccugaucucc auccugugg cugucacccu	120
40	uggucaccuc cgugcuguca cugccaucuc ccccc	155
45	<210> 80 <211> 310 <212> ARN <213> Homo sapiens	
	<400> 80	
	agagccuacu cugaugaccg ugGCCUUGGC uccuccagga aggcucagga gcccuaccuc	60
50	ccugccauua uagcugcucc ccgcccagaag ccugugccaa cucucugcau ucccugaucu	120
	ccaucuccugu ggcugucacc cuuggucacc uccgugcugu cacugccauc uccccccuga	180
	ccccucuaac ccauccucug ccuccucccc ugcagucaga ggguccuguu cccaucagcg	240
55	auucccugc uaaaacccuu ccaugacucc ccacugcccu aagcugaggu cagucuccca	300
	agccugacau	310
60	<210> 81 <211> 204 <212> ARN <213> Homo sapiens	
65	<400> 81	

ES 2 753 201 T3

	uauagcugcu ccccgccaga agccugugcc aacucucugc auucccugau cuccaucuccu	60
	guggcuguca cccuugguca ccuccgugcu gucacugcca ucucuccccu gacccucua	120
	acccaucucc ucgcucccuc ccugcaguca gaggguccug uucccaucag cgauuccccu	180
5	gcuuaaacc uuccaugaca gccc	204
	<210> 82	
	<211> 169	
	<212> ARN	
10	<213> Homo sapiens	
	<400> 82	
	ucugcauucc cugaucucca ucccuguggc ugucacccu ggucaccucc gugcugucac	60
15	ugccaucucc cccugaccc cucuaacca uccucugccu cccucccugc agucagagg	120
	uccuguuccc aucagcgauu cccugcuua agcccuucca ugacucccc	169
20	<210> 83	
	<211> 217	
	<212> ARN	
	<213> Homo sapiens	
25	<400> 83	
	cucccugcca uuauagcugc uccccgccag aagccugugc caacucucug cauucccuga	60
	ucuccauccc uguggcuguc acccuugguc accuccgugc ugucacugcc aucucucccc	120
30	ugacccucuc aaccauccu cugccucccu cccugcaguc agaggguccu guucccauca	180
	gcgauuccc ugcuaaacc cuuccaugac ucccca	217
35	<210> 84	
	<211> 272	
	<212> ARN	
	<213> Homo sapiens	
40	<400> 84	
	gccuacucug augaccgugg ccuugggucc uccaggaagg cucaggagcc cuaccucccu	60
	gccauuauag cugcuccccg ccagaagucu gugccaacuc ucugcauucc cugaucucca	120
	ucccuguggc ugucacccu ggucaccucc gugcugucac ugccaucucc cccugaccc	180
45	cucuaacca uccucugccu cccucccugc agucagaggg uccuguuccc aucagcgauu	240
	cccugcuua aaccuucca ugacucccu cu	272
50	<210> 85	
	<211> 131	
	<212> ARN	
	<213> Homo sapiens	
55	<400> 85	
	cuaccucccu gccauuauag cugcuccccg ccagaagccu gugccaacuc ucugcauucc	60
	cugaucucca ucccuguggc ugucacccu ggucaccucc gugcugucac ugccaucucc	120
60	cccugaccc c	131
	<210> 86	
	<211> 136	
	<212> ARN	
65		

ES 2 753 201 T3

	<213> Homo sapiens	
	<400> 86	
5	cugguacugc augcacgcaa ugcua gcugc cccuuucccg uccuggguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
	aguuccagac accucc	136
10	<210> 87 <211> 203 <212> ARN <213> Homo sapiens	
15	<400> 87	
	ccggcccuuc ccccguuuug acaugugua accgacaguc ugccugggcc acagcccucu	60
	caccucuggua cugcaugcac gcaaugcuag cugcccccuu cccguccugg gcaccccgag	120
20	ucucucccga ccccgggucc cagguaugcu cccaccucca ccugcccccac ucaccaccuc	180
	ugcuaguucc agacaccccc gcg	203
25	<210> 88 <211> 197 <212> ARN <213> Homo sapiens	
30	<400> 88	
	ccuuccccccg uuuugaacau guguaaccga cagucugccu gggccacagc ccucucaccc	60
	ugguacugca ugcacgcaau gcuagcugcc cccuuucccg ccugggcacc ccgagucucc	120
35	cccgacccccg ggucccaggu augcucccac cuccaccugc cccacucacc accucugcua	180
	guuccagaca ccuccac	197
40	<210> 89 <211> 273 <212> ARN <213> Homo sapiens	
45	<400> 89	
	ucugccuggg ccacagcccu cucaccucgg uacugcaugc acgcaaugcu agcugccccc	60
	uucccguccu gggcaccccc agucucuccc gaccccggu cccagguaug cucccaccuc	120
50	caccugcccc acucaccacc ucugcuaguu ccagacaccu ccacgcccac cugguccucu	180
	cccaucgcc acaaaagggg gggcacgagg gacgagcuua gcugagcugg gaggagcagg	240
	gugagggugg gcgaccaggg auucccccac ccc	273
55	<210> 90 <211> 199 <212> ARN <213> Homo sapiens	
60	<400> 90	
65		

ES 2 753 201 T3

	ugacaccuca gcugacagcg ugggcaacgc cugccgccug cucugaggcc cgauccagug	60
	ggcaggccaa ggccugcugg gccccgcgg acccaggugc ucugggucac ggucccuguc	120
5	cccgcacccc cgcuucuguc ugccccauug uggcuccuca ggcucucucc ccugcucucc	180
	caccucuacc uccaccca	199
10	<210> 91 <211> 193 <212> ARN <213> Homo sapiens	
	<400> 91	
15	cucagcugac agcgugggca acgccugccg ccugcucuga gggccgaucc agugggcagg	60
	ccaaggccug cuggggcccc gcggaccag gugcucuggg ucacgguccc uguccccgca	120
20	ccccgcuuc ugucugcccc auuguggcuc cucaggcucu cuccccugcu cucccaccuc	180
	uaccuccacu ccc	193
25	<210> 92 <211> 158 <212> ARN <213> Homo sapiens	
	<400> 92	
30	ucugaggccc gauccagugg gcaggccaag gccugcuggg cccccgcgga cccaggugcu	60
	cugggucacg gucccuugucc ccgcaccccc gcuucugucu gcccacauugu ggcuccucag	120
	gcucucuccc cugcucuccc accucuaccu ccaccccc	158
35	<210> 93 <211> 153 <212> ARN <213> Homo sapiens	
40	<400> 93 ggcccgaucc agugggcagg ccaaggccug cuggggcccc gcggaccag gugcucuggg	60
	ucacgguccc uguccccgca cccccgcuuc ugucugcccc auuguggcuc cucaggcucu	120
45	cuccccugcu cucccaccuc uaccuccacc ccc	153
50	<210> 94 <211> 128 <212> ARN <213> Homo sapiens	
	<400> 94	
55	gccugcuggg cccccgcgga cccaggugcu cugggucacg gucccuugucc ccgcaccccc	60
	gcuucugucu gcccacauugu ggcuccucag gcucucuccc cugcucuccc accucuaccu	120
	ccgcccc	128
60	<210> 95 <211> 116 <212> ARN <213> Homo sapiens	
65	<400> 95	

ES 2 753 201 T3

	cgcggaacca ggugcucugg gucacggucc cugucggcg accccgcuu cugucugccc	60
	cauuguggcu ccuagggcuc ucggggcg ucgggaccu uuaccuccac ccuac	116
5	<210> 96 <211> 190 <212> ARN <213> Homo sapiens	
10	<400> 96	
	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
15	gccugcuggg cggcgcgga cccaggugcu cugggucacg guccugucc ccgcaccccc	120
	gcuucugucu gggccauugu ggcuccucag gcucucucc cugcucucc accucuaccu	180
	ccacccccac	190
20	<210> 97 <211> 191 <212> ARN <213> Homo sapiens	
25	<400> 97	
	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
	gccugcuggg cggcgcgga cccaggugcu cugggucacg guccugucc ccgcaccccc	120
30	gcuucugucu gggccauugu ggcuccucag gcucucucc cugcucucc accucuaccu	180
	ccacccccaa c	191
35	<210> 98 <211> 191 <212> ARN <213> Homo sapiens	
40	<400> 98 cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
	gccugcuggg cggcgcgga cccaggugcu cugggucacg guccugucc ccgcaccccc	120
	gcuucugucu gggccauugu ggcuccucag gcucucucc cugcucucc accucuaccu	180
45	ccacccccaa c	191
50	<210> 99 <211> 184 <212> ARN <213> Homo sapiens	
	<400> 99	
55	cgugggcaac gccugccgcc ugucugagg cccgauccag ugggcaggcc aaggccugcu	60
	ggggcccgcc ggacccaggu gcucuggguc acgguccug ucccgccacc cccgcuucug	120
	ucugcccau uguggcuccu caggcucucu cccugcucu cccaccucua ccuccacca	180
60	cacc	184
65	<210> 100 <211> 238 <212> ARN <213> Homo sapiens	

ES 2 753 201 T3

	<400> 100	
	uccugagggga cugggacucc ccuacagcc augaccuuga caccucagcu gacagcgugg	60
5	gcaacgccug ccgccugcuc ugaggcccg auccagugggc aggccaaaggc cugcuggggcc	120
	cccgcgggacc caggugcucu gggucacggg cccugucccc gcacccccgc uucugucugc	180
	cccauugugg cuccucaggc ucucuccccu gcucucccac cucuaccucc acgcccac	238
10	<210> 101 <211> 234 <212> ARN <213> Homo sapiens	
15	<400> 101	
	cugagggacu gggacucccc uuacagccau gaccuugaca ccucagcuga cagcgugggc	60
20	aacgccugcc gccugcucug agggccgauc cagugggcag gccaaaggccu gcuggggccc	120
	cgcggaacca ggugcucugg gucacggucc cuguccccgc acccccgcuu cugucugccc	180
	cauuguggcu ccucaggcuc ucuccccugc ucuccaccu cuaccuccac accu	234
25	<210> 102 <211> 258 <212> ARN <213> Homo sapiens	
30	<400> 102	
	ugcggagcca gcuggaggcc auuuuccuga gggacuggga cuccccuac agccaugacc	60
35	uugacaccuc agcugacagc gugggcaacg ccugccgccu gcucugaggc ccgauccagu	120
	gggcaggcca aggccugcug gggccccgcg gaccagggug cucuggguca cggucccugu	180
	ccccgcaccc ccgcuucugu cugccccauu guggcuccuc aggcucucuc cccugcucuc	240
40	ccaccucuac cuccaccc	258
45	<210> 103 <211> 249 <212> ARN <213> Homo sapiens	
	<400> 103	
	cuggaggcca uuuuccugag ggacugggac uccccuaca gccaugaccu ugacaccuca	60
50	gcugacagcg ugggcaacgc cugccgccug cucugaggcc cgauccagug ggcaggccaa	120
	ggccugcugg gccccgcgg acccaggugc ucugggucac ggucccuguc cccgcacccc	180
55	cgcuucuguc ugccccauug uggcuccuca ggcucucucc ccugcucucc caccucuacc	240
	uccccccac	249
60	<210> 104 <211> 226 <212> ARN <213> Homo sapiens	
	<400> 104	
65		

ES 2 753 201 T3

	cugggacucc ccuuacagcc augaccuuga caccucagcu gacagcgugg gcaacgccug	60
	ccgccugcuc ugaggcccaa uccagugggc aggccaaaggc cugcugggcc cccgcggacc	120
	caggugcucu gggucacggu cccugucucc gcacccccgc uucugucugc cccauugugc	180
5	cuccuuaggc ucucuccccu gcucucccac cucuaccucc accccc	226
	<210> 105	
	<211> 133	
10	<212> ARN	
	<213> Homo sapiens	
	<400> 105	
15	gcagcaaugc agcucaaaac gcuuagccua gccacacccc cacgguaaac agcagugauu	60
	aacuuuuagc aauaaacgaa aguuuaacua agcuauacua accccagggu uggucaauuu	120
	cgugccagcc acc	133
20	<210> 106	
	<211> 233	
	<212> ARN	
	<213> Homo sapiens	
25	<400> 106	
	cuuucuauua gcucuuaqua agauuacaca ugcaagcauc cccguuccag ugaguucacc	60
	cucuaaaucua ccacgauaaa aagggaacag caucaagcac gcagcaaugc agcucaaaac	120
30	gcuuagccua gccacacccc cacgggaaac agcagugauu aaccuuuagc aauaaacgaa	180
	aguuuaacua agcuauacua accccagggu uggucaauuu cgugccagcu acc	233
35	<210> 107	
	<211> 190	
	<212> ARN	
	<213> Homo sapiens	
40	<400> 107	
	guuccaguga guucacccuc uaaaucacca cgaucaaaag ggacaagcau caagcacgca	60
	gcaaugcagc ucaaaacgcu uagccuagcc acacccccac gggaaacagc agugauuaac	120
45	cuuugcaau aaacgaaagu uuaacuaagc uauacuaacc ccaggguugg ucaauuucgu	180
	gccagccacc	190
50	<210> 108	
	<211> 154	
	<212> ARN	
	<213> Homo sapiens	
55	<400> 108	
	aaagggacaa gcaucaagca cgagcaaug cagcucaaaa cgcuuagccu agccacaccc	60
	ccacgggaaa cagcagugau uaaccuuuag caauaaacga aaguuaacu aagcuauacu	120
60	aaccccaggg uuggucaauu ucgugccagc cacc	154
	<210> 109	
	<211> 158	
65	<212> ARN	

ES 2 753 201 T3

	<213> Homo sapiens	
	<400> 109	
	ucaaaaggga caagcaucaa gcacgcaaca augcagcuca aaaacgcuua gccuagccac	60
5	acccccacgg gaaacagcag ugauuaaccu uuagcaauaa acgaaaguuu aacuaagcua	120
	cacuaacccc agggguugguc aaauucgugc cagccacc	158
10	<210> 110 <211> 210 <212> ARN <213> Homo sapiens	
	<400> 110	
15	uacacaugca agcauccccg uuccagugag uucaccucuc aaauaccac gaucaaaagg	60
	gacaagcauc aagcacgcag caaugcagcu caaaaacgcu uagccuagcc acacccccac	120
20	gggaaacagc agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc	180
	ccagggguugg ucaauuucgu gccagccacc	210
25	<210> 111 <211> 143 <212> ARN <213> Homo sapiens	
	<400> 111	
30	caucaagcac gcagcaaugc agcucaaaac gcuuagccua gccacacccc cacgggaaac	60
	agcagugauu aaccuuuagc aaauaacgaa aguuuaacua agcuauacua accccagggg	120
35	uggucaauuu cgugccaacc acc	143
40	<210> 112 <211> 154 <212> ARN <213> Homo sapiens	
	<400> 112	
	aaagggacaa gcaucaagca cgcagcaaug cagcucaaaa cgcuuagccu agccacaccc	60
45	ccacgggaaa cagcagugau uaaccuuuag caauaacga aaguuaacu aagcuauacu	120
	aaccccaggg uuggucaauu ucgugccaac cacc	154
50	<210> 113 <211> 140 <212> ARN <213> Homo sapiens	
	<400> 113	
55	caagcacgca acaaugcagc ucaaaacgcu uagccuagcc acacccccac gggaaacagc	60
	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccagggguugg	120
60	ucaauuucgu gccaccacc	140
65	<210> 114 <211> 144 <212> ARN	

ES 2 753 201 T3

	<213> Homo sapiens	
	<400> 114	
5	caucaagcac gcagcaaugc agcucaaaac gcuuagccua gccacacccc caugggaaac	60
	agcagugauu aaccuuuagc aaauaaacgaa aguuuaacua agcuauacua accccagggg	120
	uggucaauuu cgugccagcu cacc	144
10	<210> 115 <211> 142 <212> ARN <213> Homo sapiens	
15	<400> 115	
	caagcacgca gcaaugcagc ucaaaacgcu uagccuagcc acacccccac gggaaacagc	60
	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccagggguugg	120
20	ucaauuucgu gccagccaca cc	142
25	<210> 116 <211> 140 <212> ARN <213> Homo sapiens	
	<400> 116	
	caagcacgca gcaaugcagc ucaaaacgcu uagccuagcc acacccccac gggaaacagc	60
30	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccagggguugg	120
	ucaauuucgu gccagccacc	140
35	<210> 117 <211> 162 <212> ARN <213> Homo sapiens	
40	<400> 117	
	cacgaucaaa agggacaagc aucaagcacg cagcaaugca gcucaaaacg cuuagccuag	60
	ccacaccccc acgggaaaca gcagugauua accuuuagca auaaacgaaa guuuuacuaa	120
45	gcuaucuaa ccccagggguu ggucuuuuc gugccagcca cc	162
50	<210> 118 <211> 170 <212> ARN <213> Homo sapiens	
	<400> 118	
	uaaaucacca cgaucaaaag ggacaagcau caagcacgca gcaaugcagc ucaaaacgcu	60
55	uagccuagcc acacccccac gggaaacagc agugauuaac cuuuagcaau aaacgaaagu	120
	uuaacuaagc uauacuaacc ccagggguugg ucaauuucgu gccagccacc	170
60	<210> 119 <211> 236 <212> ARN <213> Homo sapiens	
65	<400> 119	

