



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 344 073**

51 Int. Cl.:
C12N 15/67 (2006.01)
C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06124482 .8**
96 Fecha de presentación : **14.06.2002**
97 Número de publicación de la solicitud: **1829972**
97 Fecha de publicación de la solicitud: **05.09.2007**

54 Título: **Secuencias de ADN con actividad anti-represora.**

30 Prioridad: **04.07.2001 EP 01202581**
05.07.2001 US 303199 P

45 Fecha de publicación de la mención BOPI:
17.08.2010

45 Fecha de la publicación del folleto de la patente:
17.08.2010

73 Titular/es: **CHROMAGENICS B.V.**
Archimedesweg 4
2333 CN Leiden, NL

72 Inventor/es: **Otte, Arie Pieter y**
Kruckeberg, Arthur Leo

74 Agente: **Ungría López, Javier**

ES 2 344 073 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 de ADN con actividad anti-represora.

5 La invención hace referencia a los campos de la medicina y la biología celular. La invención en particular hace referencia a los medios y métodos para la regulación de la transcripción génica. Adicionalmente la invención hace referencia a los medios y métodos para determinar si una secuencia de ADN comprende una cualidad moduladora de la transcripción génica y/o una cualidad represora de la transcripción génica.

10 Con el progreso de los diversos proyectos genoma, se han vuelto asequibles las secuencias de los genomas completos de los organismos. La avalancha de datos ha incrementado el interés de muchos investigadores. Uno de los descubrimientos más notables fue la observación de que el genoma humano no codifica significativamente más genes que el genoma de organismos simples como la mosca de la fruta. El enfoque de muchos investigadores se está desplazando ahora de la identificación de los genes a la determinación de la expresión génica y la función génica. Los
15 ejemplos de tales tecnologías son las micromatrices de ADN, las aplicaciones genómicas funcionales y la proteómica. Estas tecnologías tienen en común que están centradas en la función y expresión de secuencias codificadoras. No obstante, si bien nuestro conocimiento de los genes aumenta espectacularmente, la comprensión de cómo está regulada la expresión de los genes está limitando la capacidad de aplicar este conocimiento rápidamente creciente. Este es por ejemplo el caso en la generación de plantas y animales transgénicos y en la terapia génica humana. En
20 estas aplicaciones el ácido nucleico foráneo es introducido típicamente en células para obtener la expresión de las secuencias codificadoras. A menudo la integración del ácido nucleico foráneo en el genoma de la célula es requerida para el funcionamiento prolongado de las secuencias introducidas. No obstante, la integración de secuencias en el genoma conduce a la imprevisibilidad de la expresión debido a que el ADN circundante influye en la transcripción de las secuencias integradas. Esta imprevisibilidad es debida en parte al hecho de que las secuencias introducidas no
25 pueden ser proporcionadas todavía con la suficiente información genética para aislar funcionalmente las secuencias integradas de los efectos que influyen en la transcripción del ADN circundante. Por otra parte esto es debido al hecho de que no se sabe suficiente sobre los efectos que influyen en la transcripción del ADN circundante.

La presente invención tiene que ver con las secuencias de ADN que comprenden la capacidad de influir en la
30 transcripción de los genes en *cis*. Típicamente, aunque no necesariamente, las secuencias investigadas no codifican por sí mismas una proteína funcional. Diversas secuencias de elementos de secuencia no codifican por sí mismas una proteína funcional. Se han identificado diversos elementos de secuencia con la capacidad de afectar a la transcripción génica en *cis*. Estos elementos se extienden desde promotores, intensificadores, y silenciadores de elementos limítrofes y regiones de anclaje a la matriz.

35 El hecho de que se hayan descubierto tantos tipos diferentes de secuencias reguladoras da la impresión de que es muy fácil diseñar casetes de expresión eficaces. No obstante, más bien es al contrario. El diseño de casetes de expresión todavía está dirigido a menudo por el ensayo y error. Bastante a menudo es posible obtener alguna clase de expresión de un gen foráneo en una célula diana o en su progenie. No obstante, muy a menudo es difícil prever con
40 cualquier clase de exactitud el nivel de expresión o la persistencia de la expresión que una casete de expresión puede presentar en una célula diana.

Un método para detectar, y opcionalmente seleccionar, una secuencia de ADN con una cualidad moduladora de la transcripción génica, comprende proporcionar un sistema de transcripción con una variedad de vectores que com-
45 prenden fragmentos, comprendiendo dichos vectores i) un elemento con una cualidad represora de la transcripción génica, y ii) un promotor que dirija la transcripción de un gen informador, comprendiendo el método adicionalmente realizar una etapa de selección en dicho sistema de transcripción con el fin de identificar dicha secuencia de ADN con dicha cualidad moduladora de la transcripción génica. Dichos fragmentos pueden estar localizados entre i) dicho elemento con una cualidad represora de la transcripción génica, y ii) dicho promotor que dirige la transcripción de
50 dicho gen informador. La ARN polimerasa inicia el proceso de transcripción tras la unión a una secuencia específica, denominada promotor, que señala dónde debe comenzar la síntesis de ARN. Una cualidad moduladora puede intensificar la transcripción a partir de dicho promotor en *cis*, en un tipo de célula dado y/o un promotor dado. La misma secuencia de ADN puede comprender una cualidad intensificadora en un tipo de célula o con un tipo de promotor, mientras puede comprender o no otra cualidad moduladora de la transcripción génica en otro tipo de célula o con otro
55 tipo de promotor. La transcripción puede estar influenciada por un efecto directo del elemento regulador (o la proteína o las proteínas que se unen a él) sobre la transcripción de un promotor concreto. No obstante, la transcripción puede estar influenciada por un efecto indirecto, por ejemplo porque el elemento regulador afecta a la función de uno o más elementos reguladores distintos. La cualidad moduladora de la transcripción génica también puede comprender una cualidad de la transcripción génica estable. Con estable se quiere significar que el nivel de transcripción observado no
60 cambia significativamente a lo largo de al menos 30 divisiones celulares. Una cualidad estable es útil en situaciones en las que las características de expresión deben ser predecibles a lo largo de muchas divisiones celulares. Los ejemplos típicos son las líneas celulares transfectadas con genes foráneos. Otros ejemplos son los animales y plantas transgénicos y las terapias génicas. Muy a menudo, las casetes de expresión introducidas funcionan de manera diferente al cabo de un número creciente de divisiones celulares o generaciones de plantas o animales. En una realización preferida
65 una cualidad estable comprende la capacidad de mantener la transcripción génica en generaciones posteriores de una planta o animal transgénico. Por supuesto en el caso de que la expresión sea inducible, dicha cualidad comprende la cualidad de mantener la inducibilidad de la expresión en generaciones posteriores de una planta o animal transgénico. Frecuentemente, los niveles de expresión caen espectacularmente con el número creciente de divisiones celulares. Con

un método descrito en el presente documento es posible detectar y opcionalmente seleccionar una secuencia de ADN que sea capaz de prevenir al menos en parte la espectacular caída en los niveles de transcripción con números crecientes de divisiones celulares. Dicha cualidad moduladora de la transcripción génica puede comprender una cualidad de transcripción génica estable. Sorprendentemente, los fragmentos que comprenden una secuencia de ADN con dicha cualidad de transcripción génica estable pueden ser detectados y opcionalmente seleccionados con un método descrito en el presente documento, a pesar del hecho de que dicho método no mide necesariamente la estabilidad a largo plazo de la transcripción. En una realización preferida de la invención dicha cualidad moduladora de la transcripción comprende una cualidad intensificadora de la transcripción génica estable. Se ha observado que la incorporación de una secuencia de ADN a un vector de expresión con un gen de interés, produce un nivel superior de transcripción de dicho gen de interés, tras la integración del vector de expresión al genoma de una célula y por otra parte que dicha expresión génica superior también es más estable que en ausencia de dicha secuencia de ADN con una cualidad moduladora de la transcripción génica.

En experimentos diseñados para introducir un gen de interés en el genoma de una célula y para obtener la expresión de dicho gen de interés, se ha observado lo siguiente. Si junto con dicho gen de interés también se había introducido una secuencia de ADN con una cualidad moduladora de la transcripción génica, se podían detectar más clones que expresaban más de una cierta cantidad de producto génico de dicho gen de interés, que cuando dicha secuencia no era introducida junto con dicho gen de interés. De este modo la presente invención también proporciona un método para incrementar el número de células que expresan más de cierto nivel de un producto génico de un gen de interés tras proporcionar dicho gen de interés al genoma de dichas células, que comprende proporcionar a dicha célula una secuencia de ADN que comprende una cualidad moduladora de la transcripción génica junto con dicho gen de interés.

Las oportunidades de detectar un fragmento con una cualidad moduladora de la transcripción génica varían con la fuente de la cual derivan los fragmentos. Típicamente, no existe un conocimiento previo de la presencia o ausencia de fragmentos con dicha cualidad. En esas situaciones muchos fragmentos no comprenderán una secuencia de ADN con una cualidad moduladora de la transcripción génica. En estas situaciones se introduce una etapa de selección formal para las secuencias de ADN con dicha cualidad. Esto se realiza por medio de la selección de vectores que comprenden dicha secuencia basándose en una característica de un producto de dicho gen informador, que puede ser seleccionado a favor o en contra. Por ejemplo, dicho producto génico puede inducir fluorescencia o un depósito de color (v.g. proteína fluorescente verde y derivados, luciferasa o fosfatasa alcalina) o conferir resistencia a antibióticos o inducir apoptosis y muerte celular.

Un método como se describe en el presente documento es particularmente adecuado para detectar y opcionalmente seleccionar una secuencia de ADN que comprende una cualidad intensificadora de la transcripción génica. Se ha observado que al menos alguna de las secuencias de ADN seleccionadas, cuando son incorporadas a un vector de expresión que comprende un gen de interés, pueden incrementar espectacularmente la transcripción génica de dicho gen de interés en una célula huésped incluso cuando el vector no comprende un elemento con una cualidad represora de la transcripción génica. Esta cualidad intensificadora de la transcripción génica es muy útil en las líneas celulares transfectadas con genes foráneos o en animales y plantas transgénicos.

Para la presente invención dicho sistema de transcripción comprende células huésped. El uso de células huésped garantiza que los fragmentos son detectados y opcionalmente seleccionados con actividad en las células.

Un elemento con una cualidad represora de la transcripción génica reprimirá, en el método de la invención la transcripción de un promotor en el sistema de transcripción utilizado. Dicha represión no tiene que conducir a niveles de expresión no detectables. Es importante que la diferencia en los niveles de expresión en ausencia o presencia de represión sea detectable y opcionalmente seleccionable. En una realización, la represión de la transcripción génica en dichos vectores produce cromatina represora de la transcripción génica. En esta realización se pueden detectar secuencias de ADN, y opcionalmente seleccionar que sean capaces, al menos en parte, de contrarrestar la formación de cromatina represora de la transcripción génica. En un aspecto una secuencia de ADN capaz de contrarrestar, al menos en parte, la formación de cromatina represora de la transcripción génica comprende una cualidad de transcripción génica estable. En una realización la secuencia de ADN implicada en la represión de la transcripción génica es una secuencia de ADN que es reconocida por un complejo proteico y donde dicho sistema de transcripción comprende dicho complejo. Preferiblemente dicho complejo comprende una proteína de unión a heterocromatina que comprende HP1, una proteína del grupo Polycomb (Pc-G), una actividad histona desacetilasa o MeCP2 (proteína de unión a metil-CpG). Muchos organismos comprenden una o más de estas proteínas. Estas proteínas frecuentemente muestran también actividad en otras especies. Dicho complejo puede por tanto comprender también proteínas de dos o más especies. El grupo mencionado de complejos de proteínas asociadas con cromatina conocidos es capaz de transmitir la represión de un amplio intervalo a lo largo de muchos pares de bases. Los complejos también están implicados en la transferencia estable del estado reprimido de los genes a las células hijas tras la división celular. Las secuencias seleccionadas de esta manera son capaces de transmitir la anti-represión de amplio intervalo a lo largo de muchos pares de genes (van der Vlag *et al.*, 2000).

El vector utilizado puede ser cualquier vector que sea adecuado para la clonación de ADN y que pueda ser utilizado en un sistema de transcripción. Cuando se utilizan células huésped, se prefiere que el vector sea un vector de replicación episómica. De esta manera, se evitan los efectos debidos a los sitios de integración diferentes del vector. Los elementos de ADN que flanquean el vector en el sitio de integración pueden tener efectos sobre el nivel de transcripción del promotor y de ese modo imitar los efectos de los fragmentos que comprenden secuencias de ADN con una cualidad

moduladora de la transcripción génica. En una realización dicho vector comprende un origen de replicación del virus de Epstein-Barr (EBV), OriP, y un antígeno nuclear (EBNA-1). Tales vectores son capaces de replicar en muchos tipos de células eucarióticas y ensamblarse a la cromatina en condiciones apropiadas.

Se proporciona una secuencia de ADN que comprende i) una secuencia de ADN aislada de una planta o vertebrado, o derivados de los mismos, o ii) una secuencia de ADN sintético o una construida por medio de ingeniería genética, cuya secuencia de ADN es una secuencia inhibidora de la represión que, puede ser detectada mediante el método descrito en el presente documento, seleccionada y opcionalmente clonada. También se proporciona una secuencia de ADN que comprende i) una secuencia de ADN aislada de una planta o vertebrado, o derivados de los mismos, o ii) una secuencia de ADN sintético o una construida mediante ingeniería genética, cuya secuencia de ADN es detectada, seleccionada y opcionalmente clonada por medio del método descrito en el presente documento. Dicha secuencia de ADN puede comprender una secuencia como se representa en la Tabla 4 o un homólogo funcional de la misma. Un homólogo funcional de una secuencia representada en la Tabla 4 es una secuencia derivada con la información dada en la Tabla 4. Por ejemplo, una secuencia que puede ser derivada de una secuencia de la Tabla 4 mediante delección, modificación y/o inserción de bases en o de la secuencia enumerada en la Tabla 4, donde dicha secuencia derivada comprende la misma actividad en clase, no necesariamente en cantidad, de una secuencia como se representa en la Tabla 4. Un homólogo funcional es adicionalmente una secuencia que comprende una parte de dos o más secuencias representadas en la Tabla 4. Una secuencia de ADN sintético es una secuencia que no deriva directamente o indirectamente de una secuencia presente en un organismo. Por ejemplo una secuencia que comprende una secuencia scs o scs' de drosophila no es una secuencia sintética, ni siquiera cuando la secuencia scs o scs' es generada artificialmente.

La presente invención proporciona en particular una secuencia de ADN aislada y/o recombinante que tiene actividad antirepresora, dicha secuencia seleccionada de un grupo que consiste en:

- (a) ID SEC.: 43 en la figura 20;
- (b) un fragmento de ID SEC.: 43 en la Figura 20;
- (c) una secuencia derivada de ID SEC. N°: 43 en la figura 20 por inserción, delección y/o mutación de una o más bases;

en la que la actividad antirepresora de dicha secuencia confiere a una célula de osteosarcoma U-2 OS humana la habilidad de crecer después de 4-5 semanas de cultivo en presencia de zeocina 250 $\mu\text{g/ml}$ y doxiciplina 0,1 ng/ml , cuando dicha célula comprende una proteína de fusión represora de LexA que contiene el dominio de unión a ADN LexA y una región codificante de HP1 o HPC2 bajo control del sistema regulatorio transcripcional Tet-Off, cuando dicha secuencia de ADN aislado y/o recombinante se clona en una secuencia de policlonaje en un plásmido, situándose dicha región de policlonaje entre cuatro sitios operadores LexA y el promotor SV40 que controla el gen de resistencia a zeocina, cuando el plásmido está presente en dicha célula.

La invención se refiere al conocimiento creciente de la regulación génica de orden superior y a los medios y métodos para utilizar este conocimiento. Si bien se han caracterizado elementos, tales como los promotores y los intensificadores clásicos, que dirigen y regulan la transcripción de genes individuales, los elementos reguladores de orden superior que gobiernan las capacidades de la transcripción génica de regiones cromosómicas completas todavía han recibido poca atención. Gran parte del conocimiento de los autores de la presente invención referente a semejantes elementos de orden superior proviene del estudio de la embriogénesis. Durante la embriogénesis las células quedan comprometidas en diferentes rutas evolutivas. Una vez comprometidas, las células raramente cambian su destino, ni siquiera después de muchas divisiones celulares.

Esta quedando cada vez más claro que la transmisión estable de los patrones de transcripción génica específicos del tipo de célula no depende del contexto de un promotor, sino que en lugar de eso está mediada por cambios en la estructura del ADN y las proteínas asociadas, denominadas cromatina. La regulación génica a nivel cromosómico implica modificaciones del ADN (v.g. metilación), las histonas (v.g. acetilación y/o metilación), y las interacciones de amplio intervalo entre elementos cromosómicos distantes.

El molde de cromatina es un complejo altamente condensado de ADN, histonas, y proteínas no histónicas, que es capaz de empaquetar el genoma completo en el núcleo y simultáneamente permitir la transcripción apropiada de los genes específicos. El cromosoma eucariótico no es un molde uniforme para la activación de la transcripción génica. Se pueden distinguir diferentes tipos de cromatina y regiones de cromatina, que afectan de modo diferente a la transcripción génica. Las llamadas regiones de heterocromatina identifican estructuras de cromatina “cerradas” mientras la eucromatina está asociada con una estructura de cromatina “abierta” más difusa. La región de eucromatina puede ser sometida a cambios estructurales, dando como resultado estructuras más o menos condensadas, referidas como heterocromatina y eucromatina facultativas. Se cree que la formación de eucromatina o heterocromatina facultativa representa el mecanismo subyacente de la regulación génica mediada por cromatina, manteniendo los genes en un estado activo o reprimido, de una manera específica del tipo celular.

En todos los eucariotas se han identificado numerosos complejos asociados con cromatina que están implicados en el mantenimiento de la especificidad del tipo celular, uno de los cuales es el complejo del grupo Polycomb (PcG). El

complejo PcG está implicado en la represión estable de los genes, en los que se cree que los cambios en la estructura de la cromatina juegan un importante papel. De un modo similar, se ha identificado una segunda clase de proteínas, denominada grupo trithorax (Trg), que contrarresta la acción de las proteínas PcG. Las proteínas TrG están implicadas en el mantenimiento de la transcripción génica. Basándose en sus respectivos modos de acción, las proteínas PcG y TrG representan por lo tanto un sistema de memoria celular que es importante para la transmisión heredable de los patrones de transcripción génica.

Cómo están asociados los complejos de PcG y TrG con sus genes diana todavía no está claro. Estudios genéticos han caracterizado secuencias reguladoras que actúan en cis que mantienen estados transcripcionalmente inactivos de los genes. El silencio mediado por estas secuencias reguladoras que actúan en cis depende de la presencia de proteínas PcG funcionales, y por tanto estas secuencias han sido denominadas elementos de respuesta a PcG (PRE). Se han identificado secuencias que están implicadas en la represión de la cromatina mediada por PcG. Sin embargo, todavía no se ha encontrado (en vertebrados ni en plantas) PRE completos que comprendan toda la información de la secuencia requerida para mediar la represión de la cromatina.

Un elemento de respuesta al grupo Polycomb es un elemento que es capaz de reprimir la transcripción de un promotor en respuesta a la interacción directa y/o indirecta de una o más proteínas del grupo Polycomb con dicho elemento. Un elemento de respuesta de tipo grupo Polycomb es un elemento de respuesta al grupo Polycomb o alternativamente es un elemento capaz de reprimir la transcripción de un promotor tras la interacción directa y/o indirecta de una o más proteínas con dicho elemento, donde dichas una o más proteínas no pertenecen al grupo Polycomb, pero donde como resultado de dicha represión de la transcripción génica por la interacción se forma cromatina. Los ejemplos de tales proteínas son proteínas asociadas con cromatina tales como la proteína heterocromatina 1 (HPLJL) (Eisenberg *et al.*, 1990). Otra proteína asociada con la cromatina que reprime la actividad génica es la proteína de unión a metil-CpG, MeCP2 (Nan *et al.*, 1997). En una realización preferida un elemento sensible de tipo grupo polycomb de la invención comprende la capacidad para reprimir la transcripción de un promotor a lo largo de grandes distancias, preferiblemente a lo largo de más de 2.000 pares de bases (van der Vlag *et al.*, 2000).

Un gen informador es un gen que codifica un producto de expresión cuya presencia puede ser detectada directamente o indirectamente en una célula.

Entre los ejemplos de las secuencias de ADN con una cualidad moduladora de la transcripción génica se encuentran los denominados elementos STAR enumerados en las Tablas 1 y 2.

Los métodos como se describen en el presente documento producen la clonación e identificación de numerosos elementos que comprenden una actividad moduladora de la transcripción génica. Semejante elemento puede contener ácido nucleico irrelevante que no es instrumental en el desempeño de dicha cualidad, por ejemplo no implicado en la formación de cromatina represora de la transcripción génica. Las secuencias funcionales de tales elementos pueden ser delineadas por medio de diversos métodos conocidos en la técnica. En una realización se realizan delecciones y/o sustituciones en una secuencia de ADN con una cualidad moduladora de la transcripción génica. El ADN que es modificado de tal modo, es sometido a ensayo en cuanto a la actividad. Esto se puede realizar utilizando un ácido nucleico modificado individual o generando una colección de ácidos nucleicos de ensayo que comprende dicho ácido nucleico modificado.

La elucidación de las secuencias funcionales dentro de las secuencias de ADN de la invención permite la elucidación de las secuencias consenso para los elementos con una cualidad moduladora de la transcripción génica. Se prevé que se encontrará más de un tipo de secuencias consenso para un elemento que comprenda una cualidad moduladora de la transcripción génica. Puede proporcionarse una genoteca de ácidos nucleicos aislados y/o recombinantes que comprenden cualidades moduladoras de la transcripción génica y/o represora de la transcripción génica tales como elementos de respuesta de tipo grupo polycomb. En una realización dicha genoteca comprende ácidos nucleicos aislados y/o recombinantes que comprenden la misma secuencia consenso. En una realización dicha genoteca comprende más de un tipo de secuencias consenso. Dicha genoteca puede ser utilizada por ejemplo para determinar si una molécula de ADN dada comprende una cualidad moduladora del ADN. En una realización preferida dicha genoteca comprende esencialmente todos los elementos con una función intensificadora de la transcripción génica, elementos que comprenden una cualidad de la transcripción génica estable y/o elementos con una cualidad represora de la transcripción génica tales como elementos de respuesta de tipo grupo polycomb, de un cromosoma. Junto con el conocimiento de la localización de estos elementos en el cromosoma esto permite a un experto en la técnica generar una predicción de la regulación de orden superior de la expresión génica de genes naturalmente presentes en dicho cromosoma y de genes (ácido nucleico foráneo) introducidos en dicho cromosoma mediante medios recombinantes. Semejante predicción puede ser utilizada por ejemplo para seleccionar una localización candidato adecuada en dicho cromosoma para la inserción de ADN foráneo. Una localización adecuada puede ser la localización que se espera que sea expresada específicamente en una cierta célula, tipo celular y/o tejido. Preferiblemente, dicho cromosoma comprende el cromosoma 21 o el cromosoma 22. En una realización todas las secuencias de ADN que comprenden una cualidad moduladora de la transcripción génica o represora de la transcripción génica en una célula, están en la genoteca. En esta realización se puede utilizar el genoma completo para la predicción de una localización candidato adecuada. En una realización dicha genoteca ha sido generada en diferentes líneas celulares de especies que oscilan entre plantas y humanos. En diferentes líneas y/o especies celulares se expresarán proteínas diferentes (o complejos de proteínas) capaces de interaccionar con secuencias de ADN con una cualidad represora de la transcripción génica, dando como resultado diferentes elementos de ADN con una cualidad represora de la transcripción génica. De un modo similar se

expresarán diferentes proteínas que interaccionan directa o indirectamente con secuencias de ADN que comprenden una cualidad moduladora de la transcripción génica. Por lo tanto la elaboración de la genoteca depende del tipo celular y depende de la presencia de proteínas relevantes. Este también es el caso con los elementos de respuesta de tipo grupo polycomb. Si HP1 es expresada en el tipo celular uno, los elementos dependientes de HP1 serán detectados mediante el método de la invención. Si HP1 no es expresada en el tipo celular dos, el método de la invención no detectará el elemento que ha sido recuperado del tipo celular uno.

En un aspecto dicha genoteca comprende al menos un elemento capaz de contrarrestar, al menos en parte, la formación de cromatina represora de la transcripción génica. Junto con el conocimiento de la localización de las secuencias de ADN con una cualidad represora de la transcripción génica en un cromosoma o genoma, el conocimiento de la localización de semejantes elementos contrarrestadores permite la predicción exacta de la regulación de orden superior de la transcripción génica de genes (insertados) en dicho cromosoma o genoma. Preferiblemente dicha genoteca comprende otros elementos reguladores de la transcripción tales como intensificadores y silenciadores. Aunque tales secuencias han limitado la influencia sobre la regulación génica de orden superior, la información sobre la localización de tales otras secuencias incrementa adicionalmente la exactitud de la predicción de las localizaciones adecuadas en el genoma para la expresión de secuencias foráneas introducidas allí. Preferiblemente, dicha genoteca comprende esencialmente todas las secuencias de ADN que comprenden una cualidad moduladora de la transcripción génica y/o todas las demás secuencias reguladoras de un cromosoma.

Considerando que un cromosoma ya consta típicamente de varias decenas de millones de bases, se prefiere que la información que la genoteca pueda proporcionar sobre la regulación génica de orden superior sea incorporada en un sistema al menos parcialmente automatizado.

Otro uso de una genoteca es la generación de una predicción de la transcripción de genes tras la modificación elegida como diana de las secuencias de un cromosoma de manera que las secuencias reguladoras de “orden superior” sean mutadas. Por ejemplo, uno o más elementos sensibles de tipo grupo polycomb de la invención, y/o otros elementos reguladores de dicho cromosoma pueden ser mutados. Se espera que esto cambie los niveles de transcripción de los genes que están en las proximidades de los elementos sensibles de tipo grupo polycomb y/o otros elementos moduladores de la expresión.

Otro uso más de una genoteca o sistema es la predicción de la expresión génica resultante de las mutaciones del genoma. En los casos en los que una mutación da como resultado una transcripción génica alterada, la detección de semejante transcripción génica alterada puede indicar la presencia de dicha mutación de origen natural. Este enfoque es útil por ejemplo al limitar el número de secuencias de proteínas que van a ser sometidas a ensayo en un análisis de diagnóstico. Esto es particularmente importante en los enfoques de micromatrices debido a que en estos enfoques el número de secuencias expresadas que va a ser sometido a ensayo, está limitado por el número de secuencias que puede contener una matriz como máximo. Con los medios y métodos como se describen en el presente documento es posible limitar el número de secuencias que van a ser sometidas a ensayo en los enfoques de micromatrices.

Otro uso más de un sistema o genoteca es el descubrimiento de las dianas de fármaco. Los elementos reguladores, sean elementos de “orden superior” o no, funcionan debido a la proteína (complejos) que se puede unir a ellos. Se puede utilizar un sistema para determinar si el redireccionamiento de los fármacos para interferir en la unión o la función de una proteína concreta (complejo) proporciona la esperanza de la alteración de un gen concreto.

De acuerdo con la invención también se proporciona un constructo de ADN con una secuencia de ADN de la invención. En una realización preferida se proporciona un constructo de ADN que comprende un promotor conectado operablemente con un ácido nucleico de interés. Preferiblemente, la cantidad de actividad de una cualidad de dicha secuencia de ADN con una cualidad moduladora de la transcripción génica, depende de la orientación de dicha secuencia de ADN en dicho constructo, en comparación con dicho promotor. Preferiblemente dicha cualidad moduladora de la transcripción génica depende de la presencia de una señal. Preferiblemente, dicha señal comprende una proteína de unión a ADN. Preferiblemente, dicha señal comprende una proteína TAT del virus de la inmunodeficiencia humana.

Uno de los usos de una secuencia de ADN que comprende una cualidad moduladora de la transcripción génica es por supuesto la regulación de la transcripción de un gen de interés. La transcripción de un gen de interés puede ser alterada alterando las secuencias en las proximidades de dicho gen de manera que las secuencias de ADN con cualidades moduladoras de la transcripción génica se pueden generar de cualquier clase o cantidad o ambas.

Asimismo es posible diseñar secuencias de ADN con una cualidad moduladora de la transcripción génica deseada. Las proteínas de unión a ADN junto con otras proteínas y secuencias de ADN determinan las cualidades de la secuencia de ADN. Es posible insertar una o más secuencias de ADN de unión a proteínas en una secuencia de ADN con una cualidad. Permitiendo la unión de la proteína o las proteínas de unión es posible interferir en, o dirigir, la cualidad, permitiendo de este modo la generación de secuencias de ADN con cualidades constructoras. Por supuesto también es posible separar sitios de unión a la proteína de una secuencia de ADN con una cualidad moduladora de la transcripción génica alternado de ese modo la cualidad de las secuencias de ADN resultantes. La combinación de la adición y eliminación también es posible. Se pueden seleccionar cualidades moduladoras de la transcripción génica concretas poniendo a punto los métodos de detección descritos en la presente invención. Por ejemplo es posible sintetizar secuencias de ADN con cualidades moduladoras de la transcripción génica inducibles. Se encuentran disponibles proteínas de unión a ADN que solamente se unen a su secuencia diana en ausencia o presencia de una señal. Los

ejemplos no limitantes de tales proteínas son el represor TET y las diversas mutaciones del mismo, el represor lac, los receptores de hormonas esteroides, el receptor del ácido retinoico, y sus derivados. Es posible por ejemplo diseñar una secuencia de ADN con una cualidad modulador de la transcripción génica específica del tipo celular. La secuencia de ADN se puede hacer específica para un complejo de proteína que sea expresado de una manera específica del tipo celular.

Los constructos de expresión que comprenden una secuencia de ADN que comprende una cualidad moduladora de la transcripción génica son adecuados para obtener la expresión de dicho constructo en células que comprenden más de una copia de dicho constructo de expresión. Asimismo cuando el constructo de expresión está presente en el genoma de dicha célula y, asimismo cuando la casete de expresión está presente en más de una copia de dicha célula. Por otra parte, incluso funcionan cuando son integrados en la misma posición en más de una copia.

De acuerdo con la invención dicha secuencia de ADN que tiene actividad anti-represora comprende una secuencia denominada STAR (Stabilizing Anti-Repression). Una secuencia STAR según se utiliza aquí hace referencia a una secuencia de ADN que comprende una o más de las cualidades moduladoras de la transcripción génica mencionadas.

Se presentan aquí numerosos métodos para determinar si una secuencia comprende actividad STAR. La actividad STAR es confirmada si la secuencia es capaz de realizar al menos una de las siguientes funciones: (i) inhibir al menos en parte el efecto de la secuencia que comprende un elemento de represión de la transcripción génica de la invención, (ii) bloquear al menos en parte la represión asociada con la cromatina, (iii) bloquear al menos en parte la actividad de un intensificador, (iv) conferir a un ácido nucleico conectado operablemente que codifica una unidad de transcripción en comparación con el mismo ácido nucleico solo (iv-a) un predictibilidad de la transcripción superior, (iv-b) un transcripción superior, y/o una estabilidad de la transcripción a lo largo del tiempo superior.

El gran número de secuencias que comprenden actividad STAR identificadas en la presente invención abre una amplia variedad de posibilidades para generar e identificar las secuencias que comprenden la misma actividad en clase, no necesariamente en cantidad. Por ejemplo, está dentro del alcance del experto en la técnica alterar las secuencias identificadas en la presente invención y someter a ensayo las secuencias alteradas en cuanto a la actividad STAR. Tales secuencias alteradas son también por lo tanto parte de la presente invención. La alteración puede incluir la delección, inserción y mutación de una o más bases en las secuencias.

Las secuencias que comprenden actividad STAR fueron identificadas en tramos de 400 bases. No obstante, no se espera que sean necesarias las 400 bases para conservar la actividad STAR. Los métodos para delimitar las secuencias que confieren una cierta propiedad a un fragmento de entre 400 y 5.000 bases son bien conocidos. Se estima que la longitud de la secuencia mínima de un fragmento que comprende actividad STAR es de aproximadamente 50 bases.

La actividad STAR es un rasgo compartido por las secuencias enumeradas en la Figura 20.

Como se ha mencionado antes, una secuencia STAR puede ejercer su actividad de un modo direccional, es decir, conteniéndola un lado del fragmento más que otro. Por otra parte, la actividad STAR puede ser amplificada en cantidad multiplicando el número de elementos STAR. Esto último sugiere que un elemento STAR puede comprender uno o más elementos que comprendan actividad STAR.

El término cualidad en relación con la secuencia hace referencia a una actividad de dicha secuencia. El término STAR, secuencia STAR o elemento STAR, según se utiliza aquí, hace referencia a una secuencia de ADN que comprende una o más de las cualidades moduladoras de la transcripción génica mencionadas. El término "secuencia de ADN" según se utiliza aquí, a menos que se especifique de otro modo, no hace referencia a un listado de ordenamiento específico de bases sino más bien a una porción física de ADN. Una cualidad de transcripción con referencia a una secuencia de ADN hace referencia a un efecto que dicha secuencia de ADN tiene sobre la transcripción de un gen de interés. Según se utiliza aquí, "cualidad" hace referencia a las propiedades detectables o atributos de un ácido nucleico o proteína en un sistema de transcripción.

Ejemplos

Ejemplo 1

Métodos para aislar elementos STAR

Materiales y métodos

Plásmidos y cepas. El vector de selección para los elementos STAR, pSelect-SV40-zeo ("pSelect", Figura 1) es construido como sigue: el vector pREP4 (Invitrogen V004-50) es utilizado como esqueleto plasmídico. Este proporciona el origen de replicación oriP de Epstein Barr y el antígeno nuclear EBNA-1 para la replicación episómica de elevado número de copias en líneas celulares de primate; el gen de resistencia a la higromicina con el promotor de la timidina quinasa y el sitio de poliadenilación, para la selección en células de mamífero; y el gen de resistencia a ampicilina y el origen de replicación colE1 para el mantenimiento en *Escherichia coli*. El vector contiene cuatro sitios consecutivos para el operador LexA entre los sitios de restricción XbaI y NheI (Bunker and Kingston, 1994). Embebido entre los operadores LexA y el sitio NheI está un poliligador que consta de los siguientes sitios de restric-

ción: *HindIII*-*AscI*-*BamHI*-*AscI*-*HindIII*. Entre el sitio *NheI* y un sitio *SalI* está el gen de resistencia a zeocina con el promotor de SV40 y el sitio de poliadenilación, derivado de pSV40/Zeo (Invitrogen V502-20); este es el marcador seleccionable para el rastreo de STAR.

El vector pSDH (Figura 2) es construido como sigue: El gen informador de la luciferasa de pGL3-Control (Promega E1741) es amplificado mediante PCR e insertado en pUHD10-3 digerido con *SacII*/*BamHI* (Gossen and Bujard, 1992). Este coloca la luciferasa bajo el control del promotor Tet-Off, y aguas arriba de la señal de poliadenilación de SV40. Se introducen múltiples sitios de clonación mediante PCR, aguas arriba del promotor Tet-Off (MCSI, *XhoI*-*NotI*-*EcoRI*-*SalI*) y aguas abajo de la señal de poliadenilación (MCSII, *NbeI*-*BglIII*-*EcoRV*-*HindIII*).

Se construyen genotecas génicas mediante digestión con *Sau3AI* de ADN genómico humano, ya sea purificado a partir de placenta (Clontech 6550-1) o portado en cromosomas artificiales bacteriano/P1 (BAC/PAC). Los clones BAC/PAC contienen ADN genómico de la región citogenética 1q12 (clones RP1154H19 y RP3328E19) o a partir de la agrupación HOX de genes homeóticos (clones RP1167F23, RP1170019, y RP11387A1). Los ADN son fraccionados por tamaño, y la fracción de tamaño 0,5-2 kb es ligada en el vector pSelect digerido con *BamHI*, mediante mecanismos normalizados (Sambrook *et al.*, 1989).

La construcción de las células huésped ha sido descrita (van der Vlag *et al.*, 2000). Brevemente, se basan en la línea celular de osteosarcoma humano U-2 OS (American Type Culture Collection HTB-96). U-2 OS es transfectada establemente con el plásmido pTet-Off (Clontech K1620-A), que codifica una quimera de proteína que consta del dominio de unión a ADN del represor Tet y el dominio de transactivación VP16. La línea celular es transfectada establemente con posterioridad con genes de la proteína de fusión que contienen el dominio de unión al ADN LexA, y las regiones codificadoras de HP1 o HPC2 (dos proteínas del grupo Polycomb de *Drosophila* que reprimen la expresión génica cuando se traban al ADN). Los genes represores de LexA están bajo el control del sistema regulador transcripcional Tet-Off (Gossen and Bujard, 1992).

Rastreo de la genoteca y caracterización del elemento STAR. Los genotecas génicas de pSelect son transfectadas en la línea celular U-2 OS/Tet-Off/represor LexA mediante precipitación con fosfato de calcio (Graham and van der Eb, 1973; Wigler *et al.*, 1978) como recomendaba el proveedor del reactivo de transfección (Life Technologies). Las células transfectadas son cultivadas bajo selección con higromicina (25 µg/ml) y represión con tetraciclina (doxiciclina, 10 ng/ml) durante 1 semana (confluencia 50%). Después la concentración de doxiciclina se reduce a 0,1 ng/ml para inducir los genes represores de LexA, y después de 2 días se añade zeocina a 250 µg/ml. Las células son cultivadas durante 4-5 semanas más, hasta que los cultivos de control (transfectados con pSelect vacío) son eliminados por la zeocina.

Las colonias resistentes a zeocina de esta transfección de la genoteca son propagadas, y el ADN plasmídico es aislado y recuperado en *E. coli* mediante mecanismos normalizados (Sambrook *et al.*, 1989). Los elementos STAR candidatos del ADN rescatado son analizados mediante mapeo con endonucleasas de restricción (Sambrook *et al.*, 1989), análisis de la secuencia de ADN (Sanger *et al.*, 1977), y actividad de STAR (resistencia a zeocina) después de la re-transfección a U-2 OS/Tet-Off/represor LexA y disminución de la concentración de doxiciclina.

Los elementos STAR candidatos que tienen una secuencia de ADN correspondiente a una secuencia conocida en el genoma humano son identificados mediante búsquedas BLAST (Altschul *et al.*, 1990) de la base de datos del genoma humano (<http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html> 20 Junio 2001). Las localizaciones cromosómicas de los elementos son registradas, junto con la proporción de ADN repetitivo y la identidad de los genes adyacentes.

Aquellos candidatos que muestran actividad STAR tras la re-transfección son caracterizados adicionalmente por la subclonación del fragmento STAR en el plásmido pSDH y la integración estable en el ADN cromosómico de U-2 OS. Los plásmidos de pSDH son co-transfectados en células U-2 OS con pBABE puro (Morgenstern and Land, 1990), y seleccionados en cuanto a la resistencia a la puomicina. Se aíslan y cultivan poblaciones de aproximadamente 30 clones individuales por elemento STAR. Los clones son analizados periódicamente en cuanto a la actividad luciferasa según las instrucciones del fabricante (Roche 1669893).

Resultados

Caracterización del elemento funcional STAR. Los rastreos de ADN genómico y de los loci HOX y 1q12 rindieron 17 elementos STAR genuinos. Los criterios son que (a) los elementos desplegaban actividad STAR tras la re-transfección de los clones basados en pSelect en la línea celular de osteosarcoma humano U-2 OS huésped (indicando que la actividad anti-represora expresada en el rastreo inicial es específica del plásmido y no debida a cambios por artefactos en las células huésped); (2) los elementos contienen una secuencia de ADN que coincide con la secuencia de la base de datos del genoma humano (indicando que el clon no contiene una secuencia de ADN contaminante, v.g. de fuentes bacterianas o de vectores).

Los elementos STAR son sub-clonados en el plásmido pSDH e integrados en el genoma de la célula huésped. La expresión de los genes informadores es analizada en poblaciones de transfectantes estables para demostrar la capacidad de los elementos STAR para proteger los genes informadores del silenciamiento después de la integración al azar en el genoma. Esto proporciona información (1) sobre la proporción de clones que despliegan una expresión elevada, y (2) sobre el grado de sobre-expresión logrado por los elementos STAR.

La expresión del gen informador de la luciferasa por un clon es considerada significativa si está dos veces por encima del nivel medio para los plásmidos que no contienen elementos STAR (nivel de referencia). Para todos los plásmidos se observa una distribución en el nivel de expresión entre los clones: desde la no expresión hasta una expresión significativamente por encima del nivel de referencia, y desde unos pocos sobre-expresadores a muchos sobre-expresadores. La actividad STAR superior es manifestada por los plásmidos que dan como resultado muchos clones de sobre-expresión, incluyendo algunos clones altamente sobre-expresados. Los resultados de un experimento representativo se muestran en la Tabla 1, y en las Figuras 3-5:

Los resultados indican que los elementos STAR humanos que son sometidos a ensayo rinden una proporción mucho mayor de clones de sobre-expresión que el gen informador no protegido, o el gen informador protegido por el elemento SCS de *Drosophila* (Kellum and Schedl, 1992). Además, el grado de sobre-expresión por estos plásmidos es mucho mayor a partir del gen informador protegido con STAR que del informador no protegido o protegido con SCS.

Secuencia del elemento STAR y datos de la posición genómica. En la Tabla 2 se enumeran las localizaciones cromosómicas de cada uno de los 17 elementos STAR, así como la identidad de los genes inmediatos y el contenido de ADN repetitivo de los elementos. Los elementos STAR se distribuyen a lo largo de numerosos cromosomas. Son diversos en su secuencia de ADN real y contenido en ADN repetitivo, y despliegan diversos grados de asociación con los genes vecinos.

Ejemplo 2

Características de la expresión del transgen que son debidas al STAR

Antecedentes: se utiliza la recombinación de sitio específico para separar precisamente los ADN heterólogos de sus localizaciones cromosómicas. Esto se lleva a cabo rutinariamente por medio de uno de dos sistemas: la recombinasa cre y la diana loxP del bacteriófago P1 (Feng *et al.*, 1999), o la recombinasa FLP y FRT (diana recombinasa FLP) de la levadura (Wigley *et al.*, 1994). En estos sistemas, una región de ADN (conteniendo normalmente un gen informador y/o un marcador seleccionable) está flanqueada en el cromosoma por la diana loxP o FRT. La actividad de la recombinasa cataliza después la separación por corte precisa de la región de ADN del cromosoma. La recombinasa resuelve sus dos secuencias de reconocimiento en un único sitio, suprimiendo la secuencia entre ellas. De este modo, un tramo de ADN debe estar flanqueado por sitios diana que van a ser suprimidos con posterioridad *in vivo* tras la introducción o activación de la recombinasa (Schwenk *et al.*, 1995; Dymecki, 1996). Las recombinasas Cre y Flp catalizan la recombinación entre dos repeticiones invertidas de 13 pares de bases, separadas por un espaciador con un mínimo de 6 (loxP) u 8 (FRT) pares de bases (Senecoff *et al.*, 1985). La secuencia loxP es ATAACCTTCGTATA y la secuencia FRT es GAAGTTCCTATAC.

Protocolo: Utilizando la clonación de ADN convencional (Sambrook *et al.*, 1989), se construye un gen informador (que codifica una proteína informadora, por ejemplo la proteína fluorescente verde (GFP) (Bierhuizen *et al.*, 1997) o la luciferasa (Himes and Shannon, 2000) que está flanqueado en un plásmido por un par de elementos STAR. En cada caso, los elementos están a su vez flanqueados por sitios diana de la recombinasa. Un elemento está flanqueado por un par de sitios loxP, y el otro está flanqueado por un par de sitios FRT (Figura 1). Tras la transfección el plásmido se integra en el cromosoma del huésped en un pequeño porcentaje de células, y los integrantes se seleccionan mediante resistencia a antibióticos.

Utilizando los mecanismos convencionales, ("SuperFect Transfection Reagent Handbook", Qiagen, Noviembre, 1997) estos plásmidos son transfectados en la línea celular de osteosarcoma humano U-2 OS, y seleccionados en cuanto a la resistencia a la higromicina. Los productos aislados resistentes a la higromicina tienen el plásmido integrado establemente en el genoma de la línea celular. Los productos aislados individuales son propagados en medio de cultivo celular, y la expresión del gen informador transgénico es analizada, por ejemplo mediante citometría de flujo (Stull *et al.*, 2000).

Después utilizando mecanismos convencionales (transfección, o estimulación con hormonas), los productos aislados anteriores son tratados con el fin de introducir o activar la actividad recombinasa. Esto se realiza sucesivamente, de manera que por ejemplo la actividad de la recombinasa cre catalice la separación por corte de STAR1, y con posterioridad la actividad recombinasa FLP catalice la separación por corte de STAR2. El nivel de expresión de este gen informador en estas células es analizado y el valor comparado con el valor de referencia del producto aislado que contiene STAR parental.

Ejemplo 3

Análisis de la secuencia de los STAR; determinación de la secuencia esencial mínima para la función del elemento; conservación de la secuencia entre elementos; y propiedades de elementos en tándem y múltiples

Antecedentes: Se aíslan fragmentos de ADN que contienen elementos STAR mediante selección genética utilizando el plásmido pSelect (Figura 1). En esta sección se describe el enfoque para caracterizar las secuencias de ADN dentro de aquellos fragmentos que tienen actividad STAR.

Protocolos:

Secuencia de ADN: Se diseñan oligonucleótidos basándose en la secuencia del plásmido pSelect para secuenciar los fragmentos de ADN. Los fragmentos son secuenciados utilizando la técnica de terminación de la cadena didesoxi (Sanger *et al.*, 1977). Las secuencias de ADN son localizadas después con respecto a su posición en el cromosoma utilizando la base de datos de la secuencia del genoma humano pública (http://www.ncbi.nlm.nih.gov:80/cgi-bin/Entrez/hum_srch?chr=hum chr.inf&query). Los genes y la densidad génica en la vecindad del fragmento son registrados a partir de la anotación de la secuencia del genoma. La actividad transcripcional de aquellos genes es determinada a partir de las bases de datos públicas de micromatrices de ADN (<http://arrays.rockefeller.edu/xenopus/links.html>) y datos SAGE (Serial Analysis of Gene Expression; <http://bioinfo.amc.uva.nl/HTM-bin/index.cgi>). Una vez recopilada la información posicional sobre las secuencias STAR, los datos son analizados en términos de las secuencias consenso subyacentes. Las secuencias consenso o las tendencias (por estas se entiende las áreas locales ricas en combinaciones de nucleótidos concretas, v.g. ricas en bases C y G) son detectadas utilizando algoritmos de búsqueda de la similitud tales como la puntuación de similitud clustalw (Higgins *et al.*, 1996) y blosum (Altschul and Gish, 1996). Después se utilizan cualquiera de los consensos subyacentes o tendencias encontrados para identificar otros STAR potenciales realizando búsquedas BLAST (Altschul *et al.*, 1990). La búsqueda previa ha identificado proteínas reguladoras de la transcripción que se unen a aislantes y elementos límite conocidos (Gaszner *et al.*, 1999; Gerasimova and Corces, 1998). En los ejemplos descritos antes, los sitios de unión a la proteína coinciden con los sitios hipersensibles a la DNasa I que son esenciales para la función del aislante o el límite. La hipótesis de que los elementos STAR también están unidos por proteínas reguladoras conocidas es examinada investigando la base de datos TRANSFAC de factores de transcripción (<http://transfac.gbf.de/TRANSFAC/>) en cuanto a los motivos secuenciales que existen en los elementos STAR. Los motivos secuenciales que son comunes entre los miembros de las colecciones de STAR son indicadores de que el factor de transcripción correspondiente se une a ese elemento.

Secuencia esencial mínima: Utilizando este conocimiento, los elementos STAR son truncados y sometidos a ensayo en cuanto a su funcionalidad. Esto se realiza utilizando la reacción en cadena de la polimerasa (PCR) para clonar sub-fragmentos de los fragmentos que contienen STAR en pSelect mediante mecanismos normalizados (Sambrook *et al.*, 1989). Los plásmidos que contienen los sub-fragmentos son transfectados en células U-2 OS y sometidos a ensayo en cuanto a la funcionalidad analizando la resistencia a antibióticos.

Direccionalidad: Los elementos STAR son sometidos a ensayo en cuanto a su direccionalidad utilizando el plásmido pSelect. Por ejemplo, la dirección de los elementos STAR aislados mediante el rastreo de pSelect es referida como orientación 5'3'. La orientación del elemento es invertida mediante mecanismos de recombinación de ADN convencionales (Sambrook *et al.*, 1989). Los plásmidos resultantes son transfectados en la línea celular U-2 OS y la expresión del gen informador es analizada (Bierhuizen *et al.*, 1997; Himes and Shannon, 2000). El nivel de expresión del plásmido con el elemento de orientación inversa es comparado con el de orientación 5'3'. Si el plásmido de orientación inversa tiene niveles de expresión similares, el elemento STAR no muestra direccionalidad.

Combinaciones y múltiplos de elementos: Para determinar si los elementos STAR son capaces de funcionar en pares mezclados, se combinan elementos diferentes y se someten a ensayo. El análisis se realiza en el plásmido pSDH insertando un elemento STAR en MCSI y un STAR diferente en MCSII mediante mecanismos de ADN recombinante (Sambrook *et al.*, 1989). Los plásmidos resultantes son transfectados, y la expresión del gen informador es analizada (Bierhuizen *et al.*, 1997; Himes and Shannon, 2000); los resultados son comparados con la expresión de plásmidos que contienen el mismo elemento en MCSI y MCSII; si la expresión es similar para los dos tipos de plásmidos, se concluye que los elementos STAR diferentes no interfieren entre sí.

La resistencia de los elementos STAR individuales se compara con las repeticiones en tándem de los elementos. Esto se realiza mediante concatamerización de los elementos STAR de interés con ADN ligasa e inserción del producto de ligadura en el plásmido pSDH mediante mecanismos de ADN recombinante (Sambrook *et al.*, 1989). Los plásmidos resultantes son transfectados en células U-2 OS, y la expresión del gen informador es analizada (Bierhuizen *et al.*, 1997; Himes and Shannon, 2000); Los resultados se comparan con la expresión de plásmidos que contienen elementos STAR individuales.

Ejemplo 4

Determinación de la distancia a lo largo de la cual funciona STAR

Antecedentes: se utilizan elementos STAR para optimizar la expresión de transgenes individuales y múltiples. Para determinar si un único par de elementos STAR puede proteger transgenes grandes o múltiples del silenciamiento es necesario determinar el intervalo a lo largo del cual actúan los elementos STAR.

Protocolo: Los elementos STAR son sometidos a ensayo en cuanto a su funcionalidad a lo largo de una distancia utilizando plásmidos derivados basados en pSelect, como sigue. Se ensambla una genoteca de fragmentos de ADN al azar de 500 pb a 10 kb mediante mecanismos de clonación de ADN normalizados (Sambrook *et al.*, 1989). De esta genoteca se seleccionan los fragmentos que no poseen actividad STAR, sometiéndolos a ensayo en el plásmido pSelect como se ha descrito antes. Para los elementos STAR, estos fragmentos son insertados entre el sitio de clonación y el promotor del gen informador en el plásmido pSelect apropiado (Figura 1). Este grupo de plásmidos es transfectado

a la línea celular U-2 OS, y la expresión es medida como se ha descrito antes. La resistencia de la expresión del gen informador se corresponde con la longitud del fragmento de ADN al azar que separa el elemento STAR y el promotor.

5 Ejemplo 5

Determinación de la longitud máxima de los elementos STAR

10 *Antecedentes:* los STAR son clonados como fragmentos de ADN recuperados utilizando el plásmido pSelect, lo que se realiza con fragmentos de ADN genómico de menos de 2 kb. No obstante, estos podrían ser porciones de un elemento STAR más ampliado. La actividad de STAR ampliada se examina mediante los siguientes experimentos.

15 *Protocolo:* los elementos STAR clonados en pSelect son mapeados para la secuencia del genoma humano. Con el fin de determinar si son porciones de un elemento STAR más ampliado, las regiones de 4 kb que abarcan los clones son amplificadas mediante PCR y clonadas en el plásmido pSelect y/o pSDH mediante mecanismos de ADN recombinante normalizados (Sambrook *et al.*, 1989). Los plásmidos resultantes son transfectados en células U-2 OS y analizados en cuanto a la expresión del gen informador como se ha descrito antes; los plásmidos que contienen el elemento STAR de 2 kb original son incluidos como control. Se pueden esperar tres posibles resultados: (1) una expresión similar por los productos aislados de control y ampliados, demostrando que el elemento STAR está confinado al fragmento de 2 kb original; (2) una expresión inferior por los productos aislados de STAR ampliado, sugiriendo que el elemento STAR está contenido en el fragmento de 2 kb y no actúa eficazmente a lo largo de una distancia o que el fragmento ampliado contiene un elemento con una cualidad represora de la transcripción génica; (3) una expresión superior por los productos aislados de STAR ampliado, sugiriendo que la región ampliada contiene un elemento STAR más completo. En el caso del resultado (3), el ejercicio se reitera con un fragmento de PCR más grande de 6 kb.

25 Un elemento STAR también puede estar compuesto por sitios a los cuales se unen diversas proteínas. Por lo tanto los fragmentos de ADN grandes con una actividad STAR podrían ser divisibles en fragmentos más pequeños con actividad STAR (ver ejemplo 3). Los elementos que son mayores de 2 kb son reconocidos como elementos STAR si todavía presentan actividad STAR tras el truncamiento a menos de 2 kb (incluyendo por medio de delección interna).

30 Ejemplo 6

Estados de metilación y acetilación de histona de elementos STAR y de transgenes adyacentes

35 *Antecedentes:* Las propiedades reguladoras de los elementos STAR están asociadas con la estructura de la cromatina local, que es determinada por el propio ADN y por las proteínas asociadas al ADN. Los cambios en la estructura de la cromatina que están asociados con los cambios en la expresión génica son producidos a menudo por modificaciones secundarias de las macromoléculas, especialmente la metilación del ADN o la acetilación de las proteínas histónicas. La identificación de las modificaciones secundarias que se producen en los elementos STAR y en los transgenes adyacentes proporciona sellos para estos elementos.

40 *Protocolo:* Metilación de ADN: los elementos STAR son clonados en el plásmido pSelect mediante mecanismos normalizados (Sambrook *et al.*, 1989). Las células U-2 OS son transfectadas establemente con estos plásmidos, y con pSelect que carece de un elemento STAR como control para determinar la metilación de ADN basal en el gen informador.

45 Las células son cosechadas y la cromatina purificada mediante procedimientos normalizados (Thomas, 1998). El ADN es digerido con las endonucleasas de restricción *HpaII* y *MspI* en reacciones separadas (Sambrook *et al.*, 1989). Ambas enzimas de restricción son capaces de cortar la secuencia no metilada CCGG. Cuando la C externa está metilada, ni *MspI* ni *HpaII* pueden escindir. No obstante, a diferencia de *HpaII*, *MspI* puede escindir la secuencia cuando la C interna está metilada. El ADN es sometido a transferencia Southern y la transferencia es analizada mediante marcaje terminal indirecto (Pazin and Kadonaga, 1998). Como control, el correspondiente plásmido pSelect en forma de ADN no metilado, desnudo, también es cortado con las enzimas descritas y sometido a transferencia Southern. La comparación de los diferentes tamaños de los fragmentos de ADN revela si el ADN está metilado *in vivo* o no.

50 *Acetilación de las histonas:* Las mismas líneas celulares transfectadas utilizadas para el análisis de metilación del ADN se utilizan para estos experimentos. El método descrito más abajo rinde un mapa de alta resolución del patrón de acetilación de la histona sobre los elementos STAR y el gen informador (Litt *et al.*, 2001). Los productos digeridos con nucleasa Micrococcica de los núcleos son fraccionados sobre gradientes de sacarosa, y los monómeros y dímeros de nucleosoma purificados son enriquecidos en histonas acetiladas mediante inmunoprecipitación con anticuerpos anti-acetilhistona. La fracción de nucleosoma y los productos inmunoprecipitados son sometidos a análisis, por ejemplo, mediante PCR en tiempo real (Jung *et al.*, 2000) utilizando cebadores y una sonda Taqman que hibrida con el gen informador o el elemento STAR para rendir productos de 0,2 kb, con una ventana móvil de 0,1 kb. La tasa de incremento de la señal fluorescente de la sonda Taqman durante la PCR (que es proporcional a la abundancia del ADN molde en la muestra) es medida después. La proporción de la abundancia del ADN molde en la fracción de nucleosoma y los productos inmunoprecipitados proporciona una mapa fino del patrón de acetilación de las histonas para cada 0,1 kb sobre el gen informador y el elemento STAR (o sobre el gen informador en ausencia de un elemento).

Ejemplo 7

Posicionamiento del nucleosoma In vivo y sitios hipersensibles a la ADNasa I

Antecedentes: La cromatina está formada por ADN, histonas, y proteínas no histónicas. Las histonas forman una partícula núcleo que está envuelta por ~150 pb de ADN para formar el nucleosoma. Los nucleosomas están separados por 50-75 pb de ADN conector.

Los nucleosomas posicionados establemente sobre el ADN cromosómico reprimen la expresión génica, y los factores que excluyen los nucleosomas o la cromatina remodelada de otro modo pueden superar esta represión. El posicionamiento de los nucleosomas en una región cromosómica es analizado mediante análisis con nucleasa microcólica (MNasa); la MNasa corta la cromatina preferentemente en el ADN conector. De un modo similar, algunas áreas de ADN están constitutivamente expuestas a las proteínas no histónicas, y estas son frecuentemente regiones reguladoras, es decir, sitios donde se unen los factores reguladores que actúan en cis. Experimentalmente, estos sitios son hipersensibles a la digestión por la enzima ADNasa I.

Protocolo: Para determinar la posición de los nucleosomas sobre el gen informador y sobre los elementos STAR, se utiliza la MNasa (Saluz and Jost, 1993). Los núcleos son purificados a partir de las células U-2 OS cultivadas y digeridas con MNasa como se ha descrito antes (acetilación de histonas). Para investigar los sitios hipersensibles a la ADNasa I en los elementos STAR o el gen informador, los núcleos purificados son tratados con ADNasa I a una concentración apropiada (v.g. 100 µg/ml de ADN genómico en 20-100 U/ml de ADNasa I), como se ha descrito (Wallrath *et al.*, 1998). El ADN desnudo es digerido con ADNasa I como control. Para ambas técnicas, el gen informador y los elementos STAR son sometidos a mapeo fino utilizando la prolongación del cebador o el marcaje terminal indirecto y la transferencia Southern, como se ha descrito (Tanaka *et al.*, 1996; van der Vlag *et al.*, 2000). El análisis de la MNasa revela una escala de bandas discretas en un autorradiograma correspondiente a las posiciones de los nucleosomas sobre los elementos STAR o el gen informador. Los sitios hipersensibles a la ADNasa I se manifiestan como bandas discretas en el autorradiograma resultante que están ausentes o son menos prominentes en el control de ADN desnudo.

Ejemplo 8

Tipo celular, dependencia del tejido, y dependencia del promotor de los elementos STAR

Antecedentes: Se ha informado de que algunos elementos aislantes o límite pueden mostrar especificidad hacia el tejido (Takada *et al.*, 2000). Los elementos STAR tienen muchos rasgos en común con los elementos aislantes y límite. Tanto los elementos STAR promiscuos como los específicos de tejidos tienen un valor biotecnológico en las aplicaciones transgénicas. El análisis descrito más abajo se realiza para evaluar la dependencia del tipo de célula. La especificidad de la célula y del tejido de los elementos es examinada adicionalmente examinando la expresión de los genes en la vecindad de los elementos en el genoma humano, utilizando las bases de datos públicas de micromatrices de ADN (<http://arrays.rockefeller.edu/xenopus/links.html>) y datos SAGE (Serial Analysis of Gene Expression; <http://bioinfo.amc.uva.nl/HTM-bin/index.cgi>).

Protocolo: los elementos STAR son sometidos a ensayo en el plásmido pSDH. Se transfectan tres líneas celulares utilizando protocolos normalizados: la línea celular de osteosarcoma U-2 OS (Heldin *et al.*, 1986), la línea celular Vero de riñón de mono verde Africano (Simizu *et al.*, 1967), y la línea celular CHO de ovario de hámster Chino (Kao and Puck, 1968). Los elementos capaces de funcionar en los tres tipos celulares son categorizados como promiscuos. Aquellos que solamente despliegan actividad en una o dos de las líneas celulares se categorizan como restringidos en su funcionalidad en el tipo celular.

Especificidad del promotor: los elementos STAR son actualmente seleccionados y sometidos a ensayo en el contexto de la función con dos promotores, el promotor de citomegalovirus completo (CMV) o el Elemento de Respuesta a la Tetraciclina y el promotor mínimo de CMV (combinado con el activador transcripcional tTA). Para evaluar la especificidad del promotor, la función STAR es sometida a ensayo con otros promotores virales comúnmente utilizados, a saber los promotores temprano y tardío del virus de simios de tipo 40 (SV40), y los promotores principal y tardío de EIA adenoviral, y la repetición larga terminal del virus del sarcoma de Rous (RSV) (Doll *et al.*, 1996; Smith *et al.*, 2000; Weaver and Kadan, 2000; Xu *et al.*, 1995). Cada uno de estos promotores es clonado por separado en el plásmido pSelect mediante mecanismos normalizados (Sambrook *et al.*, 1989) junto con los elementos STAR. Los plásmidos resultantes son transfectados en la línea celular U-2 OS y analizados en cuanto a la expresión del gen informador, como se ha descrito antes. La capacidad de los elementos STAR para protegerse frente al silenciamiento, es determinada mediante comparación con los plásmidos que carecen de elementos STAR.

Ejemplo 9

Métodos para mejorar los elementos STAR

Antecedentes: Se desarrollan elementos STAR mejorados. Las mejoras rindieron una actividad anti-represiva de resistencia incrementada, y elementos con actividad inducible y específica del tejido. Estas mejoras se realizan por medio de una combinación de técnicas.

Protocolos:

Evolución forzada: Se utiliza una PCR propensa a errores (Cherry *et al.*, 1999; Henke and Bornscheuer, 1999) para introducir una media de una o dos mutaciones puntuales por elemento. Los elementos mutagenizados son rastreados utilizando plásmidos pSelect conteniendo proteínas de fusión informador-marcador seleccionable por ejemplo mediante clasificación celular activada por fluorescencia y resistencia a antibióticos (Bennett *et al.*, 1998). Posteriores rondas de PCR propensas a errores y selección se llevan a cabo para derivar elementos con mejoras adicionales en la actividad.

Combinaciones en tándem y heterólogas: Como se ha descrito antes, las combinaciones en tándem y heterólogas de los elementos son sometidas a ensayo en cuanto a la actividad en comparación con los elementos individuales (ejemplo 3).

La dominancia relativa de los elementos STAR es sometida a ensayo caso por caso. Esto se utiliza para someter a ensayo la resistencia de un elemento; por ejemplo, si un nuevo elemento STAR es dominante para un elemento fuerte, conocido con una cualidad represora de la transcripción génica, en ese caso el STAR es clasificado como muy fuerte. También se considera la posibilidad de que la relación de dominancia entre un STAR y un elemento con una cualidad de represión de la transcripción génica sea específica del tipo celular, del tejido, o del promotor (ejemplo 8). En el ensayo de dominancia se utiliza el plásmido pSelect, con elementos individuales con una cualidad represora de la transcripción génica situada aguas arriba de los elementos STAR individuales mediante mecanismos de ADN recombinante (Sambrook *et al.*, 1989). Los plásmidos son transfectados a células U-2 OS y la expresión del gen informador es analizada. La dominancia de STAR es manifestada por una expresión superior que el plásmido con sólo un elemento con una cualidad represora de la transcripción génica.

Introducción de sitios de unión para otras proteínas de unión a ADN a elementos STAR para añadir características novedosas (v.g. inducibilidad, especificidad de tejido).

Antecedentes: se crean elementos STAR regulables combinándolos con sitios de unión para proteínas de unión al ADN dependientes de la señal. En un ejemplo esto implicaría la yuxtaposición de un STAR y un elemento de respuesta a glucocorticoides (GRE). En ausencia de estimulación por glucocorticoides el elemento STAR funcionaría como se describe. Tras la estimulación, el receptor de glucocorticoides de origen natural se une a GRE e interfiere en la función de STAR.

Protocolo: Utilizando la clonación de ADN convencional (Sambrook *et al.*, 1989), se introduce un GRE en el vector pSelect adyacente a los elementos STAR. El plásmido es transfectado en células U-2 OS como se ha descrito antes. Las células se dividen en dos cultivos; uno se trata con glucocorticoides (10 μ M). La expresión del gen informador se mide y se compara entre los dos cultivos. Las diferencias en la expresión demuestran la capacidad para regular la función de STAR mediante la acción de una proteína de unión a ADN dependiente de la señal.

Elementos STAR promiscuos: El ensayo o la intensificación de estas características implica el cultivo en diferentes líneas celulares, y el cultivo a largo plazo sin selección con antibiótico (ejemplos 8 y 10).

Ejemplo 10

Los elementos STAR obvian la necesidad de una selección continua para el mantenimiento del transgen

Antecedentes: En la transgénesis, la dependencia de los marcadores de selección tiene dos desventajas: el agente de selección es normalmente caro y conlleva un coste metabólico para las células, y existen objeciones regulatoras y éticas para incluir marcadores seleccionables en aplicaciones transgénicas, especialmente si el propio transgen está en el producto (v.g. plantas de cultivo, vectores de terapia génica). Los elementos STAR reducen o eliminan la necesidad de mantener la selección después de establecer el producto aislado transgénico. Por consiguiente, el gen de resistencia puede ser separado del genoma transgénico mediante recombinación de sitio específico con una pérdida disminuida de la expresión del transgen.

Protocolo: Se producen líneas celulares U-2 OS transfectadas establemente conteniendo elementos STAR integrados cromosómicamente flanqueando genes informadores por medio de co-transfección del plásmido pSDH con un plásmido de resistencia a antibióticos que actúan en trans como se ha descrito antes. El experimento implica someter a ensayo la estabilidad del nivel de expresión del gen informador en estas líneas celulares durante un cultivo prolongado (3-6 meses) en ausencia de selección. Este se somete a ensayo con elementos STAR flanqueando los genes informadores de la luciferasa o GFP en plásmidos pSDH. El gen de resistencia a antibióticos es separado construyendo un plásmido de expresión (basado en pSDH) en el cual el marcador de selección del antibiótico está flanqueado por sitios diana para la recombinasa. El marcador seleccionable es separado por corte con posterioridad por medio de la actividad de la recombinasa, como se ha descrito antes (ejemplo 2).

Ejemplo 11

La Pronosticabilidad y el rendimiento resultan mejorados por la aplicación de elementos STAR a los sistemas de expresión

Los elementos STAR funcionan bloqueando las influencias de la represión transcripcional sobre las unidades de expresión de los transgenes. Estas influencias de represión pueden estar debidas a la heterocromatina ("efectos sobre la posición", (Boivin & Dura, 1998)) o a copias adyacentes del transgen ("silenciamiento génico inducido por repeticiones", (Garrick *et al.*, 1998)). Dos de los beneficios de los elementos STAR para la producción de proteínas heterólogas son el incremento de la pronosticabilidad del descubrimiento de células huésped recombinantes primarias con un elevado nivel de expresión y el incremento del rendimiento durante los ciclos de producción. Estas ventajas se ilustran en este ejemplo.

Materiales y Métodos

Construcción de vectores pSDH y derivados que contienen STAR: El vector pSDH-Tet fue construido mediante amplificación por reacción en cadena de la polimerasa (PCR) del marco de lectura abierto de la luciferasa a partir del plásmido pREP4-HSF-Luc (van der Vlag *et al.*, 2000) utilizando los cebadores C67 y C68 (todos los cebadores de la PCR y oligonucleótidos mutagénicos se enumeran en la Tabla 5), e inserción del fragmento SacII/BamHI en pUHD10-3 digerido con SacII/BamHI (Gossen & Bujard, 1992). La unidad de expresión de la luciferasa fue re-amplificada con los cebadores C65 y C66, y re-insertada en pUHD10-3 con el fin de flanquearla con dos sitios de clonación múltiple (MCSI and MCSII). Después se introdujo un sitio AscI en MCSI mediante digestión con EcoRI e inserción de un conector (que constaba de oligonucleótidos hibridados D93 y D94). El promotor de CMV fue amplificado a partir del plásmido pCMV-Bsd (Invitrogen K510-01) con los cebadores D90 y D91, y utilizado para remplazar el promotor Tet-Off en pSDH-Tet mediante digestión con SalI/SacII y ligadura para crear el vector pSDH-CMV. El marco de lectura abierto de la luciferasa en este vector fue remplazado por SEAP (Fosfatasa Alcalina Secretada) como sigue: el vector pSDH-CMV fue digerido con SacII y BamHI y sus extremos fueron convertidos en romos; el marco de lectura abierto SEAP fue aislado de pSEAP-basic (Clontech 6037-1) por medio de digestión con EcoRI/SalI, sus extremos fueron convertidos en romos y fue ligado en pSDH-CMV para crear el vector pSDH-CS. El gen de resistencia a la puomicina bajo el control del promotor de SV40 fue aislado del plásmido pBabe-Puro (Morgenstern & Land, 1990) mediante PCR, utilizando los cebadores C81 y C82. Este fue ligado en el vector pGL3-control (sitio BamHI eliminado) (Promega E1741) digerido con NcoI/XbaI, para crear pGL3-puro. pGL3-puro fue digerido con BglII/SalI para aislar el gen de resistencia a SV40 puro, sus extremos fueron convertidos en romos y fue ligado en pSDH-CS digerido con NheI, con los extremos romos. El vector resultante, pSDH-CSP, se muestra en la Fig 6. Todas las etapas de clonación se llevaron a cabo siguiendo las instrucciones proporcionadas por los fabricantes de los reactivos, según los métodos conocidos en la técnica (Sambrook *et al.*, 1989).

Los elementos STAR fueron insertados en MCSI y MCSII en dos etapas, mediante digestión del elemento STAR y el vector pSDH-CSP con una enzima de restricción apropiada, seguido de ligadura. La orientación de los elementos STAR en los vectores pSDH recombinantes fue determinada mediante mapeo de restricción. La identidad y la orientación de los insertos fueron verificadas mediante análisis de la secuencia de ADN. La secuenciación fue realizada mediante el método didesoxi (Sanger *et al.*, 1977) utilizando un secuenciador de ADN automatizado Beckman CEQ2000, según las instrucciones del fabricante. En resumen, el ADN fue purificado de *E. coli* utilizando QIAprep Spin Miniprep and Plasmid Midi Kits (QIAGEN 27106 y 12145, respectivamente). La secuenciación del ciclo se llevó a cabo utilizando los oligonucleótidos C85, E25, y E42 (Tabla 5) de costumbre, en presencia de terminadores coloreados (CEQ Dye Terminator Cycle Sequencing Kit, Beckman 608000).

Transfección y cultivo de células CHO con plásmidos pSDH: La línea celular de Ovario de Hámster Chino CHO-K1 (ATCC CCL-61) en medio HAMS-F12 + Suero de Ternera Fetal al 10% conteniendo glutamina 2 mM, 100 U/ml de penicilina, y 100 microgramos/ml de estreptomicina a 37°C/CO₂ al 5%. Las células fueron transfectadas con el vector pSDH-CSP, y sus derivados conteniendo STAR6 o STAR49 en MCSI y MCSII, utilizando SuperFect (QIAGEN) como describe el fabricante. En resumen, las células fueron sembradas en recipientes de cultivo y se hicieron crecer durante la noche hasta una confluencia del 70-90%. El reactivo SuperFect se combinó con el ADN plasmídico (linealizado en este ejemplo mediante digestión con PvuI) a una proporción de 6 microlitros por microgramo (v.g. para una placa Petri de 10 cm, 20 microgramos de ADN y 120 microlitros de SuperFect) y se añadieron a las células. Tras incubación durante la noche la mezcla de transfección fue remplazada por medio de nueva aportación, y las células transfectadas fueron incubadas adicionalmente. Tras cultivar durante la noche, se añadieron 5 microgramos/ml de puomicina. La selección con puomicina se completó en 2 semanas, después de lo cual se aislaron clones CHO/pSDH resistentes a la puomicina individuales al azar y se cultivaron adicionalmente.

Análisis con Fosfatasa Alcalina Secretada (SEAP): La actividad SEAP (Berger *et al.*, 1988, Henthorn *et al.*, 1988, Kain, 1997, Yang *et al.*, 1997) en el medio de cultivo de los clones CHO/pSDH-CSP fue determinada como describe el fabricante (estuche Clontech Great EscAPe #K2041). Brevemente, una alícuota del medio fue inactivada con calor a 65°C, después fue combinada con tampón de análisis y sustrato quimioluminiscente CSPD e incubada a la temperatura ambiente durante 10 minutos. La tasa de conversión de sustrato fue determinada después en un luminómetro (Turner 20/20TD). La densidad celular fue determinada sometiendo a recuento las células tratadas con tripsina en un contador celular Coulter ACT10.

