



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 344 071**

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

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06124475 .2**

96 Fecha de presentación : **14.06.2002**

97 Número de publicación de la solicitud: **1829971**

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: **Kruckeberg, Arthur Leo y**
Otte, Arie Peter

74 Agente: **Ungría López, Javier**

ES 2 344 071 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 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 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 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.

15 La presente invención tiene que ver con las secuencias de ADN que comprenden la capacidad de influir en la 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.

20 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 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.

25 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 comprenden 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. Los 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 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 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 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 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

ES 2 344 071 T3

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.

5 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 de la invención. Preferiblemente dicha secuencia de ADN comprende 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:

- 25 (a) ID SEC.: 27 en la figura 20;
- (b) un fragmento de ID SEC.: 27 en la Figura 20;
- 30 (c) una secuencia derivada de ID SEC. N°: 27 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.

40 En un aspecto 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.

50 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

55 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.

65 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.

5 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
10 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 en vertebrados ni en plantas) no se han encontrado PRE completos que comprendan toda la información de la secuencia requerida para mediar la represión de la cromatina.

15 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
20 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).

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

30 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.

35 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 deleciones 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. Se puede proporcionar una genoteca de ácidos nucleicos aislados y/o recombinantes que comprenden cualidades moduladoras de la transcripción génica y/o represora de la transcripción génica
45 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.
55 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 interactuar 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 interactúan 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
65

ES 2 344 071 T3

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.

5 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
10 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.

15 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.

20 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
25 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.

35 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.

40 De acuerdo con la invención se proporciona también 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.

45 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 moduladora 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.

ES 2 344 071 T3

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.

En una realización preferida de la invención dicha secuencia de ADN con una cualidad moduladora de la transcripción génica comprende un constructo denominado STAR (Stabilizing Anti-Repression) 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 colE1 de derivado de pSV40/Zeo (Invitrogen V502- 20); este es el marcador seleccionable para el rastreo de STAR.

ES 2 344 071 T3

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-*
5 *EcoRI-SalI*) y aguas abajo de la señal de poliadenilación (MCSII, *NbeI-BglII-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
10 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
15 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
20 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
25 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
30 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
35 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
40

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 puromicina. Se aíslan y cultivan poblaciones de aproximadamente 30 clones individuales por elemento STAR. Los clones son analizados periódicamente en cuanto a la actividad luciferasa
50 según las instrucciones del fabricante (Roche 1669893).

Resultados

55 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.
65

ES 2 344 071 T3

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 estables 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 que 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=humchr.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

55 *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.

ES 2 344 071 T3

Ejemplo 5

Determinación de la longitud máxima de los elementos STAR

5 *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.

10 *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.

20 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).

Ejemplo 6

30

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.

55 *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 Microcócica 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).

65

ES 2 344 071 T3

Ejemplo 7

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

5 *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.

10 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 microcócica (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.

15 *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

30 *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>).

35 *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.

40 *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 STAR 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

65 *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.

ES 2 344 071 T3

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 SacIII/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 reemplazar 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 reemplazado 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 BglIII/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 dideoxi (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 estreptomina 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 incubar durante la noche la mezcla de transfección fue reemplazada 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.

ES 2 344 071 T3

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

ES 2 344 071 T3

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 genotecas 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.

Materiales 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.

5 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).

20 *Materiales 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-SV4 0-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 puromicina, y los análisis con luciferasa se llevaron a cabo como se describe en el Ejemplo 11.

30 *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-SV4 0, 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).

65 *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-STAR,66-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.

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 apli-

caciones 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.

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 (Knofler *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 blo-

queda 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/blastview) 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 fragmentos 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 puromicina) 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 *S a11*; 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 puromicina (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 puromicina.

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 puromicina.

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.

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.

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.

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).

Panel A: Análisis de bloqueo del intensificador.

Panel B: Análisis del intensificador.

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.

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.

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)

Referencias

Altschul, S.F. and Gish, W. (1996) Local alignment statistics. *Methods Enzymol*, 266, 460-480.

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.

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.

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.

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.

Foekking, 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 USA*, 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 USA* 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 USA*, 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.
- Knofler, 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 USA*, 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 transactivator to its target *DNA Embo J* 14, 5660-71.

ES 2 344 071 T3

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

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 *Drosophila*). 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.

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*		
STAR-6	2p21	Quinasa supuesta desconocida L5 kb Proteína asociada microtúbulos R 20 kb	19% SINE (MIR) 29% LINE

ES 2 344 071 T3

TABLA 2 (continuación)

Clon	Ubicación cromosómica ¹	Genes adyacentes ²	Secuencia repetida
STAR-7	Lq32.2		12% Alu 4% MIR (SINE) LINE 1 2,5% 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 2 344 071 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

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.

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.

* Posición ambigua, diversos aciertos

ES 2 344 071 T3

TABLA 4

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

5
10
15
20
25
30
35
40
45
50
55
60
65

STAR3 directo

**ACGTNCTAAGNAAACCATTATTATCATGACATTAACCTATAAAAATAGGC
GTATCACGAGGCCCTTTCGTCTTCACTCGAGCGGCCAGCTTGGATCTCGA
GTA CTGAAATAGGAGTAAATCTGAAGAGCAAATAAGATGAGCCAGAAAAC
CATGAAAAGAACAGGGACTACCAGTTGATTCCACAAGGACATTCCCAAGG
TGAGAAGGCCATATACCTCCACTACCTGAACCAATTCTCTGTATGCAGATT
TAGCAAGGTTATAAGGTAGCAAAGATTAGACCCAAGAAAATAGAGA ACT
TCCAATCCAGTAAAAATCATAGCAAATTTATTGATGATAACAATTGTCTCC
AAAGGAACCAGGCAGAGTCGTGCTAGCAGAGGAAGCACGTGAGCTGAAA
ACAGCCAAATCTGCTTTGTTTTTCATGACACAGGAGCATAAAGTACACACCA
CCA ACTGACCTATTAAGGCTGTGGTAAACCGATT CATAGAGAGAGGTTCT
AAATACATTGGTCCCTCATAGGCAAACCGCAGTTCACTCCGAACGTAGTC
CCTGGAAATTTGATGTCCAGNATAGAAAAGCANAGCAGNCNNNNNTAT
ANATNNNGNTGANCCANATGNTNNCTGNNC**

ES 2 344 071 T3

TABLA 4 (continuación)

STAR3 inverso

5
10
15
20
25
30
35
40
45
50
55
60
65

**GAGCTAGCGGCGCGCCAAGCTTGGATCCCGCCCCGCCCCCTCCGCCCTCG
AGCCCCGCCCCTTGCCTAGAGGCCCTGCCGAGGGGCGGGGCCTGTCCC
TCCTCCCCTTTCCCCCGCCCCCTACCGTCACGCTCAGGGGCAGCCTGACC
CCGAGCGGCCCGCGGTGACCCTCGCGCAGAGGCCCTGTGGGAGGGGCGT
CGCAAGCCCCTGAATCCCCCCCCGTCTGTTCCCCCTCCCGCCCAGTCTC
CTCCCCCTGGGAACGCGCGGGGTGGGTGACAGACCTGGCTGCGCGCCAC
CGCCACCGCGCCTGCCGGGGGCGCTGCCGCTGCCTGAGAACTGCGGCT
GCCGCTGGAGGAGGTGCCGTCGCCTCCGCCACCGCTGCCGCCGCCGCC
AGGGGTAGGAGCTAAGCCGCCCATTTTGTGTCCCCCTGTTGTTGTGCT
TGACATGAATCCGACATGACACTGATTACAGCCCAATGGAGTCTCATTAA
ACCCGAGTCGCGGTCCCGCCCCGCCGCTGCTCCATTGGAGGAGACCAAAG
ACACTTAAGGCCACCCGTTGGCCTACGGGTCTGTCTGTCACCCACTCACT**

**AACCACTCTGCAGCCCATTGGGGCAGGTTCCCTGCCGGTCATNTCGCTTCC
AATAAACACACCCCTTCGACCCCATNATCCCCCCTTCGGGAACCAACC
CCGGGGGAGGGGTCCACTGGNCAATACCAATTNAANAGAACCGCTNGGG
TCCGCCTNTTNCGGGCNCCCTATTGGGTT**

ES 2 344 071 T3

TABLA 4 (continuación)

STAR4 directo

5
10
15
20
25
30
35
40
45
50
55
60
65

**GGGGAGGATTCTTTTGGCTGCTGAGTTGAGATTAGGTTGAGGGTAGTGAA
GGTAAAGGCAGTGAGACCACGTAGGGGTCATTGCAGTAATCCAGGCTGG
AGATGATGGTGGTTCAGTTGGAATAGCAGTGCATGTGCTGTAACAACCTC
AGCTGGGAAGCAGTATATGTGGCGTTATGACCTCAGCTGGAACAGCAATG
CATGTGGTGGTGTAATGACCCAGCTGGGTAGGGTGCATGTGATGGAACA
ACCTCAGCTGGGTAGCAGTGTACTTGATAAAATGTTGGCATACTCTACATT
TGTTATGAGGGTAGTGCCATTAAATTTCTCCACAAATTGGTTGTCACGTAT
GAGTGAAAAGAGGAAGTGATGGAAGACTTCAGTGCTTTTGGCCTGAATAA
ATAGAAGACGTCATTTTCAGTAATGGAGACAGGGAAGACTAANGNAGGGT
GGATTCAGTAGAGCAGGTGTTTCAGTTTTGAATATGATGAACTCTGAGAGA
GGAAAACTTTTTCTACCTCTTAGTTTTTGNGNCTGGACTTAANATTAAAG
GACATANGACNGAGANCAGACCAAATNTGCGANGTTTTTATATTTACTT
GCNGAGGGAATTTNCAAGAAAAAGAAGACCCAANANCCATTGGTCAAAA
CTATNTGCCTTTTAANAAAAAGANAATTACAATGGANANANAAGTGTTGN
CTNNGCAAAAATTGGG**

ES 2 344 071 T3

TABLA 4 (continuación)

STAR4 inverso

5
10
15
20
25
30
35
40
45
50
55
60
65

**GGATTNGAGCTAGCGGCGCGCCAAGCTTGGATCTTAGAAGGACAGAGTG
GGGCATGGAAATGCACCACCAGGGCAGTGCAGCTTGGTCACTGCCAGCTC
CNCTCATGGGCAGAGGGCTGGCCTCTTGCAGCCGACCAGGCACTGAGCG
CCATCCCAGGGCCCTCGCCAGCCCTCAGCAGGGCCAGGACACACAAGCCT
TTGACTTCCTCCTGTCAGTCTGCTGCCATTCTGTTTTGTGGTCATCACT
CCTTCCCTGTCCTCAGACTGCCCAGCACTCAAGGATGTCCTGTGGTGGCA
TCAGACCATATGCCCTGAANAGGAGTGAGTTGGTGTTTTTTGCCGCGCC**

**CANAGAGCTGCTGTCCCCTGAAAGATGCAAGTGGGAATGATGATGNTCAC
CATCNTCTGACACCAAGCCCTTTGGATAGAGGCCCAACAGTGAGGATGG
GGCTGCACTGCATTGCCAAGGCAACTCTGTNNTGACTGCTACANGACANT
CCCAGGACCTGNGAAGNNCTATANATNTGATGCNAGGCACCT**

ES 2 344 071 T3

TABLA 4 (continuación)

STAR6 directo

5
10
15
20
25
30
35
40
45
50
55
60
65

CCACCACAGACATCCCCTCTGGCCTCCTGAGTGGTTTCTTCAGCACAGCTT
CCAGAGCCAAATTAACGTTCACTCTATGTCTATAGACAAAAAGGGTTTGTG
ACTAAACTCTGTGTTTTAGAGAGGGAGTTAAATGCTGTAACTTTTTAGGG
GTGGGCGAGAGGAATGACAAATAACAACCTTGTCTGAATGTTTTACATTTT
TCCCCACTGCCTCAAGAAGGTTCAACAACGAGGTCATCCATGATAAGGAGT
AAGACCTCCAGCCGGACTGTCCCTCGGCCCCAGAGGACACTCCACAGA
GATATGCTAACTGGACTTGGAGACTGGCTCACACTCCAGAGAAAAGCATG
GAGCACGAGCGCACAGAGCANGGGCCAAGGTCCAGGGACNGAATGTCT
AGGAGGGAGATTGGGGTGAGGGTANTCTGATGCAATTACTGNGCAGCTC
AACATTCAAGGGAGGGGAAGAAAGAAACNGTCCCTGTAAGTAAGTTGTNC
ANCAGAGATGGTAAGCTCCAAATTTNACTTTGGCTGCTGGAAAGTTTNN
GGGCCNANANAANAACANAAANATTTGAGGTTTANACCCACTAACCCN
TATNANTANTTATTAATACCCCTAATTANACCTTGGATANCCTTAAAATAT
CNTNTNAAACGGAACCCTCNTTCCNTTTNNAATNNNAAGGCCATTN
NGNNCNAGTAAAAATCTNNNTTAAGNNNTGGGCCNAACAAACNTNTTC
CNAGACACNTTTTTNTCCNGGNATTTNTAATTTATTTCTAANCC

ES 2 344 071 T3

TABLA 4 (continuación)

STAR6 inverso

ATCGTGTCTTTCCAGGGACATGGATGAAGCTGGAAGCCATCATCCTCAG
CAAACAAACACAGGAACAGAAAACCAAATACCACATGTTCTCACTCATAAG
TGGGAGCTGAACAGTGAGAACACATGGACACAGGGAGGGGAACATCACA
CACCAAGGCCTGTCTGGTGTGGGGAGGGGAGGAGCATCAGGACAAA
TAGCTAATGCATGTGGGGCTTAAACCTAGATGACGGGTTGATAGGTGCAG

CAATCCACTATGGACACATATACCTATGTAACAACCCNACCTTNTTGACAT
GTATCCAGAACTTAAAGGAAAATAAAAATTAATAAAAATTNCCCTGGAA
TAAAAAAGAGTGTGGACTTTGGTGAGATN

STAR8 directo

GGATCACCTCGAAGAGAGTCTAACGTCCGTAGGAACGCTCTCGGGTTCAC
AAGGATTGACCGAACCCAGGATACGTGCTCTCCATCTGAGGCTTGNTC
CAAATGGCCCTCCACTATTCAGGCACGTGGGTGTCTCCCCTAACTCTCC
CTGCTCTCCTGAGCCCATGCTGCCTATCACCCATCGGTGCAGGTCTTTCT
GAANAGCTCGGGTGGATTCTCTCCATCCCCTTCTTTCCCAAGAAAGAA
GCCACCGTTCCAAGACACCCAATGGGACATTCCCNNTCCACCTCCTTNTC
NAAAGTTNGCCAGGTGTTCNAAACAGGTTAGGGAGAGAANCCCCAGG
TTTNAAGTTNCAAGGCATAGGACGCTGGCTTGAACACACACACACNCTC

ES 2 344 071 T3

TABLA 4 (continuación)

STAR8 inverso

**GGATCCCGACTCTGCACCGCAAACCTCTACGGCGCCCTGCAGGACGGCGGC
CTCCTGCCGCTTGGACGCCAGNCAGGAGCTCCCCGGCAGCAGCAGAGCA
GAAAGAAGGATGGCCCCGCCCACTTCGCCTCCCGGCGGTCTCCCTCCCG
CCGGCTCACGGACATAGATGGCTGCCTAGCTCCGGAAGCCTAGCTCTTGT
TCCGGGCATCCTAAGGAAGACACGGTTTTTTCCTCCCGGGGCCTCACCACA
TCTGGGACTTTGACGACTCGGACCTCTCTCCATTGAATGGTTGCGCGTTC
TCTGGGAAAAG**

STAR18 directo

**TGGATCCTGCCGCTCGCGTCTTAGTGTTTCTCCCTCAAGACTTTCCTTCTG
TTTTGTTGTCTTGTGCAGTATTTTACAGCCCCTCTTGTGTTTTTCTTTATTT
CTCGTACACACACGCAGTTTTAAGGGTGATGTGTGTATAATTAAGGAC
CCTTGGCCATACTTTCCTAATTCTTTAGGGACTGGGATTGGGTTTGACTG
AAATATGTTTTGGTGGGGATGGGACGGTGGACTTCCATTCTCCCTAAACT
GGAGTTTTGGTCGGTAATCAAACTAAAAGAAACCTCTGGGAGACTGGAA**

ES 2 344 071 T3

TABLA 4 (continuación)

5

ACCTGATTGGAGCACTGAGGAACAAGGGAATGAAAAGGCAGACTCTCTGA
ACGTTTGATGAAATGGACTCTTGTGAAAATTAACAGTGAATATTCACTGTT
10 GCACTGTACGAAGTCTCTGAAATGTAATTAAGTTTTTTATTGAGCCCCCG
AGCTTTGGCTTGC GCGTATTTTTCCGGTCGCGGACATCCCACCGCGCAGA
GCCTCGCCTCCCCGCTGNCCTCAGCTCCGATGACTTCCCCGCCCCCGCCC
15 TGCTCGGTGACAGACGTTCTACTGCTTCCAATCGGAGGCACCCTTCGCGG

15

20

STAR18 inverso

25

TGGATCCTGCCGCTCGCGTCTTAGTGTTTCTCCCTCAAGACTTTCCTTCTG
TTTTGTTGTCTTGTGCAGTATTTTACAGCCCCTCTTGTGTTTTTCTTTATTT
25 CTCGTACACACACGCAGTTTTAAGGGTGATGTGTGTATAATTAAGGAC
CCTTGGCCCATACTTTCCTAATTCCTTAGGGACTGGGATTGGGTTTGACTG
AAATATGTTTTGGTGGGGATGGGACGGTGGACTTCCATTCTCCCTAAACT
30 GGAGTTTTGGTCGGTAATCAAACTAAAAGAAACCTCTGGGAGACTGGAA
ACCTGATTGGAGCACTGAGGAACAAGGGAATGAAAAGGCAGACTCTCTGA
ACGTTTGATGAAATGGACTCTTGTGAAAATTAACAGTGAATATTCACTGTT
35 GCACTGTACGAAGTCTCTGAAATGTAATTAAGTTTTTTATTGAGCCCCCG
AGCTTTGGC

40

45

50

55

60

65

ES 2 344 071 T3

TABLA 5

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

5

10

15

20

25

30

35

40

45

50

55

60

65

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 2 344 071 T3

TABLA 5 (continuación)

	Número	Secuencia
5	E16	GTTCGGGATCCACCTGGCCG
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	GTTCGTGGACACGACCTCCG
	D70	TACAAGCCAACCACGGCCT
60	D71	CGGAAGTGCTTGACATTGGG
65		

ES 2 344 071 T3

TABLA 6

Elementos STAR de la invención, incluyendo localización y longitud genómica

5

10

15

20

25

30

35

40

45

50

55

60

65

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 2 344 071 T3

TABLA 6 (continuación)

	STAR	Localización ¹	Longitud ²
5			
	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 2 344 071 T3

TABLA 6 (continuación)

5	STAR	Localización ¹	Longitud ²
	49	Clon al azar Chr 12, y 3q26.32	1015
10	50	6p21.31	2361
	51	5q21.3	2289
15	52	7p15.2	1200
	53	Xp11.3	1431
20	54	4q21.1	981
	55	15q13.1	501
25	56	Incluye 3p25.3	741
	57	4q35.2	1371
30	58	21q11.2	1401
	59	17 clon al azar	872
35	60	4p 16.1 y 6q27	2068
	61	7p14.3 y 11q25	1482
40	62	14q24.3	1011
	63	22q13.3	1421
45	64	17q11.2	1414
	65	7q21.11=28.4	1310
50	66	20q13.33 y 6q14.1	~2800
55	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/		
60			
65			

ES 2 344 071 T3

TABLA 6 (continuación)

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

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

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

25 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
30 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.