ES 2 753 201 T3

	agccuuucua uuagcucuua gaaagauuac acaugcaagc aucccccguuc cagugaguuc	60
	accucuaaaa ucaccacgau caaaaggagc aagcaucaag cacgcagcaa ugcagcucaa	120
	aacgcuuagc cuagccacac cccacggga aacagcagug auuaaccuuu agcaauaaac	180
5	gaaaguuaa cuaagcuua cuaacccag gguuggucaa uuucgugcca gccacc	236
	<210> 120	
	<211> 151	
10	<212> ARN	
	<213> Homo sapiens	
	<400> 120	
15	gggacaagca ucaagcacgc agcaaugcag cucaaaacgc uuagccuagc cacaccccca	60
	cgggaaacag cagugauuaa cccuuagcaa uaaacgaaag uuuaacuaag cuauacuaac	120
	cccagggguug gucaauuucg ugccagccac c	151
20	<210> 121	
	<211> 151	
	<212> ARN	
	<213> Homo sapiens	
25	<400> 121	
	gggacaagca ucaagcacgc agcaaugcag cucaaaacgc uuagccuagc cacaccccca	60
	cgggaaacag cagugauuaa cccuuagcaa uaaacgaaag uuuaacuaag cuauacuaac	120
30	cccagggguug gucaauuucg ugccagccac c	151
	<210> 122	
	<211> 253	
35	<212> ARN	
	<213> Homo sapiens	
	<400> 122	
40	uuucgccccca gcuuugcagg augaaacacu uccccgcuug gcucucuauuc uuccacaaga	60
	gagaccuuuc uccggaccug guugcuacug guucagcaac ucugcagaaa auguccuccc	120
	cuguggcugc cucagcucau gccuuuggcc ugaaguccca gcauugaugg cagccccuca	180
45	ucuuccaagu uuugugcucc cccuuaccua acgcuuccug ccucccaugc aucuguacuc	240
	cuucugugcc acu	253
	<210> 123	
	<211> 253	
50	<212> ARN	
	<213> Homo sapiens	
	<400> 123	
55	uuucgccccca gcuuugcagg augaaacacu uccccgcuug gcucucuauuc uuccacaaga	60
	gagaccuuuc uccggaccug guugcuacug guucagcaac ucugcagaaa auguccuccc	120
	cuguggcugc cucagcucau gccuuuggcc ugaaguccca gcauugaugg cagccccuca	180
60	ucuuccaagu uuugugcucc cccuuaccua acgcuuccug ccucccaugc aucuguacuc	240
	cuucugugcc acu	253
65	<210> 124	

ES 2 753 201 T3

	<211> 250	
	<212> ARN	
	<213> Homo sapiens	
5	<400> 124	
	ccccagcuuu gcaggaugaa acacuucccc gcuuggcucu cauucuucca caagagagac	60
	cuuucucccg accugguugc uacugguuca gcaacucugc agaaaauguc cuccccugug	120
10	gcugccucag cucaugccuu ugGCCUGAAG ucccagcauu gauggcagcc ccucaucuuc	180
	caaguuuugu gcuccccuuu accuaacgcu uccugccucc caugcaucug uacuccuccu	240
	gugccacaaa	250
15	<210> 125	
	<211> 250	
	<212> ARN	
	<213> Homo sapiens	
20	<400> 125	
	ccccagcuuu gcaggaugaa acacuucccc gcuuggcucu cauucuucca caagagagac	60
	cuuucucccg accugguugc uacugguuca gcaacucugc agaaaauguc cuccccugug	120
25	gcugccucag cucaugccuu ugGCCUGAAG ucccagcauu gauggcagcc ccucaucuuc	180
	caaguuuugu gcuccccuuu accuaacgcu uccugccucc caugcaucug uacuccuccu	240
	gugccacaaa	250
30	<210> 126	
	<211> 233	
	<212> ARN	
	<213> Homo sapiens	
35	<400> 126	
	cuuugcagga ugaaacacuu ccccgcuugg cucucauucu uccacaagag agaccuuucu	60
40	ccggaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucccc uguggcugcc	120
	ucagcucaug ccuuuggccu gaagucccag cauugauggc agccccucau cuuccaagu	180
	uugugcuccc cuuuaccuaa cgcuuuccugc cucccaugca ucuguacucc ucc	233
45	<210> 127	
	<211> 228	
	<212> ARN	
	<213> Homo sapiens	
50	<400> 127	
	ugcaggauga aacacuuccc cgcuuuggcuc ucauucuucc acaggagaga ccuucuccg	60
55	gaccugguug cuacugguuc agcaacucug cagaaaaugu ccuccccugu ggcugccuca	120
	gcucaugccu uuggccugaa gucccagcau ugauggcagc ccucaucuuc ccaaguuuug	180
	ugcuccccuu uaccuaacgc uuccugccuc ccaugcaucu guacucca	228
60	<210> 128	
	<211> 280	
	<212> ARN	
	<213> Homo sapiens	
65	<400> 128	

ES 2 753 201 T3

	gugaagauga ccacauuca ggaagaaccu ucugccccag cuuugcagga ugaaacacuu	60
	ccccgcuugg cucuccuucu uccacaagag agaccuuucu ccggaccugg uugcuacugg	120
5	uucagcagcu cugcagaaaa uguccucccu uguggcugcc ucagcucgua ccuuuggccu	180
	gaagucccag cauuaauggc agccccucau cuuccaaguu uugugcuccc cuuuaccuaa	240
	ugcuuccugc cucccaugca ucuguacucc ugcugugcca	280
10	<210> 129 <211> 207 <212> ARN <213> Homo sapiens	
15	<400> 129 uccacaagag agaccuuucu ccggaccugg cugcuacugg uucagcagcu cugcagaaaa	60
	uguccucccu uguggcugcc ucagcucgua ccuuuggccu gaagucccag cauuaauggc	120
20	agccccucau cuuccaaguu uugugcuccc cuuuaccuaa ugcuuccugc cucccaugca	180
	ucuguacucc ugcugugcca caaacac	207
25	<210> 130 <211> 207 <212> ARN <213> Homo sapiens	
30	<400> 130 uccacaagag agaccuuucu ccggaccugg cugcuacugg uucagcagcu cugcagaaaa	60
	uguccucccu uguggcugcc ucagcucgua ccuuuggccu gaagucccag cauuaauggc	120
35	agccccucau cuuccaaguu uugugcuccc cuuuaccuaa ugcuuccugc cucccaugca	180
	ucuguacucc ugcugugcca caaacac	207
40	<210> 131 <211> 163 <212> ARN <213> Homo sapiens	
45	<400> 131 gcuacugguu cagcagcucu gcagaaaug uccucccuug uggcugccuc agcucguacc	60
	uuuggccuga agucccagca uuaauggcag ccccucaucu uccaaguuuu gugcuccccu	120
50	uuaccuaaug cuuccugccu cccaugcauc uguacuccug cgu	163
55	<210> 132 <211> 249 <212> ARN <213> Homo sapiens	
	<400> 132	

ES 2 753 201 T3

	agaaccuucu gccccagcuu ugcaggauga aacacuuccc cgcuggcuc ucauucuucc	60
	acaagagaga ccuucuccg gaccugguug cuacugguuc agcagcucug cagaaaugu	120
5	ccuccuugu ggcugccuca gcucguaccu uuggccugaa gucccagcau uaauggcagc	180
	ccucaucuu ccaaguuuug ugcuccccuu uaccuaaugc uuccugccuc ccaugcaucu	240
	guacuccug	249
10	<210> 133 <211> 249 <212> ARN <213> Homo sapiens	
15	<400> 133	
	agaaccuucu gccccagcuu ugcaggauga aacacuuccc cgcuggcuc ucauucuucc	60
20	acaagagaga ccuucuccg gaccugguug cuacugguuc agcagcucug cagaaaugu	120
	ccuccuugu ggcugccuca gcucguaccu uuggccugaa gucccagcau uaauggcagc	180
	ccucaucuu ccaaguuuug ugcuccccuu uaccuaaugc uuccugccuc ccaugcaucu	240
25	guacuccug	249
	<210> 134 <211> 264 <212> ARN <213> Homo sapiens	
30	<400> 134	
	ugaagaugac cacauucaag gaagaaccuu cugccccagc uuugcaggau gaaacacuuc	60
35	cccgcuggc ucucauucuu ccacaagaga gaccuucuc cggaccuggu uguucagcag	120
	cucugcagaa aauguccucc cuuguggcug ccucagcucg uaccuuuggc cugaaguccc	180
40	agcauuaug gcagcccuc aucuuccaag uuuugugcuc cccuuuaccu aaugcuuccu	240
	gccuccaug caucuguacu ccug	264
	<210> 135 <211> 270 <212> ARN <213> Homo sapiens	
45	<400> 135	
50	ugaagaugac cacauucaag gaagaaccuu cugccccagc uuugcaggau gaaacacuuc	60
	cccgcuggc ucucauucuu ccacaagaga gaccuucuc cggaccuggu ugcuaucuggu	120
	ucagcagcuc ugcagaaaau guccucccuu guggcugccu cagcucguac cuuuggccug	180
55	aaugcccagc auuaauggca gcccucauc uuccaaguuu ugugcucccc uuuaccuaau	240
	gcuuccugcc ucccaugcau cuguacuccc	270
60	<210> 136 <211> 245 <212> ARN <213> Homo sapiens	
65	<400> 136	

ES 2 753 201 T3

	ccacauucaa ggaagaaccu ucugccccag cuuugcagga ugaaacacuu ccccgcuugg	60
	cucucauucu uccacaagag agaccuuucu ccggaccugg uugcuacugg uucagcagcu	120
5	cugcagaaaa uguccucccu uguggcugcc ucagcucgua ccuuuggccu gaagucccag	180
	cauuaauggc agccccucau cuuccaaguu uugugcuccc cuuuaccuaa ugcuuuccugc	240
	cccc	245
10	<210> 137 <211> 257 <212> ARN <213> Homo sapiens	
15	<400> 137 gugaagauga ccacauucaa ggaagaaccu ucugccccag cuuugcagga ugaaacacuu	60
	cccgcuugg cucucauucu uccacaagag agaccuuucu ccggaccugg uugcuacugg	120
20	uucagcagcu cugcagaaaa uguccucccu uguggcugcc ucagcucgua ccuuuggccu	180
	gaagucccag cauuaauggc agccccucau cuuccaaguu uugugcuccc cuuuaccuaa	240
	ugcuuccugc cccccau	257
25	<210> 138 <211> 209 <212> ARN <213> Homo sapiens	
30	<400> 138 agaaccuucu gcccagcguu ugcaggauga aacacuuccc cgcuggcuc ucauucuucc	60
35	acaagagaga ccuuucuccg gaccugguug cuacugguuc agcagcucug cagaaaugu	120
	ccucccuugu ggcugccuca gcucguaccu uuggccugaa gucccagcau uaauggcagc	180
	cccucaucu ccaaguuuug ugcuccccc	209
40	<210> 139 <211> 276 <212> ARN <213> Homo sapiens	
45	<400> 139 cugaagugaa gaugaccaca uucaaggaag aaccuucugc ccagcuuug caggaugaaa	60
50	cacuuccccg cuuggcucuc auucuuccac aagagagacc uuucuccgga ccugguugcu	120
	acugguucag cagcucugca gaaaauugucc ucccuugugg cugccucagc ucguaccuuu	180
	ggccugaagu ccagcauua auggcagccc cucaucuucc aaguuuugug cuccccuuua	240
55	ccuaaugcuu ccugccuccc augcaucugu acuccu	276
60	<210> 140 <211> 276 <212> ARN <213> Homo sapiens <400> 140	
65		

ES 2 753 201 T3

	cugaagugaa gaugaccaca uucaaggaag aaccuucugc cccagcuuug caggaugaaa	60
	cacuuccccg cuuggcucuc auucuuccac aagagagacc uuucuccgga ccugguugcu	120
5	acugguucag cagcucugca gaaaauugucc ucccuugugg cugccucagc ucguaccuuu	180
	ggccugaagu cccagcauua auggcagccc cucaucuucc aaguuuugug cuccccuuua	240
	ccuaaugcuu ccugccuccc augcaucugu acuccu	276
10	<210> 141 <211> 270 <212> ARN <213> Homo sapiens	
15	<400> 141	
	aagaugacca cauucaagga agaaccuucu gcccagcuu ugcaggauga aacacuuccc	60
	cgcuuggcuc ucauucuucc acaagagaga ccuuucuccg gaccugguug cuacugguuc	120
20	agcagcucug cagaaaaugu ccucccuugu ggcugccuca gcucguaccu uuggccugaa	180
	gucccgcau uaauggcagc cccucaucu ccaaguuuug ugcuccccuu uaccuaaugc	240
	uuccugccuc ccaugcaucu guacuccugc	270
25	<210> 142 <211> 211 <212> ARN <213> Homo sapiens	
30	<400> 142	
	ccccgcuugg cucucauucu uccacaagag agaccuuucu ccggaccugg uugcuacugg	60
35	uucagcagcu cugcagaaaa uguccuccu uguggcugcc ucagcucgua ccuuuggccu	120
	gaagucccag cauuaauggc agcccucacu cuuccaaguu uugugcuccc cuuuaccuaa	180
	ugcuuccugc cucccaugca ucuguacucc u	211
40	<210> 143 <211> 211 <212> ARN <213> Homo sapiens	
45	<400> 143	
	ccccgcuugg cucucauucu uccacaagag agaccuuucu ccggaccugg uugcuacugg	60
50	uucagcagcu cugcagaaaa uguccuccu uguggcugcc ucagcucgua ccuuuggccu	120
	gaagucccag cauuaauggc agcccucacu cuuccaaguu uugugcuccc cuuuaccuaa	180
	ugcuuccugc cucccaugca ucuguacucc u	211
55	<210> 144 <211> 286 <212> ARN <213> Secuencia artificial	
60	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 144	
65		

ES 2 753 201 T3

	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60
	gacuucccac ugcucugccu gccucucccc acugcacugg cacagccccg ccuugccgcu	120
	gcugauccau ugccggugug accugcccgu ccucaccaag acugacugcc ugcugcuuug	180
5	cuacugcccc ggcccaugag acugacuucc cacugcucug ccugccucuc cccacugcac	240
	uggcacagcc ccgccuugcc gcugcugauc cauugccggu gugacc	286
10	<210> 145 <211> 292 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 145	
20	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60
	gacuucccac ugcucugccu gccucucccc acugcacugg cacagccccg ccuugccgcu	120
	gcugauccau ugccggugug accuuccagc cagacacccg ccccccgggc cuggcuaaga	180
25	aguugcuucc uguugccagc augaccuacc cucgccucuu ugaugccauc cgugccacc	240
	uccuuuugcu ccuggacccu uuagccucuc ugcccuucca cucucugacc cc	292
30	<210> 146 <211> 298 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 146	
40	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60
	gacuucccac ugcucugccu gccucucccc acugcacugg cacagccccg ccuugccgcu	120
	gcugauccau ugccggugug accgccuugg cuccuccagc aaggcucagg agcccuaccu	180
	cccugccauu auagcugcuc cccgccagaa gccugugcca acucucugca uucccugauc	240
45	uccaucccug uggcugucac ccuuggucac cuccgugcug ucacugccau cucccccc	298
50	<210> 147 <211> 279 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 147	
60	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60
	gacuucccac ugcucugccu gccucucccc acugcacugg cacagccccg ccuugccgcu	120
	gcugauccau ugccggugug acccugguac ugcaugcacg caaugcuagc ugcccuuuc	180
	ccguccuggg uaccccagau cuccccccgac cucggguccc agguaugcuc ccaccuccac	240
65	cugccccacu caccaccucu gcuaguucca gacaccucc	279

	<210> 148	
	<211> 333	
	<212> ARN	
	<213> Secuencia artificial	
5	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
	<400> 148	
10	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60
	gacuucccac ugcucugccu gccucucucc acugcacugg cacagccccg ccuugccgcu	120
	gcugauccau ugccggugug acccugacag cgugggcaac gccugccgcc ugcucugagg	180
15	cccgauccag ugggcaggcc aaggccugcu gggcccccg ggaccaggu gcucuggguc	240
	acggucccug uccccgcacc cccgcuucug ucugccccau uguggcuccu caggcucucu	300
20	ccccugcucu cccaccucua ccuccacccc cac	333
	<210> 149	
	<211> 285	
	<212> ARN	
	<213> Secuencia artificial	
25	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
	<400> 149	
30	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60
	gacuucccac ugcucugccu gccucucucc acugcacugg cacagccccg ccuugccgcu	120
35	gcugauccau ugccggugug accgagagcu cgcuuucuu cuguccaaau ucuaauaaag	180
	guuccuuugu ucccuaaguc caacuacuaa acugggggau auuaugaagg gccuugagca	240
	ucuggauucu gccuaauaaa aaacauuuau uuucuuugcu gcguc	285
40	<210> 150	
	<211> 285	
	<212> ARN	
	<213> Secuencia artificial	
45	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
	<400> 150	
50	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60
	gacuucccac ugcucugccu gccucucucc acugcacugg cacagccccg ccuugccgcu	120
	gcugauccau ugccggugug acccaagcac gcagcaaugc agcucaaaac gcuuagccua	180
55	gccacacccc cacgggaaac agcagugauu aaccuuuagc aaauaacgaa aguuaacua	240
	agcuauacua accccagggg uggucauuu cgugccagcc acacc	285
60	<210> 151	
	<211> 376	
	<212> ARN	
	<213> Secuencia artificial	
65	<220>	

	<223> Combinación de elementos de secuencia UTR-3'		
	<400> 151		
5	ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc ccaugagacu	60	
	gacuucccac ugcucugccu gccucucucc acugcacugg cacagccccg ccuugccgcu	120	
	gcugauccau ugccggugug acccuuugca ggaugaaaca cuuccccgcu uggcucucau	180	
10	ucuuccacaa gagagaccuu ucuccggacc ugguugcuac ugguucagca acucugcaga	240	
	aaauguccuc ccugugggcu gccucagcuc augccuuugg ccugaagucc cagcauugau	300	
	ggcagccccc caucuuccaa guuuugugcu cccuuuacc uaacgcuucc ugccucccau	360	
15	gcaucuguac uccucc	376	
	<210> 152		
	<211> 292		
20	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> Combinación de elementos de secuencia UTR-3'		
25	<400> 152		
	uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu ugccagcaug	60	
30	accuaccuc gccucuuuga ugccaucgc ugccaccucc uuuugcuccu ggacccuuua	120	
	gccucucugc ccuuccacuc ucugaccccu gcccguccuc accaagacug acugccugcu	180	
	gcuuugcuac ugcccgggcc caugagacug acuuccacu gcucugccug ccucucccca	240	
35	cugcacuggc acagccccgc cuugcccgug cugauccau gccgguguga cc	292	
	<210> 153		
	<211> 298		
40	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> Combinación de elementos de secuencia UTR-3'		
45	<400> 153		
	uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu ugccagcaug	60	
	accuaccuc gccucuuuga ugccaucgc ugccaccucc uuuugcuccu ggacccuuua	120	
50	gccucucugc ccuuccacuc ucugaccccu uccagccaga caccgcccc ccggcccugg	180	
	cuaagaagu gcuuccuguu gccagcauga ccuaccucg ccucuugau gccauccgcu	240	
55	gccaccuccu uuugcuccug gacccuuuag ccucucugcc cuuccacucu cugacccc	298	
	<210> 154		
	<211> 304		
60	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> Combinación de elementos de secuencia UTR-3'		
65	<400> 154		