Transfección y cultivo de células U-2 OS con plásmidos pSDH: La línea celular de osteosarcoma humano U-2 OS (ATCC #HTB-96) fue cultivada en Medio de Eagle Modificado de Dulbecco + Suero de Ternera Fetal al 10% conteniendo glutamina, penicilina, y estreptomina (*supra*) a 37°C/CO₂ al 5%. Las células fueron co-transfectadas con el vector pSDH-CMV vector, y sus derivados conteniendo STAR6 o STAR8 en MCSII y MCSII, (junto con el plásmido pBabe-Puro) utilizando SuperFect (*supra*). La selección con Puomicina se completó en 2 semanas, tiempo tras el cual se aislaron clones U-2 OS/pSDH-CMV resistentes a la puomicina al azar y se cultivaron adicionalmente.

Análisis con luciferasa: la actividad luciferasa fue analizada (Himes & Shannon, 2000) en células resuspendidas según las instrucciones del fabricante del estuche de análisis (Roche 1669893), utilizando un luminómetro (Turner 20/20TD). La concentración de proteína celular total fue determinada mediante el método del ácido bicínico según las instrucciones del fabricante (Sigma B-9643), y utilizada para normalizar los datos de la luciferasa.

Resultados

Los clones de las células CHO recombinantes que contienen el vector pSDH-CSP, o los plásmidos pSDH-CSP que contienen STAR6 o STAR49 (Tabla 6), fueron cultivados durante 3 semanas. La actividad SEAP de los sobrenadantes de cultivo fue determinada después, y es expresada basándose en el número de células (Fig. 7). Como se puede observar, se aislaron clones con elementos STAR en las unidades de expresión que expresan una actividad SEAP 2-3 veces superior que los clones cuyas unidades de expresión no incluían elementos STAR. Además, el número de clones que contienen elementos STAR que expresan actividad SEAP a la actividad máxima o por encima de ella de los clones STAR negativos es bastante elevado: del 25% al 40% de las poblaciones de clones STAR excedían la expresión SEAP más elevada de los clones pSDH-CSP.

Los clones de las células U-2 OS recombinantes que contenían el vector pSDH-CMV, o los plásmidos pSDH-CMV que contenían STAR6 o STAR8 (Tabla 6), fueron cultivados durante 3 semanas. La actividad luciferasa de las células huésped fue determinada después, y expresada como unidades de luciferasa relativas (Fig. 8), normalizadas para la proteína celular total. Los clones de U-2 OS recombinantes con elementos STAR flanqueando las unidades de expresión que tenían rendimiento mayores que los clones STAR negativos: la expresión más elevada observada a partir de los clones STAR8 era 2-3 veces superior que la expresión observada a partir de los clones STAR negativos. Los clones STAR6 tenían niveles de expresión máxima 5 veces superiores que los clones STAR negativos. Los elementos STAR conferían también una mayor pronosticabilidad: para ambos elementos STAR, de 15 al 20% de los clones presentaban expresión luciferasa a niveles comparables o mayores que los clones STAR negativos con el nivel de expresión más elevado.

Estos resultados demuestran que, cuando se utilizan con un promotor CMV potente, los elementos STAR incrementan el rendimiento de proteínas heterólogas (luciferasa y SEAP). Los tres elementos STAR introducidos en este ejemplo proporcionan rendimientos elevados. El aumento de pronosticabilidad conferido por los elementos STAR es manifestado por la gran proporción de los clones con rendimiento iguales o mayores que los rendimientos más elevados presentados por los clones STAR negativos.

Ejemplo 12

Los elementos STAR mejoran la estabilidad de la expresión del transgen

Durante el cultivo de las células huésped recombinantes, es una práctica común mantener la selección mediante antibióticos. Esto está destinado a evitar el silenciamiento transcripcional del transgen, o la pérdida del transgen a partir del genoma mediante procedimientos tales como la recombinación. No obstante no es deseable para la producción de proteínas heterólogas, por numerosas razones. Primero, los antibióticos que se utilizan son bastante caros, y contribuyen significativamente al coste unitario del producto. Segundo, para el uso biofarmacéutico, la proteína debe ser demostrablemente pura, sin trazas del antibiótico en el producto. Una ventaja de los elementos STAR para la producción de proteínas heterólogas es que confieren una expresión estable a los transgenes durante el cultivo prolongado, incluso en ausencia de selección con antibiótico; esta propiedad se demuestra en este ejemplo.

Materiales y Métodos

La línea celular U-2 OS fue transfectada con el plásmido pSDH-Tet-STAR6 y cultivada como se describe en el Ejemplo 11. Los clones resistentes a la puomicina individuales fueron aislados y cultivados adicionalmente en ausencia de doxiciclina. A intervalos semanales las células fueron transferidas a recipientes de cultivo nuevos a una dilución 1:20. La actividad luciferasa fue media a intervalos periódicos como se describe en el Ejemplo 11. Al cabo de 15 semanas los cultivos fueron divididos en dos productos replicados; un producto replicado continuó recibiendo puomicina mientras el otro producto replicado no recibió antibiótico durante el resto del experimento (25 semanas en total).

Resultados

La Tabla 7 presenta los datos de expresión de la luciferasa mediante la unidad de expresión flanqueada por STAR6 durante el crecimiento prolongado con o sin antibióticos. Como se puede observar, la expresión del transgen informador, luciferasa, permanece estable en las células huésped U-2 OS durante el experimento. Una vez que los cultivos se dividieron en dos tratamientos (más antibiótico y sin antibiótico) la expresión de la luciferasa era esencialmente estable en ausencia de selección con antibiótico. Esto demuestra la capacidad de los elementos STAR para proteger los transgenes del silenciamiento o pérdida durante el cultivo prolongado. También demuestra que esta propiedad es independiente de la selección con antibiótico.

Por lo tanto la producción de proteínas heterólogas es posible sin que incurra en los costes del antibiótico o la dificultad del procesamiento aguas abajo.

Ejemplo 13

Secuencias esenciales mínimas de elementos STAR

Los elementos STAR fueron aislados del rastreo genético descrito en el Ejemplo 1. En el rastreo se utilizan genotipas construidas con ADN genómico humano que estaba fraccionado por tamaño entre aproximadamente 0,5-2 kilobases (*supra*). Los elementos STAR oscilan entre 500 y 2.361 pb. (Tabla 6). Es probable que, para muchos de los elementos STAR que han sido aislados, la actividad STAR sea conferida por un fragmento de ADN más pequeño que el clon aislado inicialmente. Es útil determinar estos tamaños de fragmentos mínimos que son esenciales para la actividad STAR, por dos razones.

Primera, los elementos STAR funcionales más pequeños serían ventajosos en el diseño de vectores de expresión compactos, puesto que los vectores más pequeños transfectan las células huésped con mayor eficacia.

Segunda, determinar las secuencias STAR esenciales mínimas permite la modificación de aquellas secuencias para una funcionalidad intensificada. Dos elementos STAR han sido sometidos a mapeo fino para determinar sus secuencias esenciales mínimas.

Materiales y métodos

STAR10 (1.167 pares de bases) y STAR27 (1.520 pares de bases) han sido sometidos a mapeo fino. Han sido amplificados mediante PCR para rendir sub-fragmentos de una longitud aproximadamente igual (leyenda de la Fig. 9). Para el ensayo inicial, estos han sido clonados en el vector pSelect en el sitio BamHI, y transfectados en células U-2 OS/Tet-Off/LexA-HP1 como se describe en el Ejemplo 1. Tras la selección en cuanto a la resistencia a la higromicina, LexA-HP1 fue inducido disminuyendo la concentración de doxiciclina. Las células transfectadas fueron incubadas después con zeocina para someter a ensayo la capacidad de los elementos STAR para proteger la unidad de expresión SV40-Zeo de la represión debida a la unión a LexA-HP1.

Resultados

En este experimento STAR10 y STAR27 confieren una buena protección frente al silenciamiento génico, como se esperaba (Fig. 9). Esto se manifiesta por un fuerte crecimiento en presencia de zeocina.

De los 3 sub-fragmentos de STAR10, 10A (~400 pares de bases) confiere a las células transfectadas un crecimiento vigoroso en presencia de zeocina, que excede el del elemento STAR completo. Las células transfectadas con los constructos pSelect que contenían los otros 2 sub-fragmentos no crecen en presencia de zeocina. Estos resultados identifican el fragmento 10A de ~400 pares de bases que abarca la secuencia de ADN responsable de la actividad anti-represión de STAR10. STAR 27 confiere un crecimiento moderado en zeocina a las células transfectadas en este experimento (Fig. 9). Uno de los sub-fragmentos de este STAR, 27B (~500 pares de bases), permite un crecimiento débil de las células huésped en el medio que contiene zeocina. Esto sugiere que la actividad anti-represión de este STAR está parcialmente localizada en el sub-fragmento 27B, pero la actividad completa requiere también las secuencias 27A y/o 27C (cada una de ~ 500 pares de bases).

Ejemplo 14

Los elementos STAR funcionan en diversas cepas de células de mamífero cultivadas. La elección de la línea de células huésped para la expresión de la proteína heteróloga es un parámetro crítico para la cualidad, el rendimiento, y el coste por unidad de la proteína. Consideraciones tales como las modificaciones post-traduccionales, la capacidad de la ruta secretora, y la inmortalidad de la línea celular dictan la línea celular apropiada para un sistema de producción biofarmacéutico concreto. Por esta razón, las ventajas proporcionadas por los elementos STAR en términos de rendimiento, pronosticabilidad, y estabilidad deben ser obtenibles en diversas líneas celulares. Esto fue sometido a ensayo comparando la función de STAR6 en la línea celular U-2 OS humana en la cual éste había sido clonado originalmente, y en la línea celular CHO que es ampliamente aplicada en biotecnología.

Material y Métodos

Se refieren los experimentos del ejemplo 11.

Resultados

La expresión del gen informador SEAP en células CHO es presentada en la Fig. 7; la expresión del gen informador de luciferasa en las células U-2 OS se presenta en la Fig. 8. Por medio de la comparación de los resultados de estos dos experimentos, es evidente que el elemento STAR6 es funcional en ambas líneas celulares: la expresión del gen informador era más predecible en ambas, y los clones de cada línea celular presentaban rendimientos más altos, cuando el gen informador era protegido de los efectos de la posición por STAR6. Estas dos líneas celulares derivan de diferentes especies (humano y hámster) y diferentes tipos de tejidos (hueso y ovario), reflejando la amplia gama de huéspedes en la cual puede ser utilizado este elemento STAR para la mejora de la expresión de proteínas heterólogas.

Ejemplo 15

Los elementos STAR funcionan en el contexto de diversos promotores transcripcionales. La transcripción del transgen se realiza colocando el marco de lectura abierto del transgen bajo el control de un promotor exógeno. La elección del promotor está influenciada por la naturaleza de la proteína heteróloga y el sistema de producción. En la mayoría de los casos, se prefieren los promotores constitutivos fuertes debido a que pueden proporcionar elevados rendimientos. Algunos promotores virales tienen estas propiedades; promotor/intensificador del gen temprano inmediato de citomegalovirus ("promotor de CMV") se considera generalmente como el promotor más fuerte en uso biotecnológico común (Boshart *et al.*, 1985, Doll *et al.*, 1996, Foecking & Hofstetter, 1986). El promotor del virus de simios SV40 también es moderadamente fuerte (Boshart *et al.*, 1985, Foecking & Hofstetter, 1986) y es utilizado frecuentemente para la expresión ectópica en vectores de células de mamífero. El promotor Tet-Off es inducible: el promotor es reprimido en presencia de tetraciclina o antibióticos relacionados (comúnmente se utiliza doxiciclina) en líneas celulares que expresan el plásmido tTA (Clontech K1620-A), y la eliminación del antibiótico produce una inducción transcripcional (Deuschle *et al.*, 1995, Gossen & Bujard, 1992, Izumi & Gilbert, 1999, Umana *et al.*, 1999).

Material y Métodos

La construcción de los vectores pSDH-Tet y pSDH-CMV se describe en el Ejemplo 11. pSDH-SV40 fue construido mediante amplificación por PCR del promotor SV40 (cebadores D41 y D42) del plásmido pSelect-SV40-Zeo (Ejemplo 1), seguido de digestión del producto de la PCR con SacII y SalI. El vector pSDH-CMV fue digerido con SacII y SalI para separar el promotor de CMV, y el vector y el fragmento de SV40 fueron ligados entre sí para crear pSDH-SV40. STAR 6 fue clonado en MCSI y MCSII como se describe en el Ejemplo 11. Los plásmidos pSDH-Tet, pSDH-Tet-S TAR 6, pSDH-Tet-STAR7, pSDH-SV40 y pSDH-SV40-STAR6 fueron co-transfectados con pBabe-Puro en U-2 OS utilizando SuperFect como describe el fabricante. El cultivo celular, la selección con puomicina, y los análisis con luciferasa se llevaron a cabo como se describe en el Ejemplo 11.

Resultados

Las Figs. 8, 10, y 11 comparan la expresión del gen informador luciferasa de 3 promotores diferentes: dos promotores virales fuertes y constitutivos (CMV y SV40), y el promotor Tet-Off inducible. Los tres promotores fueron sometidos a ensayo en el contexto del elemento STAR6 en células U-2 OS. Los resultados demuestran que el rendimiento y la pronosticabilidad de los 3 promotores aumentaban con STAR6.

Como se describe en los Ejemplo 11 y 14, STAR6 es beneficioso en el contexto del promotor CMV (Fig. 8). Se observan mejoras similares en el contexto del promotor de SV40 (Fig. 10): el rendimiento del clon STAR6 más altamente expresado es 2-3 veces mayor que el mejor de los clones pSDH-SV40, y los clones STAR 6 (20% de la población) tienen rendimientos más altos que el mejor de los clones STAR negativos. En el contexto del promotor Tet-Off a concentraciones inductoras (poca doxiciclina), STAR6 también mejora el rendimiento y la pronosticabilidad de la expresión del transgen (Fig. 11): el clon STAR6 más altamente expresado tiene un rendimiento 20 veces superior que el mejor clon pSDH-tet, y los 9 clones STAR6 (35% de la población) tienen rendimientos más altos que el mejor clon STAR negativo. Se concluye que este elemento STAR es versátil en sus propiedades protectoras del transgen, puesto que funciona en el contexto de diversos promotores de la transcripción biotecnológicamente útiles.

Ejemplo 16

La función del elemento STAR puede ser direccional

Si bien las secuencias de ácido nucleico cortas pueden ser simétricas (v.g. palindrómicas), las secuencias de origen natural más largas son típicamente asimétricas. Como resultado, el contenido de información de las secuencias de ácido nucleico es direccional, y las propias secuencias pueden ser descritas con respecto a sus extremos 5' y 3'. La direccionalidad de la información de la secuencia de ácido nucleico afecta a la disposición en la cual son ensambladas las moléculas de ADN recombinante utilizando mecanismos de clonación normalizados conocidos en la técnica

(Sambrook *et al.*, 1989). Los elementos STAR son secuencias de ADN largas, asimétricas, y tienen una direccionalidad basada en la orientación en la cual son clonadas originalmente en el vector pSelect. En los ejemplos dados antes, utilizando dos elementos STAR en vectores pSDH, esta direccionalidad es preservada. Esta orientación es descrita como la orientación natural o 5'-3', relativa al gen de resistencia a la zeocina (ver la Fig. 12). En este ejemplo se somete a ensayo la importancia de la direccionalidad para la función STAR en el vector pSDH-Tet. Puesto que los genes informadores de los vectores pSDH están flanqueados en ambos lados por copias del elemento STAR de interés, la orientación de cada copia STAR debe ser considerada. En este ejemplo se compara la orientación natural con la orientación contraria (Fig. 12).

10 *Materiales y Métodos*

El elemento STAR6 fue clonado en pSDH-Tet como se describe en el Ejemplo 11. Las células U-2 OS fueron co-transfectadas con los plásmidos pSDH-Tet-STAR₆₆-natural y pSDH-Tet-STAR₆₆-contrario, y cultivadas como se describe en el Ejemplo 11. Los clones individuales fueron aislados y cultivados; el nivel de expresión de luciferasa fue determinado como se ha descrito (*supra*).

Resultados

Los resultados de la comparación de la actividad STAR₆₆ en la orientación natural y la orientación contraria se muestran en la Fig. 13. Cuando STAR₆₆ está en la orientación contraria, el rendimiento de un solo clon es razonablemente elevado (60 unidades luciferasa). En contraste, el rendimiento del clon de expresión más elevada cuando STAR₆₆ está en la orientación natural es considerablemente mayor (100 unidades luciferasa), y la pronosticabilidad también es mucho mayor: 7 clones de la población de orientación natural (30%) expresan la luciferasa por encima del nivel del clon de expresión más elevada de la población de orientación contraria, y 15 de los clones de la población de orientación natural (60%) expresan la luciferasa por encima de aproximadamente 10 unidades luciferasa relativas. Por lo tanto se demuestra que la función STAR₆₆ es direccional.

Ejemplo 17

La expresión del transgen en el contexto de los elementos STAR es dependiente del número de copias.

Las unidades de expresión del transgen para la expresión de la proteína heteróloga son integradas generalmente en el genoma de la célula huésped para asegurar la retención estable durante la división celular. La integración puede producir una o múltiples copias de la unidad de expresión que esté siendo insertada en el genoma; las copias múltiples pueden estar presentes o no en forma de disposiciones en tándem. El rendimiento incrementado demostrado para los transgenes protegidos por elementos STAR (*supra*) sugiere que los elementos STAR son capaces de permitir que las unidades de expresión del transgen funcionen independientemente de las influencias sobre la transcripción asociadas con el sitio de integración en el genoma (independencia de los efectos de la posición (Boivin & Dura, 1998)). Esto sugiere adicionalmente que los elementos STAR permiten que cada unidad de expresión funcione independientemente de las copias vecinas de la unidad de expresión cuando están integradas en forma de una ordenación en tándem (independencia del silenciamiento del gen inducido por la repetición (Garrick *et al.*, 1998)). La dependencia del número de copias se determina a partir de la relación entre los niveles de expresión del transgen y el número de copias, como se describe más abajo en el ejemplo.

Materiales y Métodos

Las células U-2 OS fueron co-transfectadas con pSDH-Tet-STAR₁₀ y cultivadas bajo selección con puromicina como se ha descrito (*supra*). Ocho clones individuales fueron aislados y cultivados adicionalmente. Después las células fueron cosechadas, y una porción fue analizada en cuanto a la actividad luciferasa como se ha descrito (*supra*). Las células restantes fueron sometidas a lisis y el ADN genómico fue purificado utilizando el DNeasy Tissue Kit (QIAGEN 69504) como describe el fabricante. Las muestras de ADN fueron cuantificadas mediante espectrometría de UV. Tres microgramos de cada muestra de ADN genómico fueron digeridos con PvuII y XhoI durante la noche como describe el fabricante (New England Biolabs), y resueltos mediante electroforesis en gel de agarosa. Los fragmentos de ADN fueron transferidos a una membrana de nailon como se ha descrito (Sambrook *et al.*, 1989), e hibridados con una sonda marcada radiactivamente al gen de la luciferasa (aislado de pSDH-Tet digerido con BamHI/SacII). La transferencia fue lavada como se ha descrito (Sambrook *et al.*, 1989) y expuesta a una pantalla de formadora de imágenes con fósforo (Personal F/X, BioRad). El autorradiograma resultante (Fig. 14) fue analizado mediante densitometría para determinar la fuerza relativa de las bandas de ADN de luciferasa, que representa el número de copias del transgen.

Resultados

Las actividades de la enzima y los números de copias (intensidades de las bandas de ADN) de la luciferasa en los clones de la población de clones pSDH-STAR₁₀ se muestran en la Fig. 15. El número de copias del transgen está altamente correlacionado con el nivel de expresión de luciferasa en estos clones pSDH-STAR₁₀ ($r = 0,86$). Esto sugiere que STAR₁₀ confiere una dependencia del número de copias de las unidades de expresión del transgen, haciendo la expresión del transgen independiente de otras copias del transgen en disposiciones en tándem, e independiente de las influencias del silenciamiento del gen en el sitio de integración.

Ejemplo 18

Los elementos STAR funcionan como bloqueadores intensificadores pero no como intensificadores

5 Los promotores génicos están sujetos tanto a influencias positivas como negativas sobre su capacidad para iniciar la transcripción. Una clase importante de elementos que ejercen influencias positivas son los intensificadores. Los intensificadores son característicamente capaces de afectar a los promotores incluso cuando están localizados lejos (muchos kilopares de bases) del promotor. Las influencias negativas que actúan mediante la formación de heterocromatina (v.g. proteínas del grupo Polycomb) han sido descritas antes, y estas son la diana de la actividad STAR. La base bioquímica para la función intensificadora y para la formación de heterocromatina es fundamentalmente similar, ya que ambas implican la unión de proteínas al ADN. Por lo tanto es importante determinar si los elementos STAR son capaces de bloquear las influencias positivas así como las influencias negativas, en otras palabras, proteger los transgenes de los intensificadores genómicos en la vecindad del sitio de integración. La capacidad para proteger los transgenes de la actividad intensificadora asegura un funcionamiento estable y predecible de los transgenes en las aplicaciones biotecnológicas. Este ejemplo examina el funcionamiento de los elementos STAR en un análisis de bloqueo del intensificador.

Otro rasgo de la actividad STAR que es importante para su función es el aumento de rendimiento que confieren a los transgenes (Ejemplo 11). Los STAR son aislados basándose en su capacidad para mantener elevados niveles de expresión de zeocina cuando las proteínas formadoras de heterocromatina se unen adyacentes a los elementos STAR candidato. Se pronostica que se producirá una elevada expresión debido a que se prevé que los STAR bloquearán la diseminación de la heterocromatina en la unidad de expresión de zeocina. No obstante, un segundo escenario es que los fragmentos de ADN de los clones resistentes a zeocina contienen intensificadores. Se ha demostrado que los intensificadores tienen la capacidad de superar los efectos represivos de las proteínas del grupo Polycomb tales como las utilizadas en el método del rastreo con STAR (Zink & Paro, 1995). Los intensificadores aislados mediante este fenómeno serían considerados falsos positivos, puesto que los intensificadores no tienen las propiedades reivindicadas aquí para los STAR. Con el fin de demostrar que los elementos STAR no son intensificadores, estos han sido sometidos a ensayo en un análisis de intensificador.

30 El análisis de bloqueo del intensificador y el análisis de intensificador son metodológicamente y conceptualmente similares. Los análisis se muestran esquemáticamente en la Fig. 16. La capacidad de los elementos STAR para bloquear los intensificadores se realiza utilizando el sistema intensificador E47/caja E. La proteína E47 es capaz de activar la transcripción por medio de promotores cuando se une a una secuencia de ADN de la caja E localizada en la vecindad de esos promotores (Quong *et al.*, 2002). E47 está implicada normalmente en la regulación de la diferenciación de linfocitos B y T (Quong *et al.*, 2002), pero es capaz de funcionar en diversos tipos celulares cuando es expresada ectópicamente (Petersson *et al.*, 2002). La caja E es una secuencia de ADN palindrómica, CANNTG (Knöfler *et al.*, 2002). En el análisis de bloqueo del intensificador, se coloca una caja E aguas arriba del gen informador de la luciferasa (incluyendo un promotor mínimo) en un vector de expresión. Se coloca un sitio de clonación para los elementos STAR entre la caja E y el promotor. La proteína E47 es codificada en un segundo plásmido. El análisis se realiza transfectando tanto el plásmido E47 como el vector de expresión de la luciferasa en las células; la proteína E47 es expresada y se une a la caja E, y el complejo E47/caja E es capaz de actuar como intensificador. Cuando el vector de expresión de la luciferasa no contiene un elemento STAR, el complejo E47/caja E intensifica la expresión de la luciferasa (Fig. 16A, situación 1). Cuando los elementos STAR son insertados entre la caja E y el promotor, su capacidad para bloquear el intensificador es demostrada mediante la expresión reducida de la actividad luciferasa (Fig. 16A, situación 2); si los STAR no pueden bloquear los intensificadores, la expresión de la luciferasa es activada (Fig. 16A, situación 3).

La capacidad de los elementos STAR para actuar como intensificadores utiliza el mismo vector de expresión de la luciferasa. En ausencia de E47, la propia caja E no afecta a la transcripción. Al contrario, el comportamiento intensificador de los elementos STAR dará como resultado la activación de la transcripción de la luciferasa. El análisis es realizado transfectando el vector de expresión de la luciferasa sin el plásmido E47. Cuando el vector de expresión no contiene elementos STAR, la expresión de la luciferasa es baja (Fig. 16B, situación 1). Si los elementos STAR no tienen propiedades intensificadoras, la expresión de la luciferasa es baja cuando un elemento STAR está presente en el vector (Fig. 16B, situación 2). Si los elementos STAR tienen propiedades intensificadoras, la expresión de la luciferasa será activada en los vectores que contienen STAR (Fig. 16B, situación 3).

Materiales y Métodos

El vector de expresión de la luciferasa fue construido insertando la caja E y un promotor mínimo de la fosfatasa alcalina humana a partir del plásmido mu-E5+E2x6-cat(x) (Ruezinsky *et al.*, 1991) aguas arriba del gen de la luciferasa en el plásmido pGL3-básico (Promega E1751), para crear pGL3-caja-E-luciferasa (donación de W. Romanow). El plásmido de expresión de E47 contiene el marco de lectura abierto de E47 bajo el control de un promotor de la beta-actina en el plásmido pHBAPr-1-neo; E47 es expresado constitutivamente a partir de este plásmido (donación de W. Romanow). Los elementos STAR 1, 2, 3, 6, 10, 11, 18, y 27 han sido clonados en el vector de expresión de la luciferasa. Los clones que contenían el elemento scs de *Drosophila* y el núcleo HS4-6x de la beta-globina de pollo ("HS4") han sido incluidos como controles positivos (se sabe que bloquean los intensificadores, y no tienen propiedades intensificadoras intrínsecas (Chung *et al.*, 1993, Kellum & Schedl, 1992)), y el vector de expresión de luciferasa vacío ha sido incluido como control negativo. Todos los análisis fueron realizados utilizando la línea celular

U-2 OS. En el análisis de bloqueo del intensificador, el plásmido E47 fue co-transfectado con los vectores de expresión de la luciferasa (vector vacío, o conteniendo elementos STAR o de control positivo). En el análisis del intensificador, el plásmido E47 fue co-transfectado con el vector de expresión de la luciferasa sin STAR como control positivo para la actividad intensificadora; todas las demás muestras recibieron un supuesto plásmido durante la co-transfección. Las células transfectadas transitoriamente fueron analizadas en cuanto a la actividad de la luciferasa 48 horas después de la transfección del plásmido (*supra*). La actividad de la luciferasa expresada a partir del plásmido que no contenía los elementos de la caja E o STAR/control fueron sustraídos, y las actividades de la luciferasa fueron normalizadas para el contenido de proteína como se ha descrito (*supra*).

Resultados

La Fig. 17 muestra los resultados del análisis de bloqueo de intensificador. En ausencia de elementos STAR (o de los elementos scs y HS4 de bloqueo del intensificador conocido), el complejo intensificador E47/caja E activa la expresión de la luciferasa ("vector"); este nivel de expresión intensificado ha sido normalizado a 100. La actividad intensificadora es bloqueada por todos los elementos STAR sometidos a ensayo. La actividad intensificadora es bloqueada también por los elementos HS4 y scs, como se esperaba (Bell *et al.*, 2001, Gerasimova & Corces, 2001). Estos resultados demuestran que además de su capacidad para bloquear la diseminación del silenciamiento transcripcional (influencias negativas), los elementos STAR son capaces de bloquear la acción de los intensificadores (influencias positivas).

La Fig. 18 muestra los resultados del análisis de intensificador. El nivel de expresión de luciferasa debido a la intensificación por el complejo E47/caja-E se ajusta a 100 ("E47"). En comparación, ninguno de los elementos STAR ocasiona una activación significativa de la expresión de la luciferasa. Como se esperaba, los elementos scs y HS4 tampoco ocasionan la activación del gen informador. Por lo tanto se concluye que al menos los elementos STAR sometidos a ensayo no poseen propiedades intensificadoras.

Ejemplo 20

Los elementos STAR son conservados entre ratón y humano

El análisis BLAT de la secuencia de ADN STAR frente a la base de datos del genoma humano (<http://genome.ucsc.edu/cgi-bin/hgGateway>) revela que algunas de estas secuencias tienen una elevada conservación de la secuencias con otras regiones del genoma humano. Estas regiones duplicadas son elementos STAR candidatos; si muestran actividad STAR, serían considerados parálogos de los STAR clonados (se dice que dos genes o elementos genéticos son parálogos si derivan de un evento de duplicación (Li, 1997)).

El análisis BLAST de los STAR humanos frente al genoma de ratón (http://www.ensembl.org/Mus_musculus/bblastview) también revela regiones de elevada conservación de la secuencia entre ratón y humano. Esta conservación de la secuencia ha sido demostrada para fragmentos de 15 de los 65 elementos STAR humanos. La conservación oscila entre el 64% y el 89%, sobre longitudes de 141 pares de bases a 909 pares de bases (Tabla 8). Estos grados de conservación de la secuencia son notables y sugieren que estas secuencias de ADN pueden conferir actividad STAR también dentro del genoma de ratón. Algunas de las secuencias de los genomas de ratón y humano de la Tabla 8 podrían ser definidas estrictamente como ortólogos (se dice que dos genes o elementos genéticos son ortólogos si derivan de un evento de especiación (Li, 1997)). Por ejemplo, STAR6 está entre los genes SLC8A1 y HAAO en los genomas tanto humano como de ratón. En otros casos, un STAR humano clonado tiene un parálogo dentro del genoma humano, y su ortólogo ha sido identificado en el genoma de ratón. Por ejemplo, STAR3a es un fragmento de la región 15q11.2 del cromosoma humano 15. Esta región es idéntica en un 96,9% (paróloga) con un fragmento de ADN en 5q33.3 del cromosoma humano 5, que está cerca del gen de la interleuquina IL12B. Estos ADN humanos comparten aproximadamente una identidad del 80% con un fragmento de la región 11B2 del cromosoma de ratón 11. El fragmento 11B2 también está cerca del gen de la interleuquina IL112B (de ratón). Por lo tanto STAR3a y el fragmento 11B2 de ratón pueden ser definidos estrictamente como parálogos.

Con el fin de someter a ensayo la hipótesis de que la actividad STAR es compartida entre regiones de elevada conservación de la secuencia en el genoma de ratón y humano, uno de los STAR humanos con una secuencia conservada en el ratón, STAR18, ha sido analizada con mayor detalle. La conservación de la secuencia en el genoma de ratón detectada con el clon STAR18 original se extiende hacia la izquierda del cromosoma 2 humano durante aproximadamente 500 pares de bases (la Fig. 19; izquierda y derecha hacen referencia a la descripción normalizada de los brazos del cromosoma 2). En este ejemplo los autores de la presente invención examinan si la región de conservación de la secuencia define un elemento STAR de "origen natural" en humanos que es más extensa en longitud que el clon original. Los autores de la presente invención también examinan si la función STAR de este elemento STAR está conservada entre ratón y humano.

Materiales y Métodos

La región de conservación de la secuencia de ratón/humano en torno a STAR 18 fue recuperada del clon BAC humano RP11-387A1 mediante amplificación por PCR, en tres fragmentos: la región completa (cebadores E93 y E94), la mitad izquierda (cebadores E93 y E92), y la mitad derecha (cebadores E57 y E94). Los correspondientes fragmen-

tos de la región de ratón homóloga fueron recuperados del clon BAC RP23-400H17 de la misma manera (cebadores E95 y E98, E95 y E96, y E97 y E98, respectivamente). Todos los fragmentos fueron clonados en el vector pSelect y transfectados en una línea celular U-2 OS/Tet-Off/LexA-HP1 (*supra*). Tras la transfección, se llevó a cabo la selección con higromicina para seleccionar las células transfectadas. La proteína LexA-HP1 fue inducida disminuyendo la concentración de doxiciclina, y la capacidad de las células transfectadas para resistir al antibiótico zeocina (una medida de la actividad STAR) fue evaluada controlando el crecimiento celular.

Resultados

El clon STAR18 original fue aislado de ADN humano digerido con Sau3AI ligado en el vector pSelect basándose en su capacidad para prevenir el silenciamiento de un gen resistente a la zeocina. El alineamiento del clon STAR18 humano (497 pares de bases) con el genoma de ratón reveló una elevada similitud de secuencia (72%) entre las regiones STAR18 de humano y de ratón ortólogas. También dejaba patente una elevada similitud (73%) en la región que se prolongaba durante 488 pares de bases inmediatamente hacia la izquierda del sitio Sau3AI que define el extremo izquierdo de la región clonada (Fig. 19). Fuera de estas regiones la similitud de secuencia entre el ADN humano y de ratón cae por debajo del 60%.

Como se indica en la Fig. 19, los elementos STAR18 tanto humanos como de ratón confieren supervivencia en zeocina a las células huésped que expresan la proteína represora lexA-HP1. El clon STAR18 de 497 pares de bases y su ortólogo de ratón confieren ambos capacidad de crecimiento (Fig. 19, a y d). Las regiones de 488 pares de bases adyacentes de elevada similitud de ambos genomas también confieren capacidad para crecer, y de hecho su fenotipo de crecimiento es más vigoroso que el del clon STAR18 original (Fig. 19, b y e). Cuando la región completa de similitud de secuencia era sometida a ensayo, estos ADN tanto de ratón como de humano confieren crecimiento, y el fenotipo de crecimiento es más vigoroso que los dos sub-fragmentos (Fig. 19, c y f). Estos resultados demuestran que la actividad STAR de STAR18 humano es conservada en su ortólogo de ratón. La elevada conservación de la secuencia entre estas regiones ortólogas es particularmente notable debido a que no son secuencias codificadoras de proteínas, conduciendo a la conclusión de que tienen alguna función reguladora que ha evitado su divergencia evolutiva por medio de mutaciones.

Este análisis demuestra que los elementos STAR clonados identificados mediante el programa de rastreo original pueden representar en algunos casos elementos STAR parciales, y que el análisis del ADN genómico en el cual están embebidos puede identificar secuencias con una actividad STAR más fuerte.

Ejemplo 21

Materiales y Métodos

Utilizando el rastreo genético descrito en la solicitud de patente original, se aislaron inicialmente sesenta y seis (66) elementos STAR a partir de ADN genómico humano y se caracterizaron con detalle (Tabla 6). El rastreo se realizó sobre genotecas génicas construidas mediante digestión con Sau3AI de ADN genómico humano, ya sea purificado de placenta (Clontech 6550-1) o llevado a cabo en cromosomas artificiales bacteriano/P1 (BAC/PAC). Los clones BAC/PAC contienen ADN genómico de regiones del cromosoma 1 (clones RP1154H19 y RP3328E19), de la agrupación HOX de genes homeóticos (clones RP1167F23, RP1170019, y RP11387A1), o del cromosoma humano 22 (Research Genetics 96010-22). Los ADN fueron fraccionados por tamaño, y la fracción de tamaño 0,5-2 kb fue ligada en el vector pSelect digerido con BamHI, mediante mecanismos normalizados (Sambrook *et al.*, 1989). Los plásmidos pSelect que contenían ADN genómico que confería resistencia a zeocina a bajas concentraciones de doxiciclina fueron aislados y propagados en *Escherichia coli*. Los rastreos que rendían los elementos STAR de la Tabla 6 han analizado aproximadamente el 1-2% del genoma humano.

Los insertos de ADN genómico humano en estos 66 plásmidos fueron secuenciados mediante el método didesoxi (Sanger *et al.*, 1977) utilizando un secuenciador de ADN automático Beckman CEQ2000, empleando las instrucciones del fabricante. Brevemente, el ADN fue purificado a partir de *E. coli* utilizando QIAprep Spin Miniprep y Plasmid Midi Kits (QIAGEN 27106 y 12145, respectivamente). La secuenciación en ciclos se llevó a cabo utilizando los oligonucleótidos de costumbre correspondientes al vector pSelect (cebadores D89 y D95, Tabla 5), en presencia de terminadores coloreados (CEQ Dye Terminator Cycle Sequencing Kit, Beckman 608000). Las secuencias de ADN STAR ensambladas fueron localizadas en el genoma humano (construcciones de bases de datos Agosto y Diciembre 2001) utilizando BLAT (Basic Local Alignment Tool (Kent, 2002); <http://genome.ucsc.edu/cgi-bin/hgGateway>; Tabla 6). Además, las secuencias STAR combinadas comprenden 85, 6 kilopares de bases, con una longitud media de 1,3 kilopares de bases.

Breve descripción de los dibujos

Fig. 1. La familia de plásmidos pSelect para la selección y caracterización de elementos STAR. Un marcador de resistencia (zeocina o puomicina) o un gen informador (GFP o luciferasa) bajo el control del promotor de SV40 promiscuo es adyacente a un sitio de clonación BamHI flanqueado por los sitios AscI y HindIII. Aguas arriba del sitio de clonación están los operadores lexA a los cuales se puede unir la proteína lexA. La unión de las proteínas del grupo lexA-Polycomb quiméricas a los operadores ocasiona la represión del gen marcador o informador. Los fragmentos de

ADN insertados en el sitio de clonación que bloquean la represión son identificados mediante la expresión persistente del gen marcador o informador. El plásmido replica episómicamente en células de mamífero cultivadas debido a la secuencia de oriP.

Fig. 2. La familia pSDH de plásmidos para someter a ensayo los elementos STAR. Dos sitios de clonación múltiple (MCSI y MCSII) flanquean un gen informador (GFP o luciferasa) cuya expresión es conducida por un promotor aguas arriba (CMV, Tet-off, o SV40). Los elementos STAR a someter a ensayo son insertados en MCSI y MCSII. Estos contienen sitios de restricción únicos (MCSI: XhoI, NotI, EcoRI, y SalI; MCSII: HindIII, EcoRV, BglII, y NheI). El plásmido replica después de la integración al azar en el genoma de células de mamífero.

Fig. 3. Proporción de clones que sobre-expresan la luciferasa. Las células de osteosarcoma humano U-2 OS fueron transfectadas establemente con plásmidos pSDH (conteniendo el gen informador de la luciferasa bajo el control del promotor tet-off), y los clones transfectados individuales fueron aislados y cultivados. La expresión de la luciferasa fue medida enzimáticamente. Se determinó la expresión de luciferasa media por los clones que contenían el pSDH sin STAR ("nivel de referencia"). Los clones de los grupos de todos los plásmidos fueron puntuados como "sobre-expresión" si su actividad luciferasa era más de dos veces más alta que el nivel de referencia. Se traza el porcentaje de clones con sobre-expresión en cada grupo de plásmidos.

Fig. 4. Veces de sobre-expresión por los clones con sobre-expresión. El intervalo de sobre-expresión en los plásmidos pSDH que contienen STAR integrado en el ADN genómico fue determinado dividiendo las actividades de la luciferasa de cada clon por el nivel de referencia. Para aquellos que presentaban una expresión significativa (más de dos veces por encima del nivel de referencia), se observaron incrementos en veces reales; los valores mínimo y medio de estos datos se trazan para cada plásmido.

Fig. 5. Veces de sobre-expresión por los clones con sobre-expresión. El intervalo de sobre-expresión en los plásmidos pSDH que contienen STAR integrado en el ADN genómico fue determinado dividiendo las actividades de la luciferasa de cada clon por el nivel de referencia. Para aquellos que presentaban una expresión significativa (más de dos veces por encima del nivel de referencia), se observaron incrementos en veces reales; los valores máximos de estos datos se trazan para cada plásmido.

Fig 6. El plásmido pSDH-CSP utilizado para someter a ensayo la actividad STAR. El gen informador de la Fosfatasa Alcalina Secretada (SEAP) está bajo el control del promotor CMV, y el marcador seleccionable de resistencia a la puomicina (puro) está bajo el control del promotor de SV40. Flanqueando estos dos genes se encuentran múltiples sitios de clonación en los cuales los elementos STAR pueden ser clonados. El plásmido también tiene un origen de replicación (ori) y un gen de resistencia a ampicilina (ampR) para la propagación en *Escherichia coli*.

Fig. 7. STAR6 y STAR49 mejoran la pronosticabilidad y el rendimiento de la expresión del transgen. Se determinó la expresión de SEAP a partir del promotor CMV por las células CHO transfectadas con pSDH-CSP, pSDH-CSP-STAR6, o pSDH-CSP-STAR49. Los constructos que contienen STAR confieren mayor pronosticabilidad y un elevado rendimiento en relación con el constructo pSDH-CSP solo.

Fig 8. STAR6 y STAR8 mejoran la pronosticabilidad y el rendimiento de la expresión del transgen. Se determinó la expresión de luciferasa a partir del promotor de CMV por las células U-2 OS transfectadas con pSDH-CMV, pSDH-CMV-STAR6, o pSDH-CMV-STAR8. Los constructos que contienen STAR confieren mayor pronosticabilidad y un elevado rendimiento en relación con el constructo pSDH-CMV solo.

Fig 9. Secuencias esenciales mínimas de STAR10 y STAR27. Las porciones de los elementos STAR fueron amplificadas mediante PCR: STAR10 fue amplificado con los cebadores E23 y E12 para rendir el fragmento 10A, E13 y E14 para rendir el fragmento 10B, y E15 y E16 para rendir el fragmento 10C. STAR27 fue amplificado con los cebadores E17 y E18 para rendir el fragmento 27A, E19 y E20 para rendir el fragmento 27B, y E21 y E22 para rendir el fragmento 27C. Estos sub-fragmentos fueron clonados en el vector pSelect. Tras la transfección en células U-2 OS/Tet-Off/LexA-HP1, se verificó el crecimiento de los cultivos en presencia de zeocina. Las tasas de crecimiento variaban de vigoroso (+++) a escaso (+/-), mientras algunos cultivos no lograban sobrevivir al tratamiento con zeocina (-) debido a la ausencia de actividad STAR en el fragmento de ADN sometido a ensayo.

Fig 10. Función del elemento STAR en el contexto del promotor SV40. pSDH-SV40 y pSDH-SV40-STAR6 fueron transfectados en la línea celular de osteosarcoma humano U-2 OS, y la expresión de la luciferasa fue analizada con o sin protección del silenciamiento del gen por STAR6 en clones resistentes a la puomicina.

Fig 11. Función del elemento STAR en el contexto del promotor Tet-Off. pSDH-Tet y pSDH-Tet-STAR6 fueron transfectados en la línea celular de osteosarcoma humano U-2 OS, y la expresión de la luciferasa fue analizada con o sin protección del silenciamiento del gen por STAR6 en clones resistentes a la puomicina.

Fig 12. Diagrama esquemático de la orientación de elementos STAR a medida que son clonados en el vector pSelect (panel A), a medida que son clonados en vectores pSDH para conservar su orientación natural (panel B), y a medida que son clonados en el vector pSDH en la orientación contraria (panel C).

Fig 13. Direccionalidad de la función STAR66. El elemento STAR66 fue clonado en pSDH-Tet en orientación natural (STAR66 natural) o en orientación contraria (STAR66 contraria), y transfectado en células U-2 OS. La actividad de la luciferasa fue analizada en clones resistentes a la puromicina.

5 Fig 14. Dependencia del número de copias de la función STAR. Transferencia Southern de unidades de expresión de la luciferasa en pSDH-Tet-STAR10, integrado en ADN genómico de U-2 OS. La sonda de ADN de luciferasa radiactiva para detectar la cantidad de ADN de transgen en el genoma de cada clon, que después fue cuantificado con un aparato para la formación de imágenes con fósforo.

10 Fig 15. Dependencia del número de copias de la función STAR. Se determinó el número de copias de unidades de expresión pSDH-Tet-STAR10 en cada clon mediante formación de imágenes con fósforo, y se comparó con la actividad de la enzima informadora luciferasa expresada por cada clon.

15 Fig 16. Análisis de bloqueo del intensificador e intensificador. Los vectores de expresión de luciferasa utilizados para someter a ensayo los STAR en cuanto a la actividad de bloqueo del intensificador y del intensificador se muestran esquemáticamente. El sitio de unión a la caja-E para la proteína intensificadora E47 está aguas arriba de un sitio de clonación para los elementos STAR. Aguas abajo del sitio de clonación STAR está el gen de la luciferasa bajo el control de un promotor mínimo de la fosfatasa alcalina humana (mp). Los histogramas indican los resultados esperados para las tres posibles situaciones experimentales (ver el texto).

20 Panel A: Análisis de bloqueo del intensificador.

Panel B: Análisis del intensificador.

25 Fig 17. Análisis de bloqueo del intensificador. La expresión de la luciferasa de un promotor mínimo es activada por el intensificador E47/caja-E en el vector vacío (vector). La inserción de los bloqueadores del intensificador (scs, HS4) o de los elementos STAR (elementos STAR 1, 2, 3, 6, 10, 11, 18, y 27) bloquean la activación de la luciferasa por el intensificador E47/caja-E.

30 Fig 18. Análisis de intensificador. La expresión de la luciferasa de un promotor mínimo es activada por el intensificador E47/caja-E en el vector vacío (E47). La inserción de los elementos scs y HS4 o diversos elementos STAR (STAR 1, 2, 3, 6, 10, 11, 18, y 27) no activan la transcripción del gen informador.

35 Fig 19. Conservación de la secuencia de STAR18 entre ratón y humano. La región del genoma humano que contiene STAR18 de 497 pares de bases (recuadros negros); el elemento existe entre los genes del homeodominio HOXD8 y HOXD4 del cromosoma humano 2. Está alineado con una región en el cromosoma de ratón 2 que comparte una identidad de secuencia del 72%. La región del cromosoma 2 humano inmediatamente a la izquierda de STAR18 está altamente conservada con el cromosoma 2 de ratón (identidad del 73%, recuadros grises); más allá de esta región, la identidad cae por debajo del 60%. Se indica la capacidad de estas regiones de humano y de ratón, ya sea por separado o combinadas, para conferir crecimiento en zeocina: -, sin crecimiento; +, crecimiento moderado; ++, crecimiento vigoroso; +++, crecimiento rápido.

Fig. 20. Secuencias que comprenden STAR1 - STAR65 (SEQ ID:1 - 65).

45 Referencias

- Altschul, S.F. and Gish, W. (1996) Local alignment statistics. *Methods Enzymol*, 266, 460-480.
- 50 Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol*, 215, 403-410.
- Bell, AC, West, AG, and Felsenfeld, G. (2001) Insulators and boundaries: versatile regulatory elements in the eukaryotic genome *Science* 291, 447-50.
- 55 Bennett, R.P., Cox, C.A. and Hoeffler, J.P. (1998) Fusion of green fluorescent protein with the Zeocin-resistance marker allows visual screening and drug selection of transfected eukaryotic cells. *Biotechniques*, 24, 478-482.
- Berger, J, Hauber, J, Hauber, R, Geiger, R, and Cullen, BR. (1988) Secreted placental alkaline phosphatase: a powerful new quantitative indicator of gene expression in eukaryotic cells *Gene* 66, 1-10.
- 60 Bierhuizen, M.F., Westerman, Y., Visser, T.P., Wognum, A.W. and Wagemaker, G. (1997) Green fluorescent protein variants as markers of retroviral-mediated gene transfer in primary hematopoietic cells and cell lines. *Biochem Biophys Res Commun*, 234, 371-375.
- 65 Boivin, A, and Dura, JM. (1998) *In vivo* chromatin accessibility correlates with gene silencing in *Drosophila* *Genetics* 150, 1539-49.

- Boshart, M, Weber, F, Jahn, G, Dorsch-Hasler, K, Fleckenstein, B, and Schaffner, W. (1985)** A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus *Cell* 41, 521-30.
- Bunker, C.A. and Kingston, R.E. (1994)** Transcriptional repression by Drosophila and mammalian Polycomb group proteins in transfected mammalian cells. *Mol Cell Biol*, 14, 1721-1732.
- Cherry, J.R., Lamsa, M.H., Schneider, P., Vind, J., Svendsen, A., Jones, A. and Pedersen, A.H. (1999)** Directed evolution of a fungal peroxidase. *Nat Biotechnol*, 17, 379-384.
- Chung, JH, Whiteley, M, and Felsenfeld, G. (1993)** A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila *Cell* 74, 505-14.
- Deuschle, U, Meyer, WK, and Thiesen, HJ. (1995)** Tetracycline-reversible silencing of eukaryotic promoters *Mol Cell Biol* 15, 1907-14.
- Doll, R.F., Crandall, J.E., Dyer, C.A., Aucoin, J.M. and Smith, F.I. (1996)** Comparison of promoter strengths on gene delivery into mammalian brain cells using AAV vectors. *Gene Ther*, 3, 437-447.
- Feng, Y.Q., Seibler, J., Alami, R., Eisen, A., Westerman, K.A., Leboulch, P., Fiering, S. and Bouhassira, E.E. (1999)** Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange. *J Mol Biol*, 292, 779-785.
- Foecking, MK, and Hofstetter, H. (1986)** Powerful and versatile enhancer-promoter unit for mammalian expression vectors *Gene* 45, 101-5.
- Garrick, D, Fiering, S, Martin, DI, and Whitelaw, E. (1998)** Repeat-induced gene silencing in mammals *Nat Genet* 18, 56-9.
- Gaszner, M., Vazquez, J. and Schedl, P. (1999)** The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. *Genes Dev*, 13, 2098-2107.
- Gerasimova, T.I. and Corces, V.G. (1998)** Polycomb and trithorax group proteins mediate the function of a chromatin insulator. *Cell*, 92, 511-521.
- Gerasimova, TI, and Corces, VG. (2001)** Chromatin insulators and boundaries: effects on transcription and nuclear organization *Annu Rev Genet* 35, 193-208.
- Gossen, M. and Bujard, H. (1992)** Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A*, 89, 5547-5551.
- Graham, F.L. and van der Eb, A.J. (1973)** Transformation of rat cells by DNA of human adenovirus 5. *Virology*, 54, 536-539.
- Heldin, C.H., Johnsson, A., Wennergren, S., Wernstedt, C., Betsholtz, C. and Westermark, B. (1986)** A human osteosarcoma cell line secretes a growth factor structurally related to a homodimer of PDGF A-chains. *Nature*, 319, 511-514.
- Henke, E. and Bornscheuer, U.T. (1999)** Directed evolution of an esterase from *Pseudomonas fluorescens*. Random mutagenesis by error-prone PCR or a mutator strain and identification of mutants showing enhanced enantioselectivity by a resorufin-based fluorescence assay. *Biol Chem*, 380, 1029-1033.
- Henthorn, P, Zervos, P, Raducha, M, Harris, H, and Kadesch, T. (1988)** Expression of a human placental alkaline phosphatase gene in transfected cells: use as a reporter for studies of gene expression *Proc Natl Acad Sci U S A* 85, 6342-6.
- Higgins, D.G., Thompson, J.D. and Gibson, T.J. (1996)** Using CLUSTAL for multiple sequence alignments. *Methods Enzymol*, 266, 383-402.
- Himes, S.R. and Shannon, M.F. (2000)** Assays for transcriptional activity based on the luciferase reporter gene. *Methods Mol Biol*, 130, 165-174.
- Izumi, M, and Gilbert, DM. (1999)** Homogeneous tetracycline-regulatable gene expression in mammalian fibroblasts *J Cell Biochem* 76, 280-9.
- Jung, R., Soondrum, K. and Neumaier, M. (2000)** Quantitative PCR. *Clin Chem Lab Med*, 38, 833-836.
- Kain, SR. (1997)** Use of secreted alkaline phosphatase as a reporter of gene expression in mammalian cells *Methods Mol Biol* 63, 49-60.

Kao, F.T. and Puck, T.T. (1968) Genetics of somatic mammalian cells, VII. Induction and isolation of nutritional mutants in Chinese hamster cells. *Proc Natl Acad Sci U S A*, 60, 1275-81.

Kellum, R. and Schedl, P. (1992) A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Mol Cell Biol*, 12, 2424-2431.

Kent, WJ. (2002) BLAT-the BLAST-like alignment tool *Genome Res* 12, 656-64.

Knöfler, M, Meinhardt, G, Bauer, S, Loregger, T, Vasicek, R, Bloor, DJ, Kimber, SJ, and Husslein, P. (2002) Human Hand1 basic helix-loop-helix (bHLH) protein: extra-embryonic expression pattern, interaction partners and identification of its transcriptional repressor domains *Biochem J* 361, 641-51.

Li, W-H (1997) Molecular Evolution, Sinauer Associates, **Sunderland MA.**, **Meinke, DW, Cherry, JM, Dean, C, Rounsley, SD, and Koornneef, M. (1998)** Arabidopsis thaliana: a model plant for genome analysis *Science* 282, 662, 679-82.

Litt, M.D., Simpson, M., Recillas-Targa, F., Prioleau, M.N. and Felsenfeld, G. (2001) Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J*, 20, 2224-2235.

Morgenstern, J.P. and Land, H. (1990) Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic Acids Res*, 18, 3587-3596.

Nan, X., Javier Campoy, F., and Bird A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* 88, 471-481.

Petersson, K, Ivars, F, and Sigvardsson, M. (2002) The pT alpha promoter and enhancer are direct targets for transactivation by E box-binding proteins *Eur J Immunol* 32, 911-20.

Pazin, M.J. and Kadonaga, J.T. (1998) Transcriptional and structural analysis of chromatin assembled *in vitro*. In Gould, H. (ed.) Chromatin: A Practical Approach. *Oxford University Press*, Oxford, pp. 172-194.

Quong, MW, Romanow, WJ, and Murre, C. (2002) E protein function in lymphocyte development *Annu Rev Immunol* 20, 301-22.

Ruezinsky, D, Beckmann, H, and Kadesch, T. (1991) Modulation of the IgH enhancer's cell type specificity through a genetic switch *Genes Dev* 5, 29-37.

Saluz, H.P. and Jost, J.P. (1993) Approaches to characterize protein-DNA interactions *in vivo*. *Crit Rev Eukaryot Gene Expr*, 3, 1-29.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press*, Plainview NY.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 74, 5463-5467.

Simizu, B., Rhim, J.S. and Wiebenga, N.H. (1967) Characterization of the Tacaribe group of arboviruses. I. Propagation and plaque assay of Tacaribe virus in a line of African green monkey kidney cells (Vero). *Proc Soc Exp Biol Med*, 125, 119-123.

Smith, R.L., Traul, D.L., Schaack, J., Clayton, G.H., Staley, K.J. and Wilcox, C.L. (2000) Characterization of promoter function and cell-type-specific expression from viral vectors in the nervous system. *J Virol*, 74, 11254-11261.

Stull, R.A., Hyun, W.C. and Pallavicini, M.G. (2000) Simultaneous flow cytometric analyses of enhanced green and yellow fluorescent proteins and cell surface antigens in doubly transduced immature hematopoietic cell populations. *Cytometry*, 40, 126-134.

Takada, T., Iida, K., Akasaka, K., Yasue, H., Torii, R., Tsujimoto, G., Taira, M. and Kimura, H. (2000) Evaluation of heterologous insulator function with regard to chromosomal position effect in the mouse blastocyst and fetus. *Mol Reprod Dev*, 57, 232-237.

Tanaka, S., Livingstone-Zatchej, M. and Thoma, F. (1996) Chromatin structure of the yeast URA3 gene at high resolution provides insight into structure and positioning of nucleosomes in the chromosomal context. *J Mol Biol*, 257, 919-934.

Thomas, J.O. (1998) Isolation and fractionation of chromatin and linker histones. In Gould, H. (ed.) Chromatin: A Practical Approach. *Oxford University Press*, Oxford, pp. 1-34.

Umana, P, Jean-Mairet, J, and Bailey, JE. (1999) Tetracycline-regulated overexpression of glycosyltransferases in Chinese hamster ovary cells *Biotechnol Bioeng* 65, 542-9.

van der **Vlag, J., den Blaauwen, J.L., Sewalt, R.G., van Driel, R. and Otte, A.P. (2000)** Transcriptional repression mediated by polycomb group proteins and other chromatin-associated repressors is selectively blocked by insulators. *J Biol Chem*, 275, 697-704.

Wallrath, L.L., Swede, M.J. and Elgin, S.C.R. (1998) Mapping chromatin structure in Drosophila. In Gould, H. (ed.) *Chromatin: A Practical Approach. Oxford University Press*, Oxford, pp. 59-77.

Weaver, L.S. and Kadan, M.J. (2000) Evaluation of adenoviral vectors by flow cytometry. *Methods*, 21, 297-312.

Wigler, M, Pellicer, A., Silverstein, S. and Axel, R. (1978). Biochemical transfer of single-copy eukaryotic genes using total cellular DNA as donor. *Cell* 14, 725-731.

Wigley, P., Becker, C., Beltrame, J., Blake, T., Crocker, L., Harrison, S., Lyons, I., McKenzie, Z., Tearle, R., Crawford, R. and et al. (1994) Site-specific transgene insertion: an approach. *Reprod Fertil Dev*, 6, 585-588.

Xu, Z.Z., Krougliak, V., Prevec, L., Graham, F.L. and Both, G.W. (1995) Investigation of promoter function in human and animal cells infected with human recombinant adenovirus expressing rotavirus antigen VP7sc. *J Gen Virol*, 76, 1971-1980.

Yang, TT, Sinai, P, Kitts, PA, and Kain, SR. (1997) Quantification of gene expression with a secreted alkaline phosphatase reporter system *Biotechniques* 23, 1110-4.

Zink, D, and Paro, R. (1995) Drosophila Polycomb-group regulated chromatin inhibits the accessibility of a trans-activator to its target DNA *Embo J* 14, 5660-71.

Tabla 1. Elementos STAR mejoran expresión transgen.

Plásmido	Clones con sobre-expresión, %	Veces sobre-expresión (intervalo)	Número de clones
Vacío	12	3-11	25
SCS (control positivo)	24	3-160	21
STAR-6	62	2-200	26
STAR-3	39	5-820	23
STAR, -8	63	7-315	19
STAR-4	31	25-1500	13
STAR-1	57	5-80	23

ES 2 344 073 T3

TABLA 1 (continuación)

Tabla 1. Elementos STAR mejoran expresión transgen.			
Plásmido	Clones con sobre-expresión, %	Veces sobre-expresión (intervalo)	Número de clones
<p>La expresión del gen informador luciferasa es medida en líneas celulares que contienen plásmidos pSDH integrados, sin elementos STAR ("vacío", control negativo) o conteniéndolos (incluyendo el elemento control positivo, SCS de <i>Drosophila</i>). El nivel medio de expresión del control negativo se define como el nivel de referencia, y los clones son considerados con sobre-expresión si su nivel de expresión es >2 veces superior al nivel de referencia. Se informa sobre el porcentaje de clones con sobre-expresión para cada plásmido y el número de veces de sobre-expresión, junto con el número de clones analizados para cada plásmido.</p>			

TABLA 2

Elemento STAR clonado

Clon	Ubicación cromosómica ¹	Genes adyacentes ²	Secuencia repetida
STAR-1	N.d.		
STAR-2	N.d.		
STAR-3	For 5q33.3 Rev 10q22.2	Parte Chr 10 en histona. Gen acetiltransferasa	
STAR-4	For 1p31.1 Rev 14q24.1	Sin genes dentro del intrón de 10 kb de señalización de la proteína G	83% repetitiva LINE2 & LTR ERV Clase I
STAR-5	For 3q13.1 Rev 10q22.1*		

ES 2 344 073 T3

TABLA 2 (continuación)

Clon	Ubicación cromosómica ¹	Genes adyacentes ²	Secuencia repetida
STAR-6	2p21	Quinasa supuesta desconocida L5 kb Proteína asociada microtúbulos R 20 kb	19% SINE (MIR) 29% LINE
STAR-7	Lq32.2		12% Alu 4% MIR (SINE) LINE 12,5% L31CR1 11,5% MER1 7% de un complejo bajo al 2%
STAR-8	9q32	Caja ZFP KRAB conteniendo proteína dedo de cinc	35% ERV_Clase I (LTR) 2% repetición simple
STAR-9	Ver STAR4		
STAR-10	N.d.		
STAR-11	2p25.1	Inhibidor de proteína de enlace de ADN desconocido R 15 kb (tipo Myc)	12% (ALU) (SINE) 26% MarRs (LINE)
STAR-12	5q35.3	Familia metaloproteinasa TS2 ADAM desconocida R 15 kb	Complejidad baja 3%
STAR-13	Ver STAR4 y 9		
STAR-14	F N.d. R 20q13.33		
STAR-15	1p36.36	Subunidad de canal de K con compuerta a voltaje L 6 kb R 4 kb desconocido	14% LTR (MaLRs)

ES 2 344 073 T3

TABLA 2 (continuación)

Clon	Ubicación cromosómica ¹	Genes adyacentes ²	Secuencia repetida
STAR-16	F 8p23.1 R 8p22 etc.		
STAR-17	2q31.1	Factor transcripción L 6 kb R 40 kb HNRNP	10% de complejidad simple y baja
<p>¹ La localización cromosómica es determinada mediante una búsqueda BLAST de los datos de la secuencia de ADN de los clones STAR frente a la base de datos del genoma humano. La localización se proporciona de acuerdo con la nomenclatura normalizada que se refiere al ideograma citogenético de cada cromosoma; por ejemplo 1p2.3 es la tercera sub-banda citogenética de la segunda banda citogenética del brazo corto del cromosoma 1 (http://www.ncbi.nlm.nih.gov/Class/MLACourse/Genetics/chrombanding.html). F. Resultado de la reacción de secuenciación directa; R, resultado de la reacción de secuenciación inversa. N.d. no determinado todavía.</p> <p>² Basado en Human Genoma Map View Build 22 (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch?chr=hum_chr.inf&query April 2001). L, izquierdo; R, derecho.</p> <p>* Posición ambigua, diversos aciertos</p>			

TABLA 4

Secuencia de diversos elementos STAR en una hebra (directa) o la hebra opuesta (inversa)

STAR3 directo

ACGTNCTAAGNAAACCATTTATTATCATGACATTAACCTATAAAAAATAGGC
 GTATCACGAGGCCCTTTTCGTCTTCACTCGAGCGGCCAGCTTGGATCTCGA
 GTACTGAAATAGGAGTAAATCTGAAGAGCAAATAAGATGAGCCAGAAAAAC
 CATGAAAAGAACAGGGACTACCAGTTGATTCCACAAGGACATTCCCAAGG
 TGAGAAGGCCATATACCTCCACTACCTGAACCAATTCTCTGTATGCAGATT
 TAGCAAGGTTATAAGGTAGCAAAAGATTAGACCCAAGAAAATAGAGAACT
 TCCAATCCAGTAAAAATCATAGCAAATTTATTGATGATAACAATTGTCTCC
 AAAGGAACCAGGCAGAGTCGTGCTAGCAGAGGAAGCACGTGAGCTGAAA
 ACAGCCAAATCTGCTTTGTTTTTCATGACACAGGAGCATAAAGTACACACCA
 CCAACTGACCTATTAAGGCTGTGGTAAACCGATTTCATAGAGAGAGGTTCT
 AAATACATTGGTCCCTCATAGGCAAACCGCAGTTCCTCCGAACGTAAGTC
 CCTGGAAATTTGATGTCCAGNATAGAAAAGCANAGCAGNCNNNNNNNTAT
 ANATNNNGNTGANCCANATGNTNNCTGNNC

ES 2 344 073 T3

TABLA 4 (continuación)

STAR3 inverso

GAGCTAGCGGCGCGCCAAGCTTGGATCCCGCCCCGCCCCCTCCGCCCTCG
AGCCCCGCCCCCTTGCCCTAGAGGCCCTGCCGAGGGGCGGGGCCTGTCCC
TCCTCCCCTTTCCCCCGCCCCCTACCGTCACGCTCAGGGGCAGCCTGACC
CCGAGCGGCCCCGCGGTGACCCTCGCGCAGAGGCCTGTGGGAGGGGCGT
CGCAAGCCCCTGAATCCCCCCCCGTCTGTTCCCCCTCCCGCCCAGTCTC
CTCCCCCTGGGAACGCGCGGGGTGGGTGACAGACCTGGCTGCGCGCCAC
CGCCACCGCGCCTGCCGGGGGCGCTGCCGCTGCCTGAGAACTGCGGCT
GCCGCCTGGAGGAGGTGCCGTGCCTCCGCCACCGCTGCCGCCGCCGCC
AGGGGTAGGAGCTAAGCCGCCGCCATTTTGTGTCCCCCTGTTGTTGTCGT
TGACATGAATCCGACATGACACTGATTACAGCCCAATGGAGTCTCATTA
ACCCGAGTCGCGGTCCCGCCCCGCCGCTGCTCCATTGGAGGAGACCAAAG
ACACTTAAGGCCACCCGTTGGCCTACGGGTCTGTCTGTCACCCACTCACT

AACCACTCTGCAGCCCATTGGGGCAGGTTCTTGCCGGTCATNTCGCTTCC
AATAAACACACCCCTTCGACCCCATNATTCCCCCTTCGGGAACCACCC
CCGGGGGAGGGGTCCACTGGNCAATACCAATTNAANAGAACCGCTNGGG
TCCGCCTNTTTNCGGGCNCCCTATTGGGTT

ES 2 344 073 T3

TABLA 4 (continuación)

STAR4 directo

GGGGAGGATTCTTTTGGCTGCTGAGTTGAGATTAGGTTGAGGGTAGTGAA
GGTAAAGGCAGTGAGACCACGTAGGGGTCATTGCAGTAATCCAGGCTGG
AGATGATGGTGGTTCAGTTGGAATAGCAGTGCATGTGCTGTAACAACCTC
AGCTGGGAAGCAGTATATGTGGCGTTATGACCTCAGCTGGAACAGCAATG
CATGTGGTGGTGTAAATGACCCCAGCTGGGTAGGGTGTCATGTGATGGAACA
ACCTCAGCTGGGTAGCAGTGTAAGTGTGATAAAATGTTGGCATACTCTACATT
TGTTATGAGGGTAGTGCCATTAAATTTCTCCACAAATTGGTTGTCACGTAT
GAGTGAAAAGAGGAAGTGATGGAAGACTTCAGTGCTTTTGGCCTGAATAA
ATAGAAGACGTCATTTTCAGTAATGGAGACAGGGAAGACTAANGNAGGGT
GGATTCAGTAGAGCAGGTGTTTCAGTTTTGAATATGATGAACTCTGAGAGA
GGAAAACTTTTTCTACCTCTTAGTTTTTGNGNCTGGACTTAANATTAAAG
GACATANGACNGAGANCAGACCAAATNTGCGANGTTTTATATTTTACTT
GCNGAGGGAATTTNCAAGAAAAAGAAGACCCAANANCCATTGGTCAAAA
CTATNTGCCTTTTAANAAAAAGANAATTACAATGGANANANAAGTGTTGN
CTNGGCAAAAATTGGG

ES 2 344 073 T3

TABLA 4 (continuación)

5	STAR4 inverso
10	GGATTNGAGCTAGCGGCGCGCCAAGCTTGGATCTTAGAAGGACAGAGTG GGGCATGGAAATGCACCACCAGGGCAGTGCAGCTTGGTCACTGCCAGCTC CNCTCATGGGCAGAGGGCTGGCCTCTTGCAGCCGACCAGGCACTGAGCG CCATCCCAGGGCCCTCGCCAGCCCTCAGCAGGGCCAGGACACACAAGCCT TTGACTTCCTCCTGTCACTGCTGCTGCCATTCCTGTTTTGTGGTCATCACT CCTTCCCTGTCCTCAGACTGCCCAGCACTCAAGGATGTCCTGTGGTGGCA TCAGACCATATGCCCTGAANAGGAGTGAGTTGGTGTTTTTTGCCGCGCC
20	
25	CANAGAGCTGCTGTCCCCTGAAAGATGCAAGTGGGAATGATGATGNTCAC CATCNTCTGACACCAAGCCCTTTGGATAGAGGCCCAACAGTGAGGATGG GGCTGCACTGCATTGCCAAGGCAACTCTGTNNTGACTGCTACANGACANT CCCAGGACCTGNGAAGNNCTATANATNTGATGCNAGGCACCT
30	
35	
40	
45	
50	
55	
60	
65	

ES 2 344 073 T3

TABLA 4 (continuación)

STAR6 directo

CCACCACAGACATCCCCTCTGGCCTCCTGAGTGGTTTCTTCAGCACAGCTT
 CCAGAGCCAAATTAAACGTTCACTCTATGTCTATAGACAAAAAGGGTTTTG
 ACTAAACTCTGTGTTTTAGAGAGGGAGTTAAATGCTGTAACTTTTTAGGG
 GTGGGCGAGAGGAATGACAAATAACAACCTGTCTGAATGTTTTACATTC
 TCCCCACTGCCTCAAGAAGGTTCAACACGAGGTCATCCATGATAAGGAGT
 AAGACCTCCCAGCCGGACTGTCCCTCGGCCCCCAGAGGACACTCCACAGA
 GATATGCTAACTGGACTTGGAGACTGGCTCACACTCCAGAGAAAAGCATG
 GAGCACGAGCGCACAGAGCANGGGCCAAGTCCCAGGGACNGAATGTCT
 AGGAGGGAGATTGGGGTGAGGGTANTCTGATGCAATTACTGNGCAGCTC
 AACATTCAAGGGAGGGGAAGAAAGAAACNGTCCCTGTAAGTAAGTTGTNC
 ANCAGAGATGGTAAGCTCCAAATTTNAACTTTGGCTGCTGGAAAGTTTNN
 GGGCCNANANAANAACANAAANATTTGAGGTTTANACCCACTAACCCN
 TATNANTANTTATTAATACCCCTAATTANACCTTGGATANCCTTAAAATAT
 CNTNTNAAACGGAACCCTCNTTCCCNNTTNNAATNNNAAAGGCCATTN
 NGNNCNAGTAAAAATCTNNNTTAAGNNNTGGGCCCNAAACAAACNTNTTC
 CNAGACACNTTTTTTNTCCNGGNATTTNTAATTTATTTCTAANCC

ES 2 344 073 T3

TABLA 4 (continuación)

STAR6 inverso

ATCGTGTCTTTCCAGGGACATGGATGAAGCTGGAAGCCATCATCCTCAG
 CAACTAACACAGGAACAGAAAACCAAATACCACATGTTCTCACTCATAAG
 TGGGAGCTGAACAGTGAGAACACATGGACACAGGGAGGGGAACATCACA
 CACCAAGGCCTGTCTGGTGTGGGGAGGGGAGGGAGAGCATCAGGACAAA
 TAGCTAATGCATGTGGGGCTTAAACCTAGATGACGGGTTGATAGGTGCAG

CAATCCACTATGGACACATATACCTATGTAACAACCCNACCTTNTTGACAT
 GTATCCCAGAACTTAAAGGAAAATAAAAATTAAAAAAAATTNCCCTGGAA
 TAAAAAAGAGTGTGGACTTTGGTGAGATN

STAR8 directo

GGATCACCTCGAAGAGAGTCTAACGTCCGTAGGAACGCTCTCGGGTTCAC
 AAGGATTGACCGAACCCCAGGATACGTCGCTCTCCATCTGAGGCTTGNTC
 CAAATGGCCCTCCACTATTCCAGGCACGTGGGTGTCTCCCCTAACTCTCC
 CTGCTCTCCTGAGCCCATGCTGCCTATCACCCATCGGTGCAGGTCCTTTCT
 GAANAGCTCGGGTGGATTCTCTCCATCCCACTTCCTTTCCCAAGAAAGAA
 GCCACCGTTCCAAGACACCCAATGGGACATTCCCNNTCCACCTCCTTNTC
 NAAAGTTNGCCAGGTGTTCNNTAACAGGTTAGGGAGAGAANCCCCCAGG
 TTTNAGTTNCAAGGCATAGGACGCTGGCTTGAACACACACACACNCTC

ES 2 344 073 T3

TABLA 4 (continuación)

STAR8 inverso

**GGATCCCGACTCTGCACCGCAAACCTCTACGGCGCCCTGCAGGACGGCGGC
CTCCTGCCGCTTGACGCCAGNCAGGAGCTCCCCGGCAGCAGCAGAGCA
GAAAGAAGGATGGCCCCGCCCACTTCGCCTCCCGGCGGTCTCCCTCCCG
CCGGCTCACGGACATAGATGGCTGCCTAGCTCCGGAAGCCTAGCTCTTGT
TCCGGGCATCCTAAGGAAGACACGGTTTTTCTCCCGGGGCCTCACCACA
TCTGGGACTTTGACGACTCGGACCTCTCTCATTGAATGGTTGCGCGTTC
TCTGGGAAAG**

STAR18 directo

**TGGATCCTGCCGCTCGCGTCTTAGTGTTTCTCCCTCAAGACTTTCCTTCTG
TTTTGTTGTCTTGTGCAGTATTTTACAGCCCCTCTTGTGTTTTTCTTTATTT
CTCGTACACACACGCAGTTTTAAGGGTGATGTGTGTATAATTAAAAGGAC
CCTTGCCCATACTTTCCTAATTCTTTAGGGACTGGGATTGGGTTTGACTG
AAATATGTTTTGGTGGGGATGGGACGGTGGACTTCCATTCTCCCTAAACT
GGAGTTTTTGGTCGGTAATCAAACTAAAAGAAACCTCTGGGAGACTGGAA**

ES 2 344 073 T3

TABLA 4 (continuación)

5

ACCTGATTGGAGCACTGAGGAACAAGGGAATGAAAAGGCAGACTCTCTGA
ACGTTTGATGAAATGGACTCTTGTGAAAATTAACAGTGAATATTTCACTGTT
10 GCACTGTACGAAGTCTCTGAAATGTAATTAAGTTTATTGAGCCCCCG
AGCTTTGGCTTGCGCGTATTTTTCCGGTCGCGGACATCCCACCGCGCAGA
GCCTCGCCTCCCCGCTGNCCTCAGCTCCGATGACTTCCCCGCCCCCGCCC
15 TGCTCGGTGACAGACGTTCTACTGCTTCCAATCGGAGGCACCCTTCGCGG

15

20

STAR18 inverso

TGGATCCTGCCGCTCGCGTCTTAGTGTTTCTCCCTCAAGACTTTCCTTCTG
25 TTTTGTGTCTTGTGCAGTATTTTACAGCCCCTCTTGTGTTTTTCTTTATTT
CTCGTACACACACGCAGTTTAAAGGGTGATGTGTGTATAATTAAGGAC
CCTTGGCCCATACTTTCCTAATTCTTTAGGGACTGGGATTGGGTTTGACTG
30 AAATATGTTTTGGTGGGGATGGGACGGTGGACTTCCATTCTCCCTAAACT
GGAGTTTGGTGGTAATCAAACTAAAAGAAACCTCTGGGAGACTGGAA
ACCTGATTGGAGCACTGAGGAACAAGGGAATGAAAAGGCAGACTCTCTGA
35 ACGTTTGATGAAATGGACTCTTGTGAAAATTAACAGTGAATATTTCACTGTT
GCACTGTACGAAGTCTCTGAAATGTAATTAAGTTTATTGAGCCCCCG
AGCTTTGGC

40

45

50

55

60

65

ES 2 344 073 T3

TABLA 5

*Oligonucleótidos utilizados para las reacciones en cadena de la polimerasa (cebadores de PCR)
o mutagénesis de ADN*

Número	Secuencia
C65	AACAAGCTTGATATCAGATCTGCTAGCTTGGTCGAGCTGATAC TTCCC
C66	AAACTCGAGCGGCCGCGAATTCGTCGACTTTACCACTCCCTAT CAGTGATAGAG
C67	AAACCGCGGCATGGAAGACGCCAAAAACATAAAGAAAGG
C68	TATGGATCCTAGAATTACACGGCGATCTTTCC
C81	AAACCATGGCCGAGTACAAGCCCACGGTGCGCC
C82	AAATCTAGATCAGGCACCGGGCTTGCGGGTCATGC
C85	CATTTCCCCGAAAAGTGCCACC
D30	TCACTGCTAGCGAGTGGTAAACTC
D41	GAAGTCGACGAGGCAGGCAGAAGTATGC
D42	GAGCCGCGGTTTAGTTCCCTCACCTTGTCG
D51	TCTGGAAGCTTTGCTGAAGAAAC
D89	GGGCAAGATGTCGTAGTCAGG
D90	AGGCCCATGGTCACCTCCATCGCTACTGTG
D91	CTAATCACTCACTGTGTAAT
D93	AATTACAGGCGCGCC
D94	AATTGGCGCGCCTGT
D95	TGCTTTGCATACTTCTGCCTGCCTC
E12	TAGGGGGGATCCAAATGTTC
E13	CCTAAAAGAAGATCTTTAGC
E14	AAGTGTTGGATCCACTTTGG
E15	TTTGAAGATCTACCAAATGG

ES 2 344 073 T3

TABLA 5 (continuación)

5	Número	Secuencia
	E16	GTTTCGGGATCCACCTGGCCG
10	E17	TAGGCAAGATCTTGGCCCTC
	E18	CCTCTCTAGGGATCCGACCC
15	E19	CTAGAGAGATCTTCCAGTAT
	E20	AGAGTTCCGGATCCGCCTGG
20	E21	CCAGGCAGACTCGGAACTCT
	E22	TGGTGAAACCGGATCCCTAC
25	E23	AGGTCAGGAGATCTAGACCA
	E25	CCATTTTCGCTTCCTTAGCTCC
30	E42	CGATGTAACCCACTCGTGCACC
	E57	AGAGATCTAGGATAATTTTCG
35	E92	AGGCGCTAGCACGCGTTCTACTCTTTTCCTACTCTG
	E93	GATCAAGCTTACGCGTCTAAAGGCATTTTATATAG
40	E94	AGGCGCTAGCACGCGTTCAGAGTTAGTGATCCAGG
	E95	GATCAAGCTTACGCGTCAGTAAAGGTTTCGTATGG
45	E96	AGGCGCTAGCACGCGTTCTACTCTTTCATTACTCTG
	E97	CGAGGAAGCTGGAGAAGGAGAAGCTG
50	E98	CAAGGGCCGCAGCTTACACATGTTC
	D58	CCAAGTTGACCAGTGCC
55	D80	GTTTCGTGGACACGACCTCCG
	D70	TACAAGCCAACCACGGCCT
60	D71	CGGAAGTGCTTGACATTGGG
65		

ES 2 344 073 T3

TABLA 6

Elementos STAR de la invención, incluyendo localización y longitud genómica

STAR	Localización ¹	Longitud ²
1	2q31.1	750
2	7p15.2	916
3 ³	15q11.2 y 10q22.2	2132
4	1p31.1 y 14q24.1	1625
5 ⁴	20q13.32	1571
6	2p21	1173
7	1q34	2101
8	9q32	1839
9 ⁴	10p15.3	1936
10	Xp11.3	1167
11	2p25.1	1377
12	5q35.3	1051
13 ⁴	9q34.3	1291
14 ⁴	22q11.22	732
15	1p36.31	1881
16	1p21.2	1282
17	2q31.1	793
18	2q31.3	497
19	6p22.1	1840
20	8p13.3	780
21	6q24.2	620
22	2q12.2	1380
23	6p22.1	1246

ES 2 344 073 T3

TABLA 6 (continuación)

5	STAR	Localización ¹	Longitud ²
	24	1q21.2	948
10	25 ⁵	1q21.3	1067
	26	1q21.1	540
15	27	1q23.1	1520
	28	22q11.23	961
20	29	2q13.31	2253
	30	22q12.3	1851
25	31	9q34.11 y 22q11.21	1165
	32	21q22.2	771
30	33	21q22.2	1368
	34	9q34.14	755
35	35	7q22.3	1211
	36	21q22.2	1712
40	37	22q11.23	1331
	38	22q11.1 y 22q11.1	~1000
45	39	22q12.3	2331
	40	22q11.21	1071
50	41	22q11.21	1144
	42	22q11.1	735
55	43	14q24.3	1231
	44	22q11.1	1591
60	45	22q11.21	1991
	46	22q11.23	1871
65	47	22q11.21	1082
	48	22q11.22	1242

ES 2 344 073 T3

TABLA 6 (continuación)

STAR	Localización ¹	Longitud ²
49	Clon al azar Chr 12, y 3q26.32	1015
50	6p21.31	2361
51	5q21.3	2289
52	7p15.2	1200
53	Xp11.3	1431
54	4q21.1	981
55	15q13.1	501
56	Incluye 3p25.3	741
57	4q35.2	1371
58	21q11.2	1401
59	17 clon al azar	872
60	4p 16.1 y 6q27	2068
61	7p14.3 y 11q25	1482
62	14q24.3	1011
63	22q13.3	1421
64	17q11.2	1414
65	7q21.11=28.4	1310
66	20q13.33 y 6q14.1	~2800

¹ La localización cromosómica es determinada mediante una búsqueda BLAST de los datos de los elementos STAR frente a la base de datos del genoma humano. La localización se da de acuerdo con la nomenclatura normalizada que se refiere al ideograma citogenético de cada cromosoma; por ejemplo 1p2.3 es la tercera sub-banda citogenética de la segunda banda citogenética del brazo corto del cromosoma 1 (<http://www.ncbi.nlm.nih.gov/Class/MLACourse/Genetics/>

TABLA 6 (continuación)

chrombanding.html). En los casos en los que la reacción de secuenciación directa e inversa identificaban ADN de diferentes loci genéticos, se muestran ambos loci.