35 ⁵ Idéntico a la Tabla 4 "STAR18".

40

45

50

55

60

65

ES 2 344 071 T3

TABLA 7

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

5

	<u>Divisiones Celulares²</u>	<u>Expresión Luciferasa³</u>
10	STAR6 más Puromicina	18.000
15	60	23.000
	84	20.000
	108	16.000
20	STAR6 sin Puromicina ⁴	12.000
25	84	12.000
	108	15.000
	144	12.000

30

¹El 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.

35

²El 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.

40

³La luciferasa fue analizada como se describe en el Ejemplo 1.

45

⁴Al 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.

50

55

60

65

ES 2 344 071 T3

TABLA 8

Elementos STAR humanos y sus supuestos ortólogos y parálogos de ratón

5

10

15

20

25

30

35

40

45

50

55

60

65

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%

ES 2 344 071 T3

TABLA 8 (continuación)

5
10
15
20
25
30
35
40
45
50
55
60
65

¹Ubicación citogenética del elemento STAR en el genoma humano.

²Ubicación citogenética del elemento STAR ortólogo en el genoma de ratón.

³Longitud 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 importante.

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: 27 de la Figura 20;
- (b) un fragmento del SEQ ID: 27 de la Figura 20;
- (c) una secuencia derivada del SEQ ID: 27 de la Figura 20 por delección, inserción 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 $\mu\text{g/ml}$ 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 ARN 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 2 344 071 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