ES 2 753 201 T3

	uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu ugccagcaug	60
	accuaccuc gccucuuuga ugccauccgc ugccaccucc uuuugcuccu ggacccuuua	120
5	gccucucugc ccuuccacuc ucugaccccc ccuuggcucc uccaggaagg cucaggagcc	180
	cuaccucccu gccauuauag cugcucucccg ccagaagccu gugccaacuc ucugcauucc	240
	cugaucucca ucccuguggc ugucacccuu ggucaccucc gugcugucac ugccaucucc	300
10	cccc	304
	<210> 155	
	<211> 285	
15	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
20	<400> 155	
	uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu ugccagcaug	60
	accuaccuc gccucuuuga ugccauccgc ugccaccucc uuuugcuccu ggacccuuua	120
25	gccucucugc ccuuccacuc ucugaccccc ugguacugca ugcacgcaau gcuagcugcc	180
	ccuuucccg ccuggguacc ccgagucucc cccgaccucg ggucccaggu augcuccac	240
30	cuccaccugc cccacucacc accucugcua guuccagaca ccucc	285
	<210> 156	
	<211> 339	
	<212> ARN	
35	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
40	<400> 156	
	uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu ugccagcaug	60
	accuaccuc gccucuuuga ugccauccgc ugccaccucc uuuugcuccu ggacccuuua	120
45	gccucucugc ccuuccacuc ucugaccccc ugacagcgug ggcaacgccu gccgccugcu	180
	cugagggcccg auccaguggg caggccaagg ccugcugggc ccccgcggaac ccaggugcuc	240
	ugggucacgg ucccuguccc cgcacccccg cuucugucug ccccauugug gcuccucagg	300
50	cucucucccc ugucuccca ccucuaccuc cccccccac	339
	<210> 157	
	<211> 291	
55	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
60	<400> 157	
65		

ES 2 753 201 T3

	uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu ugccagcaug	60
	accuaccuc gccucuuga ugccaucgc ugccaccucc uuuugcuccu ggacccuua	120
5	gccucucugc ccuuccacuc ucugacccc agagcucgcu uucugcugu ccaauucua	180
	uuaaagguuc cuuuguucc uaaguccaac uacuaaacug ggggauuua ugaaggccu	240
	ugagcaucug gauucugccu auaaaaaac auuuuuuuc auugcugcgu c	291
10	<210> 158 <211> 291 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 158	
20	uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu ugccagcaug	60
	accuaccuc gccucuuga ugccaucgc ugccaccucc uuuugcuccu ggacccuua	120
	gccucucugc ccuuccacuc ucugacccc aagcacgcag caaugcagcu caaacgcuu	180
25	agccuagcca cccccacg ggaacagca gugauuacc uuuagcaua aacgaaugu	240
	uacuaagcu auacuaacc cagguuggu caauuucug ccagccacac c	291
30	<210> 159 <211> 382 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 159	
40	uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu ugccagcaug	60
	accuaccuc gccucuuga ugccaucgc ugccaccucc uuuugcuccu ggacccuua	120
	gccucucugc ccuuccacuc ucugacccc uuugcaggau gaaacacuuc cccgcuuggc	180
45	ucucauucuu ccacaagaga gaccuuucuc cggaccuggu ugcuacuggu ucagcaacuc	240
	ugcagaaaau guccucccu guggcugccu cagcucaugc cuuuggccug aagucccagc	300
	auugauggca gcccucauc uucaaguuu ugugcuccc uuuaccuac gcuuccugcc	360
50	uccaugcau cuguacuccu cc	382
	<210> 160 <211> 298 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3'	
60	<400> 160	
65		

ES 2 753 201 T3

	gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuaua gcugcucccc	60
	gccagaagcc ugugccaacu cucugcauuc ccugaucucc aucccugugg cugucacccu	120
5	uggucaccuc cgugcuguca cugccaucuc cccccugccc guccucacca agacugacug	180
	ccugcugcuu ugcuaucucc cgggcccacg agacugacuu cccacugcuc ugccugccuc	240
	uccccacugc acuggcacag ccccgccuug ccgcugcuga uccauugccg gugugacc	298
10	<210> 161 <211> 304 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 161	
20	gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuaua gcugcucccc	60
	gccagaagcc ugugccaacu cucugcauuc ccugaucucc aucccugugg cugucacccu	120
25	uggucaccuc cgugcuguca cugccaucuc cccccuucca gccagacacc cggcccccg	180
	cccuggcuua gaaguugcuu ccuguugcca gcaugaccua cccucgccuc uuugaugcca	240
	uccgcugcca ccuccuuuug cuccuggacc cuuuagccuc ucugcccuuc cacucucuga	300
30	cccc	304
	<210> 162 <211> 310 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 162	
40	gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuaua gcugcucccc	60
	gccagaagcc ugugccaacu cucugcauuc ccugaucucc aucccugugg cugucacccu	120
45	uggucaccuc cgugcuguca cugccaucuc cccccgccuu ggcuccucca ggaaggcuca	180
	ggagcccuac cucccugcca uuauagcugc uccccgccag aagccugugc caacucucug	240
	cauucccuga ucuccauccc uguggcuguc acccuugguc accuccgugc ugucacugcc	300
50	aucucccccc	310
	<210> 163 <211> 291 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3'	
60	<400> 163	
65		

ES 2 753 201 T3

	gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuaua gcugcucucc	60
	gccagaagcc ugugccaacu cucugcauuc ccugaucucc auccugugg cugucacccu	120
5	uggucaccuc cgugcuguca cugccaucuc cccccuggu acugcaugca cgaaugcua	180
	gcugcccuu ucccguccug gguaccccg gucucccccg accucggguc ccagguaugc	240
	ucccaccucc accugcccca cucaccaccu cugcuaguuc cagacaccuc c	291
10	<210> 164 <211> 345 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 164	
20	gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuaua gcugcucucc	60
	gccagaagcc ugugccaacu cucugcauuc ccugaucucc auccugugg cugucacccu	120
25	uggucaccuc cgugcuguca cugccaucuc cccccugac agcgugggca acgccugccg	180
	ccugcucuga gggccgauc agugggcagg ccaaggccug cugggcccc gcggaccag	240
	gugcucuggg ucacgguccc ugucucccgca cccccgcuuc ugucugcccc auuguggcuc	300
30	cucaggcucu cccccugcu cucccaccuc uaccuccacc cccac	345
	<210> 165 <211> 297 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
40	<400> 165	
	gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuaua gcugcucucc	60
	gccagaagcc ugugccaacu cucugcauuc ccugaucucc auccugugg cugucacccu	120
45	uggucaccuc cgugcuguca cugccaucuc cccccgagag cugcguuucu ugcuguccaa	180
	uuucuaauaa agguuccuuu guucccuaag uccaacuacu aaacugggg auauuaugaa	240
50	gggccuugag caucuggauu cugccuaaua aaaaacauu auuuucauug cugcguc	297
	<210> 166 <211> 297 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3'	
60	<400> 166	
65		

ES 2 753 201 T3

	gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuaua gcugcucucc	60
	gccagaagcc ugugccaacu cucugcauuc ccugaucucc auccugugg cugucacccu	120
5	uggucaccuc cgugcuguca cugccaucuc cccccaagc acgcagcaau gcagcucaa	180
	acgcuuagcc uagccacacc cccacgggaa acagcaguga uuaaccuuua gcaauaacg	240
	aaaguuaac uaagcuauac uaaccccagg guuggucaau uucgugccag ccacacc	297
10	<210> 167 <211> 388 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 167	
20	gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuaua gcugcucucc	60
	gccagaagcc ugugccaacu cucugcauuc ccugaucucc auccugugg cugucacccu	120
25	uggucaccuc cgugcuguca cugccaucuc cccccuuug caggauaaa cacuuccccg	180
	cuuggcucuc auucuuccac aagagagacc uuucuccgga ccugguugcu acugguucag	240
	caacucugca gaaaauugucc ucccugugg cugccucagc ucaugccuuu ggccugaagu	300
30	cccagcauug auggcagccc cucaucuucc aaguuuugug cuccccuuua ccuaacgcu	360
	ccugccuccc augcaucugu acuccucc	388
35	<210> 168 <211> 279 <212> ARN <213> Secuencia artificial	
40	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 168	
45	cugguacugc augcacgcaa ugcuaugcugc cccuuucccg uccuggguaac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
	aguuccagac accuccugcc cguccacacc aagacugacu gccugcugcu uugcuacugc	180
50	ccgggcccau gagacugacu ucccacugcu cugccugccu cuccccacug cacuggcaca	240
	gccccgccuu gccgcugcug auccauugcc ggugugacc	279
55	<210> 169 <211> 285 <212> ARN <213> Secuencia artificial	
60	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 169	
65		

ES 2 753 201 T3

	cugguacugc augcacgcaa ugcuaagcugc cccuuucccg uccugggguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacacac caccucugcu	120
5	aguuccagac accuccuucc agccagacac ccgcccccg gccuggcua agaaguugcu	180
	uccuguugcc agcaugaccu acccucgccu cuuugaugcc auccgcugcc accuccuuu	240
	gcuccuggac ccuuuagccu cucugcccu ccacucucug acccc	285
10	<210> 170 <211> 291 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 170	
20	cugguacugc augcacgcaa ugcuaagcugc cccuuucccg uccugggguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacacac caccucugcu	120
25	aguuccagac accuccgccu uggcuccucc aggaaggcuc aggagccua ccuccugcc	180
	auuauagcug cccccgccca gaagccugug ccaacucucu gcauuccug aucuccauc	240
	cuguggcugu caccuuggu caccuccug cugucacugc caucuccccc c	291
30	<210> 171 <211> 272 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 171	
40	cugguacugc augcacgcaa ugcuaagcugc cccuuucccg uccugggguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacacac caccucugcu	120
	aguuccagac accucccugg uacugcaugc acgcaaugcu agcugcccu uucccguccu	180
45	ggguaccccg agucuccccc gaccucgggu cccagguaug cucccaccuc caccugcccc	240
	acucaccacc ucugcuaguu ccagacaccu cc	272
50	<210> 172 <211> 326 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 172	

ES 2 753 201 T3

	cugguacugc augcacgcaa ugcuaugcugc cccuuucccg uccugggguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
5	aguuccagac accucccuga cagcgugggc aacgccugcc gccugcucug aggcccgau	180
	cagugggcag gccaaaggccu gcugggcccc cgcggaccga ggugcucugg gucacggucc	240
	cuguccccgc acccccgcuu cugucugccc cauuguggcu ccucaggcuc ucuccccugc	300
10	ucucccaccu cuaccuccac cccac	326
	<210> 173	
	<211> 278	
15	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
20	<400> 173	
	cugguacugc augcacgcaa ugcuaugcugc cccuuucccg uccugggguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
25	aguuccagac accuccgaga gcucgcuuuc uugcugucca auuucuaaua aagguuccuu	180
	uguucccuua guccaacuac uaaacugggg gauauuauga agggccuuga gcaucuggau	240
30	ucugccuaau aaaaaacauu uauuuuauu gcucguc	278
	<210> 174	
	<211> 278	
	<212> ARN	
35	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
40	<400> 174	
	cugguacugc augcacgcaa ugcuaugcugc cccuuucccg uccugggguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
45	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
	ccccacggga aacagcagug auuaaccuuu agcaauaaac gaaaguuaa cuaagcuaua	240
50	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
	<210> 175	
	<211> 369	
	<212> ARN	
55	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
60	<400> 175	

ES 2 753 201 T3

	cugguacugc augcacgcaa ugcuaagcugc cccuuucccg uccugggguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
5	aguuccagac accucccuuu gcaggaugaa acacuucccc gcuuggcucu cauucuucca	180
	caagagagac cuuucuccgg accugguugc uacugguuca gcaacucugc agaaauguc	240
	cucccugug gcugccucag cucaugccuu ugccugaaag ucccagcauu gauggcagcc	300
10	ccucaucuuu caaguuuugu gcuccccuuu accuaacgcu uccugccucc caugcaucug	360
	uacuccucc	369
15	<210> 176 <211> 333 <212> ARN <213> Secuencia artificial	
20	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 176	
25	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
	gccugcuggg ccccgcgga cccaggugcu cugggucacg gucccuugucc ccgcaccccc	120
	gcuucugucu gcccacuugu ggcuccucag gcucucuccc cugcucuccc accucuaccu	180
30	ccacccccac ugcccguccu caccaagacu gacugccugc ugcuuugcua cugcccgggc	240
	ccaugagacu gacuucccac ugcucugccu gccucucccc acugcacugg cacagccccg	300
	ccuugccgcu gcugauccau ugccggugug acc	333
35	<210> 177 <211> 339 <212> ARN <213> Secuencia artificial	
40	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 177	
45	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
	gccugcuggg ccccgcgga cccaggugcu cugggucacg gucccuugucc ccgcaccccc	120
	gcuucugucu gcccacuugu ggcuccucag gcucucuccc cugcucuccc accucuaccu	180
50	ccacccccac uuccagccag acacccgccc cccggcccug gcuaagaagu ugcuuccugu	240
	ugccagcaug accuaccuc gccucuuuga ugccaucgc ugccaccucc uuuugcuccu	300
55	ggacccuuua gccucucugc ccuuccacuc ucugacccc	339
60	<210> 178 <211> 345 <212> ARN <213> Secuencia artificial	
	<220> <223> Combinación de elementos de secuencia UTR-3'	
65	<400> 178	

ES 2 753 201 T3

	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
	gccugcuggg cccccgcgga cccaggugcu cugggucacg gucccugucc ccgcaccccc	120
5	gcuucugucu gcccacauugu ggcuccucag gcucucuccc cugcucuccc accucuaccu	180
	ccacccccac gccuuggcuc cuccaggaag gcucaggagc ccuaccuccc ugccauuua	240
	gcugcucccc gccagaagcc ugugccaacu cucugcauuc ccugaucucc aucccugugg	300
10	cugucacccu uggucaccuc cgugcuguca cugccaucuc ccccc	345
	<210> 179	
	<211> 326	
15	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
20	<400> 179	
	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
25	gccugcuggg cccccgcgga cccaggugcu cugggucacg gucccugucc ccgcaccccc	120
	gcuucugucu gcccacauugu ggcuccucag gcucucuccc cugcucuccc accucuaccu	180
	ccacccccac cugguacugc augcacgcaa ugcuaugcugc cccuuucccg uccuggguac	240
30	cccagagucuc ccccagaccuc gggucccagg uaugcuccca ccuccaccug ccccacucac	300
	caccucugcu aguuccagac accucc	326
	<210> 180	
35	<211> 380	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
40	<223> Combinación de elementos de secuencia UTR-3'	
	<400> 180	
	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
45	gccugcuggg cccccgcgga cccaggugcu cugggucacg gucccugucc ccgcaccccc	120
	gcuucugucu gcccacauugu ggcuccucag gcucucuccc cugcucuccc accucuaccu	180
50	ccacccccac cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg	240
	gcaggccaag gccugcuggg cccccgcgga cccaggugcu cugggucacg gucccugucc	300
	ccgcaccccc gcuucugucu gcccacauugu ggcuccucag gcucucuccc cugcucuccc	360
55	accucuaccu ccacccccac	380
	<210> 181	
	<211> 332	
60	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
65	<400> 181	

ES 2 753 201 T3

	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
	gccugcuggg cccccgcgga cccaggugcu cugggucacg gucccugucc ccgcaccccc	120
	gcuucugucu gccccauugu ggcuccucag gcucucuccc cugcucuccc accucuaccu	180
5	ccacccccac gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc	240
	cuaaguccaa cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc	300
	uaauaaaaaa cauuaauuuu cauugcugcg uc	332
10	<210> 182 <211> 332 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 182	
20	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
	gccugcuggg cccccgcgga cccaggugcu cugggucacg gucccugucc ccgcaccccc	120
	gcuucugucu gccccauugu ggcuccucag gcucucuccc cugcucuccc accucuaccu	180
25	ccacccccac caagcacgca gcaaugcagc ucaaaacgcu uagccuagcc acacccccac	240
	gggaaacagc agugauuaac cuuugcaau aaacgaaagu uuaacuaagc uauacuaacc	300
30	ccaggguugg ucaauuucgu gccagccaca cc	332
	<210> 183 <211> 423 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 183	
40	cugacagcgu gggcaacgcc ugccgccugc ucugaggccc gauccagugg gcaggccaag	60
	gccugcuggg cccccgcgga cccaggugcu cugggucacg gucccugucc ccgcaccccc	120
45	gcuucugucu gccccauugu ggcuccucag gcucucuccc cugcucuccc accucuaccu	180
	ccacccccac cuuugcagga ugaaacacuu ccccgcuugg cucucauucu uccacaagag	240
	agaccuuucu cccgaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucucc	300
50	uguggcugcc ucagcucaug ccuuuggccu gaagucccag cauugauggc agccccucau	360
	cuuccaaguu uugugcuccc cuuuaccuaa cgcuuuccugc cucccaugca ucuguacucc	420
55	ucc	423
	<210> 184 <211> 285 <212> ARN <213> Secuencia artificial	
60	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 184	
65		

ES 2 753 201 T3

	gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc cuaaguccaa	60
	cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc uaauaaaaaa	120
5	cauuuauuuu cauugcugcg ucugcccugc cucaccaaga cugacugccu gcugcuuugc	180
	uacugcccgg gcccaugaga cugacuuccc acugcucugc cugccucucc ccacugcacu	240
	ggcacagccc cgccuugccg cugcugaucc auugccggug ugacc	285
10	<210> 185 <211> 291 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 185	
20	gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc cuaaguccaa	60
	cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc uaauaaaaaa	120
25	cauuuauuuu cauugcugcg ucuuccagcc agacacccgc ccccgccccc uggcuaagaa	180
	guugcuuccu guugccagca ugaccuaccc ucgccucuuu gaugccaucc gcugccaccu	240
	ccuuuugcuc cuggacccuu uagccucucu gcccuuccac ucucugaccc c	291
30	<210> 186 <211> 297 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 186	
40	gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc cuaaguccaa	60
	cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc uaauaaaaaa	120
	cauuuauuuu cauugcugcg ucgccuuggc uccuccagga aggcucagga gcccuaccuc	180
45	ccugccauua uagcugcucc ccgccagaag ccugugccaa cucucugcau ucccugaucu	240
	ccaucccugu ggcugucacc cuuggucacc uccgugcugu cacugccauc ucccccc	297
50	<210> 187 <211> 278 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 187	
60	gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc cuaaguccaa	60