² Se determinan las longitudes precisas mediante análisis de la secuencia de ADN; las longitudes aproximadas son determinadas mediante mapeo de restricción.

³ Secuencia y localización de STAR3 ha sido refinada a partir del ensamblaje de las Tablas 2 y 4.

⁴ Los STAR con estos números en las Tabla 2 y 4 se han dejado a un lado (referido en adelante como "oldSTAR5" etc.) y sus números asignados a los elementos STAR mostrados en el apéndice de secuencias de ADN. En el caso de oldSTAR5, oldSTAR14, y oldSTAR16, los ADN clonados eran quimeras de más de dos localizaciones cromosómicas; en el caso de oldSTAR9 y oldSTAR13, los ADN clonados eran idénticos a STAR4.

⁵ Idéntico a la Tabla 4 "STAR18".

ES 2 344 073 T3

TABLA 7

Los elementos STAR otorgan estabilidad con el tiempo a la expresión de los transgenes¹

	<u>Divisiones Celulares²</u>	<u>Expresión Luciferasa³</u>
STAR6 más Puromicina	42	18.000
	60	23.000
	84	20.000
	108	16.000
STAR6 sin Puromicina ⁴	84	12.000
	108	15.000
	144	12.000
¹ El plásmido pSDH-Tet-STAR6 fue transfectado en células U-2 OS, y los clones fueron asilados y cultivados en medio sin doxiciclina como se describe en el Ejemplo 1. Las células fueron transfectadas a recipientes de cultivo nuevos a una dilución de 1:20. ² El número de divisiones celulares se basa en la estimación de que en una semana el cultivo alcanza la confluencia celular, lo que representa ~6 divisiones celulares. ³ La luciferasa fue analizada como se describe en el Ejemplo 1. ⁴ Al cabo de 60 divisiones celulares las células fueron transfectadas a dos recipientes de cultivo; uno estaba provisto de medio de cultivo que contenía puromicina, para las 60 primeras divisiones celulares, y el segundo estaba provisto de medio de cultivo que carecía de antibiótico.		

ES 2 344 073 T3

TABLA 8

Elementos STAR humanos y sus supuestos ortólogos y parálogos de ratón

SEQ: ID	STAR	Humano ¹	Ratón ²	Similitud ³
1	1	2q31.1	2D	600 pb 69%
2	2	7p 15.2	6B3	909 pb 89%
3	3a	5q33.3	11B2	248 pb 83%
4	3b	10q22.2	14B	1.363 pb 89% 2.163 pb 86%
5	6	2p21	17E4	437 pb 78%
6	12	5q35.3	11b1.3	796 pb 66%
7	13	9q34.3	2A3	753 pb 77%
8	18	2q31.3	2E1	497 pb 72%
9	36	21q22.2	16C4	166 pb 79%
10	40	22q11.1	6F1	1.270 pb 75% 2.309 pb 70%
11	50	6p21.31	17B1	1.451 pb 72% 2.188 pb 80% 3.142 pb 64%
12	52	7p15.2	6B3	1.846 pb 74% 2.195 pb 71%
13	53	Xp11.3	XA2	364 pb 64%
14	54	4q21.1	5E3	1.174 pb 80% 2.240 pb 73% 3.141 pb 67% 4.144 pb 68%
15	61a	7p14.3	6B3	188 pb 68%

TABLA 8 (continuación)

5	¹ Ubicación citogenética del elemento STAR en el genoma humano.
10	² Ubicación citogenética del elemento STAR ortólogo en el genoma de ratón.
15	³ Longitud de las regiones que presentan una alta similitud de secuencias, y una similitud de porcentaje. En algunos casos existe más de un bloque de elevada similitud; en otros casos, cada bloque se describe de manera separada. La similitud de <60% no se considera
20	importante.

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Una secuencia de ADN aislada y/o recombinante que tiene actividad anti-represora, seleccionada dicha secuencia del grupo que consiste en:

- (a) el SEQ ID: 44 de la Figura 20;
- (b) un fragmento del SEQ ID: 44 de la Figura 20;
- (c) una secuencia derivada del SEQ ID: 44 de la Figura 20 por inserción, delección y/o mutación de una o más bases;

donde la actividad anti-represora de dicha secuencia confiere a una célula de osteosarcoma U-2 OS humana la capacidad para crecer después de 4-5 semanas de cultivo en presencia de 250 μ g/ml de Zeocina y 0,1 ng/ml de doxiciclina, cuando dicha célula comprende una proteína de fusión represora de LexA que contiene el dominio de unión a ADN de LexA y una región codificadora de HP1 o HPC2 bajo el control del sistema regulador de la transcripción Tet-Off, cuando dicha secuencia de ADN aislada y/o recombinante es clonada en una secuencia poliligadora en un plásmido, estando situado dicho poliligador entre cuatro sitios operadores de LexA y el promotor de SV40 que controla el gen de resistencia a zeocina, cuando el plásmido está presente en dicha célula.

2. Un constructo de ADN recombinante provisto de una secuencia de ADN según la reivindicación 1.

3. Un constructo de ADN según la reivindicación 2, que comprende adicionalmente un promotor conectado operablemente a una secuencia de ácido nucleico de interés.

4. Un constructo de ADN según la reivindicación 3, donde dicha secuencia de ácido nucleico de interés es un marco de lectura abierto de un transgen.

5. Un constructo de ADN según la reivindicación 4, donde dicho promotor es un promotor exógeno.

6. Un constructo de ADN según una cualquiera de las reivindicaciones 3-5, donde dicho promotor es un promotor constitutivo fuerte, preferiblemente un promotor viral, o un promotor inducible.

7. Un constructo de ADN según una cualquiera de las reivindicaciones 3-6, donde dicho promotor es un promotor de CMV, un promotor de SV40 o un promotor Tet-Off.

8. Un constructo de ADN según una cualquiera de las reivindicaciones 3-7, que comprende en el siguiente orden:

- (i) una secuencia de ADN según la reivindicación 1,
- (ii) el promotor conectado operablemente a la secuencia de ácido nucleico de interés, y
- (iii) una secuencia de ADN según la reivindicación 1, preferiblemente en orientación opuesta a (i).

9. Un método para obtener una célula anfitriona que comprende la etapa de transfectar la célula anfitriona con un constructo de ADN según una cualquiera de las reivindicaciones 3-8.

10. Una célula que comprende un constructo de ADN según una cualquiera de las reivindicaciones 3-8.

11. Una célula según la reivindicación 10, que comprende múltiples copias de dicho constructo de ADN.

12. Una célula según la reivindicación 10 u 11, que es una célula CHO.

13. Un método para producir un producto génico en una célula que comprende proporcionar una casete de expresión que comprende:

- (i) un transgen que codifica dicho producto génico, y
 - (ii) una secuencia de ADN según la reivindicación 1,
- y permitir la transcripción de dicha casete de expresión en una célula.

14. Un método según la reivindicación 13, donde dicha célula es una célula CHO.

ES 2 344 073 T3

15. Un método según la reivindicación 13 o 14, donde la casete de expresión comprende en el siguiente orden:

- (i) una secuencia de ADN según la reivindicación 1,
- 5 (ii) el transgen que comprende un promotor conectado operablemente a un marco de lectura abierto que codifica dicho producto génico, y
- (iii) una secuencia de ADN según la reivindicación 1, preferiblemente en orientación opuesta a (i).

10

16. Un método según una cualquiera de las reivindicaciones 13-15, donde se proporciona una multiplicidad de dichas casetes de expresión y se permite la transcripción de las mismas en dicha célula, y la expresión de dicho transgen depende del número de copias.

15 17. El uso *in vitro* de una secuencia de ADN según la reivindicación 1, para regular la transcripción de un ácido nucleico de interés.

20

25

30

35

40

45

50

55

60

65

Fig. 1

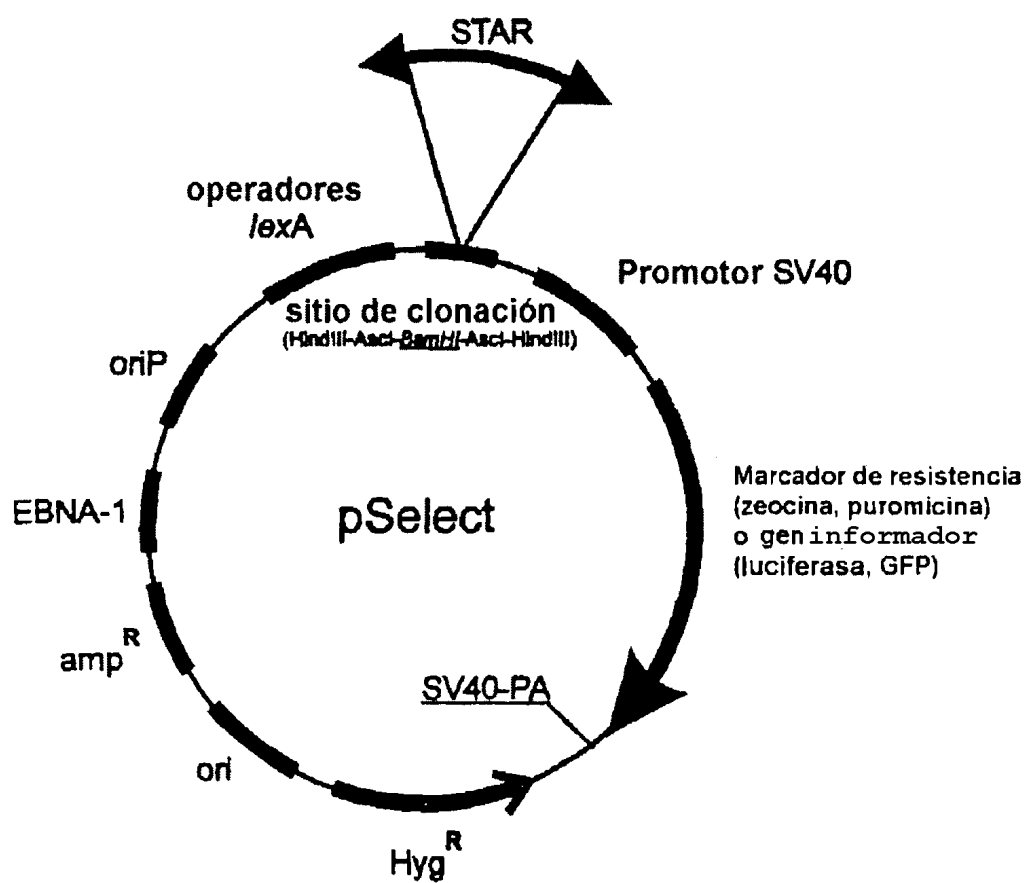


Fig. 2

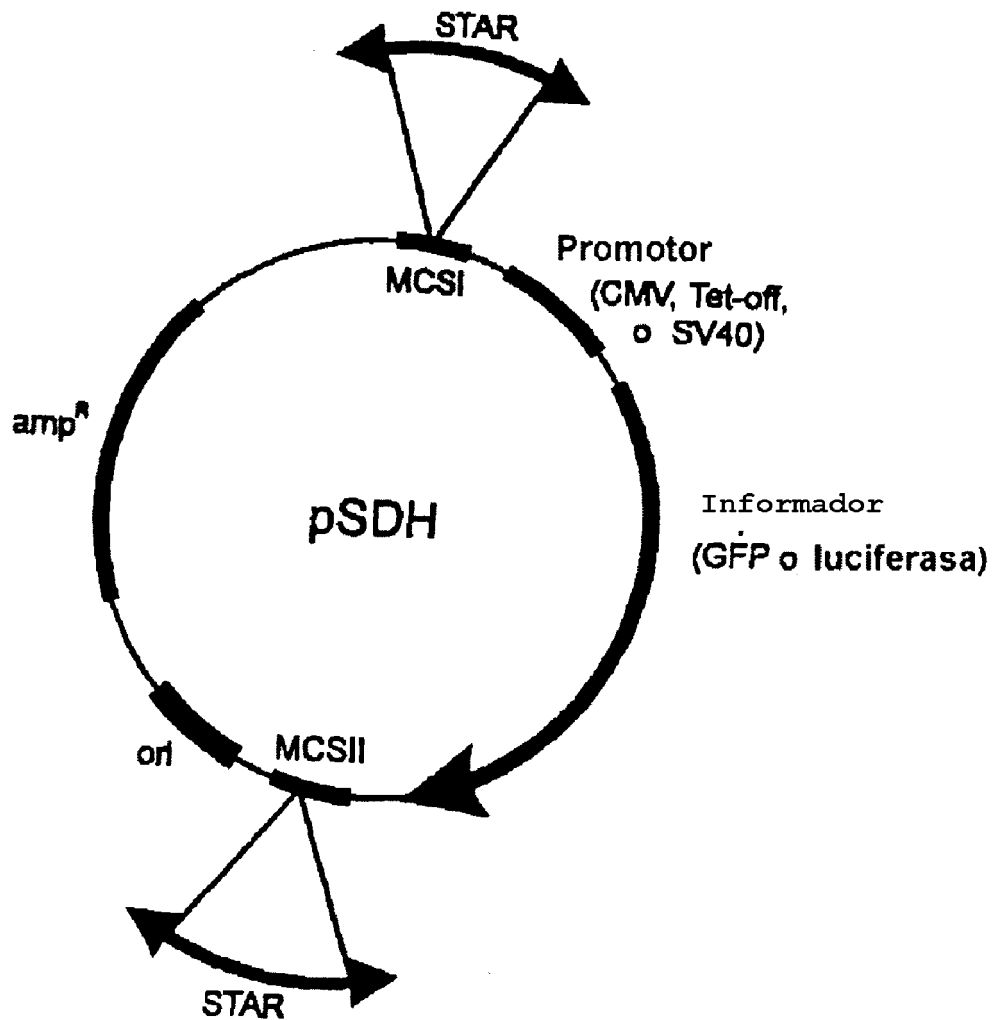


Fig. 3

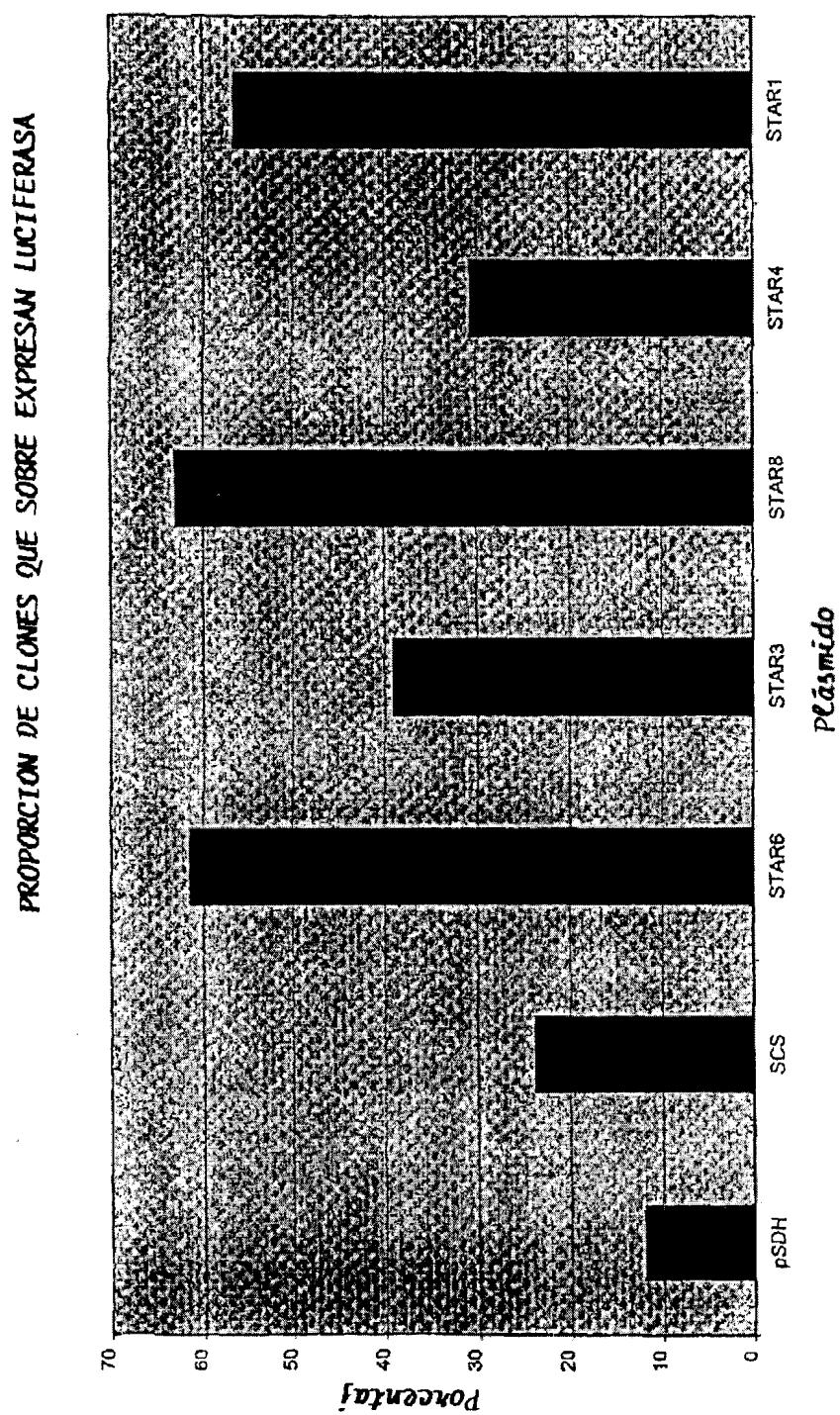


Fig. 4

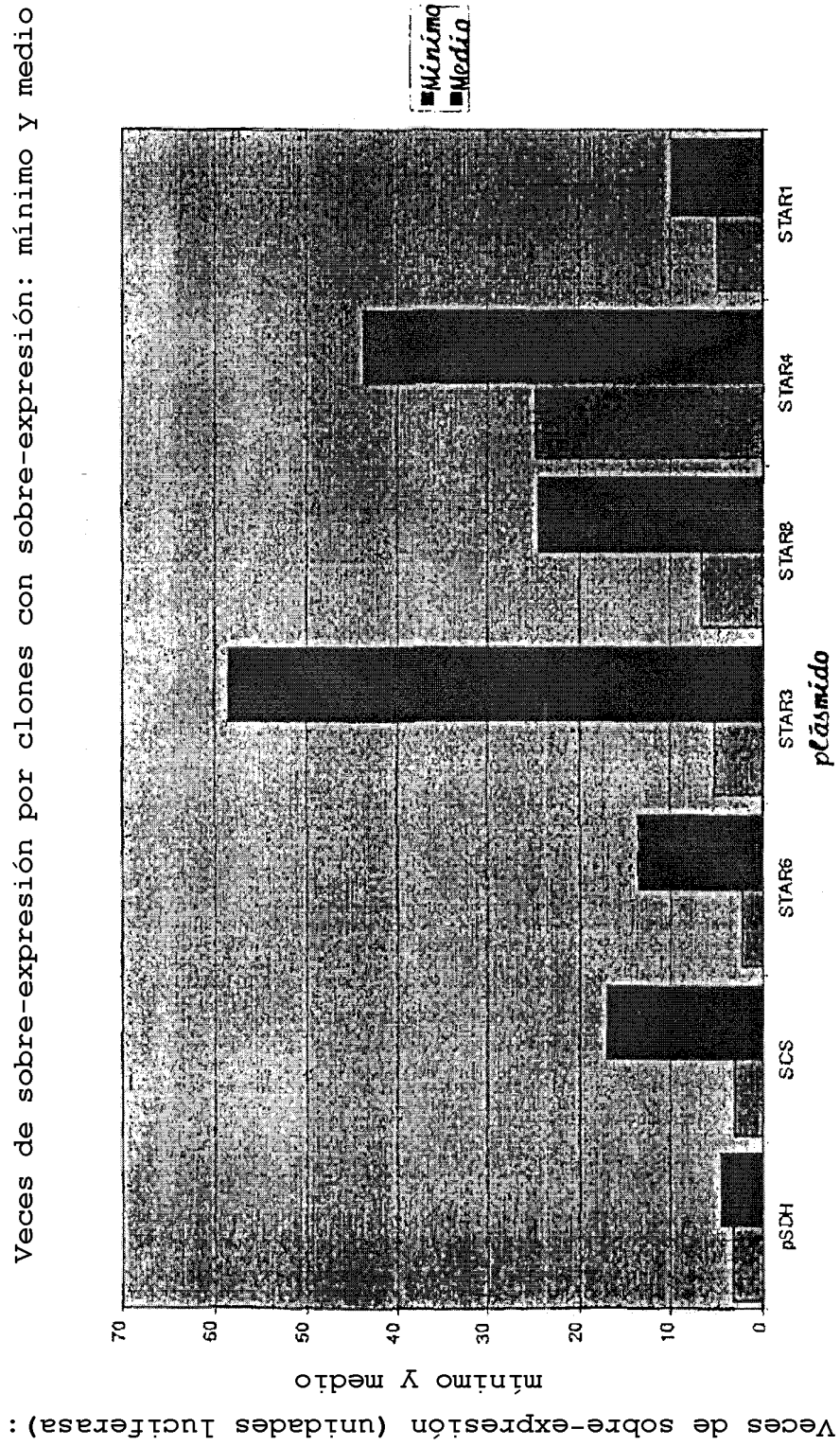


Fig. 5

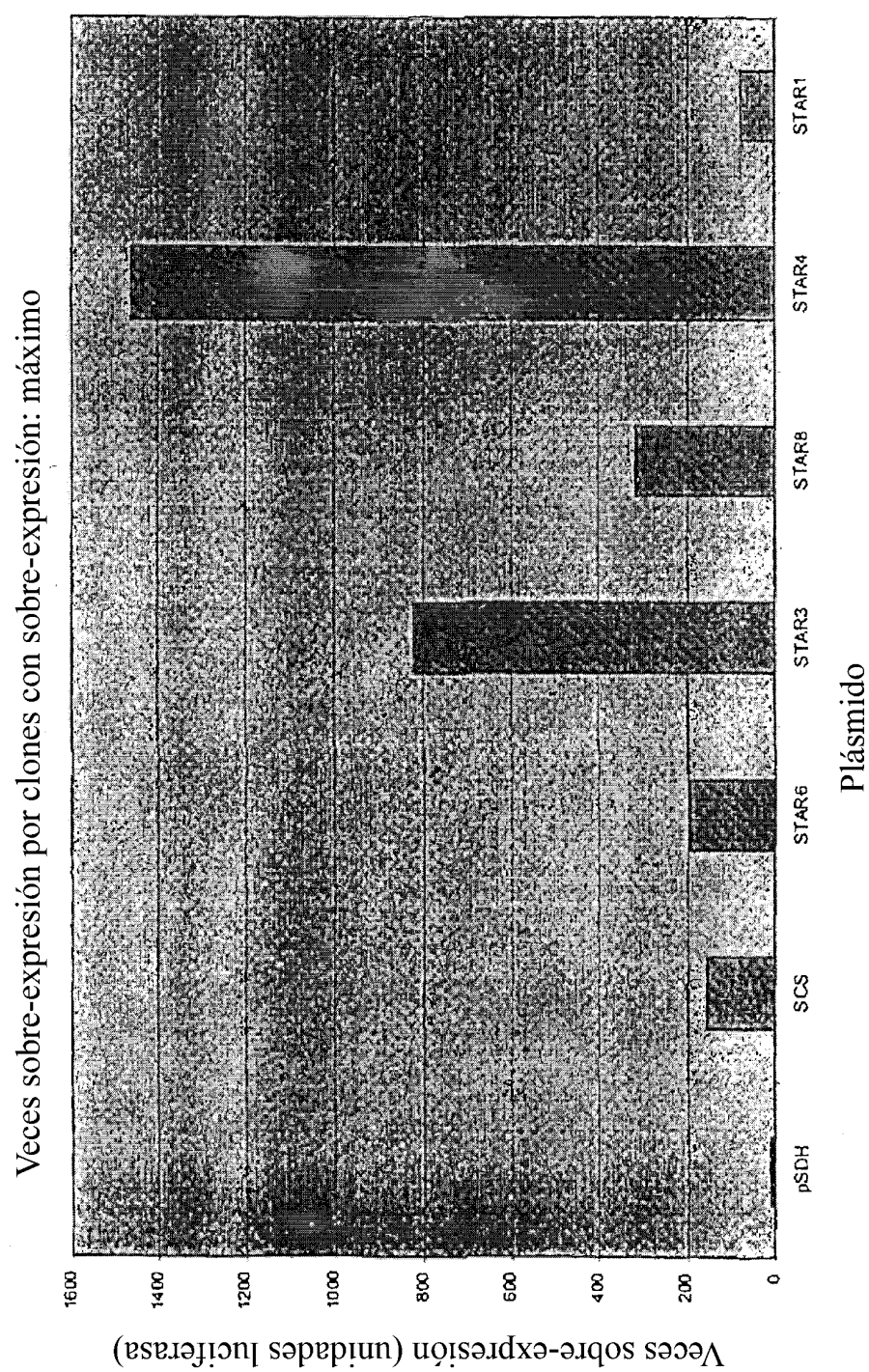


FIG. 6

Vector para someter a ensayo actividad
STAR

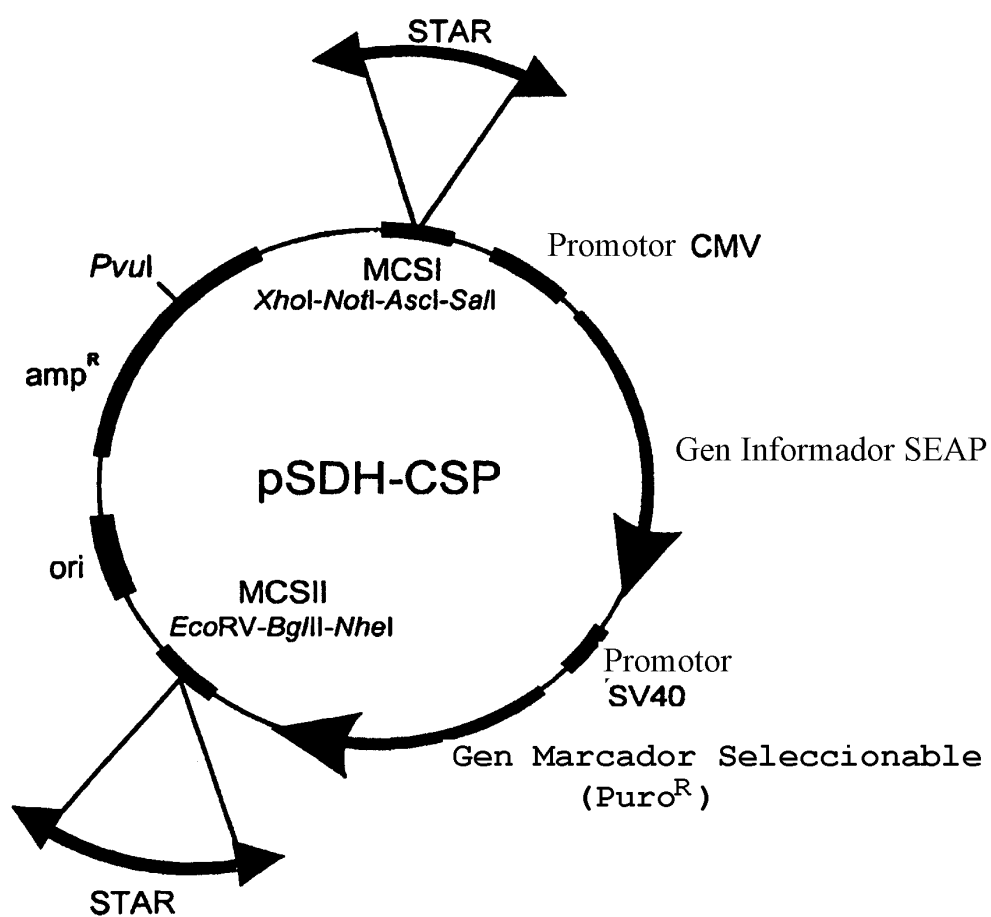


FIG 7

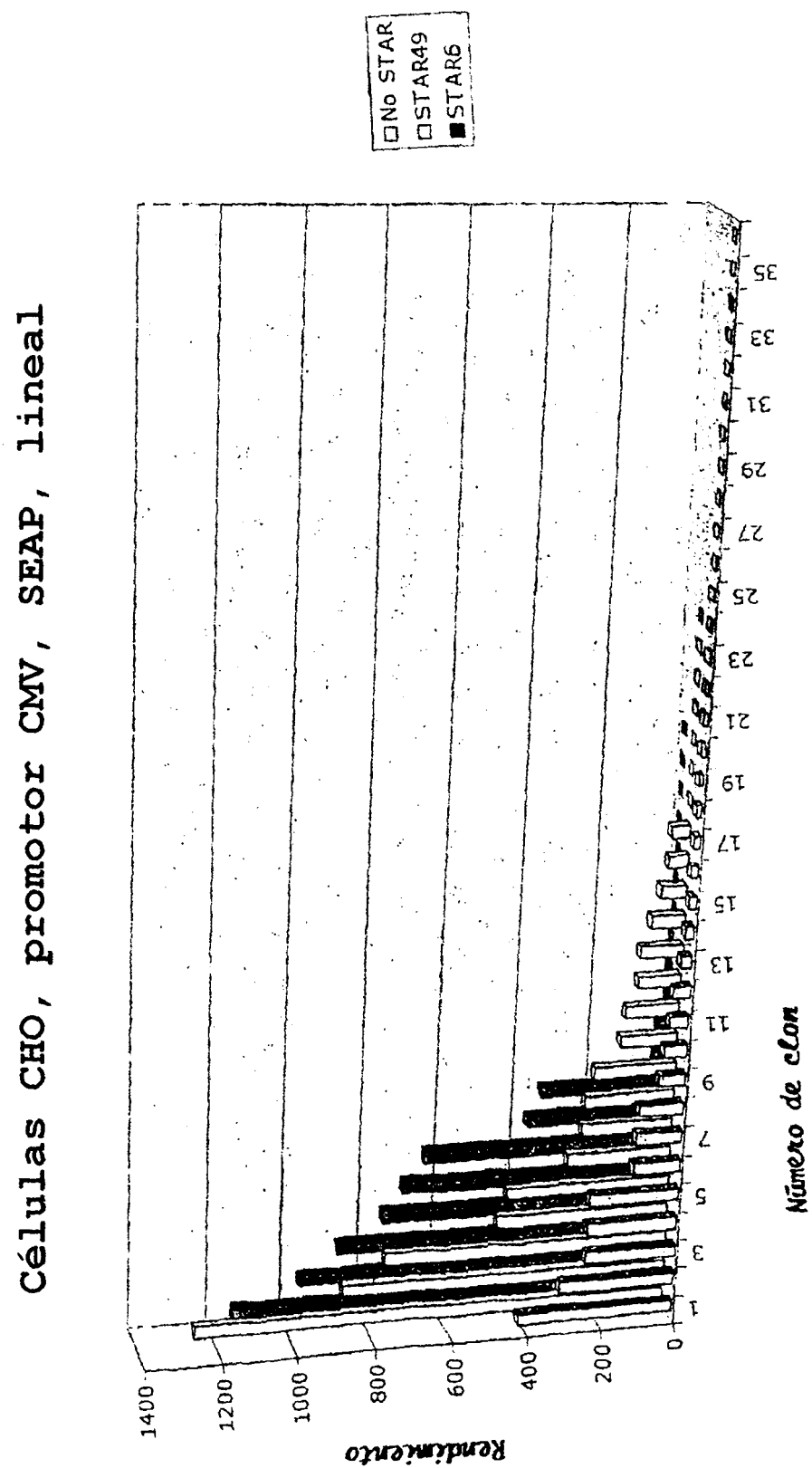


FIG. 8

Células U-2 OS, promotor CMV, luciferasa, circular

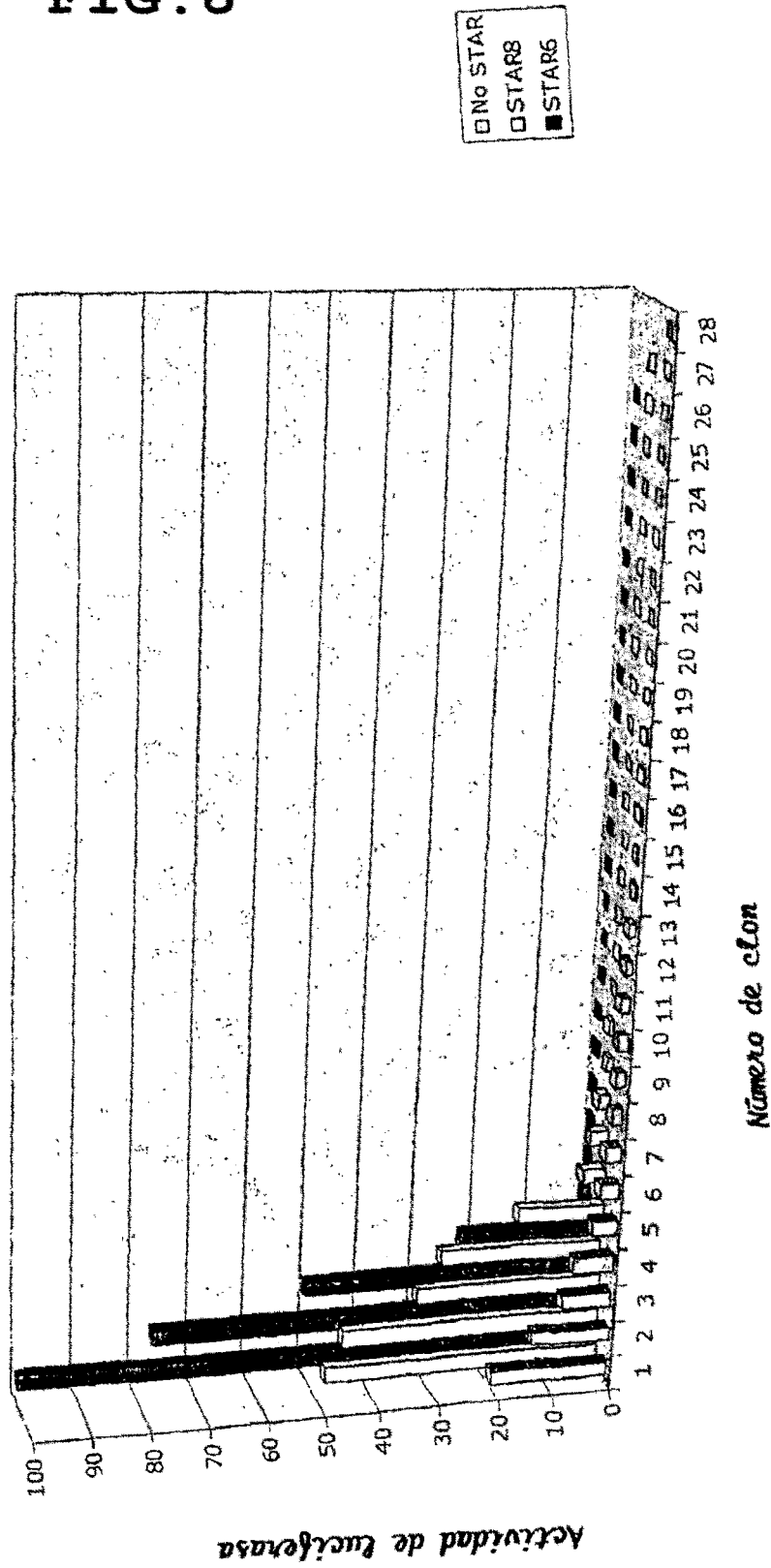


FIG. 9

Mapeo fino STAR









		<u>Crecimiento</u>
STAR10		++
10A		+++
10B		-
10C		-
STAR27		+
27A		-
27B		+/-
27C		-

FIG. 10

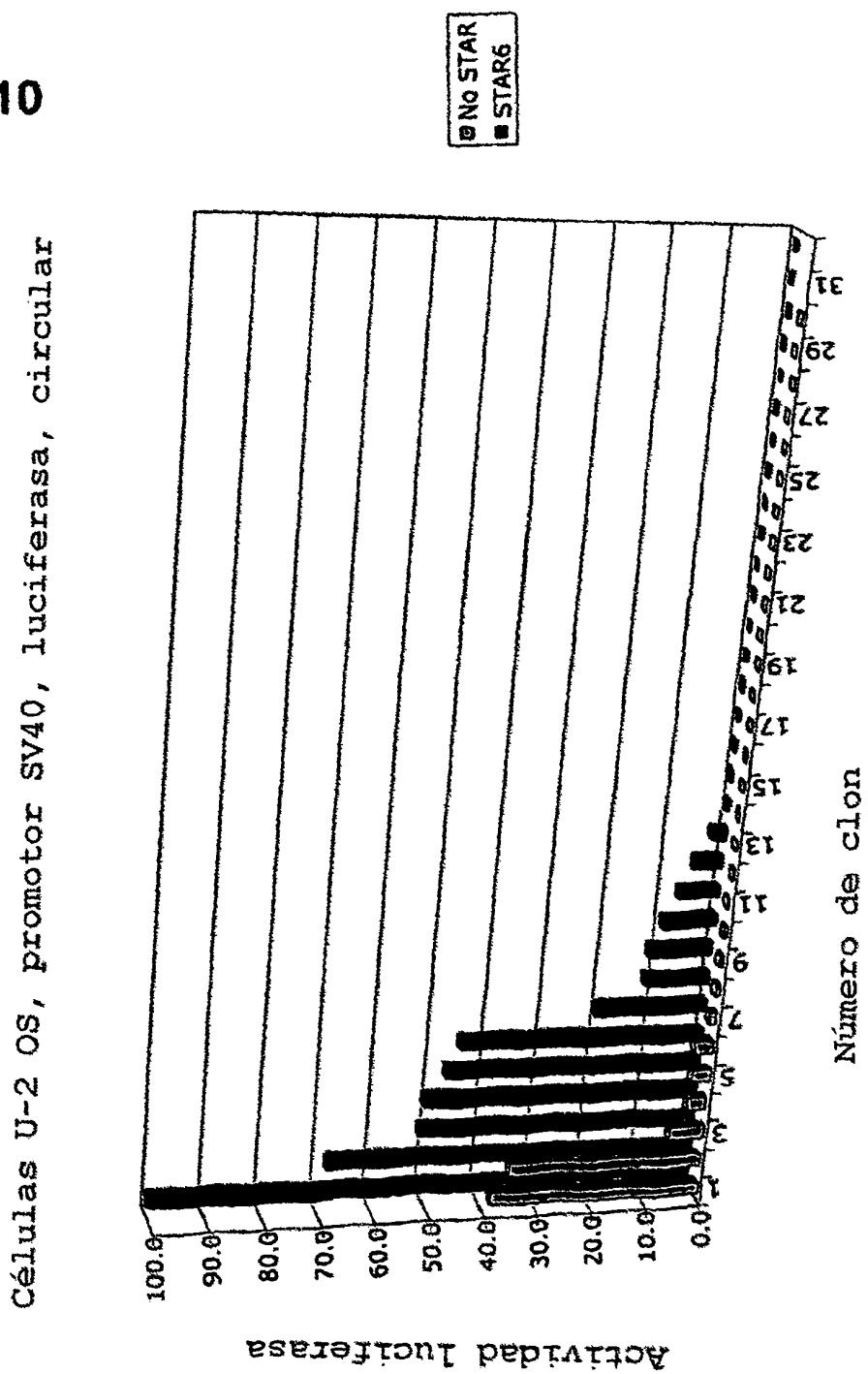


FIG. 11

Células U-2 OS, promotor Tet-Off, luciferasa, circular

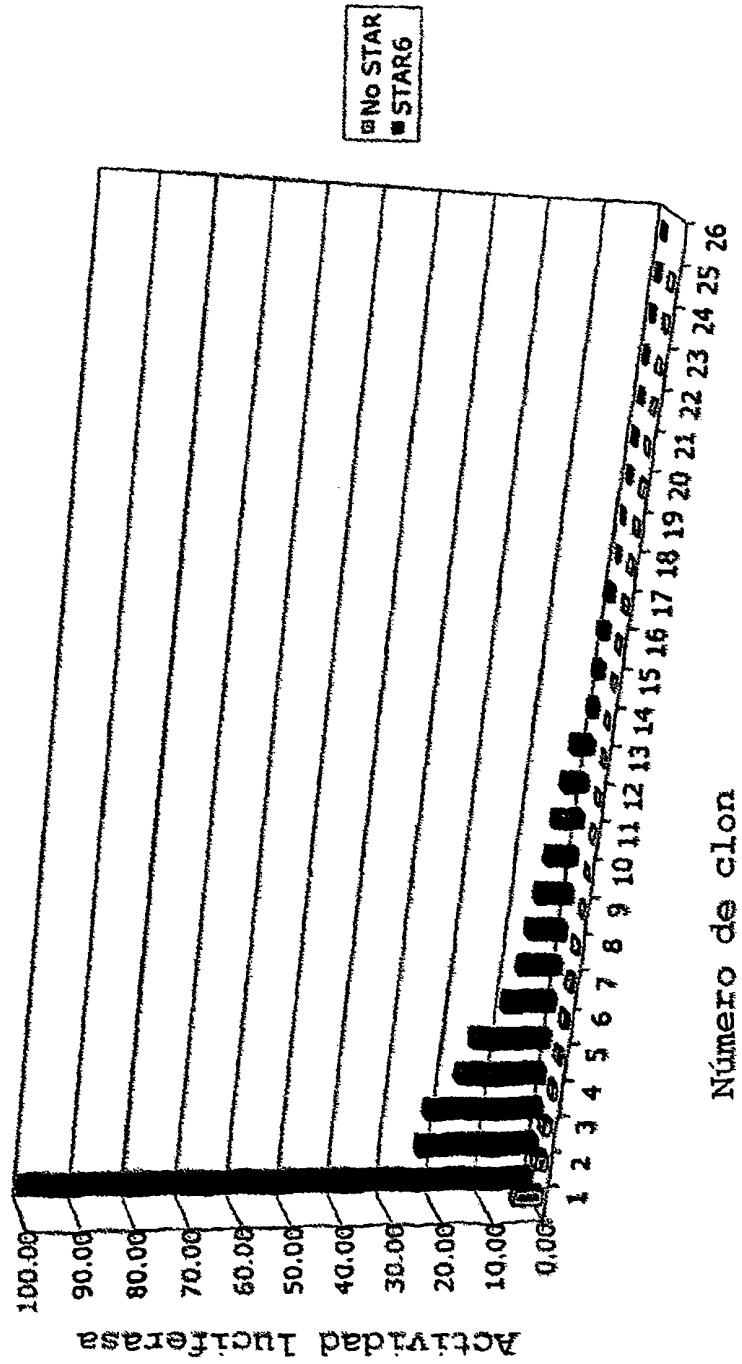


FIG. 12

Orientación del elemento STAR

A. Vector pSelect con elemento STAR clonado:



B. Vector pSDH, STAR en orientación natural:



C. Vector pSDH, STAR en orientación contraria:



FIG. 13

Direccionalidad de la función STAR



FIG. 14

Dependencia del número de copias STAR

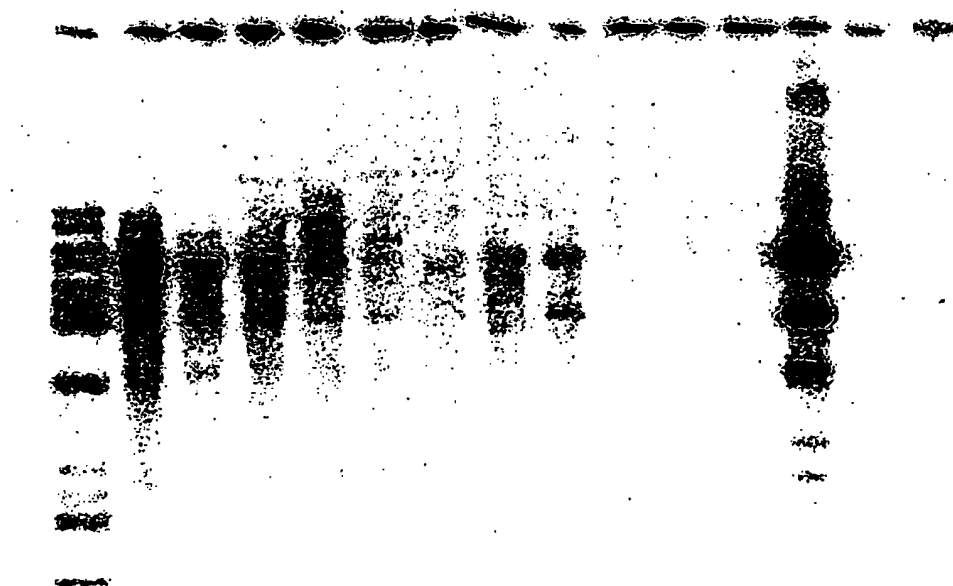


FIG. 15

Dependencia del número de copias STAR

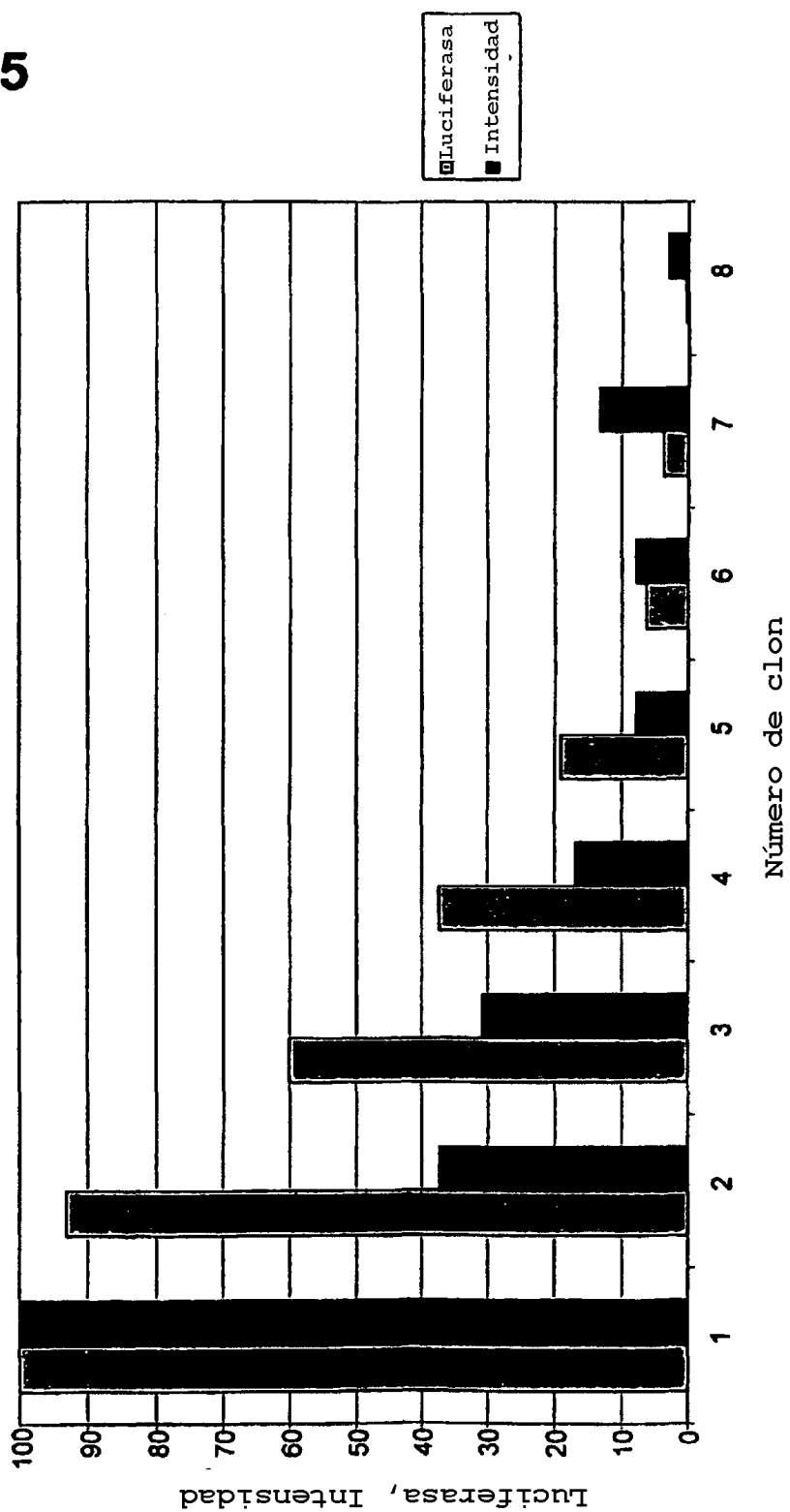
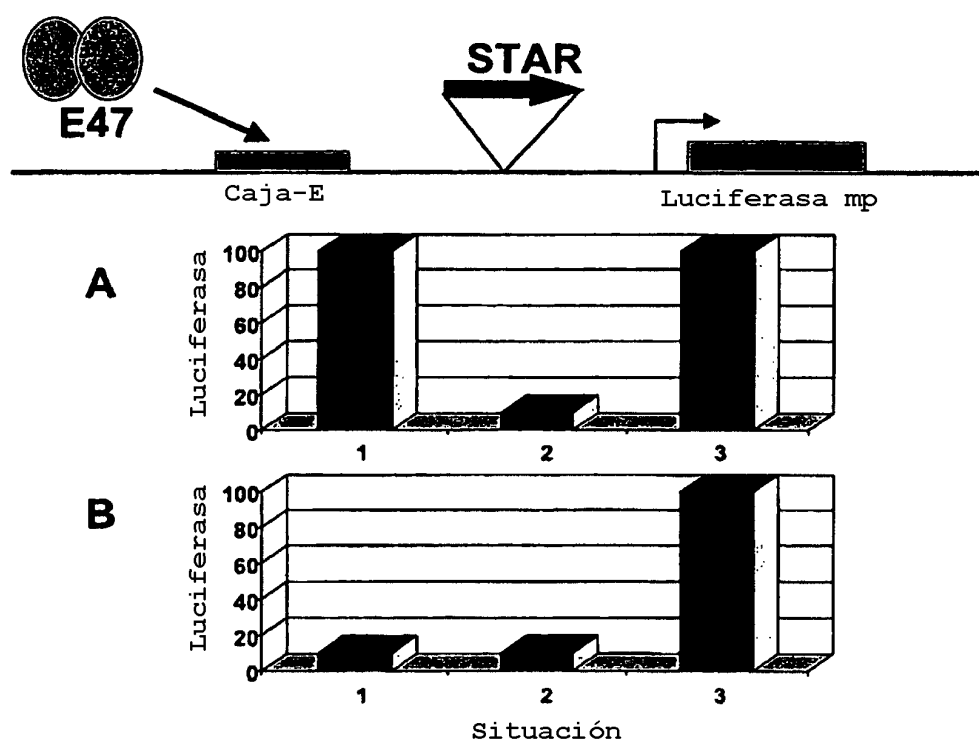


FIG. 16

Diagrama esquemático de Análisis de Intensificador
y bloqueo de intensificador



Análisis de bloqueo de intensificador

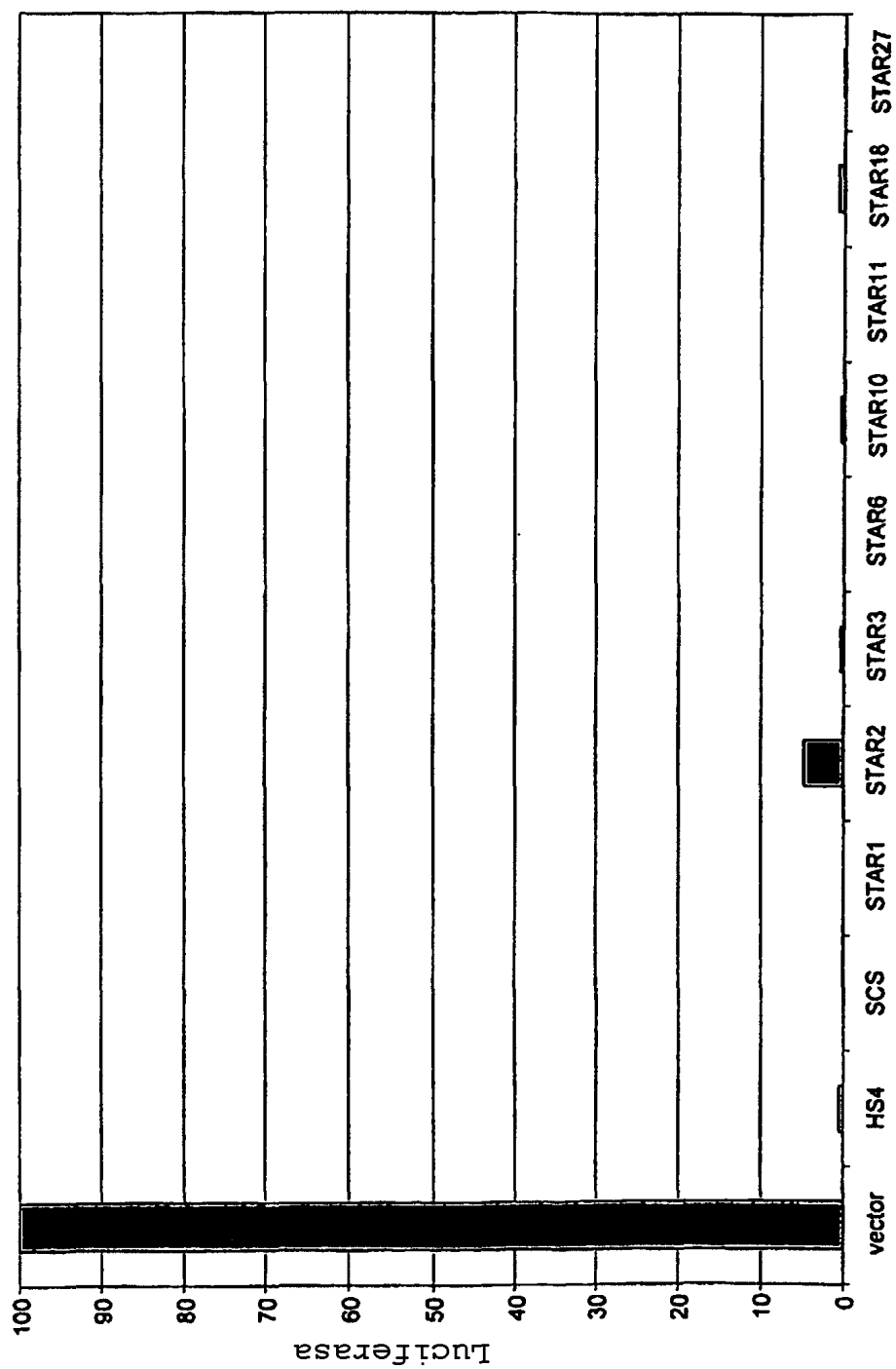


FIG. 17

FIG. 18

Análisis intensificador

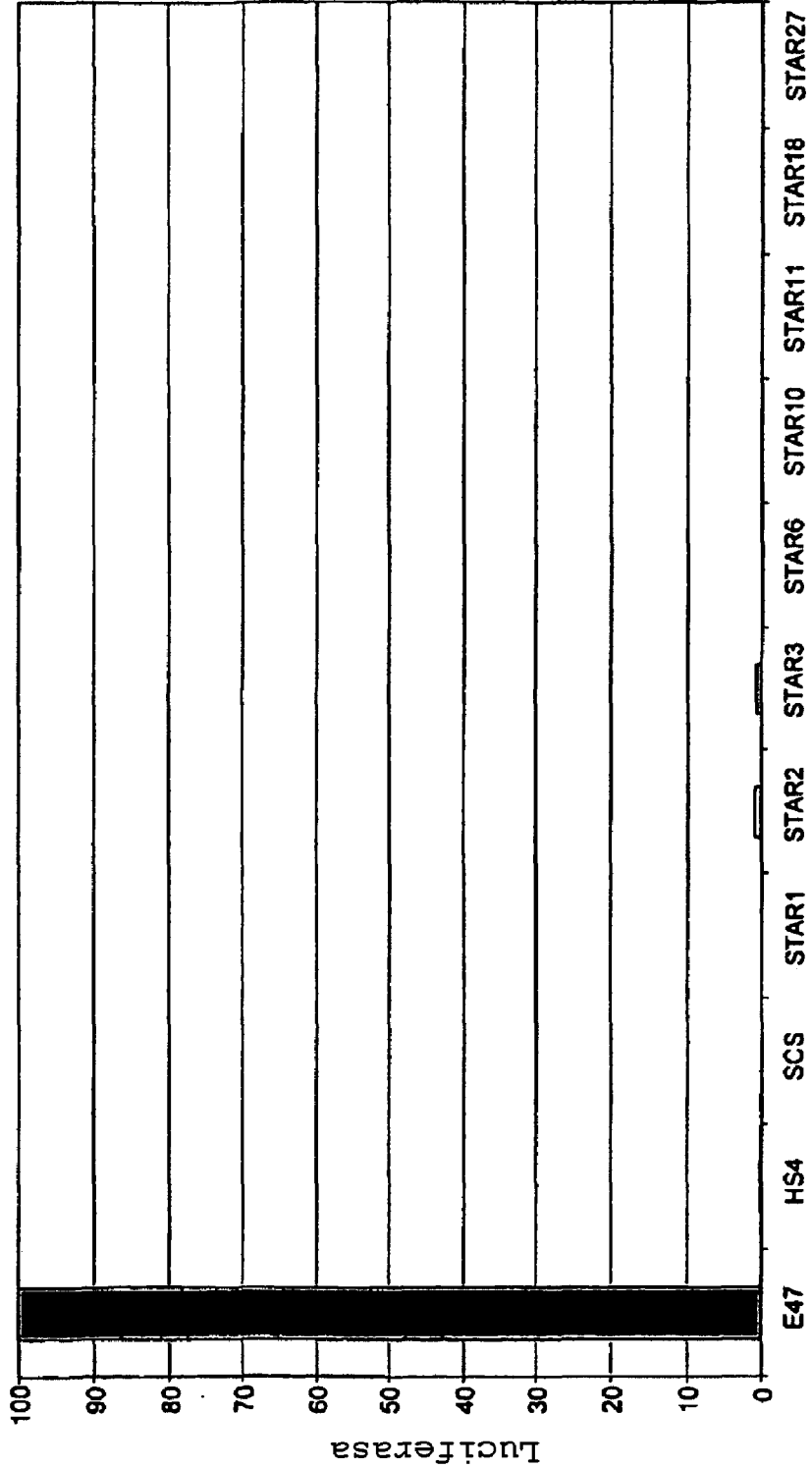


FIG. 19

La secuencia y la función de STAR18 es conservada entre ratón y humano

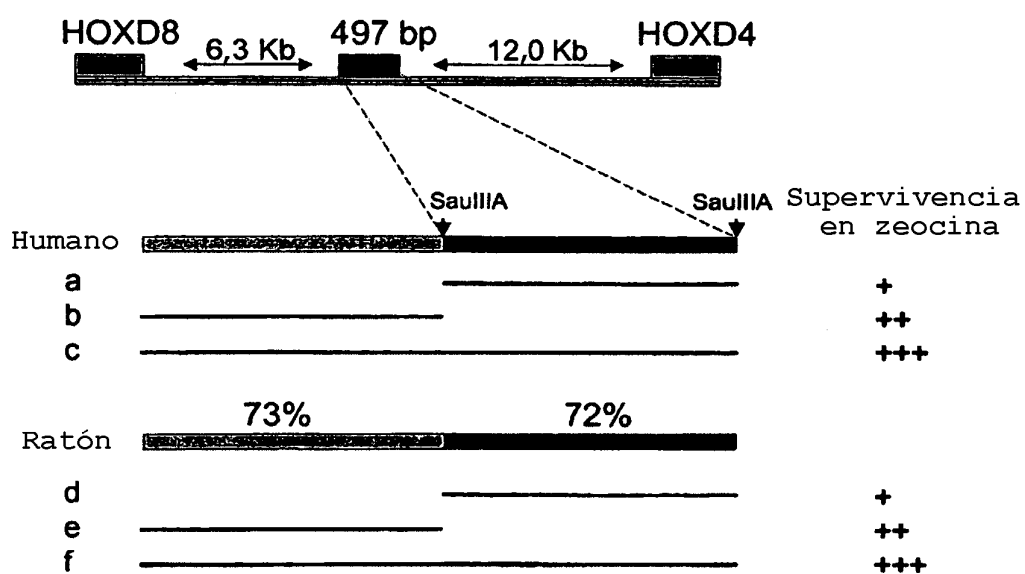


FIG. 20

```

SEQ ID:1
>STAR1
ATGCGGTGGGGGCGGCCAGAGACTCGTGGGATCCTTGGCTTGGATGTTTGGATCTTCTGAGTTGCTGTGCCGCGAAAG
ACAGGTACATTTCTGATTAGGCTCTGTAAGCTCCTGGAGGACCATCTCATTAAGACGATGGTATTGGAGGGAGAGTCACA
GAAAGAACTGTGGCCCCCTCCTCACTGCAAAACGGAAGTGAATTTATTTTAAATGGGAGTTGGAAATGTGAGGGCTGCAGG
AACCAGTCTCCTCCTTCTTGGTTGGAAAAGCTGGGGCTGGCCCTCAGAGACAGGTTTTTTGGCCCCGCTGGGCTGGGCAGT
CTAGTCGACCCCTTGTAGACTGTGCACACCCCTAGAAGAGCAACTACCCCTATACACCAGGCTGGCTCAAAGTGAAGGGGC
TCTGGGCTCCAGTCTGAAAAATCTGGTGTCTTGGGGACCTCTGGTCTTGGCTTCTCTCTCCCTGCACTGGCTCTGGGTGC
TTATCTCTGCAGAAAGCTTCTGCTAGCAAAACCCACATTCAGCGCCCTGTAGCTGAACACAGCACAAAAAGCCCTAGAGATC
AAAAGCATTAGTAAGGAGTTGAGCGGGAGGTGAATATTTAACGCTTTTGTTCATCAATAACTCGTTGGCTTTGACCTGT
CTGAACAAGTGCAGCAATAAGGTGAATGTCAGGTTCACAGCGTCAACAAATATGAAAAATGTGTATATTACCCCGTCTCC
AGCGGGCGCGCCAGGCTCCC

SEQ ID:2
>STAR2
GGGTGCTTCCGTAATCTTCCCTGAGAAGGATGGTGGCCGGTAAGTCCGTGTAGGTGGGGTGGCGCTCCCAGGCCCCGG
CCCGTGGTGGTGGCCGCTGCCAGCGGGCCGGGCACCCCATAGTCCATGGCGCCCGAGGACAGGTGGGGGAGGTGAGTTAG
ACCAAGAGGGCTGGCCCCGAGTTGCTCATGGGCTCCACATAGCTGCCCCCAAGAGACGGGGCTTCCCTGTATGTGTGG
GGTCCATAGCTGCGCTTGCCTTGCAGGCCATGAGCGTGGGGTCTATAGTCCGGGGTGGCCCCCTGCCCGCGCCCTGCCGC
CGTGTAGCGCTTCTGTGGGGGTGGCGGGGTGGCGAGCTGGGCAGGGAGCGAGGGTAGGAGGCGGGGGGAGCGCCGTAGGT
ACCTTGGGGGGCTTGGAGAAGGGCGGGGGCGAGCTGGGGCTCATACGGGACCTGTGTACAGCGAAATGCAATAGAGTTACG
ATAGCCACCGGCTCCGGGGGGGACGGGGCTGGGACTTGGAGACTGGCCCGCGGATGACGTTAGCATGCCCTTGGCTTCTG
ATCCTTTTGTACTTCTATCGCGGATTCCTGGAACAGATCTTGATCTGGCGCTCAGTGAGGTTCAGCAGATTGGCCATCTC
CACCGCGCGCGCGCCACAGGTAGCGGTGAAGTGAAGTCTTCTTCCAGCTCCACAGCTGCGCGCTCGTGTAGGCGGT
GCGCGCGCGCTTGGAGCAAGGCTGCCCGGGGGCTCTTGTGCCAGCGCAGCTTTCGCCCTGCGAGGACAGAGAGGGAAG
AGCGCGCTCAGGGCTGCCCGGGCCCGCCAGCCCTGACCCAGCCCGGGCCCTCTCCACAGGCCCA

SEQ ID:3
>STAR3
ATCTCGAGTACTGAAAATAGGAGTAAATCTGAAGAGCAAAATAGATGAGCCAGAAAACCATGAAAAGAACAGGCACTACCAG
TTGATTCCACAAAGGACATTTCCCAAGGTGAGAAGGCCATATACTTCCACTACCTGAAACCAATTTCTCTGTATGACAGATTAGC
AAGGTTATAGGTTAGCAAAAGATTAGACCCAAAGAAAATAGAGAACTTCCAAATCCAGTAAJAAATCATAGCAAAATTTATTGAT
GATAACAAATTTCTTCAAAGGAACAAAGGCAGAGTCGTGCTAGCAGAGGAAGCAGCTGAGCTGAAAACAGCAGAAAATCTGCTT
TGTTTCTATGACAGCAGGACATAAAGTACACACCACCACTGACCTATTAAGGCTGTGGTAAACCGATTCTATAGAGAGAGG
TTCTAAATACATTTGGTCCCTACAGGCAAACTGCAATTCGCTCGGAACGATGTCCTGGAATAATTTGATGTCAGATATAGAA
AAGCAGAGCAGTCAAAAATATAGATAAAGCTGAACAGATGTTCCTGGGCAATGTAGCAGCACCACACTTAAGATATA
ACCTCAGGCTGTGGACTCCTCCTCGGGAGCGGTGCTGCCGGCGCGGGCGGGCTCCGCAACTCCCGGCTCTCTCGCCC
GCCCTCCGTTCTCTCGGGCGCGCGCGGGGGCGGGGCTGCGCGCTCACAGCGCGGGCTCTTCTGCGCGCGGGCTCGGA
GGCAGTGGCGGTGGCGGCCATGGCTTCTGCGTTGCGCGATGTGAGCAATTCGAACCTGAGGGTCACTCTCTTGGGACTGGT
TAGCAGTGGGTGCGGCCACGGAGGGCGAGTTGAAGCAGGGTGGGGTGCACCTCCCCAGGAAGTCCAGTGGGTGAGG
AATCCCTCCCTTAGCCAAAGGAGGCGGTGAGGGACTGTGCCCGGTGAGAGACTGTGCCCTGAGGAAAGGTGCACCTTGGC
CCAGATATACACTTTTCCACCGTCTTCAAAACCCGACAGCAGGAGATTCCTCGGGTTCTTACACCACAGGACCTTG
GGTTTCAACCACAAAACCGGCCATTTGGGCAGACACCCAGCTAGCTGCAAGAGTTGTTTTTTTTTTTATCTCTCTGTGGC
ACCTGGAACCGCCAGCGAGAGAGCACCTTTCACTCCTCCCTGGAAGGGGGCTGAAGGCAGGGACCTTTAGCTGCGGGCTAGGG
GGTTTGGGGTTGAGTGGGGGAGGGGAGGGGAAAGGCCCTGCTCATTTGGCGTCTGCAAGCCAAATAGGCTACGCTCCTC
TGTGCGAGTAGACCCAAATCTTTCTTAGAGGTGAGGGGGCGGGTAGGTGGAAGTAGAGGTGGCGCGTATCTAGGAGAG
AGAAAAGGGCTGGACCAATAGGTGCCCGGAAGAGGGGACCCAGCGGCTGTGTGATTGGTATTGGGAGTGGACCTCCCC
CGGGGTGGTGGCGAGGGGGGATGATGGGTGAGGGGTGTGTTATGTGGAAGCGAGATGACCGGCAGGAACCTGCCCA
ATGGGCTGCAGAGTGGTTAGTGAAGTGGGTGACAGACAGACCCGTAGGCCAAAGGGTGGCTTAAAGTGTCTTTGGTCTCCTC
CAATGGAGCAGCGCGGGCGGGGACCGGACTCGGGTTTAAATGAGACTCCAATTTGGGCTGTAAATCAGTGTCAATGTCGGAATC
ATGTCAACGACAACAAACAGGGGGACACAAAATGGCGGGGGCTTAGTCTTACCCCTGGCGCGGGCGGAGCGGTGGCGGAGG
CGACGGCACTCCTCAGGCGGGCAGCGCAGTTTCTCAGGCGAGCGGAGCGCCCCGGCAGGCGCGGTGGCGGTGGCGGCGG
GCCAGGTCTGTACCCACCCCGCGCTTCCAGGGGGAGGAGACTGGCGGGGAGGGGGGAACAGACGGGGGGGATTTCAGG
GGCTTGCAGCGCCCCCTCCACAGGCGCTTGGCGGAGGGTCAACCGGGGGCGCTCGGGGTGAGGCTGCCCTGAGCGTGAC
GGTAGGGGGCGGGGAAAGGGAGGAGGACAGCCCGCCCTTGGCAGGGCTCTTAGGGCAAGGGGGCGGGGTTCGAGG
AGCGAGGGGGCGGGGGCGG