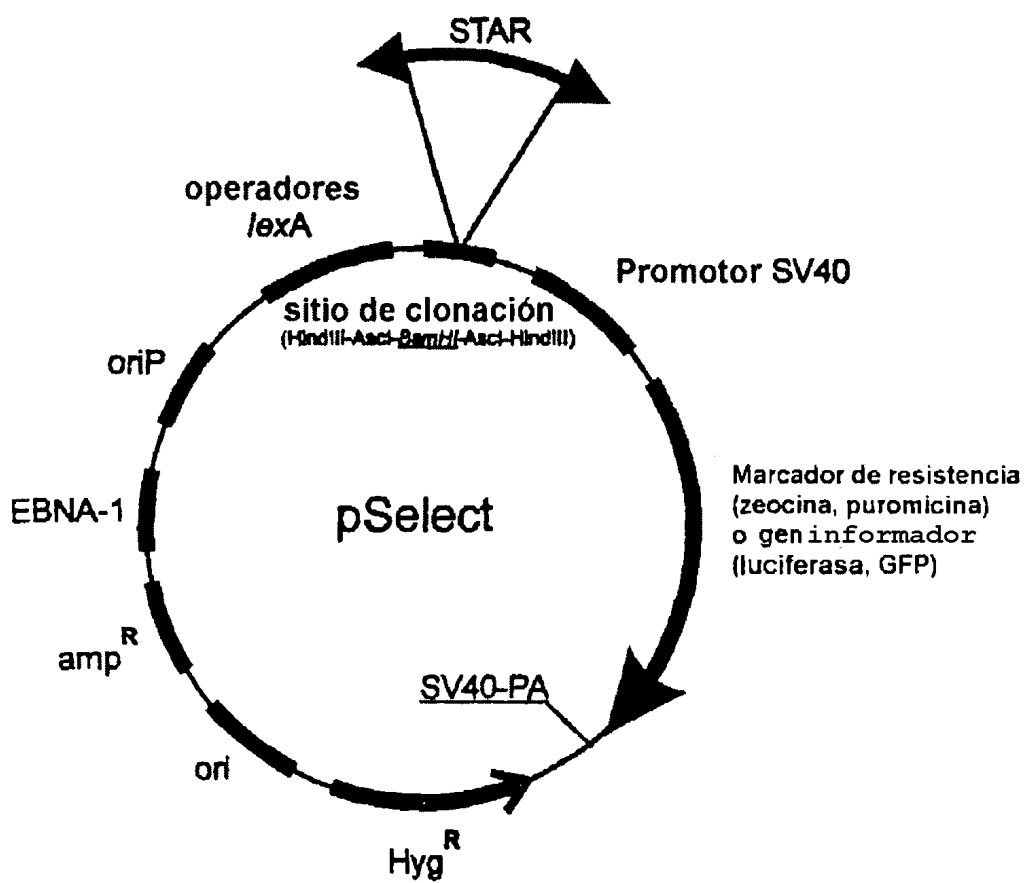


Fig. 2

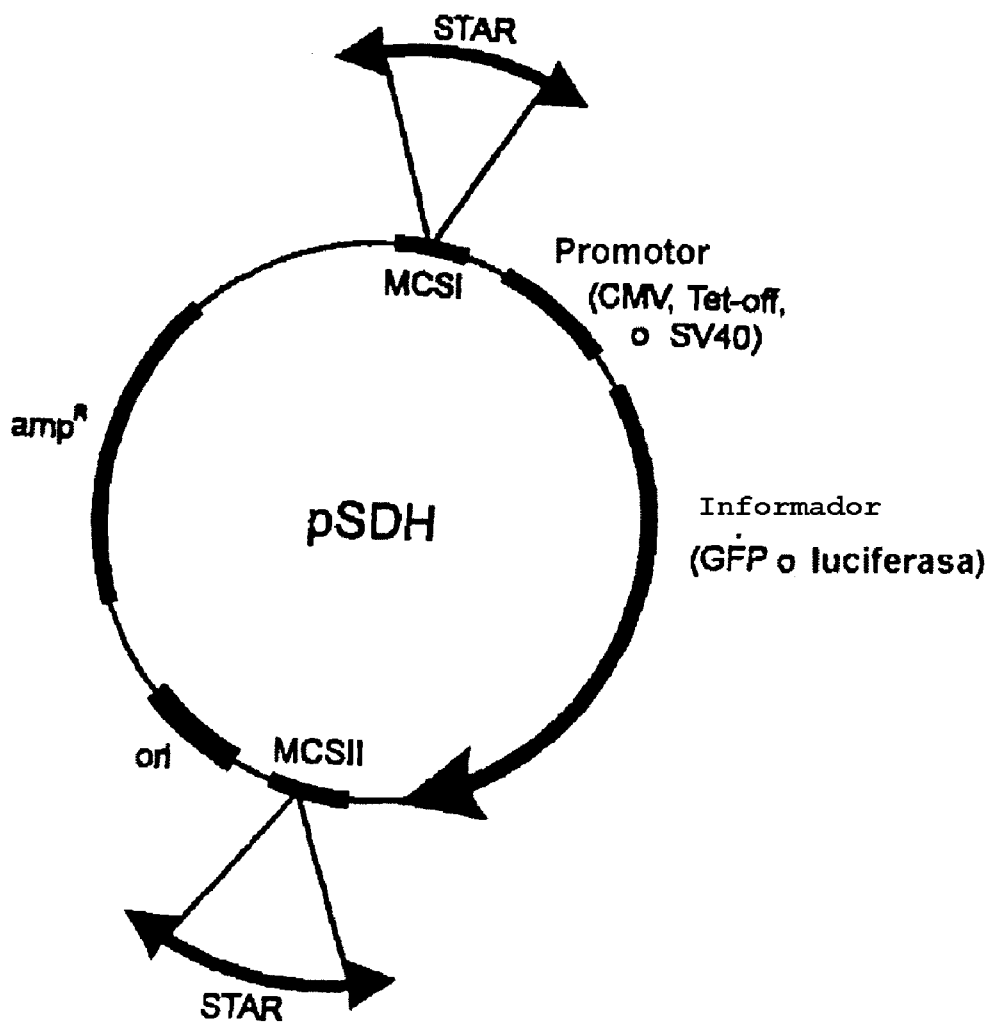


Fig. 3

PROPORCIÓN DE CLONES QUE SOBREVIVEN EXPRESAN LUCIFERASA

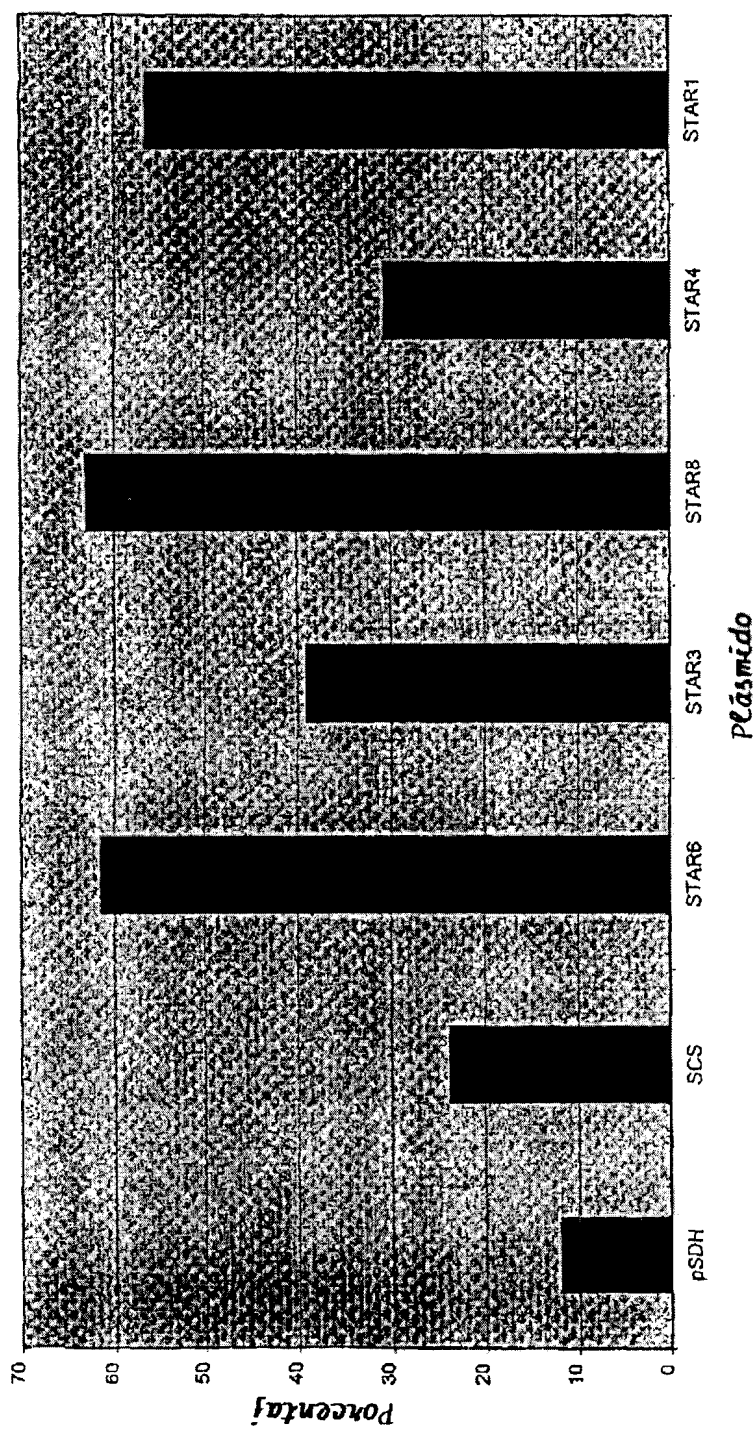


Fig. 4

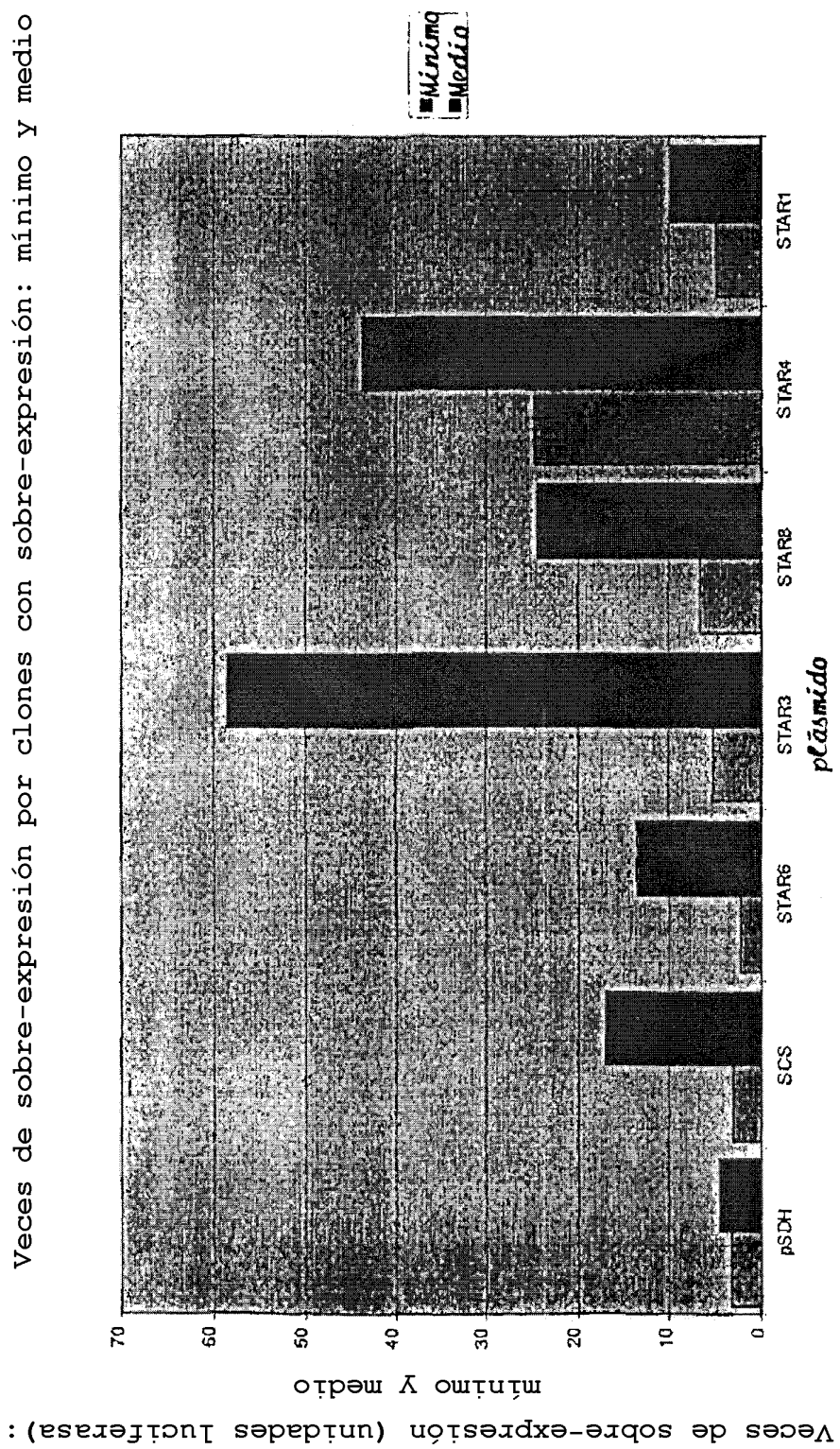


Fig. 5

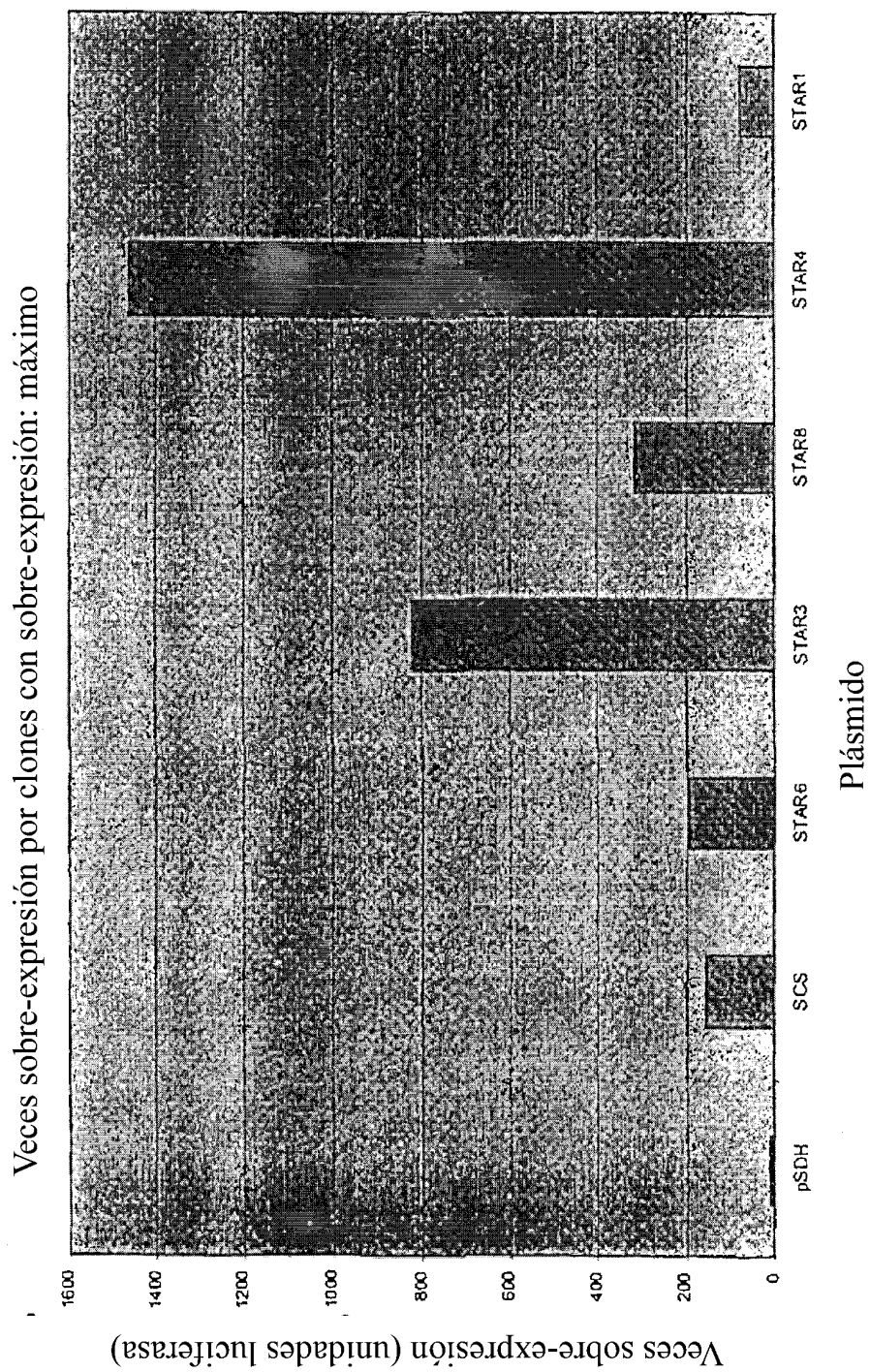


FIG. 6

Vector para someter a ensayo actividad
STAR

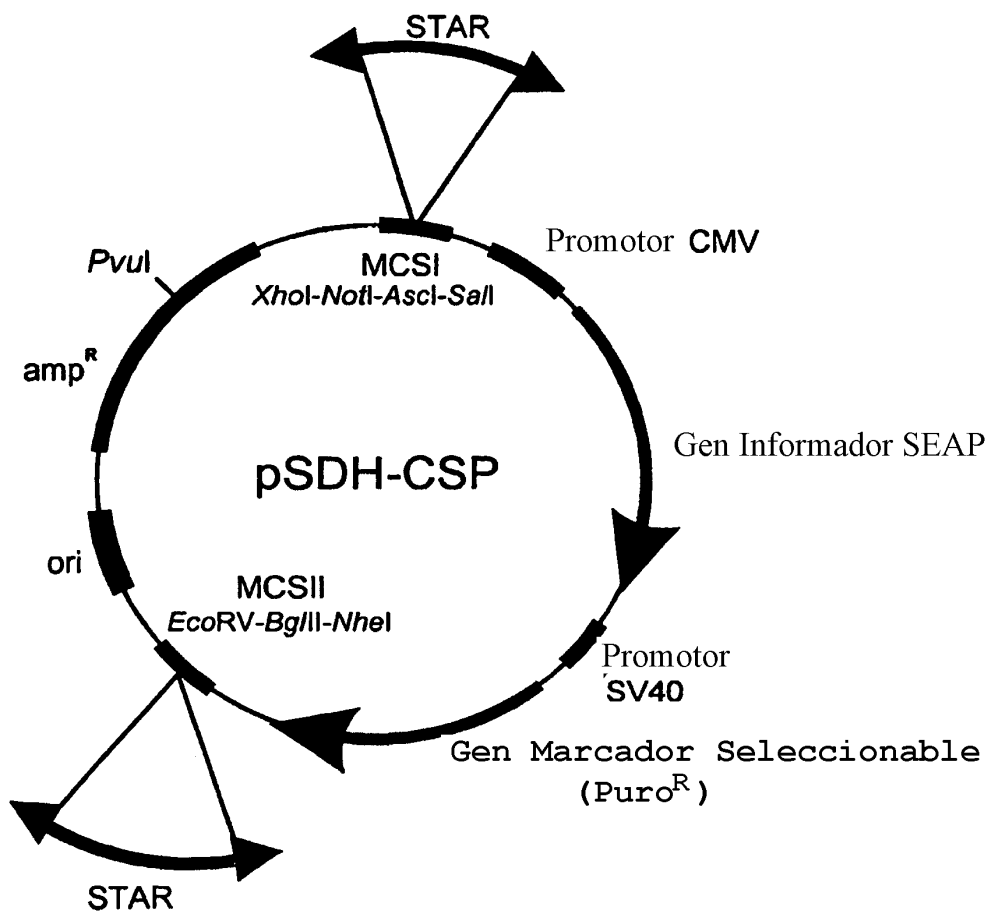


FIG 7

Células CHO, promotor CMV, SEAP, lineal

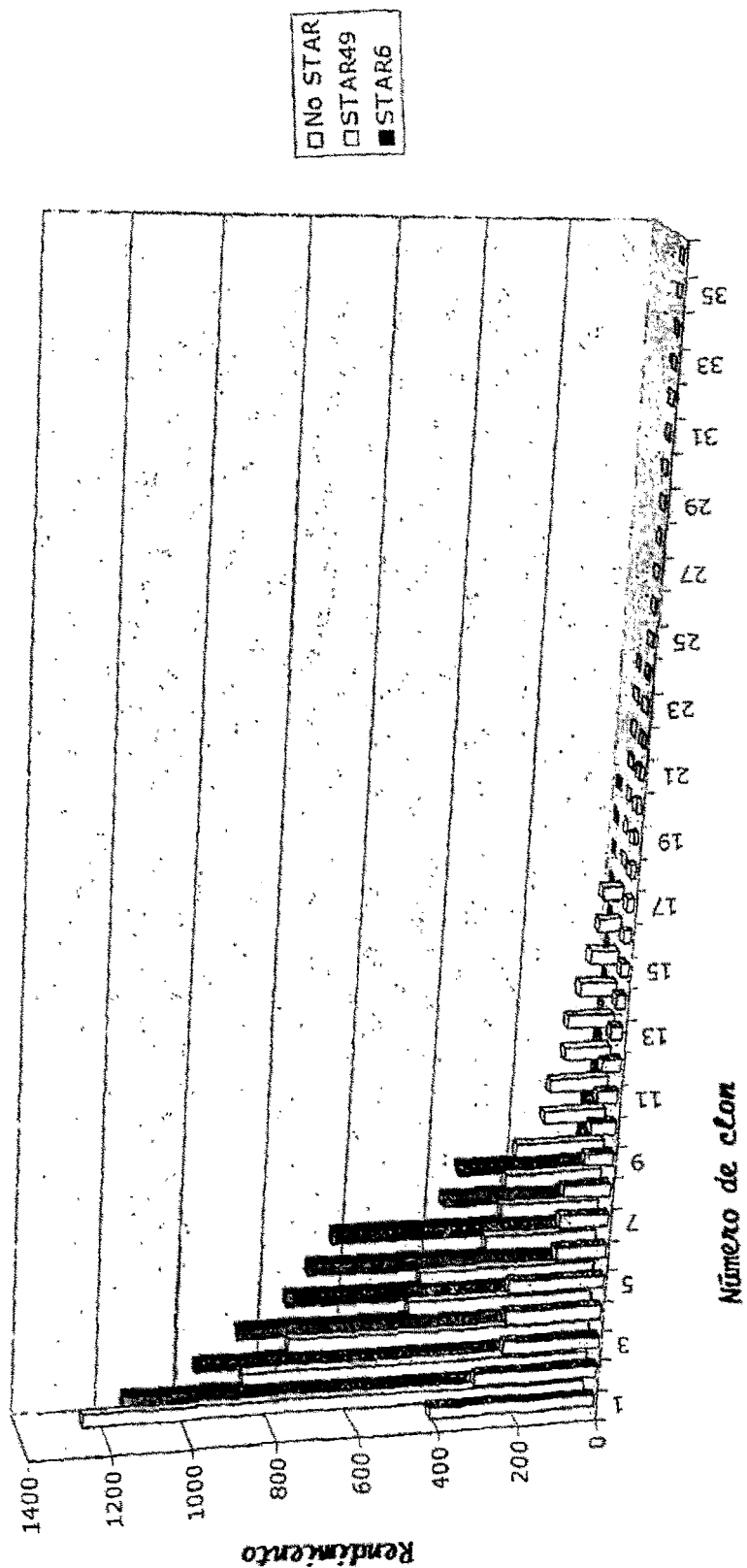


FIG. 8

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

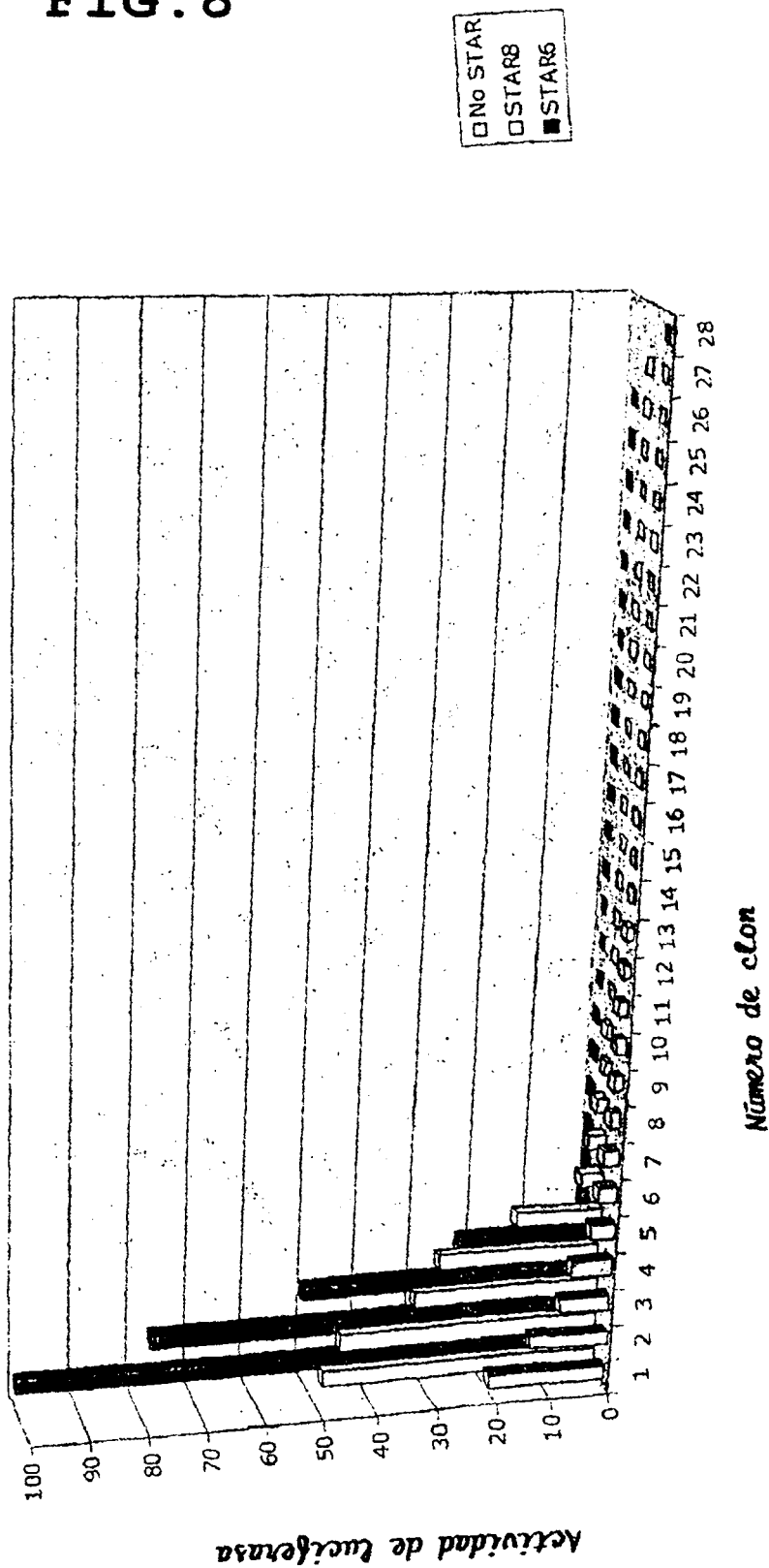


FIG. 9

Mapeo fino STAR









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

FIG. 10

Células U-2 OS, promotor SV40, luciferasa, circular

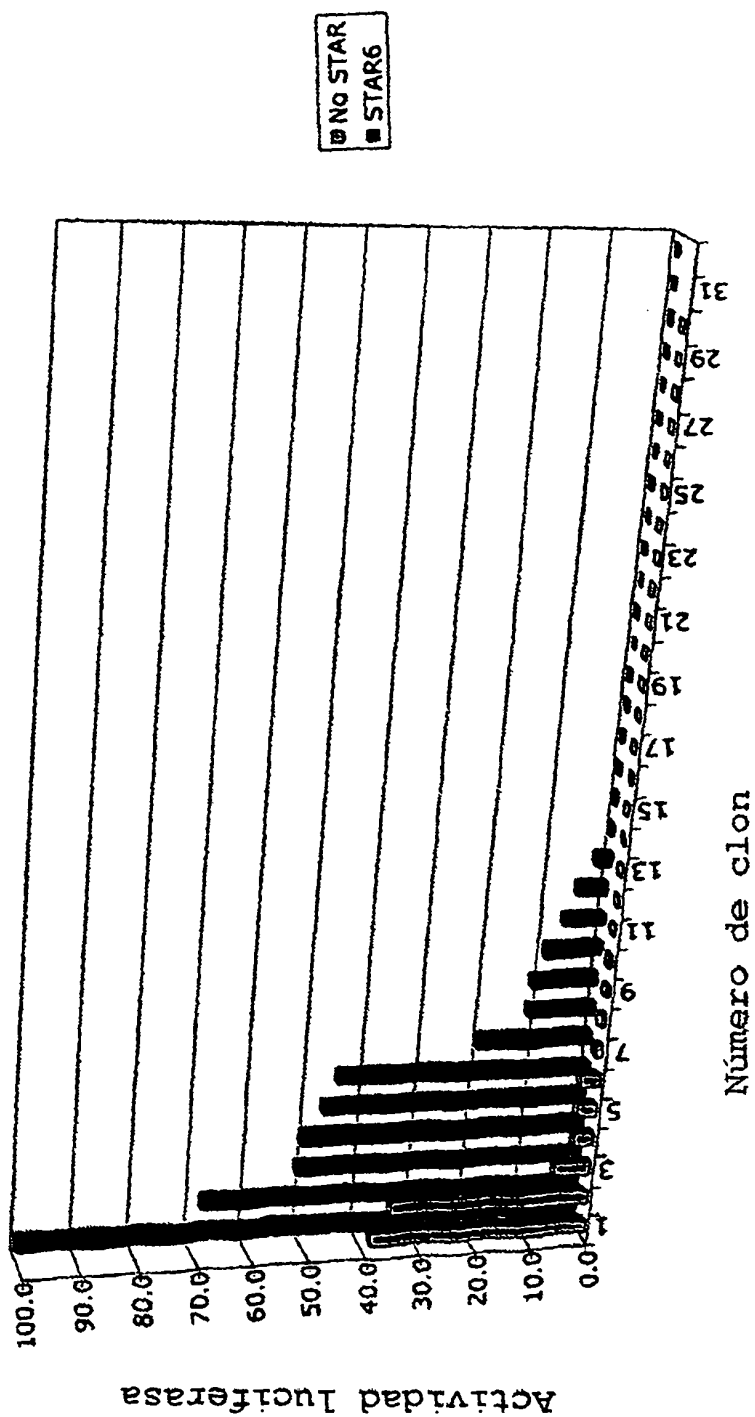
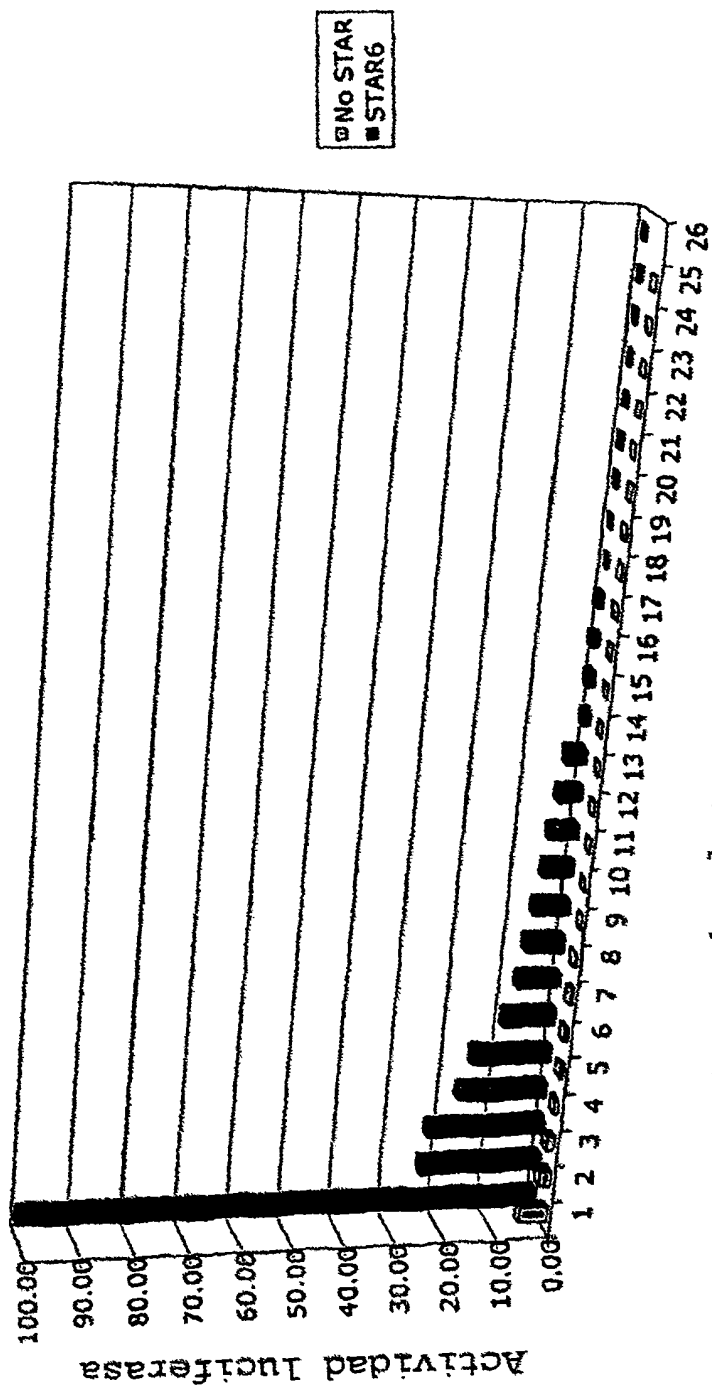