ES 2 753 201 T3

	cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc uaauaaaaa	120
	cauuuauuuu cauugcugcg uccugguacu gcaugcacgc aaugcuagcu gccccuuucc	180
5	cguccugggg accccgaguc ucccccgacc ucggguccca gguaugcucc caccuccacc	240
	ugccccacuc accaccucug cuaguuccag acaccucc	278
10	<210> 188 <211> 332 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 188	
20	gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc cuaaguccaa	60
	cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc uaauaaaaa	120
	cauuuauuuu cauugcugcg uccugacagc gugggcaacg ccugccgccu gcucugaggc	180
25	ccgauccagu gggcaggcca aggccugcug gggccccgcg gaccagggug cucuggguca	240
	cgguccugc ccccgacccc ccgcuucugu cugccccauu guggcuccuc aggcucucuc	300
	cccugcucuc ccaccucuac cuccaccccc ac	332
30	<210> 189 <211> 284 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 189	
40	gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc cuaaguccaa	60
	cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc uaauaaaaa	120
	cauuuauuuu cauugcugcg ucgagagcuc gcuuucuugc uguccaauuu cuauuaaagg	180
45	uuccuuuguu ccuaagucc aacuacuaaa cugggggaua uuaugaaggg ccuugagcau	240
	cuggauucug ccuaauaaaa aacauuuauu uucauugcug cguc	284
50	<210> 190 <211> 284 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 190	
60	gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc cuaaguccaa	60

ES 2 753 201 T3

	cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc uaauaaaaaa	120
	cauuuauuuu cauugcugcg uccaagcacg cagcaaugca gcucaaaacg cuuagccuag	180
5	ccacaccccc acgggaaaca gcagugauua accuuuagca auaaacgaaa guuuuacuaa	240
	gcuaauacuaa ccccagggguu ggucaauuuc gugccagcca cacc	284
10	<210> 191 <211> 375 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 191	
20	gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc cuaaguccaa	60
	cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc uaauaaaaaa	120
	cauuuauuuu cauugcugcg uccuuugcag gaugaaacac uuccccgcuu ggcucucauu	180
25	cuuccacaag agagaccuuu cuccggaccu gguugcuacu gguucagcaa cucugcagaa	240
	aauguccucc ccuguggcug ccucagcuca ugccuuuggc cugaaguccc agcauugaug	300
	gcagccccuc aucuuccaag uuuugugcuc cccuuuaccu aacgcuuccu gccucccaug	360
30	caucuguacu ccucc	375
35	<210> 192 <211> 285 <212> ARN <213> Secuencia artificial	
	<220> <223> Combinación de elementos de secuencia UTR-3'	
40	<400> 192	
	caagcacgca gcaaugcagc ucaaaacgcu uagccuagcc acacccccac gggaaacagc	60
45	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccagggguug	120
	ucaauuucgu gccagccaca ccugcccugc cucaccaaga cugacugccu gcugcuuugc	180
	uacugcccgg gcccaugaga cugacuuccc acugcucugc cugccucucc ccacugcacu	240
50	ggcacagccc cgccuugccg cugcugaucc auugccggug ugacc	285
55	<210> 193 <211> 291 <212> ARN <213> Secuencia artificial	
	<220> <223> Combinación de elementos de secuencia UTR-3'	
60	<400> 193	
	caagcacgca gcaaugcagc ucaaaacgcu uagccuagcc acacccccac gggaaacagc	60

ES 2 753 201 T3

	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccagggguugg	120
	ucaauuucgu gccagccaca ccuuccagcc agacacccgc ccccgggccc uggcuaagaa	180
5	guugcuuccu guugccagca ugaccuaccc ucgccucuuu gaugccaacc gcugccaccu	240
	ccuuuugcuc cuggacccuu uagccucucu gcccuuccac ucucugaccc c	291
10	<210> 194 <211> 297 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 194	
20	caagcacgca gcaaugcagc uaaaaacgcu uagccuagcc acacccccac gggaaacagc	60
	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccagggguugg	120
	ucaauuucgu gccagccaca ccgccuuggc uccuccagga aggcucagga gcccuaccuc	180
25	ccugccauua uagcugcucc ccgccagaag ccugugccaa cucucugcau ucccugaucu	240
	ccaucuccgu ggcugucacc cuugrucacc uccgugcugu cacugccauc uccccc	297
30	<210> 195 <211> 278 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 195	
40	caagcacgca gcaaugcagc uaaaaacgcu uagccuagcc acacccccac gggaaacagc	60
	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccagggguugg	120
	ucaauuucgu gccagccaca ccugguacu gcaugcacgc aaugcuagcu gcccuuucc	180
45	cguccugggu accccgaguc ucccccagacc ucggguccca gguaugcucc caccuccacc	240
	ugccccacuc accaccucug cuaguuccag acaccucc	278
50	<210> 196 <211> 332 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 196	
60	caagcacgca gcaaugcagc uaaaaacgcu uagccuagcc acacccccac gggaaacagc	60
	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccagggguugg	120
65		

ES 2 753 201 T3

	ucaauuucgu gccagccaca ccugacagc gugggcaacg ccugccgccu gcucugaggg	180
	ccgauccagu gggcaggcca aggccugcug ggcccccgcg gaccaggug cucuggguca	240
5	cgguccugc ccccgacccc ccgcuucugu cugcccauu guggcuccuc aggcucucuc	300
	ccugcucuc ccaccucua cuccaccccc ac	332
10	<210> 197 <211> 284 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3' <400> 197	
20	caagcacgca gcaaugcagc uaaaaacgcu uagccuagcc acacccccac gggaaacagc	60
	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccaggguugg	120
	ucaauuucgu gccagccaca ccgagagcuc gcuuucugc uguccaauuu cuauuaaagg	180
25	uuccuuuguu ccuaagucc aacuacuaaa cugggggaua uuaugaaggg ccuugagcau	240
	cuggauucug ccuaauaaaa aacauuuauu uucauugcug cguc	284
30	<210> 198 <211> 284 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3' <400> 198	
40	caagcacgca gcaaugcagc uaaaaacgcu uagccuagcc acacccccac gggaaacagc	60
	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccaggguugg	120
	ucaauuucgu gccagccaca cccaagcacg cagcaaugca gcuaaaaacg cuuagccuag	180
45	ccacaccccc acgggaaaca gcagugauua accuuuagca auaaacgaaa guuuacuaa	240
	gcuaucuaa cccagggguu gguaauuuc gugccagcca cacc	284
50	<210> 199 <211> 375 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3' <400> 199	
	caagcacgca gcaaugcagc uaaaaacgcu uagccuagcc acacccccac gggaaacagc	60
	agugauuaac cuuuagcaau aaacgaaagu uuaacuaagc uauacuaacc ccaggguugg	120
60	ucaauuucgu gccagccaca cccuuugcag gaugaaacac uuccccgcuu ggcucucauu	180
65		

ES 2 753 201 T3

	cuuccacaag agagaccuuu cuccggaccu gguugcuacu gguucagcaa cucugcagaa	240
	aauguccucc ccuguggcug ccucagcuca ugccuuuggc cugaaguccc agcauugaug	300
5	gcagcccuc aucuuccaag uuuugugcuc cccuuuaccu aacgcuuccu gccucccaug	360
	caucuguacu ccucc	375
10	<210> 200 <211> 376 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 200	
20	cuuugcagga ugaacacuu ccccgcuugg cucucauucu uccacaagag agaccuuucu	60
	ccggaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucccc uguggcugcc	120
	ucagcucaug ccuuuggccu gaagucccag cauugauggc agcccucacu cuuccaaguu	180
25	uugugcuccc cuuuaccuaa cgcuccugc cucccaugca ucuguacucc uccugcccgu	240
	ccucaccaag acugacugcc ugcugcuuug cuacugcccg ggcccagag acugacuucc	300
	cacugcucug ccugccucuc cccacugcac uggcacagcc ccgccuugcc gcugcugauc	360
30	cauugccggu gugacc	376
35	<210> 201 <211> 382 <212> ARN <213> Secuencia artificial	
	<220> <223> Combinación de elementos de secuencia UTR-3'	
40	<400> 201	
	cuuugcagga ugaacacuu ccccgcuugg cucucauucu uccacaagag agaccuuucu	60
	ccggaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucccc uguggcugcc	120
45	ucagcucaug ccuuuggccu gaagucccag cauugauggc agcccucacu cuuccaaguu	180
	uugugcuccc cuuuaccuaa cgcuccugc cucccaugca ucuguacucc uccuuccagc	240
50	cagacacccg ccccccggcc cuggcuaaga aguugcuucc uguugccagc augaccuacc	300
	cucgccucuu ugaugccauc cgcugccacc uccuuuugcu ccuggacccu uuagccucuc	360
	ugcccuucca cucucugacc cc	382
55	<210> 202 <211> 388 <212> ARN <213> Secuencia artificial	
60	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 202	
65		

ES 2 753 201 T3

	cuuugcagga ugaaacacuu ccccgcuugg cucucauucu uccacaagag agaccuuucu	60
	ccggaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucucc uguggcugcc	120
5	ucagcucaug ccuuuggccu gaagucccag cauugauggc agccccucau cuuccaaguu	180
	uugugcuccc cuuuaccuaa cgcuccucgc cuccaugca ucuguacucc uccgccuugg	240
	cuccuccagg aaggcucagg agcccuaccu cccugccauu auagcugcuc cccgccagaa	300
10	gccugugcca acucucugca uucccugauc uccaucuccug uggcugucac ccuuggucac	360
	cuccgugcug ucacugccau ccccccc	388
15	<210> 203 <211> 369 <212> ARN <213> Secuencia artificial	
20	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 203	
25	cuuugcagga ugaaacacuu ccccgcuugg cucucauucu uccacaagag agaccuuucu	60
	ccggaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucucc uguggcugcc	120
	ucagcucaug ccuuuggccu gaagucccag cauugauggc agccccucau cuuccaaguu	180
30	uugugcuccc cuuuaccuaa cgcuccucgc cuccaugca ucuguacucc ucccugguac	240
	ugcaugcacg caaugcuagc ugcccccuuu cgcuccuggg uaccccgagu ccccccgac	300
	cucgggucucc aggaugcuc ccaccuccac cugcccccacu caccaccucu gcuaaguucca	360
35	gacaccucc	369
40	<210> 204 <211> 423 <212> ARN <213> Secuencia artificial	
	<220> <223> Combinación de elementos de secuencia UTR-3'	
45	<400> 204	
	cuuugcagga ugaaacacuu ccccgcuugg cucucauucu uccacaagag agaccuuucu	60
	ccggaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucucc uguggcugcc	120
50	ucagcucaug ccuuuggccu gaagucccag cauugauggc agccccucau cuuccaaguu	180
	uugugcuccc cuuuaccuaa cgcuccucgc cuccaugca ucuguacucc ucccugacag	240
55	cuggggcaac gccugccgcc ugcucugagg cccgauccag ugaggcaggcc aaggccugcu	300
	gggccccgc ggaccaggu gcucuggguc acggucccug uccccgcacc cccgcuucug	360
	ucugccccaau uguggcuccu caggcucucu cccugcucu cccaccucua ccuccacccc	420
60	cac	423
65	<210> 205 <211> 375 <212> ARN <213> Secuencia artificial	

<220>	
<223> Combinación de elementos de secuencia UTR-3'	
<400> 205	
5	cuuugcagga ugaaacacuu ccccgcuugg cucucauucu uccacaagag agaccuuucu 60
	ccggaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucucc uguggcugcc 120
	ucagcucaug ccuuuggccu gaagucccag cauugauggc agccccucau cuuccaaguu 180
10	uugugcuccc cuuuaccuaa cgcuuccugc cucccaugca ucuguacucc uccgagagcu 240
	cgcuuucuug cuguccaaau ucuauuaaag guuccuuugu ucccuaaguc caacuacuaa 300
	acugggggau auuaugaagg gccuugagca ucuggauucu gccuaaaaa aaacauuuau 360
15	uuucauugcu gcguc 375
<210> 206	
<211> 375	
<212> ARN	
<213> Secuencia artificial	
<220>	
<223> Combinación de elementos de secuencia UTR-3'	
<400> 206	
25	cuuugcagga ugaaacacuu ccccgcuugg cucucauucu uccacaagag agaccuuucu 60
30	ccggaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucucc uguggcugcc 120
	ucagcucaug ccuuuggccu gaagucccag cauugauggc agccccucau cuuccaaguu 180
	uugugcuccc cuuuaccuaa cgcuuccugc cucccaugca ucuguacucc ucccaagcac 240
35	gcagcaaugc agcucaaaac gcuuagccua gccacacccc cacgggaaac agcagugauu 300
	aaccuuuagc aaauaacgaa aguuuuacua agcuauacua acccaggggu uggucaauuu 360
	cgugccagcc acacc 375
<210> 207	
<211> 466	
<212> ARN	
<213> Secuencia artificial	
<220>	
<223> Combinación de elementos de secuencia UTR-3'	
<400> 207	
50	cuuugcagga ugaaacacuu ccccgcuugg cucucauucu uccacaagag agaccuuucu 60
	ccggaccugg uugcuacugg uucagcaacu cugcagaaaa uguccucucc uguggcugcc 120
55	ucagcucaug ccuuuggccu gaagucccag cauugauggc agccccucau cuuccaaguu 180
	uugugcuccc cuuuaccuaa cgcuuccugc cucccaugca ucuguacucc ucccuuugca 240
	ggaugaaaca cuuccccgcu uggcucucau ucuuccacaa gagagaccuu ucuccggacc 300
60	ugguugcuac ugguucagca acucugcaga aaauguccuc cccuguggcu gccucagcuc 360
	augccuuugg ccugaagucc cagcauugau ggcagccccc caucuuccaa guuuugugcu 420
	ccccuuuacc uaacgcuucc ugccucccau gcaucuguac uccucc 466
65	

	<210> 208	
	<211> 278	
	<212> ARN	
	<213> Secuencia artificial	
5	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
	<400> 208	
10	cugguacugc auggacgcaa ugcuaagcugc cccuuucccg uccuggguac cccgagucac	60
	ccccgaccuc gggucccagg uaucguccca ccuccaccug cccacucac caccucugcu	120
15	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc auagccacac	180
	ccccacggga aacagcagug auuaaccuuu agcaauaau gaauguuuaa cuaagcuaua	240
	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
20	<210> 209	
	<211> 278	
	<212> ARN	
	<213> Secuencia artificial	
25	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
	<400> 209	
30	cucguacugc auggacgcaa ugcuaagcugc cccuuucccg uccuggguac cccgagucac	60
	caccgaccuc gggucccagg uaucguccca ccuccaccug cccacucac caccucugcu	120
	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc auagccacac	180
35	ccccacggga aacaguagug auuaaccuuu agcaauaau gaaugucuaa cuaagcuaua	240
	cuaaccccag gguugaucaa uuacgugcca gccacacc	278
40	<210> 210	
	<211> 278	
	<212> ARN	
	<213> Secuencia artificial	
45	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
	<400> 210	
50	cucguacugc auggacgcaa ugcuaagcugc cccuuucccg uccuggguac cccgagucac	60
	caccgaccuc gggucccagg uaucguccca ccuccaccug cccacucac caccuuugcu	120
	aguuccagac accucccaug cacgcagcaa ugcagaucaa aacgcuuagc auagccacac	180
55	ccccacggga aacaguagug auuaaccuuu agcaauaau gaaugucuaa cuaagcuaua	240
	cuaaccccag gguugaucaa uuacgugcca gccacacc	278
60	<210> 211	
	<211> 278	
	<212> ARN	
	<213> Secuencia artificial	
65	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	

ES 2 753 201 T3

	<400> 211	
	gucguacugc auggacgcaa ugcuagcagc accuuucccg uccugggguac cccgagucac	60
5	caccgaccuc gggucccagg uaucguccca ccuccacgug cccacccac caccuuugcu	120
	aguuccagag accucccaug cacgcagcaa ugcagaucaa aacgcuuagc auagccacac	180
	cgccacggga aacaguagug aucaaccuuu agcuauaauc gaaugucuaa cuaagcuauu	240
10	cuaaccacag gguugaucaa uuacgugcca gccagacc	278
	<210> 212	
	<211> 278	
15	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
20	<400> 212	
	cugguacugc augcacgcaa ugcuagcugc cccaaagggc uccugggguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
25	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
	ccccacggga aacagcagug auuaaccuuu agcaauaaac gaaaguuaa cuaagcuaua	240
30	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
	<210> 213	
	<211> 278	
	<212> ARN	
35	<213> Secuencia artificial	
	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
40	<400> 213	
	cugguacugc augcacgcaa ugcuagcugc cccuuucccg ugguccguac cccgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
45	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
	ccccacggga aacagcagug auuaaccuuu agcaauaaac gaaaguuaa cuaagcuaua	240
50	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
	<210> 214	
	<211> 278	
	<212> ARN	
	<213> Secuencia artificial	
55	<220>	
	<223> Combinación de elementos de secuencia UTR-3'	
60	<400> 214	

ES 2 753 201 T3

	cugguacugc augcacgcaa ugcuaagcugc cccuuucccg uggaccguac ggcgagucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
5	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
	ccccacggga aacagcagug auuaaccuuu agcaauaaac gaaaguuuaa cuaagcuaua	240
	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
10	<210> 215 <211> 278 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 215	
20	cugguacugc augcacgcaa ugcuaagcugc cccuuugccg uggaccguac ggcgugucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
25	ccccacggga aacagcagug auuaaccuuu agcaauaaac gaaaguuuaa cuaagcuaua	240
	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
30	<210> 216 <211> 278 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 216	
40	cugguacugc augcacgcaa ugcuaagcugc cccuuugggc uggaccguac ggcgugucuc	60
	ccccgaccuc gggucccagg uaugcuccca ccuccaccug cccacucac caccucugcu	120
	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
45	ccccacggga aacagcagug auuaaccuuu agcaauaaac gaaaguuuaa cuaagcuaua	240
	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
50	<210> 217 <211> 278 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3'	
	<400> 217	