```

```

SEQ ID:4
>STAR4
GATCTGAGTCAATGTTTAAAGGGAGGATTTCTTTGGCTGCTGAGTTGAGATTAGGTTGAGGGTAGTGAAGGTAAAGGCAGT
GAGACCAGTAGGGGTCAATTCAGTAAATCCAGGCTGGAGATGATGGTGGTTGAGTTGGAAATAGCAGTGCAATGTGCTGAAC
AACCCTCAGCTGGGAAGCAGTATATGTTGGCTGTGACCTCAGCTGGAACAGCAATGCAATGTTGGTGGTGAATGACCCCGAG
TGGTAGGGTGCAATGTTGTTAACGACCTCAGCTGGGTAGCAGTGTGTGATGTAAACACCTCAGCTGGGTAGCAGTGTAT

```

FIG. 20, cont

CTGTGATAAAATGTTGGCACTACTAGATTTTGTTATGAGGGTAGTGCCATTAAATTCCTCCACAAATGGTTGTGCACGTATG
 GTGAAGAAAGGAAGTAGTGAAGAAATCTCAGTGCTTTTGGCCGTGAATTAAGTAAGACGCTCATTTCCAGTTAAATGGAGACA
 GGGAAAGCTAAAGGTAGGGTGGGATTCAGTAGACAGGTGTCAGTTTGAATATGATGAATCTTGAGAGAGGAAAACTT
 TTTCACCTCTTTAGTTTGTGACCTGGAATTAAGAAATTAAGATGACATTAAGACAGATTAACAGACAAAAATATGCCAGGT
 TATTTAAATATTTTACTTGCGAGGGAAATCTCAAAAGAAAAATGAAGACCAAGAGGACATAGGGCTCAAAGACTCAT
 ATGCCCTTTTAAAGTAGAAAAATGATAAAATTTTAACAATGTGAGGAAGACAAAGGTGTTTGAGCTGAGGGCAAATAAAATTTGGCT
 ACAGTGATTAAGAAATAATATGGGGAAATGATTAATTTAGTAGATTTATCTTCATATCTATTTTGGCTTCA
 ACTTCAGCTCTCTAGTAGTAAGAAATGTTCTCTCTCTCGTGATCAGAGAGGACCTTTCTCATGGGAAATTTATGACCT
 TGTCTTAAGTAGTAAGAGGGAAGATCGATCGATCTTTTCCAGACATCAGGATCGCAAAATTTCCCTCCATCTCAGCTCTCAA
 CCGCATGCTGGGGCTCATGCGATCTCAGCACTCGTATGAGTGACCTCTTCTCGACGGCTGCCCTGGGTAGCTGGTGCACT
 GCTAGGTCACTCATGTGACAGGAGCTGGGCTCTGGGCAATGCCAATTCGAGGACGCCCATCTCCCTTCACTGGCTGGGGGG
 TCTCATCCGAAAGGGCTTGGTGTGCAGAAAGTGGTGCACCACTCATTTCCCACTTGCCATCTTCAAGGGGACAGCCAG
 CTGCTTTGGGGCGGGCAAAAAACCCAACTCACTCCTCTTCAGGGGCTCTGGTCTGATGCCACCAAGGACATCTTGA
 GTCTGGGGCACTCTGAGGACAGGAAAGGATGATGACCAACAAACAGGAAATGGCAGCTGAGTGCAGGAGGAGTCAAAAG
 GCTTGTTGCTCTGGGGCTTCTGATGAGGCTGGCGAGCCCACTTGGATGGGCGCTCAAGTCCCTGGTGGCTGCAAGAGGCAAGC
 CCTCTGCCCATGAGGGGAGCTGGCAGTGACCAAGCTGCACCTGCCCTGGTGGTGCATTTCTTGCCCCACTCTTCTCTCTAA
 GATCC

SEO ID: 5

SEQ ID
3-STAR5

[illegible]

SEQ ID: 6

STAR6

TGAACCCACCCAGACATCCCCCTCTGGCCCTCTTGAGTGGTTCTTCAGCAGACAGCTTCCAGAGCCAAATTAAGCGTTCACTCT
 ATGCTATATAGACA AAAAGGCTTTTGAC TAAACTCTGTGTTTATAGAGGGAGTAAATCGTTTAACCTTTTAGGGGGTGGG
 CAGAGGGGATGACAAAATAACAATCTGTCTGAATGTTTACATTTCTCCCACTGCTGCAAGAAGGTTTCAACAGAGGTCAT
 CCA TGA TAAGGAGTAAGAACCTCCAGCGCGCACTCTCCCTGGGCCCCAGAGGACACTCCAGACAGA TATGCTTAACCTGGAC
 TGGAGATGGCTTCACTACTCCAGAAAAGACA TGGAGCAGCGAGCTACA CAGACAGGCCCAAGGTC CCAAGGACAGAATGCT
 AGGAGGGAGATTGGGGTGAAGGTAATCTGATGCAATTA CTCTGTGCACCTCAACAATTTCAAGGGAGGGGGAAAGAAAGAAAC
 TCCCTGCTCAAGTAAGTTGTGCAGCAGACAGATGGTAAGCTCCAAAATTTGAAACTTTTGCTCTGCGAAAGTTTAGGGGGGCG
 AGATAAGAAGACATAAGAAGCTTTAGAGGTTTACTACACATGAGACGCTTATGCAATTTATTTATTTATTTCTCTTATTT
 ATTAATTTGTA TAACTCTATAATAATCTTATGA AAAAGCGAAAGCCCTCATATACCCATTTTACAGATAGAAAAGTACAGAA
 TTTTGAAGCAGCATACCTAAGAAATAGCTAGTAAATGAAGAAGCTTGGCACTTAACCAACCC TATCTCAACAGAGTACACACT
 CTTTTTTTTTTTCCAGTGTGAACTTTTAA TTTTATTTATTTATTTAAAGTTCTGGGATATAGTGCAGAAAGTATAGTTTGT
 TTACATAGGTATATGTGTAATAGTTGGATTGCTGCACCTTATCAACCGCTCATCTAGGTTTACAGGCCCACTGCATATAGCT
 ATTTGCTGTATGCTCTCCCTCCCTCCCAACAGCAGGCGCTTGGTGTGATGTTCCCTCCCTGCTGTGTCATAGTGT
 CTCACGTGTCAGCTCCCACTATGAGTGAGAAAGCTGTGATTTGGTTTTCGTCTCTGTGTAGTTTGTCTGAGGATGATG
 GCTTCCAGCTTCTATCTATGCTCCGATGCAAAAGACAGCAT

SEO ID:7

>START

ATCATGCCAGCTTAGGCGACAGAGTGAGACTGGACATAATAACAATAATAATAAAAAATAATAATAAAACAATATCTGAGAGAAAAAATTTGATCATATAATAAGAAAAATAAGGTTTGTGGCGTGTTGTTTGTGTTTACCTAAGACAGAGCTGTCCCTCTCATTTGGGTAGTTTTATTTGCAAGACAGAAATCATCTCCGACATGATTCGAGGTTGATGAAAAAGATGAATATGAATCCACCTTCTGCCATCTATTCACTTGTCACATTTAATAAGACACTTACATGCCATTATTAGCATGTTTCTCTCCATACCAATGAGTATGTAACAATCAAGAGATTAATAATAACACAAATAAGAACATGAAGGTTATCAAAATGTTACATACAAATAATTAACACAAATATATTAATAAATTTCTCGGAAAAAGCATATGGCTTACTCTCAAGGTAAAGCTCATTTTCTGTGTTTAAACCTTTTCTCAA

FIG. 20, cont

GTTGAAGTCTCTAAGTTTCAACCCGTAATCTATTTGCAAGTTTGTGCAAAATTTAGGGATTGAATCCATAGTAATTAGTG
 ATTTATTGTGGTGTAGGGAGACAAGTCAAAAGAAATCAGGACTGCTAGGTAGATGACTAAGGAAAGGATGGTTACAGAGGTG
 ACATAAAGCACTCAGAAAGAAAAGGTGAGGAAACGGAGGACAGAAAAAACCTAAGTTCTGCTGGGTGATGCTGAATTTGT
 CATCACAAAATCTGCATTGTGGAAGCTTTAGCTATTGAGGAGATTGCTCAAGTGTAGAACTGAGAACAAATAGGCAGTGAAC
 CCGAGAGAACATCAAGAGACTGAGAGAAAATGAACCAAGCTTCCAGGTGCTCCATGTTCCAAACCAACATTTGTATTGTCA
 GAAGGAATTGAGAGGCAAAAGGAAACCAATAAAAAATAAACAGGAAAGGGCATACATGATTACCAACCCCTTTTCTCACC
 AGCTGCTCATGGACCAGCTTCTCCTAGTGCTATTTTCTTGGTCACTGCATCACTCTGCTAACATAGTTTCCCACTAGCT
 CTGAGGCTGTCCAGAGGGGAGCCAGCTGTCTCTCTCTTCCACACTCTGTTGGAGGAACTGTCTATTAGCAGCTCC
 TACTAAACGCATTTATGACAAACAGGAGGAGATAAATACTAGAAAGTGAACAACTCAAACCTCAGAGCCTCTCATTGT
 TATGAATGCCCTTGTAGGTCTTGGGCTATTTTAAATTTATAAATGTGTATTTTCTTCAAAGAAAACCAACCAATTTG
 TATAAGTTCAGAAATCTGCAAACTGAGGTCCATCCATGCACTCAGGATACATTCATAGCATCTCTGAGCTGGAAAATATC
 TTAAGGTCATATATGTCTCCAACTGCAAGAATCTCTGCGCAGCTTCTTTAAATCATCATCTAAAGAGGGGAAA
 TCCCAAGCTGTGTTTGGATTTTGCTCTGTCACTTTGCCAGTTTCCCACTCCATAAAAGGGCAACAATATGAATTTCTGTAT
 AAGGTAGTTGTTAATAATAACAAAGTGCAGTCCACTTCCCTAAGAAAAATATGGGTTTCTGCTTCAAGTCTAGGGA
 GAGGAAAAAAGGGGGGTGAGAGTGATTATTTATCATTTATATTTGGAATGTTTTCAGACATAAAAGCTCACCAGC
 TCTTAGGCCACAGACAGATGCATTATGAAAGTTAAGCTAAGTCTTCTCATCATGAGCTGCACCTATATCCCCATTACTCTT
 CTAGAACTGCATAATTTATTTATTTCTTCTTCAAAGTTTGAAGAGCCATTCTGTCTCTAAGATTTTTTTTTTTTTTT
 TTGGAGACAGAGTCTCCGCTGTTGCCAGGCTGGAGTGCAATGGCACTATCTCAGCTCACTGCAACCTCTGCCTCCGAGA
 TCAAGTGATTCTCTGCTCAGCCTCCGAGTAGCTGGGATTACAAGCACGCCACCAACCAACAGCTAAATTTTTCTGTAT
 TTTTAGTAGAGACGAGGTTTACCATTGTGGCCAGGCTGGTCTTGAACCTCTGACCTCGGGTGATCCACCCACCT

SEQ ID:8

>STAR8

GAGATCACCTCGAAGAGAGTCTAACGTCGTAGGAACGCTCTCGGGTTCAAGGATTGACCGAACCCAGGATACGTCGC
 TCTCATCTGAGGCTTGCTCCAAATGGCCCTCCACTATTCCAGGACGCTGGGTGCTCCCCCTAACTCTCCCTGCTCTCCTG
 AGCCCATGCTGCTATCACCCATCGGTGCAGGTCCTTTCTGAAGAGCTCGGGTGGATTCTCTCCATCCCACTTCTTTCC
 AAGAAGAGCCAGCTTCCAAGACCCCAATGGGACATTTCCCTTCCACCTCTCTTCCAAAGTTGCCAGGTGTTCTATC
 ACAGTTAGGGAGAGAGCCGCCAGGTTTCAAGTACAAGCATAGGACGCTGGCATGAACACACACACACACACACACAC
 CACACACACACACACAGGCTCGAAGAGGTAGCCACAAGGTCATTAAACACTTGACGACTGTTTCCAAAAACGTGGA
 TGCAGTTCATCCAGCCCAAGGTCGAAGCAACACGGAATGGTGGAGAGATTCCAGAGGCTCACCACAAACCTCT
 CAGGAATATTTTCTGACCTGGGGGAGAGGTTGGAAACATTGAGGACATTTCTGGGACACAGGAGAGCTGACCGAC
 CAGGATTTTCTTCCACTGCAAAATGACCTATGGCGGGGGCATTTCACTTTCCCTGCAAAATCACCTATGGCAGGTACC
 TCCCAAGCCCCCACCCTCACTTCCGCAATCGGCATGGCTCGGCTCTATCCGGGTGCTCACTCCAGGTAGGCTTCTCAAC
 GCTCTCGGCTCAAAGAGACAAATCACAGGTCCAAGCCCAAGCCCACTCTCTCTTTGTATACCCACAGAAGTTAG
 AGAAACGCCACACTTTGAGACAAATTAAGAGTCTTTATTTAAGCCGGCGGCAAGAGATGGCTAACGCTCAAAATCT
 CTGGGCCCGAGGAAGGGCTTGACTAACTTTCTATACCTTGGTTAGGAAGGGAGGGGAACCAATGCGGTAATTTCTAG
 AGAAGTAAAAACATCGAGGAATCAAAAGAACCAATGGTTATAGAGAGATAAACAGTTTAAAGGCAAAATGGTTACAAA
 GGCACCGGTACAGGTGCGGGCTCTAAATCTTATGACACTTAGATATAGGTGCTATGCTGGACACGAACCTCAAGGCTT
 TATGTTGTATCTCTCGAGAAAAATCTGGGAACCTCATGCACTGTTTGTGCCAGTATCTTATCAGTTGATTGGGCTCCC
 TTGAATCTGAGTATCTGCTTACACAGGTCAACTCTTGGGAAGGGGTTGGTAAGGAGCCCTTCTGTCTCGTAAAT
 TAAGGGGTCGATTGGAGTTTGTCCAGCATTTCCAGCTACAGAGAGCTTATTTACATGAGAAGCAAGGCTAGGTGATTA
 GAGACCAACAGGGAAGATTCAAGATAGCGACTTAGAGTAAAAACAAGTTAGGCAATTTCACTTTCCAGAGAACGCGCAA
 CATTCATCAATGGGAGAGAGGTCCCGAGTCTGCAAGTCCCAGATGTGGCGAGCCCCGGGAGGAAAAACCGTGTCTTCTTAC
 GATGCCCGGAACAGAGCTAGGCTTCCGAGCTAGGCAGCCATCTATGCTCGTGAGCCGGCGGGAGGAGACCGCGGGAG
 GCGAAGTGGGCGGGGCTCTCTTCTTCTGCTCTGCTGCTGCGGGAGCTCTGCTGCGCTCAAGCGCAGAGGCGC
 GCGCTCTGACGGGCGCGTAGAGTTTGGGTCAGAGT

SEQ ID:9

>STAR9

ATGAGCCCCCAAAATGATCCTCTGGCTTATGACAACCTGATGCAGCCAGGAAATGCTGCAACATGCCCACTAGCAGCT
 GCGAACCCCTCTGTGAGGAGAGAACGTTTACATTAAGAAACCTTTGTTTGCAGCAGAGACTATTAGGTCACACATG
 TGTGGCTCTCAGTCTTTGAGCCTTTGAAGTTCTCTATCCTTGGTGGAGGCTGAGCTCTCCATGGAAACCTGGTCCGA
 TAGTGAGAGGAGCAGACCTCTGGAAACACCTTTTACACCTGACCAAGCAGCCAGTCTATGGGCCAGTGATGCAACAAAG
 TCAACCGGTGATTTCTGGCCCTCAGAAAAGCAGCCCCGGGAAGGTGAGGAGGAGCTGCTGACTCCCTCTTCCCTGCA
 GCCGCCCAAGCACACCCAGGAGCCCTGCAGGTTTGGGTTCAACAGGTGCCAGGAGGCTCCACGATGCTGCATTTCTTACG
 AGCTCTGGAGGTCAGATGGTCTGGTCAGAGGCTGCATTCTGAGTATCAGGAGCCATGGGGCAACGTTTCTGCGATTG
 AGGAAGGGGCAATTTCTGGGCTGGGCGAGAACAAAGGTTTGGCTGAGCTGGAGCATCCGCTCCATCAGTGTTTTCCGGCA
 ACTGTACTATCCATCGTCTTCCCTTCCACAGCTGACCATGGCTTTGGAAAATGCTCTGAACTTTCTTTTCAAGAGATT
 GACTCCCAACTCCACACTTAGGGGAAGTCAAGCTTACTTCTCAGAAATCAGAGAAGGCATAAAAAAGAAATCATTCTAAA
 GGCCTTTAGAACTTAACTTCAAGTCTGACAGCGGCCAGCTAATTTCTGGTGCCTTCCAGGAATCTTCTGACTGCAAAAA
 AAGCATTTACCACTGAACACAAACCCAGTTACAGATAGAAAAACATAGTCATTAAATAGAAATATAAGCATCTGGCCTC
 TGCCCATCATATAGAGTCTTGGCTGGATGGAATTTGTAATCAAGAGGATGAGACAGGATTGTAACCTTGTGCCAATGTGAAA
 GGGTTTGTCTAGGTATCATTTTGTCTAAATGCATGGGTAATTTCCAAAGTTCTTTGGAGCTGAATTTCACAATTTAG
 TGCAGGCTCTGGTGAGCCCACTTGACTTATCTCAGAGTACAATGCAAGTGGCGTGCTACAATGCTGGGCAAGAGAACCA
 ATGTCAACAGCCAGAGGTGGTGGGCTTACCAGGCTCCAGGCAATGCTTATGGTGGGCTTGGGCTGGGAGAACAG
 CACCTTTGCTGGTCCATGAGTATCTGGGTCAACTCTCTGTGACACAGAGGCCATGGCGACAGGCTTCCAGGAAA

FIG. 20, cont

AGAAAAGGGCAGCAGCTGAAATCGTCAGGTGGAGAGGCAGTCATCCTTGCTCAGTCAACTCTAATCCGGCTGCCCTCCTCC
TCAGCTTCAGGGTGAACTCTCCTAAGCTGTGTCTTTGGTATCTGATGGGCATTAGGTGCTGGTGAAGAGCTGGAGGGTCT
CTTTGGGATATTACAGAAAGCCAACTAGCCTTGTAATCAATACTAGGCACCTCACCCCTGAAGTCTACGTTTCCAGA
TTCTGAAAAATGGGAAAGCAATGTGTGATGCTGTGAGGTCCCTCAGCCTTGGTGATGGGTAGGAGGGCTCTAAAG
GGTGGCAGCTCCAGTGTCCAGTGGGGCTGAAGTTGGTCCCTTCCCTTCCAGCTCCCATCTATGGTTAGCCCAATCCC
TTCCGTACCTAAGAGTATGCACATGGATGCTCCACGCAGAGCCTCTGCTCCACTCCAGGAAGTG

SEQ ID:10

>STAR10

AGGTCAGGAGTTCAAGACCAGCCTGGCCAACTGGTGAACCCCTGTCCCTACAAAAATACAAAAATTAGCCGGGGCTGGT
GGGGGGGGCTATAATCCAGCTACTCAGGATGCTGAGACAGGAGAATGTGTGAACCCGGAGGTGGAGGTTCCAGTGAA
CTGAGATCCGGCCACTGCACCTCCAGCCTGGTGACAGAGAGAGACTCCGCTCAACACAGACAAACAAACAAACAAAC
AACAAAAATGTTTACTGCAGCTTTATTGAGATAAAATTCACATGCCATAAAGGTCACTTCTACAGTATACAAATTCAGTG
GATTAGTATGTTACAAAGTTGTACGTTGTTTACCAATCTACTCCAGAACATTTACATCACCCCTAAAGAGAGCTCTTAG
CAGTCACTTCTCATTTCTCCAGCCCTGCOAACCCAGAACTACTNTCTGTCTCTATTCTGAATATTTTATATAAAGGAG
TCTTATCATATGGGCCCTTTAGCTTACCTTCTTACCTTAGCATCATGTTTAAAGATTCAATCCACAGTGTAGCACGTGT
CAGTTAATTCATTTCATCTTATGGCTGGATAATGCTCTATTGTATGCATATCCCTCACTTTGCTTATCCATTTCATCAACTG
ATTGACATTGGGGTTATTCTACTTTTGGACTATTATGAGTAAATGCTGCTATGAACATTCCTGTACCAATCGTTACGTGGA
CATATGCTTTCAATTCCTCGATGATGTAACTTAGGGTTGGAGTTGCTGGGTCAATATGTTAACTCAGTGTTCATTTTGTG
AAGAACTACCAAAATGGTTTCCAAAGTGGATGCAACACTTACATTCACCAGCAAGATATGAAGGTTCCAAATGCTCTCTA
CATTTTGGCAACACTGTGTATTTCTTTTATTATTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
AGGCTGGAGTGCAGTGGCACAATTTCACTCAGTGCATCTCCACCTCTCGGGCTCAAGCGATACTCTGCTCAACCTCC
CGAGTAACGGGATTACAGGCGCCACCACCAAGCTAAATTTTGTATTTTATGATAGAGACGGGGTTTCATCATGT
CGGCCAGGTGTACTCGAATCTGACCTCAAGT

SEQ ID:11

>STAR11

GATTCGGGTGGGTTTGTATGATCTGAGAGTCCCTTGAATAAAAAGAAATTCAGAAAAGCTGTGAACCTTCACTTTCCCT
ATTCTTAACCTTACTTGCTTTGGGGAGGCTGAGGCAGGAGGATGACTTAAGGCCAGGAGTTTGAAGATGTAGTGAGCTATG
ACCACACCGGTTACACTCAAGCCTGGGGAGACCAACAACAAACCTTACCTGCCAATGCTTCCATGCTGGAAAATTTATTT
CGTTTCTGGATTGTGAAAGAACTGGCTTACTGAAACCAACACTTCTCAAAACCCCTTCTTCCAGTTAGGTGTTAAGATT
TTAACAGCCTTTCTTATCTGAATAAAATGCAACAAAGTAACTTAAGAGATGTCAACAACTCACTGTTTGTGTACAAAG
ATGAGTCTCCATGCTTCACTGCCCTGTGGGAACTCTCATCAGGCTCTAGTGGCAAGACTCCTGTGTGCTACCGGAAACCG
TCCCTTCTCTCAGGGCAGACAGTCAATGGAATTTCCATGCACTCCAGCTCAGCAGGAGTCCATGACTTAAGAAGGC
CAATGAGCTGTGGGTGAAGTCTGTGGACGGGGAGCCACATGCGTCACTTCCAGGCTGGGCGTGTGCACTCTCCACTCTC
TTCCCTGTGGGTGCAAGGGCGGGCAGAGGGCCCTGAAACCTTGGAGGTGCGTGGAGGCCAAAAAGAGGAGCGTGGGC
CTCTGGGCTTCTATGTAATTTAGGTAACTGCACTGAACTGTCCAGGTGAACAAGAAATAAAGCTCAAAATGTAATTCAGTCGATTA
GATTTGGTGATGGTTGTACAGCGGTACCTTCCCTCAACATAATAAATTTTCAACAACTCAATAATGGCTCACATGTA
TAAATAATTCATATGAAATCCCGGGAACAATGCTTATTTAGCTCAAGCTTAATCAGAGTAGTCCATCTGAGGGAGGAG
ATAGTAGAGGGCAGCAAGGGGTTGCTCACTGAAGATAAAGTGGCTTGGTAAAGAAATGGTTGAAGAGTGAAGTACAGATAG
GGTAAATCCACATCTCAGACATTTCTGTGATGGTCCGTATTAATCTCAAGATGAAATGATAGAGTTGAACCAATTTAATTAG
ATTCTAGAAATCTAATTAATTAAGATGGGCAATTTCCACAAAGGACTAAACAAAGTACAAGAGGATTAATAATAATCAATCA
CATGGGAGGACCGCCTTGCACTTTAAAAATGATGGAGCTTATCAAGACTGGCTGTGGATATCTGTCTCCTGGGAGGGTTTTT
TCCCCATTTTCTTTTGTAGACATGTTCTGCTATGTTGCCAGGCTGGTCTTGAACCTCTGGGCTCAAGTGATCTT

SEQ ID:12

>STAR12

ATCCTGCTTCTGGGAAGAGAGTGGCTCCCTTGCTGAGGTGACTTTGGCAGGACAGCAGAAACCCAGGTTTCTGTGAGG
AGGAAGTGCTCAGCTTATCTCTGTGAAGGGTCGTGATAAGGCCAGGAGGAGGAGGGGCTTGCCAGGATGTTGCCCTTTCTGT
GCCATATGGGACATCTCAGCTTACGTTGTTAAGAAATATTTGCCAAGAGATGCACACAGAAATTTCTGTAAAGAAATAGGAT
GGAGTTTAAAGGGTTACTACGAAAAAAGAAACTACTGGAGAGAGGGAAGCCAAACACCAACCAAGTTTGAATTCGATTT
TATTGGACGAATGCTCACTTTAAATTTAAATGGAGTCCAACTTCTTTCTCACCCAGAGCTGAGAAGGTGGCAATTCAA
AATGTTTACACTTGTTCATCTGCCCTTTTGCTAAGTCCGTGGTCCCTTACCTCTTCCCTTCTTACATTTGTCGTTTCT
ATCGCACACATATGCTCACTTTTATATTTTACATATATAAATTTTATATATGCTTGTGAATAATGCCAGACGAGGGAATG
AAATAGTCTTGAACAGCTGGAATAATATGCAACAGTGGGAGATTTGGGCACATGTACATTTCTGTACTGCAAGTTGAC
AACAGACCAAGTTGTTATAGTGAGGCTGGGTGGTTTTTATTTTCTCTAGGACACAGCTTGCCTGGTGGAGTAGGCC
TCTTGCAAGAGGCAATTTCTTAGGAGGCTCAACTTCCCAAGAGAGGAGAGGGCGAGACTGGAGTTGTGCTGGCAGCACA
GAGCAAGGGGGCAGCGCAGGACTGCACCTGACAGAGGGGCTGGAGAAAGCGGAGGCTGGCACCCAGTGGCCAGCGAGGCC
AGGTCCAAGTCCAGCGAGGTGAGGCTAGAGTACAGCAAGGCCAAGGTCCAAGGTCACTGAGTCTAAGTCCATGGTTCAG
TGAGGCTGAGACCCAGGTTCAATGAGGCCAAGTCCAGAGTCCAGTAAGGCCAGATCCAGGTTCCAGGAGGTCAGG

SEQ ID:13

>STAR13

CTGCCCTGATCCCTTAATGCTTTTGGCCAGAGCACCCCGTAAGTCCAACCCAGAGGGGCTCATCCGCAAGCCCTCGG
GAAGAGGACAGTACGAGGAGCGGCTGCCCTGTGAGCTGCACGGGGCAGAAATGCTTTTGGGCTCATGTTGGATGTCACACA
CATCCATATGGGGTCAGTCTTATTAGGATTCCTTCCGGGAAGAGGTAGAGGTTAGGAGGGTTAAGCCACGACAGAGGCAAT
CGAGAGGGGTGGCTGGATGGGTCTGCATGCTGTCTCATGACACGGGGAGGCTTGCAAAATTTGCTTCCAGCCCATAGT

FIG. 20, cont

GCCCCACAGAGGAGCCCGGAGTCCCTGGTGGGCGTCTGTGTCTCTGCAAGGAGCCAGTGGAGATGGCCCCGTGAACCTCT
CATCCCCCTTGGCTTGGTGGGGTCTCTGGCAGGTTTATGGAGCCGTACATCTTTGGGAGCCGCTTGAACACGACATCATC
GACCTGGAACAGACAGCCACGACCTCCAGCTGGCCTTGAACCTTACCGCCACATGGCTTACCGCAAGGGCATCATCTTG
TTTATAAGCCGCAACCGGCGAGTCTCGTACCTGATTGAGAACATGGCCCGTGACTGTGGGAGTACGCCACACCTCGCTAC
TTCAGGGGGCGCATGCTGACCAACGCGCGCTCCCTCTTTGGCCCCAGGTCCGCTGCCGGACCTCATCATCTTCCGAC
ACGCTCAACAACATCTTTGAGCCACAGTGGCCGTGAGAGAGCGAGCAAGATGAACATCCCCACAGTGGGCACTGGTGGAC
ACCAACTGCAACCCCTGCCCTCATACCTTACCTGTATCCCGGCAATGACGACTCTCCGCTGGCTGTGCACCTCTACTGCAAG
CTCTTCCAGACGGCCATCACCCGGGCAAGGAGAGCGGCGAGGTTGAGGCTCTCTATCGCCTGCAAGGGCCAGAGGAG
CCCGGGGACCAAGGGCCAGCCACCTCTCTGGGGCTGACATGAGCCATTCCTGTGATGTCTACTCTCTCCCAAAGCAAA
CCACAGCCAAAGCTGTCTGAGCTGGGAGTCCCTTCCCCAGCCCTGGGTACGCGGCATCTCTAGTCTGTGTACTTACTCA
GCTGATGTACAGTGCAGACATCCACCGTTCACCCACAGAACAGTGGCTGAGCGGACCAAGCTTGCCATGTGCGTTTGTCT
CTGTGGGAAACAGAGCACAGAGGGTGAAGCAGATGTGCAGAACGCCCTTGGCTGCAGTTAGGACCTCAGTGGCT

SEQ ID:14

>STAR14

AGCAAGGACCAAGGCTCTGCTTCCAGTCAGCATGAGCAGAGCAGACTCCCTTTGAGCAGAGCATCAGGGCAGAAATAGAA
CAGTTTCTGAA TGAGAAAAGACAGCATGAGACCCAAATA TGATGGGTGAGTGGAGAGAAACAGACACACATGAAAT
TOGGCGAAGTCACTCTCGAAATCCCAACAGAGCGCGCTACAAAGGTGGTGACACCGGAGGGCTTGTGGGCGTCCAGAG
GAGTTTGGCTTCTGACAGCTTCCCGGTAGCAAGAGCAAAACGTCAGCCAGAGAGCTTCAAGTCCAAAGTCAAGACAC
GACCAAGCAGGAGAGGAGGCGAGCAAAAGCCAGCAACCCCAACCGCTTTCAGAGAGCAGTACAGAAATAAAGTGGGATT
AAAAGGAAGCCAGCAGCCGCAAGGAGCGGAAAGCGAGTACGAGCGCGGTACAGGCGCCGAGGCGTCCGACTCCAGCAGC
GACGAGGCAATGAGGAGGCCATCCAGCTGTACAGGTGCAGAAACACAGAGGAGGCGAGGGGAGCCGCCACAGAGG
GTCCAGCTTCAAGAGGAAGAGCAGCTGCCCTCCCGCACAGCAGCAAGCAGCGCCACAAAGAGTGGCTTGGCAGAGACC
CACAGGAACACCCAGCAAGAAAGCCAGTGGCCACCAAGACCAAGACCGACCTTGGTCCAGGG

SEQ ID:15

>STAR15

CAGTACATGCAGAACTGAGTCCAAACGAGACGGACAGCAAAACCGGAGTGGGCTCCAGACATTCCTGGGGGAAAGGGAT
CCTAACCCAGAGCAGTTAAAGTCATCTCCCTCCAACTCTATGACACAGGCTGTGGCTGTCTATTTAAAGCTGAGTGAAG
TTTAAAGCTTTCCATTAGAAACAAAGCGCAGCTGGCTGCCAGCACTCAATTAATTTACATAAAGTGCCTTTGA
GGCTGAAGCAATTCGACTGATTTTCAATGTGAAATAAATAATGTAATAAAGTGTCTTTGAAATTAATTTCTAAACAGAACATC
AGAACTGCTGAAATCATCAGAAATCGGCTATTTTGAATAAATGGAATTCATCAAAAGCAATCTTCGGCCAAACATGTTAGAG
AAGCATGTTAAACACCGCATAGGAATGTTACATTTCTAGAAATTTGACATTTTCAATGACGGAATAATACGTATCTTGT
ATAATGAAATACCACTACTAAACATATGCTATAAATAGAAATGATGCTTTTGTTCAAAGTCAATATACCTGAGCA
TGCAAAATAATAAATAAAGTGAGATCTTCAATGGCAAGCTGCCGAGGATAAATATGACCCACAGTGGCCCCAGTA
TTCTCGGGGCAAACTGGAAAGGGCTAACAGGCAACATTTTCAATGTTATTTCTACTGAGTGCAGTAAATTTTAAATA
TACATGAAATAAATAAAGCTGTGGTATGGTTTAAAGAAATTTCCATAACCTGGTGAATCTTTCACACAGGGTAATAGG
TTCAATAAGCCTTGGTCTCTGCAAAACAGCATCAACTTGACAATGACTAAAGAAAGCAACAGCAAAATCTGACGCAAT
TGGAGCCAATGGCTTGGGTGGGCGGCTGTAAGCTCTCCGCTCTGGAGCAAGTCTGGGCCAGCGGCTGGCATGTGGG
CACTGCAAGGCGCTGGGTGGGCGAGGTGTCAGCTCTCCGCTCATCTGAGGCTAGTCTGAGGCTTGGTGGCTGGCAGTGGG
CCTGCAAGGCGCTTACTTCTCACCCAGCTCCAC TTCCCTCCCTGCGCTCACTGGGCTTCACAGAGCCAAATGAACACTGGG
GTCAAGATTCAAGGCCAGCATCCAC TGCAGTGGGCACTGCCCTTCCACAGGCTTGGCTCCAGGAAGCAACCCCACTTCA
GCCACAGATTAGGGCAACAGGAAATCCCAATCCCACTGCCAGTGACTAACCAAGGAAGGGGCTCACGTGAGGCTGGGCT
CAGGCTTGTGTGAGACCGCTTGTCTATGAGCTTGGATTTAAGGAATTTGGGAGCAAGAGCTTTCTTTTACAGGGCC
ACCAGCAGGGAATAAAGTTAGCCCAACGCGAGTTGACAGTCAACCCCAACAGGACCCAGGGCCACAGAAAGCGGAGAG
GACAACAGAGGA TGAGGTGGGGCCAGCAGAGGGACAGAGAGAGCTGCCCTGGGCAACAGGCAAGAAAGCATCCACAGTG
CAAGAAAGTAGGGCAGCTAGACTTAAATCAGAACTACCGCTCATCAAAAGATAGTGTAAATTTGGGGTGTCTAATA
TTAACTATGTCGCCCAAAAGGCAATGTGTGGAATAATTAATCCCAACAAACAGGGCTGGGAGGTGGAGCTCATGAGAGGT
GGTGAGGCCATGAGGGTGGAGTGAATGGAATGCCATTTGCTCGGGAATGGGCTCTTCTACAGGAAGAGTTACAGCC
CCCTTCTCTTGTCTACCTCTCTTTGCCCTTTCCGTAGGGAGTGACGTAAACAGAAAGGCCCTCACAAAGATGCTGGCAGCT
TGATCTTGGACTC

SEQ ID:16

>STAR16

CGCCACCTCGGCTTTCCAAAGTCTGGGATTACAGGCATGAGTCACTGCGCCATCTCTGATTCCAAAGTCTTTAGATAAATA
ACTTAACTTTTCGACCAATGGCAATCAGGCAATCTTTGAATCTGCCATAGACCTAGGACATCCCTCTCCCTACAAGTTG
CCCGCGCTTCCAGACCAAAOCAAATGATCATCTTACATGTAATTGATTGAAGTTTACATCTCCCTAAACATATAAAACCA
AGCTATAGTCTGACCACTCAGGCAGGTGTTCTCAGGACCTCCCTGGGGCTATGGCATGGGTCTGGTCTCAGATTGGC
TCAGAAATAATCTTCAATAATTTTCAGAAATTTACTCTTTTCAATCACCATTACCTATCACCATAAGTCAGAGTTTTC
CACACCCCTTCTCTCAGATTACGTAATTTGCTAGAAATGGCCACCAATCAGGAAGTATTTTACTTACAAATACCAATTT
ATTAATGAAGAACTCAAAATCAGGAATAGCCAAATGGAAGAGGCAATAGGGAAGGTATGGAGGAAGGGCCAAAGCTTCCAT
GCCCTGTGTGCACACACCTCTCAGCATCTTCA TGTGTTTCAACCACTCAGAACTCTTCAAACTTTGTCTATTTAGGGGT
TTTATGGCAGTTTCACTATG TAGGCAATGTTGATAAATCACTGGTCACTGGTGTATGAACTCTGCTCCAGCTCTCTCTC
TCTCTTCCAGAAAGTCTTGAAGTGGGCTGAAGTTTCAAGGTTAGTTGCTGTGACAAACAGCCCTATCTCTGAAGCT
ATTGAGGGGTCCCAAAAGTTACCTTAGTATGTTGGAAGAGGCTTATTATGAATAACAAAGATGCTCTCTTTTATCC
ACTAGGAGCATATCCAAAGTCTTGGGGAACAAAGCATGTTACTGGTAGCAATTCATACAGGTAGATAGCAATCTCAAT
CTTGCTTCTCAGAAAGAAATTTGACCAAGGGGCAATAAGCAGAGTGAAGGACCAAGATAAGTTTATAGCAGGAGTG

FIG. 20, cont

AAAGTTTATTAAAAAGTTTATAGCCAGGAATGAAAGAAAGTAAAGTACATTTGGAAGAGGGCCAAGTGGGGCAGATGAGAGA
GTCAAACACCATGCCCTGTTTGATGTTTGCGCTGGGGTCTTATATGATGACATGCTCTGAGGGTGTGATCTCTCTCCCT
GATTCTTCCTTGGGGTGGGCTGTCGCCATGCACAAAGGCTGCCAGCAGTAGGGAGGGGGCCGATG

SEQ ID:17

>STAR17

[illegible]

SEO ID:18

>STAR10

ATCTTTTGGCTCTCTAAATGATATGATGGGTGTGTTTTTTTTCCACCCTGCTAATAAATTACATTGCAACATCTCTCC
CTCAACTCTCAAACCTGCTGAAGTGAACAATATGCTATTAAGAAAATACTTTCTGAGAGAAAAAAGGATTTTCTTCCACC
TGATTTTGAATGGCACTTCCGCTGCGATGTGCAAAATCTATTGCTATTCCTCTGAACATTTGTAAATGAAACCTTTG
GACAGTCTCCGCTATTCGCTAGACCTGCCCTGCCAAGGCAAGGTTGTGTTCCAAAGCATGCTGTTAAATCTTAAC
TTATTAATCCCTTAAATTTCAAATTAACAACTGGGGCACTTCCAAAGAGTTCTGCAACAAACATGCTCATCTTTGTGGGAG
GTGTTTTCTGTAATTAATGATGGAAATCATGCTCATTTCAAAATGGAGGTTCCAGATTTGTGGCCAGCTGATGCTGCAAAAT
TATCTT

SEO ID:19

>STAR19

[illegible]

SEO ID:20

>STAR20

GATCTCAAGTTTCAATATCATGTTTGGCAAACATTTCGATGCTCCCACTCTTACCTAAAGCTATCAGAAAGCGCTTTGG
 GAATCTGCTACGAGAGCTACAGAAAGACAGTGAAGAAGCCATGGACCTCCCAACAAAACAGCGCAAGCTTTTGGCAAAA
 AGATGACTGACGAAGACTGTTAAGCAAAAAACCTGTCTCTCGTCAGATGTGCTATCCAGAAATAGAAAAATATTTTC
 CTCTCAACTCTGTAGGCTTCGAGGATTTTGACCTGCCGTAAGGACACGGATGCACATCTCCCTTGAGGTGAAGTGCCCTC
 TCATGATCTGTGATGAGGAGAGAGACTGAAAGAGTTTTCAGCTGGGCCCCCTTCACTCTTGAGATGCCCCTCTCCAC
 CATGGAAA TCCAACTGTTGTGCACTCTTACCTTTAAGCATCTGTTGACCTTGGATTTGAAITGGCCACTCTTTGCTCTGACA
 TAGATATTTAAATTTCTTAGTCTTACGATTTTGTGTAATTTCTAATTAAGCAATATTTGTTTAAACGAAAAAAGAA
 TATATACTTAAATCCCTAAATAAAATACCAATTAAGAGAAAAACAGGAGTATAACTTAATAGGGAACAAAGGACAATA
 ATGGGAATAATACTCTAACTCAAAATGAAGACAGAAAAAGAAAGAAATGAAGAAATGAAGAAGACAGATAAATAGAAAAAATA
 GCAATATGAAGAACAAACTTGACCGGGTGAATGGTGGCTGATGGCTGTAATGCTGTAATGCTG

SEQ ID:21

FIG. 20, cont

>STAR21
 GATCAATAATTTGTAATAGTCAGTGAAACAAAGGGGTATATACATAATGCTACAGAAATCCATTCTCTGGGTATAAAATCC
 TAGACATATTTATGCATATGTACACCAAGATATATCTGCAAGAATGTTACAGCAAAATCTTTTGTAGTAGCAAAAGGCCA
 AAAGGTCATCAACAAGAAAAATTAATACATTGTGGGCACATAATGGCATCCCTTATGCCAATAAAAATGGATGAAATTTATAGT
 TAGGTTCAAAAGGCAAGCCTCCAGATAATTTATATCATATAATTTCCATGTACAACTTCAACAACAGCAAAACTAAACAT
 ATACAAATGTCAGGGAATAATGATGAACAAAGTTAGAAAAATGATTAATATAAAAAATCTGCACAGTGATAACATTTAAATGAG
 AAAAAAGAAAGGAGGCTTAGGGAGGGACCTACAGGGAACGCCAAGTTTCATGGTAAAGTACTAAATACATAATCAAGCA
 CTCAAAATAGAAAAATTTTATAGTAAATGTTTATAGTAAATATCTTACTTAAACAAGGTCAGGCCAGGCACGGTGGCT
 CACACCTGTAATCCAGCACCTTTGGGAGGCTGAGGCGGGT

SEQ ID:22

>STAR22
 COCTTGTGATCCACCCGCTTGGCTCCCAAAGTGC TGGGATTACAGGCGTGAGTCACTACGCCGGGCCACCTCCCTGTGA
 TATTATTTCTAAGTACTATTATGTTAAAAAAGTTTAAAAATATGATTTAATGAATCCAGAAAC TAGGATTTTACAT
 TGTACAGTTTCTTATTATATAAAAAATAAAATCAACAAATAATATATGGTAAAGTAAAAAGAAAAACAAAAACAAAAAGTG
 AAAAAAATAAACACACTCCTGTCAAAAAACAACAGTTGTGATAAAACTTAAAGTCCGAAAAATTTAGAAACATCCTTTCTA
 AAGAGTTCTGAATAAAATAAGGAAATAAAATAATCAACATAGTTTGGTCAATGGTTCTGTTATATGTGATGGAATATGTTTA
 TTGATTTGTGATATGTTGAACCTTATCTCAATAGATGCAGACAAGGCTTGATAAAAGTTTAAACACCTTTTCATGTTGAAA
 ACTCTCAATAGAC TAGGTATTTGATGAACATATCTCAAAATAATAGAAAGCTATTTATGATAAACCCATAGCCAAATATCATA
 CTGAGTGGGCAAAAGCTGGAAGCATTTCCCTTTGAAAACTGGCACAAGACAAGGATGCCCTCTCTCACCACTCTCTATTAAT
 GTAGTATTTGGAAGTTCTGGCCAGAGCAATCAGGCAGGAGAAAGAAAGGTTATTAATAAGGAAGAGAGGAAAGTCAAAATGTT
 CTCGTGTTGCAGTAACATGATTTGTATATTTAGAAAAACCCATTGTCTCACTTAAACCTCCTTAAGCTGATAAACCACT
 TCAGCAAAAGTCTCAGGATACAAAAATCAATGTGCAAAAAATCAACAAGCATTCCTATACACCGATAATAGACAGCAGAGAGCCA
 AATCATAGTGAAGTCCATTCAAAATTTGCTTCAAAAGAAAAATAAAATACTTAGGAAATACAACTTTCACGGGACATGAAGGA
 CATTTTCAAGGACAATAAAAAACCTCTCTCAAGGAAATGAGAGAGGACACAAAGAAATGGAATAACATTCATGCTCATG
 GAAGAAATCAATATCATGAAAAATGGCCATCTGCCCAAAGTAAATTTATAGATTCAATGCTAACCCCATCAAGCCACCATTTGA
 CTTTCTTACAGAACTAGAAAAAACTATTTTAAACTCATATGTAGTCAAAAGAGTCGGTATAGCCAAAGCAATCCATA
 GCATAAAGAACCAAGCTGGAATGCAATCAGCTGACTTCAAAACATAC TACAAGGCTACAGTAAACAAAACAGCATGGTACTG
 GTACCAAAACAGATAGATAGACCGATAGAACAGAACAGAGGCTTCGAAAA TAAACACCACACATCTACAACCCCTTTGATCTT
 CAA

SEQ ID:23

>STAR23
 ATCCCTTCATCCTTCAGGGCAGCTGAGCAGGGCTCGAGCAGCTGGGGAGCCTCACCTTAATGCTCCTGGGAGGGCAGCCA
 GGGAGCATGGGGTCTGCAGGCAATGCTCAGGGTCTGCAGGCGGCACGCAOCCATGTCAGCGCCGCCCACTGTTGCTCTG
 OCTCCGCCACCTTGCCCATGGGCTTCAGCAGCCAGCCACAAAGTCTGCAGCTGCTGTACATGGACAAGAGCCCAAGCAG
 CTAGAGGACCTTGTGTTCCACGTCGCCAGGAGCATGGGCCACAGCCAAAGACAGTCAGGAGCAGGCGGGCTTCTGG
 CAGGCCACGCTTACCTCTGTCCTCACACAGATGGGAGATTTCTGTTGTGATTTTGTAGTGAATGTCGCCCTTTGGTGACATC
 CAAGATAGTGTCTGAAGCACCGCTCTAACAATGTGTGTATTTCTGAAAACGAGAACTCTTTATCTGAAATAAATGATG
 CAAATAAATTTAGTTGGATTGGAATTTCTATTCAATGTAGGCAATGCACAAAAAGTCCAAACATTGCATATGACACAAAGAA
 AAGAAAAAGCTTGCAATTCCTTAAATACAAATATCTGTTAACTATATTGCAAAATATATTGAAATACACTTCTATTATGTA
 CATATAATATATATATGATATATATATATATATATATATATGTTACATATAATACTCTTATTATGTTACATATAAT
 ATTTATCTATAAGTAAATACATAAAATATAAAGATTTGAGTAGCTGTAGAATCTGCTTATGTTTATAGCTACTACTAC
 AAAAAATATCTCTCCACTTATGCCAGTTTGCCATATAAAATATGATCTTCTCATTTGATGGCCAGGGCAAGAGTGCAGTGGG
 TACTTATCTCTGTGAGGAGGGAGGAGAAAAAGGGAACAAGGAGAAAGTCAAAAAGGAAAACTCTGGTGTGTCAAAAATGT
 CAAGTTTACATATTTCCGAGAGGGAATGACATGTGCCACAGAAAGGACCTGCCAGCTAAATGTGTCACAGATATCTCAG
 GAAGCTTAAATGATTTTTTAAAAAGAAAGAGATGGCAATGTCACCTTTCTTTGAGCTGAGGCTGTGGGATGATGACAGA
 TTTCTGGAAGGCAAGAGCTCTGCTTTTTCACACCGAGGGACTTTCAGGAATGAGGCCAGGGTGTGAGCACTACACCA
 GGAATCCCTGGAGAGTGTTTTCTTACTTAA

SEQ ID:24

>STAR24
 ACGAGGTCACGAGTTTCAGACACGCTTGGCCAAAGATGGTGAAGCCTGTCTCTACTAAAAATACAACAAGTAGCCGGGGCG
 GGTGACGGGCGGCTGTAAATCCAGCTACTCAGGAGGCTGAAGCAGGAGAACTCTAGAAACCCAGGAGGCGGAGGTGCAGTG
 AGCTGAGACTGCCCGCTGCACCTTAGCTTGGGCAACACAGCAAGACTCTGTCTCAAAATAAAATAAAATAAAATAA
 ATAAATAAAATAAAATAAGAAAGGAGAGTTGGAAGTAGATGAAGAGAGAAAAAGAAATCTTAGATTTCCTATCTGAAG
 GCACCATGAAGATGAAGGCCACCTCTTCTGGGCCAGGTCTCTCCGTTGCAGGTGAACCGAGTCTGGCCCTCATTTGGAGAC
 CAAAGAGATGACTTTGGCTTGGCTCTTAGTGAGGAAGCCA TGCCTAGTCTGTTCTGTTGGGCTGTATCTGTATCACT
 TGAATGCTCTCTCTGGACTTTCCATGGAATTCAGGGATGCAACTGAGAAGTTTATTTTAAATGCACCTTACTTGAAGTAAGA
 GTTATTTTAAAAATTTTAGCAAGGAAATGAATTCAGCAGGTTTGCACGAAACATTCACATGTGAGGAAAAACAGGA
 AAACCACTATGCTAGAAAAAGCAAAATGCTGTGAGATTTGCTCACAAACAAAAATGGGTGCCAGCAGGTAGGTTTGAAGCC
 TCAGGTTGGGCACATTTTACCTTAAGCGCACGTGTTGGTGAACCTAAGGTGACTGTAGGACTTATATATACATACATACAT
 ATAAATATATATACATATTTATGTGATATACACACACACACACACACACAGGGTCTTGCTATCTTTGCCAGGGTG
 GTCCTCAACTCTGGGTCTCAAGGATCTCTGCCCTCCCTTCCCAAAG

SEQ ID:25

>STAR25

FIG. 20, cont

ATAAAAAATAAAAAACCTGCTCTAATTGCAAGGCTCTATCTTCTCTCCCAACACCTGAAATTTTAGTGAAAACGGG
 GCTTCCGTAGGAAGGAGTAGCTAGCTATCCGGTCCGGTACAGGTATTCAGTGCGTGAATACCTTGACTCCTAAGGCTCA
 GGATTTGACTGGGTGCTCCGACTGCCCGCCCCCAACCGGGACCCAGTCCACCGCGCCAGCCTGCGGCGTCCCT
 GACTCGCGGGATTGAGCTTCGGTGCCCAACAAACACTCCACCGCGGCTGCGTCCACTTTACCTGCGCGCGGACCAAGC
 TTCGAAAGAAAAGTGTCACCATGGTGTGAGGAGCTTCACCTCGAAAATGGTAGTCCCGGGTGGCAGAGATTCCGAAGAC
 GACCCCTCATGCCTTTTTCTCACAGCCGCTGCCCTAGATTGGCGCTACTTGGCTTCGGCCATGTTGAAGTTGAACCTCCAA
 ATCTAACTGGCCCGGCTCCCGGCTGCCGAGCTCCGATTGGCGCTCCCGCGAAGGGTGCTCCGATTGGAAGCAGTA
 GAACGCTCTCACCGAGCAGGGCGGGGGCGGGGAAGTCACTCGAGGCTGAGGCGAGCGGGGAGGCGAGGCTCTCGCGGCTG
 GGATGTCGCGCAGCGAAAAATACGCGCAAGCCAAAGCTCGGGGCTCAAATAAAACTTTTAATTACATTTTCAGAGACTTC
 GTACAGTGCAACAGTGAAATTCAGTGTAAATTTTCACAAGAGTCCATTTCATCAACAGTTTCAGAGAGTCTGCCTTTCAT
 TCCCTGTTTCCCTCAGTGCTCCAAATCAGGTTTCAGTCTCCAGAGGTTCTTTTAGTTTGAATTACCGACCAAACTCCAG
 TTTAGGAGAAATGGAAGTCCACCGTCCCATCCCAACCAAAACATATTTAGTCAAAACCAATCCAGTCCCTAAAGAAATTA
 GGAAAGTATGGGCAAGGGTCTTTTAATATACACACATCACCTTAAACTGCGTGTGTGTACGAGAAATAAGAAAAA
 CACAAGAGGGGCTG

SEQ ID: 26

>STAR26

CCCCCTGACAAGCCCGAGTGTGTGATGTTCCCACTCTGTGTCCATGCAATTCATTTGTTCAACTCCCATCTGTGAGTGAG
 AACATGTCAGTGTGTGTTTCTGTCTTGAAGTAGTTTGTGAGAAATGATGGTTTCCAGCTTCATCCATGTCCTTGCAAAG
 GAGATGAACCTATCTCTTTTATGGCTTCATAGTATTCCATGGCACATATGCGCACATTTTAAATCCAGTCTATCAT
 GATGGACATTTGGGTGGTTCCAGTCTTTGCTATTGTGAATAGCACCAATTAACATATGTGTGATGTATACATCTTT
 ATAGTAGCATGATTATAAATCTTCGGGTATATACCTGTAAATGGGATCGCTGGGTCAAATGGTATTTCTAGTTCTAGATC
 CTTGAGGAATCACCACTGCTTTCCCAATGGTTGAACATAATTTAGCTCCCAACGACAGTGTAAAGCATTTCTATTTT
 TCCAGTCTCTCCAGTATCTGTGTTTCTGACTTTTAAATGATCATCATTTCT

SEQ ID: 27

>STAR27

CTTGGCCCTCACAAAGCCTGTGGCCAGGGAACAAATTAGCGAGCTGCTTATTTTGTCTTTGTATCCCAATGCTGGGCAATAA
 GCTTGCATTTATGAGTAAATCCGGTAGAAGTATGTGTTCAAGGACCAAGTTGATAAATACCAAGAAATCCAGAGAGGGA
 GAGAACAATTGAGTAGAGGATAGTGACAGAAGAGATGGGAACCTCTGACAAGAGTTGTGAAGATGTACTAGGACGGGGGAAC
 AGCTTAAGGAGAGTCAACAGGACCGAGCTCTGTCAAGCCGGCTGCCATGGAGGCTGGGTGGGGCCATGCTAGCTTTCC
 TTCTTCTCAGAGTTCAGAGTGTACGCTTTGAACCTCTAAATCCAGAGGCAATTAATCAATGTTTCTTCTAGGGGCAATAC
 CTGCGCTGCTGTGGAAGACTTTCTTCCCTGTGGGTGCGCCAGTCCCGAGATGAGACGGTTTGGGTGAGGGCCAGGTGAC
 CGTTGGGTGTGTGCTTATGTCTGATGACAGTCTAGTCTACTGCTATTAGTCAATGAGGGAGGTGTGTTAAAGATGAGATG
 CTGGGTACATCCCTAGAGAGGTGTTCAGATATGGGCACATGGGAGGGCTGGAAGGATAGGTTACTGCTAGAGCTAGAGAA
 GCCACATCTCTTAAACACCTTGGCTTTTCCCACTGCCAAGATCCAGAAAGTCTTGTGTTTCTGCTGCTTTCTCTCTTTT
 TTTCTTTTCTTCTGAGATGGAGTCTGGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCACGATTTCCGGCTCACTGCAAGT
 TCCCGCTCTCAGGTTCAACCATCTCTCCACCTCAGGCTCCCGAGTAGCTGGGACTACAGGGCCCAACACACCCAGCTAAAT
 TTTTGTATTTTATGATAGAGAGGGGCTTCAACATGTTAGCCAGGATGGTCTTGTATCCGCTGCTCAGCTCCCAAGGTG
 CTGGGATTAACAGGCTGAGCCACCGCGCCCGGCTGCTTTCTTCTTCAATGAAGCATTCAGCTGGTGAAGAAAGCTCAGCCA
 GCGTGTCTGGAACCTCTTCACTCAAGTGAATGCTGCTGCTCAGCTCCCAAGTGTGAGATTAACAGGATGAGCCAGTCT
 CGAATGTGGCTTTTCTTTTCTTTTGTGTAACAAAGGCTCTACTGTTGCCAGGCTGAGTGCAGTGGCATACCTCAGCTCC
 ACTGCAGCTCGAATCTTGGGCTCAAGCAATCCCTCCCAACTGAGCTCCCGAGTAGCTGGGGCTACAAAGGCAATGCCACC
 ACGCTGGCTTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
 GTCTCAGCTCACTGCAAGCTCCCGCTCTCTGGGTCAAGCGATTTCTCTGCTCAGCTCCCGAGTAGCTGGGATTAAGGC
 ACCTGCCACCATGCTGGCTAAATTTTCTTATTTTATAGTAGGATGGGTTTCACCATGTT

SEQ ID: 28

>STAR28

AGGAGGTATTCCTGAGCAAAAGGCCAGCTAGTGAACGGATAAAATGCCATGTAAAGATCTGTTTACCCTGAGAAAGGCA
 TTTCTCAACTCTCCCTATAAAATGCCAAGTGGAGCACCCAGATGAAATAGCTGATATGCTTTCTATACAAGCCATCTAGG
 ACTGGCTTATCATGACCAAGGATATTCACCCACTGAATATGGCTATTACCCCAAGTATGTTAAATGCTGTAGTTAAGGGGG
 TCCCTTCCACATGCAACCCAGGTTATAACCAAGAAAGGTTCCCAATCTAGACTCCCAAGAGAGGGTTCTTAGACCTCATG
 CAAGAAAGAACTTGGGCAAGTACATAAAGTGAAGCAAGTTTATTAAGAAAGTAAAGAAACAAAAAATGGCTACTCCAT
 AAGCAAGATTATTTCTCACTTATATGATTAAATAGAGATGGATTATTCATGAGTTTCTGGGAAAGGGGTGGGCAATTTCT
 GGAACGTAGGGGTCTCTCCACTTTAGACCATATAGGGTATCTTCTGATATTTGCCATGGCAATTTGTAAGTGTCTATGGCA
 CTGATGGGAGTCTTTAGCACTCTAATGCAATATAATAGCAATATAATGAGCAGTGAAGTGAACAGAGGTCACCTTCTG
 TTGCAATATTTGTTTCTAGTGGGTTTGGTTGGCTTTTCTTTTCTTTTAAACCAACCTGTTTCTTTATTTATTTATTTATTT
 ATTT
 GGCTCACTGCAAGCTCTGCTCTCTGGTTACGCCCATTTCTGCTGCTCAGGCTCCCGAGTAGCTGGGACTACAGGTGCGCTG
 CCACCATACCCGGCTAATTTTCTATTTTCTAGTAGAGACGGGTTTCAACGTTGTAGCCAGGATGGTCT

SEQ ID: 29

>STAR29

AGCTTGGACACTTGTGTATGCCACTTTGGAATGTTGAAGGGCCGCTCTCCACACCGCTGGCCACTTTTAAATATGTCCC
 CTCCTGCCAGAGGGCCAGAGGAGGGGCTGTGAGGGTGACAGGAGTTGACTGCTCTCAGCAGGGGGTTCCCGAGGG
 ACCTTTCTCCCATTTGGGACGATAGAAGGACCTAGAAGGGCCCTCCAAGCCAGCTGGGCTGACAGGGCCAGCGATT

FIG. 20, contd

CGATGCCCTTCCCTGACTCAGGTGGCGGTGTCCTAAAGGTGTGTGTTTCTGTTCGCGAGGGGGTGGCGGATACAGTGG
AGCATCTGTCGCCGAAGTCTCTGAGCGTGGGTGAAGTGTCTCGAGGGGTCAGCGCTCTCTCCGACTCTCTCCATCAGTGC
GCTCA CAGCCTGTAGGACAGCTCCGGGAAGCTCTGTGATGAGGCATCTGGTCAATCCCTCGAGATCTCTGCCCTGGCCT
GAAGAGGAGGGGAGGAGGAGGAGGAGCCCTCTCCATGCCCGCAGGCGTGGAGGGCTCGAAGCCGGCCGGCCAGCATTTGATGCC
AGGCTTGGCTGTGCAAGAAGCAGATTGCCCTGGCCCTGGCCAGGCTTCCACAGTAGGATGTGGTATGGCAGCCAGTTGGGGT
ATTGAGGGGCTGCTGTATGCCGCCACCAACCTCCCCAGGTAGGGTGGTGAACAGTAGGCTGGACAAGTGGACCTGTTCCCATC
TGAGATTCAAGAGGCCCACTCTCGGAGGTTGCGATGTCAGCGGAGATCCCTCCACTGCATCTCAGGCTGGGCAACAGATGA
ACTCTGCTGTAAAAAACAAGACAGCAACGCAAAAAAACCTCTGGCCCACTGGCTAAATTTGTAATAAAGTTTAA
TGGCACATAGACACACCACTTCAATTACATATGCTGGCGGCTGTTTGTCAATACCTTTAGTATGACACAGACAGCAGTGG
CCATGGGAAGCAAAAAATTATCTGTCTGGCCCTTTACAGAAGTCTGCTAGAGGGAGAGCCCGGCCCATGGGGCAGGAC
CATGGGGTGCGGCAGAGAGGAGGCTCGTGCTCTCAGCGGGCTAGTTGGGTATCTAGTGCCTGTCTTCTGCACTGGAGC
ACAGGGGTGCGGCAGGTAAGTACTGGAGGAGGACGCTCTCGCTCGCCGACAGCATGGGACCCAGGAGACCTTGAGAGGCT
TAACGAATGGGAGACAGCAGGACAGCGGCTCCATGTGCTGGCCCTCAGTTTCTGCTGTAAAGTAGGAGGGGACGT
GTGAAGGTGAACCTGTGAGGCAGAGCTCTGCTCAGCAATTCGAGGAGGGGCTCTCGCCCACTCTGTGTGCACCCAGAGT
GAGGGGCGAGCGGGTGAGATCTCACCATCAGCCCATAGGGGTGTCTCTGCTGGTCCAGTCCCAAGGATGTCCCATCCCT
CTGGCTGTGTGGGGACAGCAGAGTCCCTGGGGCTGGGAGGCTCCACATGTTTTGTAGTGGTTTTCTGAACTGTAA
ATTGATGTGAAATATCTCTTTCCCTTTTACTTAGGAAGCACTCCAAAGAGGACAGTGACTGTGCTCGAGAAGTTCCAGCT
GGTCTGGAAGCTGCGCCAGAAAGCCAGGTAAGTCTCAGCAGGCGCCGGCAGGGGTGTCTGCGCAGCACTATGGGCACAG
CCAGGGGTGTGCTATCGCGCGGCTGAGGCTAGGCTCTCCGCGGCGCTCGCTGTGCTGGGAGGGGCTGCTGCTGACATGCA
AGGCGCGGTTGTGTCCGCGGCTCAGCTGTGTAGTCACTGCTGCTGTGAGTGTCTGACAGCACTGTGGTGTGGGA
AGGGGCTCTGTGATTCAGCTCTCTGGGACGGTGGCGAGGGTGGAGCTGCCCTCCTCTCCACACCGCCCTGGCCAGGGTGT
AAGCGGCTGTGGGAAGACTCAGGCGCGGGCTGCTGCTGTGAGGCTGCGGACCTCTGCCCTAGACAGAGCTGGGGCT
TCCCGGCTTAGGAGCAGCGCGGAGACACAGGGCTCGAGTAGTCCAGGCTGCCAGCTGGAGGCGCTCTCTGGCG
TCCGGCTGTGACTGACAGCCAGGAGCGGGGCTGTGTAAATTGCTGTTTCTCTTCAACACAGAACCTTTTCGGGAAGATG

SEO ID: 30

>STAR30

ATTTCTGTTATTATTAAGCCACCACCAACTGCTGGTGCTTTGTATAGCAGCTCCAGGAAAGCTAACACATGCACGTGCCCACTT
 TTCCCTTTTCAGCATTTCTTCAGGACTTGCTGGCTCTCTTTGTGTTGGGGTACACCCATAGCATGGCCCACTTCCCACATTC
 CZAAGAACCAACACCTGCATCTAGTCTTTGTGGGCTGCTAGAGCAAAAATATCATAGACTGGGTGACTTTATAAACACAGAAA
 TCTCATTTCTACATCTTGGAGGCTGGGAAAGTCCAAATACGAGGCAACCATCATATTTGGTCTCTGGCTAGGAGCCCTCTCTA
 GCTCTCTACATCTGTGCTCTACATGTCGAGGAGGCGAAGGCAGCTCTCTGGGGTCCCTTTTACAGGCCCAAAAATCCCAATCTCA
 TAGGGCTGTATGACTTCATGACTTAACTCTTAACTCTTAACTGGCCCACTCTTAAATGCTATGGGGCTTAGGATTTACCAATAAA
 TTTTGGGGGGACACACATATTTCAGACATATGCAAAACCCCAACATAAAAAAACCCTCACTTAAGGTTTCCAAATGGCATGGC
 ATTTAAATCATGTTTCATATTATCAATAAAAGAGCATTAAGTCAACAAATGTATAAACGGGTGAGATTTGTCCGCTAGTAT
 GTTCCACAAAATCTGTGGCTTTAAAGAGTGAGATCGGGAATGGGAACTATTACATCTTCTTCACTCTTTTGGTACTGSCCT
 CACTGTGTTAAAAAATAAAAGAGGACGAGCAGCATTTTACATACAGGAAAAAACAATGAGGTTATCTTCATCTCTGAAA
 AAAAAATGGAAAATGAAACAGTGGGAGTCACATCATGAAAATGCTTAATGGTACAAATTCATGTGACAATAAAACAATAAGATA
 GAGGAGCGCTGTTTATAGCATAAAGCACTGTAAAAATGACAGCGCTCGAAGAGAGATGAAACACCACTATTTTGTTAAGGTAG
 TCGAGTGGCAGGTGATTTCTCTCTTTTGAATAATTTCCATTTTCATATATGCGAGTTTGTGCATTTACTAAAACTTTGG
 TTGGTACACATGCTATAATAGATAGATATAATAGTATAGATAGATAGATATAATAGACAGGTAGGTAGATAGATAGATAGAT
 ATGAGAAATAAGTCGCTGTACTTGGCCTTTCAGGCCATAAGTACATATCCCTCTCTCTGCTCATTTGCTATGCTATGCTGCTGAG
 ACAAGGCAGTCTGTGCCCTCTGGCCGTAATTCAAATGTGCCCTGTGCTCTGGCTGTATGTCCTTTCCACCCCAATACAA
 TTTCTCCGAGGTCTACTTCTAAGTGTGAAGGCCACCCAGATCAGATGCTCTCTGTGCTGCTTACTTACCAATTTCTTAATT
 ATAACTATAAACACAAATGAGGCTCTGATAAATACCATAGACCTTCAGGCCCCCTGTGATAACTTCACTCAATTTAAACCTAAC
 AAGGAAATAACCTACCATGAATCGCGAGGCACGAGCAGCTAAGGACTCAGCAAGGTACGCGAGCTATTGGTGAGTAACCAT
 GAGTCAAGCTCTCAGACGCTGTGGCTGTAGAAATAGGTTTCCCAACTCAGCAGCTGGGACAATTTACGCTGGATATCTCT
 TCTGTTTGGGGGGCTCTGCTGTGCCCTGTAGGATAATAGGAGCAATCTGTGCCCTCTACCCCATCAGACGACGACGACCTCC
 CATGCCAGTGTGTGACACCAAGCAATGTCTGCCCACTTGGCAAGTGTCCCTGGGTGGGAAATGACACC

SEQ ID: 31

SEQ ID:
>STAR31

CACCGCGCTTGGCCCCCAGAGTGC TGGGATTAACAAGTGTAACCACCATTCCTGGCTAGATTTAAATTTTAAAAAAATAA
 AAGAGATAGTAGAAATAGTTTATTTTAGGAGAGAGCCCTTAACTGGCAGCAGGCGAGGAGCAGGGGTGAGGCTTCCCTTANTTC
 AGGCTCACTCTAAACCCACCCAGGACTGTGTGTCAATCTCCAAATAAGGAGAAAGGTTGCTGCCGCCCGCTGTGAGTGTGCT
 CAGTGGAGGGTAGAGGGCCGCTGGGCAGAGTGCTTCTATGACGTGCTCATCAAGAAAGGCTTCAAGCAATTCGCCCGGACGTCG
 TGCTCATCCCCATCTTACTTCCAGCTAGGAGAAGGCCGCTGGCCACAGTCAACCCAGCCGCGCAAGTGTACCCCTGGGTTG
 GACCCAGAGCTATGATCTGCCAGGGGTTCCAGCTGAGAATCAGGCCACGTTCTAGGCAGAGGGGCTACCTACTCTGGAC
 TCCAGTAGCTGTGTAGTGA TCGAGGACTCATGGCTGCAGAGCC TGGACATGGGTTCACTACCTGGCTGCTCTTGGGACGCC
 CATCTTCAATGCGAGGTGAGGCCCAAGCATATGATCCAGACATGCAAA TGGGTGTGAATCTCTCTTCTCCCAAGAGCCA
 CTCTCTCACTGTCTTACTTAGGAAGGCAAGGGGCA TGGTGGAACTCTGAAGCGCTGTGTAGGGGTCTCAGCGGAACTGTGAC
 ATGTCAGAGCTTCCCTTCTCTCTCTGAACTAGATTTGAGCAGAGAGCAAGGACATGAAACAGCAACCAGAAATTTTG
 GGAACGGGCTCTCA TCCAGCTCAGGCTACCTCTCTTTTAAATTTTAAATTTAAATTTAAATTTT TTTTATAGAGACA
 GAGTCTTACTGTGTGGCCAGGCTGTAGTGCAGTGGCAACAATAGTCTCACTGAGCCCTCAACATGCCACCTTCCAGCTCT
 TGCTATAGCTAGACTAGAGCTGCACCAACCAACCACTCACTCAATATTTTATTTTGTAGAGAGAGGTTTCCACATCT