FIG. 11

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



Número de clon

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

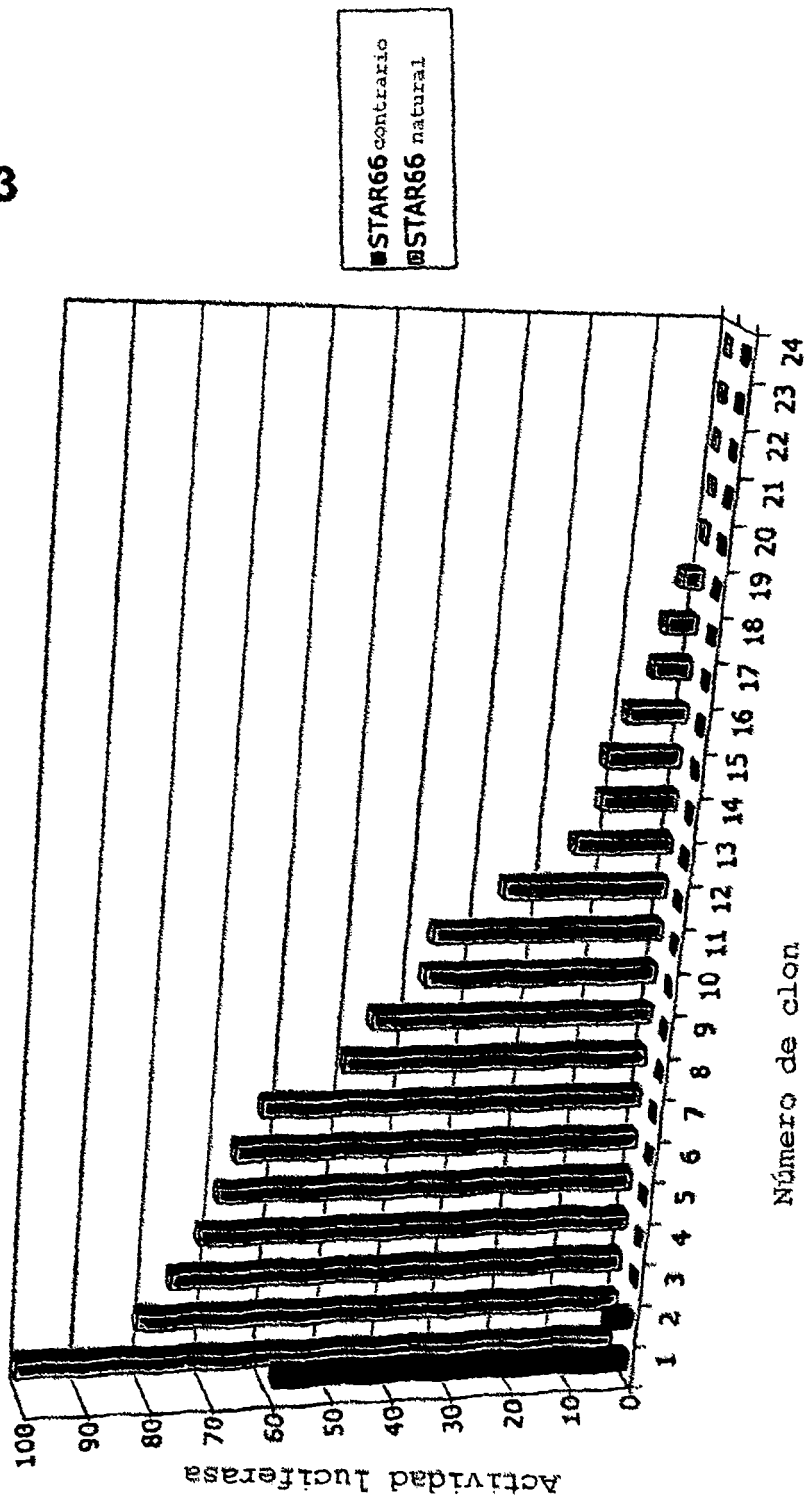


FIG. 14

Dependencia del número de copias STAR

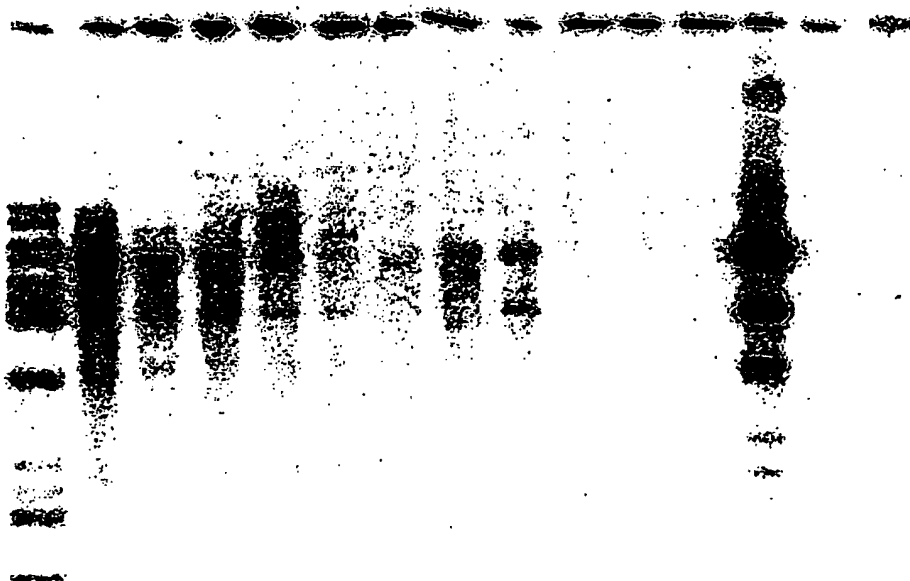


FIG. 15

Dependencia del número de copias STAR

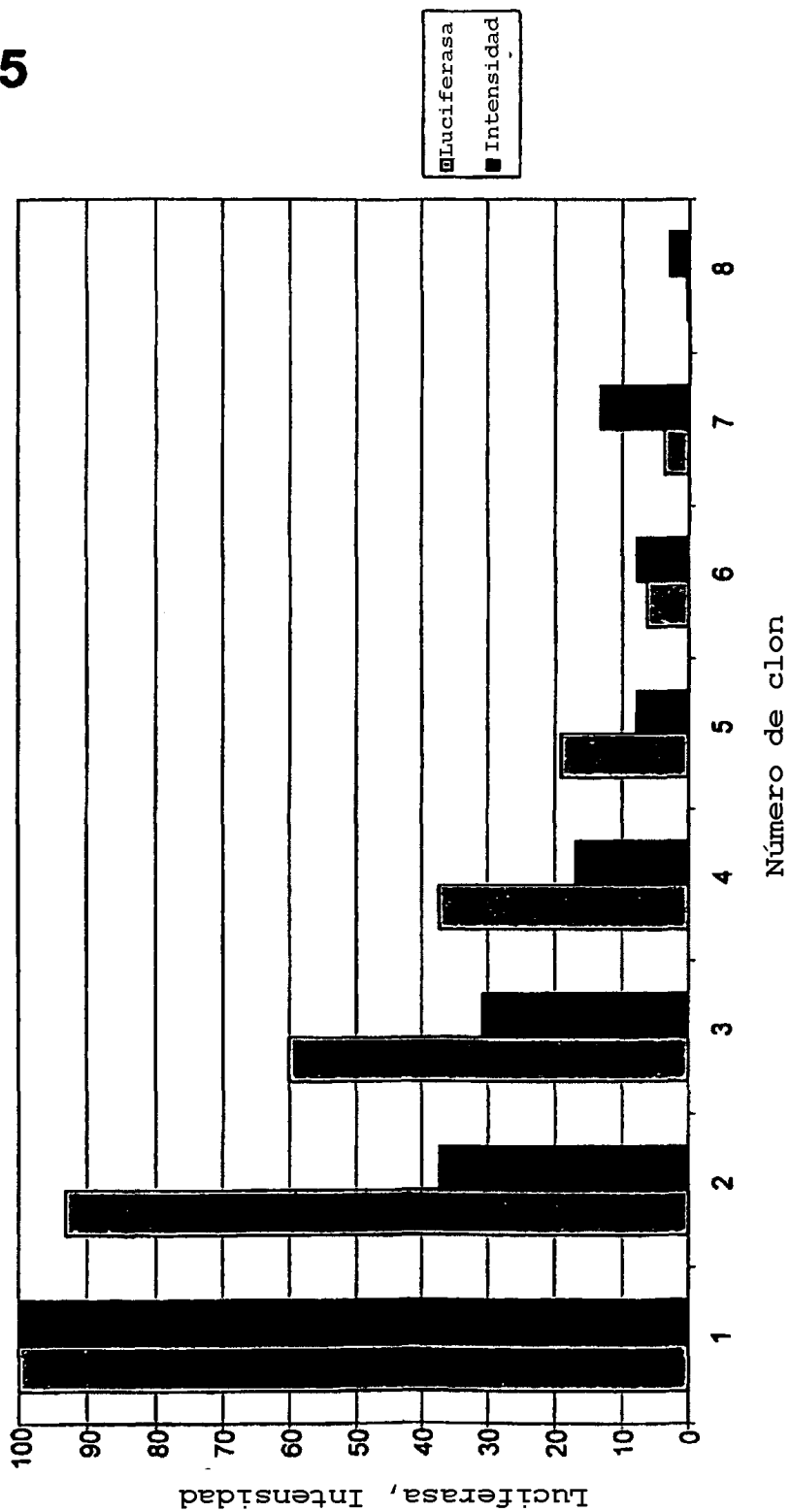
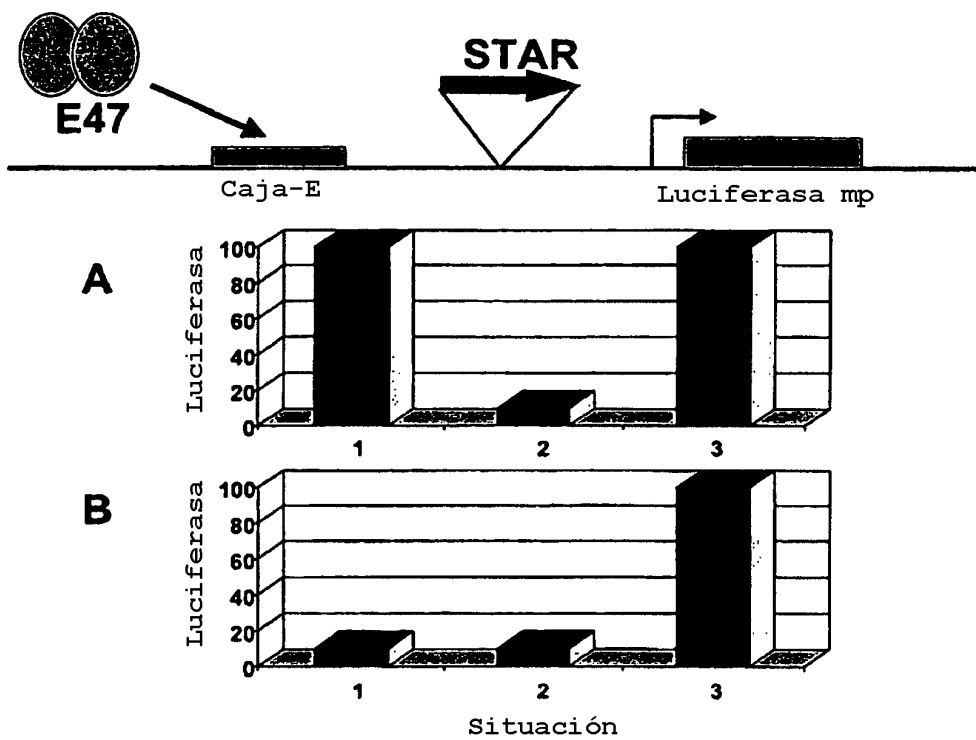