ES 2 753 201 T3

	cugguacugc augcacgcaa ugcuagcugc cccuuucccg ugguccguac cccgagucuc	60
	ccccgaccuc gggucggacc uaugcuccca ccuccaccug cccacucac caccucugcu	120
5	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
	ccccacggga aacagcagug auuaaccuuu agcaauaaac gaaaguuuaa cuaagcuaua	240
	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
10	<210> 218 <211> 278 <212> ARN <213> Secuencia artificial	
15	<220> <223> Combinación de elementos de secuencia UTR-3' <400> 218	
20	cugguacugc augcacgcaa ugcuagcugc cccuuucccg uggaccguac ggcgagucuc	60
	ccccgaccuc gccucggucc uaugcuccca ccuccaccug cccacucac caccucugcu	120
	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
25	ccccacggga aacagcagug auuaaccuuu agcaauaaac gaaaguuuaa cuaagcuaua	240
	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
30	<210> 219 <211> 278 <212> ARN <213> Secuencia artificial	
35	<220> <223> Combinación de elementos de secuencia UTR-3' <400> 219	
40	cugguacugc augcacgcaa ugcuagcugc cccuuugccg uggaccguac gggcugucuc	60
	ccccgaccag ccucggucc uaugcuccca ccuccaccug cccacucac caccucugcu	120
	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
45	ccccacggca aacagcagug auuaaccuuu agcaauaaac gaaaguuuaa cuaagcuaua	240
	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
50	<210> 220 <211> 278 <212> ARN <213> Secuencia artificial	
55	<220> <223> Combinación de elementos de secuencia UTR-3' <400> 220	

ES 2 753 201 T3

	cugguacugc augcacgcaa ugcua gcugc cccuuugggc uggaccguac gggcugucuc	60
	ccccgaccag cccucggucc uaugcuccca ccuccaccug cccacucac caccucugcu	120
5	aguuccagac accucccaag cacgcagcaa ugcagcucaa aacgcuuagc cuagccacac	180
	ccccagccca aacagcagug auuaaccuuu agcaauaaac gaaaguuaa cuaagcuaua	240
	cuaaccccag gguuggucaa uuucgugcca gccacacc	278
10	<210> 221 <211> 142 <212> ARN <213> Secuencia artificial	
15	<220> <223> Elemento de secuencia UTR-3'	
	<400> 221	
20	gagagcucgc uuucuugcug uccaauuucu auuaaagguu ccuuuguucc cuaaguccaa	60
	cuacuaaacu gggggauauu augaagggcc uugagcaucu ggauucugcc uaauaaaaaa	120
25	cauuuuuuuu cauugcugcg uc	142

REIVINDICACIONES

1. Una molécula de ácido nucleico que comprende en la dirección de transcripción 5' → 3':
 - (a) un promotor;
 - (b) una secuencia de ácido nucleico transcribible o una secuencia de ácido nucleico para introducir una secuencia de ácido nucleico transcribible; y
 - (c) una secuencia de ácido nucleico que, cuando se transcribe bajo el control del promotor (a), codifica para una región no traducida en 3' en el transcrito que no está naturalmente unida a la secuencia de ácido nucleico (b), dicha región no traducida en 3' comprende
 - (i) (c-4) la secuencia de ácido nucleico de la región no traducida en 3' del Potenciador de la Separación del Amino Terminal (AES) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 86 a 89, preferentemente la SEQ ID NO: 86, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico; o
 - (ii) (c-8) cualquier combinación de (c-4) y una o más secuencias de ácido nucleico seleccionadas del grupo que consiste en:
 - (c-1) la secuencia de ácido nucleico de la región no traducida en 3' del Fragmento Fc de IgG, Receptor, Transportador, Alfa (FCGRT) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 1 a 50, preferentemente la SEQ ID NO: 27, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-2) la secuencia de ácido nucleico de la región no traducida en 3' de la Proteína Específica de Linfocitos 1 (LSP1) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 51 a 72, preferentemente la SEQ ID NO: 52, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-3) la secuencia de ácido nucleico de la región no traducida en 3' del Ligando de Quimiocina 22 (CCL22) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 73 a 85, preferentemente la SEQ ID NO: 79, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-4) la secuencia de ácido nucleico de la región no traducida en 3' del Potenciador de la Separación del Amino Terminal (AES) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 86 a 89, preferentemente la SEQ ID NO: 86, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-5) la secuencia de ácido nucleico de la región no traducida en 3' del Miembro 3 de la Familia de fosfolipasa D (PLD3) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 90 a 104, preferentemente la SEQ ID NO: 96, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-6) la secuencia de ácido nucleico del ARN no codificante del ARN 12S codificado mitocondrialmente (MT-RNR1) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 105 a 121, preferentemente la SEQ ID NO: 115, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico, y
 - (c-7) la secuencia nucleica de la región no traducida en 3' del Complejo Principal de Histocompatibilidad Clase II DR Beta 4 (HLA-DRB4) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 122 a 143, preferentemente la SEQ ID NO: 126, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,

en donde las secuencias de ácido nucleico (b) y (c) bajo el control del promotor (a) pueden transcribirse para proporcionar un transcrito común en el que la secuencia de ácido nucleico transcrita a partir de la secuencia de ácido nucleico (c) está activa para aumentar la eficiencia de la traducción y/o la estabilidad de la secuencia de ácido nucleico transcrita a partir de la secuencia de ácido nucleico transcribible (b).
2. La molécula de ácido nucleico como se reivindicó en la reivindicación 1, en donde la secuencia de ácido nucleico (c-8) comprende una combinación de (c-4) y (c-6), en donde, preferentemente, (c-4) se ubica 5' a (c-6), y/o la combinación de (c-4) y (c-6) comprende la secuencia de ácido nucleico de la SEQ ID NO: 174, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico.
3. La molécula de ácido nucleico como se reivindicó en la reivindicación 1 o 2, que comprende además (d) una secuencia de ácido nucleico que, cuando se transcribe bajo el control del promotor (a), codifica para una secuencia de ácido nucleico que es una secuencia de poliadenilo, en donde, opcionalmente, la secuencia de poliadenilo comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A.

4. La molécula de ácido nucleico como se reivindicó en cualquiera de las reivindicaciones 1 a 3, que es una molécula circular cerrada o una molécula lineal.
 5. La molécula de ácido nucleico como se reivindicó en cualquiera de las reivindicaciones 1 a 4, en donde la secuencia de ácido nucleico transcribible comprende una secuencia de ácido nucleico codificante para un péptido o proteína y la secuencia de ácido nucleico para introducir una secuencia de ácido nucleico transcribible es un sitio de clonación múltiple.
 6. La molécula de ácido nucleico como se reivindicó en cualquiera de las reivindicaciones 1 a 5, que comprende además uno o más miembros seleccionados del grupo que consiste en: (i) un gen reportero; (ii) un marcador de selección; y (iii) un origen de replicación.
 7. La molécula de ácido nucleico como se reivindicó en cualquiera de las reivindicaciones 1 a 6, que es adecuada, en particular después de la linealización, para la transcripción in vitro del ARN, en particular del ARNm.
 8. El ARN que puede obtenerse mediante la transcripción, preferentemente, la transcripción in vitro, mediante el uso de una molécula de ácido nucleico como se reivindicó en cualquiera de las reivindicaciones 1 a 7 como un molde.
 9. El ARN que comprende en la dirección 5' → 3':
 - (a) una región no traducida en 5';
 - (b) una secuencia de ácido nucleico codificante para un péptido o proteína; y
 - (c) una región no traducida en 3' que no está unida naturalmente a la secuencia de ácido nucleico (b), dicha región no traducida en 3' comprende
 - (i) (c-4) la secuencia de ácido nucleico de la región no traducida en 3' del Potenciador de la Separación del Amino Terminal (AES) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 86 a 89, preferentemente la SEQ ID NO: 86, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico; o
 - (ii) (c-8) cualquier combinación de (c-4) y una o más secuencias de ácido nucleico seleccionadas del grupo que consiste en:
 - (c-1) la secuencia de ácido nucleico de la región no traducida en 3' del Fragmento Fc de IgG, Receptor, Transportador, Alfa (FCGRT) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 1 a 50, preferentemente la SEQ ID NO: 27, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-2) la secuencia de ácido nucleico de la región no traducida en 3' de la Proteína Específica de Linfocitos 1 (LSP1) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 51 a 72, preferentemente la SEQ ID NO: 52, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-3) la secuencia de ácido nucleico de la región no traducida en 3' del Ligando de Quimiocina 22 (CCL22) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 73 a 85, preferentemente la SEQ ID NO: 79, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-4) la secuencia de ácido nucleico de la región no traducida en 3' del Potenciador de la Separación del Amino Terminal (AES) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 86 a 89, preferentemente la SEQ ID NO: 86, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-5) la secuencia de ácido nucleico de la región no traducida en 3' del Miembro 3 de la Familia de Fosfolipasa D (PLD3) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 90 a 104, preferentemente la SEQ ID NO: 96, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-6) la secuencia de ácido nucleico del ARN no codificante del ARN 12S codificado mitocondrialmente (MT-RNR1) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 105 a 121, preferentemente la SEQ ID NO: 115, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (c-7) la secuencia nucleica de la región no traducida en 3' del Complejo Principal de Histocompatibilidad Clase II DR Beta 4 (HLA-DRB4) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 122 a 143, preferentemente la SEQ ID NO: 126, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
- en donde la secuencia de ácido nucleico (c) es activa para aumentar la eficiencia de la traducción y/o la estabilidad de la secuencia de ácido nucleico codificante para un péptido o proteína.

10. El ARN como se reivindicó en la reivindicación 9, que comprende además (d) una secuencia de ácido nucleico que es una secuencia de poliadenilo que opcionalmente comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A, en donde, preferentemente, dicha secuencia de ácido nucleico (d) está ubicada en el extremo 3' de dicho ARN, y/o las secuencias de ácido nucleico (c) y (d) están activas para aumentar la eficiencia de la traducción y/o la estabilidad de la secuencia de ácido nucleico codificante para un péptido o proteína.
11. El ARN como se reivindicó en la reivindicación 9 o 10, que además comprende (e) un casquete 5'.
12. Un método para obtener ARN, que comprende:
 - (i) proporcionar una molécula de ácido nucleico como se reivindicó en cualquiera de las reivindicaciones 1 a 7, y
 - (ii) transcribir el ARN mediante el uso de la molécula de ácido nucleico como un molde.
13. Un método para obtener un péptido o proteína, que comprende:
 - (i) obtener el ARN codificante para el péptido o proteína de acuerdo con el método de conformidad con la reivindicación 12, y
 - (ii) traducir el ARN.
14. El método como se reivindicó en la reivindicación 12 o 13, **caracterizado porque** comprende, además, antes de la transcripción de la molécula de ácido nucleico, la escisión de la molécula de ácido nucleico.
15. Un método para obtener ARN, que comprende:
 - (i) acoplar una secuencia de ácido nucleico (b) que, cuando se transcribe, codifica para una región no traducida en 3', en el extremo 3' de una secuencia de ácido nucleico transcribible (a) que comprende una secuencia de ácido nucleico codificante para un péptido o proteína, en donde la secuencia de ácido nucleico (a) no está unida naturalmente a la secuencia de ácido nucleico (b), y
 - (ii) transcribir el ácido nucleico obtenido, dicha región no traducida en 3' comprende
 - (iia) (b-4) la secuencia de ácido nucleico de la región no traducida en 3' del Potenciador de la Separación del Amino Terminal (AES) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 86 a 89, preferentemente la SEQ ID NO: 86, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico; o
 - (iib) (b-8) cualquier combinación de (b-4) y una o más secuencias de ácido nucleico seleccionadas del grupo que consiste en:
 - (b-1) la secuencia de ácido nucleico de la región no traducida en 3' del Fragmento Fc de IgG, Receptor, Transportador, Alfa (FCGRT) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 1 a 50, preferentemente la SEQ ID NO: 27, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (b-2) la secuencia de ácido nucleico de la región no traducida en 3' de la Proteína Específica de Linfocitos 1 (LSP1) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 51 a 72, preferentemente la SEQ ID NO: 52, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (b-3) la secuencia de ácido nucleico de la región no traducida en 3' del Ligando de Quimiocina 22 (CCL22) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NO: 73 a 85, preferentemente la SEQ ID NO: 79, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (b-4) la secuencia de ácido nucleico de la región no traducida en 3' del Potenciador de la Separación del Amino Terminal (AES) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NO: 86 a 89, preferentemente la SEQ ID NO: 86, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (b-5) la secuencia de ácido nucleico de la región no traducida en 3' del Miembro 3 de la Familia de Fosfolipasa D (PLD3) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NO: 90 a 104, preferentemente la SEQ ID NO: 96, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (b-6) la secuencia de ácido nucleico del ARN no codificante del ARN 12S codificado mitocondrialmente (MT-RNR1) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NOs: 105 a 121, preferentemente la SEQ ID NO: 115, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,
 - (b-7) la secuencia nucleica de la región no traducida en 3' del Complejo Principal de Histocompatibilidad Clase II DR Beta 4 (HLA-DRB4) que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEC ID NO: 122 a 143, preferentemente la SEQ ID NO: 126, un fragmento de esta, o una variante de dicha secuencia o fragmento de ácido nucleico que es al menos 90 % idéntica a dicha secuencia o fragmento de ácido nucleico,

en donde las secuencias de ácido nucleico (a) y (b) pueden transcribirse para proporcionar un transcrito común en el que la secuencia de ácido nucleico transcrita a partir de la secuencia de ácido nucleico (b) está activa para aumentar la eficiencia de la traducción y/o la estabilidad de la secuencia de ácido nucleico transcrita a partir de la secuencia de ácido nucleico transcribible (a).

- 5 16. El método como se reivindicó en la reivindicación 15, que comprende además el acoplamiento de una secuencia de ácido nucleico (c) que, cuando se transcribe, codifica para una secuencia de ácido nucleico que es una secuencia de poliadenilo que opcionalmente comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contiene nucleótidos distintos de los nucleótidos A, en el extremo 3' de la
- 10 secuencia de ácido nucleico (b), en donde, preferentemente, las secuencias de ácido nucleico (a), (b), y (c) pueden transcribirse para proporcionar un transcrito común en el que las secuencias de ácido nucleico transcritas a partir de las secuencias de ácido nucleico (b) y (c) están activas para aumentar la eficiencia de la traducción y/o la estabilidad de la secuencia de ácido nucleico transcrita a partir de la secuencia de ácido nucleico transcribible (a).
- 15 17. Un método para obtener un péptido o proteína, que comprende:
(i) obtener ARN mediante el método de conformidad con la reivindicación 15 o 16, y
(ii) traducir el ARN.
- 20 18. El método como se reivindicó en cualquiera de las reivindicaciones 12 a 17, en donde la transcripción se realiza in vitro.
19. El ARN que se obtiene mediante el método como se reivindicó en cualquiera de las reivindicaciones 12, 14 a 16 y 18.
- 25 20. Un método para obtener un péptido o proteína, que comprende traducir el ARN como se reivindicó en cualquiera de las reivindicaciones 8 a 11 y 19.
21. El ARN como se reivindicó en cualquiera de las reivindicaciones 8 a 11 y 19 para usar en un método de transfección de una célula huésped, en donde, preferentemente, la célula huésped es una célula presentadora de antígeno, en particular una célula dendrítica, un monocito o un macrófago.
- 30 22. El ARN como se reivindicó en cualquiera de las reivindicaciones 8 a 11 y 19 para usar en un método de vacunación.
23. El ARN como se reivindicó en cualquiera de las reivindicaciones 8 a 11 y 19 para usar en un método de reprogramación de células somáticas en células que tienen características de células madre.
- 35