FIG. 20, cont

[illegible]

SEO ID:32

2-STAR32

GATCACACAGCTTGTATGTGGGAGCTAGGATTGGAAACCCAGAGCTCTGGCCCAAGGTTTACATGCTCTACCCACCTGCATAC
AATGGCTCTCTCAATTAATCACTCAGTATAAACAATTAGATCTCTGTTTAAACCACTAGAAATTAGTAGGCTAAGTAAATGAAG
TCGAGGACCTACAGCTGAATGGCATCTCTCGAAAAAGCCCGCAAGGAATTTAAATCAGCTTTCTGCTTAATCTTAGAGCA
CTGAGTTAGCAATAATAGAAATACAGTGTGTCGCAGAAATAGCTTTATGCTGCACCATCTAAATCTGGCGCTGTGTTTCCAA
AAATCTAAACAAAATCTACGTACTCTTTAATATAAAAGTACATAGCTACTCTGGGAGGCTGAGGTCTGGGAGGAT
GGCTTGAGTTTGAAGCTGCAAGTATGCTACTATCAAGCTATAAAATAGCCGCTGCATCTCCAGCTTGGGCAACATACCCAGGC
CTATCTCTCAAAAAATAAAAGAAATAACATCTACATTAGAAAGAAATTAATTTTATGGGTTTTTTGCAATTTTATATATAC
ACAGCAACACAGCAGCATATGAAAAAATGGGTATGAATCTAGGCATCTAACTGGAAGACAGTACTAAATCAATGCTCAATG
TAGTACAGGCTGACTAGGTTGGGTTGTTTCTGCTTTTCTCTCTCTCTCTCTTTTCTTTTCTTTTGTAGACAGGAGCTT
TGCCTCTTTTGGCCGAGCTTGATTTGCTCAATGGCTGTAATCTCAG

SEQ ID: 33

>STAR33

[illegible]

SEQ ID: 34

>STAR34

ACTGGGCACCCCTCCTAGGCAGGGAAATGTGAGAACTGCCGCTGCTCTGGGGCTGGGGCCATGTCAACAGCAGAGGGGAGGA
CGGTGTTATACACAGCTGGGAAGACACTAGGGTGTCTAGCCACAAGAGTCGTGGTGAACAGGAGGAGCTGTGCTTCAACT
TGACGCCCTTAACACACCGAGCTGGTGTCTAGCCACTCACTCGGGGTGAGCCAGCTTAGAGTAGTGCCAGCTGGGAGCCCTC
CGCCGCCCTCCCTCTCTGAAAGAGCCAGGGGTCAAAACAGTGTAGACTAGAGGGGCTGAGGGCCACATGTTTATAGACCA
AAGGTGGGGCTGCATCAGCGGGGTGGCTTGGGAGCAGCTGCATGGTGGCTACTTGGGAGGGGTCTCCAGCCCTCTCAA
TGGTGTTCGGGCTGGTTCGCCCACTGGCGGCCACCTCGAACAGAGCTGGATATGAGGGCTCATGGCTGGCCAAATGGGAGCCA
CCGAGATGGGGACAGCAGAAATAAGACAGCAGCAGCTCTGGGGGGCAGGGGATGAGCAAAAGCAGGCCCAAGCCCCA
GCCACCTGCACACTTGGCTTCCCAACAGCCCTCGAGCCGCCAGCCACACTCATGTGCACCTAGCCGCTGATAGACTACT
GTCCTTACGGAGAAAGTCCGTGAGAACAGGCAGCTGTGTGTGTGCTGCTATAGCTGTGTGTGTGTGATCCCTGCAC
CGAGGCTGACTTGCACCTGCCCTGGG

SEQ ID:35

>STAR35

CGACTTGTGTGATGCGGGGCTCTTTTGTGTTCCATATGAACCTTTAAAGTAGTCTTTTCCAAATCTGTGGAAGAAAGCTATGG
TAGGTTGTAGTGGGATGGCATTGAATCTGTAAATACCTCTGGGCAGTATGGCAATTTCCAAATGTGATCTTCTCATCCCA
TAGATGAAGAAATGTTCTCCATAGTTGTATATCTCTTTATTTCTGTAGCAGTGGTTGTAGTCTCTTGAAGAGGTTG
CTTCACATCCCTGTGAATGAGTACCTTAGGATATTTATCTCTTGAAGCAAAATGTGAAATGGGAGTACCTACGATTT
GGCTCTCTGTGTGTCTGCTGGGTATATAAANAATGTGTGATNTGTGATCAATTTGTATGATATCTGTAGACTCTGTAA
TTTGTCTTATCTGCTTNNGGCAACCTTTTGGGCTGAAACATATGGGATTTTCTAAATATACATCATGCTGCTCCAAAG
GGAACAATTTGACTCTCTCTTTCTCCATTTGAAATACACTTTTATCTCTCTGCCTAATGTGCCCTGGGCAACCTTCA
ACACTATGNTNINQAAATAGAGNTGTGTAGAGAGGGCATCTCGTGTCTGTCTGCCAGHTTTTCAAAGGGAAATGCTTCAGTT
TTGGCCCATCTCAGATATGATATGGGCTGTGGGTGTGTGAAATGAATAGCTCTATATATTTGAAATGTGCTCCATCAATACT
AATTTATTTGAAGAGTTTGTAGCATGAAGCAATGTGTAAATTTGGCTCAAAAGGCTTTTCTGCATATGTGAAATATCAATGAT

FIG. 20, cont

GTTTTTCGCTTGGCTGNTGTTTATATGCTGGATNACATTTATTGATTGCTGATATNGAACCCAGCCCTNNCATCCACAGG
 ATGACCAACCACTAGTTCACAGCTTGGCGCGGNCNTAGCTCTGAAGCAGGCCAAAGATATGCAAAAGCATGCATCTCAATTAGT
 CAGAGCCCAATGTCGCCCGCTACCTCGCGGCATCGCCGCTCAGCTCNGCNGCTGCGCCCATTTGCGGCAATGCGCTAGTA
 TTTTTTTANNATCCAAAGCGGNCGCCGCTCGCTGCANATCTTCAAGTNNAGAGNTTGGAGGCGNAGCGCTGCAAAACCTCCG
 CAGGNCNTCTNNCGATGACCACTGTTTAAATATTCGCTTTCAGTNTTCTANNNGNG

SEQ ID:36
>STAR3C
GTACCACTCTTAGGCTCATCGATACCTCTGTGCACCTGTGTCAGTGCCCTGGAGTGTGTGTCTAGCCAGGCCCCATCCCC
TGGACATCAGGGGATCAGGACTAGTGTGGCATGTACACTTGGGCTCAGGGGATCAGGAATGTAGAGGCCACATGTACACT
TTGGGCTCAGTGGACTAGCAGACTAGTAGGAGCCCACTGTACACTTGGGCTCAGGGGATCAGGAATGTAGTACAGGCCACAT
GTACACTTGGGCTCAGGGGATCAGGAATGTAGTGGGCCCATCTACACTTGGGCTCAGGGGATCAGGACTAGTGTAGGCC
CAGACTGTACACTTGGGCTCAGGGGATCAGGAATGTAGTGGGCCCATCTACACTTGGGCTCAGGGGATCAGGAATGTAGTGT
GCCGCCCATCTGTACACTTGGGACAGGTGAACCATCTGCACTGTGCTGCAGAGGCTCAGGCCCTGTGCAGATGAAATGTGGTCACT
CGATCTTCTACAGCTGGCCAGCCCTGTGAACCATCTGCACTTCCCATCTGTGCTGTGTCTACAGCACTGCCATGTGCA
CGCGCGAGATCTCCAGCGACATCTCTGCCAGCACCCCGACACACACTCTCCAGCGAGCTCTTGTGGCCCTCCAGGCCAG
CTCCCGCGATTTCTTCTTGGGAGAGGCTCTGGACAGCAACCCCTCTCTCCAGCGATGCGCTTTTGTCTGCTCTGCCCACT
CGCGGAGCCCTGCGCTTGCATCTACATAGACACATACAGTACACTCTCTCTCAGCAAGCGCCGACAGCGGTGGACT
CTCTGGGCGGCTGCTCTGTAGGCTGTACACAGGTTGTGTTACTTGCACATCTAGTGAATATGGGCAGATGTACACT
TCTCTCTCGGTTCTCTGTGTAGAGGATCTAGTGCGATCTCTGTCTGTGGCGGCCACTTTTGTCTCGTGAATGCTATTATTT
TTTTCACAAATATTTCCCAAGTCTCTTCTGTGTGCAAGGATATAGGGCTGCAGCGCGGAGGCCACCAATGCACTCTGTCTCT
GAGATAGGCTGTGATTTGTAGTGCAGGAGATCGAATGTATACACAAATGATGTAAATGGTTAACTGCCCAAGGAGAAAG
GCCAGAAAGAGGTGGGGCATGTGTGTGTTCTTAGTGTACAAAGAGAACAGGAGGGGCTCTCTGGATGAAGTGCATCTGT
CTCTGGGATCTGGAGGAGGAGAAAAATGTCCCAAAAGAGCAGAGAGCCACCGTAGGCTGTGCACGAGGAGCAACTTGTCT
GGGCTTATGGAATCTCAGAGGCGAAGTGTAAAGCAAAAGTCTTTGGGGGCCCAATATGGAATCTTCTGTCTTTAAAGGGGCT
TCTGGCCCTCTGCTGTGTGACTTGGGCAAGTTACTTCCACTCTAGTGTCTTGTGGCCATCTGTAAAGTGGTGAGGATTA
ATGCTATCTACACTGTGTGAAATGTGAAGTAATATTGTGCAAAAGGCTTATATAGGGGTCTTAATAGTACTAGTAGT
ACTTCAATGTCTTGACAAATTTAACTATTATTTATGTGATCAGGCTCACTCTCTCAGGGGAGTAATGTCCCTGTGCTGT
CTGTCAAAATTAACATTTGTTATCTCCTGTGGGCATCTGGCGAGGTGGCTAGGAAGCGTGGACTGTTTCTGTGTAGCGT
GCTGACATGCTG

SEQ ID: 37
>STAR37
ATCTCTCTCTGCCAAAGCAACAGCGGTCCCTGCCCCAAACGAGCTACCCGACTCAGTGGGGTACGGGATGCTGCTCCAGCA
CTCTAAACACTGCCGACCTGGTGGCTGCTGCTGCTACCCACCAACCCCGAGCGGCTCTCCCTGACGGCTGGGCTGTGGCACA
CTCTGCTCGATGTAGACATCTAGGACCTCTCTGGGCAGGACGCCCATCTACTGCACCTGGCTCTTCAGCAGCCCAACGACCGG
CTCAGGGGTTGCTTCCAGCGGGCTTGATCATAGAGTCAACATGCAGCCAGTGCTCTCAGGAAGATCTTTTAAATGAAGATGT
CGCGACAGGTTGGTGTGTAGTGGTGGAGCTGAGCTCAGATTGCTGCTGTAAGCAACACCGCTACTTCCAAAGAAAGACGAGGTGT
CTCTAAACCCGACGGGCGGGAGCAAAAGAGCTCACTTCCAGCGGGGAGTCTGGAGAGCCAGCGCTGTGTGGGAGGTACACTCT
GGCATCTTAGGCTCTGGCTGCTACTGCAGAGGAGGAGCGCGGCTCCCTCTCTGCTGGAGAAAGCTGGGTTCCCCACGCCCAACCC
TCTTCTGCACACCAACAGCGCTGTGTAGGCGAGGAGGCTGGGTGAAGACTGTGCTGATAGCCAAAAGCCACCGGCTTGGGA
CCCTGGCGGAGAGGAGCATCTTAAAGACA TGGAAAAATGGGTCATACCTCCATATAGACAGACATCTCTACATAAATAA
AAAGTGTGTATGGGAGAGGAGTTGGGAGCGGAAATAAAAAATGGCCAGACATCTGATAGACTGTTCCTGATGTTCGAGTGA
ACAGATGTCACATCTAGACACAGAGGAGGACAGAGCAAGGGCTGAAATTTGGCTTTATAGCAAACTGTGCTCTCGCAGC
CGCTCAA TGTGCTGTAGACACAGACA TGGAAAGCCAGGACAGGAGGCCACCGGGCCGACATAGGTTACAGAGAGCAGAGGCT
GCTGTTCTGTCTCTGCCACCCATCAGGGGTGACCTGCTGTATAA TGTGGTGGCTCTGCTTCCGACAGAAAAAGACAGCACT
TAAATCTCACATCTCAGACAGGACACAGAGAAAGGCTTCAAGTGGCTCTGTFTTCTGCTCCCTGTCGCGGACGTCACAA
CAGAGAGAGGTGTACGGACACA TGGCCCGCTGTACGAGCTCCCGCTGAGCTGTAGGCTCAGCAGGAGCTGCGCACTGACA
CACAGGGGACCCCACTCTTCCGCACTCTGGAGCGGCTTGCAGACAGAGCCGCACTGGGTCGTGCTGTCTATCCAGGGAACC
CACACATCTCTCTTTAAATGTGATCTCT

SEQ ID:38
>STAR38
GATCTATGGGAGTAGCTTCTCTAGTGAGCTTTCCTTCAAATACCTTGC AACCCAGGTAGAGAAATTTGGAGTGAAGGTTT
TTCTTCGTGTTCTTCTACCAATATGGAATATGCA TCTCTTTTGAAATAGTTAAAGTAAATTACCTCTCTTTCCAGATAGCTGTC
GTATCTGGAGACTTGGTATATCTTTTCCATCCGACCGTCTCTATTAACCCGATAACTATCTTTTGAAACAGCTGGGTGAGAAA
GACACTTGGTCAGGAAGCTGGGACGACAGGACCAACTACCGCTCTACCCACGAGCATAGAGTAAAGGCAAAACGAAGCTCTGTA
TAAGATACCGGTTGGCATCTGTATAGTGTGAGATGACGCGCTGTCTCTCGACAGCAGGAGTACACAGACACTTTTCTATTA
TTCTTAAATGCTTTTAAATGTCTAAGTAGGAAGTCTTGACATATAAATTTGATTAAGCAATATGATATCATAGMAATTTTGTG
GAAACGGAAATACCAAAAATGCGAAATGCAATGCTGTTCTTCTTGAGTGAAGACAGAAATTTCTTGTGTTAAAATAAATTTTATTT
ATATATTTGAGGTTGACCACATGACCTTAAAGGATACATATGACAGTAAACTGGTTACTACAGTGAAGCAAAATTAACATAT
CTACCATCTGATACATAGTTACATTTTTTGTGTGACAGAACAGCTAAATCTACGTATTTAAACAAAAATCTTAAGACAAAT
ACA TTTTAAATTAATCATAGCCCTCATGATGACATATAGATCGTGTGGTGTGTTTCTTCGCTCCCGCCGACCGCTCTCTCTCTG
TGATAGAGATCTCAATTCCTACGAGGCTCTCGGAGAACTGGCCCGCTTGCAGGGTAGGTCGCCCGGAGCTCGAGGCGGAGTAC
TAATGATCGAGCTCTTGGGGCGGGCTGGGAAAATCTGGGGGAAAATTAATCTGAGCTCTCAATGCCAGCTCTGTGTAG
ATGCTCTTGGGACAGCGCTGGAAAGTCTCCCTTCAGGAGAGCTTFCGCGAGTGCTCTACGGTGGTCTCGCTCGCCCGCCAGCGCTCA
TCTCATGTGACTCAACACGCGCGCTGGAGGAGGAGGCTGCGCTGGACAGCCCGGTGCGCCCTTTCCCTGGGGGAGCGGAG

FIG. 20, cont

CTGGAGCTCTGGCGGACGGCTGGGAGCGGGGCTCGGAGGCTGGGCTGGGGACCCAAAGGTTGGGCGGGGCGCAGGAGGT
GGGCTCAGGGTTCTCCAGAGAAATCCCAATGAGCTGACCCGAGGGCGGGCGGGCCAGTAGGCACCGGGCCCCGGGTGAC
CTGGCGGACCCGAAGCTGGAGCAGCCACTGCAAAATGCTGCGCTGACCCCAAAATGCTGTGCTCTTTAAATGTTTAAATTAAGA
ATAATTAATAGGTCCGGGTGGGAGGCTCAAGCCTTAATCCCGAGCCTGGCGAGGCGGAGGAGGA

SEQ ID:39

>STAR39

TCACATGCAACCTCCACCTCCAGGTTCAAGTGATTCTCTGCTCGGCTCCCGAGTAGCTGGGACTACAGGTGCAATGACA
CCGCACTCGGTAGTTTTTGTATTTTGTAGTAGAGACAGGGTTTCACATATGTTGGCCAGGTGGTCTCGAACTCCAGACCTT
GTGATCCGCCCCACCTCGGCTCCCAAGTGCTGGGATTACAGAGTGAGCCACTCGGCTGGGCTGCACCCCTTACTATAT
ATGCTTTGCAITTTCTTTTAGATTTGAAGAACCCTCATTAATACTCTAGCACTAATCTTATGTCAGTAAATGCAATAGCAA
ATATCTCCTGACGTGGGAGAAATATATATTGCAAGTCTTCTTTGTGAACATATGTTTTCAGTTCTAGGGAGCCAGACGCTTA
TGACTGAAAGCCCTAGTCAATCGTGGAGAAATGCAATCAACTTTGTGAAGAACTGCCAAACCTTTATTCATATATGCTGTAT
AAATTTTACATTACCACCAATAATGATGAGAGTTCCAGTTGCTTCACATCCTCACCAGCAATTTGCTTTGCTGCTCTTTT
TTCCTTTGGTTATCTAGTGGGCAATAGATATATAGTATCCTTTGTGGTTAATGTAAATCCACTGAAGACTAATAACA
TTTGCATATTTCTAATTAATAAGCCTTTTAAAGTGACTTTCAAGTCTTTGCTCATTTTATTTAGATATTTGCTTCTTAT
TATTGATTTGAAAGAAATATATTTAATGCTTATATTTCTGGTTAATAAGCCCTTTGTCATATTTTCCAAAACAAATATTTGG
TTGTTCTGACTACTTTCTTCTGCTCTTTGAAATGACTTTGGTGGCTTTGGCCAAAATCAATGACACATACATGTTGGGT
GCATCTCCAGACTACCACATTCGCTTTATCTAATTTGCTCTCCTTTGTGCAATAACACTCTGCTTTGATAAATGGTAAGTTT
TGAGATTCAGGCTGTGTAAGTCTCTTAATTTTCTGGGTTTCAATATTTGCTTTTAAATAATTTGATTTTCTATTTTCA
TTACATTTTAAATAAATCTGTTAGTGGGATTTGATTTGGCAATGCACTGAACCTGATCAATTTGGGGAGATTGGACA
TTCTATATATGGAATCCGCTGGTCACTCACTTTAAGAACTCTTTCTCAATCCATAGTAACCTCAATAGGTTGAGATGCTA
CTGTTTCTGCTCAGTCTGTGCTGAGGCTCTATGCTCTTCAATTTTGTCAATCCAAATTAACCTCAGCTTTGCAATCAATAC
TATTCTTGGCTTTGGTGGCTGTACCTCTCTCTTAATCAACCAATCCACAACTTACCTCCAAATTCAGGGCTTTGCTCATTC
TTCCAGGAGGAGTGTGCTCAGTCTATCTACTAGTATATTAATTTCTGCTGGCTTGGTATCAAGGCATCCCATTTCCGG
CTTCCATGAGATGCTCAGAGGGCAATGCTGCCCGGTGTAGCTGCATGGTCAAAGCTCTTCAATATCTCTTGGCTCATCACT
AACTCAGTATTTTGTACTCTGCTTCAGCTATAGGGAGCTACTGTAGTTTCTTGAAGACATATGCTCTCTCTCTCTCT
ACATCTGGACCTGAGCAGATCTCTGTACTGCTGCTTGAACCAATGTGATCCCGAGGCACACACCATTAGCTTAGAAGCTC
CCCTGATTTCTCAAGGCTGGTTGAGTCTCTCTCTGCTCTCATGCAACAGTTGGCAATCTCTCTGCTGAGCAGCTAGC
CCATGATGCTCTTTGGAGGAGAGAGCTGAGTCTTTCTCACTATTTGAATTTCCAGCATTCATCAGAGGCTGGCATATATA
AAGCCCTCCATATATGATATTAAGTGAATGGATAAATGAATAAAGTTATATATATGATATATATGATATATATGATAT
TGATATATGATATATATGATATGATATATGATATATGATATATATATGATATATATGATATATATGATATATATGATAT
GTATATATATGATATATGATATATGATATATGATATATATATGATATATATGATATATATGATATATATGATATATAT
ATCTCTGATTTTAAACCAATATGGAATAAAGCTAGAGGAGCTTAAGGCAAAATACCAATGTTCTCACTCACAAAATCTAAA
CAAGTTGAATCTTACAAGTAGAGAGTAGGATGATGTTTACCAAGGGCTGGGGAGCGGAGAGGATGGGAAAGCATAGCT
GTCCATCAAGGGTAGAAAGTTTCAATTTAGACAAGAGGAATCAGCTTTAGTCACTATTTTAC

SEQ ID:40

>STAR40

GCTGTGATTCAAACTGTCAGCGAGATAAGGCAGCAGATCAAGAAAGCACTCCGGGCTCCAGAAGGAGCTTCAGGCGCAGC
TTTGAGCATAAAGCTGTGATGAGCAGTGAAGTCTTGAATAGTGTTCAGGCGCAGCATGTTACCATTCATGCTTGAATCTTA
GCCAGTGTGAGGAGGCTGGAGTCAAGTCTCTAGAGAGTTGAGCAGCTCCAGCCTTAGATCTCCAGTCTTATGCGGTGT
GCCCATTCGCTTTGTGCTGAGCAGTCCCTGGCCACACCCAGTAAACAGTTCTGGGATCTATGGGAGTAGCTTCTTATGATG
CTTTCCCTTCAAAACTTTTGAACAGGTAGAGAAATTTGGAGTGAAGGTTTGTCTCTGTTTCTTCAAAATAAGGATAT
GCATCTCTTTTGAATAATTAAGTAAATTAACCTCTCTTTTCAATACTGCTCTCATGCGAACTTGGTATCTGTTTCCA
TCCAGCCTTTCTATAAACCAGTAACTCTTTTGAACAGGTGGGTGAGAAAGACACCTGGTCAAGGAACGCGGACCAAG
GACAACTCAGGCTCAGCCAGGCTCAGACTAAAGGCAACAGGAGCTCTGTATAAAGTACCGGTGGCATGTGATATAGTG
GAGATGCGGCTGTGCTCTGAGACAGGGAGTCAACAGACACTTTCTATAAATTTCTTAAGTGTCTTGAATGTTCAAGTA
GAAAGCTTAACATTAATTTGATGAACAAATGATATTTCAATGGAATATTTTGAACGGAATACCAAAAAATGCGCAATAGT
GGTCTTTCTGGAAGGAGCAAACTTTCTGTTTAAATTAATTTTATATATTTGAGGTGAGCAGATGACCTT
AAGGATACATATAGACAGTAACTGGTTACTACAGTGAAGCAAAATTAACATATCTACCATCTACATAGTTACATTTTCTT
GTGTGACAGGAACAGCTAAAACTACGTTTAAACAAAAATCTTAAAGCAATACATTTTATTAATATAGCCCTCATGA
TGATACATTAGATCTCTAA

SEQ ID:41

>STAR41

TGCTCTGTTGCCAGGCTGCAGTGCAATGGCGCTGTCTCGGCTCATCGCAACCTCGGCTCCAGATTCAGGTGATTTCTC
CTGCTCAACCTCCCAAGTAGCTGGGATTACAGTATGCGAGCAACAGCCCGGCTAATTTTGTATTTGTAAATAGAGACGGG
GTTTCTTCAATGTTGGTCAAGCTGGTCTCAAAATCTCTGCGCTCAGGTGATCTGCCCACTTTGGCTCCCAAGGTGCTGGGAT
TACAGGCAATGAGCCACTGTGCCCGGCTGGGCTGGGCTTTTAAAGGGAGTGGAGGTTGAGGGCTGGAAAAATGGGAGAG
TTGATTTGGTGGGGCAAGGGGATGTAATCATCAGGGTGTACAAACTGCACCTTTGGTTTGTAGTCACTCTCTGTTGGGTCTC
TCGGAGCAGCTCAGTCACTAGCTCCATCAGTATACAGGACCCAAAGGAATATCTCAAGGGAAACAGCAATTTCTTAAGGT
TCAAGTTGTGATCTACGGAGCAGTTAGGGAACTACAACTCTGTGACAGGCTTACATGCTTCTGAGGCAATGAGACACCA
AGCAGCTACGAGGAAGCAGTCAAGAGGACACCGGACCTTAGTGACTGATGCTGATGTGCTGCGAGCTGGGTTCATTTTCAAT
TCTCCCTCCCTTCCCTTCAATTAATTTTGAAGTTTATAGGGAACATTTCAACCACTCTGCTGTGGAATCTCTGCTCACT
ACGGAGTCTGTCACTCTGGCTGTATGGGCTGTGGGCTCTGCGGTGCCATTTCTCAGGAGGTGTGAGACCAATGAGACCGG
AGGTGAGCAAGGCTAGAGACACACCCCGGCTCCATCCAAATCATGTTTCTTGGGTGCTTGGTTCTATGACGGCTGCA

FIG. 20, cont

TGTCCTTAGTCCCTGCATGGGAACAGCTCCCTGTGGTGAGCAGGCCCTGAGGAAGCCCTTGAGCGGGAAATGGAGCCTAGGC
TTAGGCTGCCCTGGTAAGAGCTGGAGGGAAACAGCGAGGCTTTGTGCTACTTTTTTCCAGAATGAAATACGTGACTGATG
TTGCTGTCTCGCAGCGCCACGTTTCCCGCCACAACCCACGGGAACGAGGATGAGGAGAACACACTCTCCGTGCATGCCACAC
G

SEQ ID: 42

>STAR42

AAGGGTGAGATCACTAGGGAGGGAGGAAGGAGCTATAAAAGAAAGAGGTCACATCACAATTTTAAACCTTTTAAACCT
TTGGTTTTTAAATGTCGGTGTCTCTCATTAGCAGTAAGCCCTGTGGAAGCAGGAGCTTTTCTCATTTGACCAACCATGACAAG
ACCCATTTATGAACATAATAGACACACAAATGTTATCGGATATTTATGAAATATAGGAATTTTCCCTCACACCTC
ATGACCACATTTCTGGTACATTTGTATGAATGAATATACCAATTTTACCTATGCTGTATTTAGGTCCTTTCTGTCAGG
CTATAAAATATGTATGGCCCGGTACAGTGACTTACGCCCGTAGTCCAGAACTTTGGGAGGCGGAGGCGGGTGGATCAC
CTGAGGTGGGAGTTCAAAACAGCCTGACCAACATGGAGAAACCCGCTCTGCTTAAAAATACAAAAATTAACCTGGACAC
GGTGGCTATGCTCTAAATCCACAGCTACCGGGAAGCTGAGGCAAGGAGAACTGCTTGAACCCAGGAGCGGAGGTGTGGT
GAGTCGAGATTTGGCCCATTTGCACTCCAGCTGGGCAACAAGAGCGAAATTCATCTCAAAAAAAGAAAAAGTATGACTG
TATTAGAGTAGTATGTGGATTTGAAAAATTAATAAGTGTGCCAACTTACCTTAGGGTTTATACCATTTATGAGGGTGTG
GGTTTC

SEQ ID: 43

>STAR43

CAATAGATCTACACAAAACAGATAATGTCTGCCCATTTTCCAAAGATAATGTGGTGAAGTGGGTAGAGAGAAATGCAT
CCATTTCTCCACCAACCTCTGCTAAATTTGCCATGTACAGTACTGAGACCGGGGCTTATTCACGCGGCAGAAATG
TGCACCAAGCACCTCTGTCTCAATTTGCACTCTAGGCCCTGCTATTTGATGGTGTGAAGGCTTGCACCTGGCATGGAAGG
TCCGTTTTGTACTTCTGCTTTAGCAGTTCAAGAGCAGGGAGAGCTGCGAGGGCCCTGCGAGCTTCAGATGGATGGTCT
AGCTGTGTGGAGGGCGCTCTGTGGTCAATATCTCCAGCCCGCTGCGGTGTGCTGTGTGCTTGGCTTGGCTGTCTGGCTCTC
CATGCCCTGTGTGGCTCCAAAATGTCACTAGCTGCAACCCAGGAAGAAATGTCAGGCGCCATCTCTTTATGTGCTTTGGGC
TATTTGATTTCCCGTGGGTATATTCCTTAGTGAAGCCAGGAACACAGGAGGTAGTGTCTTTGGGAGAGTTTGGACC
TATGGGTATGAGGTAAZAGACACAGTATCTTCTCTTTCATTTGGTGAGACTGTAGCTCTGGCCGCGGACTGAATTCACAC
CAGCTCACTTGGGAAACTTTATTTCCAAACATAGTCACTTGAACATTTGTCGAGATGAGGGACACAGAGAGGCCCTAG
ATTTGTACATCTGGGTGTATGTCTATAAATAGAAATGCTTTGGTGGTCAACTAGACTTGTTCATGTTCACATTTAGCTCTG
CCTTTTCGGTGGTGATTTAAAAATTATGTATATCTGTGTTGGAATATAGTGGAGCTATGGTGTGGCATTTCATCTGGCTT
TTGTTTAGCTACAGCCGCTCTGTATTTGGGAGGCTTTGAAGCTCAGTAGCTAAAGAGAGGTATCCCTCACTCCCTCCAGAG
AGCGGTCCCTCCAGCGCTCAATGAGAGTTTGTAGCACCTTGAATGAGTTAAAC TTGTTTATTTTAAAAACA TCTTGG
TTATGAATGTGCTTATATGAATTAAGTGAACAACTTATGTTGTGAAGATTTGATTTGGTGCTAAGGTGTATAAATTCAG
GGACCAAGTCTCTGGAAGAGTTCAATTTAGCATGAAGTCAGCCTGTGGCAGGTTGGGTGGAGCCAGGGAACAATGGAGAAAC
TTTCACTGGGTGG

SEQ ID: 44

>STAR44

TGAGTTGGGGTCCTAAGCCAGAGTTAACTATGCTTTCAZATATTC TTGCAAGTAGAAGTACAGTGTGGTGTAAATCCCG
CTTAGATGGAATAGCTAGGCCAGAGGAATAATGGTAAATGGAACCATATGACCGTATGCAATTCATGTGCATATTTATAT
CAAGAAAAGAACATTA TAGGTGGGTGAGACCCCTATTTGTTCTGACAAATGTCATCTGTATTIACATGCTCTGTTTCGGGAG
TTTGGATGCAAGGGATTTCTGTCTGGAATGTAAAGCATGTGCTTCTGCTGTATGTAGCTAC TCAAATTTGTATTCTTGAC
TAATAAAGTCATAAAACATAAATCAACCTCTGTGTGGTGTCTCTCTTCCAATTAATTTATCTTTAGCAAAAAGTATGAAAT
GTGTGTGTTATGTAAACATTTCTATATAAATATATTAATGATTTATAGCTTTATTCATAAAGTTTAAAGTGTTTCTCT
CTATGACTACATTTATTTTAAACAGAAATTTCTTTAACTGAAACCTCAAGGAAGACTATCTGGGTAACCTTTTCAAAAA
GAATTTGTCCTGTATTTTGGGATGAATATATTAATTTCTTGACTGTTTAAACAGCACATAATTTTACAGACAAAGCCAC
TTTTTCAAGCGCTGCTTCTCTCTCCATTTTCCCTATCTCTGTGATTTGACACCTCCAAACCCCTGTAGCCTGCTCTGCTCTC
TCTTAAACAGTCTCTGATGATCTACTTCTTAAGTATTTTTCAGCCCTGCTCTCTCTCCATCATGATGGATTCACCTTCCA
GTTGAAATCCTTATGTTACCTCTCCCTGGATTAAGGAGTAA TCAGAGAGCTGGTCTCTCTTAACTCAGGATTCACCTCTCT
CATCTGTTGTTTACAGTGACATCAGAAAGATATTTTAAAAATGATGAAC TAGAAATTAATATAAAAAACACATACACACA
TAAATAATAC TAAATTTTCAATGATGTTCCAAATTAATGAAAAATAAATATAGGAGCCACTTATGTCTGGCCCTCAATC
TTTCAATTCAAACTTATCTCTGCCACTATCTCTTTGAACATTTGATTTCCAGCTACTTTAGAAATAATAATAATACATAAT
ATTATAGAGCCCTTCTCTGGGTCTCTACCGTACAAAATACCTTACATATAACATTTAATCTTTGACAACCTTATATAGG
CATGCACAAATTAATCTATCTATATATCTATATATAATAAAATCTATATTTATAGATAAGAAAAATAGAGGTTAA
AAACTTGGC AAAATTAACAAAGCTTGAAGGTGAGCAGTTGGGATTTGAATCTAGGCATCTGCTGCTATAGCTACAGTGG
CTTCTTGTGCAAAAAGCCTTGCACTTCCCTAGACTTAACTTTCTCAAAATCTGTGCTTTTACATGCTCTTCCAAATGT
CTGAAAAATCTTTTCCCAACCTTCAGTCTAACTGTGGTACTCA TGTTCACCCCAAGAAATGACTCCATCTGTCCCTCTCC
ATGAAAAATTTCTTTGAATCTCAGCACTTTGGGAGGCTGAGGCAAGT

SEQ ID: 45

>STAR45

CACGCCCCAGCGTGCCTGGACTACTGCTCCGAGGACTCCTGTTCTGCTGCACTTGGACTACGGCACAGAGGACCCAG
CTCCCGCCGGCTGAGCTATGACACAGAGGACCCAGCTCCCGGAGCCTGGACTATGGCACCAGAGGACCCAGCCCGG
CTTCTGGGCTTAAGGCACAGTAGGACCTGCTCATGTGTACTCTGCTCAGGAGGACCTCGCAGGGCGGCGCACTGGA
CTAAGCTACTTAAGGAGCCCAACCCCTGCTTAACCTGGACTAAGGCACTGGAGAACTCTTCTGCTCCGACAGCCAGGACT
CTTGCAAGAGAACTTCAGCCGCGTCCCTGGACTGTGGCAGTAGGGCCACACCCAGCCCATGGACTCTGTATTT

FIG. 20, cont

GGAGGAAGAGTAGTGATAAATGTCAGGTTTACAACCTGGAAGTAGCAATCAATGTGCCACAATAGATGGATGTGATGTA
 AAATTATAAATGATGAAAACATTATGTTGTAATTCCTACCCAGAACAGTTACACAAGACAAAAGACGTAAGAAAATCCACA
 TAGGGAAGGAAGAGGTAAGATTGTTCTGTTTTTGGAAAATATAATCTTAAGATAGAGAAAACTTAAGATTCCACCAAA
 ATAAATGGTTATAGCTGATGAAAGAAATCAATAAAGTTAAATAGTTACAAAATCAACATACAAAATATCATATATGTTCTAT
 TAACATATGACAAAATATTACCTGAAAAATAAAGGCAATTCAATTTATAATAGAAATCAAAAACAGATATATAAATATATAAA
 AGACAGGAGTAAATTTAAATCAAAACCATAAAGATTACATACGAAAACATATAGCACAATTGATGAAAAAAATTAATAATGG
 CATAAATAAATGGGAAAAATCTTCATTGATGGATTCAAAAATTAGTATTTGTAAGATGTCAAATGCTACCCAAAGCAATC
 TACAGATTAAATGCAACCACTATCAAAATTCCAATGTCAATTCACAGAAAATAGAAAAATTAAGTCAAAAATTTGATGGA
 ACCACAAAAGACCTGGACCAACCAAGCAATCTTGAACAAAAGAACAAAGCTGGAGGCAATCAGACTACCTGACTCCAAAC
 TCTATTACAAAAGCTATAGGAATTAAAAACAGCATAGCAATGGCAATAAAAACAGACATGTAAAAACAGTACAAAGGGAATAGA
 ACCTGTAATAAATCCGTTGCTGCTGCTGCTCAATTGATTTTTGATAAATAAATCAAAAATACACAGTGAAGAAAAGAAATP
 ATTTTCAATAAATGGTGTAGACAAAACAGCATATCCACATACAGAAGATAAATAATTTGACTTTTATTTGCTCTTTATACA
 AGCATCAAAATCAAAAATTAAGTTTAAATGTAAACTACTACAAAGAAATATAGAAAGGAGACTGTATGACATTGGCCTGAGC
 TATGATTTTCGTAGATTATTCAAAAGGCAACAAAGCAAAACACACAAATGAGACTGCATAAACTTAAACTTTTCCA
 CAGGAAAAGAGCAATGATAGAAATTAAGAGAACCCACAAATGGGATAATATTTTAAACCATACATCAGTAAGGGGCTCA
 TATAATAATATATAGCAACTCAACTACTCAAAAATTAAGAAAAAACTATGCTTATTAATAAATAGCAAGAAATCAGAA
 TAGACATTTCTCATCATACAAAAGGCCAACCCAGGTACATGAAAAATCATAAACATTCTAATTTATCAGAGAAAGTGCA
 ATCAATGCCACAATGAGATATCACTTCACACATTTTACTAGGGCTATTAATAAAAAAGATGGAAGAATAAGTGTGGTGAGG
 ATGTGGAGAAAAAGAACCTGTACACTGTGTGGTAGGAATGGAATTAGTACAGCAATCTTGGAAAACAGTACGAAGCTTT
 CTCAGAATAATATAAATTTATTTACCTATGATCCA

SEQ ID:46

>STAR46

ATGTTTTCTCGCCCTTCGCATTTTCGCAAAATCTGTTGAATCATTCAGTTACTTAGGTTTGCCTTCGCTCCCCCAT
 TACAAACTACTTACTGGGTTTTCAACCCCTAGTTCCCTCATTTTATGATTATGCTCATTTCTTTGTACACTTCGCTCTG
 CTCCTACTCCCACTCATGGCCCTGGCTTTGGATTAATGTTTTGGTCTTTTATTTTGTCTTCCTCTACCTCAACACT
 ATCTTCTCTCCAGCTCTCCGGTACCTATCACCAAGGTTGTCATTAACCTTTTATATTTCTTCAATATCCATGATATC
 ATTTGCAATAAGCGTATATTAACAAAATCACAGGTTTATGGAGATATAATTCATATACCTTAAATTCAGGCTTTTAAAG
 TGTACCTTTTCAATGGTTTTGGTATATTCACAAAGTTATGCAATGATCACCACTCTGATTCCATAACATGTTCAATZAC
 CTCAAAAGAAAGTCTGTACTATTAGTAGTCAATTCACATTCACCACTCCCTCTGGCTCTGGGCAGTCACTGATCTTTGTG
 TCTCTATGGAATTTGCCCTAGTCTAGGTAATTTTATGTAATGGCATATACAAATGTAACCTTTTGTGGCTTTTTCAT
 TTAGCAAAATGTTATCAAGGCTCTGCTCTGTTGTAGCATGTATTAGCACTTCAATTTCTTATATGCTGAATGAZATACTTTA
 TTTGTCCATCAGTTGTTCACTGCTTTATTTGTCCATCAGTTGATGAACATTTGCGTTTTTCCACTTTGGGCTATTAAAGAA
 AATGCTACTGTGGAACAAGTGTGTACAAGTTCTCTACAAATTTTGTGTGGACATATCTTTTCACTTCTCTCAGGTTGTA
 TCTGGGAATTTGAATGCTGGGCTGTGTAGTAGTATGCTTAAACACTTTGAGAAATGCTATATAATGTTCTCCAGAGCTGTAC
 CATTTTAAATCTGTGTATGAGGATTCACAGTTCTCCACTTCTCACCAGTGTATGGATTTGGGGGTATACTTTAAAAAA
 GTGGGAATAGGCTGGGCAGTGGCTACAGTTGTAAATCCCAACACTTCAGGAAGCTCAGGTGGGAGGATCACTTTGAGCCT
 AGTAGTTTGAAGCAGCTGGGCAACATAGGGAGACCTGTCTCTACAAAAATTAATTTAAAAATAATAGCTGGGGCTGTG
 TGGCACAACCTGTAGTCCAGCTACATGGGAGGCTGAGGTGGGAAGGATTCCTGAGCCAGAAATTTGAGGTTCAGTGA
 GCCATGATGGCAGCACTATACGTAGCTTGGGTGTCAGAGCAAGACTCCGTTTCAGGGAAGAAAAAAAGTGGGATGAT
 ATTTTGCACACTTTCTCTCTGTTTTCTTAAATTTCACTCTCTGAAAATTCATTAATAATAGCTGGTACCACTCTAACTCA
 TTGTGTTTCAATGGCTGCATAGTAATTTGCAATAATAAATAATCCATTCATTCAAAGTTAGCAGATATTTGACTGTTA
 GGTGCCAGGCACTGCTTAAGCGTTAAAGAAAAACACACAAAACTTTTGCAATCTTAGAGTTTATTTTCCAAATGGAGGGG
 GTGGAGGGAGGTAAAGAAATTAAGAAATAAATTAATTTACATATATAGCATAGGGTTTACCAGTGAGTGAGCTTGAATCGT
 TGGCAGCTTTCTTAGTAGTATAAATACAGTACTAAGATGAAATTAATCTTAAATGGTGTACTTTAAATTTACTGGAAATAGGT
 ATTTACTATTAGTCACTTTGCAGGTGAAGTGGAAACCACTCGTAAAAATGTAATAAGAAACAGCTGGTTAATGTT

SEQ ID:47

>STAR47

ATCATTAGTCAATTAGGGAATGCAAAATGAAAAACACAAGCAGCCACCAATATACACCTACTAGGATGATTTAAAGGAAAAAT
 AAGTGTGAAGAAGGACGTAAAGAAATTTGAACCTGTATACATTGATGGTAGAAATGGAATAAAGTTGCAGCCACTGTGAAAA
 ACAGCTTGCAGTGGCTCAGAGGTTAAATATAGAACCCCTGTGGACCCAGGAATCTACTCTTAGGCACCCCAAGAAATA
 GAGAACAGAAATCAAAACAGATGTTTGTATACATAATGTTTGTAGCATCACTTTTACAGGAGCCAAAAGGTGGAAATATCC
 AACCATCAGTGAACAAATGAATGTAATAAAGCAAGGTGGTCTGCTGCAATGCTACATCATCCATCTGTAATAAAGCAAC
 ATCATTTTGAATAGATGATCAACATGGGTGGACATTTGAGAACATTTAGCTTAGTGAAATAAGCCAGACAAAAAGGAATAT
 ATTTGATAAATTTGATAATACATGAAGTGCCTAGAAATAGTCAAAATTCATACAGAGAAAGTGGGATAGGAATCACCATGGGCT
 GGAATAAGGGGAAGCTCTATCTCTTATTTGTCGACAAAGGTTTCTGTAAGAAATCATCAAAAATTTGCGGTGTAGATAGTG
 GTGTTGGTTATGCAACCCCTGTGAATATATGGAATGCCATGGAGTGACACTTTGGTTAAAGGTTCAAAATGATAAATATTTG
 TGTATATATATTTTCCACGATAGAAAACAGCACAGCCAAAGCCACATGCCAGTCTTTGTTAGCTTCTCTCTCTTACCTT
 CAAGAGTGGGCTGAAGCTTGTCCAACTTTCAAGGTTGCTGAAGACTGTATGATGGAAGTCACTGCAATGGGAAAGAAAT
 TAAATGGAGAGAGGAAAACTTGAGAAATCCACACTACTCACCCTGCAGGGCCAAAGAACTCTGCTCCCATGGCTTGTCTGTC
 CTGCTCAGTATTTCTGTCACCTCTCTTTTCAACTGAAGACTTTGTACCTGAAGGGTTTCCAGGTTTTTCACTTCG
 GCCTTGTGAGGACTGATCTCTCAACT

SEQ ID:48

>STAR48

FIG. 20, cont

ATCATGTAATTTGTTTCTGAAATTAATCTTAGATACATTAATGTTTATGTTACCATGAATGTGATATTATAATATAATAT
 TTTTAAATGGTTGCTACTGTTTATAAGAAATTCATTTTCTGTTTACTTTGCTTCATATCTGAAAACCTTGCTGATTTGAT
 TAGTGCCATCCACAAATTTTCTGGATTTCTATGGGTAATTACAAATCTCCACACAAATGAGTTGCTGAGCCAAAGATCA
 CACCAGTGTACTCCAGCTGGGGCAGAGTGAGACACCACTCCACAAAAACACATAAAACAAACAGAACTCCACAC
 AATGACAACTGATGTGCTTTCTTTTCTTCCCTCTCTATAATATTTCTTTGCTCTATCTTAACGAACTGGCCAGAAA
 CCCAGGACAAATGATAAATACGAGCAGTGCAACAGACATCTCATTTCCCTTTCTAGCTTTTATAAAAAATACGATTATGC
 TTCAACATTAATATGTTGTTGTCGATGGTTTGTATTATAGATAAGCTTATCAGGTTAAGAAAATTGCTGCTTTCTAGT
 TTGGTATAAAGATTATAATAAATGAATGTTGTATTTATCACTCTTATTTTCTTACATCTGCTAAGGTAACTCTGTG
 TTTTCCCTTTTCAATCTCCATAATGTTGTTGATGACATTAAAAATACCTTCTATTGTTAAAAATATCTTTGCAACGCTGTATA
 GAACCAATGCTTTATCTGTATTGCTGATGGATTTTTGAATAATATGTAAGGTGGACTTAGTTTCTAAGGGGAATAGAAT
 TTCTAATATATTAAAAATATTGTGATGATGTTCTGAAGGACATGGTGTGCTATTTCTATACCATCTGCTACTAGAGG
 AGCCGACTGAAGTCTACACTGCCGAGGAGGGAGAGGTGCTCTTCCGTTTCTGTTGCTGTTAGCCATCTCCAGTGGTAGC
 TGCAGTGATATAATGTTGCTGAGTGCCGACAGTTCTGGAAGGAGCAACACAGTGAATTCAGCAGCAGCAGTATTGCGGGAT
 CCCACGATGGAGCAAGGAAAATAATCTGGAAGCAATGACAAATATCAGCTGTTGGCTATAGCAGCTGAGATGTAGTCTCT
 ACCGTGGCAGCTTCAAGGACAGTAGTGAATGGTCCAAATGGCGCCAGACCTAGAAAATGACATTTCTCAGCACCGGCTCCA
 GATGCTGAGCTTGGACAGCTGACGCCCT

SEQ ID:49

>STAR49

AAACAGAAACCCAAACAAATGGGAGTGACATGCTAAACCCAGAAACCCAAACAAATGGGAGGCTCTGCTAAACCCAGAAA
 CCCAAACAAATGGGAGTGAAGTCTAAACCCAGAAACCCAAACAAATGGGAGTGCTCTGCTACACAGAAACCCAAACGA
 TGGGAGTGACGTGATAAAACAGACACCCAAACAAATGGGAGTGACCTGCTAAACCCAGAAACCCAAACAAATGGGAGTGAC
 GTGCTAAACCTGGAACCTAAACCAATGGGAGTGAGGTGCTAAACCCAGAAATCCATAACAAATGAGTGACGTGCTAAAC
 CAGAACCCAAACAAATGGGAGTGACGTGCTAAACCCAGGAAACCCAAACAAATGAGAGTGACGTGCTAAACCCAGAAACCCAA
 ACATGGGAAATGACCTGCTAAACCCGGAACCCAAACAAATGGGAGTGATGCTAAACCCAGAAACCCAAACAAATGGGAAAT
 GACATGCTAAACCTGGAACCCAAACCAATGGTAACTAAGAGTGATGCTAAGGCCCTACATTTTGGTACACTCTCAACTAA
 GTGAGAACTTGACTGAAGAGGAGGATTTTCTTAAAGCAGAGTTTGGTCTGTCGCCAGAGTGAGTGAGTGAGTGAGT
 GATCTGCGCTCACTGCAAGCTCTGCTTCCGGGTTGAGGCCATTTCTGCTGAGCTGCTGAGTGTGGGAAATACAGG
 CACCGCCACACACTTGCTGCTAAATTTTGTATTTTGTAGAGATGGGGTTACCACTATTAGCAAGGATGGTCTCAATC
 TCTGACCTGTGATCTGCCACCTTCCAAAGTCTGGGATTACAGGTGTAGGCCAACACCCAGCAAAAGGA
 GGAATTTTAAAGCAAAATATGGGAGGCCATTTTGTAACTAAGCTCAATGCAATAGGTGCCAACAGACCAAAACCAAC
 AAACCAAAATGGAGTCACTCATGCTAAATGTAGCATTAATCAAA

SEQ ID:50

>STAR50

CAACCATGTTCCGCAAGAGCGGCTGTATTATAACATGAATGAGGGAAAGCCTAGTAGCTCCATTGGAATGGGAAGA
 ATGGCAAGAGAGACAGGCTCATTTCTAGAAAGCAATCTCACACCTGTGGTCTCACCCATTGAATGTCTCACCCA
 ATCTCCAAACAGAGAAATGAGTGACTGTGTGTCACATGCGTGTGCAATGTGTGAAGATGAGTGGAATGTCTATATGG
 GAACATATATGTGAATGTATGTGTGTAACATATGTGTGACTGGCAGCTGGGGAGTGCTGGTTGGAGTGTGGTGTGATGTA
 GATGCAATGATGCTGTGTGTGATGATGCTGGCGGAGGCGGAAGGGAGAACAGCAGGCTCAGGTGTCCGAGAGAGGC
 TGGGAGGAACATATAACCTGGGCAATTTCTCTCTCAACAGGCTTTCTTGGCCAAATAGCGGCAGAGCTCAAAAGTTCA
 CAGAGATAGTGCTGGGAGGCAATGAGGCAAGGCGGAAGTACTGCGAGGAGGGGAGAGGGTCTGACACTTGAGGGGTTCTA
 ATGGGAAGGAAGAGCCACATGAATTTCCACTTAGCCCCAGAGCCCTGGGCGCAGCGGTGCGCGGCTTCCAAACATACCAAC
 CATTTCCAAATGTGTGGCGCAGAAATTAACCTCTCTTAGCTCAGTTTCCCCACCTGTAAAAATGGCAGAAATTAACCAAGCT
 TACCTTTCCCGCAGTGTGTGAGGATGAAGAGAGCTATGTAGCTGATGCACTTAGAAGAAAGGCTAGGGTGTGAGTGGTACT
 CGTCTGGTGGGTGTGGGAAGACATTTAGGCAATGAGGACTGGGGAGAGCCTGGCCCATGGCTTCCACTCAGCAAGGTCA
 GTCTCTGTGCTCTGCACTGCCAGCTTCCAGAGAGGACCTTCCCAACAGCACTGCCAGCTGCCAGTCAACATATAGTT
 ACACACATACAATCACATATATGTTCCCATATAGACACATTCACACTCATACCTTCACACATGCACAGGCAATGTGCACACA
 CAGTCACTCATTTCTGTGTGGAGATGGGTGAGGACATTCATGGGTGAGGACCAACAGGTGTGAAGATTGCTTTCTAGA
 AAAATGACTCTGTCTCTTTTCCCATTTCTTCCCAATCCGATGGAGCTACTAGGCTTTTCCCTCATTTCAATGTTAAATAAAC
 CTTCCTCAAATGGGCAAAATGGGCTTTCTCAAGAGTGGTGAGTGTCCCATCCCTGCGGTGGGACAGGGGTGGCAGCGGACAA
 GCGTGGCTGGAGGGAAGTGTGAGGCTGATTCOCAGTCCAACTCCAGCTTCCAAACCTTCACTCTCCAGGAGTCTTCAATTC
 TTGGCTCTAATTTTGGCTCTTTGTTTTCTTTTTTATTTTATCGAGAACTGGGTGGAGAGCTTTTGGTGTCTATGGGGAATGC
 TTTGAACCCCTCTCTGCTCAACATGGGAGCTGGCTTGAGTCAACTGGTCTCCATGGAAATTTCTTTTATGTGTATAA
 CAGCTAAGTTTTAGGCGCTGTGTGCTCCGTCAGGGTGGAAAGCAGCCGTGTGATGTGGAATGCTTGGCTCAGATTTCTT
 GGGCAACAGATGCCGTCTCTCAACTCACCATAAAGAGCCCAAAATATGTTGCTTGGAGACCAATGCTGTGGTTATG
 TCTAGTAAATTCAGCTGCGTTCACTTGGGAAGCCCTTCTGAATGTCAAAGCCATGAGATAAAGGACATATATATAGTAGCT
 AGGGTGTGCCACTCTTTAGGGGCCATCTCCGAGGTGTGAGCCTAAGTCCAGGAAGAGAGGAACTCTGTTTTGGAGC
 CAAAGCATAAAAAACCTTAGCCACAAACCATGAACATTTGTTTGTGAGGTTCTGAGTCCAGGAGGGGCTCTGAGGA
 GAGGGGAGCTGGAGCTGGTAGGATTAATGTAGATGGAGCAAGGGCCCTTAAAGAGGTGGGAGCAGCATGAGCAAGGCA
 GAGAGGTGGTAAATGATAGGATGTGCAATGGGAAGAGTTTGGCTGGAAACAGAGTTACAGAAATAGAAAAATTAACACTA
 TTAATTTAGGCTCTTACTAGCTGTCTGACATTTCTAGTCACTGAGATAGGTTTGGTATACAAAAACAAAAATCCATCTCTA
 TGGACATTTTAGTACTAACACAAATATAAATAAATAAGTGAACAAAGCTCAAAACATGCCAGGCACTATTTATTTATTT
 ATTTATTTATTTATTTATTTATTTTGAACAGAGCTCTGCTCTGTGCTCCAGGCTGGAGTGTAGTGGTGGCATCTGGC
 TCACTG

SEQ ID:51

FIG. 20, cont

>STAR51
 TCACAGGTGACACCAATCCCTGACCACGCTTTGAGAAGCACTGTACTAGATTGACTTTCTAAATGTCAGTCTTCATTTTCT
 AGCTCTGTTACAGCCATGGTCTCCATATTATCTAGTACAACACATACAAATATGTTGTATACAGTATGAATATAATATA
 AAAATATGTTTATA
 ACTGCTCTAAGCACTTTACCTGACAGGAACAGTTTTTTTTTATTTTGGTGAATACAACTAACATAAAATTTATTACAAAT
 TTTAAGCAATTTTAAGTGTATAGTTTAGTGGAGTTAAATATTCAAAAATGTTGTCAGCCGTACCACTCACTAGTCTTCAT
 AACTCTTTTCATATGTAAAAATTAAGTTTATGCTCATTTAAAAATGACTCCCAATTTCCCTCCCTCAACCTCTGGAA
 ACTACCATCTATTTTCGCTCCGTAGTTTTCGCCACTCTAAGTACCTCACA TAAGTGAATTTGTCTTATTTGCTGTGT
 TGTGACCCGCTGATTTCAATTTAGTATAATGTCCTCAAGTTTATTCAGGTTATATAGCATATGTCATAATTTCTTCACIT
 TTAAGCTTGAGTAATATTTTCATCGTATGTATCTCACATTTTGTCTATCCATTCACTCTCTCAGTGGACACTTGAGTTGCTTC
 TACATTTTAGCTGTGTGAACTACTGCTGCTATGAACATGGGTGTATAAATATCTCAAGACCTTTTATCAGTTTTTAAAA
 TATATCTCAGTAGTAGTTTAGCTGGATTAATGTTAAATTTTATTTTAAATTTTGAGGAACGTCTCTACCTTTTATTCAT
 ATAGTAGCTATACCAATTGACAAATGGCAATCTTACCAACAGGGCATTAAGGGTCTCAATTCCTCACATATTTCCCTGATC
 TTGTTATTTTCAGGTGTTTTTTTTTTTTTTTTTTTATGAGGACCATGTTAAGGGTGTAAAGGTGATATTTCAATTAAG
 TTTGATTTGCAATTTCCCTAAATGATTAGTATGTTAAGCATCTCTTCATGTGCCATTTGGCCATTTGTAZATCTTCCTTAA
 AAAATATATATATACATCTCTTTGCCATTTTGAATTAATGTTTTATTTTTTGTATTGAGTTTCAATACTTTCTATATA
 ACCTAGGTATTAATCCCTTATCAGACTTAAGATTGCAAAATATCTCTTTTCATCCACAGGTGTGTAATCTCTCTGTGG
 TAAATCTTTTGTATGCTGTGTGTGTCAGAAATGATTTCACTCTCTGTGGGTCTTGTGTCCTACTGACTTCAAGAAATAAGCTG
 CGGACCTTAGTGGTGAAGTGTACACTTCTTATAGATGGTGTTCGGGAGTTTGTCTTCAGATGTGTCCAGAGTTTCTTC
 CTTCCTAATGGGTTCAATGGTCTTTGCTGACTTCAGGAATGAAGCCGACAGCTTCGCAAGTGAGGTTTACAGCTCTTAAAGGTG
 GCGTGTCCAGAGTTGTTTTGTTCCCTCTGGTGGTTCGTGCTGCTTCAGGAATGAAGCCGACAGACCTCGCAGTG
 AGTGTCTACAGCTCATAAAGGTAGTGGCGACACAGAGTGAGCTGCAGCAAGATTACTGTGAAGAGCAAAAGCAAAAGCTT
 CCACAGCATAGAAGGACACCCAGCGGGTTCTGCTGCTGGCTCAGGTGGCCAGTTATTTTCCCTTATTTGCCCTGCGCA
 CATCTCTGTGATTTGGTCCATTTTACAGAGTACTGATTTGGTCCATTTTACAGAGTGCTGATTTGGTGCATTTACAACTCTT
 GCTAGACACAGAGTGCTGATTTGCTGCACTCTACAGAGTGCTGATTTGGTGCATTTACAGTCTTTAGCTAGATACAGAAAG
 CTGATTGCTGCGTTTTTACAGAGTCTGATTTGGTGCATTTACAACTCTTTAGCTAGACAGTCTGATTTGGTGGGTTTT
 TACAGAGTGCTGATTTGGTGGGCTTTTACAGAGTGCTGATTTGGTGCATTTACAACTCTTTAGCTAGACAGAGTGCTGAT
 GGTGGGTTTTAAATCCTCTAGCTAGACAGAAAGTTTTCCAAGTCCCCACCTGACCGAGAAAGCCCTAGCTGCTTCACTCT
 CAGTGTATATCTTTGACATTTTGTCCCTCCAAAAATCTCATGTTGAAATGTAACCCCTAAATGTTGGAACGTAGGCCAGAGTG
 GATGTGGCTGGGCCATGGGGA

SEQ ID: 52

>STAR52
 CTTATGCCATCTGCGGTGCCATGTGGAACTTCGCTGAAGAAGCTAAATTTACTGACCATCTGTGCTTAGAGCGGGTTCT
 CCAAGGAAAGGCTCTGTAAA TCTGCTCC TTTGAAAATCTAGGGGAAAAACAGCTCTCTTCACTGAGGATTAATTTAAAAA
 GGGGGAATAGGAAATTTCCATGCGTTGGAAGTCCATTTAGATTCTACATGAACCATCATATATGTGTCACATACATAATCT
 TATTTTTTTATTTTAAAAAAGGATAATTTATATCCAGTGACAAAGTTTGGGAAAGGCCAAGGCAAGCAATTTAGGTTGA
 ACATTATGTAGCGTTTATATAGACCTTGCAGAGCTCTGTGCAATATCCACCCTGAACACGTGAGGTGCTACTCAAGTCTC
 TCTGGCCCTGGTAA TGTGACTCCCTTCCTTTATTTGCATGAATCGCTTGGATTGGGTGTCAGGTTTTTAAAGCTCAAGG
 TTTACGGCTATTTGTTCAACCAATCAGCATCTTACGTTGAGCTGATTGGCTTCTACTGTAGGTGTCAATCACTCAAAATTT
 TGCACTACTCTCTCAGGCCCGCGGAGCCCTGTCTAGTGGCTGTGGCAGCTGGAAGAGAAAGGAACTGGACGGAGAAAGT
 AAAAACTACTTTGCTTTGCAAGGCAAGGAAAGTATTTCTTTCTCTATTTTAAAAAATTTGATTGATATTTTAC
 CTAAATAAAATAAACATTCATTAACCAAAAAATAGCAACTATCAAGATTGTTTACTAAATTTTCTGTAATGTTTACTGTTT
 CAATAAGTAGCCAAAGGAATATTAACACAAAAATAGAA TGTGTAATTTTATGTCTATAAAGACCAATTTTAAACATA
 AAGTGAACATGGGGTTTCTAAATAAAATTTACCGTGGTAGCGTAAACACACTGCTTTCAATACTTGGGCAATGCTGAAAGTGC
 TGCACTCTAAGATAAAAAATACACCAAGGGGGGATTTCAAGAACATTAATTTGCTTTTAAATAATCTGTTATTTCTGTCA
 CTTTGGCTTTTTATTTATTTACCGTGAACCTCACAGACAGAAATTTACTTGGAGTTTCTGAAATACTGTGTTGTACATTT
 TCTCATCTTACAGTACCCACACACCCAAAAATAAAAAACAAAGAGAG

SEQ ID: 53

>STAR53
 CCCTGAGGAAGATGAAGGTAACCTCGTAAGAGAACCTTCCACTCATCCCCACATCCCTGCAGACGTGCTATTCTGTTAT
 GATAGTGGTATCCCATCTGTCTACTTCTCCCAATCATTCCTTCTTACAATTTTCTACTGTACAGCATTGAGGCTGAAC
 GATAGAGATTTCCCATGCTCTTTCTACTCTCTGCTGTATATATCCGGGATCTCTCCCTACCCAGGATGCTGTGGGGTCT
 CCACAGCCCAAGTAAGCCCTGATATGCGGGCCACACCTTTCTCTAGCTTAGGAATTGATAACCCAGGGAGGAAAGTCACTG
 TGGCATGAACAGATGGTTCACTTCGAGGAACCGTGGAAAGCGTGTGCAGGTCTTGAGATAGGGCAGAAATCGGAGTGTGCA
 GGTCTGCAAGTCAAGGAGGAGTTGAGATTGCGTTGOCACGTGGTGGGAACCTCACTGCCACTTATTTCTCTCTCTCTGTC
 CTCAGCTCAGGGATACGACATGCCCCATGATGAGAGCAGAAAGTGGTGACCTTTTACAGAAATAGGGCATGGCTGCGGA
 CCCCCTGCTCATCAGGTGCTATAGCAAGTGAAGCAAGTGTTCACACAGTGAAAGTTGAGCGTCAATTTTCTCTAGTGTGCC
 AAGAGTTGATGTTAGCGTTTACGTTGATTTTCTTACACTGTGTCATTTCTGTTAGATACATAAATTTTCAATTTGATGAGCA
 AGACATATTTAAATGATATTTTGGTTTGTATCCATGCACTTACCTTAGAAAAAAGATTTGTCGGTTTACCTCTGCAATGG
 AACAGCATTTACCTCTCTCTCTCCCAAGATGTGACTACTGAGGGCAGTTCTGAGTGTTTAAATTTTCAAGTTTTTCTCTGCA
 TTTACACACACACGACACAAAC
 CCAGTATAAGCATCTGCACTCTGCTTTTCCCATTTGCCATGCGTCTGGTCAAGCTCCCTCACTCTGTTTCTGCTGTCAGCA
 TGTACTCCCTCATTCGATTTCCCTGTAGCAGTCACTGACAGTTAATAAACCTTTTGAACAGTTCCCTCAGTTGTTGCTGCT
 TGCCATTTATTTGTGCACACAGCTGTGTGACAGTGTGTGATTTCTTTTGAAGAAAGATTTTGAAGTGGAAATGCTGTGCT
 AAAGGAGTCATTTATTTCAACAAAAACATTAATGAGTGGCTCTGCTGCTGAGGCGCTGTTCTAGTGTGAGGCGAGCTCAGG

FIG. 20, cont

GAACAAGGCAGACAGGAGTTCC TGACCCCGTTC TAGAGGAGGATGTTTCCAGTTGTTGGGTTTGTGTTGTTTCTTC
TAGAGATGGTGGTCTTGCTCTG TOCAGGCTAGAGTGCAGTGGCATGATCATAGC

SEQ ID: 54

>STAR54

CCATAAAAGTGTCTTAAACTGCAGAAAAATCCCCCTACAGTCTTACAGTTCAAGAAATTTTCAGCATGAAATGCC TGGTAG
ATTACCTGACCTTTTTCGCCAAAAAATAGGCACAGCAGCTCTCTCTGACTCTGACTTTCTATAGTCTTACTGAAATTATA
GTCTTACTGAAATTCATTCTTCAGTGTGTGAGTCTGAAAGGACACCCACATTTCTCTTTGCTCTTTGTCATTTCTTTGTT
GTAAGGCGCAGGATGTTTAAAGTTGAAGTCAATTGACTTGCAGAAATGAGAAATTTTCAGAGGGCATTGTTCTCTAGACCAT
GTAGCTTAGAGCAGTGTTCACACTGAGGTGTGCTGCTAA TGTTCTGCACTTCTTACCAA TAGTATCAATTTACCCAGCAACA
GCATATGATAGAGGACTTCGAAACCCAGAAAA TGTTTGCATATATCCAAAGCCCTTTGGGAAATGGAAGGAATTCG
GGCTCCCATTTTTATATATGGA TAGATAGAGACCAAGAAAGACCAAGGCACTCCATGTGCTTTACATTAATAAAGTACA
AAAATGTTAAACATGTAGGAAGTCTAGGCGAAGTTATGTGAGAAATCTTTACACTAATTTTGCAACATTTTAAATGCAAGTCT
GAAATTAATGTCAGAAATAAGTAAAAA TTTTACAAGTTAAGCAGAGAAATACCAATGATTAGTCAGAGAAATAAGTAGCAAAA
TCTTCTCTCAGTATTGACTTGGTTGCTTTTCAA TCTCTGAGGACACAGCAGTCTTCGCTTCGCAAAATCCCAAGTACATCT
AGTGAGGAGACTCAGCTGAGACTTTGGCTAAATGTTGGGGGGTCCCTCTCTGCTCTCCCGAGGCGAGTGAGCTTCAGGGCC
GACCTCACTCGTGGCACACAACATAATCTGGGGGAAGCAACCCGATGCCAGCATGATGCAGATATCTCAGGGTATGATCG
GCC

SEQ ID: 55

>STAR55

CCTGAACATCATGATCCGCCACCTCAGCCTCTGAAGTGTGGGATTACAGGTGTGAGCCACCACACCCAGCCGCAACACA
CTCTTGAGCAACCAATGTGTCTATAAAGAAAAATAAATGGAATTCAGAAAGTATCTTGAGACAGACAAAAATGGAACACAA
CATACCAAAATTTATGGGACACAGCAAAAGCAGTTT TAGGAGGGAGTTTATAGTGATGAATACCTACCTCAAAATCAATTA
GCCGTGATGTGATGACACTACAGTGTATAAAATGAA TTTGAAATCCACATTTGTGCCCAATACATATATACAAATTTTATTTGTT
AATTAATAAATAAATAAATTTTAAAAAGAAAGAGCTCAATTAACAACTTAACCTTATACCTCAAGGAAATAGAAG
AGCCAGCTAAGCCCAAGTTGACAGAAAGAAAAAATAATGCGCAAGAAATGAAACAGAGACTAGAAAGCAATTTGAAG
AGATCAGCAAACTA

SEQ ID: 56

>STAR56

ACACAGGAAAAGATGCAATTTGTTTCAGCAGAGCTTTGAACCGGGGATGACGGTCTCCCTCGTTGCCCGGCAACA TGGTGZA
GCAGCCAGCCAGTTATTTCTCTGGCGTAAGCAATACCAGGAAGGAAGTCTTACTGCTGTGCCCGCCGGAACAGGTTGTT
CCTGCCCTCTGAACCTTGC TGGCGCCA TGAAGCAGA TTAAGAAGCTCCAGCGCTCTGCTCGGCAAGAAACGATGGAATAA
CTCTCTAAAGAAAGCCGTTGAATATGGACGGGCAAAAGTGA TAGCGCACCGCCCTTATTTGCCCGGGATGGGGATGA
GCTTAGTCAGCGTGTGCTCTCCGGGTGTCTGCGTGGCGAGTTGCACGTCTTCTCAGACGAACCGATGACTGGATGGAATGGCC
GCCCGACTCGTCACACTGATGATACGGATGTGCTCTCCGATATACACCATGTTATCGGAGAGCTGCCAACGTATGGTTATC
GTCCGGTATGGGGCTGTCTTCGCGAGACAGGCAGAACTTGA TGGTATGCCCTGGCATCAATGCCAAACGTGTTTACCGGATCA
TGCGCCAGAAATGCGCTGTGTCTTGAGCGAAACCTGCTGTATCCGCCATCGAAACGGGACATACAGGCAGAGTGGCGGTGA
AAGAAAGCAATCAGCGATGGTGCTCTGACCGGTTGAGTTCTGCTGTGATTAACGGAGAGAGACTGCGTGTACGTTCCGGC
TGGACTGCTGTG

SEQ ID: 57

>STAR57

TOCTTCGTGTAATAGGCAAAATGTATTTTAGTTTCCACCACATGTTCTTTCTGTAGGGCTTGATGTTGGAATTTTA
TOCAATTAATTAATTAACACTATACCAACAACTGCTAAATTCGAGATGTGGCAGTGAATAAAAAGTTATAGTTTCTGA
TTTTGTGGAGCTTTGACTTTAAATGATGGACAAAAACAACATCTTAAATATATATTTCA TCAAAATTA TAGTGGGTGAAT
TATTTATATGTGCAATTTACATGTGTATGTATACATAAA TGGCGGTTACTGGCTGCACTGAGAAATGACACGTGGCGCGAA
CGAGGCTGGGCGGTGAGAAAGGCC TCCCAAGGAGGTGGCTTTGAAGCTGAGTGGTCTTCCAGCTGAAAAGGCTGGAAG
GGCATTOCAAGAAAAGGCTGAGGCCAGCCGGGAAGAGGTTCCAGTGGCGCTCTGGAAACGGAAAGCGCACCTGCC TGAACG
AAAAATGAGTGTGCTGAAATAGGACGCTAGAAAGGAGGCGAGGCTGGCAAAAGCGACCGAGGAGGAGCTCAAAAGGAGCGA
GCGGGGAAGGCCGTGTGGAGCTGGAGGAAGCACTTCGGAAGCGCTTCTGAGCGGGTAAGGCCGCTGGGAGCATGAACCTG
CTGAGCAGGTGTGTCCAGAAATTCGTGGGTTCTTGGTCTCACTGACTTCAAGAA TGAAGAGGGACCGCGGACCTCGCGGTG
AGTGTTCAGCTCTTAAAGTGGCGGCTCTGGAGTTTGTCTCTCTGATGTTCGGAATGTTTCAGAGTTTCTTCCCTTCGTT
GGGTTCTGCTGCTCGCTGGCTCAGGAGTGAAGCTGCAAGCTTCGCGTGAAGTGTACAGCTCA TAAAGCAGGGTGGGCT
CAAGAGTGAAGCAGCAGCAAGATTTATGCAAGAAATGAAGAAACAAGCTTCCACACTGTGGAAGGGGACCCAGCGGCT
TGCCACTGCTGGCTCCGAGGCTGCTTTTATTTCTCTTATCTGGCCCCACCCACATCTGCTGATTTGGTAGAGCCGAAATGGT
CTGTTTTCAGCGGCTGATTTGGTGGCTTTTACAAATCCCTGGCTAGATACAAAGGTTCTCCAGCTCCGCCACAGATAGCTA
GATAGAGTCTCCACAAAGGTTCTCCAAGGCCACCAGAGTAGCTAGATACAGAGTTTGA TTTGGTGCAATCACAAACC
CTGAGCTAGACACAGGTTGATGACTGGTGTGTTTACAAACCTTCGCGTAGATACAGAGTATCAATTTGGCGTATTTACAACT
ACTGAGCTAGGCAATAAGGTTCTCCAGGTCGCCACAGACTCAGAGGCCAGCTGGCTTCACCCAGTG

SEQ ID: 58

>STAR58

AAGTTTACCTTAGCCCTAAATTAATTTCAATTTGATTTGGCAATTTAGGAAATATGTAATAAGGAAATGCTCTTAGGAGATAA
GGATAACATATGTCTAAGAAAAATTAATTTGAAATATTAATCA TGAATAAAAATGTTAGAACTGAAAAAAATTAATTTGTA
CTCCTTCCAGCGTAGGCAGGAGTATCTAGATACCAACTTTAACAACCTCAACTTTAACAACCTTCGAACCAACAGATGGCTA

FIG. 20, cont.