FIG. 16

Diagrama esquemático de Análisis de Intensificador y bloqueo de intensificador



Análisis de bloqueo de intensificador

FIG. 17

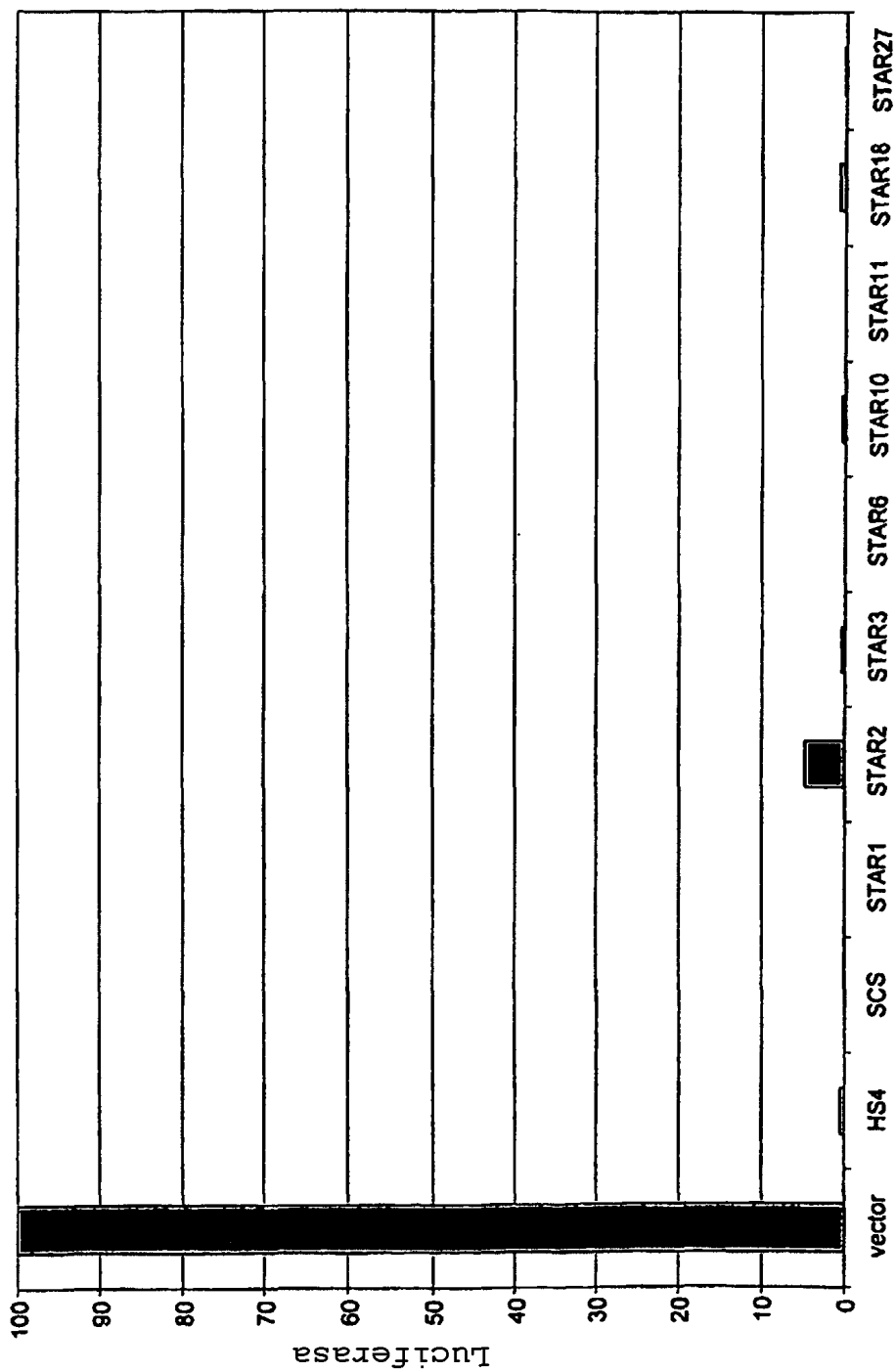


FIG. 18

Análisis intensificador

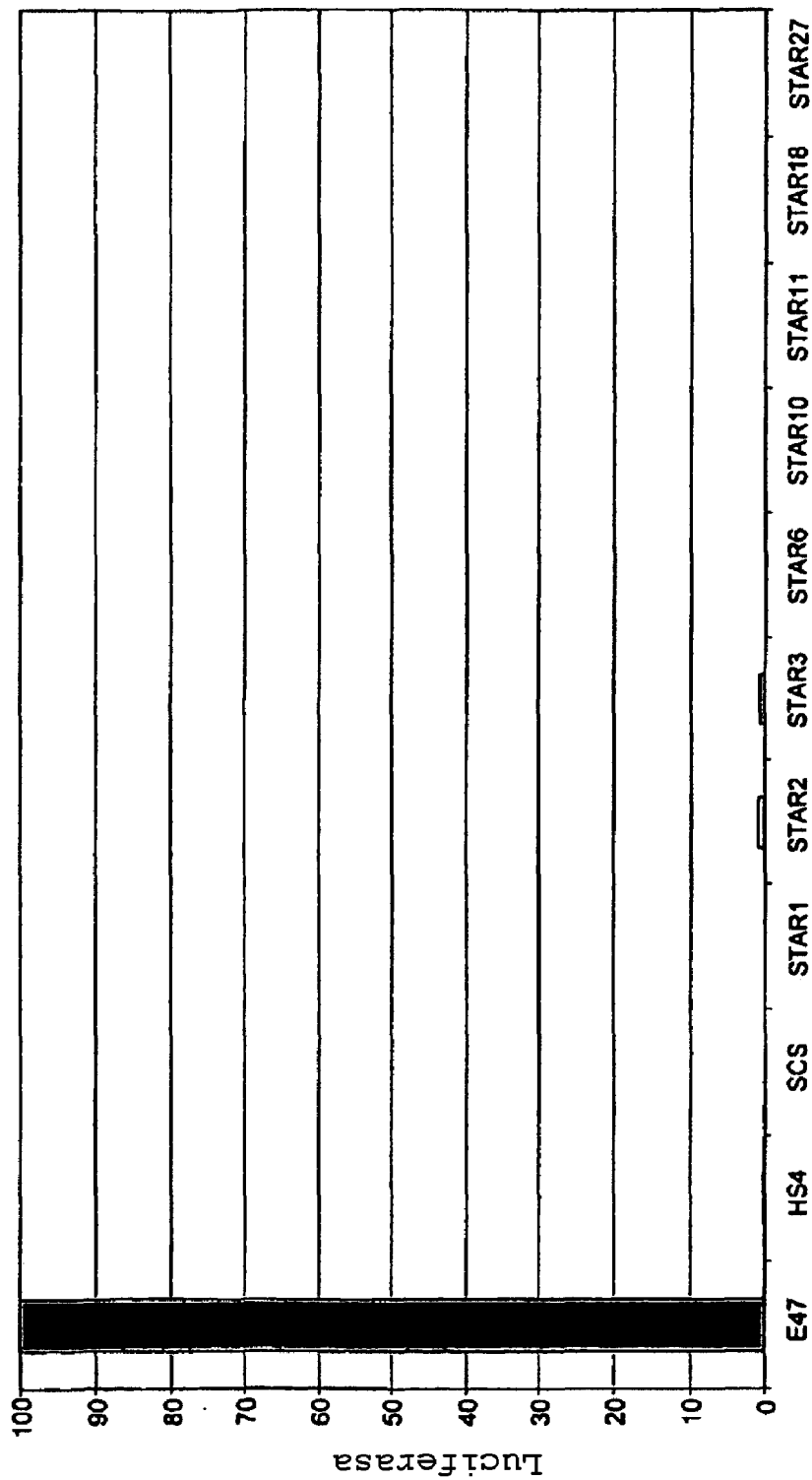


FIG. 19

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

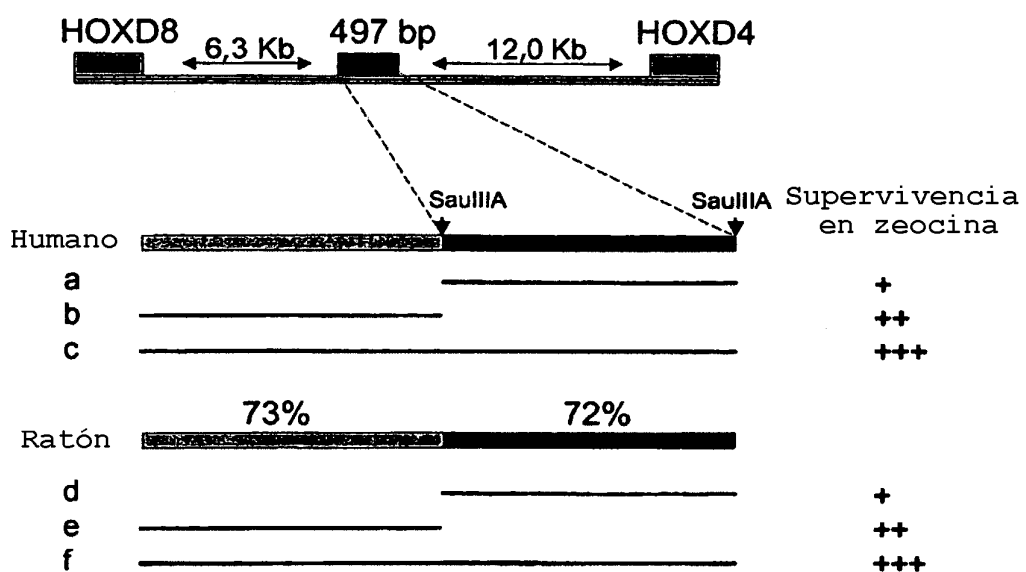


FIG. 20

```

SEQ ID: 1
>STAR1
ATGCGGTGGGGGCGCCAGAGACTCGTGGGATCCTTGCCCTGGATGTTTGAATCTTCTGAGTTGCCGTGCCCCGAAAG
ACAGGTACATTTCTGATTAGGCCCTGTGAAGCCCTCTGGAGGACCATCTCAATTAAGACGATGGTATTGGAGGGAGAGTCACA
AAAGAAGCTGTTGGCCCTCCCTCACTGCAAAAACGGAAAGTATTTATTTTAAATGGGAGTTGGAAATATGTGAGGGCTCGCAGG
AACCAGTCTCCCTCCTTCTTGGTTGGAAAAAGCTGGGGCTGGCCCTCAGAGACAGGTTTTTTGGCCCCGCTGGCTGGCCAGT
CTAGTCGACCCCTTTGTAGACTGTGCACACCCCTAGAAGAGCAACTACCCCTATACACCAGGCTGGCTCAAGTGAAGGGGGC
TCGGGGCTCCAGTCTGGAAAACTGGGTGCTCTGGGGACCTCTGGTCTGCTCTCTCTCCCTGCCACTGGCTCTGGGTGC
TTATCTCTGCAGAACCTCTCTGCTAGCAAAACCCACATTCAGCGCCCTGTAGCTGAACACAGCACAAAAGCCCTAGAGATC
AAAACCATTAGTA TGGCAGTTGAGGGGGAGGTGAAATTTAAACGCTTTTGTTCATCAATAACTCGTTGGCTTTGACCTGT
CTGAACAAGTCAGCAATAAGGTGAAA TGCAGGTCA CAGCGCTAACAAAATGAAAAATGTGTAATATCACCCCGGCTCC
AGCCGGCGCCAGGCTCC

SEQ ID: 2
>STAR2
GGTGTCTCC TGAATCTTCC TGAGAAAGGATGGTGGCCGTAAGTCCGTGTAGGTGGGGTCCGGCTCCCCAGGCCCGG
CCCGTGGTGGTGGCCGCTGCCAGCGGGCCGGCACCCCATAGTCCATGGCCGCGGAGGCAGCGTGGGGGAGGTGAGTTAG
ACCAAAAGGGGTGGCCCGGAGTTGCTCA TGGGCTCCACA TAGCTGCCCCCAAGAAAGACGGGGCTTCCCTGTA TGTGTGG
GGTCCCATAGCTGCGCTTGCCCTGCAGGCCATGAGCGTGGGGTCA TAGTGGGGGTGCCCCCTCGCCCGCCCTTCCCGC
CGTGTAGCGCTTCTGTGGGGTGGCGGGGGTGGCAGCTGGGCAGGGACCGAGGGTAGGAGGGGGGGGAGCCCGTAGGT
ACCTGGGGGGCTTGGAGAAGGGCGGGGGCGAGTGGGGCTCA TACGGGACCTGTTCACAGCGAAATGCA TAGAGTTACG
ATAGCCACCGGCTCCGGGGGCGAGGGGCTGCGACTTGGAGACTGGCCCCCGATGACGTTAGCATGCCCTTGCCCTTCTG
ATCTTTTGTACTTTCATGGCGGATTC TGGAAACAGATCTTGAATCTGGCGCTCAGTGAGGTT CAGCAGATGGCCATCTC
CACCCGGCGGGCCGCA CAGGTAGCGTGAAGTGGAACTCTTCTCCAGCTCCACAGCTGCGCGCTCGTGTAGGGCGT
GGCCCGCGCTTGGAGCAAGCTGCCCGGGGGGCTTGTGCGCCAGCGCAGCTTTCGCTGGAGGACAGAGAGAGGAAAG
AGCGGGCTCAGGGGCTGCGCGGGCCCGCCAGCCCTGACCCAGCCCGGCCCTTCTCCACAGGCCCAA

SEQ ID: 3
>STAR3
ATCTGGAGTACTGAAA TAGGAGTAAATCTGAAAGCAAA TAAGATGAGCCAGAAAACCA TGAAGAAGAACAGGACTACCAG
TTGATTCACAAAGGACATTC CCAAGGTGAGAAGGCCATATACCTCCACTACCTGAAACCAATCTCTGTA TGCAGATTTAGC
AAGGTTATAAGGTAGCAAAAGATAGACCCAAAGAAAATAGAGAACTTCCAATCCAGTAAAAATCATAGCAAAATTTATTGAT
GATAACAATTTCTCCAAAGGAACAAAGGCAGAGTCGTGTAGCAGAGGAAGCACG TGAGCTGAAAACAGCCAAA TCTGCTT
TGTTTTCA TGACACAGGACATAAAGTACACACCACCACTGACCTATTAAGGCTGTGGTAAACCGATTCATAGAGAGAGG
TCTAAAATACATGGTCCCTCACAGGCAAACTGCAGTTCGCTCCGAAAGTAGTCCCTGGAAAATTTGATGTCCAGTATAGAA
AAGCAGAGCAGTCAAAAATATAGATAAAGCTGAAACAGATGTTCCTTGGGCAATGTAGCAGCACCACTAAGATAATA
ACCTCAGCTGTGGACTCCCTCCCTGGGGAGCGGTGCTGCGGGGGGGGGGGCTCCGCAACTCCCGGGCTCTCTGCGCC
GCCCTCCGTTCTCTCGGGGGGGGGGGGGGGGGGGGACTGCGCGCTCACAGCGGGGGCTCTCTGCGCCCGGGCTCGGA
GGCAGTGGCGGTGGCGGCA TGGCTTCTGGCTTCCGCGATGT CAGCAATTCGAACTGAGGGTCACTCTCTTGGGACTGGT
TAGACAGTGGTGCAGCCACGGAGGGCGAGTTGAAGCAGGGTGGGGTGCACCTCCCCAGGAAGTCCAGTGGGTCAGGG
AACTCCCTCCCTAGCCAAGGGAGGCGGTGAGGGACTGTGCCCGGTGAGAGACTGTGCCCTGAGGAAAGGTGCACCTGCGC
CCAGATACTACATTTTCCACCGGCTTCAAAAACCCGAGACCAAGGAGATTCCTCGGGTTCTACACCACAGGACCCCTG
GGTTCAACCAAAAACCGGGCCATTTGGGCAGACACCAGCTAGCTGCAAGAGTTGTTTTTTTTTTA TACTCTCTGGC
ACCTGGAACGCCAGCGAGAGACACTTTCACTCCCTGGAAAAGGGGGCTGAAGGCAGGGACCTTTAGCTGCGGGCTAGGG
GCTTTGGGGTGTAGTGGGGAGGGGAGAGGAAAAGCCCTCGTCAATGGCGTCTCTCAGCCAAATAAGGCTACCTCTCTC
TGCTGCGAGTAGACCCAA TCCTTTCTAGAGGTGAGGGGGGGGGTGGTGGAACTAGAGGTGGCCGGTATCTAGGAGAG
AGAAAAGGGCTGGAACAA TAGGTGCCCGGAAGAGGGGACCCAGCGGCTGTGATTTGGTATTGGCAG TGGACCCCTCCC
CGGGTGTGCGCGAGGGGGGATGATGGGTGGAGGGGTGTGTTTATG TGGAAAGCAGATGACCGGCAAGAAC TGCCCCA
ATGGGCTGCAGAGTGGTTAGTGTGGTGA CAGACAGACCCGTAGCCCAACGGGTGGCCTTAAGTGTCTTTGGTCTCCTC
CAATGGAGCAGCGGGGGGGGGGAGCCGACTCGGGTTAA TGAAGCTCCATTTGGCTGTAA TCAAGTCA TGTGCGGATC
ATGTCAACGACAACAAGGGGACACAAA TGGCGGGGGCTTACTCTACCCCTGGCGGGGGGAGCGGTGCGCGAGG
CGACCGCACTCTCCAGGGGGCAGCGCAGTTTCAGGCAGCGCCAGCCCGGGCAGGGCGGGTGGCGGCA
GCCAGGCTGTCAACCCACCCCGGCTTCCAGGGGGAGGAGACTGGCGGGAGGGGGAA CAGACGGGGGGGATTCAGG
GGCTTGGAGCGCCCTCCACAGGCTCTGCGGGAGGGTCA CCGGGGGGGCTCGGGGTGAGGCTGCCCTGAGCGGTGAC
GGTAGGGGGGGGAAAGGGGAGGAGGACAGCCCGCCCTCGGCAGGGCTCTAGGGCAAGGGGGGGGGCTCGGAGG
AGCGGAGGGGGGGGGGG

```

FIG. 20, cont

CTTGATAAAATGTTGGCACTACTTAGATTGTTATGAGGGTAGTGCCATTAATAATTCTCCACAAATGGTTGTCACGTATG
 AGTGAAAAAGAGGAGTGA TGGAAGACTTCAGTGC TTTTGGCCTGAAATAA TAGAAGACGTCA TTTCCAGTTAA TGGAGACA
 GGGAGACTAAAGGTAGGGTGGGATTTCAGTAGAGCAGGTGTCAGTTTGAATATGATGAACCTCGAGAGAGGAAAAACTT
 TTTCTACCTCTTAGTTTTGACTGGACTTAAAGAA TTAAGTGACATAAGACAGAGTAACAAGACAAAAATATGCGAGGT
 TATTTAAATATTTTACTTGCAGAGGGGAATCTTCAAAGAAAAATGAAGACCCAAAGAACCCATTAGGGTCAAAGCTCAT
 ATGCOCTTTTAAAGTAGAAAA TGAATAATTTAACAA TGTGAGAAGACAAAGGTGTTGAGCTGAGGGCAA TAAATTTGTTGG
 ACAGTGA TTAAGAAA TATATGGGGAAA TGAATGATAAGTTA TTTTAGTAGATTTATCTTCATA TCTATTTGGCTTCA
 ACTTCCAGTCTCTAGTGA TAAAGATGTTCTCTCTTCCCTGGTACAGAGAGACCTTTCTCATGGGAAATTTATGACCT
 TGCTGAAGTAGAAAGGGGAAGATCGATCTCCTG TTTCCAGCATCAGGATGCAAAACATTTCCCTCCA TCCAGTTCTCAA
 CCCCATGGCTGGGCTCA TGGCATTTCCAGCATGCTATGAGTGCACCTTTCTCGCAGGC TGCTCGGGTAGCTGGTGCAC
 GCTAGGTCAGTCTATG TGCAGCAGGACTGGGCTCTGGGCAA TGCAGTTGGCAGCCCCCA TCCCTCCACTGCTGGGGCC
 TCCATCCAGAAAGGGCTGG TGTGCAGAACGATGGTGCACCA TCA TCA TCCCACTTGGCA TCTTTCAGGGGACAGCCAG
 CTGCTTTGGGGCCGCAAAAACACCAACTCAC TCCCTTTCAGGGGCTCTGGTCTGA TGCACCACAGGACATCTCTGA
 TGTCTGGCAGCTCGAGCACAGGAAAGGAGTGA TGACCA CA AAA CAGGAA TGGCAGCAGCAGTGACAGGAGAACTCAAAG
 GCTTGTGTCTCTGGCCCTGCTGAGGGCTGGCGAGGGCCCTGGGA TGGCGCTCAGTGCC TGGTGGCTGCAAGAGGCCAGC
 CCTCTGCCATGAGGGGAGCTGGCAGTGACCAAGCTGCACCTGCCCTGGTGGTGCATTTCTTGCCTCCACTTTCTCTTAA
 GATCC

SEQ ID: 5

>STAR5

AGCAGAGATCTTATTTCCCGTATTCCTTGTGGCCACAGCACCCTCCACGCCAAAGCAAACCAAGCAAAGGAGGCCCTGAT
 GAGGAGGGGCTTCCCCAACTGGCTCCACAGGCTCTACATACGTACCCACCCAGACACACAGAGCTGCTTCTCTGCT
 CTCACACAGACTGAGCTGTGCCAGACATTTCCCTTAGCACTAACCAACTCTTTCAAAAATACATTTTTCTTAAAAGA
 ACAAGTTAAACAAGTTGACTCATTTAAGAAC TGTAGAAAGATAACC TTGTTTAAATA TGTATTTGCAGAAAT
 TGGAGCAGAAAGTTACCAACATTTGCTGGTTCACGCCAGGAGGTAGAGCGTGGTGGCATCCAGAACCTTCC TCCACTC
 CTGCCCTGGCTGGTTTTTATCA TCTTTGTA TCCCAAGAACTTCTCAGTGTCTCAGGAGTGTAGGCAC TCACTACCTG
 TTTGGTAGTTACATGAA TGAATGCA TAA TGACTAAGTGAGTTAA TGA TGAAGCTAAT TGTCTCTCCCTTTGCTTTTCCA
 GAGCTTTCCAAGGTAAAGTGTGGACACTCTTTCTTCA TCTCAGATTTAATCAACTAAGAA TGTCTGCAAA TGAACACCA
 GCTCACAAA CT CAGGAA TACATGAAAAGCA TTTGCTTATTTTAACTAAC TCAAA TCTATGTCTCAGTCTCCCTTTTAT
 GCTGGA TGTGGGCTTAA TCTCAGTGGTTCCTCA TTTGOCAGACCTG TGTCCAGTTTGGGGCTTCA CATAGAGCCAC
 CCAATCACAGGAGGGAAGGGCTTGTCTTTGGCTTCCACTCCACCTCTCTCTCTCCGAGCTTTGA TGTCTCACTTT
 CCTTTTACCACCTGGGAAGCTTCTGCCATGATACATTTAGACCTCAA TGTAA TGGCAATTTGGGTTTGGGTTCTCATA
 AACTCAGAAGTCCAGGAAA TCGCC TGC TGCCTCCACACAC TCTGAGGGCA TCTGGAATCTTACCCTTACCTGGAGC
 CTGCTGGCTCAACTGTT TGAAGTCTG TGTCTGGGCA TGCAGGTAAT TGGAGGATCTCTGTGGCA TAAAAA TACCC
 GAAGTCCACCTCAAAGTTGA TGCAGGGTCTCTGCA TTTCA TTTGCAAAATGTTCTATCATTTCTATAGTTTTCAGCTTAC
 AGTCAAGGGCCAGGACTTTGCAOCTTTGGTAAACCTCAATCTCTCTCTCTCTGGCTTCTACTCTCTTCTCCCTCAATCC
 CAAATCAAGGCCCTTGA TGTCTGGAGGTAGGAAAGCCTGGTCTGCTCA TGA TATAG TCTACATCATAGCCTTTGTCAT
 CTCA TGTGATTCATCAACAA CCGTGTGTGGA TGGGGCCACCAATA TGTGCCAGGAGTTGAGGACACCGAGGTTATGATG
 ATGAAA TAGATAAGGGGCCACACTCAGGAACTCGAGGAGTGGAGCTGTGGACCCAGCATGCGAGTAAAGACCCAGT
 GAGCTCACAGACAGATCATTTAAATCAGGT

SEQ ID: 6

>STAR6

TGACCCACCACAGACATCCCTCTGGCTCTGAGTGGTTTCTTCAGCACAGCTTCCAGAGCCAAATTAACCGTTCACCT
 ATGCTATAGACAAAAGGGTTTGGACTAACTCTG TGT TTTAGAGAGGGAGTTAAATGCTGTAACTTTTATAGGGTGGG
 CGAGAGGGATGACAAAATAACAAC TGTCTGAA TGT TTTACATTTCTCCCACTGCC TCAAGAAGGTTCACAACGAGGTCA
 CCA TGA TAAAGAGTAAAGCTTCCAGCCGGAGCTGCTCTGGGCCCCAGAGGACAC TCCACAGAGATATGCTAACTGGACT
 TGGAGACTGGCTCACACTCCAGAGAAAAGCATGGAGCACAGCGCACAGAGCACGGGCAAGGTCCACGGGACAGAA TGTCT
 AGGAGGGAGATGGGGTGGGGTAAATCTGATGCAATTA CTGTGCCAGCTCAACATTCAGGGAGGGGGAAGAAAGAACAG
 TCCCTG TCAAGTAAAG TGTGCAAGCAGAGATGGTAAAGCTCCAAAATTTGAAACTTTGGCTGCTGGAAGTTTATAGGGGCG
 AGATAAGAAGACATAAGAGACTTTGAGGGTTTACTACACACTAGACGCTCTATGCA TTTATTTATTTATTTCTCTTATTT
 ATTTACTTTGATAAATCTTATAATAATCTTATGAAAACGGAAACCTCATATACCAATTTTACAGATGAGAAAAGTGACAA
 TTTTGAGAGCATAGCTAAGAA TACCTAGTAAAGTAAAGGAGCTGGGACTTAAACCAAACCTATCTCACACAGATCACACT
 CTTTTTTTTTTTCCAGTGAATTTTTTTTAA TTTTATTTTACTTTAAAGTCTTGGGATACATG TGCAGAAGGTA TGGTTTG
 TTACATAGGTATATG TGTGCCATAGTGGATGCTGCACCTATCAACCCGTCACTAGGT TTAAGCCCACTATGATAGCT
 ATTTGCTCTGATGCTCTCCCTCCCTCCCAACAGAGGCTTTGGTGTGTA TGT TCCCTCCCTGCTG TCCATGTT
 CTCACCTTTGAGCTCCACTTATGAGTGAAGCGTGTGGTATTTGGTTTTCTG TGTGTAGTTTGTGAGGATGATG
 GCTTCCAGCTTCA TCCATGTCCTGCAAGGACAGATC