Figura 1

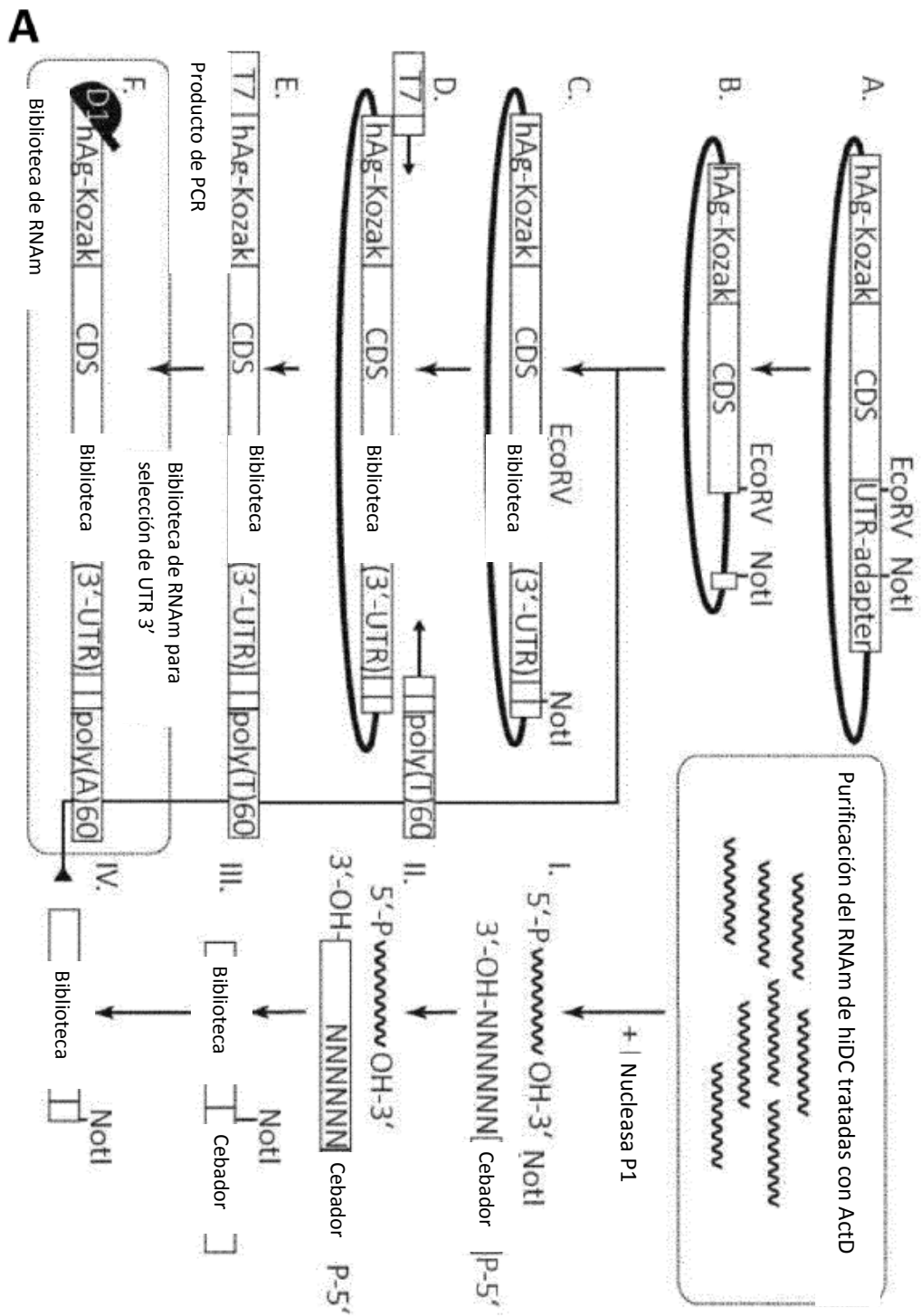


Figura 1

B

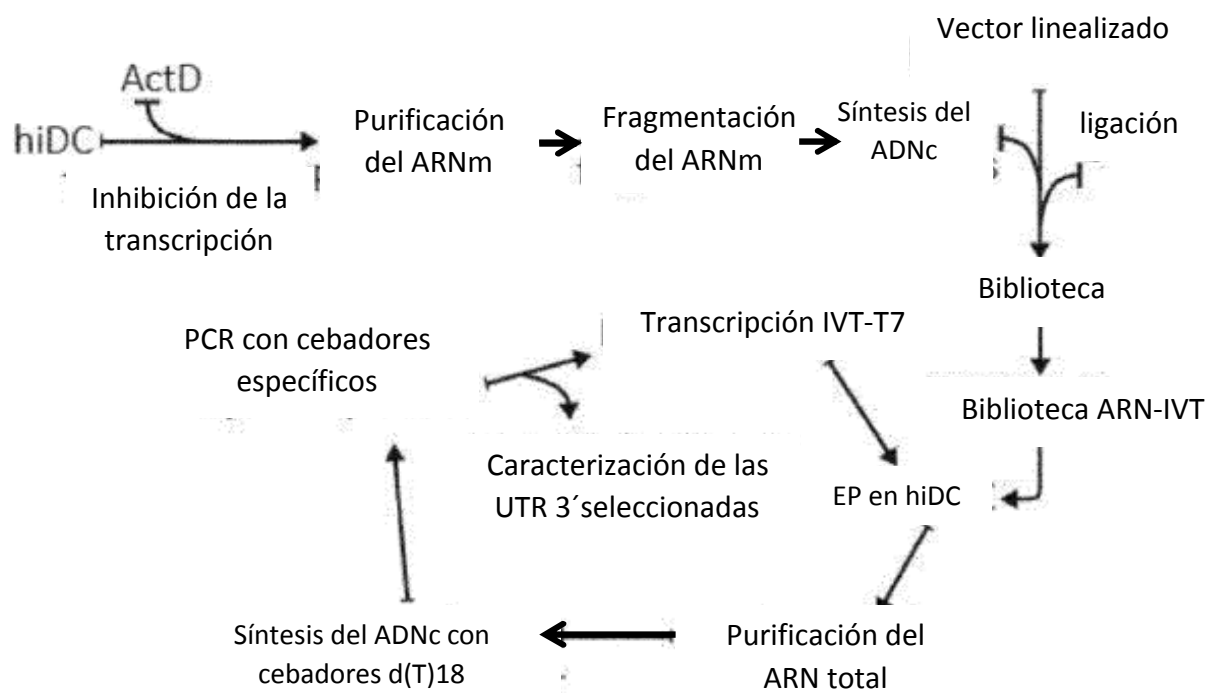


Figura 2

A

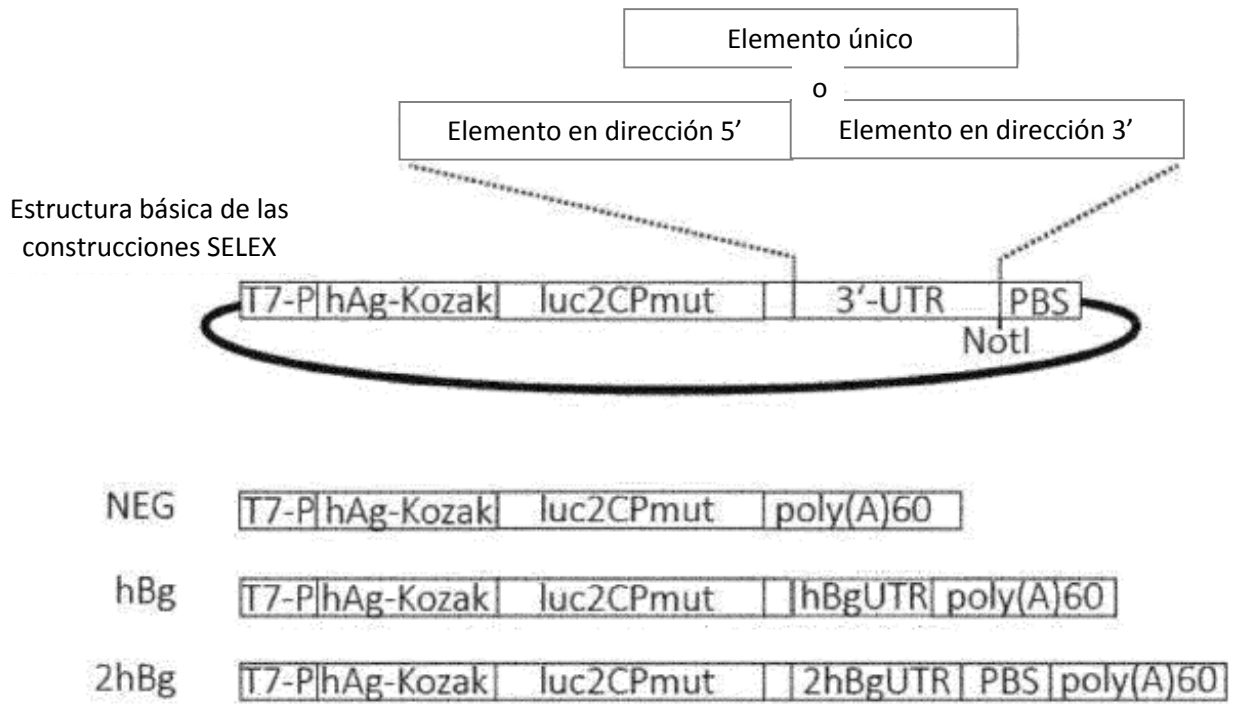


Figura 2

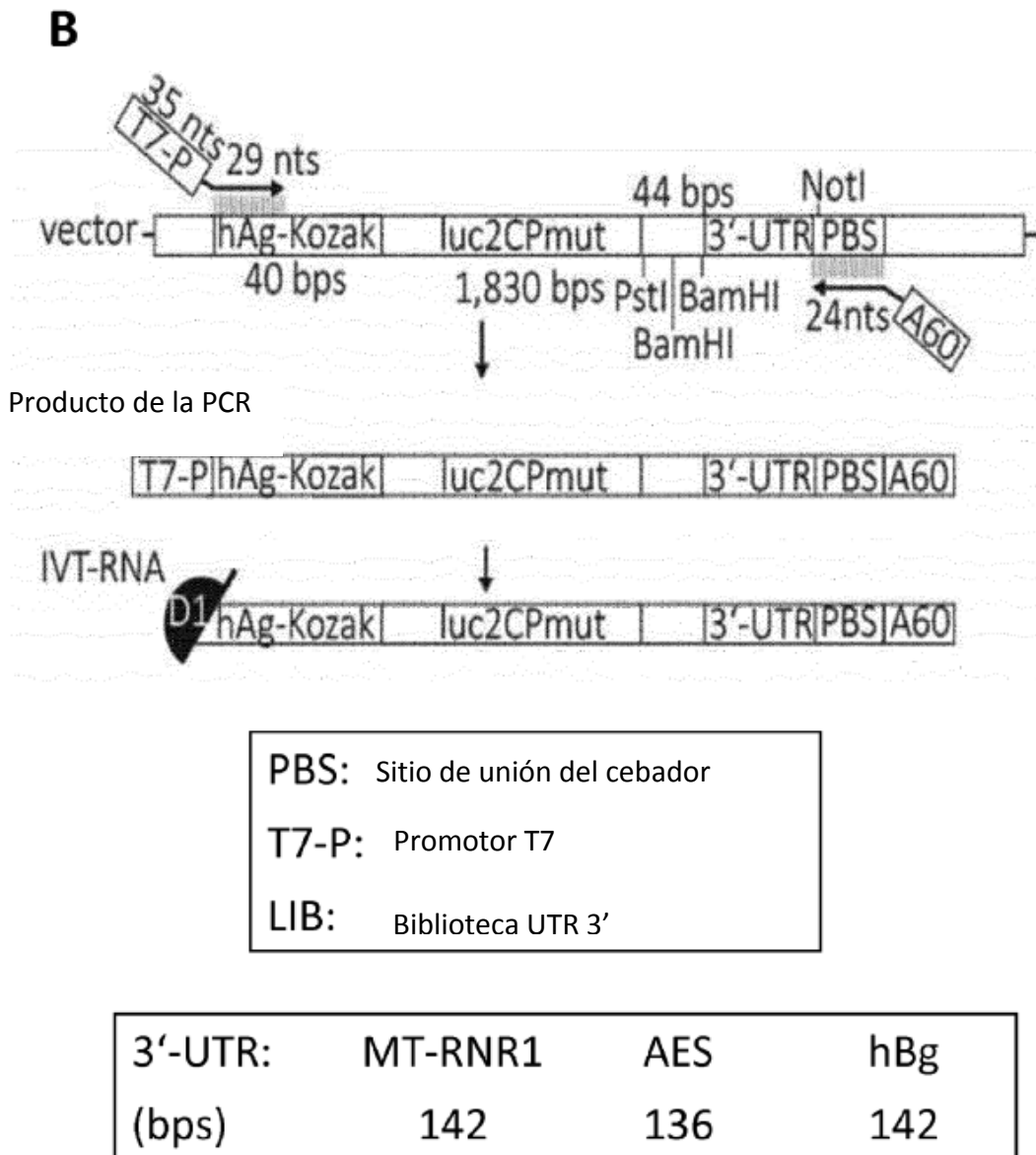


Figura 3

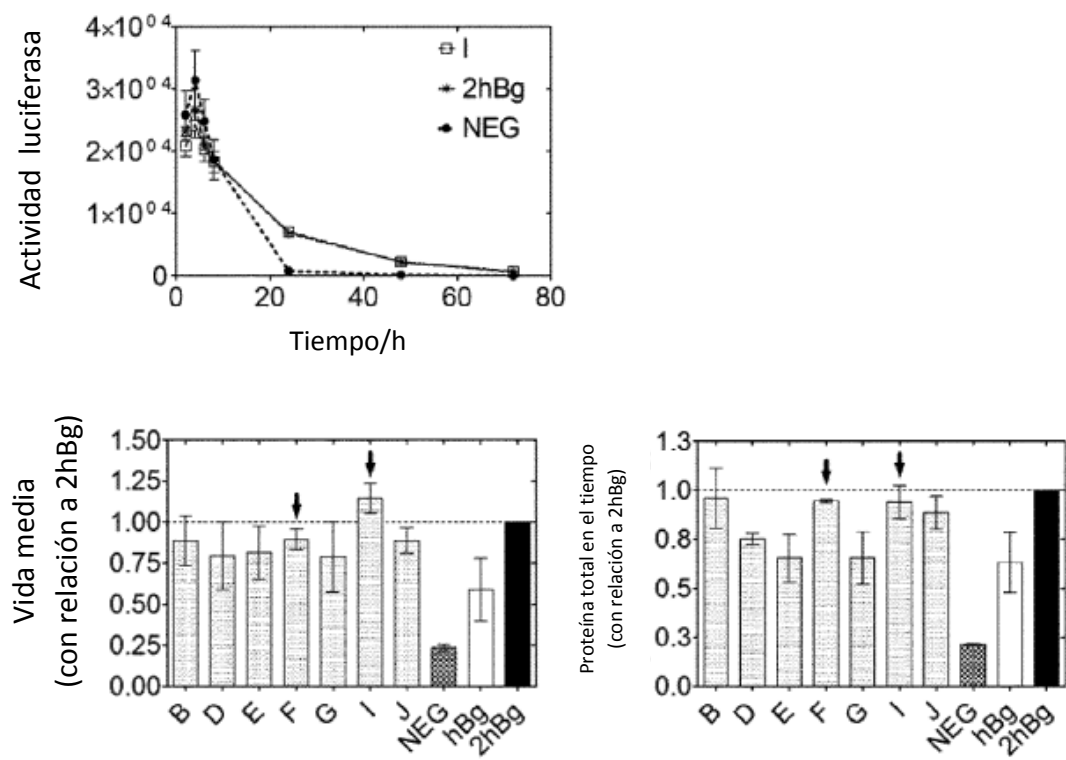


Figura 4

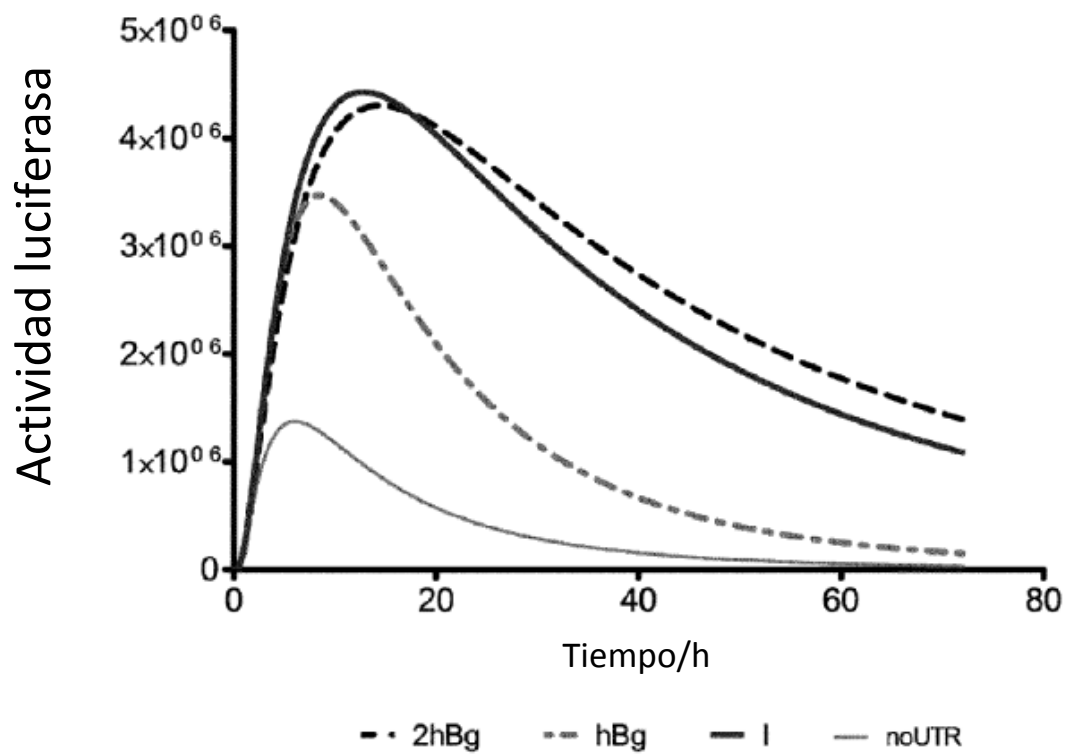


Figura 5

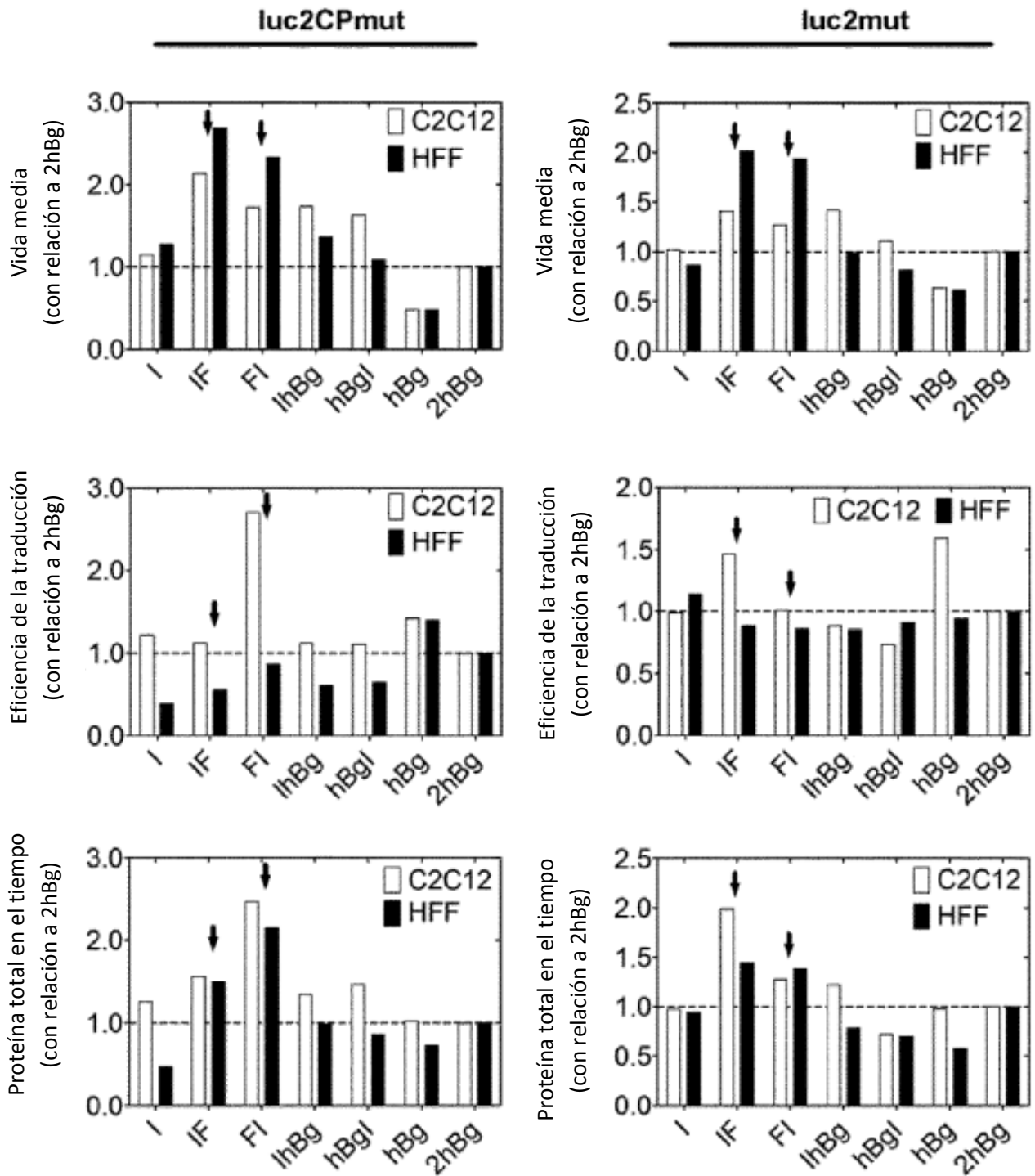


Figura 6

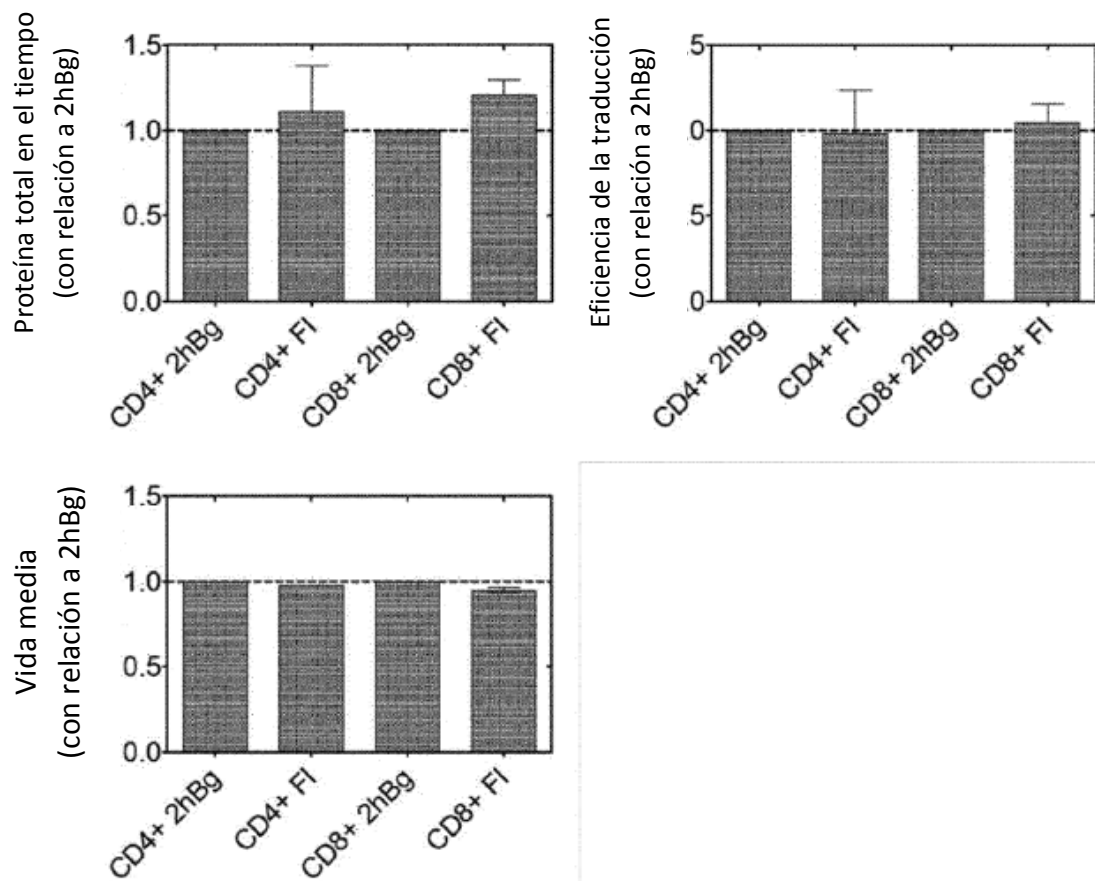
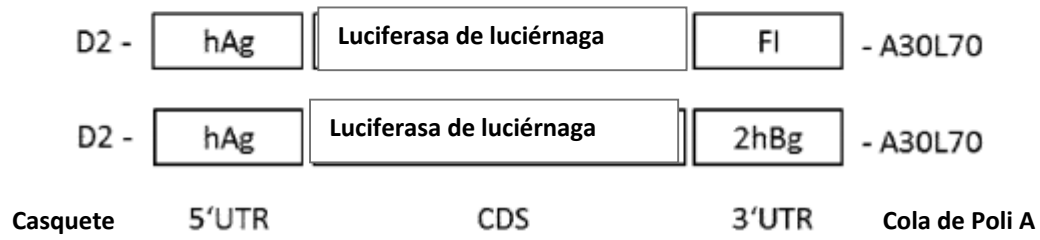


Figura 7

A



B

	Integridad del ARN	
	No modificado	Modificado m1Y
D-2-hAg-Luc-F-I-UTR-[A30-L-A70]	81	84