GGAGATTACCACTATTAGCATGATATCTTTTATTGATAAAAAAATAAAACTTCCATTAAATTTTAAAGCTACTACAATC
CTATAAAATTTTAACTTACCAGTGTCATCAAGTCACATAATTTAAAACTATGAAATCTCTGATTTTAACTCCTCAGAT
TGAAATCTACTTATTTTTAGTACATATATCACTACTCGCGCTAGTAGAAGAGCTTGGAAATTTGAGAAAACCACT
CAGACGTTTGATATTTCTCATATTCACATAATTTATTTTTAAATGAGTTCTGCAATGCATCAAGCAGTGGCATAACAGG
AGAAAATTTAAATTTGGTGGAAGAATATGTGTGCCAAACAATCCCTGAAATTTGATGAGATGCATACTCTGAGTTATTT
GTTTCAAAATGTGTACCTGTTTATACAAGGGTATCACCTTTGAAATCTCAACATTTAAATGAAATTTTATAGCAATTTGTG
TAACATGATATATATAAAATCTGATATAACATTTTTTATACCTGTTTAGAGTTTAAAGGAGAAAAGGAGTTAAGAATA
ATTACATTTTATAGATGTGCTGGGTGCAAAAACTTCAACACTATCTGAAATCTTTTCTCAATGCTCTTGACAACT
TACCCACTTGGGTATCTATTAGCACTGCAAAATCAACATTTTGATGATGCTAATTTTTCTTCGCCATAATTTATTTGTTT
CTCAGCTTTAGCCGAATTTTTCATATGACATTGCTGCAAGTATGCTGACGCTTCAATGAGCTCTCAGAGAGCTTTTCAG
TCTTCTCTGGACTACTTGCAATGCTTCCAACAAAAAGAAGCACTCTTGATGTCAGTCACTCAAATAAATGGAAATGGGCCCA
TTTACTAGGAATGTTTATACAGAAATAAAAGATAGAGCTGACACAGTCTTCAGTCTCATCTGATTCTTCTGTTAAGGCC
TGGCCATATTTCTCAGATGTAATGTGGCGCAGGCGCATGTTAAATGGCTGCTCTCTAGGATAGTTTGACTGTTGGATT
CTCATCTTCCCTCTCTCTTAGGAAGGAAGGTTACATAGTACTGTGGCTCTTGGAAATATAGATTCAAAAGAACTAAATGG
AGTATCATCTCCCACTGCTCTTGTG

SEQ ID: 59

>STAR59

GAGATCACGCCCATCTCACTACAGCCGTGGGGTCAAGGACCAAGATCCCATCTCAGAAACAACAACAACACACAACTGACATCGGTGGGGT
 GGGATTTAATTTCGACGGTGTCAGGCTCAGGAGTCTCTTGACAGAGTACATCCACGACCCTGGGGGGAAACGTCTGATCGGTGGGGT
 GGAATCTATATTTTGTGGCTCAAGGGGAGGGTTTGTAGAGGGTAGTCCCGCAAGCGGTGATGGCTTAAGGAAGAGCCCTTCCGCCCA
 AGAAGCGGATATCTATTCTTAGCCCTTAGGCCAACCAAGAGGGAGAAATCGGGCTCTCCACAGACCCCAACACCCCAACCCCA
 CCGACCGCCACACCCCTCCACCATCTGTGAATGGGCTCTCGCTCGCTTCAGCGCTTAGTTCACACCGGTGTGGTTTGTGAAACCCTCC
 AGGCTGTGTGGCTGGGTTTCGGTGTGGTGGGGTGGGGCGCTCTGTGACAGAGGAGGGGGTAAGAAGCGCGGTGTCCCGCGGGGT
 CCGCGGAGCTGGGGCGGTGGCGGTGGGGTGGGGTGGGGCGCCAGAGCTTTGGGAATCTGTCCGCTTGTGGGAGAGCTTCCCTCTCT
 GTCCCTCTCTTGACCATAGCTTCCACATCTCCGGCTTCAGTGGGGACATCTGTGGTGGTGAAGATCACCATCACTTTTGACCTT
 AGCCGACGAAAGCGGCTCCCAAGAGTCTCTCCCGGAGCGGGGGCGCTTGGGACAGCTCAAGAAGATGCTGACGGTACAGGT
 GTGTGACCGGTGATGTACTTCGGAAGCCCTCGGGCGCAATGACAGAGTATCAATTTGACCTCGGTGGGACAGGTACGCTTTGGG
 GAGTTCGGTGGCTCTCTCAGAGAGCTATCCAGCGCTACGACAGATGGTCCGAGG

SEO ID: 60

>STAR60

[illegible]

SEO ID: 61

>STAR61

CAGCCGCCACATGCCCGACCCCTGTGCTCAGCTCTGCAGCGGGGCA TGGTGGGCAGAGACACAGAGGCCAAGGCCCTGCTTC
 GGGGACGGTGGGCCCTGGGATGAGCA TGGCCCTTGGCCCTGGCCAGAGTNCCTCTTGAAGAGGGGTCAGGAGGGGCTGCT
 CGACTCGGGGAGGAGGCCATAGCGCATCTGGCANAAGTAGAANTAGTGTGGTGCTTNCACCCAGGCCAGGGCCAGGCTGCT
 GGGTATGGACCCGGGGCCNTCTGTTCTAGACAGGAAGGATATGGTGAAGCACTCAAAGGACAGCCACTGGAGAGCTCCAG
 CGACAGGACCTTGAAGAGGCCCTGGGGGCACTCTGCTCTTTCTGGGCTGTGTGTCCTTGGGCTTGGGCCCTTCTCTCTCT
 CCGAGGGCTTGAAGAGGGCTGGCCCTGGCTTGGCTCAAGAAGACACTTAGACTGGCTACCAAGTCTGGGCCATAGGCTGCT

FIG. 20, cont

GTGGGTGCAGGCGCTGTGCGGGTGACCTGAGAGCCAGGGCTGGCAGGTGAGAGTCAGGAGAGGGAATGGCAGTGGATGCCCTG
 TGCAGGATCTGCCATAATCATGGTGAAGCTGGAGGAAATCCAAAGTGGGCAATGCACCTGTCACCTATTCTTTATTCATGTGT
 GCCCATCCCAACAGCAGGGAGCCTGGCCAGGAGGGCCCCGGGAGAAAGGCATGATGGGCTGTGTTCATTTAGGAAGGA
 TGGACGGTTGTGAGACGGGTAAATCAGAACGGGCTGCCACCTCGGCCGAGAGGGCCCCGTGGTGGGTGGCACCATCTGG
 GCTGGAGAGCTGCTCAGGAGGCTCTCTAGGGCTGGGTGACCAAGNCTGGGGTACAGTAGCCATGGGAGCAGGTGCTTACC
 TGGGGCTGTCCCTGAGCAGGGGCTGCATTGGGTGCTCTGTGAGCACACACTTCTTATTACCTGAGTCCCNCTGAGTGAT
 GAGNACACCCCTTGTCTTTCAGATGAATCTGAGCATGGAGATGTTAAGTGGCTGGCTGAGCCACACAGCAGATGGAATGGTG
 TAGCTGGGACCTGAGGGCAGGAGTCCAGCCGAGGACTTCCCAAGGTTGTGGCAACTCTGACAGCATGACCCAGGGA
 ACACCCATCTCAGCTCTGGTCAGACACTGCGGAGTTGTGTGTAACCCACACAGCTGGAGACAGCCACCTAGCCCCACCC
 TTATCCTCTCCCAAGGAACCTGCCCTTTCCCTTCAATTTCTCTTACCTGCAATTGAGGGAACACACAGTGTGGCAGAAAGGA
 ACATGGGTTCAGGACCCAGATGGACTTGTTCACAGTGCAGCCCTCCTGTCTCTTGCAGAGTGGCTCTTCCACTGTGAAG
 TTGGGACAGTACACCACTCAATACTGCTGGGCCCCGTCACAGGTGGGCAGGCAACGGATGGCAGTCACTGGCTGTGGGT
 CTGCAGAGTGG

SEQ ID: 62

>STAR62

AGTGTCAAATAGATCTACACAAAACAGATAATGTCTGCCCATTTTTCAAAAGATAATGTGGTGAAGTGGGTAGAGAGAAA
 TGCATCCATTCTCCCCACCCCAACCTCTGCTAAATGTGCCATGTACAGTACTGAGACCAAGGGGCTTATTCACAGCGGGCA
 GAATGTGCACCAAGCACTCTTGTCTCAATTTGCAGTCTAGGCCCTGCTATTTGATGGTGAAGGCTTGCACCTGGCATG
 GAAGGTCCGTTTGTACTTCTTGGCTTTAGCAGTTCAAAGAGCAGGGAGAGCTGCGAGGGGCTCTGCAGCTTCAGATGGATG
 TGCTAGAGTGTGTGAGGCGGCTTCTGTGGTCCATTAATCTCCAGCCGCCCTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT
 CTCTCCATGCCCTGTGTGGCTCCAAAATGTCTAATGTCTGCAACCCAGGAAGAAATGTCAGGCCCACTCTCTTTAATGTGTGT
 TGGGCTATTGTGATTCGCCGTGGGTATATTCCTTAGGTAAAGCCAGAAAGACAGGAGGTAGTTGCTTTGGGAGAGTTT
 GGACCTATGGGTATGAGGTAAATAGACACAGTATCTTCTCTTTCAATTTGGTGAGACTGTTAGCTCTGGCCCGGACTGAAAT
 CCACACAGCTCACTTGGGAAAACTTTATTCAAAACATAGTCAATTTGAACATTTGGAGAAATGAGGGACAGAGAAAGGGC
 CTTAGATTGTACATCTGGGTGTATGTCTATAAATAGAAATGCTTTGGTGGTCAACTAGACTTGTTCATGTGACATTTAG
 TCTTGGCTTTTCGGTGGTGAATTTAAAAATTAATGTAATCTTTGTTGGAATATAGTGGAGCTATGGGTGGCAATTTTCACT
 GGGTTTTTGTATTAGCTCAGCCGCTCCTGTATAGGCGAGCCTTGAAGCTCAGTAGCTAAATGAAGAGGTATCCTCACTCCCTC
 CAGAGAGCGGTCCCTCAAGGCTCAATTGAGACTTTGTCA

SEQ ID: 63

>STAR63

GCGTCTGAGCGCTGGGAACCCATGAGCCCCGTCCAATGGAGTTGAGGAAGGGGTTGCGCCACGGGGTGGGCGCCCTCTA
 CACAGCGCGCTTCTCTTCTTCTCTGTTAGCGCGCGGGGACCAAGCTCTGGTTCTGCACCTCGCGCTCTGGGAGCAGGCGCG
 GCTTTGGCGAGGCTTCCCGGGGCTGCCAGGCTCTGCTCGGCTCGCCCGCGCAGGCGCGGCTTCCGGAAGCCGCCAGGG
 TCCAGTCCAAGGCCCGGATTCCTCAAGGCCAGGGCCCGGGGAGCAGTATGGAACAGGGCGGGAGCCAGTCTCTCCGAGCA
 TGGAGTAATCTGAGCTTTTGAAGAAAGAAAGCGGACCCACCCCATCGAGAACCGGGCGCTTGTATTAGGAGCTTCTCG
 GCGCTCAGCGAGTGTCCCGGCTTCTCGGCGCTTCTCTCTCCAAAGCCGCCACCCCGACAGCGGCGCTTCTCTGGGAGCTC
 CCGTGGGGCTGCGCTTTCAGCCCAACACAGGAGGCTTTCAGGAGCTTCCAGTCCCAACAGCAGCCAGAGAGCCCG
 ACTCCACCTGTACTGCCAAGCTTTCAGAGAGGGGCGGCTGGACATGCCCGCGCAGGAGGAGCCCGGCTCAGACCCCG
 TGCAGTGGCAGCAACCCAGAACACCCGTCAGAGGCTCTGAGCAGCCAGGAAGTGGCTGGAAGAGCATAAGGCACTTCACT
 CTCCTCTGTAAGACGAAGGACCGGAGAACACATGCTGAACCCCTGCTTTTGCAGAGGGCGGATGCTTTCAGGACAGGCGCGCT
 CAGCAGGTGTCCATCTTATTTCACACCTTTGTGTTTATATCACTTATTTTGCATTTTATGTCATTAATTAACAAATATGCAGC
 TGGCCAGGCGCAGTGGCTCAAGCTCTAAATCCAGCACTTGGGAGGCGGAGGAGGTATCACTTGGAGGCGAGGAGTTC
 GGGACCGGCTGGGCAACATAGCAAAACCCCATTTGCTAAATAAAATACAAAAATAGCCAGCCAATGGTGGCGGGCACCTGC
 AGTCCCGCTACTCCGAGCTGAAGCAGGAGAACTCACTTGAACCCAGGAGGCGGAGGTGGCAGTGAAGCTATCAAGCCATT
 ACCTCCAGCTTGGGCAACAGAGAAAGACTGTCTCAAAAAAAATTAATAGCCAGCAGAAATTAATGTGTGTCAGGCCAAGC
 AGTCCCGCCCACTCAGGCTCTGTCCCTACAGCTCCAGGCACTCCCGCAGCCCTCCCGCTGGACAGAGGTAAATGCCAGA
 GGGTGAATAATCCACCAAGGTAAAGCCAGAAACAAAAAGCTCAAAAGCTTGGGCACTCCCTCCGCTCAGACCTTAGAGCAG
 ATTCTCTCATCGACAGCACGATCAGGCTGTGG

SEQ ID: 64

>STAR64

AGAGATCTTTTAAAGGCTCAAAGAGCCCTGCGGCTCCCTGCCAATAGCTCTGCCATCGTCCCGAGAGCTTTGGAGGACCC
 TCCACCATCGGGCCCAACCCAGCTGAGCTGGGTGCTGCTCTGAGGCTCTGCTCACTTCAGCTGAGCATGAGGCTCT
 GCTGTGCTGCTTTCAGCAGCAGGAGCAGGCTGATGAGCTGGGCTTGCAGCACTTCTCTGTGCGGAATACAAATTCAC
 AGACAGAGGATTTAAATTCAGTGGAGGTGACAGGAAAGAAAGGAAACCTCCAGGTATCAGAGAAAGGAGGGGGTGTG
 AAGACAGTATGGGAGGAGGTCAGGCTGGGGCTCAGCTCTGGGAAGTGGCAGCTGAACAGGAGTACGCGCGGGTCCCA
 TGCAAGGGAAATGAGGACCGAGGCTTGCATGTGCGAGGGCTTCCGAGGCTGCCCGCTGTGTAACAGGACACCCAGAA
 AGTCTGCCTTCCAGCTTGGCAAGTGGCAAGGAACCTCTGGGTGGGAAACAAATCAACAAACAAATTTGTCAGTAAAAAAC
 AGAAACCTTCACACTTTCTTTTGAACCTTTGAAAAAAGCAAAATCCACTGCAGCTCACCAAGGCAAGAGAAAACTT
 TAAGAAATACCAGAGAGAAAGACAGTTACTTTGCAAAAGAACATCTAATGCAAGGAGATAATGAATAACAGACTCTTCA
 AAGGGCTGAAGGAAAAAAACCGTCCACTAGAAATCTATCCCAAACTGTCACTGAGAGCAAGGGCAAAACAAAGCTTT
 CTGAGCAGGCTGGACGAGGTGCGCTCACGCCCTGTAATCTTAGCACTTTGGAGGCAAGGTGGGAGGACCGCTTTAAGCCA
 GAAGTTTGAAGACAGTGTGGTAAACATAATGAGACCCCACTCTTAAGAAAAAGAAATTAATAAGCAAGACTTTTTCAGA
 CAACAGTCTCTGAGAGCTGGCTATCTTGGCTGTCTGTAAAGAAATGCTGCGAGACACCTCAATTAGGAAAGAGACTGA
 ATCTAGAAGGAAAGAGCAGAGCATGAGGTACAAATGAGGAGCAAAATAACAGGTACCATATAAGCAAAACCAAAATACACAT

FIG. 20, cont

TCACATATACGAAACAATAAAAAAGACTCATTGGGGGGTTAAACACTGTTGAACATAAAATCCTGGATAACAGCAGCATGA
AAGGTGGGGTGGTGGTCCAGGAAAGCAATCAAGGTCCATGTCTCATTTGGGAGGAGGGTAGGGAGACTCATGAACCTGA
GGCTCCCTTCAGGCAAGCACAGTGCAAAAAAATTATAATAATGGGAAACAGATACAGTAGACTGTGATGTACAACTCTCAG
AGCAGTAGAAGGGAGGGTATAAAACAAATCTGATCCA

SEQ ID: 65

>STAR65

TCGAGACCAGCCTGGCCAACATGGTGAACATTGCTCTATTACAAATACAAAAATTAGCCAGGTGTAGTGGTGCAATGCC
GCAGTCCCAGCCATTTGGGAGGCTGAAGTTGGAGAAATCGCTTAAACCTGGGAGGTGGAGGTTCATTTGAGCCGAGAAGCAC
TCCAGCCTGGATGACGGAGCAAGACTGTCTCAAAAAGAAAAAAGAAAGCAGCAGCAAAATATCCCTGTCTCTGATGGAGG
CTATATAACCAACCAACAGTGAAATGCATAAGACAAATTCAGGTTATGGTAGATACCAATAAGTGGGAGATGAACAAATGAG
AACACATGGACACAGGAGGAGAACATCACACACTGGGGCTCTCGGGGGTGGGGAAATAGGGGGTGATAGCAATTAGGAG
AAATACCTAATGTGATAACAGGTTAGTGGGTGCAGCAACACCATTGGCACGTGTATATCTATGTAAACACACTGCACGT
TCTGCACATGTATCCAGAACTTAAAGTATAATAAAAAAGACATTAATAAATTATGATATAAAATCCCAATTCAGTTGT
TTAAAAAGAGAAAACAATTATCTTTATATAATAGCGGAAAAATAGATGGCGGAAATTAAGCCTCGTCATATTTCTAAC
AGAACTTTCTGATAAACTTGATTAAATAAAAAATTTAAATACTACTAAACACATAGAAGAAATAAAATTAACCTTCACAA
AAAAAAGTACAATGAATGAAGACAAGGTGTACTTGAAAAAGAACTGAATAAATATCTACATATAAAAAAATCTGAT
GATATTGTGGTGATTTTACTTTGCTACTAGTTTCCTTTTCTCTGAAAAATTTCTTGGGATGTATTTGGTTTCAT
TAGTAAAAATTCAGTTTCCTTTGCAATCTGAACATTTGGAGCTTCATCCATAGCCAGTATGCCCTAACATTATCTTTGGACA
ACTGTAAAAATTAGAACACTGCCAGACATATTTAATGTATGATGATATCAACACTGGGACACATTTATACATCTTTAT
CCAAATCAAAATGATTCACGTGGTTTATAAATGTACATGGATATATCTCTACCTAAGCAGATAGTTAGGAGAGTTAGTAA
AAATGAGGTGGAATAAGGAGTCACTGTCCCTTCACAGGGAGAGAAATTCGCTTTTCTCCTAATAATACCCCTTTGCTTGAAC
AGACTCCAACCCCTCATCTTTTGTCTTTAAATGACCACATTTATTTTAACTTTGATAAACCAACAGAAAGATATTTGAT
CCATCAACATTAC

ES 2 344 073 T3

LISTA DE SECUENCIAS

<110> Chromagenics B.V.
 Otte, Arie P.
 5 Arthur, Kruckeberg L.

<120> Secuencias de ADN que tienen actividad anti-represora

10 <130> 0103 EP 06 DIV

<150> EP 01202581.3
 <151> 04-07-2001

15 <150> US 60/303.199
 <151> 05-07-2001

20 <160> 122

<170> PatentIn version 3.1

25 <210> 1
 <211> 749
 <212> ADN
 <213> Secuencia Artificial

30 <220>
 <223> secuencia de STAR1

35 <400> 1

	atgcggtggg ggcgcgccag agactcgtgg gatccttggc ttggatgttt ggatctttct	60
	gàgttgccctg tgccgcgaaa gacaggtaca tttctgatta ggcctgtgaa gcctcctgga	120
40	ggaccatctc attaagacga tggatttgga gggagagtca cagaaagaac tgtggcccct	180
	ccctcactgc aaaacggaag tgattttatt ttaatgggag ttggaatatg tgagggtgc	240
	aggaaccagt ctccctcctt cttggttgga aaagctgggg ctggcctcag agacaggttt	300
45	tttggccccg ctgggctggg cagtctagtc gaccctttgt agactgtgca cacccttaga	360
	agagcaacta cccctataca ccaggctggc tcaagtgaag ggggctctgg gctccagtct	420
	ggaaaatctg gtgtcctggg gacctctggt cttgcttctc tcctcccctg cactggctct	480
50	gggtgcttat ctctgcagaa gcttctcgt agcaaaccac cattcagcgc cctgtagctg	540
	aacacagcac aaaaagccct agagatcaaa agcattagta tgggcagttg agcgggaggt	600
	gaatatttaa cgcttttggt catcaataac tcgttggtt tgacctgtct gaacaagtcg	660
55	agcaataagg tgaaatgcag gtcacagcgt ctaacaaata tgaaaatgtg tatattcacc	720
	ccggtctcca gccggcgcgc caggctccc	749

<210> 2
 <211> 883
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <223> secuencia de STAR2

ES 2 344 073 T3

<400> 2

```

gggtgcttcc tgaattcttc cctgagaagg atggtggccg gtaagggtccg tgtagggtggg      60

5      gtgcggtctc ccaggccccg gcccgtggtg gtggccgctg cccagcggcc cggcaccccc      120
      atagtccatg gcgcccagag cagcgtgggg gaggtgagtt agaccaaaga gggctggccc      180
10     ggagttgtc atgggctcca catagctgcc cccacgaag acggggcttc cctgtatgtg      240
      tggggctcca tagctgccgt tgccctgcag gccatgagcg tgcgggtcat agtcgggggt      300
      gccccctgcg cccgcccctg ccgcccgtga gcgcttctgt gggggtggcg ggggtgcgca      360
15     gctgggcagg gacgcagggt aggaggcggg gggcagcccc taggtaccct gggggggctt      420
      ggagaagggc gggggcgact ggggctcata cgggacgctg ttgaccagcg aatgcataga      480
      gttcagatag ccaccggctc cggggggcac ggggctgcga cttggagact ggccccccga      540
20     tgacgttagc atgcccttgc ccttctgatc cttttgtac ttcattgcgc gattctggaa      600
      ccagatcttg atctggcgct cagtggggtt cagcagattg gccatctcca cccggcgcgg      660
      ccggcacagg tagcggttga agtggaactc tttctccagc tccaccagct gcgcgctcgt      720
25     gtaggccgtg cgcgcgcgct tggacgaagc ctgccccggc gggctcttgt cgccagcgca      780
      gctttcgctt gcgaggacag agagaggaag agcggcgctc ggggctgccg cggccccgcc      840
      cagccccctga cccagcccgg cccctccttc caccaggccc caa                        883

```

<210> 3

<211> 2126

<212> ADN

<213> Secuencia artificial

<220>

<223> secuencia de STAR3

<400> 3

```

atctcgagta ctgaaatagg agtaaactctg aagagcaaat aagatgagcc agaaaaccat      60
gaaaagaaca gggactacca gttgattcca caaggacatt cccaagggtga gaaggccata      120
45     tacctccact acctgaacca attctctgta tgcagattta gcaaggttat aaggtagcaa      180
      aagattagac ccaagaaaat agagaacttc caatccagta aaaatcatag caaatattatt      240
      gatgataaca attgtctcca aaggaacaag gcagagtcgt gctagcagag gaagcacgtg      300
50     agctgaaaac agccaaatct gctttgtttt catgacacag gagcataaag tacacaccac      360
      caactgacct attaaggctg tggtaaaccg attcatagag agaggttcta aatacattgg      420
      tccctcacag gcaaactgca gttcgctccg aacgtagtcc ctggaaattt gatgtccagt      480
55     atagaaaagc agagcagtca aaaaatatag ataaagctga accagatggt gcctgggcaa      540
      tgttagcagc accacactta agatataacc tcaggctgtg gactccctcc ctggggagcg      600
      gtgctgccgg cggcgggcgg gctccgcaac tccccggctc tctcgccgc cctcccgttc      660
60     tcctcgggcg gcggcggggg ccgggactgc gccgctcaca gcggcggtc ttctgcgccc      720
      ggcctcggag gcagtggcgg tggcgggccat ggcctcctgc gttcgccgat gtcagcatTT      780
      cgaactgagg gtcatctcct tgggactggg tagacagtgg gtgcagccca cggagggcga      840
65     gttgaagcag ggtgggggtgt cacctcccc aggaagtcca gtgggtcagg gaactccctc      900

```

ES 2 344 073 T3

ccctagccaa gggaggccgt gagggactgt gcccggtag agactgtgcc ctgaggaaag 960
 gtgcactctg gcccagatac tacacttttc ccacggtctt caaaacccgc agaccaggag 1020
 5 attccctcgg gttcctacac caccaggacc ctgggtttca accacaaaac cgggccattt 1080
 gggcagacac ccagctagct gcaagagttg tttttttttt tatactcctg tggcacctgg 1140
 aacgccagcg agagagcacc tttcactccc ctggaaaggg ggctgaaggc agggaccttt 1200
 10 agctgcgggc taggggggtt ggggttgagt gggggagggg agagggaaaa ggcctcgtca 1260
 ttggcgtcgt ctgcagccaa taaggctacg ctctctgct gcgagtagac ccaatccttt 1320
 cctagagggtg gagggggcgg gtaggtggaa gtagagggtg cgcggtatct aggagagaga 1380
 15 aaaagggtcg gaccaatagg tgcccgaag aggcggacc agcggctgtg tgattggtat 1440
 tggcagtgga ccctccccc ggggtggtgcc ggaggggggg atgatgggtc gaggggtgtg 1500
 tttatgtgga agcgagatga ccggcaggaa cctgccccaa tgggctgcag agtggtagt 1560
 20 gagtgggtga cagacagacc cgtaggccaa cgggtggcct taagtgtctt tggctctctc 1620
 caatggagca gcggcggggc gggaccgcga ctcggttta atgagactcc attgggctgt 1680
 aatcagtgtc atgtcggatt catgtcaacg acaacaacag ggggacacaa aatggcggcg 1740
 25 gcttagtcct acccctggcg gcggcggcag cgggtggcga ggcgacggca ctctccagg 1800
 cggcagccgc agtttctcag gcagcggcag cggcccgcc aggcgcggtg gcggtggcg 1860
 gcagccagggt ctgtcaccca cccgcgcgt tcccaggggg aggagactgg gcgggagggg 1920
 30 ggaacagacg gggggggatt caggggcttg cgacgcccct cccacaggcc tctgcgcgag 1980
 ggtcaccgcg gggccgctcg gggtcaggct gcccctgagc gtgacggtag ggggcggggg 2040
 aaaggggagg agggacaggc cccgcccctc ggcagggcct ctagggcaag ggggcggggc 2100
 35 tcgaggagcg gagggggggcg gggcgg 2126

<210> 4
 40 <211> 1625
 <212> ADN
 <213> Secuencia artificial
 45 <220>
 <223> secuencia de STAR4
 <400> 4

50 gatctgagtc atgttttaag gggaggattc ttttggctgc tgagttgaga ttaggttgag 60
 ggtagtgaag gtaaaggcag tgagaccacg taggggtcat tgagtaatc caggctggag 120
 55 atgatgggtg ttcagttgga atagcagtc atgtgctgta acaacctcag ctgggaagca 180
 gtatatgtgg cgttatgacc tcagctggaa cagcaatgca tgtggtggtg taatgacccc 240
 agctgggtag ggtgcatgtg gtgtaacgac ctgagctggg tagcagtggtg tgtgatgtaa 300
 60 caacctcagc tgggtagcag tgtacttgat aaaatgttgg catactctag atttggtatg 360
 agggtagtgc cattaaattt ctccacaaat tggttgtcac gtatgagtga aaagaggaag 420
 tgatggaaga cttcagtgtc tttggcctga ataaatagaa gacgtcattt ccagttaatg 480

65

ES 2 344 073 T3

gagacagga agactaaagg taggggtgga ttcagtagag cagggtgttca gttttgaata 540
 tgatgaactc tgagagagga aaaacttttt ctaccttta gtttttgtga ctggacttaa 600
 5 gaattaaagt gacataagac agagtaacaa gacaaaaata tgcgagggtta tttaatattt 660
 ttacttgagc aggggaatct tcaaaagaaa aatgaagacc caaagaagcc attaggggtca 720
 aaagctcata tgccttttta agtagaaaat gataaatttt aacaatgtga gaagacaaag 780
 10 gtgtttgagc tgagggcaat aaattgtggg acagtgatta agaaatatat gggggaaatg 840
 aaatgataag ttatttttagt agattttattc ttcatatcta ttttggcttc aacttcaggt 900
 ctctagtgat aagaatgttc ttctcttcct ggtacagaga gagcaccttc ctcatgggaa 960
 15 attttatgac cttgctgtaa gtagaaaggg gaagatcgat ctctgtttc ccagcatcag 1020
 gatgcaaaca tttccctcca ttccagttct caaccccatg gctgggcctc atggcattcc 1080
 agcatcgcta tgagtgcacc ttctctgagc gctgcctcgg gtagctgggtg cactgctagg 1140
 20 tcagtctatg tgaccaggag ctgggcctct gggcaatgcc agttggcagc ccccatccct 1200
 ccactgctgg gggcctccta tccagaaggg cttggtgtgc agaacgatgg tgcaccatca 1260
 tcattcccca cttgccatct ttcaggggac agccagctgc tttgggcgcg gcaaaaaaca 1320
 25 cccaactcac tcctcttcag gggcctctgg tctgatgcca ccacaggaca tccttgagtg 1380
 ctgggcagtc tgaggacagg gaaggagtga tgaccacaaa acaggaatgg cagcagcagt 1440
 gacaggagga agtcaaaggc ttgtgtgtcc tggccctgct gagggctggc gagggccctg 1500
 30 ggatggcgct cagtgcctgg tcggctgcaa gaggccagcc ctctgcccag gaggggagct 1560
 ggcagtgacc aagctgcact gccctggtgg tgcatttcct gcccactct ttccttctaa 1620
 gatcc 1625

35

<210> 5

<211> 1571

40

<212> ADN

<213> Secuencia Artificial

<220>

45

<223> secuencia de STAR5

<400> 5

50

agcagagatc ttatttcccg tattcccttg tggcacagca cctcccacgc caaagcaaac 60
 caaagcaag gagcccttga tgaggagggg ccttcccca acctggtctc ccacaggtcc 120
 tacatacgta cccaccccag acacacagag ctgcttcctg ctctcacacc agactgagct 180
 55 gtgcccagac atttcccta gacttaacca actctttcaa aaatacattt ttctctaaaa 240
 agaacaagtt taaacaaagt tgactcattt taagaactgt ttagaagata accttgtgtt 300
 tattaattat gtatttgagc aaattggagg cagaagggtta ccaacattgc ctggtgtcca 360
 60 gccaggaggt agagcgtggg ggcattccaga accttcctcc aactcctgcc tggcgtgggt 420
 tttattcatc tttgtattcc caagaaactt ctcagtgtct caggagtgtt aggcactcag 480
 tacgtgtttg gtagttacat gaatgaatgc ataatgacta agtgagttaa tggatgaagc 540

65

ES 2 344 073 T3

	taattgtctc tcccttttgc ttttccagag ctttccaagg tgaaagtgtt ggacactctt	600
	tcttcatctc agatttaatc aactaagaat gctgcaaatt gaacaccagt ccacaaaact	660
5	caggaataca tgaaaagcat tgtgccttat ttttaactaa ctcaaattct atgtcagtct	720
	cccttttatg ctggatgttg gcgctaaatc tcagtgggtt cctcattctg ccagacctgt	780
	gtccagtttg ggggcttcac atagagccac cccatcacag gagagggag ggtcttgctc	840
10	ttggttgcca tcaactccacc ctcttgtctt ccgagctttg atgttcactt tccttttcac	900
	cactcggaag cttcctgcca tgatacattg agacctcaat gttaatgcca attgggggtt	960
	ggggttctca taaactcaga agtccaggaa aatcgctgc tgcctccac aacactctga	1020
15	gggcattctg gaatcctacc acttacctgg agcctgctgg cctcaactgt tttgaagtct	1080
	gtgtctgggc catgcaggta aatgggagga tgttctgtgg ccataaaaat acccgaagtc	1140
	ccacctaaag ttgatgcagg gtcttctgca tttcattgca aaattgttct atcatttcta	1200
20	tagttttcag cctacagtca ggggccagga ctttgcaccc ttggtaaacc tcaatctctt	1260
	ctccttctg gcttctactc ctttctccct caatcccaa tcaaggccct tgattgtctg	1320
	gaggtaggaa agcctggttc tggctcatga tatagtctac atcatagcct ttgtcatctc	1380
25	atggattcac tcaacaaccg tgtgtggatg gggccacca atatgtgcca ggagttgagg	1440
	acacgcaggg ttatgatgat gaaatagata aggggccac actcacggac cctgcaggac	1500
	agtggagctg tggacccagc atgcgagtaa agaccagtg agctcaccag acagatcatt	1560
30	taaatcaggt g	1571

<210> 6
 35 <211> 1173
 <212> ADN
 <213> Secuencia Artificial
 40 <220>
 <223> secuencia de STAR6
 <400> 6

45	tgaccacca cagacatccc ctctggcctc ctgagtgggt tcttcagcac agcttccaga	60
	gccaaattaa acgttctact tatgtctata gacaaaaagg gttttgacta aactctgtgt	120
50	tttagagagg gagttaaatg ctgttaactt tttaggggtg ggcgagaggg atgacaaata	180
	acaacttgtc tgaatgtttt acatttctcc cactgcctc aagaagggtc acaacgaggt	240
	catccatgat aaggagtaag acctcccagc cggactgtcc ctcgcccccc agaggacact	300
55	ccacagagat atgctaactg gacttgagga ctggctcaca ctccagagaa aagcatggag	360
	cacgagcgca cagagcaggg ccaagggtccc agggacagaa tgtctaggag ggagattggg	420
	gtgagggtaa tctgatgcaa ttactgtggc agctcaacat tcaagggagg ggaagaaag	480
60	aaacagtccc tgtcaagtaa gttgtgcagc agagatggta agctccaaa tttgaaactt	540
	tggctgctgg aaagttttag ggggcagaga taagaagaca taagagactt tgaggggtta	600
	ctacacacta gacgctctat gcatttattt atttattatc tcttatttat tactttgtat	660

65

ES 2 344 073 T3

aactcttata ataatcttat gaaaacggaa accctcatat acccatttta cagatgagaa 720
 aagtgacaat tttgagagca tagctaagaa tagctagtaa gtaaaggagc tgggacctaa 780
 5 accaaaccct atctcaccag agtacacact cttttttttt ttccagtgtg atttttttta 840
 atttttattt tactttaagt tctgggatac atgtgcagaa ggtatggttt gttacatagg 900
 tatatgtgtg ccatagtggg ttgctgcacc tatcaaccg tcacttaggt ttaagcccca 960
 10 catgcattag ctatttgtcc tgatgctctc cctccccctc ccacaccaga caggccttgg 1020
 tgtgtgatgt tccccctcct gtgtccatgt gttctcactg ttcagctccc acttatgagt 1080
 gagaacgtgt ggtatttggg tttctgttcc tgtgttagtt tgctgaggat gatggcttcc 1140
 15 agcttcatcc atgtccctgc aaaggacacg atc 1173

<210> 7

20 <211> 2101

<212> ADN

<213> Secuencia artificial

25 <220>

<223> secuencia de STAR7

<400> 7

30

atcatgccag cttaggcgac agagtggagac tggacataat aacaataata ataaaaataa 60
 ataaataaaa caattatctg agaggaaaaa tttgattcat aataaagaga ataaaggttt 120
 35 ttggcgtgtt tgttttgttt tcacctaaga acagctgttc ccctcattgg gttagtttta 180
 tttgcaagca gaaatcatct ccgcatgatt tccaggggtg tggaaaactg aatatgaatc 240
 caccttctgc catctattca cttgtcacat ttaataagac actcatgcct attttagcat 300
 40 gttttcttcc ctaccaaag agttagtaac atcaagagat taaaataaca caaataagaa 360
 cattgaagggt attcaaagt tacatacaaa tattaacac aatattatta taattattcc 420
 tggaaatgac attgcctcta ctctcaagggt aaagggtcatt tttcttgatt taaacttttt 480
 45 tctcaagttt gaaatctcta agtttcaacc cgtaatctat ttgcaagttt gtgcaaattt 540
 tagggattga atccatagta attagtgtt tattgtgggt tagggagaca agtcaaaaga 600
 atcaggactg ctaggttagat gactaaggaa aggatgggtc acgagggtgac ataaagcact 660
 50 cagaagaaaa aggtcaggaa acggaggaca gaaaaaaacc taagttctgc tgggtgatgc 720
 tgaatttctc atcacaaaat ctgcattgtg gaagcttttag ctattgagga gattgctcaa 780
 gtgtagaact gagaacaata ggcagtgaac ccgagagAAC atcaagagac tgagagaaaa 840
 55 tgaaccagac ttccagggtc tccatgttcc aaccaacatt ttgtattgtc agaaggaatt 900
 gagaggcaaa aggaaaccca ataaaaata aaacaggaaa gggcatacat gattaccacc 960
 ccttttctca ccagctgtc atggaccagc tttctcctag tgctattttt ttgggtcactg 1020
 60 catcactctg ctaacatagt tccccacta gctctgaggc tgtcccagag ggggaagccag 1080
 ctgtcatctc cttcttccac actctgttgg aggaactctg cattagcagc tcctactaa 1140
 acgcatttat gacaaacagg caggagataa ttaactagaa agtgaacaaa ctcaaacttc 1200

65

ES 2 344 073 T3

agagcctctc atttgtatga atgcccttgt aaggctcttg gcctatttta atatttataa 1260
 atgtgttatt ttcttctaaa gaaaaccacc aaattgtata agctacagaa tctgcaaac 1320
 5 tgagggtccat ccatgcactc aggatacatt catagcatct ctgagctgga aaatatctta 1380
 aaggatcatat atgtcctcca acactgcaag aatctctctg gcagcattct tttaaaatca 1440
 tcatctaaaa gagggaaatc cccagctgtg tttggatttt gctctgtcac ttgtccagtt 1500
 10 tccccatcca taaaagggca acaatatgaa tttcctgata aggtagttgt taatataaat 1560
 acaaagtgcg tagccacttc cctaagaaaa atatggggtt tctgcttcac agtctagggg 1620
 gaggaaaaaa aaggggggtc agaagtgatt attattatca ttctatattg gaatgttttc 1680
 15 agacataaaa agctcaccac gtcttaggcc agacagatgc attatgaaag ttaagctaag 1740
 tcttctcat catgagctgc acctatatcc ccattacttc ttctagaact gcataattta 1800
 tttattcttt cttaaaagt ttgagagagc cattcttgct ctctaagatt tttttttttt 1860
 20 tttttggaga cagagctctc gtctgttgcc caggctggag tgcaatggca ctatctcagc 1920
 tctctgcaac ctctgcctcc cagattcaag tgattctcct gcctcagcct cccgagtagc 1980
 tgggattaca agcacgcacc accacaacca gctaattttt cgtatttttt agtagagacg 2040
 25 aggttttacc atgttggtcca ggctgggtct gaactcctga cctcgggtga tccaccacc 2100
 t 2101

30 <210> 8
 <211> 1821
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <223> secuencia de STAR8
 40
 <400> 8

gagatcacct cgaagagagt ctaacgtccg taggaacgct ctcgggttca caaggattga 60
 45 ccgaacccca ggatacgtcg ctctccatct gaggtctgct ccaaattggcc ctccactatt 120
 ccaggcacgt ggggtgtctcc cctaactctc cctgctctcc tgagcccatg ctgcctatca 180
 cccatcgggtg caggctcttt ctgaagagct cgggtggatt ctctccatcc cacttccttt 240
 50 cccaagaaag aagccaccgt tccaagacac ccaatgggac attccccctc cacctccttc 300
 tccaagtgtg cccaggtggt catcacaggt tagggagaga agccccagg tttcagttac 360
 aaggcatagg acgctggcat gaacacacac acacacacac acacacacac acacacacac 420
 55 acacgactcg aagaggtagc cacaaggggtc attaaacact tgacgactgt tttccaaaaa 480
 cgtggatgca gttcatccac gccaaagcca aggggtgcaaa gcaaacacgg aatggtggag 540
 agattccaga ggctcaccaa accctctcag gaatattttc ctgaccctgg gggcagaggt 600
 60 tggaacatt gaggaattt cttgggacac acggagaagc tgaccgacca ggcattttcc 660
 tttccactgc aaatgacctg tggcgggggc atttcacttt cccctgcaaa tcacctatgg 720
 cgagggtacct cccaagccc ccaccccccac ttccgcgaat cggcatgggt cggcctctat 780

65

ES 2 344 073 T3

	ccgggtgtca ctccaggtag gcttctcaac gctctcggct caaagaagga caatcacagg	840
	tccaagccca aagcccacac ctcttccttt tgttataccc acagaagtta gagaaaacgc	900
5	cacactttga gacaaattaa gagtccttta ttttaagccgg cggccaaaga gatggctaac	960
	gctcaaaatt ctctgggccc cgaggaaggg gcttgactaa cttctataacc ttggtttagg	1020
	aaggggaggg gaactcaaat gcggttaattc tacagaagta aaaacatgca ggaatcaaaa	1080
10	gaagcaaattg gttatagaga gataaacagt tttaaaaggc aaatgggttac aaaaggcaac	1140
	ggtaccaggt gcggggctct aaatccttca tgacacttag atataggtgc tatgctggac	1200
	acgaactcaa ggctttatgt tggtatctct tcgagaaaaa tcctgggaac ttcatgcact	1260
15	gtttgtgcca gtatcttatc agttgattgg gctcccttga aatgctgagt atctgcttac	1320
	acaggtcaac tccttgcgga aggggggttg gtaaggagcc cttcgtgtct cgtaaattaa	1380
	ggggtcgatt ggagtttgtc cagcattccc agctacagag agccttattt acatgagaag	1440
20	caaggctagg tgattaaaga gaccaacagg gaagattcaa agtagcgact tagagtaaaa	1500
	acaagggttag gcatttcact ttcccagaga acgcgcaaac attcaatggg agagaggtcc	1560
	cgagtcgtca aagtcaccaga tgtggcgagc ccccgaggag aaaaaccgtg tcttccttag	1620
25	gatgcccggg acaagagcta ggcttccgga gctaggcagc catctatgtc cgtgagccgg	1680
	cgggagggag accgccggga ggcgaagtgg ggcggggcca tccttcttct tgctctgctg	1740
	ctgccgggga gctcctggct ggcgtccaag cggcaggagg ccgccgtcct gcagggcgcc	1800
30	gtagagtttg cgggtgcagag t	1821

35 <210> 9
 <211> 1929
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> secuencia de STAR9

45

50

55

60

65

ES 2 344 073 T3

<400> 9

5	atgagccccc aaaaatgatc ctctggctta tgacaacctg atgcagccca ggaaatgcct	60
	gcaacatgcc cactagcagc tgggaacccc tctgtgagga agagaacggt ttacattaag	120
	aaaccctttg ttttgcagca gagactattc aggtcacaca tgtgtggcct ctcagttctt	180
	tgagccattt gaagttctct atccttgctg ggaggctgag ctctccatgg aaacctggtc	240
10	cgatagttag aggagcagac cctctggaac caccttttta cacctgacca aagcagccag	300
	tcattgggcca gtgatgcaac aagggtcaacc ggtgcattct ggcccctcag aaaagcagcc	360
	cccgggaagg tcaggaggag gctgctgact ccctcttccc ctgcagccgc cccaagcaca	420
15	cccaggagcc ctgcagggtt ggggtcacca ggtgccagca ggtcccacga tgctgcattt	480
	cttacgagct cctggaggat gcagatggc ctggtcagag gctgcattct gaggatcagg	540
	agccatgggg caacgtttct gcgattgagg aaggggcatt tctgggtgg gcagaacaaa	600
20	ggtctttggc tgagctggag catccgcctc catcagtgtt ttccggcaac tgtactatcc	660
	atcgtcttcc cttcccacag ctgacctagg ctttggaaaa tgctctgaaa ctttcttttc	720
25	agaagagttg actcccaact ccacacttag ggggaagtcaa gcctacttct cagaattcag	780
	agaaggcata aaaaagaatt catttctaaa ggccctttag aagtaacttc aggtctgaca	840
	gcggccagct aatttctggt cgccttcag gaattctctg actgcaaaaa aaaagcattt	900
30	accacctgaa cacaaccca gttacagata gaaaaacata gtcatttaaa tagaatataa	960
	gcatctggcc tctgcccatac ataattggagt aacacaaaaa tctattttca aaaggaaact	1020
	aaatattatt gacaaaaaca tgaatgggga gacctcaggg tgatacagct cttgcctgga	1080
35	tggaatttgt aatcaagagg atgagacagg attgtaactt gtgccaatgt gaaaggggtt	1140
	gctcagggtat cattcatttt gcttaaatgc atgggtaatt tccaaagttc tttggagctg	1200
	aatttcacaa tttagtgcag gtcctggtga gccacacttg acttatctca cagtacaatg	1260
40	cagtggcgtg gctacaatgc tgggcaagag aagccaatgt caacagccca ggagtggctg	1320
	ggtccttacc aggtcccaag gcatgcttca tgggtggccc tgggtctgga ggaacagcac	1380
	ctttgcctgg tccatgagta tctgggtcaa actctcctgt ggacacagaa ggccatggcg	1440
45	acaggcattc ccaggaaaag aaaagggcag cagctgaaat cgtcagggtg agaaggcagt	1500
	catccttgct cagtcaactc taatccggct gcctcctcct cagcttcagg gtgaacctct	1560
	cctaagctgt gtctttggta tctgatgggc attaggtgct ggtgaaaaag ctggaggggtc	1620
50	ctttgggata ttacagaagc ccaatctagc cttgtattca atatctaggc actctcacc	1680
	ctgaagttct acgtttccag atttctgaaa acatgggaaa gcatgtgtgt gatgtctgag	1740
	gtccccctca gcctctggtg taggggttag agggctctaa aggggtggcag ctccagtgtc	1800
55	ccagtggggc ctgaagttgg tcccttcctc tcccagctcc catccatggt ttagcccaat	1860
	cccttcctga cctaagagta ctgcacatgg atgctccacg cagagcctct gctccactcc	1920
	caggaagtg	1929

60

<210> 10

<211> 1167

65 <212> ADN

<213> Secuencia artificial

ES 2 344 073 T3

<220>

<223> secuencia de STAR10

5 <220>

<221> característica_misc

<222> (452)..(1143)

<223> "N" representa cualquier ácido nucleico en diferentes posiciones

10

<400> 10

```

15      aggtcaggag ttcaagacca gcctggccaa catggtgaaa ccctgtccct acaaaaaata      60
      caaaaattag ccgggcgtgg tggggggcgc ctataatccc agctactcag gatgctgaga      120
      caggagaatt gtttgaaccc gggagggtgga ggttgacagt aactgagatc gcgccactgc      180
20      actccagcct ggtgacagag agagactccg tctcaacaac agacaaacaa acaaacaaac      240

      aacaacaaaa atgtttactg acagctttat tgagataaaa ttcacatgcc ataaagggtca      300
      ctttctacag tatacaattc agtggattta gtatgttcac aaagttgtac gttgttcacc      360
25      atctactcca gaacatttac atcaccccta aaagaagctc tttagcagtc acttctcatt      420
      ctccccagcc cctgccaaac acgaatctac tntctgtctc tattctgaat atttcatata      480
      aaggagtcct atcatatggg ccttttacgt ctaccttctt tcacttagca tcatgttttt      540
30      aagattcatc cacagtgtag cacgtgtcag ttaattcatt tcattctatg gctggataat      600
      gctctattgt atgcatatcc ctcaatttgc ttatccattc atcaactgat tgacatttgg      660
      gttatttcta ctttttgact attatgagta atgctgctat gaacattcct gtaccaatcg      720
35      ttacgtggac atatgctttc aattctcctg agtatgtaac tagggttgga gttgctgggt      780
      catatgttaa ctcagtgttt cttttttttg aagaactacc aaatggtttt ccaaagtgga      840
      tgcaacactt tacattccca ccagcaagat atgaagggtc caatgtctct acatttttgc      900
40      caacacttgt gattttcttt tttttattta tttatttatt ttttttgag atggagtctc      960
      actctgtcac ccaggctgga gtgcagtggc acaatttcag ctcaactgcaa tctccacctc      1020
      tcgggctcaa gcgatactcc tgccctcaacc tcccagagtaa ctgggattac aggcgcccac      1080
45      caccacacca agctaatttt ttgtattttt agtagagacg gggtttcacg atgtcgacca      1140
      ggntgtactc gaactctgac ctcaagt                                1167

```

50 <210> 11

<211> 1377

<212> ADN

55 <213> Secuencia artificial

<220>

<223> secuencia de STAR11

60

65

ES 2 344 073 T3

<400> 11

5	gattctgggt gggtttgatg atctgagagt cccttgaata aaaagaattc tagaaaagct	60
	gtgaaacttc acctttcccc tattcttaac cttacttgcc tttgggaggc tgaggcagga	120
	ggatgactta aggccaggag tttgagaatg tagtgagcta tgaccacacc ggttacactc	180
	aagcctgggc gagaccacaa caaaaacctt acctgccaac tgctccatgc tggaaattta	240
10	tttcgtttct tggattgtgg aaagaactgg cttactgaaa accacacttc tctaaaaccc	300
	ttcttccagt taggtgttaa gattttaaca gcctttccta tctgaataaa aactgcacac	360
	aaagtaaaact taagagatgt caacaactca tctgtttgtt acaagatgag tctccatgct	420
15	tcatcgctg tggggaatcc tcatcagcgt ctagtggaac agactcctgt gtgctcaccg	480
	aaacgctccc cttcctccag ggcacacagt cacatggatt tcccatgcac cctggcagct	540
	cagcaggagt ccatgactta agaaggccaa tggactgtgg gtgaagtctg tggacgggga	600
20	agccacatgc gtcacttcca ggcctgggag tgtgcatcct ccaactctctt cccctgtggg	660
	tgcagaaggc ggggcagagg gccctgaaac cttggagggtc ggtggagccc aaaatgaagg	720
	agcgtgggcc tctgggtctt catgtaaatt taggtaacac tgaactgtca ggtgaacaag	780
25		
	aaataaacgt caaatgtatt cagtcgatta gatttggtga tggttgttac agcggttacc	840
	ctccctcaac ataataaatt ttcaaacaac tcataatggc tcaactcatgt ataaaaatatt	900
30	ccatagtaaa tcccgggata acatgcttat tctagctcaa gcttaatcag agtagtccat	960
	ctgagggagg agatagtaga gggcagcaag gggttgtcac tgaagataac tagccttgct	1020
	aaaagaatgg ttgaagaagt gagctacaga tagggtaaat ccacatctca gacattctgt	1080
35	gatggctctg atattatcct aaagtaaaat gtagagttga accattttta ttagattcta	1140
	gaattctatt aatttataag atgggcattt ccacaaagga ctaaacaag tacaagagga	1200
	ttaaataatc atccacatgg gaggcaccgc cttgcacttt aaaatgatgg agcttatcaa	1260
40	gactggctgt ggatatctgt ccctgggagg gttttttccc ccattttttt cttttttgag	1320
	acatgtttctc gctatgttgc ccaggctggt cttgaactcc tgggctcaag tgatcct	1377

45 <210> 12

<211> 1051

<212> ADN

50 <213> Secuencia artificial

<220>

<223> secuencia de STAR12

55

60

65

ES 2 344 073 T3

<400> 12

5	atcctgcttc tgggaagaga gtggcctccc ttgtgcaggt gactttggca ggaccagcag	60
	aaacccaggt ttcctgtcag gaggaagtgc tcagcttata tctgtgaagg gtcgtgataa	120
	ggcacgagga ggcaggggct tgccaggatg ttgcctttct gtgccatatg ggacatctca	180
	gcttacgttg ttaagaaata tttggcaaga agatgcacac agaatttctg taacgaatag	240
10	gatggagttt taagggttac tacgaaaaaa agaaaactac tggagaagag ggaagccaaa	300
	caccaccaag tttgaaatcg attttattgg acgaatgtct cactttaaat ttaaattggag	360
	tccaacttcc ttttctcacc cagacgtcga gaagggtggca ttcaaaatgt ttacacttgt	420
15	ttcatctgcc tttttgctaa gtcctgggtcc cctacctcct ttccctcact tcacatttgt	480
	cgtttcatcg cacacatatg ctcatcttta tatttacata tatataattt ttatatatgg	540
	cttgtgaaat atgccagacg agggatgaaa tagtcctgaa aacagctgga aaattatgca	600
20	acagtgggga gattgggcac atgtacattc tgtactgcaa agttgcacaa cagaccaagt	660
	ttgtttataag tgaggctggg tgggtttttat tttttctcta ggacaacagc ttgcctgggtg	720
	gagtaggcct cctgcagaag gcattttctt aggagcctca acttcccca gaagaggaga	780
25	ggcgagact ggagttgtgc tggcagcaca gagacaaggg ggcacggcag gactgcagcc	840
	tgcagagggg ctggagaagc ggaggctggc acccagtggc cagcgaggcc caggtccaag	900
	tccagcgagg tcgagggtcta gagtacagca aggccaaagg ccaagggtcag tgagtctaag	960
30	gtccatgggc agtgaggctg agaccaggg tccaatgagg ccaagggtcca gagtccagta	1020
	aggccgagat ccagggtcca gggagggtcaa g	1051

35 <210> 13

<211> 1291

<212> ADN

40 <213> Secuencia Artificial

<220>

<223> secuencia de STAR13

45

50

55

60

65

ES 2 344 073 T3

<400> 13

5	ctgccctgat cccttaatgc ttttgGCCCA gagcaccCG ctaagtccaa cccagaggg	60
	gcctcatccg caaagcctcg ggaagaggac agtgacggag gcggctgccc tgtgagctgc	120
	acggggcaga atgtcctttt ggcgtcatgt tggatgtcca cacatccata tggggtcagt	180
10	tctattagga ttccttcggg aagaggtaga gggtaggagg ggttaagcca cgagacgagg	240
	catgcagagg ggtggccttg atgggtctgc actgctgtcc atgcacacgg ggagcgttgc	300
	aaattgtgct tcccagccca tagtgccccc acagaggagc ccgggagtcc ctggtgggcg	360
15	tctgtgttcc tgcaaggagc cagtggagat ggccccgtga actctcatcc cccttgccct	420
	ggtgggggtct ctggcagggt tatggagccg tacatctttg ggagccgcct ggaccacgac	480
	atcatcgacc tggaacagac agccacgcac ctccagctgg ccttgaactt caccgccac	540
20	atggcctacc gcaagggcat catcttggtt ataagccgca accggcagtt ctcgtaacctg	600
	attgagaaca tggcccgtga ctgtggcgag tacgccaca ctcgctactt caggggcggc	660
	atgctgacca acgcgcgcct cctctttggc cccacgggtc gcctgccgga cctcatcatc	720
25	ttcctgcaca cgctcaacaa catctttgag ccacacgtgg ccgtgagaga cgcagccaag	780
	atgaacatcc ccacagtggg catcgtggac accaactgca acccctgcct catcacctac	840
	cctgtacccg gcaatgacga ctctccgtg gctgtgcacc tctactgcag gctcttcag	900
	acggccatca cccgggccaa ggagaagcgg cagcaggtg aggctctcta tcgcctgcag	960
30	ggccagaagg agcccgggga ccaggggcca gcccaccctc ctggggctga catgagccat	1020
	tccctgtgat gttcactctc ctcccaaagc aaaccacagc caagcctgtc tgagctggga	1080
	gtcccccttc ccagccctgg gtcagcggca tcctcagtcg ttgttactta ctcagctgat	1140
35	gtcacagtgc agacatccac cgttccacca cagaaccagt ggctgagcgg accaacgttg	1200
	ccatgtgcgt ttgctctgtg gggaaacagag cacagagggg gagcgacatg tgcagaacgg	1260
40	cccccttggt gcagttagga cctcagtggtc t	1291

<210> 14

<211> 711

45 <212> ADN

<213> Secuencia artificial

<220>

50 <223> secuencia de STAR14

<400> 14

55	agcaaggacc agggctctgc ctccccagtc agcatgagca gagcagactc ctttgagcag	60
	agcatcaggg cagaaataga acagtttctg aatgagaaaa gacagcatga gacccaaaaa	120
60	tgtgatgggt cagtggagaa gaaaccagac acacatgaaa attcggcgaa gtcactctcg	180

65

ES 2 344 073 T3

5 aaatcccacc aagagccggc tacaagggtg gtgcaccggc agggcctgat gggcgctccag 240
 aaggagttcg ctttctgcag acctcccccg gttagcaaaag acaaactgtc agcccagaag 300
 cctcaggtcc aaggtcacga ccacgaccac gcaggagaag gagggcagca caaagccagc 360
 aacccccacc gcccttcaga agcagtacag aataaaagtg ggattaaaag gaacgccagc 420
 accgcaagga ggggaaagcg agtcacgagc gccgtacagg cggccgaggc gtccgactcc 480
 10 agcagcgacg acggcattga ggaggccatc cagctgtacc aggtgcagaa aacacacaag 540
 gaggccgacg gggacccgcc ccagaggggtc cagctccaag aggaaagagc acctgcccct 600
 cccgcacaca gcacaagcag cgccacaaaa agtgccttgc cagagaccca caggaaaaca 660
 15 cccagcaaga agaagccagt gcccaccaag accacggacc ctggtccagg g 711

<210> 15

20 <211> 1876

<212> ADN

<213> Secuencia artificial

25 <220>

<223> secuencia de STAR16

<400> 15

30

 cagtacatgc agaactgagt ccaaacgaga cggacagcaa acccggcagt gggctcccag 60
 acattcctgg gggaaagggg tcctaaccac aggcagttaa agtcattctc tccaaccctc 120
 35 tatgacacag gctgtgctgct gtcattttaa agctgagtga aatttaaccc ttttcccatt 180
 tagaaaaaca aagcgcagct ggctgccagc actcatttaa ttttacataa acgtgctctt 240
 tgaggctgaa gcaaactctga ctgattttca atgtgaaaat aaaatgtaa aactgttctt 300
 40 ggaattatct ctaaacagaa catcagaatc gtctgaatca tcagaatcgg ctattttgga 360
 aaaaatcgat tcatcaaacg aatcttcggc caacaactgt tagagaacga tgtaaacacc 420
 acgcatagga atgttacatt ttctagaatt tgacattttc attgacggaa aattactgta 480
 45 tcttgatatat ggaaatacca ctactaaaaa cataatgcta taaatagaat gatgtctttt 540
 gtttccaaag tcaatatact cgagcaatgc aaaaataata ataaaagtga gatacttcat 600
 ggcaaagctg ccgcaggata aacattgcag ccacaagtgc cccaggtatt ctcggggcaa 660
 50 actggaaaag ggctaacagg caacattttc atgttattct actgagtgc gtaattatct 720
 ttaaaaaatat acatgaataa tgaaaaaact gtggtatggt tttaaagaaa tttccataac 780
 ctggtgaaac tcttcacaca gggtaatagg ttcataaagc cttggtcctc tgcaaaacaa 840
 55 gcatcaactt gacaatgact aaaagaagca acagcaaac tgtcacgcat ttggagccat 900
 ggcctgggtt gggccgggtg aaagctctcc gccctctgga gcaagtctgg gccccagcgg 960
 ctggcatgtg ggcaactgcag ggcctgggtt gggcaggtgt gcagctctcc gtcattctgag 1020
 60 cctagtctga ggcctgggtg ctggcacgtg ggcctctcag ggcctctact tctcacccca 1080
 gtccacttc cctccctgcc ctactgggt ctacagagc caatgaacac tggggtcaga 1140
 ttcaggggcc agcatccact gcagtgggca ctgcccttcc acaaggcctg gctccaggaa 1200

65

ES 2 344 073 T3

	gcaacccccca cctcagccac acagtagggc aacaggaaat cccattcccc catgccagt	1260
	actacaccag ggaaggggct cacgtgaggc tggccccagg cctgctgtga gaccgcgttg	1320
5	tctatgagct tggatttaag gaacttggga gcaagaagct ttctttcatt acggggccacc	1380
	agcagggaaa aaagttagcc caacgcagtt gacagtcaca cccccaccag gaccccagg	1440
	cacagaagga gggaagagga caacagagga tgaggtggg ccagcagagg gacagagaag	1500
10	agctgcctgc cctggaacag gcagaaagca tcccacgtgc aagaaaaagt aggccagcta	1560
	gacttaaaat cagaactacc gctcatcaaa agatagtgtg acatttggg tgctataatt	1620
	ttaacatgtc ccccaaaagg catgtgttg aaatttaatc cccaacaaac cagggctggg	1680
15	aggtggagcc tcatgagagg tggtagggcc atgaggggtg agtgaatgga tgaatgccat	1740
	tgtctcggga atgggcctct tctacaagga cgagttcagc cccctttct cttgtccacc	1800
	ctctctttgc cctttcgcta gggagtgcg taacaagaag gccctcaca gatgctggca	1860
20	ccttgatctt ggactc	1876

<210> 16
 25 <211> 1282
 <212> ADN
 <213> Secuencia artificial
 30 <220>
 <223> secuencia de STAR17
 35 <400> 16

	cgcccacctc ggctttccaa agtctgagg ttacaggcat gagtcactgc gcccatcctg	60
	attccaagtc tttagataat aacttaactt tttcgaccaa ttgccaatca ggcaatcttt	120
40	gaatctgcct atgacctagg acatccctct ccctacaagt tgccccgcgt ttccagacca	180
	aaccaatgta catcttcat gtattgattg aagttttaca tctccctaaa acatataaaa	240
	ccaagctata gtctgaccac ctccagcacg tgttctcagg acctccctgg ggctatggca	300
45	tgggtcctgg tcctcagatt tggctcagaa taaatctctt caaatatttt ccagaatttt	360
	actcttttca tcaccattac ctatcaccca taagtcagag ttttccaca ccccttcctc	420
	agattcagta atttgctaga atggccacca aactcaggaa agtattttac ttacaattac	480
50	caatttatta tgaagaactc aaatcaggaa tagccaaatg gaagaggcat agggaaagg	540
	atggaggaag gggcacaag cttccatgcc ctgtgtgcac accaccctct cagcatcttc	600
	atgtgttcac caactcagaa gctcttcaaa ctttgtcatt taggggtttt tatggcagtt	660
55	ccactatgta ggcatggttg ataaatcact ggtcatcgg gatagaactc tgtctccagc	720
	tcctctctct ctccctccca gaagtcctga ggtggggctg aaagtttcac aaggtagtt	780
	gctctgacaa ccagccccta tcctgaagct attgaggggt ccccaaaaag ttaccttagt	840
60	atggttgga gaggttatt atgaataaca aaagatgctc ctatttttac cactagggag	900
	catatccaag tcttgcgga acaagcatg ttaactgtag caaattcata caggtagata	960
65	gcaatctcaa ttctgcctt ctccagaaga agaatttgac caagggggca taaggcagag	1020

ES 2 344 073 T3

	tgagggacca agataagttt tagagcagga gtgaaagttt attaaaaagt tttaggcagg	1080
	aatgaaagaa agtaaagtac atttggaaga gggccaagtg ggcgacatga gagagtcaaa	1140
	caccatgccc tgtttgatgt ttggcttggg gtcttatatg atgacatgct tctgaggggt	1200
5	gcatccttct cccctgattc ttccttggg gtgggctgtc cgcattgcaca atggcctgcc	1260
	agcagtaggg aggggccgca tg	1282

<210> 17

10 <211> 793

<212> ADN

<213> Secuencia Artificial

15 <220>

<223> secuencia de STAR17

<400> 17

20	atccgagggg aggaggagaa gaggaaggcg agcagggcgc cggagcccga ggtgtctgcg	60
	agaactgttt taaatggttg gcttgaaaat gtcactagt gtaagtggct tttcggattg	120
	tcttatttat tactttgtca ggtttcctta aggagagggg gtgttggggg tgggggagga	180
25	ggtggactgg ggaaacctct gcgtttctcc tcctcggctg cacagggatga gtaggaaacg	240
	cctcgtgcc acttaacaat ccctctatta gtaaattctac gcggagactc tatgggaagc	300
	cgagaaccag tgtcttcttc cagggcagaa gtcacctgtt gggaacggcc cccgggtccc	360
30	cctgctgggc tttccggctc ttctaggcgg cctgatttct cctcagccct ccacccagcg	420
	tccctcaggg acttttcaca cctccccacc ccattttcca ctacagtctc ccagggcaca	480
	gcacttcatt gacagccaca cgagccttct cgttctcttc tcctctgttc cttctctttc	540
	tcttctcttc tgttcttct ctttctctgt cataatttcc ttggtgcttt cgccacctta	600
35	aacaaaaaag agaaaaaat aaaataaaaa aaaccattc tgagccaaag tattttaaga	660
	tgaatccaag aaagcgacc acatagccct cccacccac ggagtgcgcc aagacgcacc	720
	caggctccat cacagggccg agagcagcgc cactctgggc gtacttttgg gtcaagagat	780
40	cttgcaaaag agg	793

<210> 18

<211> 492

45 <212> ADN

<213> Secuencia Artificial

<220>

50 <223> secuencia de STAR18

<400> 18

	atctttttgc tctctaaatg tattgatggg ttgtgttttt tttccacct gctaataaat	60
55	attacattgc aacattcttc cctcaacttc aaaactgctg aactgaaaca atatgcataa	120
	aagaaaatcc ttgcagaag aaaaaagct attttctccc actgatttg aatggcactt	180
	gcggatgcag ttcgcaaatc ctattgccta ttcctcatg aacattgtga aatgaaacct	240
60	ttggacagtc tgccgcattg cgcattgagac tgccctgcga aggcaagggt atggttccca	300
	aagcaccag tggtaaatcc taacttatta ttcctttaa attccaatgt aacaacgtgg	360
	gccataaaag agtttctgaa caaaacatgt catctttgtg gaaagggtgt tttcgttaatt	420
	aatgatggaa tcatgctcat ttcaaatgg aggtccacga tttgtggcca gctgatgcct	480
65	gcaattatc ct	492

ES 2 344 073 T3

<210> 19

<211> 1840

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> secuencia de STAR19

10

<400> 19

15	tcacttcctg atatatttaca ttcaaggcta gctttatgca tatgcaacct gtgcagttgc	60
	acagggcctt gtgttcagaa agactagctc ttggtttaat actctgttgt tgccatcttg	120
	agattcatta taatataatt tttgaatttg tgttttgaac gtgatgtcca atgggacaat	180
20	ggaacattca cataacagag gagacaggtc aggtggcagc ctcaattcct tgccaccctt	240
	ttcacataca gcattggcaa tgcccatga gcacaaaatt tgggggaacc atgatgctaa	300
	gactcaaagc acatataaac atgttacctc tgtgactaaa agaagtggag gtgctgacag	360
25	ccccagagg ccacagtta tgttcaaacc aaaacttgct taggggtcag aaagaaggca	420
	atggcagggt ctaagaaaca gcccatcata tccttgttta ttcattgttac gtccctgcat	480
	gaactaatca cttacactga aaatattgac agaggaggaa atggaaagat agggcaaccc	540
30	atagttcttt ttccttttag tctttcctta tcagtaaacc aaagatagta ttggtaaaat	600
	gtgtgtgagt taattaatga gttagtttta ggcagtgttt ccaactgttg ggtaagaaca	660
	aaatatatag gcttgtattg agctattaaa tgtaaattgt ggaatgtcag tgattccaag	720
35	tatgaattaa atatccttgt atttgcattt aaaattggca ctgaacaaca aagattaaca	780
	gtaaaattaa taatgtaaaa gtttaatttt tacttagaat gacattaaat agcaaataaa	840
	agcaccatga taaatcaaga gagagactgt ggaaagaagg aaaacgtttt tatttttagta	900
40	tatttaattg gactttcttc ctgatgtttt gttttgtttt gagagagagg gatgtggggg	960
	cagggagggtc tcattttgtt gcccaggctg gacttgaact cctgggctcc agctatcctg	1020
	ccttagcttc ttgagtagct gggactacag gcacacacca cagtgtctga cattttctgg	1080
45	atTTTTTTTT tttttttatt ttttttgta gacaggttct ggctctgtta ctcaggttgc	1140
	agtgcagtgg catgatagcg gctcactgca gcctcaacct cctcagctta agctactctc	1200
	ccacttcagc ctcctgagta gccaggacta cagtgtgttg ccaccacacc tgtggctaata	1260
	ttttgtagag atggggcttc tccacgttgc cgaggctggt ctccaactcc tgggtctcaag	1320
50	cgaacctcct gacttggcct ccggaagtgc tgggattaca ggcttgagcc actgcatcca	1380
	gcctgtcctc tgtgttaaac ctactccaat ttgtctttca tctctacata aacggctctt	1440
	ttcaaagttc ccatagacct cactgttgct aatctaataa taaattatct gccttttctt	1500
55		
	acatggttca tcagtagcag cattagattg ggctgctcaa ttcttcttgg tatattttct	1560
60	tcatttggtc tctggggcat cacactctct ttgagttact cattcctcat tgatagcttc	1620
	ttcctagtct tctttacttg ttcttctct tctcctgac tctttaatat tgtttttctc	1680
	cccaggcttt agttcttagt cctcttctgt tatctattta caccaattc tttcagagtc	1740
65	tcatccagag tcatgaactt aaacctgttt ctgtgcagat aattcacatt attatatctc	1800
	cagcccagac tctcccgcaa actgcagact gatcctactg	1840

ES 2 344 073 T3

<210> 20

<211> 780

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> secuencia de STAR20

10

<400> 20

	gatctcaagt ttcaatatca tgttttggca aaacattcga tgctcccaca tccttaccta	60
15	aagctaccag aaaggctttg ggaactgtca acagagctac agaaaagtca gttaaagacca	120
	atggaccctt caaacaaaaa cagccaagct tttctgcaa aaagatgact gagaagactg	180
	ttaaagcaaa aaactctgtt cctgcctcag atgatggcta tccagaaata gaaaaattat	240
20	ttcccttcaa tcctctaggc ttcgagagtt ttgacctgcc tgaagagcac cagattgcac	300
	atctccccctt gagtgaagtg cctctcatga tacttgatga ggagagagag cttgaaaagc	360
	tgtttcagct gggccccctt tcacctttga agatgccctc tccaccatgg aaatccaatc	420
25	tgttgcagtc tcctttaagc attctgttga ccctggatgt tgaattgcc a cctgtttgct	480
	ctgacataga tatttaaatt tcttagtgct ttagagtttg tgtatatttc tattaataaa	540
	gcattatttg tttaacagaa aaaaagatat atacttaaat cctaaaataa aataaccatt	600
30	aaaaggaaaa acaggagtta taactaataa gggaacaaag gacataaaat gggataataa	660
	tgcttaatcc aaaataaagc agaaaatgaa gaaaaatgaa atgaagaaca gataaataga	720
	aaacaaatag caatatgaaa gacaaacttg accgggtgtg gtggctgatg cctgtaatcc	780

35 <210> 21

<211> 607

<212> ADN

40 <213> Secuencia Artificial

<220>

<223> secuencia de STAR21

45

<400> 21

	gatcaataat ttgtaatagt cagtgaatac aaaggggtat atactaaatg ctacagaaat	60
50	tccattcctg ggtataaatc ctagacatat ttatgcatat gtacaccaag atatatctgc	120
	aagaatgttc acagcaaadc tctttgtagt agcaaaaggc caaaaggctt atcaacaaga	180
	aaattaatac attgtggcac ataatggcat ccttatgcc aataaaatgg atgaaattat	240
55	agttaggttc aaaaggcaag cctccagata atttatatca tataattcca tgtacaacat	300
	tcaacaacaa gcaaaactaa acatatatac atgtcaggga aaatgatgaa caagggttaga	360
60	aaatgattaa tataaaaaata ctgcacagtg ataacattta atgagaaaaa aagaagggaag	420
	ggcttaggga gggacctaca gggaactcca aagttcatgg taagtactaa atacataatc	480
	aaagcactca aaatagaaaa tatttttagta atgttttagc tagttaatat cttacttaaa	540
65	acaagggtcta ggccaggcac ggtggctcac acctgtaatc ccagcacttt gggaggctga	600
	ggcgggt	607