SEQ ID: 7

>STAR7

ATCATGCCAGCTTAGGCGACAGAGTGGACTGGACATAATAACAATAAATAAAAAATAAATAAACAATATCTGA
 GAGGAAAAATTTGATTCATAATAAAGAGAAATAAGGTTTTTGGCGTGTGTTTGTGTTTTCACCTAAGAACAGCTGTCCCT
 CTCATTTGGGTTAGTTTTATTTGCAAGCAGAAA TCACTCCGCA TGAATTTCCAGGGTGTGGAAGAACTGAAATATGAA TCCAC
 CTTCTGCCATCTATTCATTTGTCACATTTAA TAAAGACACTCATGCCATTTTGA TCA TTTCTCTCCCTACCAAAATGAGTT
 AGTAACTCAAGAGATTTAAAATAACAAAAATAAGAACATTAAGGATTTCAAATGTTTACATAAATAAATAAACAATAA
 TATTAATAATTTCTTGGAAATGACATTTGCCCTACTCTCAAGGTAAGGTCATTTTTCTTGA TTTAACTTTTTCTCAA

FIG. 20, cont

GTTGAAGTCTCTAAGTTCAACCCGTAATCTATTGCAAGTTGTGCAAAATTTAGGGATTGAATCCATAGTAATTAGTG
 ATTTATTTGGGTGAGGAGACAAGTCAAAAGAAATCAGGACTGCTAGGTAGATGACTAAGGAAAGGATGGTTCACGAGGGT
 ACATAAAGCACTCAGAAGAAAAGGTGAGAAACGGAGGACAGAAAAAACCTAAGTTCTGCTGGGCTGATGCTGAAATTTGT
 CATCACAAAATCTGCATTGTGGAAGCTTTAGCTATTGAGGAGATTGCTCAAGTGTAGAACTGAGAACAAATAGGCAGTGAAC
 CCGAGAGAACATCAAGAGACTGAGAGAAAATGAACCAGACTTCCAGGTGCTCCATGTTCCAAACCAACATTTTGTATTGCA
 GAAGGAATTGAGAGGCAAAAGGAAACCAATAAAAAATAAACAGGAAAGGGCATAACATGATTACCACCCCTTTTTCACC
 AGCTGCTCATGGACCAGCTTTCTCCTAGTGTCTTTCTTGGTCACTGATCACTCTGCTAACATAGTTTCCCACTAGCT
 CTGAGGCTGTCCAGAGGGGAAGCCAGCTGTCTCTCTTCCACACTCTGTTGGAGGAACCTGCTATTAGCAGCTCC
 TACTAAACGCATTTATGCAAAACAGGAGGAGATAATTAACAGAAAGTGAACAACTCAAACTTCAGAGCTCTCATTG
 TATGAATGCCCTTGAAGGCTTGGGCTATTTAATATTTATAAATGTGTTATTTCTTCTAAAGAAAACCCCAAATG
 TATAAGCTACAGAACTGCAAAACTGAGGTCATCCATGCACTCAGGATACATTCATAGCATCTCTGAGCTGAAAAATATC
 TTAAGGTCATATATGCTCCCAACACTGCAAGAATCTCTGCGCAGCATCTTTAAATCATCATCTAAAGAGGGGAAA
 TCCCAGCTGTGTTGGATTTTGTCTGTGCACTTTCCAGTTTCCCATCAAAAGGGCAACATATGAATTTCTGTAT
 AAGGTAGTTGTTAATAAATACAAGTGGTAGCCACTTCCCTAAGAAAATATGGGGTTTCTGCTTACAGTCTAGGGA
 GAGGAAAAAAGGGGGGTCAGAAGTGATTATATATCATTTATATTTGAAATGTTTTCAGACATAAAAAAGCTCACCAG
 TCTTAGGCCAGACAGATGCATTATGAAAGTTAAGCTAAGTCTTCTCATCATGAGCTGCACCTATATCCCATTACTCTT
 CTAGAACTGCATAATTTATTTATCTTTCTTCAAAAGTTGAGAGAGCCATTCTTGTCTTAAAGATTTTTTTTTTTTT
 TTGGAGACAGAGTCTCCGTCTGTTGCCAGGCTGGAGTGAATGGCACTATCTCAGCTCACTGCAACCTCTGCTCCAGA
 TTCAGTATTTCTCTGCCCTCAGCCTCCCGAGTAGCTGGGATTACAGCACGCCACCAACCAACAGCTAATTTTTGCTAT
 TTTTATGAGAGCAGGTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCTGACCTCGGGTGTATCCACCACCT

SEQ ID: 8

>STAR8

GAGATCACCTCGAAGAGAGTCTAACGTCCTAGGAACGCTCTCGGGTTCAACAAGGATTGACCGAACCCAGGATACGTCGC
 TCTCCATCTGAGGCTTGCTCAAAATGGCCCTCCACTATTCCAGGCACGTGGGTGTCTCCCTAACTCTCCCTGCTCTCCTG
 AGCCATGCTGCTATCACCCATCGGTGCAAGGCTCTTCTGAAGAGCTCGGTGGATTCTCTCCATCCACCTCTCTTTCCC
 AAGAAGAGCCACCCTTCAAGACACCAATGGGACATTTCCCTTCCACCTCTTCTCCAAAGTTGCCAGGTTGTCATC
 ACAGTTTAGGGAGAGAGAGCCCAAGGTTTCAAGTTACAAGGCATAGGACCTGGCATGAACACACACACACACACACACA
 CACACACACACACACACACACTCGAAGAGGTAGCCCAAGGGTCAATTAACACTTGACGACTGTTTCCAAAACGTTGA
 TGCAGTTCATCCACGCCAAAGCCAAAGGTTGCAAAAGCAACCGAATGGTGGAGAGATTCCAGAGGCTCACCAACCTCT
 CAGGAATATTTCTGACCTGGGGCAGAGGTTGGAACATTGAGGACATTTCTTGGGACACAGGAGAGCTGACCCGAC
 CAGGCAATTTCTTCCACTGCAAAATGACCTATGGCCGGGCAATTTCACTTCCCTGCAAAACACCTATGGCAGGTACC
 TCCCAAGCCCCACCCCTTCCCGGAATCGGCATGGCTCGGCTCTATCCGGTGTGACTCCAGGTAGGCTTCTCAAC
 GCTCTGGCTCAAGAAAGGACAAATCACAGGTCCAAGCCCAAGCCACACCTCTCTCTTTGTTATACCCACAGAAGTTAG
 AGAAAACGCCACACTTTGAGACAAATTAAGATGCTTTAATTTAAGCCGGCGGCAAAAGAGATGGCTAACGCTCAAAAATCT
 CTGGGCCCCGAGGAGGGGCTTGAATACTTATACCTTGGTTAGGAAGGGGAGGGGAACTCAAATGGGTAACTTAC
 AGAATTAATAATGAGGAAATCAAAAGAAAGCAAAATGGTTATAGAGAGATAAACAAGTTTAAAAGGCAAAATGGTTACAAA
 GGCAACGTTACAGGTCGGGGCTTAAATCTTATGACACTTAGATATAGGTGCTATGCTGGACACGAACCTAAGGCTT
 TATGTTGTTATCTTTCAGAAAAATCTGGGAACCTCATGCACTGTTTGTGCGAGTATCTTATCAGTTGATTGGGCTCCC
 TTGAAATGCTGAGTATCTGCTTACACAGGTCAACTCTTGGGAAAGGGGTTGGTAAAGGAGCCCTTCTGTGCTGAAAT
 TAAGGGTTCGATTTGGAGTTTGTCCAGCACTCCAGCTACAGAGGCTTATTTACATGAGAAGCAAGGCTAGGTGATTA
 GAGCAACACAGGAGAGATTCAAAAGTACGACTTAGAGTAAAAACAAGGTTAGGCATTTCACTTCCAGAGAACCCGCAAA
 CATTCAATGGGAGAGGTTCCGAGTCTCAAAGTCCAGATGTGGCGAGCCCCGGGAGGAAAAACCGTGTCTTCTTAG
 GATGCCCGGAACAAGAGCTAGGCTTCCGAGCTAGGACGCTATATGCTGAGCCGGCGGAGGAGACCCCGGGAG
 GCGAAGTGGGGGGGGCCATCTTCTTCTGCTGCTGCGGGAGCTCTGCTGGCTGCGCTCAAGCCGAGGAGGGC
 GCGTCTGACGGGCGCTAGAGTTTGGGTGACAGT

SEQ ID: 9

>STAR9

ATGAGCCCCAAAAATGATCCTCTGGCTTATGACAACCTGATGACGCCAGGAAATGCTGCAACATGCCACTAGCAGT
 GGAACCCCTCTGAGGAGAGAACGTTTACATTAAGAAACCCCTTTGTTTGCAGCAGAGACTATTGAGTCAACATG
 TGTGGCTCTCAGTCTTTGAGCCATTTGAAGTTCTCTATCCTTGTGGAGGCTGAGCTCTCATGGAAACCTGGTCCGA
 TAGTGAGAGGAGCAGACCCTCTGGAACACCTTTTACACCTGACCAAAGCAGGCTCATGGGGCAGTGTGCAACAAGG
 TCAACCGGTGATTTGGCCCTCAGAAAAGCAGCCCGGGAGGTCAGGAGGAGGCTGCTGACTCCTCTTCCCTGCA
 GCCCCCAAGCACCCAGGAGCCCTGCAAGTTTGGGTTCCAGGTTGCCAGGTCACAGGTCACAGTGTCTTCTTAGC
 AGCTCCTGGAGGATGAGATGGTCTGGTCAGAGGCTGATCTGAGTATCAGGAGCAATGGGGCAACGTTTCTGCGATTG
 AGGAAAGGGGCAATTTCTGGGTGGGAGAACAAAGGCTTTGGCTGAGCTGGAGCATCCGCTCCATCAGTGTTTCCGGCA
 ACTGTACTATCCATGCTCTTCCCTTCCACAGCTGACCATGGCTTTGGAATAATGCTCTGAAACTTCTTTTCAAGAGAT
 GACTCCCACTCCACACTTAGGGGAAGTCAAGCTACTTCTCAGAAATCAGAGAAGGCATAAAAAAGAAATTCATTCTAAA
 GGCCCTTTAGAAGTAACTTAGGTTCTGACAGCGGCCAGCTAATTTCTGGTGCCTTCCAGGAATCTTCTGACTCAAAAA
 AAAGCAATTAACACTGAACACAAACCCAGTTACAGATAGAAAACATAGTCAATTTAAATAGAATATAAGCATCTGGCCCT
 TCCCATCAATTAAGGATAACACAAAAATCTATTTCAAAAGGAACTAAATATTTGACCAAAACATGAATGGGGAGAC
 CTCAGGTTGATACAGCTCTTGCCTGGATGGAAATTTGTAATCAAGAGGATGAGACAGGATTGTAATTTGCTGCAATGTGAAA
 GGGTTTGGTCAAGTATCAATTTGCTTAAATGCATGGTAATTTCCAAAGTTCTTTGGAGCTGAATTTCAAAATTTAG
 TGCAGGTCCTGGTGAAGCCACTTGAATCTTCAAGTCAACTGCAAGTGGCGTGGCTCAATGCTGGGCAAGAGAGCCA
 ATGTCACAGCCAGGAGTGGCTGGTCTTACCAGGCTCCAGGCTGCTTCAATGGTGGGCCCTGGGCTGGGAGAACG
 CACTTTTGGCTGATGAGTATCTGGTCAAACTCTCCTGTGGACACAGAAAGCCATGGCGACAGGCATTTCCAGGAAA

FIG. 20, cont

AGAAAAGGCCAGCAGCTGAAATCGTCAGGTGGAGAAGGCAGTCAATCCCTGCTCAGTCAACTCTAATCCGGCTGCCTCCCTCC
 TCAGCTTCAGGGTGAACCTCTCCTAAGCTGTGCTTTGGTATCTGATGGGCATTAGGTGCTGGTGA AAAAGCTGGAGGGTCT
 CTTTGGGATAATCAGAAAGCCAACTAGCCTTGTAATCAATACTAGGCACCTCACCCCTGAAGTCTACGTTTCCAGA
 TTCTGAAAACATGGGAAAGCAATGTGTGTGATGCTGAGGTCCOCCFCAGCCTCTGGTGTAGGGTTAGGAGGGCTCTAAAG
 GGTGCGACCTCCAGTCTCCAGTGGGGCTGAAAGTTGGTCCCTCCCTTCCAGCTCCCAATCAATGGTTAGCCAAATCC
 TTCCTACCTAAGAGTACTGCACATGGATGCTCCAGCGAGGCTCTGCTCCACTCCAGGAAGTG

SEQ ID:10

>STAR10

AGGTCAGGAGTTCAAGACCAGCTGGCCAACATGGTGA AACCTGTCCCTACAAAAATACAAAAATAGCCGGGGCGTGGT
 GGGGGGGCCCTATAAATCCAGCTACTCAGGATGCTGAGACAGGAGAATGTTTGAACCCGGGAGGTGGAGGTTGCAGTGAA
 CTGAGATCGGCCACTGCACCTCCAGCTGGTGCACAGAGAGACTCCGTC TCAACACAGACAAAACAAAACAAACAAAC
 AACAAAAATGTTTACGACAGCTTATTTGAGATAAAAATTCACATGCCATAAAGGTCACCTTCTACAGTATACAAATTCAGTG
 GATTTAGTATGTTTACAAAAGTTGTACGTTGTTCACCACTACTCCAGAACATTTACATCACCCCTAAAAGAAGCTTTTAG
 CAGTCACTTCTACTTCCAGCCCTGGCAACCAAGAACTACTNTCTGCTCTATTTCTGAAATTTTCAATAAAGGAG
 TCCATCATATGGGCCCTTTACGCTACTCTTTTCCACTTAGCATCATGTTTAAAGATTCATCCACAGTGTAGCACGCTG
 CAGTTAATTCATTTCACTTATGGCTGGATAAATGCTCTATTTGATGCTATGCCCTCACTTTGCTTATCCATTCATCAACTG
 ATGACATTTGGGTTAATTTCTACTTTTTGACTATTA TGAGTAAATGCTGCTATGAACATTCCTGTACCAATCGTTACGGTGA
 CATATGCTTTCAAATTCCTCGAGTATGTAAC TAGGGTTGGAGTTGCTGGGTCAATGTTAACTCAGTGTTCATTTTTTTG
 AAGAACTACCAAAATGGTTTCCAAAAGTGAATGCAACACTTTACATTCACCACGCAAGATATGAAGTTCCAAATGCTCTTA
 CATTTTGGCAACACTTTGATTTCTTTTATTTATTTAATTTAATTTAATTTATTTTGGAGTGGAGTCTCAC TCTGTCAACC
 AGGCTGGAGTGCAGTGGCAAAATTCAGCTCACTGCAATCTCCACCCTCTCGGGCTCAAGOGATACTCTGCCCTCAACCTCC
 CGAGTAACTGGGATACAGGGCCCCACCACCAAGCTAA PTTTTGTTATTTTAGTAGAGAGGGGTTTCAATCATGT
 CGGCAGGNTGTACTCGAATCTGACCTCAAGT

SEQ ID:11

>STAR11

GATTCGGGTGGGTTTGAATGATGAGAGTCCCTTGAATAAAAAGAAATTCAGAAAAGCTGTGAAACTCACCTTTCCOCT
 ATTCTTAACCTTACTTGCCTTTGGGAGGCTGAGGCAGGAGGATGACTTAAGGCCAGGAGTTTGAGAATGATGAGTGTATG
 ACCACACCGGTTACACTCAAGCCTGGGGGAGACCACAACAAAACCTTACCTGCCAATGCTCCATGCTGGAAATTTATTT
 CGTTTCTGGGATTTGGAAGAACTGGCTTACTGAAAACCACTTCTCTAAAACCCCTTCTCCAGTTAGGTGTTAAGATTT
 TTAACAGCCTTTCTATCTGAAATAAAAAGTGCACACAAAGTAAACTTAAAGATGTCACAACTCACTCTGTTTGTACAAG
 ATGAGTCTCAACTGCTCACTGCCCTGGGGAAATCCTCACTAGCGTCTAGTGGCAAGACTCCCTGTGTGCTCACCGAAACGC
 TCCOCTTCTCCAGGCCACACAGTCACTGGAATTTCCATGCAACCCTGGCAGCTCAGCAGGAGTCCATGACTTAAAGAGCC
 CAATGGACTGTGGGTGAAGTCTGTGGACGGGAAGCCACATGCGTCACTTCCAGGCCCTGGGGCTGTGCACTCTCCACTCTC
 TTCOCTGTGGGTGCAGAAGGCCGGGGCAGAGGGCCCTGAAACCTTTGGAGGTGGTGGAGCCCAAAAATGAAGGAGCGTGGC
 CTTCTGGGTCTTCAATGTAATTTAGGTAACACTGAACTGTTCAGGTGAACAAGAAAATAAAGCTCAAAATGATTCAGCTGATTA
 ATTTGGTGAATGGTTTACAGCGGTTACCTCCCTCAACATAATAAATTTTCAACAACTCAATAAATGGCTCACTCAATGA
 TAAAAATTTCCATAATGAAATCCCGGGAATAACTGCTTATTTCTAGCTCAAGCTTAAATCAGAGTGTGCTATGAGGGAGGAG
 ATAGTAGAGGGCAGCAAGGGGTTGCTACTGAAGATAAATAGCTTGTCTAAAAGAAATGGTTGAAGAGTGAAGTACAGATAG
 GGTAAATCCACATCTCAGACATTTCTGTGATGGTCTGATATTAATCTTAAAGTAAAATGATAGAGTTGAACCAATTTTAAATG
 ATTCFAGAAATTTAATTAATTAAGATGGGCAATTTCCACAAAGGACTTAAACAAAGTACAGAGGATTAATAATATCAATCCA
 CATGGGAGGCCACCGCTTGCACTTAAAAATGATGGAGCTTATCAAGACTGGCTGTGGATACTGTCTCCCTGGGAGGGTTTTT
 TCCOCCATTTTTTCTTTTTTGGAGACATGTTCTGCTATGTTGCCAGGC TGGTCTTGAACCTCTGGGCTCAAGTATCTCT