D-2-hAg-Luc-2hBgUTR-[A30-L-A70]	82	84

Figura 8

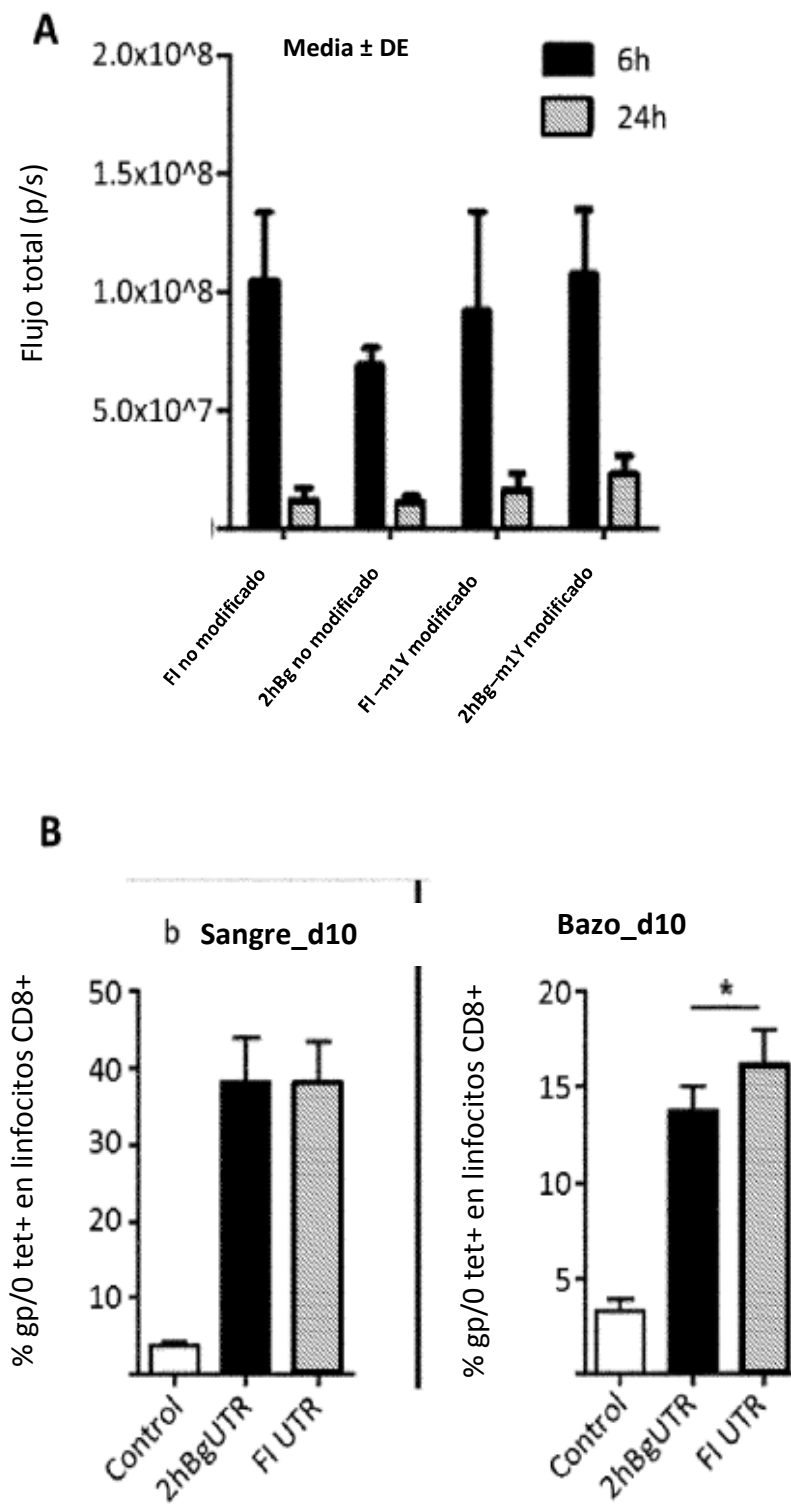


Figura 9

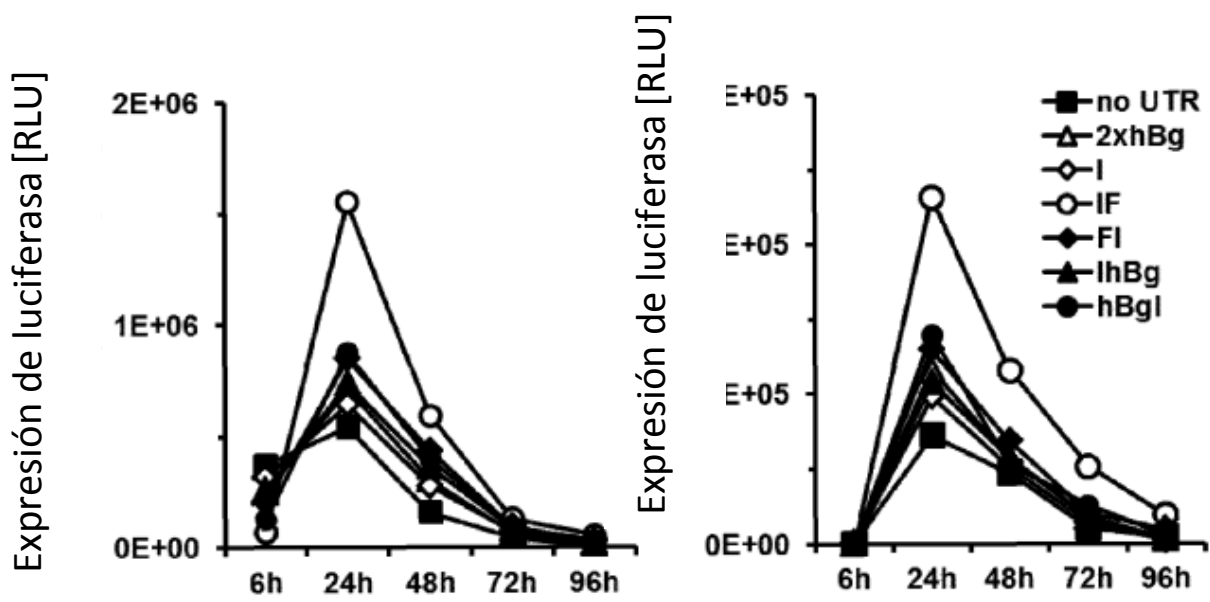


Figura 10

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGU
CUCCCCGACCUCGGGUCCCAGGUAUGCUCCCACCUCACCUGCCCCACUCACCACC
UCUGCUAGUUCCAGACACCUCcCaagcacgcagcaaugcagcucaaaacgcuuagcc
uagccacacccccacgggaaacagcagugauuaaccuuuagcaauaaacgaaaguuu
aacuaagcuauacuaacccccagggguuggucaauuucgugccagccacacc

Figura 11

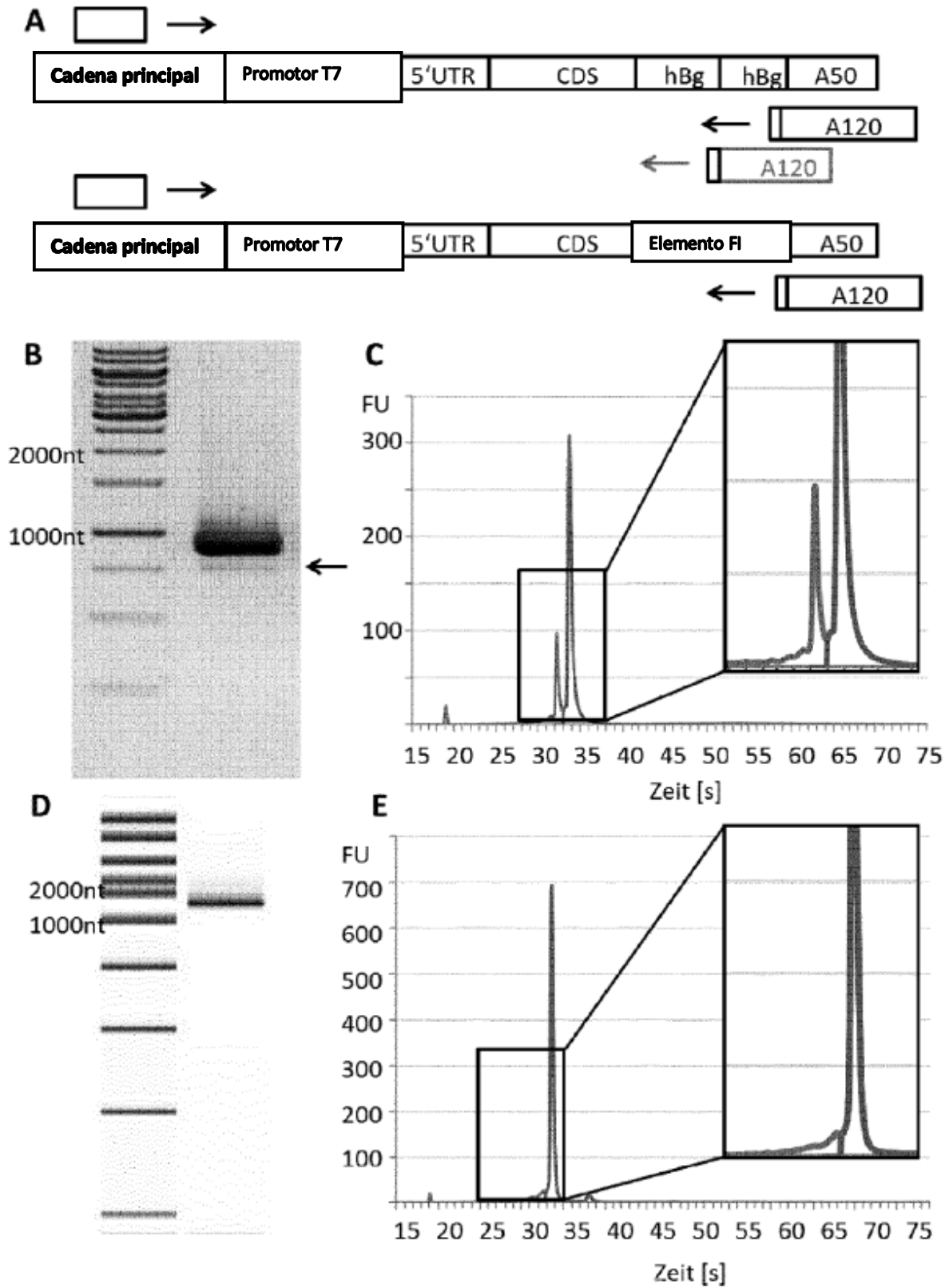


Figura 12 A

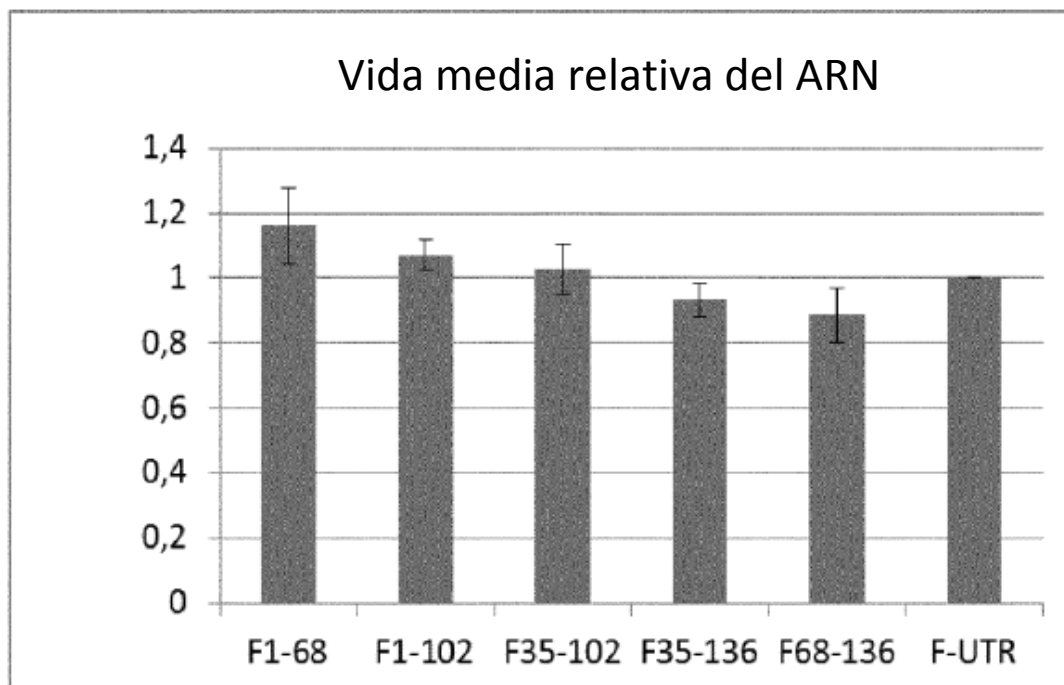
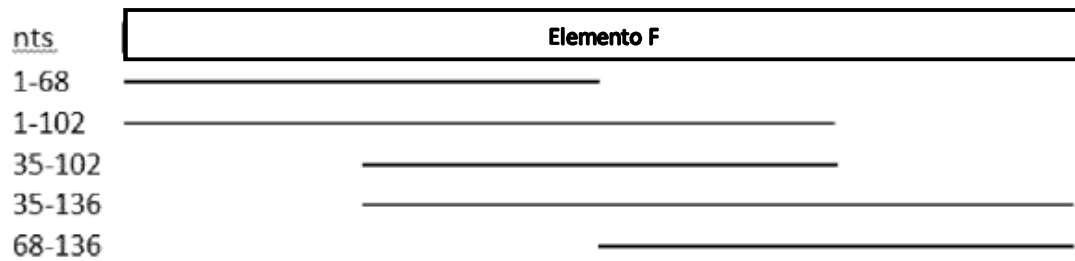