ES 2 344 073 T3

	<210> 22	
	<211> 1380	
	<212> ADN	
5	<213> Secuencia Artificial	
	<220>	
	<223> secuencia de STAR22	
10	<400> 22	
15	cccttgtgat ccacccgcct tggcctccca aagtgtctggg attacaggcg tgagtcacta	60
	cgcccgccca ccctccctgt atattatttc taagtatact attatgttaa aaaaagttaa	120
	aaaatattga tttaatgaat tcccagaaac taggatttta catgtcacgt tttcttatta	180
20	taaaaataaa aatcaacaat aaatatatgg taaaagtaaa aagaaaaaca aaaacaaaaa	240
	gtgaaaaaaa taaacaacac tcctgtcaaa aaacaacagt tgtgataaaa ctttaagtgcc	300
	tgaaaattta gaaacatcct tctaaagaag ttctgaataa aataaggaat aaaataatca	360
25	catagttttg gtcattgggt ctgtttatgt gatggattat gtttattgat ttgtgtatgt	420
	tgaacttata tcaatagatg cagacaaggc cttgataaaa gtttttaaca ctttttcattg	480
	ttgaaaactc tcaatagact aggtattgat gaaacatatc tcaaaataat agaagctatt	540
	tatgataaac ccatagccaa tatcactatg agtgggcaaa agctggaagc attccctttg	600
30	aaaactggca caagacaagg atgccctctc tcaccactcc tattaaatgt agtattggaa	660
	gttctggcca gagcaatcag gcaggagaaa gaaaaggat taaaatagga agagaggaag	720
	tcaaatgtc tctgtttgca gtaaactatg ttgtatatatt agaaaacccc attgtctcat	780
35	cctaaaaact ccttaagctg ataaacaact tcagcaaagt ctcaggatac aaaatcaatg	840
	tgcaaaaatc acaagcattc ctatacaccg ataatagaca gcagagagcc aaatcatgag	900
	tgaagtccca ttcacaattg cttcaaagaa aataaaaatac ttaggaatac aactttcacg	960
40	ggacatgaag gacattttca aggacaacta aaaaccactg ctcaaggaaa tgagagagga	1020
	cacaaagaaa tggaaaaaca ttccatgctc atggaagaat caatatcatg aaaatggcca	1080
	tactgcccac agtaatttat agattcaatg ctaaccccat caagccacca ttgactttct	1140
45	tcacagaact agaaaaaac tatttttaaaa ctcatatgta gtcaaaaaga gtcggtatag	1200
	ccaagacaat cctaagcata aagaacaaag ctggatgcat cacgtgact tcaaaccata	1260
	ctacaaggct acagtaacca aaacagcatg gtactggtag caaacagat agatagaccg	1320
50	atagaacaga acagaggcct cggaaataac accacacatc tacaaccctt tgatcttcaa	1380

	<210> 23
55	<211> 1246
	<212> ADN
	<213> Secuencia Artificial
60	<220>
	<223> secuencia de STAR23

65

ES 2 344 073 T3

<400> 23

```

5      atccccctcat ccttcagggc agctgagcag ggcctcgagc agctggggga gcctcactta      60
      atgctccttg gagggcagcc agggagcatg gggctctgag gcatgggtcca gggctctgca      120
      ggcggcacgc accatgtgca gccgccccca cctgttgctc tgcctccgcc acctggccat      180
      gggcttcagc agccagccac aaagtctgca gctgctgtac atggacaaga agcccacaag      240
10     cagctagagg accttggtt ccacgtgccc agggagcatg gccacagcc caaagaccag      300
      tcaggagcag gcaggggctt ctggcaggcc cagctctacc tctgtcttca cacagatggg      360
      agatttctgt tgtgattttg agtgatgtgc ccctttggtg acatccaaga tagttgctga      420
15     agcaccgctc taacaatgtg tgtgtattct gaaaacgaga acttctttat tctgaaataa      480
      ttgatgcaaa ataaattagt ttggatttga aattctattc atgtaggcat gcacacaaaa      540
      gtccaacatt gcatatgaca caaagaaaag aaaaagcttg cattccttaa atacaaatat      600
20     ctgttaacta tatttgcaaa tatatttgaa tacacttcta ttatgttaca tataatatta      660
      tatgtatatg tatatataat atacatatat atgttacata taatatactt ctattatggt      720
      acatataata tttatctata agtaaataca taaatataaa gatttgagta gctgtagaac      780
25     attgtcttat gtgttatcag ctactactac aaaaatatct cttccactta tgccagtttg      840
      ccatataaat atgatcttct cattgatggc ccagggaag agtgcagtgg gtacttattc      900
      tctgtgagga gggaggagaa aagggaacaa ggagaaagtc acaaaggga aactctggtg      960
30     ttgccaaaat gtcaagtttc acatattccg agacggaaaa tgacatgtcc cacagaagga      1020
      ccctgcccag ctaatgtgtc acagatatct caggaagctt aatgatttt tttaaaagaa      1080
      aagagatggc attgtcactt gtttcttgta gctgaggctg tgggatgatg cagatttctg      1140
35     gaaggcaaag agctcctgct ttttccacac cgagggactt tcaggaatga ggccagggtg      1200
      ctgagcacta caccaggaaa tccctggaga gtgtttttct tactta      1246

```

40 <210> 24

<211> 939

<212> ADN

45 <213> Secuencia artificial

<220>

<223> secuencia de STAR24

50

<400> 24

```

55     acgaggtcac gagttcgaga ccagcctggc caagatggtg aagccctgtc tctactaaaa      60
      atacaacaag tagccgggcg cgggtgacggg cgcctgtaat cccagctact caggaggctg      120
      aagcaggaga atctctagaa cccaggaggc ggagggtgag tgagctgaga ctgccccgct      180

```

60

65

ES 2 344 073 T3

	gcactctagc ctgggcaaca cagcaagact ctgtctcaaa taaataaata aataaataaa	240
	taaataaata aataaataaa tagaaaggga gagttggaag tagatgaaag agaagaaaag	300
5	aaatcctaga tttcctatct gaaggcacca tgaagatgaa ggccacctct tctgggccag	360
	gtcctcccgt tgcaggtgaa ccgagttctg gcctccattg gagaccaaag gagatgactt	420
	tggcctggct cctagtgagg aagccatgcc tagtcctgtt ctgtttgggc ttgatcctgt	480
10	atcacttgat tgtctctcct ggactttcca tggattccag ggatgcaact gagaagtta	540
	ttttaaatgc acttacttga agtaagagtt attttaaaac attttagcaa aggaaatgaa	600
	ttctgacagg ttttgactg aagacattca catgtgagga aaacaggaaa accactatgc	660
15	tagaaaaagc aaatgctgtt gagattgtct cacaaacaca aattgcgtgc cagcaggtag	720
	gtttgagcct caggttgggc acattttacc ttaagcgac tgttggtgga acttaagggtg	780
	actgtaggac ttatatatac atacatacat ataatatata tacatattta tgtgtatata	840
20	cacacacaca cacacacaca cacacagggc cttgctatct tgcccagggt ggtctccaac	900
	tctgggtctc aagcgatcct ctgcctcccc ttcccaaag	939

25 <210> 25
 <211> 1067
 <212> ADN
 30 <213> Secuencia artificial
 <220>
 <223> secuencia de STAR25
 35 <400> 25

40	ataaaaaaat aaaaaaccct gctctaattt gcaaaggctc tatctttcct cccaaccacc	60
	tgaattttta gtgaaaacgg ggcttcctgt aggaaggagt agctagctat cccggtccgc	120
	tacaggttat cagtgcgtga ataccctgac tcctaaggct caggatttga ctgggtcgcc	180
45	tcgtccgact gccccgcccc caacgcggac ccacgtcacc gcgcgccagc ctgcggccgt	240
	cctgacctcg cgggatttga gcttcggtgc caacaaacac tcccaccgcg gctgcgtcca	300
	ctttacctgc cggcggcgac cagcttctga agaaaagtgt ccaccatggt gtcgaggagc	360
50	ttcaccctcg aaatggtagt gccgggtggc acagattccg aagacgacct ctcatgcctt	420
	ttttctcac agccgctgcc tagattggcg ctacttgctt cggccatggt gaagtgaac	480
	ctccaaatct aactggcccc gcctccccgc ctgccggagc tcccgattgg ccgctccgc	540
55	gaaggggtgcc tccgattgga agcagtagaa cgtctgtcac cgagcagggc gggggcgggg	600
	aagtcattcg aggctgaggg cagcggggag gcgaggctct gcgcgggtgg atgtccgcga	660
	ccggaaaaat acgcgcaagc caaagctcgg gggctcaata aaaactttta attacatttc	720
60	agagacttcg tacagtgcaa cagtgaatat tcactgttaa ttttcacaag agtccatttc	780
	atcaaacgtt cagagagtct gccttttcat tcccttgctc ctgagtgtc caatcaggtt	840
	tccagtctcc cagaggtttc ttttagtttt gattaccgac caaaactcca gtttagggag	900
65	aatggaagtc caccgtccca tccccaccaa aacatatttc agtcaaacc aatcccagtc	960

ES 2 344 073 T3

cctaaagaat taggaaagta tgggccaagg gtccttttaa ttatacacac atcaccctta 1020
aaactgcgtg tgtgtacgag aaataaagaa aaacacaaga ggggctg 1067

5 <210> 26
 <211> 540
 <212> ADN
 <213> Secuencia Artificial
 10 <220>
 <223> secuencia de STAR26

15 <400> 26
 ccccttgaca agccccagtg tgtgatgttc ccactctgt gtccatgcat tctcattggt 60
 caactcccat ctgtgagtga gaacatgcag tgtttgggtt tctgtccttg agatagtttg 120
 20 ctgagaatga tggtttccag cttcatccat gtccttgcaa aggaagtga cttatccttt 180
 tttatggctt catagtattc catggcacat atgtgccaca ttttttaaat ccagtctatc 240
 attgatggac atttggggtg gttccaagtc tttgctattg tgaatagcac cacaattaac 300
 25 atatgtgtgc atgtatacat ctttatagta gcatgattta taatccttcg ggtatatacc 360
 ctgtaatggg atcgtctggg caaatgggtat ttctagttct agatccttga ggaatcacca 420
 cactgctttc cacaatggtt gaactaattt acgctccac cagcagtgtg aaagcattcc 480
 30 tatttctcca cgtcctctcc agtatctgtt gtttcctgac tttttaatga tcatcattct 540

<210> 27
 <211> 1520
 <212> ADN
 <213> Secuencia Artificial
 35 <220>
 <223> secuencia de STAR27
 40 <400> 27

45 cttggccctc acaaagcctg tggccaggga acaattagcg agctgcttat tttgctttgt 60
 atccccaatg ctgggcataa tgcctgccat tatgagtaat gccggtagaa gtatgtgttc 120
 aaggaccaa gttgataaat accaaagaat ccagagaagg gagagaacat tgagtagagg 180
 50 atagtgcag aagagatggg aacttctgac aagagtgtg aagatgtact aggcaggggg 240
 aacagcttaa ggagagtcac acaggaccga gctcttgca agccggctgc catggaggct 300
 ggggtggggc atggtagctt tcccttcctt ctcaggttca gagtgcagc cttgaacttc 360
 55 taattcccag aggcatttat tcaatgtttt cttctagggg catacctgcc ctgctgtgga 420
 agactttctt ccctgtgggt cgccccagtc ccagatgag acggtttggg tcagggccag 480
 gtgcaccgtt ggggtgtgtc ttatgtctga tgacagttag ttactcagtc attagtcatt 540
 60 gagggagggt tggtaaagat ggagatgctg ggtcacatcc ctagagagg gttccagtat 600
 gggcacatgg gagggctgga aggatagggt actgctagac gtagagaagc cacatccttt 660
 aacaccctgg cttttccac tgccaagatc cagaaagtcc ttgtggtttc gctgctttct 720
 65 cctttttttt tttttttttt tttctgagat ggagtctggc tctgtcgccc aggctggagt 780

ES 2 344 073 T3

gcagtggcac gatttcggct cactgcaagt tccgcctcct aggttcatac cattctccca 840
 cctcagcctc ccgagtagct gggactacag gcgccaccac acccagctaa ttttttgtat 900
 5 ttttagtaga gacggcgttt caccatgtta gccaggatgg tcttgatccg cctgcctcag 960
 cctcccaaag tgctgggatt acaggcgta gccaccgcgc ccggcctgct ttctttcttc 1020
 atgaagcatt cagctggtga aaaagctcag ccaggctggt ctggaactct tgacctcaag 1080
 10 tgatctgcct gcctcagcct cccaaagtgc tgagattaca ggcatgagcc agtccgaatg 1140
 tggttttttt tgttttgttt tgaaacaagg tctcactggt gcccaggctg cagtgcagtg 1200
 gcatacctca gctccactgc agcctcgacc tcctgggctc aagcaatcct cccaactgag 1260
 15 cctccccagt agctggggct acaagcgcat gccaccacgc ctggctatct tttttttttt 1320
 tttttttttt gagaaggagt ttcatcttg ttgccaggc tggagtgcaa tggcacagtc 1380
 tcagctcact gcagcctccg cctcctgggt tcaagcgatt ctctgcctc agcctcccga 1440
 20 gtagctggga ttataggcac ctgccaccat gcctggctaa tttttttgta ttttagtag 1500
 ggatgggggt tcaccatggt 1520

25 <210> 28
 <211> 961
 <212> ADN
 <213> Secuencia Artificial
 30 <220>
 <223> secuencia de STAR28
 35 <400> 28

aggaggttat tcctgagcaa atggccagcc tagtgaactg gataaatgcc catgtaagat 60
 40 ctgtttaccc tgagaagggc atttcctaac tctccctata aaatgccaag tggagcacc 120
 cagatgaaat agctgatatg ctttctatac aagccatcta ggactggctt tatcatgacc 180
 aggatattca cccactgaat atggctatta cccaagttat ggtaaagtct gtagttaagg 240
 45 ggggtccctc cacatggaca ccccaggta taaccagaaa gggttcccaa tctagactcc 300
 aagagagggt tcttagacct catgcaagaa agaactggg gcaagtacat aaagtgaaag 360
 caagtttatt aagaaagtaa agaaacaaaa aaatggctac tccataagca aagttatttc 420
 50 tcacttatat gattaataag agatggatta ttcatgagtt ttctgggaaa ggggtgggca 480
 attcctggaa ctgagggttc ctcccacttt tagaccatat agggatatct cctgatattg 540
 ccatggcatt tgtaactgt catggcactg atgggagtg cttttagcat tctaatagcat 600
 55 tataattagc atataatgag cagtgaggat gaccagaggt cacttctgtt gccatattgg 660
 tttcagtggg gtttggttgg cttttttttt tttttaacca caacctgtt tttatttatt 720
 tattttatta tttatttatt tatatttttt attttttttt agatggagtc ttgctctgtc 780
 60 acccaggta gagtgcagtg gcaccatctc ggctcactgc aagctctgcc tccttgggtc 840
 acgccattct gctgcctcag cctcccaggt agctgggact acaggtgcct gccaccatac 900
 ccggctaatt ttttctattt ttcagtagag acgggggttc accgtgttag ccaggatggt 960
 65 c 961

ES 2 344 073 T3

<210> 29

<211> 2233

<212> ADN

5 <213> Secuencia artificial

<220>

<223> secuencia de STAR29

10

<400> 29

	agcttggaca cttgctgatg ccacttttga tgttgaaggg ccgccctctc ccacaccgct	60
15	ggccactttt aaatatgtcc cctctgcccc gaagggcccc agaggagggg ctggtgaggg	120
	tgacaggagt tgactgctct cacagcaggg gggtccggag ggaccttttc tccccattgg	180
	gcagcataga aggacctaga agggccccct ccaagcccag ctgggcgtgc agggccagcg	240
20	attcgatgcc tttccctgac tcaggtggcg ctgtcctaaa ggtgtgtgtg ttttctgttc	300
	gccagggggg ggcggataca gtggagcatc gtgcccgaag tgtctgagcc cgtggttaagt	360
	ccctggaggg tgcacgggtct cctccgactg tctccatcac gtcaggcctc acagcctgta	420
25	ggcaccgctc ggggaagcct ctggatgagg ccatgtggtc atccccctgg agtcctggcc	480
	tggcctgaag agggagggag gaggaggcca gcccctccct agccccaagg cctgcgaggg	540
	tgcaagcccc gccccacatt ctagtccagg cttggctgtg caagaagcag attgcctggc	600
30	cctggccagg cttcccagct aggatgtggt atggcagggg tgggggacat tgaggggctg	660
	ctgtagcccc cacaacctcc ccaggtaggg tgggtgaacag taggctggac aagtggacct	720
	gttcccatct gagattcaag agcccacctc tcggagggtg cagtgcagcc agatccctcc	780
35	actgcactcc agcctgggca acagagcaag actctgtctc aaaaaaacag aacaacgaca	840
	acaaaaaacc cactctggc ccactgccta actttgtaaa taaagtttta ttggcacata	900
	gacacaccca ttcatattaca tactgctgcg gctgcttttg cattaccctt gagtagacga	960
40	cagaccacgt ggccatggaa gccaaaaata ttactgtctt ggccctttac agaagtctgc	1020
	tctagagggg gaccccgcc catggggcag gaccactggg cgtgggcaga agggaggcct	1080
	cggtgcctcc acgggcctag ttgggtatct cagtgcctgt ttcttgcatg gagcaccagg	1140
45	ggtcagggca agtacctgga ggaggcaggc tgttgcccgc ccagcactgg gaccaggag	1200
	accttgagag gctcttaacg aatgggagac aagcaggacc agggctccca ttggctgggc	1260
	ctcagtttcc ctgcctgtaa gtgagggagg gcagctgtga aggtgaactg tgaggcagag	1320
50	cctctgctca gccattgcag gggcggtctt gcccactcc tgttggtcac ccagagtga	1380
	gggcacgggg tgagatgtca ccatcagccc ataggggtgt cctcctggtg ccagggtccc	1440
	aagggatgtc ccatcccccc tggctgtgtg gggacagcag agtcctggg gctgggaggg	1500
55	ctccacactg ttttgtcagt ggtttttctg aactgttaaa tttcagtga aaattctctt	1560
	ttccctttta ctgaaggaac ctccaaagga agacctgact gtgtctgaga agttccagct	1620
	ggtgctggac gtcgcccaga aagcccaggt actgccacgg gcgccggcca ggggtgtgtc	1680

60

65

ES 2 344 073 T3

5 tgcgccagcc atgggcacca gccaggggtg tgtctacgcc ggccaggggt aggtctccgc 1740
 cggcctccgc tgctgcctgg ggagggccgt gcctgacact gcaggcccgg tttgtccgcg 1800
 gtcagctgac ttgtagtac cctgcccttg gatggtcgtt acagcaactc tgggtggttg 1860
 ggaagggggc tcctgattca gcctctgcgg acggtgcgag aggggtggagc tccccctcct 1920
 cccacccgcc cctggccagg gttgaacgcc cctgggaagg actcaggccc gggctctgctg 1980
 10 ttgctgtgag cgtggccacc tctgccctag accagagctg ggccttcccc ggcctaggag 2040
 cagccgggca ggaccacagg gctccgagtg acctcagggc tgcggacct ggaggccctc 2100
 ctggcgctgc ggtgtgactg acagcccagg agcgggggct gttgtaattg ctgtttctcc 2160
 15 ttcacacaga accttttcgg gaagatggct gacatcctgg agaagatcaa gaagtaagtc 2220
 ccgcccccca ccc 2233

20 <210> 30
 <211> 1851
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> secuencia de STAR30
 30 <400> 30

cctccccctg agccttcaga aggagcatgg cataggagtc ttgatttcag acgtctggctc 60
 35 cccagaatga tgggagaatg aatttctgtt atttaagcca cccaacctgt ggtgctttgt 120
 tatagcagcc tcaggaaact aacacactgc acgtgccac tattccccct tccagtatct 180
 ttcaggactt gctggcttcc tttgttctgg cgtacaccca tgcattggccc cattccccac 240
 40 ttcctaaaac aacaaccctg acttagtctg tttgggctgc tagaacaata tactatagac 300
 tgggtgactt ataaacaaca gaaattcatt tctcacattc tggaggtctg gaagtccaat 360
 atcgaggcac catcacattt ggtctctgct gaggccccct tcctagctcc tcaactgtgtc 420
 45 cttacatggc agaaggggca aggcagctct ctggggctcc ttttcaaggc cacaatccc 480
 attcattagg gctgatgact tcatgactta atcacctctt aatggcccca cctcctaatt 540
 gcattgggag ttaggattca acataaattt tgggggggaca cacatattca gaccatagca 600
 50 aaccccaaca ataaaaaac ttcactttaa ggttccaaat ggactggcag ttaaattcatg 660
 ttcataattt cataaaagaa ggagtaagtc aacaaattga taaacgcgtg gagatttggt 720
 cggatggatg ttcacaaaaa tgctggcctt aaagagttag atgggaaatg ggaactatta 780
 55 cattcttctt catacttttt ggtactgcct gcattgttaa aaaaaaaaaa aaagagcaca 840
 gagcattttt acaatcagga aaaaaacaat gaggttatct tcattctgga aaaaaatgga 900
 aatgaaaca gtggagtcac atcatggaaa atgcttatgg tacaatttca tgtgacataa 960
 60 aacaatagaa tagaggacct gttttatgac taaagcactg taaaaatgac aggcctggaa 1020
 ggagagatga aaaccactca tttgttaagg tagtcagggt gcagggtgatt tctcttcttt 1080
 tgaaaatttc cattttcatt atatcgaggt ttgtgcattt actaaaactt tcggttggtg 1140

65

ES 2 344 073 T3

cacatgcata aatagataga taaataagta gatagatgat agataaatag acggtaggta 1200
 gatagataga tagatatgag aaataagtcc cctgtacttg gccttgacgc cataactagt 1260
 5 cattccccctt cctctgtcca ttgctatgcc tgatggacaa ggcagtctgt gccctctggc 1320
 cccaattcca atgtgccctc tgctcctggc tgtagtccc tttccacccc aatacaattg 1380
 ctccgaggtc acttctaagt gtgaagcccc cagatcagat ggcttcttct gtgtccttac 1440
 10 cttacccaat ttctaattat aactaaaaca caatgaggct ctagtaaaat accatgagac 1500
 ttcaggccct ctgtataact tactcattt aaacctaaca aggaaaacct accatgaatc 1560
 cgaggcacag agcagctaag gaactacca aggtcacgca gctattggtg atggaaccat 1620
 15 gagtcaagct tcacagcctg ttggctctag aatagggtt cccaacctca gcactgtgga 1680
 ctttttcagg ctggataatt ctctgttggt gggggctgtt ctgtgccttg taggatatta 1740
 ggagcatctc tggcctctac ccactagacg cagcagcact cccatgccca gttgtgacaa 1800
 20 caagcaatgt ctcccaccat tgccaagtgt cccctgggtg gaaatgcacc c 1851

<210> 31

<211> 1701

25 <212> ADN

<213> Secuencia artificial

<220>

30 <223> secuencia de STAR31

<220>

35 <221> característica-misc

<222> (159)..(1696)

<223> "N" representa cualquier ácido nucleico en diferentes posiciones

40 <400> 31

caccgcctt ggccccccag agtgctggga ttacaagtgt aaaccacat tcctggctag 60
 atttaatttt ttaaaaaata aagagaagta ggaatagttc attttaggga gagcccccta 120
 45 actgggacag gggcaggaca ggggtgaggc ttccttant tcaagctcac ctcaaacca 180
 cccaggactg tgtgtcacat tctccaataa aggaaagggt gctgcccccg cctgtgagt 240
 ctgcagtggg gggtagaggg ccgtgggcag agtgcttcat ggactgtca tcaagaaagg 300
 50 cttcatgaca atcgccccag ctgctgtcat cccacattct acttcagct aggagaaggc 360
 ggcttgccca cagtcacca gccggcaagt gtcaccctg ggttggaacc agagctatga 420
 tcctgcccag ggggtccagct gagaatcagg cccacgttct aggcagagg gctcacctac 480
 55 tgggactcca gtagctgtag tgcatggagg catcatggct gcagcagcct ggacctggct 540
 tcacactggc tgtccctgtg ggcaggccat cctcaatgcc aggtcaggcc caagcatgta 600
 tcccagacaa tgacaatggg gtggaatcct ctcttgccc agaagccact cctactgtt 660
 ctacctgagg aaggcagggg catggtggaa tcctgaagcc tgctgtgagg gtctccagcg 720
 60 aacttgcaca tggtcagccc tgccttctcc tccctgaact agattgagcg agagcaagaa 780
 ggacattgaa ccagcaccca aagaattttg gggaacggcc tctcatccag gtcaggctca 840

65

ES 2 344 073 T3

cctccttttt aaaatttaat taattaatta attaatTTTT ttttagagac agagtcttac 900
 tgtgtggccc aggcgttagt gcagtggcac aatcatagtt cactgcagcc tcaaactccc 960
 5 cacctcagcc tctggattag ctgagactac aggtgcacca ccaccacacc cagctaatat 1020
 ttttattttt gtagagagag ggtttcacca tcttgcacag gctggtctca aactcctggg 1080
 ctcaagtgat cccgcccagg tctgaaagcc cccaggctgg cctcagactg tgggggtttc 1140
 10 catgcagcca cccgagggcg cccccaagcc agttcatctc ggagtccagg cctggccctg 1200
 ggagacagag tgaaccaggt ggtttttatg aacttaactt agagtttaaa agatttctac 1260
 tcatcactt gtcaagatgc gccctctctg gggagaaggg aacgtgactg gattccctca 1320
 15 ctgttgatc ttgaataaac gctgctgctt catcctgtgg gggcgtggc cctgtccctg 1380
 tgtgggtggg gcctcttcca tttccctgac ttagaaacca cagtccacct agaacagggg 1440
 ttgagaggct tagtcagcac tgggtagcgt tttgactcca ttctcggtt tcttctttt 1500
 20 ctttccagga tttttgtgca gaaatgggtt tttgttgcc gtgttagtcc tccttggag 1560
 gcagctcaga agggccgtga aatgtcgggg gacaggaccc ccaggagggg aaccccaggc 1620
 tacgcacttt aggggttcgtt ctccaggagg ggcgacctga ccccgcnatc cgtcgngcg 1680
 25 cgngnnnacn aanntttcc c 1701

<210> 32

<211> 771

30 <212> ADN

<213> Secuencia artificial

<220>

35 <223> secuencia de STAR32

<400> 32

40 gatcacacag cttgtatgtg ggagctagga ttggaacccc agaagtctgg cccaggttc 60
 atgtctcac cactgcata caatggcctc tcataaatca atccagtata aaacattaga 120
 atctgttta aaaccataga attagtagcg taagtaataa atgcagagac catgcagtga 180
 45 atggcattcc tggaaaaagc cccagaagg aattttaaat cagctttcgt ctaattctga 240
 gcagctagtt agcaaatatg agaatacagt tgttcccaga taatgcttta tgtctgacca 300
 tcttaactg gcgtgtttt tcaaaaactt aaaaacaaa tccatgactc ttttaattat 360
 50 aaaagtgata catgtctact tgggaggctg aggtggtggg aggatggctt gagtttgagg 420
 ctgcagtatg ctactatcat gcctataaat agccgctgca ttccagcttg ggcaacatac 480
 ccaggcccta tctcaaaaaa ataaaaagta atacatctac attgaagaaa attaatTTTA 540
 55 ttgggttttt ttgcattttt attatacaca gcacacacag cacatatgaa aaaatgggta 600
 tgaactcagg cattcaactg gaagaacagt actaaatcaa tgtccatgta gtcagcgtga 660
 ctgaggttgg tttgtttttt ctttttctt ctcttctctt ctctttctt ttttttgag 720
 60 acggagcttt gctctttttg cccaggcttg attgcaatgg cgtgatctca g 771

<210> 33

<211> 1368

65 <212> ADN

<213> Secuencia Artificial

ES 2 344 073 T3

<220>

<223> secuencia de STAR33

5 <400> 33

	gcttttatcc tccattcaca gctagcctgg cccccagagt acccaattct ccctaaaaaa	60
	cggtcatgct gtatagatgt gtgtggcttg gtagtgctaa agtggccaca tacagagctc	120
10	tgacaccaaa cctcaggacc atgttcatgc cttctcactg agttctggct tgttcgtgac	180
	acattatgac attatgatta tgatgacttg tgagagcctc agtcttctat agcactttta	240
	gaatgcttta taaaaacat ggggatgtca ttatatctta acctgttagc acttctgttc	300
15	gtattacca tcacatccca acatcaattc tcatatatgc aggtacctct tgtcacgcgc	360
	gtccatgtaa ggagaccaca aaacaggctt tgtttgagca acaaggtttt tatttcacct	420
	gggtgcagggt gggctgagtc tgaaaagaga gtcagtgaag ggagacaggg gtgggtccac	480
20	tttataagat ttgggtagggt agtggaataa tacaatcaaa gggggttgtt ctctggctgg	540
	ccagggtggg ggtcacaagg tgctcagtgg gagagccttt gagccaggat gagccagaag	600
	gaatttcaca aggtaatgtc atcagttaag gcagggactg gccattttca cttcttttgt	660
25	gggtggaatgt catcagttta ggaggaacc ggccattttc acttcttttg tgattcttca	720
	cttgcttcag gccatctgga cgtatagggt cagggtcacag tcacagggga taagatggca	780
	atggcatagc ttgggtcag aggcctgaca cctctgagaa actaaagatt ataaaaatga	840
30	tggtcgcttc tattgcaaat ctgtgtttat tgtcaagagg cacttatttg tcaattaaga	900
	accagtggt agaatcgaat gtccgaatgt aaaacaaaat acaaacctc tgtgtgtgtg	960
	tgtgtgtgag tgtgtgtgta tgtgtgtgtg tgtgtattag agaggaaaag cctgtatttg	1020
35	gagggtgat tcttagattc taggttcttt cctgcccacc ccatatgcac ccacccaca	1080
	aaagaacaaa caacaaatcc caggacatct tagcgcaaca tttcagtttg catattttac	1140
	atatttactt ttcttacata ttaaaaaact gaaaatttta tgaacacgct aagtttagatt	1200
40	ttaaattaag ttgttttta cactgaaaat aatttaatat ttgtgaagaa tactaataca	1260
	ttggtatatt tcattttctt aaaattctga acccctcttc cttattttcc ttttgacctg	1320
	attggtgtat tgggtcatgt actcatggat ttgccttaag gcaggagg	1368

45

<210> 34

<211> 755

<212> ADN

50

<213> Secuencia Artificial

<220>

<223> secuencia de STAR34

55

<400> 34

	actgggcacc ctcctaggca ggggaatgtg agaactgccg ctgctctggg gctgggcgcc	60
60	atgtcacagc aggaggagg acggtgttac accacgtggg aaggactcag ggtggtcagc	120

65

ES 2 344 073 T3

5 cacaaagctg ctggtgatga ccaggggctt gtgtcttcac tctgcagccc taacacccag 180
 gctgggttcg ctaggctcca tcctgggggt gcagaccctg agagtgatgc cagtgggagc 240
 10 ctcccgcccc tccccttcct cgaaggccca ggggtcaaac agtgtagact cagaggcctg 300
 agggcacatg tttatttagc agacaagggtg gggctccatc agcggggtgg cctggggagc 360
 agctgcatgg gtggcactgt ggggaggggtc tcccagctcc ctcaatggtg ttcgggctgg 420
 15 tgcggcagct ggcggcaccc tggacagagg tggatatgag ggtgatgggt ggggaaatgg 480
 gaggcacccg agatggggagc agcagaataa agacagcagc agtgcagggg ggcaggggga 540
 tgagcaaagg caggcccaag acccccagcc cactgcaccc tggcctcca caagccccct 600
 20 cgcagccgcc cagccacact cactgtgcac tcagccgtcg atacactggt ctgttaggga 660
 gaaagtccgt cagaacaggc agctgtgtgt gtgtgtgctg gtatgagtgt gtgtgtgtga 720
 tccctgactg ccaggtcctc tgcactgccc ctggg 755

25 <210> 35
 <211> 1193
 <212> ADN
 <213> Secuencia artificial
 30 <220>
 <223> secuencia de STAR35
 <220>
 35 <221> característica-misc
 <222> (355)..(1191)
 <223> "N" representa cualquier ácido nucleico en diferentes posiciones
 40 <220>
 <221> característica_misc
 <222> (312)..(1191)
 <223> "N" representa cualquier ácido nucleico en diferentes posiciones
 45

50

55

60

65

ES 2 344 073 T3

<400> 35

5	cgacttggtg atgcgggctc ttttttggtt ccatatgaac tttaaagtag tcttttccaa	60
	ttctgtgaag aaagtcattg gtaggttgat ggggatggca ttgaatctgt aaattacctt	120
	gggcagtatg gccattttca caatgttgat tcttcctatc catgatgatg gaatgttctt	180
10	ccattagttt gtatcctctt ttatttcctt gagcagtggt ttgtagttct ccttgaagag	240
	gtccttcaca tcccttgtaa gttggattcc taggtatttt attctctttg aagcaaattg	300
	tgaatgggag tncactcacg atttggtctt ctgtttgtct gctgggtgta taaanaatgt	360
15	ngtgaatntn gtacattgat ttngtatccn tgagacttng ctgaatttgc ttnatcngct	420
	tnngggaacc ttttgggctg aaacnatggg attttctaaa tatacaatca tgcgtctgc	480
	aaacagggaa caatttgact tcctcttttc ctaattgaat acactttatc tccttctcct	540
20	gcctaattgc cctgggcaaa acttccaaca ctatgntngn aataggagnt ggtgagagag	600
	ggcatccctg ttcttggtgc cagnttttca aagggaatgc ttccagtttt ggcccattca	660
	gtatgatatg ggctgtgggt nggtgcataa atagctctta tnattttgaa atgtgtccca	720
25		
	tcaataccta atttattgaa agtttttagc atgaangcat ngttgaattt ggtcaaaggc	780
	tttttctgca tctatggaaa taatcatgtg gtttttgtct ttggctcntg tttatatgct	840
30	ggatnacatt tattgatttg tgtatatnga acccagcctn ncatcccagg gatgaagccc	900
	acttgatcca agcttggcgc gcngnctagc tcgaggcagg caaaagtatg caaagcatgc	960
	atctcaatta gtcagacccc atagtccgcc cctacctcg cccatccgcc cctaactcng	1020
35	nccgttcgcc cattctcgcc catggctgac taatnttttt annatccaag cgngnccgcc	1080
	ctgcttganc attcagagtn nagagnnttg gaggccnagc cttgcaaaac tccggacngn	1140
	ttctnnggat tgaccccnnt taaatatttg gttttttgtn ttttcannng nga	1193

<210> 36

<211> 1712

<212> ADN

<213> Secuencia artificial

<220>

<223> secuencia de STAR36

ES 2 344 073 T3

<400> 36

5	gatcccatcc ttagcctcat cgatacctcc tgctcacctg tcagtgcctc tggagtgtgt	60
	gtctagccca ggcccatccc ctggaactca ggggactcag gactagtggg catgtacact	120
	tggcctcagg ggactcagga ttagtgagcc ccacatgtac acttggcctc agtggactca	180
10	ggactagtga gccccacatg tacacttggc ctcaggggac tcaggattag tgagcccca	240
	catgtacact tggcctcagg ggactcagga ttagtgagcc ccacatgtac acttggcctc	300
	aggggactca ggactagtga gccccacatg tacacttggc ctcaggggac tcagaactag	360
15	tgagccccac atgtacactt ggcttcaggg gactcaggat tagtgagccc cacatgtaca	420
	cttggacacg tgaaccacat cgatgtgctg cagagctcag ccctctgcag atgaaatgtg	480
	gtcatggcat tccttcacag tggcaccctt cgttccctcc ccacctcatc tcccattctt	540
20	gtctgtcttc agcacctgcc atgtccagcc ggcagattcc accgcagcat cttctgcagc	600
	acccccgacc acacacctcc ccagcgctg cttggcctc cagcccagct cccgccttcc	660
	ttccttgggg aagctccctg gacagacacc ccctcctccc agccatggct tttcctgct	720
25	ctgccccacg cgggaccctg ccctggatgt gctacaatag acacatcaga tacagtctt	780
	cctcagcagc cggcagacc ccagggtggact gctcggggcc tgctgtgag gtcacacagg	840
	tgctgttaac ttgccatctc agcaactagt gaatatgggc agatgctacc ttccttccgg	900
30	ttccctggtg agaggtagct gtggatgtcc tgtgttgccg gccacctttt gtccctggat	960
	gccattttatt tttttccaca aatatttccc aggtctcttc tgtgtgcaag gtattagggc	1020
	tgcagcgggg gccaggccac agatctctgt cctgagaaga cttggattct agtgcaggag	1080
35	actgaagtgt atcacaccaa tcagtgtaaa ttgttaactg ccacaaggag aaaggccagg	1140
	aaggagtggg gcatggtggg gttctagtgt tacaagaaga agccaggag ggcttctg	1200
	atgaagtggc atctgacctg ggatctggag gaggagaaaa atgtcccaa agagcagaga	1260
40	gcccacccta ggctctgcac caggaggcaa cttgctgggc ttatggaatt cagagggcaa	1320
	gtgataagca gaaagtcctt gggggccaca attaggattt ctgtcttcta aagggcctct	1380
	gccctctgct gtgtgacctt gggcaagtta cttcacctct agtgctttgg ttgcctcatc	1440
45	tgtaaagtgg tgaggataat gctatcacac tggttgagaa ttgaagtaat tattgctgca	1500
	aagggttat aagggtgtct aatactagta ctagtaggta cttcatgtgt cttgacaatt	1560
50	ttaatcatta ttattttgtc atcacctgca ctctccagg ggactaatgt ccctgctgtt	1620
	ctgtccaaat taaacattgt ttatccctgt gggcatctg caggtggt aggaaagcct	1680
	ggagctgttt cctgttgacg tgccagacta gt	1712

55 <210> 37

<211> 1321

<212> ADN

60 <213> Secuencia artificial

<220>

<223> secuencia de STAR37

65

ES 2 344 073 T3

<400> 37

5	atctctctct gccaaagcaa cagcgggtccc tgccccaacc agactacccc actcagtggg	60
	gttacggatg ctgctccagc atcctaacac tgcccagctg gtgcctgcct gtgctcacc	120
	acaacccccca ggccggcctt ccctgcagcc tgggcttggc caccttggcc tgattgagca	180
	ctgaggcctc ctgggcaccc agcccatca ctgcacctgc tgcttccagc cccacccac	240
10	cggctcaggg gttcttccca gcggcgctga tcatgaagtc aacatgcacg caagtcgtct	300
	caggaaactt tttaatgaaa gtgtcggcca cgggtggtgt taggtggctg agctcagatt	360
	gcagctgcta agacaccagc cacttacaa gagaaagcca ggctgcttca aaccagggc	420
15	cggaggcaaa aaagcatcac ttccagccgg ggagtctgga agccacgcct tgtgggaggt	480
	cacactggca tctaggcctt cgctgcact gcagaaggag agccgggtcc ccctcctgga	540
	gaacgctgcg tttcccagcc ccacaccggc tttgccacca cacaggctgt tgaggcagga	600
20	ggcgggtaag acgtagctgt agacccaaag caaccaccag ccctgggacc ctgcgggaga	660
	ggagcacttt tagaacatgg aaaaatgtgg tcatcccatc attagacagc acacatccta	720
	cataaataaa aagtcgtatg gggaaggagg ttggggaggg aataaaaaat tggcacagac	780
25	attgatagac tggtttccag tttcaaggta acagatgcac atcatgagac cagaggaggc	840
	agagacaagg gctgaatttg gcttttctaa gcaacatgtg ttcctgcgca gggctgaatg	900
	gtcgtgaga cagagatgga agccaggaca agggagccca ccgggccag ataggtacag	960
30	agagcagagg ctctgttct gtcctgcga cccatgaggg tgacactgt tgtaaatggt	1020
	ggctgtgctc tcccagcaag aaaaagcac aactaaatcc aactgcaca cagacgcaga	1080
	cagaaagcct tcaagtggct ctgttttctg ctccctgcct tgccagggtcc acaagcagag	1140
35	aggagtgtca ggcacatggc cccgctgtca ggctcccag tgagctgtag gctcagcagg	1200
	agctgcccac tgacacacag gggacacca ctctgccac cttgggagcg gttgccagac	1260
40	agagccgcac tgggtgctgg tgtcatccag ggacccaca cacttcctta aatgtgatcc	1320
	t	1321

45 <210> 38

<211> 1445

<212> ADN

50 <213> Secuencia Artificial

<220>

<223> secuencia de STAR38

55 <220>

<221> característica_misc

<222> (348)..(949)

60 <223> "N" representa cualquier ácido nucleico en diferentes posiciones

65

ES 2 344 073 T3

<400> 38

5	gatctatggg agtagcttcc ttagtgagct ttcccttcaa atactttgca accaggtaga	60
	gaattttgga gtgaagggtt tgttcttcgt ttcttcacaa tatggatatg catcttcttt	120
	tgaaaatggt aaagtaaatt acctctcttt tcagatactg tcttcatgcg aacttggtat	180
10	cctgtttcca tcccagcctt ctataaccca gtaacatctt ttttgaaacc agtgggtgag	240
	aaagacacct ggtcaggaac gcggaccaca ggacaactca ggctcaccga cggcatcaga	300
	ctaaaggcaa acaaggactc tgtataaagt accggtggca tgtgtatnag tggagatgca	360
15	gcctgtgctc tgcagacagg gagtcacaca gacacttttc tataatttct taagtgtctt	420
	gaatgttcaa gtagaaagtc taacattaaa ttgtattgaa caattgtata ttcatggaat	480
	atthttggaac ggaataccaa aaaatggcaa tagtggttct ttctggatgg aagacaaact	540
20	tttcttggtt aaaataaatt ttattttata tatttgaggt tgaccacatg accttaagga	600
	tacatataga cagtaaactg gttactacag tgaagcaaat taacatatct accatcgtac	660
	atagttacat ttttttgtgt gacaggaaca gctaaaatct acgtatttaa caaaaatcct	720
	aaagacaata catthttatt aactatagcc ctcatgatgt acattagatc gtgtggttgt	780
25	ttcttccgtc cccgccacgc ctctctctctg ggatggggat tcattcccta gcagggtgctg	840
	gagaactggc gcccttgtag ggtagggtgcc ccggagcctg aggcgggnac tttaanatca	900
	gacgcttggg ggccggctgg gaaaaactgg cggaaaatat tataactgna ctctcaatgc	960
30	cagctgttgt agaagctcct gggacaagcc gtggaagtcc cctcaggagg ctccgcgat	1020
	gtcctagggt gctgtctccg cccccacggt catttccatt gactcacacg cgccgcctgg	1080
	aggaggaggc tgcgctggac acgcccgtgg cgcctttgcc tgggggagcg cagcctggag	1140
35	ctctggcggc agcgtctggg gcggggcctc ggaggctggg cctggggacc caagggtggg	1200
	cggggcgtag gaggtgggct cagggttctc cagagaatcc ccatgagctg acccgaggg	1260
	cggccggggc agtaggcacc gggccccgc ggtgacctgc ggaccgaag ctggagcagc	1320
40	cactgcaaat gctgctgtag ccccaaatgc tgtgtccttt aaatgtttta attagaata	1380
	attaataggt ccgggtgtgg aggtcaagc cttaatcccc agcacctggc gagggcgagg	1440
45	aggga	1445

<210> 39

50 <211> 2331

<212> ADN

<213> Secuencia artificial

55 <220>

<223> secuencia de STAR39

60

65

ES 2 344 073 T3

<400> 39

5	tcaactgcaac ctccacctcc caggttcaag tgattctcct gcctcggcct cccgagtagc	60
	tgggactaca ggtgcatgac accgcacctg gctagttttt gtatttttag tagagacagg	120
	gtttcactat gttggccagg ttggtctcga actcctgacc ttgtgatccg cccacctcgg	180
10	cctcccaaag tgctgggatt acagagtgag ccactgcgcc tggcctgcac cccttactat	240
	tatatgcttt gcattttctt ttagatttga agaacctcat tataaactct agcactaatc	300
	ttatgtcagt taaatgcata gcaaatactt cctgacgtgg gagaatatat atttgcaagt	360
15	cttcttgatga acatatgttt tcagttctag ggagccagac gcctatgagt gaaaagccta	420
	gtcatcgtgg agaagtgcac tcaactttgt aagaaactgc caaaccttta ttcataatgg	480
	ttgtataaat ttacattac caccaataat gtatgagagt tccagttgct tcacatcctc	540
20	accagcattt tgttttgtct gtcttttttc ctttggttat tctagtgggc ataagatata	600
	atagtatccc ttgtggttta atgtaaattc cactgaagac taataacatt tgcataattc	660
	taattaataa gcctttttta gtgacttttc aagtctttgc tcatttttat tagatatattg	720
25	ccttcttatt attgatttga aagaattata tttatatgct tatattctgg ttataagccc	780
	tttgtcatta ttttccaaaa caatatttgg ttgtttctgt actactttcc ttgctccttt	840
	gaattgactt ggtgccttgg ccaaaaatca attgaccaca tacatgtggg tgcactctca	900
	gactaccaca ttccgtttat ctatttgtct ctccttggtg caataacact ctgtcttgat	960
30	aatggtaagt ttgagatca ggttggtgta gtcctcctaa tttttcctgg gttttcaata	1020
	ttgctttgct ttttaaaaaa tttgtatttt catttacatt ttaaaataaa cttgttagtg	1080
	ggattttgat tggcattgca ctgaactcgt ggatcaattt ggggagattg gacattctta	1140
35	tatatggatc ccgtggtcat caactttaag aactctttct catccattag taactcaatc	1200
	taggttcaga tgctactcgt tttctgctca gtctgtgtct gagcccccta tgctcttcat	1260
	tttgtcatcc aattaacctc agctttgcat caatactatt tcttgctttg gtgcctgtta	1320
40	cctctcctct aatcaccaat ccacaactta cctccaaatt cagggttggc ctcatctctc	1380
	ccaggaggag tgctgctcag tctatctact tagtattata atttctctgg cttgggtatca	1440
	aggcactccc atttccggct tccatgagat gtctcagagg gcatgctgcc cgggtgtagct	1500
45	gcatggtcaa gcttcttcat atctcttgcc tcatcactta aactcactat tttgtactcc	1560
	tgcttcagct atagggagct actgttagtt tcttgaagac atatgctctc tctctctctc	1620
	acatctggac ctgagcacat cctgttactg ctgcttgaaa caatgtgatc cccaggcaca	1680

50

55

60

65

ES 2 344 073 T3

caccattagc ttagaagcct cccctgattc ttcaaggctg gttgagtccc ttctctgtgc 1740
 tctcatgaca acagttggca attcctcggt gcagcaccta gcccattgat ctctttggag 1800
 5 gcagagactg agtctttctc actattgaat ttccagcatt catcacagag cctggcatat 1860
 ataaagccct ccatcatatg tattaagtga atggataaat gaaaaaaagt tatatatatg 1920
 tacatatatg tgtatatatg tatatgtata tatgtgtata tatgtgtgta tatgtgtgtg 1980
 10 tatatatgta catatatatg tatctatgta catatatgta tatatgtata tatatgtgtg 2040
 tgtatatgtg tgtgtgtatg tatatatatt acaatgaaat actattcagc cttaaaaagg 2100
 cagggaatcc tgtcatttaa cacaatatgg ataaacctag aggactctaa aggcaaatac 2160
 15 cacatgttct cactcacaaa atctaaacaa gttgaactcc tacaagtaga gagtaggatg 2220
 atggttacca agggctgggg gacgggagag gatggggaaa gcatagctgt ccatcaaagg 2280
 gtagaaagtt tcatttagac aagaggaatc agctttagt atctatttca c 2331
 20

<210> 40
 <211> 1071
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 30 <223> secuencia de STAR40
 <400> 40

35 gctgtgattc aaactgtcag cgagataagg cagcagatca agaaagcact ccgggctcca 60
 gaaggagcct tccaggccag ctttgagcat aagctgtgta tgagcagtga gtgtcttgag 120
 40 tagtgttcag ggcagcatgt taccattcat gcttgacttc tagccagtgt gacgagaggc 180
 tggagtccag tctctagaga gttgagcagc tccagcctta gatctcccag tcttatgcgg 240
 tgtgcccatt cgctttgtgt ctgcagtccc ctggccacac ccagtaacag ttctgggatc 300
 45 tatgggagta gcttccttag tgagctttcc cttcaaatac tttgcaacca ggtagagaat 360
 tttggagtga aggttttgtt cttcgtttct tcacaatatg gatatgcatc ttcttttgaa 420
 aatgttaaag taaattacct ctcttttcag atactgtctt catgcgaact tggatcctg 480
 50 tttccatccc agccttctat aaccagta catctttttt gaaaccagtg ggtgagaaag 540
 acacctggtc aggaacgcgg accacaggac aactcaggct caccacggc atcagactaa 600
 aggcaacaa ggactctgta taaagtaccg gtggcatgtg tattagtgga gatgcagcct 660
 55 gtgctctgca gacagggagt cacacagaca cttttctata atttcttaag tgctttgaat 720
 gttcaagtag aaagtctaac attaaatttg attgaacaat tgtatattca tggaatattt 780
 tggaacggaa taccaaaaaa tggcaatagt ggttctttct ggatggaaga caaacttttc 840
 60 ttgtttaaaa taaattttat tttatatatt tgaggttgac cacatgacct taaggataca 900
 tatagacagt aaactgggta ctacagtga gcaaattaac atatctacca tcgtacatag 960
 ttacattttt ttgtgtgaca ggaacagcta aaatctacgt atttaacaaa aatcctaaag 1020
 65 acaatacatt tttattaact atagccctca tgatgtacat tagatctcta a 1071

ES 2 344 073 T3

<210> 41

<211> 1135

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> secuencia de STAR41

10

<400> 41

	tgctcttggt gccaggctg cagtgaatg gcgctgtctc ggctcatcgc aacctccgcc	60
15	tcccagattc aagtgattct cctgcctcac cctcccaagt agctgggatt accagtatgc	120
	agcaacacgc ccggctaatt ttgtatttgt aatagagacg gggtttcttc atgttggtca	180
	ggctggcttc aaattcctgc cctcagggtga tctgcccacc ttggcctccc aaagtgtctg	240
20	gattacaggc atgagccact gtgcccggcc tgggctgggg cttttaaggg gactggagg	300
	tgaggggctg gaaaattggg agagttgatt ggtggggcaa gggggatgta atcatcagg	360
	tgtacaaact gcactcttgg tttagtacgc tcctcgtagg gtccttcgga gcagctcagt	420
25	cagtagctcc atcagtatac aggacccaaa ggaatatctc aaagggaata cagcatttcc	480
	taaggttcaa gttgtgatct acggagcagt taggggaact acaatcttgt gacagggtct	540
	acatgcttct gaggcaatga gacaccaagc agctacgagg aagcagtcag agagcacgcc	600
30	gacctagtga ctgatgctga tgtgctgcga gctgggttca ttttcatttc tcccctccc	660
	ctgccctcat taattttgta aagtttatag ggaacatttc accactctg ctgtggatcc	720
	ctgtcactta cggagtctgt catcttggct gtatgggctg tggcctctgc ggtgcccatt	780
35	ctcaggaggt gtgagacca tgaggaccgg aggtggacaa ggctagagac cacaccccc	840
	cgctccatcc aatcatgttt tcctgggtgc ttggtttcta tgcaggctgc atgtccttag	900
	tccctgcatg ggaacagctc ctgtggtgag caggcccctg aggaaggcct tgagcgggaa	960
40	tggagcctag gcttaggctg cctggtaaga gctggaggga accagccgag gcttgtgcta	1020
	cttttttttc cagaatgaaa tacgtgactg atgttgggtg cctgcagcgc cacgtttccc	1080
	gccacaacca ccggaacgag gatgaggaga acacactctc cgtggactgc acacg	1135

45 <210> 42

<211> 735

<212> ADN

50 <213> Secuencia Artificial

<220>

<223> secuencia de STAR42

55

<400> 42

	aagggtgaga tctactagga gggaggaagg agctataaaa gaaagaggtc actcatcaca	60
60	tcttacacac tttttaaaac cttgggtttt taatgtccgt gttcctcatt agcagtaagc	120
	cctgtggaag caggagtctt tctcattgac caccatgaca agaccctatt tatgaaacat	180
	aatagacaca caaatgttta tcggatattt attgaaatat aggaattttt cccctcacac	240
65	ctcatgacca cattctggtg cattgtatga atgaatatac cataatttta cctatggctg	300

ES 2 344 073 T3

	tatatttagg tcttttcgtg caggctataa aaatatgtat gggccggtca cagtgactta	360
	cgcccgtagt cccagaactt tgggaggccg aggcgggtgg atcacctgag gtcgggagtt	420
5	caaaaccagc ctgaccaaca tggagaaacc ccgtctctgc taaaaataca aaaattaact	480
	ggacacggtg gcgtatgcct gtaatcccag ctactcggga agctgaggca ggagaactgc	540
	ttgaaccag gaggcggagg ttgtggtgag tcgagattgc gccattgcac tccagcctgg	600
10	gcaacaagag cgaaattcca tctcaaaaaa aagaaaaaag tatgactgta tttagagtag	660
	tatgtggatt tgaaaaatta ataagtgttg ccaacttacc ttaggggtta taccatttat	720
	gaggggtgctg gtttc	735

15

<210> 43

<211> 1227

20 <212> ADN

<213> Secuencia Artificial

<220>

25 <223> secuencia de STAR43

<400> 43

30	caaatagatc tacacaaaac aagataatgt ctgccattt ttccaaagat aatgtggtga	60
	agtgggtaga gagaaatgca tccattctcc ccaccaacc tctgctaaat tgtccatgtc	120
35	acagtactga gaccagggg cttattccca gcgggcagaa tgtgcaccaa gcacctcttg	180
	tctcaatttg cagtctaggc cctgctattt gatggtgtga aggcttgac ctggcatgga	240
	aggtccgttt tgtacttctt gctttagcag ttcaaagagc agggagagct gcgaggccct	300
40	ctgcagcttc agatggatgt ggtcagcttg ttggaggcgc cttctgtggt ccattatctc	360
	cagccccct gcggtgttgc tgtttgcttg gcttgtctgg ctctccatgc cttgttggt	420
	ccaaaatgtc atcatgtctg accccaggaa gaatgtgcag gccatctct tttatgtgct	480
45	ttgggctatt ttgattcccc gttgggtata ttccctaggt aagaccaga agacacagga	540
	ggtagtgtct ttgggagagt ttggacctat gggtagagg taatagacac agtatcttct	600
	ctttcatttg gtgagactgt tagctctggc cgcgactga attccacaca gctcacttg	660
	gaaaacttta ttccaaaaca tagtcacatt gaacattgtg gagaatgagg gacagagaag	720
50	aggccctaga tttgtacatc tgggtgttat gtctataaat agaatgcttt ggtggtaac	780
	tagacttggt catgttgaca tttagtcttg ccttttcggt ggtgatttaa aaattatgta	840
55	tatcttgttt ggaatatagt ggagctatgg tgtggcattt tcatctggct tttgttttag	900
	ctcagcccgt cctgttatgg gcagccttga agctcagtag ctaatgaaga ggtatcctca	960
	ctccctccag agagcgggtcc cctcacggct cattgagagt ttgtcagcac cttgaaatga	1020
	gtttaaactt gtttattttt aaaacattct tggttatgaa tgtgcctata ttgaattact	1080
60	gaacaacctt atggttgtga agaattgatt tgggtgctaag gtgtataaat ttcaggacca	1140
	gtgtctctga agagttcatt tagcatgaag tcagcctgtg gcagggtggg tggagccagg	1200
65	gaacaatgga gaagctttca tgggtgg	1227

ES 2 344 073 T3

<210> 44

<211> 1586

<212> ADN

5 <213> Secuencia artificial

<220>

<223> secuencia de STAR44

10

<400> 44

	tgagttgggg tcctaagcca gaagttaact atgctttcat atattcttgc aagtagaagt	60
15	acagtgttgg tgtaaattcc ccttagatgg atagctaagc ccagaggaaa taatggtaat	120
	tggaaccata tgaccgtatg caattcatgt gcatatttat atcaagaaaa gaacattata	180
	ggtcgggtga gacctatctt tgttctgaca atgtcatctg tatttacatg tctgtttcgg	240
20	gagtttggat gtcaagggat tctgtgctgg attgtaaagc atgtgcttct gcttgatgta	300
	gctactcaat tttgtattct tgactaataa agtcataaac ataattcaac ctctgtgtgc	360
	gtgctctcct tccattaatt tatacttttag caaaaagtat tgaatgtgtg tgttatgtaa	420
25	caatttccta taaattatat taaatgattt attagcttta ttcaataaag ttttaagtgt	480
	tttcttctat gactacatta tttgttaaca agaaatttct ttaactgaaa acttcaagga	540
	agactatctg ggtaactctt tcaaaaagaa ttgtccctgt attttgggat tgaatatatt	600
30	aatttcttgt actgttttaa cagcacataa ttttacaaga caagccactt tttcaaagcc	660
	tgcttctcct cccattttcc ctatctctgt gattgacacc tccaaccctt gtagcctgcc	720
	tctgctctct ctttaaccagt cctactgata ctacttccta agtatttttc agccctgtcc	780
35	ttcctctcca tcatgatgga ttcacttcca gttgaaatcc ttatggtacc ctccctggat	840
	tatggcagta atcagagagc tgggtctcctt aactcaggat tcacttcttc tcatctgttg	900
	ttcacagtga catcagaaag atatttttaa atgatgaact agaattaatt atataaaca	960
40	cacatacaca cataaataat acttaaatctt ttcaatgatg ttccaattat gtaaaatata	1020
	atataggagg cactttatgt tctggcctca atctttcaat tcaaacttat ctccctgccac	1080
	tatctccttt gaacattgta ttccagctac tttagaataa taataatata taatattcat	1140
45	agagcccttc ctgggttcct atcacctgac aaaatacttc acatataaca tttaatcttt	1200
	gacaacttta ttaggcattg acaattatta tctatctata tatctatata tatatatata	1260
	aaatctatat tttatagata agaaaataga gggtaaaaac ttgccaaaat tacaaagctt	1320
50	agaagtgtag cagttgggat ttgaatctag gcatcctgcc tctatagtct acagtggctt	1380
	tcttgtgcc aagccttgc agttccctag acttaacatt tctcaaaatc tgtgtctttc	1440
	acatgctctt ccaattgtct ggaaaatctt tcccaacctc agtctaactg tggactcat	1500
55	gttcaccca caagaattga ctccatctgt cccctctcca tgaaaatttc tttgaatctc	1560
	agcactttgg gaggtgagg cagggtg	1586

<210> 45

60

<211> 1981

<212> ADN

<213> Secuencia artificial

65

<220>

<223> secuencia de STAR45

ES 2 344 073 T3

<400> 45

	cacgccccag	cgtgccctgg	actactgctc	cgcaggactc	ctgttctgct	gcaccctgga	60
5	ctacggcacc	agaggaccca	gctcccgccg	gcctgagcta	tggcaccaga	ggaccagct	120
	cccggcagcc	tggactatgg	caccagagga	cccagcccc	cgcttcctgg	gctaaggcac	180
	agtaggaccc	tgcctcatcg	tgtactcctg	ctcaggagga	ccctcgcagg	gcggcgcact	240
10	ggactaagct	actgaaggag	ccccaccctt	gcctaaccct	ggactaaggc	actggagaac	300
	tcttgctccg	cagagccacg	gactcttgca	caagagaacc	tcagcccagc	cgtgccctgg	360
	actgtggcac	agtagggccc	acaccacgcc	atggactcct	gtattggagg	aagagtagtg	420
15	ataaatgtcc	aggtttaca	cttgaaaagt	agcaatcaat	gtgccacaat	agatggatgt	480
	gatgtaaaat	tataaatgat	gaaaacatta	tgtgtaattg	cctagccaga	acagttacac	540
	aagacaaaga	cgtaaaagaa	atccacatag	ggaaggaaga	ggtaagattg	tttctgtttt	600
20	ttgaaaatat	aatcttaaga	tagagaaaat	cttaaagatt	ccacccaaat	aaatggttat	660
	agctgatgaa	gaaattcaat	aaagttaata	gttacaataa	caacatacaa	atatcattat	720
	tgtttctatt	aactaatgac	aaactattac	ctgaaaaata	aaggcaattc	aattttataat	780
25	agaatcaaaa	cagatatata	aatatataaa	agacaggagt	aaattttaatc	aaaaccataa	840
	aagatttaca	tactgaaaac	tatagcacat	tgatgaaaaa	aattaaaatg	gcataaataa	900
	atggagaaac	atccttcatt	gatggattca	aaaattagta	ttgtaaaagt	gtcaatgcta	960
30	cccaaagcaa	tctacagatt	aaatgcaacc	actatcaaat	tccaatgtca	ttcttcacag	1020
	aaatagaaaa	attactgcta	aaattttgat	ggaaccacaa	aagacctgga	ccaaccaaaag	1080
	caatcttgaa	caaaaagaac	aaagctggag	gcacagact	acctgactcc	aaactctatt	1140
35	acaaagctat	aggaattaaa	acagcatagc	aatggcataa	aaacagacat	gtaaaacagt	1200
	acaaagggat	atagaacctg	taaataaatc	cgtgtgtctg	tgggtcaattg	atTTTTtgat	1260
	aaaataacta	aaaatacaca	gtgaagaaag	aaaattatTT	tcaataaatg	gtgtagacaa	1320
40	aactgactat	ccacatacag	aagaataaaa	tttgactttt	atTTTgctct	ttatacaagc	1380
	atcaaatcaa	aattaaagtt	taaatgtaaa	actactacaa	ggaaatatag	aaggagactg	1440
	tatgacattg	gcctgagcta	tgattttctg	tagattattc	caaaaggcaa	caaaagcaaa	1500
45	acacacaaat	gagactgcat	aaaacttaaa	acttttccac	aggaaaagaa	gcaatgatag	1560
	aattaagaga	accacacaaat	gggataatat	ttttaaacca	tacatcaggt	aaggggctca	1620
	tataataata	tataagcaac	tcaacctact	caaaaataag	aaaaaaacta	tgcttattaa	1680
50	aaaataagca	aagaatcaga	atagacattt	cctacatcat	acaaaaggcc	aaccagggtac	1740
	atgaaaaaat	cataaacatt	cctaattatc	agagaagtgc	aatcaatgc	cacaatgaga	1800
	tatcacctca	cacattttac	tagggctatt	ataaaaaaag	atggaagata	agtgttggtg	1860
55	aggatgtgga	gaaaaagaaa	ccctgtacac	tgttggtagg	aatggaaatt	agtacagcca	1920
	tcttggaaaa	cagtacgaag	ctttctcaag	aaattataaa	tttatttacc	ctatgatcca	1980
60	t						1981

<210> 46

<211> 1859

<212> ADN

<213> Secuencia artificial

ES 2 344 073 T3

<220>

<223> secuencia de STAR46

5 <400> 46

	attgtttttc tgcgcccttct gcatttttctg caaattctgt tgaatcattg cagttactta	60
	ggtttgcttc gtctcccca ttacaaacta cttactgggt ttttcaacc tagttccctc	120
10	atTTTTatga tttatgtcga tttctttgta cacttcgtct tgctccatct cccaactcat	180
	ggccccctggc tttggattat tgttttggtc ttttattttt tgtcttcttc tacctcaaca	240
	cttatcttcc tctcccagtc tccgggtaccc tatcaccaag gttgtcatta acctttcata	300
15	ttattcctca ttatccatgt attcatttgc aaataagcgt atattaacaa aatcacagggt	360
	ttatggagat ataattcaca taccttaaaa ttcaggcttt taaagtgtac ctttcatgtg	420
	gtttttggtg tattcacaaa gttatgcatt gatcaccacc atctgattcc ataacatgtt	480
20	caatacctca aaaaagaagtc tgtactcatt agtagtcatt tcacattcac cactccctct	540
	ggctctgggc agtactgat ctttgtgtct ctatggattt gcctagtcta ggtattttta	600
	tgtaaatggc atcatacaac atgtgacctt ttgtttggct tttttcattt agcaaaatgt	660
25	tatcaaggtc tgtccctgtt gtagcatgta ttagcacttc atttcttata tgctgaatga	720
	tatactttat ttgtccatca gttgttcattg ctttatttgt ccatcagttg atgaacattt	780
	gcgtttttgc cactttgggc tattaagaat aatgctactg tgaacaagtg tgtacaagtt	840
30	cctctacaaa tttttgtgtg gacatatcct ttcagttctc tcagggtgat atctgggaat	900
	tgaattgctg ggtcgtgtag tagctatgtt aaacactttg agaaactgct ataatgttct	960
	ccagagctgt accattttta attctgtgta tgaggattcc acgttctcca cttcctcacc	1020
35	agtgtatgga tttgggggta tactttttta aaagtgggat taggctgggc acagtggctc	1080
	acacctgtaa tcccaacact tcaggaagct gaggtgggag gatcacttga gcctagtagt	1140
	ttgagaccag cctgggcaac atagggagac cctgtctcta caaaaaataa tttaaaataa	1200
40	attagctggg cgttgtggca cacacctgta gtcccagcta catgggaggc tgagggtggaa	1260
	ggattccctg agcccagaag tttgaggttg cagtgaacca tgatggcagc actatactgt	1320
	agcctgggtg tcagagcaag actccgtttc agggaagaaa aaaaaaagtg ggatgatatt	1380
45	tttgacactt ttcttcttgt tttcttaatt tcatacttct ggaaattcca ttaaattagc	1440
	tggtaccact ctaactcatt gtgtttcatg gctgcatagt aatattgcat aatataaata	1500
	taccattcat tcatcaaagt tagcagatat tgactgttag gtgccaggca ctgctctaag	1560
50		
	cgttaaagaa aaacacacaa aaacttttgc attcttagag tttattttcc aatggagggg	1620
	gtggagggag gtaagaattt aggaaataaa ttaattacat atatagcata gggtttcacc	1680
55	agtgagtgca gcttgaatcg ttggcagctt tcttagtagt ataaatacag tactaaagat	1740
	gaaattactc taaatggtgt tacttaaaatt actggaatag gtattactat tagtcacttt	1800
	gcagggtgaaa gtggaaacac catcgtaaaa tgtaaaatag gaaacagctg gttaatgtt	1859

60

<210> 47

<211> 1082

65 <212> ADN

<213> Secuencia artificial

ES 2 344 073 T3

<220>

<223> secuencia de STAR47

5 <400> 47

	atcattagtc attagggaaa tgcaaatgaa aaacacaagc agccaccaat atacacctac	60
10	taggatgatt taaaggaaaa taagtgtgaa gaaggacgta aagaaattgt aaccctgata	120
	cattgatggt agaaatggat aaagttgcag ccactgtgaa aaacagtctg cagtggctca	180
	gaaggttaaa tatagaaccc ctgttggacc caggaactct actcttaggc accccaaaga	240
15	atagagaaca gaaatcaaac agatgtttgt atactaatgt ttgtagcatc acttttcaca	300
	ggagccaaaa ggtggaaata atccaacccat cagtgaacaa atgaatgtaa taaaagcaag	360
	gtggtctgca tgcaatgcta catcatccat ctgtaaaaaa cgaacatcat tttgatagat	420
20	gatacaacat ggggtggacat tgagaacatt atgcttagtg aaataagcca gacacaaaag	480
	gaatatattg tataattgta attacatgaa gtgcctagaa tagtcaaatt catacaagag	540
	aaagtgggat aggaatcacc atgggctgga aataggggga aggtgctata ctgcttattg	600
	tggacaaggt ttcgtaagaa atcatcaaaa ttgtgggtgt agatagtggg gttggttatg	660
25	caaccctgtg aatatattga atgccatgga gtgcacactt tggttaaaag gttcaaatga	720
	taaatattgt gttatatata tttccccacg atagaaaaca cgcacagcca agcccacatg	780
	ccagtcttgt tagctgcctt cctttacctt caagagtggg ctgaagcttg tccaatcttt	840
30	caaggttgct gaagactgta tgatggaagt catctgcatt gggaaagaaa ttaatggaga	900
	gaggagaaaa cttgagaatc cacactactc accctgcagg gccaagaact ctgtctccca	960
	tgctttgctg tcctgtctca gtatttcctg tgaccacctc ctttttcaac tgaagacttt	1020
35	gtacctgaag gggttcccag gtttttcacc tcggcccttg tcaggactga tcctctcaac	1080
	ta	1082

40

<210> 48

<211> 1242

<212> ADN

45

<213> Secuencia artificial

<220>

50

<223> secuencia de STAR48

55

60

65

ES 2 344 073 T3

<400> 48

	atcatgtatt tgttttctga attaatcttt agatacatta atgttttatg ttaccatgaa	60
5	tgtgatatta taatataata tttttaattg gttgctactg tttataagaa tttcattttc	120
	tgtttacttt gccttcatat ctgaaaacct tgctgatttg attagtgcac ccacaaattt	180
	tcttgatttt tctatgggta attacaaatc tccacacaat gaggttgacg tgagccaaga	240
10	tcacaccact gtactccagc ctgggagaca gagtgagaca ccattctaca aaaacacata	300
	aacaaacaaa cagaaactcc acacaatgac aacgtatgtg ctttcttttt ttcttctctt	360
	ttctataata tttctttgtc ctatcttaac tgaactggcc agaaacccca ggacaatgat	420
15	aaatacgagc agtgtcaaca gacatctcat tccctttcct agcttttata aaaataacga	480
	ttatgcttca acattacata tgggtggtgtc gatggttttg ttatagataa gcttatcagg	540
	ttaagaaatt tgtctgcgtt tcctagtgtg gtataaagat tttaatataa atgaatgttg	600
20	tattttatca tcttattttt ttctacatc tgctaaggta atcctgtgtt tccccctttt	660
	caatctccta atgtgggtgaa tgacattaaa ataccttcta ttgttaaaat attcttgcaa	720
	cgctgtatag aaccaatgcc tttattctgt attgctgatg gatttttgaa aaatatgtag	780
25	gtggacttag ttttctaagg ggaatagaat ttctaataa tttaaaatat ttgcatgta	840
	tgttctgaag gacattggtg tgtcatttct ataccatctg gctactagag gagccgactg	900
	aaagtcacac tgccggagga ggggagaggt gctcttccgt ttctggtgtc tntagccatc	960
30	tccagtggta gctgcagtga taataatgct gcagtgccga cagttctgga aggagcaaca	1020
	acagtgattt cagcagcagc agtattgcgg gatccccacg atggagcaag ggaaataatt	1080
	ctggaagcaa tgacaatatc agctgtggct atagcagctg agatgtgagt tctcacggtg	1140
35	gcagcttcaa ggacagtagt gatggtccaa tggcgccag acctagaaat gcacatttcc	1200
	tcagcaccgg ctccagatgc tgagcttgga cagctgacgc ct	1242

<210> 49

40 <211> 1015

<212> ADN

<213> Secuencia artificial

45 <220>

<223> secuencia de STAR49

<400> 49

50	aaaccagaaa cccaaaacaa tgggagtgac atgctaaaac cagaaaccca aaacaatggg	60
	aggggtcctgc taaaccagaa acccaaaaca atgggagtgga agtgctaaaa ccagaaaccc	120
	aaaacaatgg gagtgtcctg ctacaccaga aacccaaaac gatgggagtg acgtgataaa	180
55	accagacacc caaaacaatg ggagtgacgt gctaaaccag aaacccaaaa caatgggagt	240
	gacgtgctaa aacctggaac cctaaaacaa tgcgagtgag gtgctaacac cagaatccat	300
	aacaatgtga gtgacgtgct aaaccagaac ccaaaacaat gggagtgacg tgctaaaaca	360
60	ggaacccaaa acaatgagag tgacgtgcta aaccagaaac ccaaaacaat gggaatgacg	420
	tgctaaaacc ggaacccaaa acaatgggag tgatgtgcta aaccagaaac ccaaaacaat	480

65

ES 2 344 073 T3

	gggaatgaca tgctaaaact ggaacccaaa acaatggtaa ctaagagtga tgctaaggcc	540
	ctacattttg gtcacactct caactaagtg agaacttgac tgaaaaggag gatttttttt	600
5	tctaagacag agttttggtc tgtcccccag agtggagtgc agtggcatga tctcggctca	660
	ctgcaagctc tgcctcccgg gttcaggcca ttctcctgcc tcagcctcct gagtagctgg	720
	gaatacaggc acccgccacc acacttggct aattttttgt atttttagta gagatggggg	780
10	ttcaccatat tagcaaggat ggtctcaatc tcctgacctc gtgatctgcc cacctcaggc	840
	tcccaaagtg ctgggattac aggtgtgagc caccacaccc agcaaaaagg aggaattttt	900
	aaagcaaaat tatgggaggc cattgttttg aactaagctc atgcaatagg tcccaacaga	960
15	ccaaacaaaa ccaaacaaaa atggagtcac tcatgctaaa tgtagcataa tcaaa	1015

<210> 50
 20 <211> 2355
 <212> ADN
 <213> Secuencia Artificial
 25 <220>
 <223> secuencia de STAR50
 <400> 50