SEQ ID:12

>STAR12

ATCTTGCTTCTGGGAAGAGATGGCCCTCCCTGTGTCAGTGTACTTTGGCAGGACCAGCAGAAAACCCAGGTTTCCGTTCAGG
 AGGAAGTGTCTCAGCTTATCTCTGTGAAGGGTCTGTA TAAGGCACAGGAGGAGCCAGGGCTTCCAGGATGTTGCCCTTCTGT
 GCCATA TGGGACATCTCAGCTTACGTTGTTAAGAAAATTTGGCAAGAAGATGCACACAGAAATTTCTGTAAAGCAATAGGAT
 GGAGTTTAAAGGTTACTACGAAAAGAAAAGTACTGGAGAGAGGGGAGCCAAACCAACCAAGTTTGAATTCATTT
 TATTTGGAGCAATGCTCACTTTAAATTTAAATGGAGTCCAACTTCCCTTTCTCACCCAGACGTCGAGAAGGTGGCATTCAA
 AATGTTTACACTGTTTCACTGCTTTTTTGTCTAAGTCTGTTCCOCTTACTCTCTTCCCTCACTTCACATTTGCTGTTT
 ATCCACACATATGCTCACTTTATATTTACATATAATAAATTTTATATAATGGCTTGTGAAATA TGCCAGACGAGGGAATG
 AAATAGTCTCTGAAAACAGCTGGAAAATTTGCAACAGTGGGAGATTTGGCCACATGTACATTTCTACTGCAAAAGTTGCAC
 AACAGACCAAGTTTGTATAAGTGAAGCTGGGTGGTTTTTATTTTCTCTAGGCAACAGCTTGCCTGGTGGAGTAGGCC
 TCCCTGAGAAGGCATTTCTTAGGAGCTCAACTTCCCAAGAAGAGGAGAGGGGAGACTGAGTGTGCTGCGCAGCACA
 GAGACAAGGGGCCACGGCAGGACTGCAGCTGCAGAGGGGGTGGAGAAGCGGAGGCTGGCACCCAGTGGCCAGCGAGGCC
 AGGTCAGTCCAGCTCGAGTCTAGAGTACAGCAAGGCCAAGGTCAGGTCAGTGTGAGTCTAAGGTCATGCTGCTGAG
 TGAGCTGAGACCAGGGTCCAAATGAGGCCAAGGTCAGAGTCCAGTAAAGGCCGAGATCCAGGGTCCAGGGAGGTCAG

SEQ ID:13

>STAR13

CTGCCCTGATCCCTTAATGCTTTTGGCCACAGACCCCGCTAAGTCCAAACCCAGAGGGGCTTCAATCCGCAAAAGCCTCGG
 GAAGGACAGTGCAGGAGGGCGGCTGCCCTGTGAGCTGCACGGGGCAGAAATGCTTTTGGCGTCAATGTTGATGTCCACA
 CATCCATA TGGGGTCAATGCTTATTAGGATTCCTTGGGAAGAGGTAGAGGGTAGGAGGGTTAAGCCACGAGACGAGGCAAT
 GCAGAGGGGTGGCTGGATGGGTTGCACTGCTGTCCA TGCACACGGGGAGCGTTGCAAAATGTGCTTCCAGCCCATAGT

FIG. 20, cont

CGCCCCACAGAGGAGCCCGGAGTCCCTGGTGGGGCTGTGTTCCTGCAAGGAGCCAGTGGAGATGGCCCGTGAACCTC
 CATCCCCCTTGGCTGGTGGGGTCTCTGGCAGGTTTATGGAGCCGTACATCTTTGGGAGCCCGCTGGACACAGACATCATC
 GACCTGGAAACAGACAGCCAGCCACCTCCAGCTGGCCTTGAACCTTACCAGCCACATGGCCACCGCAAGGGCATCATCTTG
 TTTAAGCCGCAACCCGAGTCTCTGTACCTGATTTGAGAAATGGCCCGTGACTGTGGGGAGTACGCCACACATCCGCTAC
 TTCAGGGCCGCACTGTGACCAADCGCCGCCCTCTTTGGCCCAAGGTCGCCCTGCCGGACCTCATCATCTCCGTCAC
 ACCCTCAACAATCTTTGAGCCACAGTGGCCGTGAGAGAGCCAGCCAAAGTAAACATCCCCACAGTGGGCATGGTGGAC
 ACCAATGCAACCCCTGCCTCATACCTACCCCTGTACCCGGCAATGAGGACTCTCCCTGGCTGTGCACCTCTACTCCAGS
 CTCCTTCAGACGGCCATCACCCGGCCAAAGGAGAGCGCCAGCAGGTTGAGGCTCTCTATCGCCCTGACGGCCAGAAAGGAG
 CCCCAGGACCCAGGGCCAGCCACCCCTCTGGGGCTGACATGAGCCATTCCTGTGATGTTCACTCTCCCAAAGCAA
 CCACAGCAAGCCCTGTCTGAGCTGGGATCCCTCTCCAGCCCTGGGTCAGCGGCACTCCAGTCTGTGTACTTACTCA
 GCTGATGTCAGTGCAGACATCCACCGTTCACCAACAGAACAGTGGCTGAGCGGACCAACGTTGCCATGTCGGTTTGTCT
 CTGTGGGAAACAGAGCACAGAGGTTAGCCACATGTCAGAACGCCCTTGGCTGCAGTTAGGACTCAGTGGCT

SEQ ID:14

>STAR14
 AGCAAGGACAGGGCTGTGCCTCCAGTACGATGAGCAGAGCAGACTCCTTTGAGCAGAGCATCAGGGCAGAAATAGAA
 CAGTTTCTCAATGAGAAAAGACAGCATGAGACCCAAATGTAATGGTCACTGGAGAAAGAACAGACACATGAAAT
 TGGCGAAGTCACTCTCGAAATCCACCAAGAGCCGCTACAAAGTGTGTGCACCCGACAGGCTGTATGGGGCTCCAGAA
 GAGTTCCCTCTGTGAGACCTCCCCCGTTAGCAAGACAAACGTGCAGCCAGAAAGCCCTCAGGTCAGAGTCAAGCCAC
 GACCAGCAGGAGAGGAGGGCAGCACAAGCCAGCAACCCCAACCCCTTCAGAAAGCAGTACAGAAATAAAGTGGGAT
 AAAAGGAAAGCCAGCAGCCAGGAGGGGAAAGCGAGTACAGAGCCGCTACAGGCGCCGAGGCTCCGACTCCAGCAGC
 GACGACGGCATTTGAGGAGGCCATCCAGCTGTACAGGTCAGAAACACACAAGGAGCCGAGCGGGACCCGCCACAGG
 CTCAGCTCCAAAGGAAAGAGCACCTGCCCCCTCCGACACAGCAAGCAGCGCCACAAAAGTGGCTTGCAGAGACC
 CACAGGAAACACCCAGCAAGAAGAGCCAGTCCCCACCAAGCAAGGACCCGACCCGCTGGTCCAGG

SEQ ID:15

>STAR15
 CAGTACATGCGAAGTGAAGTCCAAACGAGACGGCAGCAAAACCCGAGTGGGCTCCAGACATCTCTGGGGAAAGGGAT
 CCTAACCCAGCCAGTAAAGTCACTCCCAACCCCTCTATGACACAGGCTGTGGCTGTCATTTAAAGCTGAGTGA
 TTTAACCCCTTTCCCATTTAGAAAAAAGAGCCAGCTGGCTGCCAGCACTATTTAATTTTACATAAAGTGTCTTTGA
 GCTTGAAGCAAACTGTGACTGATTTTCAAATGTAATAAATAAATGTAATAAATGTAATAAATGTAATAAATGTAATAA
 AGAATTCCTGTAATCATCAGAAATGGCTATTTTGAATAAATGCGGATTCATAAAGCAATCTCCGCAACAACTGTAGAG
 AACGATGTTAACACCAAGCAGCAAGGAAATGTTACATTTTCAAGAAATTTGACATTTTCAATGACGAAATATACGTA
 ATATGAAAAATACCACACTAAACACATAATGCTATAAATGAAATGATGCTTTTGTTCACAAAGTCAATATACTCGAGCA
 TGCAAAAAATAAATAAAGTCAGATACTTTCAATGCAAAAGCTGCCGACAGATAAACAATTCAGCCACAAGTGGCCAGTA
 TTCCTGGGGCAAACTGGAAAAGGGCTAACAGGCAACATTTTCAATGTTATTCCTACTGAGTGCAGTAAATTTTAAAAATA
 TACAATAAATAAATAAAGTGTGGTATGGTTTTTAAAGAAATTTCCATAACCTGGTGAACCTCTTCCACAGGGTAAATAG
 TTCATAAAGCCCTGGTCCCTGCAAAAACAGCAATCAACTTGACAAATGACTAAAAGAAGCAACAGCAAAACTGTCCACCA
 TTTGGAGCCAATGGCTGGGTTGGCCGGTGTAAAGCTCTCCGCTCTGGAGCAAGTCTGGGCCCCAGCGGCTGGCATGGG
 CACTGCAAGGCTGGGTTGGCCAGGTTGCAAGCTCTCCGCTCATCTGAGCTAGTCTGAGGCCCTGGTGGCTGGCAGTGGG
 CCTGCAAGGCTCTACTTCTCACCCAGCTCCACTTCCCTCCCTGCCCTCACAGGCTCTCACAGGCAATGAACACTGGG
 GTACAGATTCAGGGCCAGCATCCACTGCAAGTGGGCACTGCCCTTCCAAAGGCTGGCTCCAGGAAGCAACCCCACTCA
 GCCACAGTATGGGCAACAGGAAATCCCAATCCOCCAATGCCAGTACTACAGGGAAGGGGCTCACGTGAGGCTGGCC
 CAGGCTGTCTGAGACCCGCTGTCTATGAGCTTGGATTTAAGGAATCTGGGAGCAAGAGCTTTCTTTCAATACGGGCC
 ACCAGCAGGCAAAAGTTAGCCCAACGCAAGTGTGACAGTCAACCCCAAGGACCCAGGCAAGGAGGAGGAGAG
 GACAACAGAGGATGAGGTTGGGCAAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
 CAAGAAAAGTAGGCGACTAGACTTAAAAATCAGAACTACCCCTCAATCAAAAAGATAGTGAACAATTTGGGGTCTATAAT
 TTAACAATGTCCTCCAAAGGCAATGTTGGAAATTTAATCCCAACAAACAGGCTGGGAGGTTGGAGCTCATGAGAGGT
 GGTGAGCCCATGAGGTTGGAGTGAATGGAATGAATGCCAATGCTGCGGAAATGGGCTCTCTTACAAGGAGGATTCAGCC
 CCCCCTCTCTGCTCACCTCTCTTTGCCCTTTGCCTAGGAGTACGTAACAAGAGGGCCCTCAACAAGATGCTGGCACCT
 TGATCTGGACTC

SEQ ID:16

>STAR16
 CGCCCACTCGGCTTTCCAAAGTCTGGGATACAGGCAATGAGTCACTGCGCCCATCTGATTCCAAAGTCTTTAGATAATA
 ACTTAACTTTTCGACCAATGCGCAATCAGGCAATCTTTGAATCTGCCATGACCTAGGACATCCCTCTCCCTACAAGTGT
 CCCCCTGTTCCAGACCAAAACCAATGTACACTTACATGTAATGATGAAAGTTTACATCTCCCTAAAACATATAAAACCA
 AGCTATAGTCTGACCCACTCAGGCAAGTGTCTCAGGACCTCCCTGGGCTATGCGCATGGGCTCTGCTCAGATTTGGC
 TCAGAAATAATCTCTCAAAATTTCTCAGAAATTTACTCTTTTCAATACCAATTAACCTATCACCCATAAGTCAGAGTTTC
 CACAACCCCTTCCCTCAGATTCAGTAAATTTGCTAGAAATGGCCACCAAACTCAGGAAAGTATTTACTTACAAATTAACAA
 ATTTATGAAAGAACTCAAAATCAGGAAATAGCCAAATGAAAGAGGCAATGGGAAAGGATGAGGAAAGGGCCACAAGCTTCCAT
 GOCCTGTGTGCACACCCCTCTCAGCATCTCAATGTTTCACTCAACTCAGAAAGCTTTCAAACTTTGTCAATTTAGGGGTT
 TTTATGGCAGTTCACATAATGAGGCAATGGTTGATAAATCACCTGGTCACTGGTGTATAGAACTCTGTCGCCAGCTCTCTC
 TCTCTCCCCAGAGTCCGAGGTTGGGCTGAAAGTTTCAAGAGTTAGTTGCTCTGCAACAGCCCTATCCGTAAGGT
 ATTTAGGGGTTCCCAAAGTTACTTGAATGTTGGAAAGGCTTATTAAGAAATAAAGAAAGATGCTCCATTTTACTTACC
 ACTAGGAGCATAATCAAGTCTTGGGGAACAAAGCATTTACTGAGTAAATTAATACAGGATAGATAGCAATCTCAAT
 CTGGCTTCTCAGAAAGAAATTTGACCAAGGGGCAATAGCCAGAGTGGGACCAAGATAAGTTTATAGCAGGAGT

FIG. 20, cont

AAAGTTTATTA AAAAGTTT AGGCAGGAATGAAAGAAAGTAAAGTACATTTGGAAGAGGGCCAAAGTGGCCGACATGAGAGA
 GTCAAACACCATGCCCTGTTGATGTTGGCTTGGGGTCTTATATGATGACATGCTTCTGAGGGTTGCATCCTTCTCCCT
 GATTTCTCCCTGGGGTGGCTGTCCGCAATGCAAAATGGCTGCCAGCAGTAGGGAGGGGCCGATG

SEQ ID:17

>STAR17

ATCCGAGGGGAGGAGGAGAAGGAGGAGGAGCAGGGCCGGAGCCGAGGCTGTCGGAGAAGCTGTTTAAAATGGTTGG
 CTTGAAAAATGCTACTAGTCTAAGTGGCTTTTCGGATGCTTATTTATTACTTTGTCAGGTTTCCCTAAGGAGAGGGGT
 GTTGGGGTGGGGAGGAGGTGGACTGGGGAAACCTCGCGTTCTCCCTCCGCTGCACAGGGTGAATAGGAAACGGCT
 CGCTGCCACTTAACAATCCCTCTATAGTAAATCTACGGGAGACTCTATGGGAAGCCGAGAACCCAGTGTCTTCTCCAGG
 GCAGAAGTCACCTGTGGGAACGCCCCCGGGTCCCTCCGCTGGCTTTCCGGCTCTTCAGCGCCCTGAATTTCTCCCA
 GCCCTCCACCCAGCTCCCTCAGGGACTTTTACACCTCCCTCCACCCCAATTTCCACTACAGTCTCCAGGGCACAGCACT
 CATTTGACAGCCACAGAGCTTCTCGTCT
 TGTCATAATTTCTTGGTGTCTTCCGCACTTAAACA AAAAGAGAAAAAATAAAAATAAAAAAACCCATTTGAGCCAA
 AGTATTTAAGATGAATCCAAGAAAGCAGCCACATAGCCCTCCACCCAGGAGTGCACCAAGACGCACCCAGGCTCCA
 TCACAGGCCGAGAGCAGGCCACTCTGGTCTGACTTTTGGGTCAAGAGATCTTGC AAAAGAGG

SEQ ID:18

>STAR18

ATCTTTTGTCTCTAAAATGATGATGGGTGTGTTTTTTTCCACCCTGCTAATAAAATATTACATGCAACATTTCTTCC
 CTCACCTCAAACCTGCTGAACCTGAAACAAATATGCAATAAAAGAAAAATCTTTGACAGAAAAAAAAGCTATTTCTCCAC
 TGAATTTGAAATGGCACTGCGGATGCAATTCGCAAAATCTTATTTGCCATTTCCCTCAATGAACTATGAAATGAAACCTTTG
 GACAGTCTGCCCATTTGCCATGAGACTCCCTCCCAAGCCAGGGTATGTTCCCAAGCACCCAGTGTAAAATCTTAAC
 TTAATTTCCCTTAAAATTTCAAATGTAACAACTGGGGCATAAAGAGTTCTGAAACAAACATGTCATCTTTGTTGAAAG
 GTGTTTTCTGTAAATTAATGATGGAATCATGCTCAATTTCAAATGGAGGCTCCAGATTTGTTGGCCAGCTGATCCCTGCAAA
 TATCTT

SEQ ID:19

>STAR19

TCACTTCTGTATATTTTACATTCAGGGTAGCTTATGCAATGCAACCTGTGCAGTTGCACAGGGCTTTGTTGTCAGAAA
 GACTAGCTCTTGGTTAAATCTCTGTTGTTGCCATCTTGAGATTCAATATAATAAATTTTGAATTTGTTGTTTGAACGT
 GATGTCCAATGGGACAAATGGAACATTCACATAACAGAGGAGACAGGTCAGGTGGCCAGCCCAAATTTCTTCCACCCCTTTT
 ACAATACAGCATTTGGCAATGCCCATGAGCACAAAAATTTGGGGAAACATGATGCTAAGACTCAAAGCACATAAACAATG
 TACTCTGTGACTAAAAGAAATGGAGGTGCTGACAGCCCCAGAGGCCACAGTTATGTTCAAACCAAACCTTGTGAGGG
 TGCAGAAAGAGGCAATGGCAGGGTCTAAGAAACAGCCCATCAATCTCTGTTTATCAATGTTACGTCCTGCAATGAACATA
 ATCACTTACACTGAAAAATTTGACAGAGGGAGAAAAATGAAAGATAGGGCAACCCATAGTTCTTTTCCCTTTTAGTCTTTCC
 TATACACTAAACCAAAGATAGTATTTGGTAAAAATGTTGTTGAGTTAAATTAATGAGTTAGTTTGGCAGTCTTTCCACTGTT
 GGGGTAAGAAACAAAATAATAGGCTGTATTGAGCTATTAATGTAATTTGGAATGTCAGTGAATCCAAAGTATGAATTA
 AATATCTTTGATTTGCAATTAATAATTTGGCACTGAACAAACAAAGATTAAACAGTAAAAATTAATGATGAAAGTTAAATTT
 TACTTAGAATGACATTAATATAGCAAAATAAAGCACCAATGATAAATCAAGAGAGAGACTGTGGAAGAAGGAAAAACGTTTT
 ATTTTAGTATATTTAAATGGACTTTCTTCTCTGATGTTTGTGTTTTGTTTTGAGAGAGAGGGATGTGGGGCAGGGAGCTCT
 ATTTTGTGTCAGGGCTGGACTTGAACCTCTGGGCTCCAGCTATCTGCTTAGCTTCTTGAGTAGCTGGGACTACAGGCA
 CACACCACAGTCTGACATTTCTGGATTTTTTTTTTTTTTTTTTTTTTTGAGACAGGTTCTGGCTCTGTTACTCA
 GGTGTCAGTGCATGGCAATGATAGCGGCTCACTGACGCCCAACCTCCTCAGCTAAGCTACTCTCCACTTCAGCCCTCT
 GAGTAGCCAGGACTACAGTTGTGTGCCACCACACTGTGGCTAAATTTGTTAGAGATGGGGTCTTCCACGTTGCCGAGGC
 TGGTCTCCACTCTCTGCTCAAGCGAACCTCTCTGACTTGGCTCCCGAAGTCTGGGATACAGGCTTGAGCCACTGCAAT
 CCAGGCTGTCCTCTGTTTAAACCTACTCCAAATTTGTCCTTCACTCTACATAAAGGGCTCTTTTCAAAGTTCCCAATAGAC
 CTCACCTGTGCTAAATCAATAATAAATAATCTGCTTTCTTACATGGTTCATCAGTAGCAGCAATGATTTGGGCTGCTCA
 ATCTTCTTGGTATAATTTCTTCAATTTGGCTCTTGGGCTACACACTCTCTTTGAGTTACTCATCTCCATGATGACTTCT
 TTCTAGTCTTCTTTACTGGTTCTTCT
 CTCTCTCTTATCTATTTACACCAAATCTTTCAGAGTCTCAATCCAGAGTCAATGAACCTTAAACCTGTTCTGTCAGATAA
 TTCACATTTATATATCTCCAGCCAGACTCTCCCCCAAATGACAGACTGATCTACTG