Figura 12 B

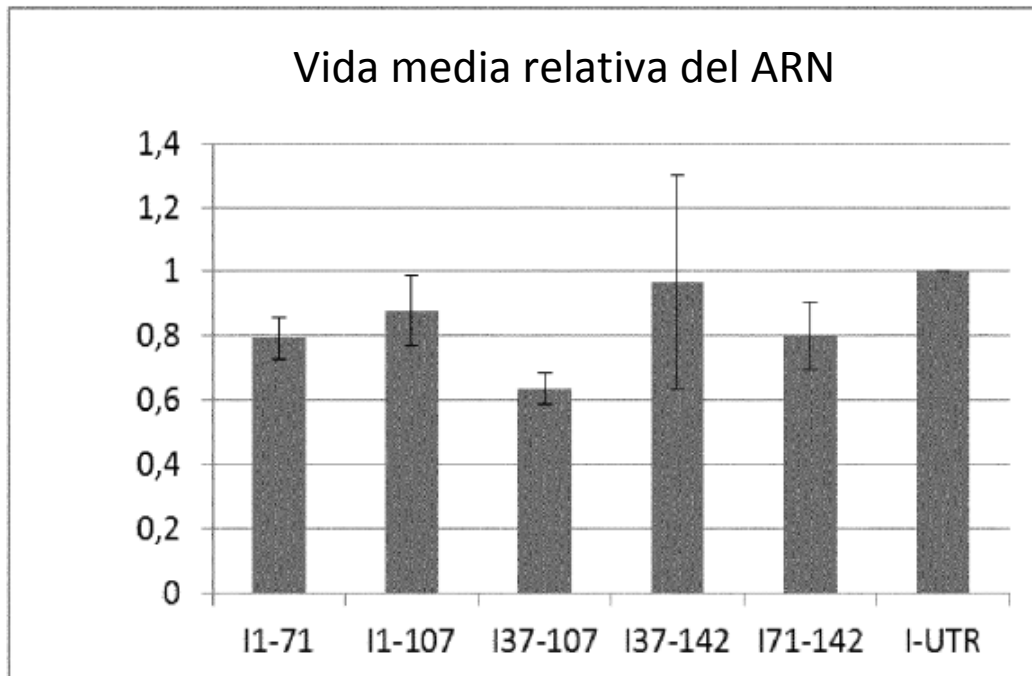
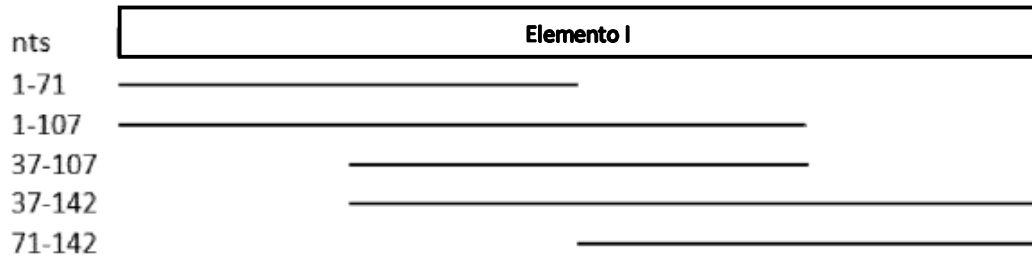


Figura 13

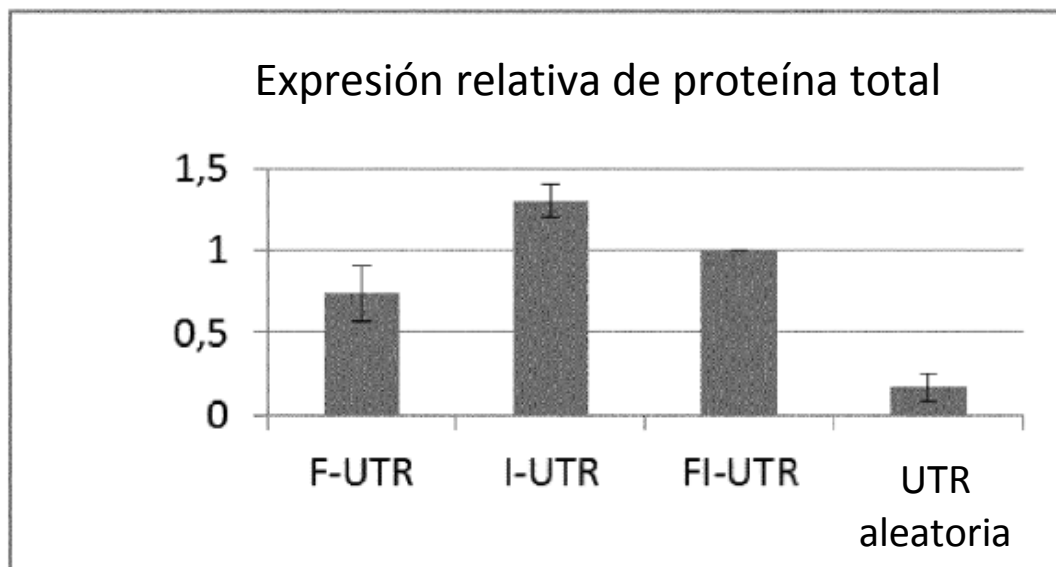
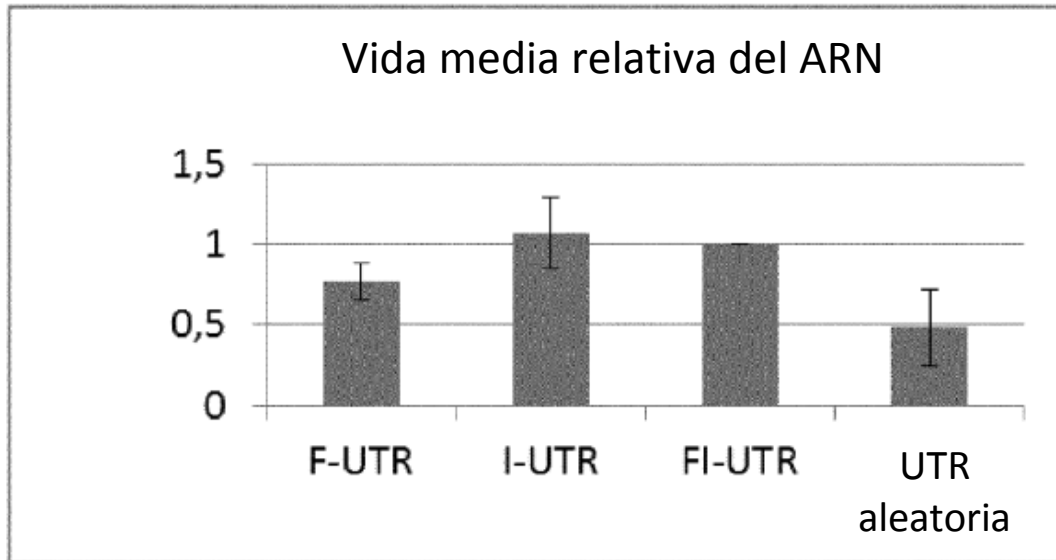


Figura 14

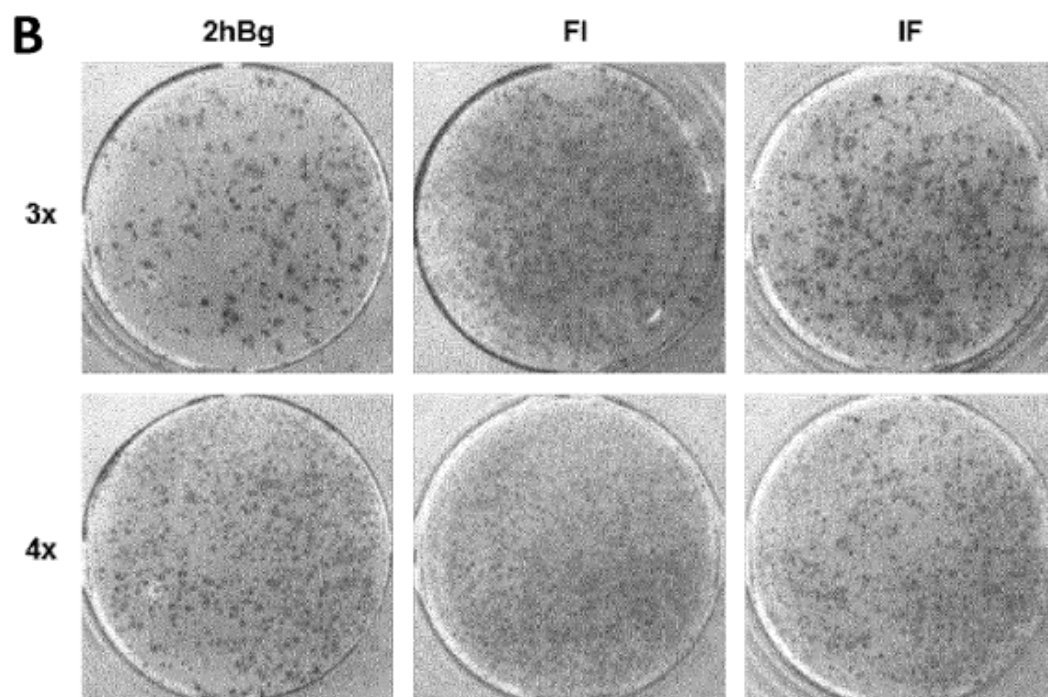
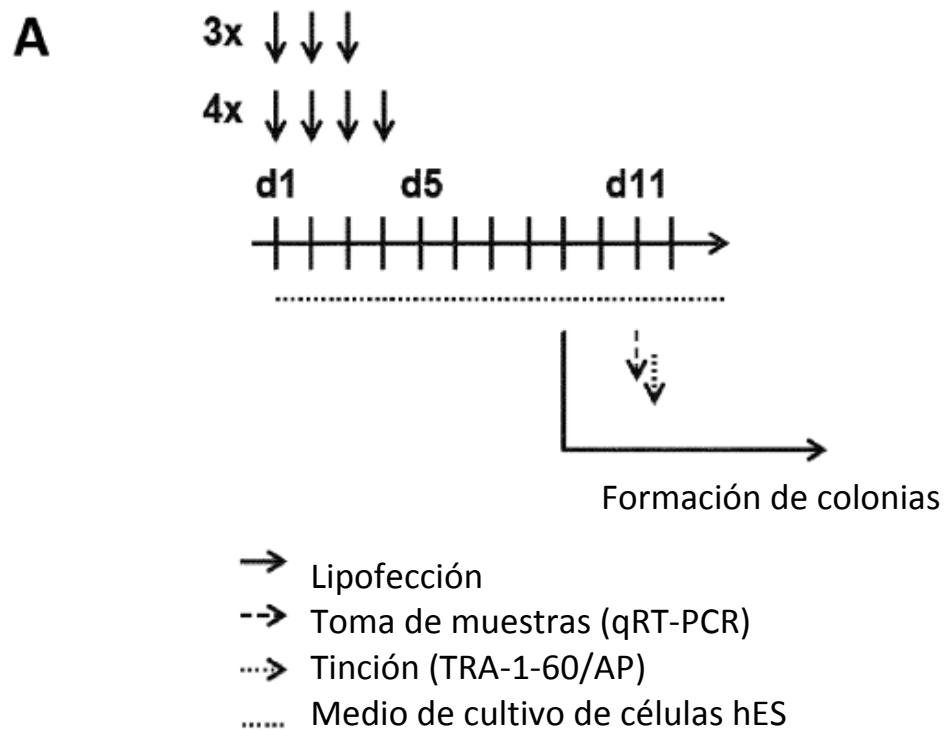


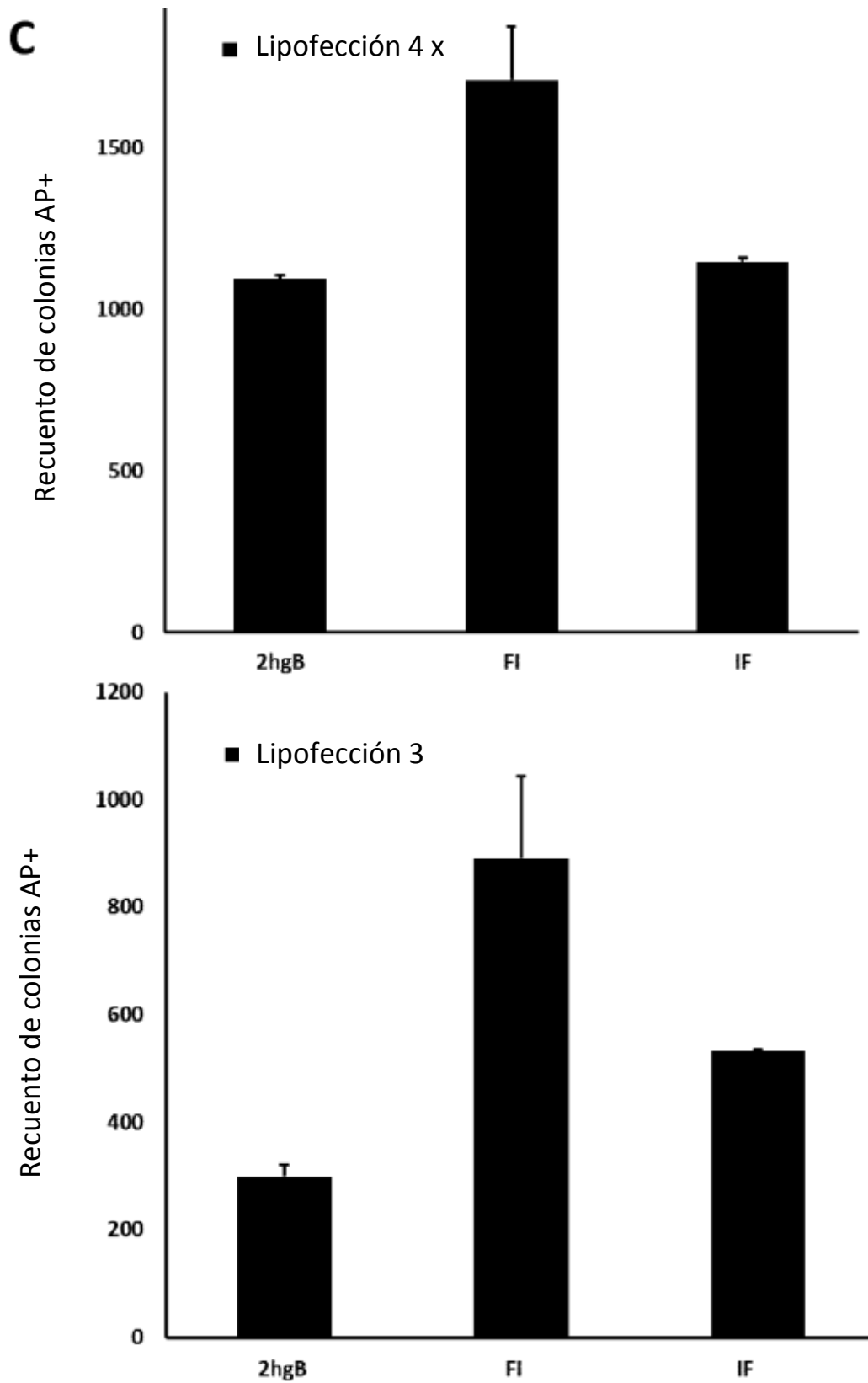
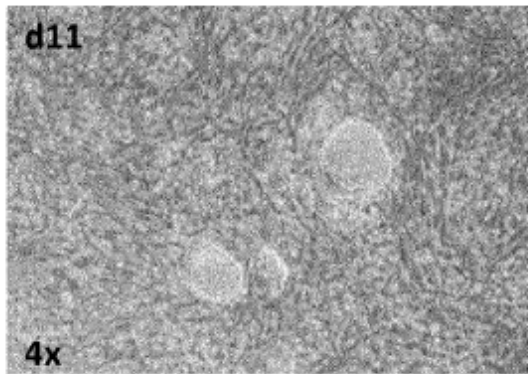
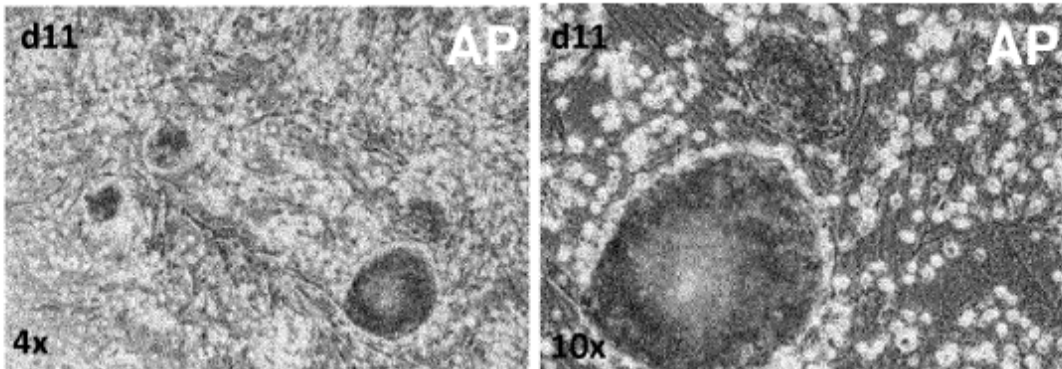
Figura14

Figura 14

D



E



F

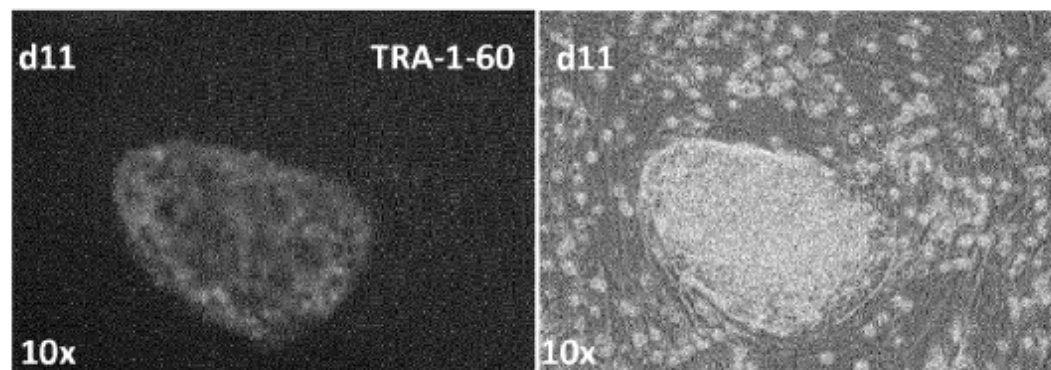


Figura 14 G