30	caaccatcgt tccgcaagag cggcttgttt attaaacatg aaatgagggg aaagcctagt	60
	agctccattg gattgggaag aatggcaaaag agagacaggc gtcattttct agaaagcaat	120
35	cttcacacct gttggtcctc acccattgaa tgtcctcacc caatctccaa cacagaaatg	180
	agtgactgtg tgtgcacatg cgtgtgcatg tgtgaaagta tgagtgtgaa tgtgtctata	240
	tgggaacata tatgtgattg tatgtgtgta actatgtgtg actggcagcg tggggagtgc	300
40	tgggttgagt gtggtgtgat gtgagtatgc atgagtggct gtgtgtatga ctgtggcggg	360
	aggcggaagg ggagaagcag caggctcagg tgtcgccaga gaggctggga ggaaactata	420
	aacctgggca atttcctcct catcagcgag cctttcttgg gcaatagggg cagagctcaa	480
45	agttcacaga gatagtgcct gggaggcatg aggcaaggcg gaagtactgc gaggaggggc	540
	agagggctctg acacttgagg ggttctaata ggaaaggaaa gaccacact gaattccact	600
	tagccccaga ccctgggccc agcggtgccg gcttccaacc ataccaacca tttccaagt	660
50	ttgccggcag aagttaacct ctcttagcct cagtttcccc acctgtaaaa tggcagaagt	720
	aaccaagctt accttcccgg cagtgtgtga ggatgaaaag agctatgtac gtgatgcact	780
	tagaagaagg tctaggggtg gagtgggtact cgtctggtgg gtgtggagaa gacattctag	840
55	gcaatgagga ctggggagag cctggcccat ggcttccact cagcaaggct agtctcttgt	900
	cctctgcact cccagccttc cagagaggac cttccaacc agcactcccc acgctgccag	960
	tcacacatag ttacacacat acaatcacat atatgttccc atatagacac attcacactc	1020
60	ataccttcac acatgcacac gcatgtgcac acacagtcac tcatttctgt gttggagatt	1080
	gggtgaggac attcaatggg tgaggaccaa cagggtgtgaa gattgcttcc tagaaaatga	1140
	ctcctgtctc tctttgccat tcttccaat ccgatggagc tactaggctt ttccctcatt	1200

65

ES 2 344 073 T3

	tcattgttttaa taaaccttcc caatggcgaa atgggctttc tcaagaagtg gtgagtgtcc	1260
	catccctgcg gtggggacag ggggtggcagc ggacaagcct gcctggaggg aactgtcagg	1320
5	ctgattccca gtccaactcc agcttccaac acctcatcct ccaggcagtc ttcattcttg	1380
	gctctaattt cgctcttggt ttctttttta tttttatcga gaactgggtg gagagctttt	1440
	gggtgtcattg gggattgctt tgaaaccctt ctctgcctca cactgggagc tggcttgagt	1500
10	caactgggtct ccatggaatt tcttttttta gtgtgtaaac agctaagttt taggcagctg	1560
	ttgtgccgtc cagggtggaa agcagcctgt tgatgtggaa ctgcttggtc cagatttctt	1620
	gggcaaacag atgccgtgtc tctcaactca ccaattaaga agcccagaaa atgtggcttg	1680
15	gagaccacat gtctggttat gtctagtaat tcagatggct tcacctggga agccctttct	1740
	gaatgtcaaa gccatgagat aaaggacata tatatagtag ctagggtggg ccacttctta	1800
	ggggccatct ccggaggtgg tgagcactaa gtgccaggaa gagaggaaac tctgttttgg	1860
20	agccaaagca taaaaaacc ttagccacaa accactgaac atttgttttg tgcaggttct	1920
	gagtcaggag agggcttctg aggagagggg cagctggagc tggtaggagt tatgtgagat	1980
	ggagcaaggg ccctttaaga ggtgggagca gcatgagcaa aggcagagag gtggtaatgt	2040
25	ataagggtatg tcatgggaaa gagtttggct ggaacagagt ttacagaata gaaaaattca	2100
	acactattaa ttgagcctct actacgtgct cgacattggt ctagtcactg agataggttt	2160
	ggtatacaaa acaaaatcca tcctctatgg acattttagt gactaacaac aatataaata	2220
30	ataaaagtga acaaaagctc aaaacatgcc aggcactatt atttatttat ttatttattt	2280
	atttatttat tttttgaaac agagtctcgc tctgttgccc aggctggagt gtagtggtgc	2340
	gatctcggct cactg	2355

35

<210> 51

<211> 2289

<212> ADN

40

<213> Secuencia Artificial

<220>

<223> secuencia de STAR51

45

<400> 51

50	tcacagggtga caccaatccc ctgaccacgc tttgagaagc actgtactag attgactttc	60
	taatgtcagt cttcattttc tagctctggt acagccatgg tctccatatt atctagtaca	120
	acacacatac aaatatgtgt gatacagtat gaatataata taaaaatatg tgttataata	180
55	taaatataat attaaaatat gtcttttatac tagataataa tacttaataa cgttgagtgt	240
	ttaactgctc taagcacttt acctgcagga aacagttttt tttttatttt ggtgaaatac	300
	aactaacata aattttattta caattttaag catttttaag tgtatagttt agtggagtta	360
	atatattcaa aatgtttgtc agccgtcacc atcatcagtc ttcataactc ttttcatatt	420
60	gtaaaattaa aagtttatgc tcatttaaaa atgactccca atttcccccc tcctcaacct	480
	ctggaaacta ccattctatt ttctgcctcc gtagttttgc ccactctaag tacctcacat	540

65

ES 2 344 073 T3

	aagtggaatt tgtcttattt gcctgtttgt gaccggctga tttcatttag tataatgtcc	600
	tcaagtttta ttcacgttat atagcatatg tcataatattt cttcactttt aagcttgagt	660
5	aatatttcat cgtatgtatc tcacattttg cttatccatt catctctcag tggacacttg	720
	agttgcttct acatttttagc tgttggaat actgctgcta tgaacatggg tgtataaata	780
	tctcaagacc tttttatcag ttttttaaaa tatatactca gtagtagttt agctggatta	840
10	tatggtaatt ttatttttaa tttttgagga actgtcctac ctttttattc aatagtagct	900
	ataccaattg acaattggca ttcctaccaa cagggcataa gggttctcaa ttctccacat	960
	attccctgat acttggttatt ttcaggtggt tttttttttt tttttttttt atgggagcca	1020
15	tgttaatggg tgtaagggtga tatttcatta tagttttgat ttgcatttcc ctaatgatta	1080
	gtgatgttaa gcatctcttc atgtgcctat tggccatttg tatatcttct ttaaaaatat	1140
	atatatactc attcctttgc ccatttttga attatgttta ttttttgta ttgagtttca	1200
20	atacttttct atataacctt ggattaatc ctttatcaga ctttaagattt gcaaatattc	1260
	tctttcattc cacaggttgc taattctctc tgttggaat atcttttgat gctgttggt	1320
	ccagaattga ttcattcctg tgggttcttg gtctcactga cttcaagaat aaagctgcgg	1380
25	accctagtgg tgagtgttac acttcttata gatggtgttt ccggagtgtt ttccttcaga	1440
	tgtgtccaga gtttcttcct tccaatgggt tcatggtctt gctgacttca ggaatgaagc	1500
	cgcagacctt cgcagtgagg tttacagctc ttaaagggtg cgtgtccaga gttgtttgtt	1560
30	ccccctggtg ggttcgtggt cttgctgact tcaggaatga agccgcagac cctcgcagtg	1620
	agtgttacag ctcataaagg tagtgcgagc acagagttag ctgcagcaag atttactgtg	1680
	aagagcaaaa gaacaaagct tccacagcat agaaggacac cccagcgggt tcctgctgct	1740
35	ggctcaggtg gccagttatt attcccttat ttgccctgcc cacatcctgc tgattggtcc	1800
	attttacaga gtactgattg gtccatttta cagagtgtg attggtgcat ttacaatcct	1860
	ttagctagac acagagtgtt gattgctgca ttcttacaga gtgctgattg gtgcatttac	1920
40	agtccttttag ctagatacag aacgctgatt gctgcgtttt ttacagagtg ctgattggtg	1980
	catttacaat ccttttagcta gacacagtgc tgattggtgg gtttttacag agtgctgatt	2040
	ggtgcgtctt tacagagtgc tgattggtgc atttacaatc ctttagctag acacagagtg	2100
45	ctgattggtg cgtttataat cctctagcta gacagaaaag tttccaagt cccacactga	2160
	ccgagaagcc cactggctt cacctctcac tgttatactt tggacatttg tcccccaaa	2220
	atctcatgtt gaaatgtaac ccctaattgtt ggaactgagg ccagactgga tgtggctggg	2280
50	ccatgggga	2289

<210> 52

<211> 1184

55 <212> ADN

<213> Secuencia Artificial

<220>

60 <223> secuencia de STAR52

65

ES 2 344 073 T3

<400> 52

	cttatgccat ctggcgggtgc catgtggaac ttcgctgaag aagctaaatt tactgaccat	60
5	ctgtgcctag agcgggtttc tccaaggaaa ggctctgtaa atctcgtcct tttgaaatct	120
	aggggaaaac agcctccttc actgaggatt aatttaaaga aagggggaaa taggaaaatt	180
	ccatgcgttg gaagtccatt tagatttcta catgaaccat catatatgtg cactacataa	240
10	ttcttatttt tttattttta aaaaagggat aatttatatt ccagtgacaa gtttgggaaa	300
	ggccaaggca agcaattgag ttgaacatta tgtagcgttt atatagacct tgcagacgtc	360
	tgtgcaatat ccaccactga acacgtgagg tcgtactcaa gtctctctgg cccctggtaa	420
15	tgtgactccc ttcctttatt tgcatgaatc gcctggattg ggtgtcaggt ttttaaacg	480
	tcaaggttta cgcctattgt tgtcaaccaa tcagcatcct actttgacgt gattggcttc	540
	tactgtaggt gtcaatcatc caaaatttgc atactactcc tcaggccgcc gggagcctgt	600
20	cagtcggctg tggcagctgg aagagaagga atcggacgga gaagaatgaa aaatcacttt	660
	gctttcgcaa agcgaaagaa aagtattctt ttcctcatta tttttaaata aatttgattg	720
	tatatattacc taataaaata aacattcaat taaacaaaaa taagcaacta tcaaagattt	780
25	gtttactaat tttcgtaatg tttactgttt caataagtag ccaaaggaat attaaacac	840
	aaaaatatga atgctgataa ttttatgtca taaagaccat tttaaaacta aaagtgaaca	900
	tggggtttct aaataaaatt accgtggtag cgtaaaaaca ctgctttcaa tacttgggca	960
30	tgctgaaagt gctgcatcct aagataaaaa atacaccaag ggggggattt caaagaacat	1020
	tattttgctt ttaataatcc tgtatttctg tcactttgcc ctttttattt atttaccgtg	1080
	aactcacaga cagaatatta cttggagttt ctgaaatact tgtgtttgta catttctcat	1140
35	cttacacgta cccacacacc ccaaaataaa aaaacaaaga agag	1184

<210> 53

<211> 1431

<212> ADN

<213> Secuencia artificial

<220>

<223> secuencia de STAR53

<400> 53

	ccctgaggaa gatgacgagt aactccgtaa gagaaccttc cactcatccc ccacatccct	60
50	gcagacgtgc tattctgtta tgatactggt atcccatctg tcacttgctc cccaaatcat	120
	tcccttctta caattttcta ctgtacagca ttgaggctga acgatgagag atttcccatg	180
	ctctttctac tccctgccct gtatatatcc ggggatcctc cctacccagg atgctgtggg	240
55	gtcccaaacc ccaagtaagc cctgatatgc gggccacacc tttctctagc ctaggaattg	300
	ataaccagg cgaggaagtc actgtggcat gaacagatgg ttcacttcca ggaaccgtgg	360
	aaggcgtgtg caggtcctga gatagggcag aatcggagtg tgcagggctc gcaggtcagg	420
60	aggagttgag attgcgttgc cacgtggtgg gaactcactg ccacttattt ctttctctct	480

65

ES 2 344 073 T3

tcttgcctca gcctcagggg tacgacacat gcccatgatg agaagcagaa cgtggtgacc 540
 ttccacgaac atgggcatgg ctgaggaccc ctctcatca ggtgcatagc aagtgaagc 600
 5 aagtgttcac aacagtgaag agttgagcgt ctttttctt agtgtgcaa gagttcgatg 660
 ttagcgttta cgttgatttt tcttacctg tgctattctg ttagatacta acattttcat 720
 tgatgagcaa gacatactta atgcatattt tggtttgtgt atccatgcac ctaccttaga 780
 10 aaacaagtat tgtcggttac ctctgcatgg aacagcatta cctcctctc tccccagatg 840
 tgactactga gggcagttct gagtgtttaa ttccagattt ttccctctgc atttacacac 900
 acacgcacac aaaccacacc acacacacac acacacacac acacacacac acacacacac 960
 15 acacaccaag taccagtata agcatctgcc atctgctttt cccattgcca tgcgtcctgg 1020
 tcaagctccc ctactctgt ttcctgggtc gcatgtactc cctcatccg attcccctgt 1080
 agcagtcact gacagttaat aaacctttgc aaacgttccc cagttgtttg ctctgacct 1140
 20 tattgtgcac acagctctgt gcacgtgtgt gcatatttct ttaggaaaga ttcttagaag 1200
 tggaattgct gtgtcaaagg agtcatttat tcaacaaac actaatgagt gcgtcctcgt 1260
 gctgagcgt gttctagggt ctggagcgac gtcagggaac aaggcagaca ggagttcctg 1320
 25 acccccgttc tagaggagga tgtttccagt tgttggtttt tgtttgtttg tttcttctag 1380
 agatggtggt cttgctctgt ccaggctaga gtgcagtggc atgatcatag c 1431

30 <210> 54
 <211> 975
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <223> secuencia de STAR54
 40
 <400> 54

ccataaaagt gtttctaaac tgcagaaaaa tccccctaca gtcttacagt tcaagaattt 60
 45 tcagcatgaa atgcctggtg gattacctga ctttttttgc caaaaataag gcacagcagc 120
 tctctcctga ctctgacttt ctatagtcct tactgaatta tagtccttac tgaattcatt 180
 cttcagtggt gcagctctgaa ggacaccac attttctctt tgtctttgtc aattctttgt 240
 50 gttgtaaggg caggatgttt aaaagttgaa gtcattgact tgcaaaatga gaaatttcag 300
 agggcatttt gttctctaga ccatgtagct tagagcagtg ttcacactga ggttgctgct 360
 aatgtttctg cagttcttac caatagtatc atttaccag caacaggata tgatagagga 420
 55 cttcgaaaac cccagaaaat gttttgccat atatccaaag ccctttggga aatggaaagg 480
 aattgctggc tcccattttt atatatggat agatagagac caagaaagac caaggcaact 540
 ccatgtgctt tacattaata aagtacaaaa tgtaacatg taggaagtct aggcgaagtt 600
 60 tatgtgagaa ttcttttacac taattttgca acattttaat gcaagtctga aattatgtca 660
 aaataagtaa aaatttttac aagttaagca gagaataaca atgattagtc agagaaataa 720
 gtagcaaaat cttcttctca gtattgactt ggttgctttt caatctctga ggacacagca 780

65

ES 2 344 073 T3

5 **gtcttcgctt ccaaatccac aagtcacatc agtgaggaga ctcagctgag actttggcta 840**
 atgttggggg gtccctcctg tgtctcccca ggcgcagtga gcctgcaggc cgacctcact 900
 cgtggcacac aactaaatct ggggagaagc aacccgatgc cagcatgatg cagatatctc 960
 agggatatgat cggcc 975

10

<210> 55

<211> 501

<212> ADN

15

<213> Secuencia Artificial

<220>

20

<223> secuencia de STAR55

<400> 55

25

cctgaactca tgatccgccc acctcagcct cctgaagtgc tgggattaca ggtgtgagcc 60
 accacaccca gccgcaacac actcttgagc aaccaatgtg tcataaaaga aataaaatgg 120
 aaatcagaaa gtatcttgag acagacaaaa atggaaacac aacataccaa aatttatggg 180
 acacagcaaa agcagtttta ggagggaggt ttatagtgat gaatacctac ctcaaaatca 240
 ttagcctgat tggatgacac tacagtgtat aaatgaattg aaaaccacat tgtgccccat 300
 acatatatac aatTTTTatt tgTTaattaa aaataaaata aaactTTaaa aaagaagaaa 360
 gagctcaaat aaacaaccta actTTatacc tcaaggaaat agaagagcca gctaagccca 420
 aagttgacag aaggaaaaaa atattggcag aaagaaatga aacagagact agaaagacaa 480
 ttgaagagat cagcaaaact a 501

40

<210> 56

<211> 741

45

<212> ADN

<213> Secuencia artificial

<220>

50

<223> secuencia de STAR56

55

60

65

ES 2 344 073 T3

<400> 56

5	acacaggaaa agatcgcaat tgttcagcag agctttgaac cggggatgac ggtctccctc	60
	gttgcccggc aacatggtgt agcagccagc cagttatttc tctggcgtaa gcaataccag	120
	gaaggaagtc ttactgctgt cgccgccgga gaacagggtg ttcctgcctc tgaacttgct	180
	gccgccatga agcagattaa agaactccag cgcttgcctg gcaagaaaac gatggaaaat	240
10	gaactcctca aagaagccgt tgaatatgga cgggcaaaaa agtggatagc gcacgcgccc	300
	ttattgcccg gggatgggga gtaagcttag tcagccggtg tctccgggtg tcgcgtgcgc	360
	agttgcacgt cattctcaga cgaaccgatg actggatgga tggccgccgc agtcgtcaca	420
15	ctgatgatac ggatgtgctt ctccgtatac accatgttat cggagagctg ccaacgtatg	480
	gttatcgctg ggtatgggag ctgcttcgca gacaggcaga acttgatggt atgcctgcga	540
	tcaatgccaa acgtgtttac cggatcatgc gccagaatgc gctgttgctt gāgcgaaaac	600
20	ctgctgtacc gccatcgaaa cgggcacata caggcagagt ggccgtgaaa gaaagcaatc	660
25	agcgatgggtg ctctgacggg ttcgagttct gctgtgataa cggagagaga ctgcgtgtca	720
	cgttcgcgct ggactgctgt g	741

30 <210> 57

<211> 1365

<212> ADN

35 <213> Secuencia Artificial

<220>

<223> secuencia de STAR57

40

45

50

55

60

65

ES 2 344 073 T3

<400> 57

5	tccttctgta aataggcaaa atgtatttta gtttccacca cacatgttct tttctgtagg	60
	gcttgtatgt tggaaatfff atccaattat tcaattaaca ctataccaac aatctgctaa	120
	ttctggagat gtggcagtga ataaaaaagt tatagtttct gattttgtgg agcttggact	180
10	ttaatgatgg acaaaacaac acattcttaa atatatafff catcaaaatt atagtgggtg	240
	aattatttat atgtgcattt acatgtgtat gtatacataa atgggcggtt actggctgca	300
	ctgagaatgt acacgtggcg cgaacgaggc tgggcggtca gagaaggcct cccaaggagg	360
15	tggctttgaa gctgagtggg gcttccacgt gaaaaggctg gaaagggcat tccaagaaaa	420
	ggctgaggcc agcgggaaag aggttccagt gcgctctggg aacggaaagc gcacctgcct	480
	gaaacgaaaa tgagtgtgct gaaataggac gctagaaagg gaggcagagg ctggcaaaag	540
20	cgaccgagga ggagctcaaa ggagcgagcg gggaaggccg ctgtggagcc tggaggaagc	600
	acttcggaag cgcttctgag cgggtaaggc cgctgggagc atgaactgct gagcaggtgt	660
	gtccagaatt cgtgggttct tgggtctact gacttcaaga atgaagaggg accgcggacc	720
25	ctcgcggtga gtgttacagc tcttaagggt gcgcgtctgg agtttgttcc ttctgatgtt	780
	cggatgtgtt cagagtttct tccttctggg ggggttcgtg tctcgtggc tcaggagtga	840
	agctgcagac cttcgcggtg agtggttacag ctcataaaag cagggtggac tcaaagagtg	900
	agcagcagca agattttattg caaagaatga aagaacaaag cttccacact gtggaagggg	960
30	acccagcgg gttgccactg ctggctccgc agcctgcttt tattctctta tctggcccca	1020
	cccacatcct gctgattggg agagccgaat ggtctgtttt gacggcgctg attggtgcgt	1080
	ttacaatccc tgcgctagat acaaagggtt tccacgtccc caccagatta gctagataga	1140
35	gtctccacac aaaggttctc caaggcccca ccagagtagc tagatacaga gtgttgattg	1200
	gtgcattcac aaaccctgag ctagacacag ggtgatgact ggtgtgttta caaaccttgc	1260
	ggtagatata gagtatcaat tggcgtattt acaatcactg agctaggcat aaagggttctc	1320
40	cagggtcccca ccagactcag gagcccagct ggcttcaccc agtgg	1365

<210> 58

45 <211> 1401

<212> ADN

<213> Secuencia artificial

50 <220>

<223> secuencia de STAR58

55

60

65

ES 2 344 073 T3

<400> 58

	aagtttacct tagccctaaa ttatttcatt gtgattggca ttttaggaaa tatgtattaa	60
5	ggaatgtctc ttaggagata aggataacat atgtctaaga aaattatatt gaaatattat	120
	tacatgaact aaaatgtag aactgaaaaa aaattattgt aactccttc agcgtaggca	180
	ggagtatcta gataccaact ttaacaactc aactttaaca acttcgaacc aaccagatgg	240
10	ctaggagatt cacctattta gcatgatatc ttttattgat aaaaaaatat aaaacttcca	300
	ttaaattttt aagctactac aatcctatta aattttaact taccagtgtt ctcaatgcta	360
	cataatttaa aatcattgaa atcttctgat ttaactcct cagtcttgaa atctacttat	420
15	ttttagttag atatatatcc aatctactgc cgctagtaga agaagcttg aatttgagaa	480
	aaaaatcaga cgttttgtat attctcatat tcactaattt attttttaaa tgagtttctg	540
	caatgcatca agcagtggca aaacaggaga aaaattaaaa ttggttgaaa agatatgtgt	600
20	gccaacaat cccttgaaat ttgatgaagt gactaatcct gagttattgt ttcaaatgtg	660
	tacctgttta tacaagggtg tcacctttga aatctcaaca ttaaatgaaa ttttataagc	720
	aatttggtgt aacatgatta ttataaaatt ctgatataac attttttatt acctgtttag	780
25	agtttaaaga gagaaaagga gttaagaata attacatttt cattagcatt gtccgggtgc	840
	aaaaacttct aacactatct tcaaatcttt ttctccattg ccttctgaac ataccactt	900
	gggtatctca ttagcactgc aaattcaaca ttttcgattg ctaatttttc tccctaaata	960
30	tttatttggt ttctcagctt tagccaatgt ttactattg accatttgct caagtatagt	1020
	gacgcttcaa tgacctcag agagctgttt cagtccttc tggactactt gcatgcttc	1080
	aacaaaatga agcactcttg atgtcagtca ctcaataaa tggaaatggg cccatttact	1140
35	aggaatgtta acagaataaa aagatagacg tgacaccagt tgcttcagtc catctccatt	1200
	tacttgctta aggcctggcc atatttctca cagttgatat ggcgcagggc acatgtttaa	1260
	atggctgttc ttgtaggatg gtttgactgt tggattcctc atcttccctc tccttaggaa	1320
40	ggaagggttac agtagtactg ttggctcctg gaatatagat tcataaagaa ctaatggagt	1380
	atcatctccc actgctctg t	1401

<210> 59

45 <211> 866

<212> ADN

<213> Secuencia artificial

50 <220>

<223> secuencia de STAR59

<400> 59

55	gagatcacgc cactgcactc cagcctgggg gacagagcaa gactccatct cagaaacaaa	60
	caaacacaca aagccagtca aggtgtttta ttcgacgggt tcaggctcag gtctcttgac	120
	aggatacatc cagcaccggg gggaaacgtc gatgggtggg gtggaatcta ttttggtggc	180
60	tcaaggagg gtttgagagg tagtcccga agcgggtgat gcctaaggaa gccctccgc	240

65

ES 2 344 073 T3

ccaagaagcg atattcattt ctagcctgta gccaccaag agggagaatc gggctcgcca 300
 cagacccac aaccccaac ccacccacc cccaccctc ccacctcgtg aaatgggctc 360
 5 tcgctccgtc aggtcttagt cacaccgtgt ggttttggaa cctccagcgt gtgtgctgg 420
 gttgcgtggt ggggtggggc cggctgtgga cagaggagg gataaagcgg cgggtgccc 480
 cgggtgccc ggcgtgggg cgtggggcgt ggggtgggtg gccagagcct tgggaactcg 540
 10 tcgctgctcg ggcgtctcc cctcctggtc ccctctctga cctacgctcc acatcttcgc 600
 cgttcagtgg ggccttctg ggtggaagtc accatccctt tggactttag ccgacgaagg 660
 ccgggctccc aagagtctcc ccggaggcgg ggccttgggc aggctcaca ggatgctgac 720
 15 ggtgacgggt ggtgacggt atgtacttcg gaggcctcgg gccaatgcag aggtatccat 780
 ttgacctcgg tgggacaggt cagctttcgt gagtcccgtg cgtccttcca gagactcatc 840
 cagcgctagc aagcatggtc ccgagg 866

20

<210> 60

<211> 2067

25 <212> ADN

<213> Secuencia artificial

<220>

30 <223> secuencia de STAR60

<220>

<221> característica_misc

35 <222> (92)..(1777)

<223> Todos los "N" en diferentes posiciones representan cualquier ácido nucleico

40

45

50

55

60

65

ES 2 344 073 T3

<400> 60

5	agcagtgag aactggggaa gaagaagagt ccctacacca ctttaatactc aaaagtactc	60
	gcaaaaaata acacccctca ccagggtggca tnattactct ccttcattga gaaaattagg	120
	aaactggact tcgtagaagc taattgcttt atccagagcc acctgcatac aaacctgcag	180
10	cgccacctgc atacaaacct gtcagccgac cccaaagccc tcagtcgcac caagcctctg	240
	ctgcacaccc tcgtgccttc aactggccg tcccccaagc ctggggcata ctntccagct	300
	ctgagaaatg tattcatcct tcaaagccct gctcatgtgt cctnntcaac aggaaaatct	360
15	cccatgagat gctctgctat ccccatctct cctgccccat agcttaggca nacttctgtg	420
	gtggtgagtc ctgggctgtg ctgtgatgtg ttcgcctgcn atgtntgttc tccccacaa	480
	tgatggggcc ctgaattctc tatctctagc acctgtgtct agtaaaggct tgggaaacca	540
20	ggctcaaagc ctggcccaga tgccaccttt tccagggtgc ttccgggggc caccaaccag	600
	agtgcagcct tctcctccac caggaactct tgcagcccca cccctgagca cctgcacccc	660
	attacccatc tttgtttctc cgtgtgatcg tattattaca gaattatata ctgtattctt	720
25	aatacagtat ataattgtat aattattctt aatacagtat ataattatac aaatacaaaa	780
	tatgtgttaa tggaccgttt atgttactgg taaagcttta agtcaacagt gggacattag	840
	ttaggttttt ggcgaagtca aaagttatat gtgcattttc aacttcttga ggggtcggta	900
30		
	cntctnacc ccatgttggt caanggtcaa ctgtctacac atatcatagc taattcacta	960
	cagaaatggt agcttgtgtc actagtatct ccccttctca taagcttaat acacatacct	1020
35	tgagagagct cttggccatc tctactaatg actgaagttt ttatttatta tagatgtcat	1080
	aataggcata aaactacatt acatcattcg agtgccaatt ttgccacctt gacctctttt	1140
	tgcaaaacac caacgtcagt acacatatga agaggaaact gcccgagaac tgaagtctct	1200
40	gagaccagga gctgcaggcg ttagatagaa tatggtgacg agagttacga ggatgacgag	1260
	agtaaatact tcatactcag tacgtgccaa gcaactgctat aagcgtctg tatgtgtgaa	1320
	gtcatttaat cctcacagca tcccacggtg taattatttt cattatcccc atgagggaaac	1380
45	agaaactcag aacggttcaa cacatatgcg agaagtcgca gccgggtcagt gagagagcag	1440
	gttcccgctc aagcagtcag accccgagtg cacactctcg acccctgtcc agcagactca	1500
	ctcgtcataa ggcggggagt gntctgtttc agccagatgc tttatgcac tcagagtacc	1560
50	caaaccatga aagaatgagg cagtattcan gagcagatgg ngctgggcag taaggctggg	1620
	cttcagaata gctggaaagc tcaagtnatg ggacctgcaa gaaaaatcca ttgtttngat	1680
	aaatagccaa agtccctagg ctgtaagggg aagggtgtgcc aggtgcaagt ggagctctaa	1740
55	tgtaaaatcg cacctgagtc tcctggtctt atgagtnctg ggtgtacccc agtgaaaggt	1800
	cctgctgcca ccaagtgggc catggttcag ctgtgtaagt gctgagcggc agccggaccg	1860
	cttctcttaa cttcacctcc aaaggcacag tgcacctggt tcctccagca ctcagctgcg	1920
60	aggcccctag ccagggtccc ggccccggc ccccggcagc tgctccagct tccttcccca	1980
	cagcattcag gatggtctgc gttcatgtag acctttgttt tcagtctgtg ctccgaggtc	2040
	actggcagca ctagccccgg ctctgt	2067

65

ES 2 344 073 T3

<210> 61

<211> 1470

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> secuencia de STAR61

10

<220>

<221> característica-misc

<222> (130)..(976)

15

<223> "N" representa cualquier ácido nucleico en diferentes posiciones

<400> 61

20

```

cagccccac atgccagcc ctgtgctcag ctctgcagcg gggcatggtg ggcagagaca    60
cagaggccaa ggcctgctt cggggacggt gggcctggga tgagcatggc cttggccttc    120
gccgagagtn ctcttgtaa ggaggggtca ggaggggctg ctgcagctgg ggaggagggc    180
gatggcactg tggcangaag tgaantagtg tgggtgcctn gcacccagg cacggccagc    240
ctggggtatg gacccggggc cntctgttct agagcaggaa ggtatggtga ggacctcaa    300
aggacagcca ctggagagct ccaggcagag gnacttgaga ggcctgggg ccatcctgtc    360

```

30

```

tcttttctgg gtctgtgtgc tctgggcctg ggccttcct ctgctcccc gggcttggag    420
agggttgccc ttgcctcgtg caaaggacca ctctagactg gtaccaagtc tggcccatgg    480
cctcctgtgg gtgcaggcct gtgcgggtga cctgagagcc agggctggca ggtcagagtc    540
aggagaggga tggcagtggg tgccctgtgc aggatctgcc taatcatggt gaggctggag    600
gaatccaaag tgggcatgca ctctgcactc atttctttat tcatgtgtgc ccatcccaac    660
aagcaggagg cctggccagg agggcccctg ggagaaggca ctgatgggct gtgttcatt    720
taggaaggat ggacggttgt gagacgggta agtcagaacg ggctgcccac ctcggccgag    780
taggaaggat ggacggttgt gagacgggta agtcagaacg ggctgcccac ctcggccgag    780
agggccccgt ggtgggttgg caccatctgg gcctggagag ctgctcagga ggctctctag    840
ggctgggtga ccaggntctg ggtacagtag ccatgggagc aggtgcttac ctggggctgt    900
ccctgagcag gggctgcatt ggggtgctctg tgagcacaca cttctctatt cacctgagtc    960
ccnctgagtg atgagnacac ccttgttttg cagatgaatc tgagcatgga gatgttaagt   1020
ggcttgccctg agccacacag cagatggatg gtgtagctgg gacctgaggg caggcagtc    1080
cagcccgagg acttcccaag gttgtggcaa actctgacag catgaccca ggaacacccc   1140
atctcagctc tggtcagaca ctgcggagtt gtgttgtaac ccacacagct ggagacagcc   1200
accctagccc cacccttatt ctctcccaa ggaacctgcc cttcccttc attttcctct   1260
tactgcattg agggaccaca cagtgtggca gaaggaacat ggggtcagga cccagatgga   1320
cttgcttcac agtcagccc tcctgtcctc ttgcagagtg cgtcttcac tgtgaagttg   1380
ggacagtcac accaactcaa tactgctggg ccgctcacac ggtgggcagg caacggatgg   1440
cagtcactgg ctgtgggtct gcagaggtgg                                1470

```

65

ES 2 344 073 T3

<210> 62

<211> 1011

<212> ADN

5 <213> Secuencia artificial

<220>

<223> secuencia de STAR62

10

<400> 62

15	agtggtcaaat agatctacac aaaacaagat aatgtctgcc cttttttcca aagataatgt	60
	gggtgaagtgg gtagagagaa atgcatccat tctccccacc caacctctgc taaattgtcc	120
	atgtcacagt actgagacca gggggccttat tcccagcggg cagaatgtgc accaagcacc	180
20	tcttgtctca atttgcagtc taggccctgc tatttgatgg tgtgaaggct tgcacctggc	240
	atggaagggtc cgttttgtac ttcttgcttt agcagttcaa agagcagggg gagctgagag	300
	ggcctctgca gcttcagatg gatgtggtca gcttggtgga ggcgccttct gtggtccatt	360
25	atctccagcc cccctgcggt gttgctgttt gcttggttg tctggctctc catgccttgt	420
	tggctccaaa atgtcatcat gctgcacccc aggaagaatg tgcaggccca tctcttttat	480
	gtgctttggg ctattttgat tccccgttg gtatattccc taggtaagac ccagaagaca	540
30	caggaggtag ttgctttggg agagtttgga cctatgggta tgaggttaata gacacagtat	600
	cttctctttc atttgggtgag actgttagct ctggccgcgg actgaattcc acacagctca	660
35	cttgggaaaa ctttattcca aaacatagtc acattgaaca ttgtggagaa tgagggacag	720
	agaagaggcc ctagatttgt acatctgggt gttatgtcta taaatagaat gctttgggtg	780
	tcaactagac ttgttcatgt tgacatttag tcttgccctt tcggtggtga tttaaaaatt	840
40	atgtatatct tgtttggaat atagtggagc tatggtgtgg ctttttcac tggcttttg	900
	tttagctcag cccgtcctgt tatgggcagc cttgaagctc agtagctaata gaagaggtat	960
	cctcactccc tccagagagc ggtccctca cggctcattg agagtttgc a	1011

45

<210> 63

<211> 1410

<212> ADN

50 <213> Secuencia Artificial

<220>

<223> secuencia de STAR63

55

60

65

ES 2 344 073 T3

<400> 63

5	gcgtctgagc cgctgggaac ccatgagccc cgtccatgga gttgaggaag ggggttcgcc	60
	ccacgggggtg ggcgccctct acacagcgcg cttcctcttc tctcgtttagc gccgcgggac	120
	cagcctcttg ttctgcacct cgcgctctgg gagcagcgcc cggctttggc gagcgcttcc	180
	ccggggctgc ccagcctctg ctccgctcgc cccgccaggc ccggctccgc gaagcccca	240
10	gggtccagtc caaggccccg attccccaag gccaggggccc cggggcagca ttggaacagg	300
	gcgcggacgc cagtcctccg agcatggagt aactgcagct tttgagaaaa gaaagcggac	360
	cccaccccat cgagaacgcg gcgccttggt tagggacgtt cctgggccgt cacggagtgt	420
15	cgccggctcc tcggccctc cctcctccaa gccccaccc ccgacagcgg cctccctggg	480
	gacctcccct cgggctgcgc tttcagccca aacacaggga ggtcttcag gagcctgccc	540
	agtccccaca gcagcccaga gacccccact cccacctgta cctgccaagc cttcagagag	600
20	ggcggccttg acatgccccg cacgggagga gccccgcctc agcaccctg caagtggcag	660
	caaccagaa caccctgag aggcctctga gcagcccagg aagtggctgg aagacgcata	720
	ggcagctcac tcctctgtaa gagcaaggac cggagaacac atgctgacct ctgcttttgc	780
25	agaggggcca tgcttcagga caggcgcgct cagcaggtgt ccattctatt tcacaccttt	840
	gtgtttatat catcttattt tgcattttat gtctaattaa caatatgcag ctggccaggc	900
	gcagtggctc aagcctctaa tcccagcact ttgggaggcc gaggcagggtg taccacttga	960
30	gggcaggagt tcgggaccgg cctgggcaac atagcaaaac ccattgcta ataaaaatac	1020
	aaaaattagc cagccatggg ggcgggcacc tgcagtccca gctactccgg aagctgaagc	1080
	aggagaatca cttgaacca ggaggcggag gtggcagtga gctatcaagc cattacactc	1140
35	cagcctgggc aacagagaaa gactgtctca aaaaaaaatt aatacgcagc agaatattat	1200
	gtggtcagcc caagcagtc cccccactca gccctctgtc cctacagctc caggcactcc	1260
	cccagcccct cccctggaca agaggtaatg cccagagggt gaaaatccac caagggttaag	1320
40		
	ccagaaacaa aaagctcaaa gcttcggcat ctccctccgc tcagaccctt agagcagatt	1380
45	cctctcatcg acagcacgat caggctgtgg	1410

<210> 64

<211> 1414

<212> ADN

<213> Secuencia Artificial

<220>

<223> secuencia de STAR64

ES 2 344 073 T3

<400> 64

5	agagatcttt taagggctca aaagaccctg cggctcccct gccaatagct ctgccatcgt	60
	ccccagagct ttcgaggacc ctccaccatc ggcgccaacc ccagctgagc tgggtgctcg	120
	tctgcaggcc tctgtccat ctgagcctga gcatgaggct ctgctgtgct gcttccagca	180
	gcagggacag ggctgatgag cctggccctt gcaagcatct tcctgtgccg aatacaattc	240
10	cacagacaga ggatttaaaa tccaagtgga ggtgacagga aagaaaggaa aacctccagg	300
	tatcagaaga aaggaggggg tgtgaagaca gtatgggagg aaggtcaggc tggggctcag	360
	ctctgggaag tgccagcctg aacaggagtc acgcccgggt ccacatgcaa gggaatgagg	420
15	accgagggcc tgcatgtggc agggccttcc gcaggctgcc ccgtctgtga acaggacacc	480
	agaagaagtc tgctttccag cctggcaaag tggcaaggaa cctctgggtg ggaaaacaaa	540
	tcaacaaaca aattgtcagt aaaaaacaga aacctcacac tttcctttct cttgacctct	600
20	tgaaaaaagc aaatccactg cagctcacca aaggcaaaga gaaaacctta agaataccca	660
	gagagaaaag acacgttact tgcaaaagaa catctaagtc agggagataa tgaaaataca	720
	gactcttcaa agggctgaag gaaaaaaacc gtccacctag aattctatcc ccaaactgtc	780
25	atctgagagc aaggggcaaaa caaacgcttt ctgagacagg ctggacgagg tcgctcacgc	840
	ctgtaatcct agcacttttg gagggcaagg tgggaggacc gctttaagcc agaagtttga	900
	gaccagtgtg ggtaacataa tgagacccca tctctaagaa aaagaaatta aataagacaa	960
30	gactttttca gacaacaagt gctctgagag ctggcctatc ttggctgtct tgtaaagaat	1020
	tgctgcgaga cacctcatta ggaaagagac tgaatctaga aggaaagagc agagcatgag	1080
	gtacaatgag gagcaaataa acaggtcacc atataagcaa acccaaatac acattcacta	1140
35	tacgaaacaa taaaaatgac tcatttgggg gggttaaaaca ctgttgaact aaaatcctgg	1200
	ataacagcag catgaaagggt ggggtggtgg tcccaggaaa gcattcaaag gtccatgtct	1260
	catttgaggag gagggtaggg agactcatga acttgaggct cccttcaggc aagcacagt	1320
40	caaaaaaatt ataataatgg gaaacagata cagtagactg tgatgtacaa ctctcagagc	1380
	agtagaaggg aggggtataaa acaaacttga tcca	1414

45 <210> 65

<211> 1310

<212> ADN

50 <213> Secuencia artificial

<220>

<223> secuencia de STAR65

55

60

65

ES 2 344 073 T3

<400> 65

	tcgagaccag cctggccaac atggtgaaac attgtctcta ttacaaatac aaaaattagc	60
5	caggtgtagt ggtgcatgcc tgcagtccca gccatttggg aggctgaagt tggagaatcg	120
	cttaaacctg ggaggtggag gttgcattga gccgagaagc actccagcct ggatgacgga	180
	gcaagactgt ctcaaaaaga aaaaaaaaag aagcagcagc aaatatccct gtcctgatgg	240
10	aggctatata acaaccaaac aagtgaatgc ataagacaat ttcaaggtta tggtagatag	300
	cataagtggtg agatgaacaa tgagaacaca tggacacagg gaggagaaca tcacacactg	360
	gggcctctcg gggggtgggg aaataggggg tgatagcatt aggagaaata cctaattgctg	420
15	ataacaggtt agtgggtgca gcaaacacc atggcacgtg tatactctatg taacacacct	480
	gcacgttctg cacatgtatc ccagaactta aagtataata aaaaaagaca ttaaaaaatt	540
	atgatataaa atcccaattc aagttgtttt aaaaagagaa aacaattatc tttatataat	600
20	agcggaaaat atagatggcg gaattaaagc ctctcatat tttctaacag aactttctga	660
	taaacttgat taaataaaaa ttttaaata cactaaacac atagaagaaa taaatttaaa	720
	ccttcacaaa aaataaagta caatgaatga agacaagggtg tacttgaaaa aagaactgaa	780
25	taaatattct acatataaaa aaaatctgat gatattgtgg tgattcttta ctttgctact	840
	agtttctctt ttttcttct gaaaaatttc ttgggatgta ttgggtttca ttagtaaaat	900
	tctaagtttc tttgcaatct gaacattgga gcttcatcca tagccagtat gccctaakat	960
30	tatctttgga caactgtaaa attagaacac tgccagacat atttaagtga tgatgtatat	1020
	caacactggg acacatttta tactatcttt attccaaaat caaatgattc actgtggttt	1080
	ataaatgtac atggatatat ctctacctaa gcagatagtt aggagagtta gtaaaaatga	1140
35	ggtggaaaat aggagtcact gtcccttcac agggagagaa ttctgctttt ctccataat	1200
	accctttgct tgaacagact ccaaccctc atcttttgtc ctttaaatga ccacatttat	1260
	tttaactttg ataaacaaca cagaaagata tttgatccat caacattcac	1310

40 <210> 66

<211> 2500

<212> ADN

45 <213> Secuencia artificial

<220>

<223> secuencia de T2F (STAR66F)

50 <400> 66

	gcaggttgga tgggtctgac ccctcctcgg gttggcttcc tgtctccagg tggacgtcct	60
	gtactccagg gtctgcaagc ctaaaaggag ggaccacagga cccaccacag acccgctgga	120
55	ccccaaaggc caggagcga ttctggccct ggcgggtgac ctggcctacc agaccctccc	180
	gctcagggcc ctggatgtgg acagcggccc cctggaaaac gtgtatgaga gcatccggga	240
60	gctgggggac cctgctggca ggagcagcac gtgcggggct gggacgcccc ctgcttcag	300

65

ES 2 344 073 T3

	ctgccccagc ctagggaggg gctggagacc cctccctgcc tccctgccct gaacactcaa	360
	ggacctgtgc tccttcctcc agagtgaggc ccgtcccccg ccccgccccg cctcacagct	420
5	gacagcgcca gtcccagggtc cccgggctgc cagcccgtga ggtccgtgag gtcctggccg	480
	ctctgacagc cgcggcctcc ccgggctcca gagaaggccc gcgtctaaat aaagcgccag	540
	cgcaggatga aagcggccag cctcgcagcc tgctcttctt gaaagctggg cgggttgggg	600
10	cgggggggctt ctctggaagg cttggagctg tccccctctg ccttggggga ctggctgccc	660
	ccggggcgcc cgggcctagc cgaggcgggtg ctcttgccgg ccagactctc ggtcagtgcg	720
	ggcacgggggt cccagccact cctagggggc agcgcagccg gcagggtggc cgccccggg	780
15	tgggacttgg accctggact ccacgggagg gctccgccac ccagcctggt gttacataag	840
	gggtggtgga ggtgggcagt cgagcgttaa agagtaacct gctgccggga agcccggcaa	900
	gcaatcgcgg ccccttcccc ggctctggca gctctgcgag cgcgccctg gggaacgggc	960
20	cctccccggc ggggcgcgcg ggcgcgcgag gtgggcggag gcctcggagc tgtgccgggc	1020
	cgggcctccc tccctaggcc agcgcgggag cgacccggag ggggcgggcc cggggcgggg	1080
	cctcgaagcg ctggccggcg ggagcgcggc cggccgggcc cgcccgccctg cgggtgtggac	1140
25	gccgcgcggc caatgcgcgc gccgggacgg gacgggacgg ggcggggcgg ggcgggacga	1200
	gacggggcgg ggcggggcgg gccgggcagc ctccgggcgg cgcggcgcgg gcggcgccg	1260
	gatccagggc gggggctggc ggcccggcca gcccgggccc gcccggggcc gcgtcctgag	1320
30	agtcagccct cgccgctgca gcctcggcgc ccggccggcc ggccatggag cgcccccg	1380
	cccgcgccgc cggccgggac cccagtgcgc tgcgggccga ggcgccgtgg ctgcgcgcgg	1440
	aggggtccggg gccgcgcgc gcgccgtga cggtgccac gccgccgag gtaccgggcg	1500
35	ccggtgggcg ggggcgccga ccaagtttct ctcgctgcaa agatggcgtc agtgctgccc	1560
	aaacttcggg cccccggggg cggggcagcg gggaggcgg ccgcgtcggc ccgcgcgtgt	1620
	ccgtgggtcc cgccggggct ggcgggggc gccggggagc ctttcccgc gcgccgggct	1680
40	gggggcgggg ccgggggcgg ggccgcgcgc tccacaccgg ccgcagccgg ttttcgaggc	1740
	gggcgcggag cggatccgcg gcggagggtt agggaccccc ctccccggc caccgcctcc	1800
	gctgagtctg cccctcccc atccgcagg ctcttccgtg ggcggcggct tcgcgggctt	1860
45	ggagttcgcg cgccgcagg agtcggagcc gcgggcctcg gacctggggg cccccggac	1920
	gtggacgggg gcggcggcg ggccccggac tccgtcggcg cacatccccg tcccagcgca	1980
	gaggtgagcg ggaggcccg tgcttcggga ctcggtgtgc gcaggggcgg tgggtgggg	2040
	gcggagacac cggccccgac ggaggccagg tcaggggccc aggtttgtaa ttaccagcca	2100
50	cccccaagct cttcagccct ggaggagctg agcagaaatg atcgatgact gggagtccct	2160
	acacctccct ccaccgcagt tcctcggggc tagagctcag aaccgggagc gggtggtgt	2220
	gcgtctctgt gcagaagagg ctgcgcggtc ggcattgggc gactgtccag gaatccctg	2280
55	ggctcctgac cgccacctc caaccctgc caggccggac acctcggctt ggctgccagg	2340
	gcaggggcgg gccctggcct ggctcgttg ggcctgggga gctgccctg cttccagccc	2400
60	agtctcccc tggctgctgc cggtgctg ccactccac ctcccaggcc tggcgtgagg	2460
	cccacagctg ctgttgaca accctggtta atgtgtgatg	2500

65 <210> 67
<211> 2500

ES 2 344 073 T3

<212> ADN

<213> Secuencia artificial

5 <220>

<223> secuencia de T2R (STAR66R)

<400> 67

10

gtttggggtg gagagaacat actgattatg ggactttgct ttgcagctta gtgctgtcct 60

gtcagtggga agcaacaggg ggcagaactc agcttgtgcc catagaggga atgtttatac 120

15

taggcctgtc cagaggcaaa tcatccatcc tagcaattgg aacctgactt ttggcaagtc 180

ctgccaccat gggctaaagt gttctggggg tctaaataaa catgaaaggc aacctagacc 240

acaaggactg caattcctgc acaagtcctg gtgctgtgtt gggcttggag ccagggaact 300

20

tggagtgcac ggaacctagt gagataccag ctgagacaac caaggaagtg cttgtgtcac 360

ccctccacca accccaggca gtacagattg tacctccaag accccttcca tctgcttgag 420

gaaggtggag ggaagagga ctttgttttg caacttggat tccagcccat ccacagtaga 480

25

ataaggcaac gggcagactc ctaaggcccc catcccagac cctagctcct ggatgacatt 540

tctaaacaca ccatgggcca gaagggaacc cattgccttg aagggaaggg cccagtcctg 600

gcagaattta tcatgtgctg aataaacagc ccttggggcc tgaataatta gtattggtag 660

30

ccaggcagta tttaccacag gccttgggtg agaccagag ccatgttggc ttcaggtgtg 720

accagcaca ttcccagctg tggtaacttt ggggagagac cacttctgct tgagaaaagg 780

agacagaaga gtaaaggggt ctttatcttg cagcctggta ccagcttggc cgcagtgggg 840

35

tagagcacca agagagcacc tgggataaac aaaatcaaaa aaccttttag tagactaaga 900

gtaaagagag aagacccaag taaatataat caaagacaaa aaaggagaga cattacaacc 960

aatacctcag aaattcaaag tatcattagc agctactttg aacaactata tgccagtaaa 1020

40

ttggaaaacc tagaagaatt atataaattc ctaacatata caacctacca agattgaacc 1080

atgaagaaat ttaaagcctg aataggccaa taacaagcaa tgagattgga gccctaatac 1140

aaagtttaca atgagaaaca ttgctcaaac aaatcataga tgacacaaac aaatggaaaa 1200

45

catccaatgc tcatggacag gaaaaaatat ttaaatttct atactgcca aagcagttta 1260

tacattcaat gctattcctg tcaaaatacc aatcttattc ttcacaaaaa aaaaattaaa 1320

aattacacag aaccaaaaaa gagcccaaat acccaaggca attttaagca aaaagaacaa 1380

50

agctggaggc atcacgttac ctgtgatcca cactataggg ctacagtaaa tgaaacagca 1440

aggtgtggt atacaaacag acacataaac caatggaata gaataaagag cttagaaata 1500

atgctccaca cctccagcca tccgatgttt gagaaagtag acataaacia gcaatgagga 1560

55

60

65

ES 2 344 073 T3

	gaggactccc tattcattaa atcaactcaa gacggaccaa aaacctaataa gtaaaacaaa	1620
	caaacaaaaa aaataactgc taaaaccctg ggagatgacc taggaaatac cattctggac	1680
5	agtacctggt gaaaatttca tgctgaagac accaaaaaca attgcagcaa aagaaaaaat	1740
	tgacatatgg gatcaaatta aacttttagag cttttgcaca gcaaaataaa ctatcaacag	1800
	agtaaatagg caccctacag gaagggagaa aatattttca atctgtgctc tgacaaagtc	1860
10	ctaatatcca gagcctataa ggaacttaaa caaattttaca aacaaaaaac aaacaacact	1920
	attacgagtt ggaaaaggac atgaatcgac actttttcaa agaagacata catgtggcta	1980
	acaagcatat gaaaaaatg ctcaacatta ctaatcatta gagaaatgca aatcaaaacc	2040
15	acaatgagat accatctcaa ccagtctgaa tggctgttat taaaaaatc agaaaaaac	2100
	agatgctggc aagggtgtgg agaaaaggaa acacttatac attgttgggg ggagtgtaaa	2160
	ttaattcagc cattgtggaa agtattgtgg tgattttcta aagaactaaa aaggaattac	2220
20	tattttacct ggaaatttca ttattgggta tatacccaaa gaaatatgaa ttattttact	2280
	ataaagacag atgcatgcat gtgttcattg tagcactatt cacagtagca aagacatgtt	2340
	atcaacctaa atgcccatta acagtaaaact ggataaggaa aatatggtac atatacactg	2400
25	tggaatacta tgcagtcata aaaagaatga gataatgttc attgcagcaa catggatgga	2460
	actggagacc attatccttg ggaaactaac aaagcaacag	2500

30 <210> 68
 <211> 13
 <212> ADN
 <213> Secuencia Artificial

35 <220>
 <223> secuencia loxP

40 <400> 68

ataacttcgt ata

13

45 <210> 69
 <211> 13
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> secuencia FRT

55 <400> 69

gaagttccta tac

13

60 <210> 70
 <211> 631
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <223> elementos STAR en STAR3 directo

ES 2 344 073 T3

<220>

<221> característica_misc

<222> (5)..(630)

5 <223> "N" representa cualquier ácido nucleico en diferentes posiciones

<400> 70

10	acgtncctaag naaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag	60
	gcccttttcgt cttcactcga gcggccagct tggatctcga gtactgaaat aggagtaaat	120
	ctgaagagca aataagatga gccagaaaac catgaaaaga acagggacta ccagttgatt	180
15	ccacaaggac attccaagg tgagaaggcc atatacctcc actacctgaa ccaattctct	240
	gtatgcagat ttagcaagg tataaggtag caaaagatta gaccaagaa aatagagaac	300
	ttccaatcca gtaaaaatca tagcaaattt attgatgata acaattgtct ccaaaggaac	360
20	caggcagagt cgtgctagca gaggaagcac gtgagctgaa aacagccaaa tctgctttgt	420
	tttcatgaca caggagcata aagtacacac caccaactga cctattaagg ctgtggtaaa	480
	ccgattcata gagagagggt ctaaatacat tggctcctca taggcaaacc gcagttcact	540
25	ccgaacgtag tccctggaaa tttgatgtcc agnatagaaa agcanagcag ncnnnnnnta	600
	tanatnnnngn tganccanatt gntnnctggn c	631

<210> 71

<211> 774

30 <212> ADN

<213> Secuencia Artificial

35 <220>

<223> elementos STAR en STAR3 inverso

<220>

40 <221> característica_misc

<222> (637)..(762)

<223> "N" representa cualquier ácido nucleico en diferentes posiciones

45 <400> 71

	gagctagcgg cgcgccaagc ttggatcccg ccccgcccc tccgcccctcg agccccgccc	60
	cttgcccctag agggccctgcc gagggggcggg gcctgtccct cctccccctt ccccgcccc	120
50	ctaccgtcac gctcaggggc agcctgacct cgagcggccc cgcggtgacc ctccgcgcaga	180
	ggcctgtggg aggggctgct caagcccctg aatccccccc cgtctgttcc cccctcccgc	240
	ccagtctcct cccctggga acgcgcgggg tgggtgacag acctggctgc gcgccaccgc	300
55	caccgcgcct gccgggggct ctgccgctgc ctgagaaact gcggctgccg cctggaggag	360
	gtgccgtcgc ctccgccacc gctgccgccc ccgccagggg taggagctaa gccgccgcca	420
	ttttgtgtcc cctgttgtt gtcgttgaca tgaatccgac atgacactga ttacagccca	480
60	atggagtctc attaaacccg agtcgcggtc ccgccccgcc gctgtcccat tggaggagac	540
	caaagacact taaggccacc cgttggccta cgggtctgtc tgtcacccac tcaactaaca	600
	ctctgcagcc cattggggca ggttcctgcc ggtcatntcg cttccaataa acacaccctt	660
65	tcgaccccat nattcccccc cttcggggaa ccccccggg ggagggggtcc actggncaat	720
	accaattnaa nagaaccgct ngggtccgcc tntttncggg cncctattg gggt	774

ES 2 344 073 T3

<210> 72
 <211> 717
 <212> ADN
 5 <213> Secuencia Artificial

 <220>
 <223> elementos STAR en STAR4 directo
 10
 <220>
 <221> característica_misc
 <222> (444)..(704)
 15 <223> “N” representa cualquier ácido nucleico en diferentes posiciones

 <400> 72

 20 ggggaggatt cttttggctg ctgagttgag attaggttga gggtagtgaa ggtaaaggca 60
 gtgagaccac gtaggggtca ttgcagtaat ccaggctgga gatgatggtg gttcagttgg 120
 aatagcagtg catgtgctgt aacaacctca gctgggaagc agtatatgtg gcgttatgac 180
 25 ctgagctgga acagcaatgc atgtggtggt gtaatgaccc cagctgggta ggggtgcatgt 240
 gatggaacaa cctcagctgg gtagcagtggt acttgataaa atgttggcat actctacatt 300
 tgttatgagg gtagtgccat taaattttct cacaaattgg ttgtcacgta tgagtgaaaa 360
 30 gaggaagtga tggaagactt cagtgccttt ggcctgaata aatagaagac gtcattttca 420
 gtaatggaga caggaagac taangnaggg tggattcagt agagcagggtg ttcagttttg 480
 aatatgatga actctgagag aggaaaaact ttttctacct cttagttttt gngnctggac 540
 35 ttaanattaa aggacatang acngaganca gaccaaatnt gcgangtttt tatattttac 600
 ttgcngaggg aatttncaag aaaaagaaga cccaanancc attggtcaaa actatntgcc 660
 ttttaanaaa aaganaatta caatgganan anaagtgttg nctnggcaaa aattggg 717
 40

 <210> 73
 <211> 541
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 50 <223> elementos STAR en STAR4 inverso

 <220>
 <221> característica_misc
 55 <222> (6)..(533)
 <223> “N” representa cualquier ácido nucleico en diferentes posiciones

 60

 65

ES 2 344 073 T3

<400> 73

	ggattngagc tagcggcgcg ccaagcttgg atcttagaag gacagagtgg ggcattggaaa	60
5	tgcaccacca gggcagtgca gcttgggtcac tgccagctcc nctcatgggc agagggctgg	120
	cctcttgtag cgcaccaggc actgagcgcc atcccagggc cctcgccagc cctcagcagg	180
	gccaggacac acaagccttt gacttcctcc tgteactgct gctgccattc ctgtttttgtg	240
10	gtcatcactc cttccctgtc ctcagactgc ccagcactca aggatgtcct gtggtggcat	300
	cagaccatat gcccctgaan aggagtgagt tgggtgtttt tgccgcgccc anagagctgc	360
15	tgtcccctga aagatgcaag tgggaatgat gatgntcacc atcntctgac accaagccct	420
	ttggatagag gcccacacag tgaggatggg gctgcactgc attgccaagg caactctgtg	480
	ntgactgcta cangacantc ccaggacctg ngaaghncta tanatntgat gcnaggcacc	540
20	t	541

<210> 74

25 <211> 794

<212> ADN

<213> Secuencia artificial

30 <220>

<223> elementos STAR en STAR6 directo

<220>

35 <221> característica_misc

<222> (374)..(792)

<223> "N" representa cualquier ácido nucleico en diferentes posiciones

40 <400> 74

	ccaccacaga catccccctc ggccctcctga gtggtttctt cagcacagct tccagagcca	60
45	aattaaacgt tcaactctatg tctatagaca aaaaggggtt tgactaaact ctgtgtttta	120
	gagagggagt taaatgctgt taacttttta ggggtgggcg agaggaatga caaataacaa	180
	cttgtctgaa tgttttacat ttctccccac tgcctcaaga aggttcacaa cgaggtcac	240
50	catgataagg agtaagacct ccagccgga ctgtccctcg gcccacagag gacactccac	300
	agagataatg taactggact tggagactgg ctacactcc agagaaaagc atggagcacg	360
	agcgcacaga gcangggcca aggtcccagg gacngaagt ctaggaggga gattggggtg	420
55	agggtantct gatgcaatta ctgngcagct caacattcaa gggaggggaa gaaagaaacn	480
	gtccctgtaa gtaagtgtg cancagagat ggtaagctcc aaatttnaac tttggctgct	540
	ggaaagttn ngggccnana naanaaacan aanatttga ggtttanacc cactaaccn	600
60	tatnantant tattaatacc cctaattana ccttggatan ccttaaaata tcntntnaaa	660
	cggaaccctc nttccnntt nnaaatnnaa aaggccattn ngnnchnagta aaaatctnnn	720
	ttaagnnttg ggcccnaaca aacntnttcc nagacacntt tttnttccng gnatttntaa	780
65	tttatttcta ancc	794

ES 2 344 073 T3

<210> 75

<211> 379

<212> ADN

5 <213> Secuencia Artificial

 $\langle 220 \rangle$

<223> elementos STAR en STAR6 inverso

10

 $\langle 220 \rangle$

<221> característica misc

<222> (287)..(379)

15

<223> “N” representa cualquier ácido nucleico en diferentes posiciones

<400> 75

20	atcgtgtcct ttccagggac atggatgaag ctggaagcca tcatcctcag caaactaaca	60
	caggaacaga aaaccaata ccacatgttc tactcataa gtgggagctg aacagtgaga	120
	acacatggac acagggaggg gaacatcaca caccaaggcc tgtctggtgt ggggagggga	180
25	gggagagcat caggacaaat agctaattgca tgtggggctt aaacctagat gacgggttga	240
	taggtgcagc aatccactat ggacacatat acctatgtaa caaccnacc ttnttgacat	300
	gtatcccaga acttaaagga aaataaaaaat taaaaaaaat tnccttgaa taaaaaagag	360
30	tgtggacttt ggtgagatn	379

<210> 76

<211> 398

35 <212> ADN

<213> Secuencia artificial

 $\langle 220 \rangle$

40 <223> elementos STAR en STAR8 directo

 $\langle 220 \rangle$

<221> característica misc

45 $\langle 222 \rangle$ (98)..(395)

<223> “N” representa cualquier ácido nucleico en diferentes posiciones

<400> 76

50	ggatcacctc gaagagagtc taacgtccgt aggaacgctc tcgggttcac aaggattgac	60
	cgaaccccag gatacgtcgc tctccatctg aggcttgntc caaatggccc tccactattc	120
	caggcacgtg ggtgtctccc ctaactctcc ctgctctcct gagcccatgc tgccatcac	180
55	ccatcgggtgc aggtcctttc tgaanagctc ggggtggattc tctccatccc acttcctttc	240
	ccaagaaaga agccaccggt ccaagacacc caatgggaca ttcccnttcc acctccttnt	300
	cnaaagttnng cccaggtygt cntaacaggt tagggagaga anccccagg tttnagttnc	360
60	aagqcatagg acgtctggctt gaacacacac acacnctc	398

<210> 77

<211> 309

65

<212> ADN

<213> Secuencia artificial

ES 2 344 073 T3

<220>

<223> elementos STAR en STAR8 inverso

5 <220>

<221> característica_misc

<222> (72)..(72)

<223> "N" representa cualquier ácido nucleico

10

<400> 77

```

ggatcccgac tctgcaccgc aaactctacg gcgcccctgca ggacggcggc ctccctgccgc      60
ttggacgcca gncaggagct ccccggcagc agcagagcag aaagaaggat ggccccgccc      120
cacttcgcct cccggcggtc tccctcccgc cggctcacgg acatagatgg ctgcctagct      180
ccggaagcct agctcttgtt ccgggcattc taaggaagac acggtttttc ctcccggggc      240

ctcaccacat ctgggacttt gacgactcgg acctctctcc attgaatggt tgcgcgttct      300
ctgggaaag                                     309

```

20

25

<210> 78

<211> 606

<212> ADN

30

<213> Secuencia artificial

<220>

<223> elementos STAR en STAR18 directo

35

<220>

<221> característica_misc

<222> (524)..(524)

40

<223> "N" representa cualquier ácido nucleico

<400> 78

45

```

tggatcctgc cgctcgcgtc ttagtggttc tccctcaaga ctttccttct gttttgttgt      60
cttgtgcagt attttacagc cctcttgtg. tttttcttta tttctcgtac acacacgcag      120
ttttaagggt gatgtgtgta taattaaaag. gacccttggc ccatactttc ctaattcttt      180
agggactggg attgggtttg actgaaatat gttttggtgg ggatgggacg gtggacttcc      240
attctcccta aactggagtt ttggtcggtg atcaaaacta aaagaaacct ctgggagact      300
ggaaacctga ttggagcact gaggaacaag ggaatgaaaa ggcagactct ctgaacgttt      360
gatgaaatgg actcttgtga aaattaacag tgaatattca ctgttgcaact gtacgaagtc      420
tctgaaatgt aattaaaagt ttttattgag ccccgagct ttggcttgcg cgtatttttc      480
cggtcgcgga catcccaccg cgcagagcct cgccctcccc ctgncctcag ctccgatgac      540
ttccccgccc ccgcccgtct cgggtgacaga cgttctactg cttccaatcg gaggcaccct      600
tcgcgg                                     606

```

50

55

60

<210> 79

65

<211> 465

<212> ADN

<213> Secuencia artificial

ES 2 344 073 T3

	<220>		
	<223> elementos STAR en STAR18 inverso		
5	<400> 79		
	tggaacctgc cgctcgcgtc ttagtgtttc tccctcaaga ctttccttct gttttgttgt	60	
	cttgtgcagt attttacagc ccctcttggtg tttttcttta tttctcgtac acacacgcag	120	
10	ttttaagggt gatgtgtgta taattaaaag gacccttggc ccatactttc ctaattcttt	180	
	agggactggg attgggtttg actgaaatat gttttgggtg ggatgggacg gtggacttcc	240	
	attctcccta aactggagtt ttggtcggta atcaaaacta aaagaaacct ctgggagact	300	
15	ggaaacctga ttggagcact gaggaacaag ggaatgaaaa ggcagactct ctgaacgttt	360	
	gatgaaatgg actcttgtga aaattaacag tgaatattca ctgttgcact gtacgaagtc	420	
	tctgaaatgt aattaaaagt ttttattgag cccccgagct ttggc	465	
20	<210> 80		
	<211> 48		
	<212> ADN		
25	<213> Secuencia artificial		
	<220>		
	<223> oligonucleótido C65		
30	<400> 80		
	aacaagcttg atatcagatc tgctagcttg gttcgagctga tacttccc	48	
35	<210> 81		
	<211> 54		
	<212> ADN		
40	<213> Secuencia artificial		
	<220>		
	<223> oligonucleótido C66		
45	<400> 81		
	aaactcgagc ggccgcgaat tcgtcgactt taccactccc tatcagtgat agag	54	
50	<210> 82		
	<211> 39		
	<212> ADN		
55	<213> Secuencia artificial		
	<220>		
	<223> oligonucleótido c67		
60	<400> 82		
	aaaccgcggc atggaagacg ccaaaaacat aaagaaagg	39	
65	<210> 83		
	<211> 32		

ES 2 344 073 T3

	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> oligonucleótido c68	
	<400> 83	
10	tatggatcct agaattacac ggcgatcttt cc	32
	<210> 84	
15	<211> 33	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> oligonucleótido C81	
	<400> 84	
25	aaaccatggc cgagtacaag cccacgggtgc gcc	33
	<210> 85	
30	<211> 35	
	<212> ADN	
	<213> Secuencia Artificial	
35	<220>	
	<223> oligonucleótido C82	
	<400> 85	
40	aaatctagat caggcaccgg gcttgcgggt catgc	35
	<210> 86	
45	<211> 22	
	<212> ADN	
	<213> Secuencia artificial	
50	<220>	
	<223> oligonucleótido c85	
	<400> 86	
55	catttccccg aaaagtgccca cc	22
	<210> 87	
60	<211> 24	
	<212> ADN	
	<213> Secuencia Artificial	
65	<220>	
	<223> oligonucleótido D30	

ES 2 344 073 T3

	<400> 87	
	tcactgctag cgagtggtaa actc	24
5	<210> 88	
	<211> 28	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido D41	
15	<400> 88	
	gaagtcgacg aggcaggcag aagtatgc	28
20	<210> 89	
	<211> 29	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido D42	
30	<400> 89	
	gagccgcggt ttagttctc acctgtcg	29
35	<210> 90	
	<211> 23	
	<212> ADN	
40	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido D51	
45	<400> 90	
	tctggaagct ttgctgaaga aac	23
50	<210> 91	
	<211> 21	
	<212> ADN	
55	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido D89	
60	<400> 91	
	gggcaagatg tcgragtcag g	21
65	<210> 92	
	<211> 30	

ES 2 344 073 T3

	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> oligonucleótido D90	
	<400> 92	
10	aggcccatgg tcacctccat cgctactgtg	30
	<210> 93	
15	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> oligonucleótido D91	
	<400> 93	
25	ctaatcactc actgtgtaat	20
	<210> 94	
30	<211> 15	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> oligonucleótido D93	
	<400> 94	
40	aattacaggc gcgcc	15
	<210> 95	
45	<211> 15	
	<212> ADN	
	<213> Secuencia Artificial	
50	<220>	
	<223> oligonucleótido D94	
	<400> 95	
55	aattggcgcg cctgt	15
	<210> 96	
60	<211> 25	
	<212> ADN	
	<213> Secuencia artificial	
65	<220>	
	<223> oligonucleótido D95	

ES 2 344 073 T3

	<400> 96	
	tgctttgcat acttctgcct gcctc	25
5	<210> 97	
	<211> 20	
	<212> ADN	
10	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido E12	
15	<400> 97	
	taggggggat ccaaatgttc	20
20	<210> 98	
	<211> 20	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido E13	
30	<400> 98	
	cctaaaagaa gatctttagc	20
35	<210> 99	
	<211> 20	
	<212> ADN	
40	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido E14	
45	<400> 99	
	aagtgttga tccacttgg	20
50	<210> 100	
	<211> 20	
	<212> ADN	
55	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido E15	
60	<400> 100	
	tttgaagatc taccaaatgg	20
65	<210> 101	
	<211> 20	

ES 2 344 073 T3

	<212> ADN	
	<213> Secuencia Artificial	
5	<220>	
	<223> oligonucleótido E16	
	<400> 101	
10	gttcgggatc cacctggccg	20
	<210> 102	
15	<211> 20	
	<212> ADN	
	<213> Secuencia Artificial	
20	<220>	
	<223> oligonucleótido E17	
	<400> 102	
25	taggcaagat cttggcctc	20
	<210> 103	
30	<211> 20	
	<212> ADN	
	<213> Secuencia Artificial	
35	<220>	
	<223> oligonucleótido E18	
	<400> 103	
40	cctctctagg gatccgaccc	20
	<210> 104	
45	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
50	<220>	
	<223> oligonucleótido E19	
	<400> 104	
55	ctagagagat cttccagtat	20
	<210> 105	
60	<211> 20	
	<212> ADN	
	<213> Secuencia Artificial	
65	<220>	
	<223> oligonucleótido E20	

ES 2 344 073 T3

	<400> 105	
	agagttccgg atccgcctgg	20
5	<210> 106	
	<211> 20	
	<212> ADN	
10	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido E21	
15	<400> 106	
	ccaggcagac tcggaactct	20
20	<210> 107	
	<211> 20	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido E22	
30	<400> 107	
	tggtgaaacc ggatccctac	20
35	<210> 108	
	<211> 20	
	<212> ADN	
40	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido E23	
45	<400> 108	
	aggtcaggag atctagacca	20
50	<210> 109	
	<211> 22	
	<212> ADN	
55	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido E25	
60	<400> 109	
	ccattttcgc ttccttagct cc	22
65	<210> 110	
	<211> 22	

ES 2 344 073 T3

	<212> ADN	
	<213> Secuencia Artificial	
5	<220>	
	<223> oligonucleótido E42	
	<400> 110	
10	cgatgtaacc cactcgtgca cc	22
	<210> 111	
15	<211> 20	
	<212> ADN	
	<213> Secuencia Artificial	
20	<220>	
	<223> oligonucleótido E57	
	<400> 111	
25	agagatctag gataatttcg	20
	<210> 112	
30	<211> 36	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> oligonucleótido E92	
	<400> 112	
40	aggcgctagc acgcgttcta ctcttttcct actctg	36
	<210> 113	
45	<211> 35	
	<212> ADN	
	<213> Secuencia artificial	
50	<220>	
	<223> oligonucleótido E93	
	<400> 113	
55	gatcaagctt acgcgtctaa aggcatttta tatag	35
	<210> 114	
60	<211> 35	
	<212> ADN	
	<213> Secuencia artificial	
65	<220>	
	<223> oligonucleótido E94	

ES 2 344 073 T3

	<400> 114	
	aggcgctagc acgcgttcag agttagtgat ccagg	35
5	<210> 115	
	<211> 35	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido E95	
15	<400> 115	
	gatcaagctt acgcgtcagt aaaggttctg tatgg	35
20	<210> 116	
	<211> 36	
	<212> ADN	
25	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido E96	
30	<400> 116	
	aggcgctagc acgcgttcta ctcttcatt actctg	36
35	<210> 117	
	<211> 26	
	<212> ADN	
40	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido E97	
45	<400> 117	
	cgaggaagct ggagaaggag aagctg	26
50	<210> 118	
	<211> 25	
	<212> ADN	
55	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido E98	
60	<400> 118	
	caagggccgc agcttacaca tggtc	25
65	<210> 119	
	<211> 17	

ES 2 344 073 T3

	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> oligonucleótido D58	
	<400> 119	
10	ccaagttgac cagtgcc	17
	<210> 120	
15	<211> 20	
	<212> ADN	
	<213> Secuencia Artificial	
20	<220>	
	<223> oligonucleótido D80	
	<400> 120	
25	gttcgtggac acgacctcgg	20
	<210> 121	
30	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> oligonucleótido D70	
	<400> 121	
40	tacaagccaa ccacgcct	19
	<210> 122	
45	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
50	<220>	
	<223> oligonucleótido D71	
	<400> 122	
55	cggaagtgct tgacattggg	20
60		
65		