SEQ ID:20

>STAR20

GACTCAAGTTTCAATATCAATGTTTGGCAAAACATTCGATGCTCCACATCTTACCTAAAGCTACCAGAAAGGCTTTGG
 GAAGTGTCAACAGAGCTACAGAAAATCAGTAAAGACCAATGGACCCCTCAAACAAAACAGCCAAAGCTTTCTGCCCCAAA
 AGATGACTGAGAAGACTGTTAAAGCAAAAACCTCTGTTCCCTGCCATGATGATGCTATCCAGAAATAGAAAAATTAATTC
 CCTCAATCCCTAGGCTTCGAGAGTTTGGACCTGCCGGAAGACACAGATTGCACATCTCCCTTTGAGTGAAGTGCTC
 TCAATGATCTGAATGAGGAGAGAGACTTGAAGAGCTTTTCAAGCTGGCCCTCCCTTCACTTTGAAAGATGCCCTTCCAC
 CATGGAAAATCCAACTGTTGCACTCTCTTTAAAGCAATCTGTTGACCTGGATGTTGAAATGGCCACTGTTGCTCTGACA
 TAGATATTTAAATCTTTAGTCTTTAGAGTTTGTGTAATTTCTATAAATAAGCAATTTTGTTTTAAACAGAAAAAAGA
 TATATACTTAAATCTTAAATAAATAAATAAATTAATAAGGAAAAACAGGAGTTAATCAATTAAGGGAACAAAGGACATAAA
 ATGGGATAATAATGCTTAAATCCAAAATAAAGCAGAAAAATGAAGAAAAATGAAGTAAGAAACAGATAAATAAGAAAACAAATA
 GCAATAATGAAGACAAAATTCACCGGGTGTGGTGGCTGATGCCGTGTAATCC

SEQ ID:21

FIG. 20, cont

>STAR21
 GATCAATAAATTTGTAATAGTCAGTGAATACAAAGGGGTATACTATAATGCTACAGAAAATCCATTCTGGGTATAAATCC
 TAGACATATTTATGCATATGTACACCAAGATATACTGCAAGAAATGTTACAGCAAAATCTTTGTAGTACAAAAGGCCA
 AAAGGTCATCAACAAGAAAATTAATACATTTGGGCACATAATGGCATCCTTATGCCAATAAAAATGGATGAAAATATAGT
 TAGGTTCAAAGGCCAAGCCTCCAGATAAATTTATATCATATAAATCCATGTACAACAATCAACAACAAGCAAAACTAAACAT
 ATACAAATGTCAGGGAAAATGATGAACAAGTTAGAAAATGATTAATAAAAAATCTGCACAGTGAATAACATTTAATGAG
 AAAAAAGAAAGGAGGCTTAGGGAGGCACTACAGGGAATCCAAAGTTCAATGTAAGTACTAAATACATAATCAAGCA
 CTCAAAATAGAAAATTTTAGTAAATGTTTACCTAGTTAAATCTTACTTAAAACAAAGGTCAGGCCAGGCAGGTCGCT
 CACACCTGTAATCCAGCAGCTTTGGGAGGCTGAGGCGGGT

SEQ ID:22
 >STAR22
 CCGTTGATGCCACCGCCCTGGCCGCCAAAGTGC TGGGATTACAGGCGTGAATCACTACGCCCGGCCACCCCTCCCTGTA
 TATTAATTTCTAAGTACTATTAATGTTAAAAAAGTTTAAAAATATGATTTAATGAAATGCCAGAACTAGGATTTTACA
 TGTCACGTTTCTTATTAATAAAAATAAAAATCAACAATAAATATATGGTAAAAGTAAAAGAAAAACAAAAACAAAAGTG
 AAAAAATAAACAACACTCCCTGTCAAAAACAACAGTTGATGATAAACTTAAAGTCCCTGAAAATTTAGAAAACAATCTCTA
 AAGAAGTTCTGATAAAAATAGGAAATAAATAACACATAGTTTGGTCATTTGGTCTGTTATGATGATGGATATGTTTA
 TTGATTTGATGATGTTGAACCTTATCTCAA TAGATGCAGACAAGGCC TTGATAAAGTTTTAACACCTTTTCAATGTTGAAA
 ACTCTCAATAGACTAGGATTTGATGAAACATACTCAAAAATA TAGAAGCTATTTATGATAAACCCATAGCCAAATACATA
 CTGAGTGGCCAAAAGCTGGAAGCATTTCCCTTTGAAAAC TGGCACAAGACAAGGATGCCCCCTCTCCACCCTCTTATAAT
 GTAGTATTTGGAAGTTCTGGCCAGAGCAA TCAAGCAGGAGAAAGAAAGGATTTAAAAATAGGAAGAGAGGAAGTCAAATGT
 CTCGTTTTCAGTAAACAATGATTTGATATA TTTAGAAAACCCCATGTCCTCACTTAAAAACTCTTAAGCTGATAAACAAT
 TCAGCAAAAGTCTCAGGATACAAAATCAA TGTGCAAAAATCACAAAGCATTCCTATACCCGATAA TAGACAGCAGAGGCCA
 AATCATGAGTGAAGTCCCAATCAAAATGCTTCAAAGAAAATAAAA TACTTAGGAATACAAC TTTACCGGGACATGAAGCA
 CATTTTCAAGGACAATAAAAACCACTCC TCAAGGAAATGAGAGAGGACACAAGAAATGGAAAACATTTCCATGCTCATG
 GAAGAA TCAATATCA TGAATAATGGCCATACTGCCCAAAGTAAATTA TAGATTCATGCTAACCCCAAGCCCACTTGA
 CTTTCTCACAGAAC TAGAAAAAACTA TTTTAAACTCAATGATAGTCAAAAAGAGTCCGGTATAGCCAAGACAATCC TAA
 GCA TAAAGAAACAAAGCTGGATGCATCACGCTGAC TTTCAAACCA TACTACAAGGCTACAGTAAACAAAACAGCA TGGTACTG
 GTACAAAACAGATAGATAGACCGATAGAACAGAACAGAGGCC TCGGAAATAACACCACACATCTACAACCCCTTGTACTT
 CAA

SEQ ID:23
 >STAR23
 ATCCCCATCCCTTCAGGGCAGCTGAGCAGGGCC TCGAGCAGCTGGGGGAGCCCTCACTTAATGCTCCTGGGAGGGCAGCCA
 GGGAGCATGGGGTCTGCAGGCA TGGTCCAGGGTCTCGCAGGCGGCACGCCAOCATGTCAGCCGCCCCCACTGTTGCTCTG
 CCTCCGCCACTGGCCATGGCC TFCAGCAGCCAGCCACAAGTCTGCAGCTGCTGTACATGGCAAGAGCCCAAGCAG
 CTAGAGGACCTTGTGTTCCAGTGGCCAGGGAGCA TGGCCACAGCCCAAGACCCAGTTCAGGAGCAGCCAGGGCTTCTGG
 CAGGCCAGCTTACTCTCTGCTTCACACAGATGGGAGATTCTGTTG TGAATTTGAGTGTATGCCCTTTGGTGACATC
 CAAGATAGTTGCTGAAGCACCGCTTAA CAA TGTG TGTATCTGAAAACGAGAACTCTTTATTTGAAAATAATGATG
 CAATAAATTTAGTTTGGATTTGAAATTTCTATTCATGTAGGCA TGCACACAAAAGTCCAAACA TTGCA TATGACACAAAGAA
 AAGAAAAGCTTGATCTCTTAAATACAAATACTGTTAAGCTATATTTGCAAAATATA TTTGAAATACACTCTATA TGTZA
 CATATAATA TATATGTA TATG TATATA TAA TATACATA TATA TGT TACA TATAATA TACTTCTAT TATGTTACATATAA
 ATTTATCTATAAGTAAATACATAAATAAAGATTTGAGTACCTGTAGAACAATGCTTTATGTTATCAGCTACTACTAC
 AAAAATATCTCTCCACTTA TGCAGTTTGGCCATA TAAATA TGA TCTCTCATTTGATGGCCAGGGCAAGAGTGCAGTGGG
 TACTTA TCTCTGTGAGGAGGGAGGAGAAAAGGGAACAAGGAGAAAGTCAAAAAGGGAAC TCGGTTTGGCAAAAATG
 CAAGTTTACATA TCCGAGACGGAAA TGA CATG TCCACAGAAAGGCCCTGGCCAGCTAA TGTGTCACAGATATCTCAG
 GAAGCTTAAA TGA TTTT TAAAAGAAAAGAGATGGCA TGTCTACTTGTCTTGTAGCTGAGGCTGTTGGGATGATGACAGA
 TTTCTGGAAGGCAAGAGCTCTGCTTTTTCACACCGAGGACTTT CAGGAA TGAAGCCAGGGTCTGAGCAC TACAACA
 GGAATCCCTGGAGAGTGT TTTCTTACTTA

SEQ ID:24
 >STAR24
 ACGAGGTCACGAGTTGAGACACCGCTGGCCAAAGTGGTGAAGCC TGTCTCTACTAAAAATACAACAGTAGCCGGGGCC
 GGTGACGGGGCCCTGTAATCCAGCTACTCAGGAGGCTGAAGCAGGAGAACTCTAGAACCCAGGAGCGGAGTGCAGTG
 AGCTGACACTGCCCGCTGCAC TCTAGCCTGGGCAACACAGCAAGACTGTCTCAATAAATAAATAAATAAATAAATAA
 AATAAATAAATAAATAAATAAAGAGGAGAGT TGAAGTAGA TGAAGAGAGAAAAGAAAATCTAGATTTCTCATCTGAAG
 GCACCA TGAAGATGAAGCCACTCTTCTGGGCCAGGTCCTCCCGTTGCAAGTGAACCGAGTCTGGCCCTCAATGGAGAC
 CAAAGGAGATGACTTTGGCC TGGCTCCTAGTGAAGGAGCA TGCCTAGTCTGTTCTGTTGGCTTGTATCCTGATCACT
 TGATGTTCTCTCC TGGACTTTCCA TGGATTCAGGGATGCAACTGAGAAGTTTATTTTAAATGCACTTACTTGAAGTAA
 GTTATTTTAAAACATTTAGCAAAGGAAATGAATCTGCAGGTTTTCACATGAAGACATTCACATGTTGAGGAAAACAGGA
 AAACCACTATGCTAGAAAAGCAAAATGCTGTTGAGATTTGCTCACAAAACAAAATGGCTGCCAGCAGGTTAGGTTGAGCC
 TCAGGTTGGGCACATTTTACCTTAAAGCCACTGTTGGTGGAACTTAAAGGTGACTG TAGGACTTATA TATACATACATACAT
 ATAAATATA TACATATTTATGTTATATACACACACACACACACACACAGGGTCTTGCATCTTGGCCAGGGTG
 GTCTCCAACCTGGGTC TCAAGCGATCCTCTGCCTCCCTTCCCAAAG

SEQ ID:25
 >STAR25

FIG. 20, cont

ATAAAAAATAAAAAACCTGCTCAATTTGCAAAGGCCTATCTTTCTCCCAACCACCTGAAAATTTAGTGAACCGGG
 GCTTCCGTAGGAAGGAGTAGCTAGCTATCCCGGTCGCTACAGGTTATCAGTGCCTGAATACCTGACTCCTAAGGCTCA
 GGAATTGACTGGGTGCGCTCGTCGACTGCCCCGCCCCAAACCGGGACCCAGTCCAGCGGCCAGCTCGGGCGCTCT
 GACCTCGGGGATTTGAGCTTCGGTGCACAAACACCCACCGCGGCTGCGTCCACTTTACCTGCGCGGGCGGACAGC
 TTCGAAAGAAAGTGCACCAATGCTGTCGAGGAGCTTACCCCTCGAAAATGCTAGTTCGGGGTGGCACAGATCCGAGAC
 GACCCCTCATGCTTTTTTCTCACAGCCGCTGCTAGATTGGGCTACTTGTCTGGCCATGTTGAAGTTGAACCTCCAA
 ATCTAACTGGCCCGGCTCCCGGCTGCGGAGCTCCCGATGTCGCGCTCCCGGGAAGGTTGCTCCGATGGAAGCAGTA
 GAACGCTCTCACCGAGCGGGCGGGGGGGAAGTCACTCGAGGCTGAGGGCAGCGGGGAGCGAGGCTCTGCGCGGTG
 GGATGTCGCGGACCGGAAAAATACGGCAAGCCAAAGCTCGGGGCTCAAATAAAACTTTTAAATACATTTAGAGACTTC
 GTACAGTGCACAGTGAATATCACTGTTAAATTTCAAGAGTCCATTTCAFAAACGTTACAGAGTCTGCTTTTTAT
 TCCCTGTCTCCAGTCTCAATCAGTTCACAGTCTCCAGAGTTCCTTTTATGTTTGAATACCGACCAAACTCCAG
 TTTAGGGAGATGGAAGTCCACCGTCCATCCCAACCAAAACATTTTCACTCAAAACCAATCCAGTCTTAAAGAAATA
 GGAAAGTATGGGCAAGGGTCTTTTAAATATACACACATCACCTTAAACTGCGTGTGTGTACGAGAAAATAAGAAAA
 CACAAGAGGGCTG

SEQ ID:26

>STAR26

CCCCCTGACAAGCCCAAGTGTGATGTTCCCACTCTGTGTCACATCTCAATTTGTTCAACTCCCACTCTGTAGTGTAG
 AACATGCAGTGTGGTGTTCCTGCTCTGAGATAGTTGCTGAGAAATGATGGTTCCAGCTTCACTCAATGCTTGCAGAA
 GAAGTGAACCTATCTTTTATGCTTCAATAGTATTTCCAAGGACATAATGTCACACTTTTTTAACTCAGTCACTCAT
 GATGGCAATTTGGTGGTTCAGTCTTTGCTATGTGAAATAGCACCAAAATTAACATATGTGTGATGATATACATCTTT
 ATAGTACATGATTTATAAATCTTCGGTATATACCTGTAATGGGATCGCTGGGTCAAATGTTATTTCTAGTTCTAGATC
 CTGAGGAAATCACCACTGCTTTCCCAATGTTGAACTAATTTACGCTCCCAACAGCAGTGTAAAAGCAATTCCTATTTT
 TCCACCTCTCTCCAGTATCTGTGTTTCTGACTTTTTAAATGATCAATCACT

SEQ ID:27

>STAR27

CTGGCCCTCACAAAGCCCTGCGCCAGGGAACAATAGCGAGCTGCTTATTTTGTCTTGTATCCCAATGCTGGGCAATA
 GCTTGCATATGAGTAAAGCCGTTAGAGTATGTTTCAAGGACCAAGTGTATAAATACCAAGAAATCCAGAGAAAGGA
 GAGAACAATGAGTATGAGGATAGTGCACAGAGAGATGGGAATCTTGACAAAGGTTGTGAAGATGTACTAGCCAGGGGGAAC
 AGCTAAAGGAGTACACAGGACCGAGCTCTTGTCAAGCCGGCTGCCATGGAGGCTGGGTGGGCGCAATGCTAGCTTTCC
 TTTCTCTCAGGTTACAGAGTGCAGCCTTGAACCTCTAATCCAGAGGCAATTAATCAATGTTTTCTTAGGGCAATAC
 CTGCCCCTGCTGGAAGACTTTCTCCCTGTGGCTCGCCAGTCCCAAGATGAGACGGTTTGGGTACAGGGCCAGGTGCAC
 CGTTGGGTGTGCTTATGCTGATGACAGTATGTTACTCAGTCAATAGTCAATGAGGGAGGTGTGTTAAAGATGGAGATG
 CTGGTCACTATCCCTAGAGAGGTGTTCCAGTATGGGCAATGGGAGGCTGGAAGGATAGGTTACTGCTAGAGCTAGAGAA
 GCCACATCTTAAACCCCTGCTTTTCCCACTGCCAAGATCCAGAAAGTCTCTGTGGTTTCGCTGCTTTCTCTTTTTTT
 TTTTTTTTTTTTTCTGAGATGGAGTCTGGCTCTGTCGCCAGGCTGAGTCCAGTGGCAGGATTTGGGCTCACATGCAAGT
 TCCGCTCTCAGGTTCAATCAATCTCCCACTCAGCTTCCCGAGTAGCTGGACTACAGGCGCCACCAACCCAGCTAAT
 TTTTGTATTTTAGTAGAGAGCGGCTTCAACCAATGTTAGCCAGGATGGTCTTGAATCGGCTGCTTACGCTCCCAAGGTG
 CTGGATACAGGCGTGAACCCCGCCCGGCTGCTTTCTTCTTCAATGAAGCAATTCAGCTGGTGAAGAAAGCTCAGCA
 GGCCTGGTGGAACTCTTGACCTCAAGTGAATGCTGCTGCTCAGCTCCCAAGTGTGAGATACAGGCAATGAGCCAGTC
 CGAATGTGGCTTTTTTTGTTTTTTGTTTTGAAACAAAGTCTCAGTGTGCCCCAGGCTGCAATGCAATGCAATACCTCAGCTCC
 ACTCAGGCTCGAACCCTGGGCTCAAGCAATCCCTCCCACTGAGCTCCCCAGTAGCTGGGCTTACAAGCGCAATGCCAC
 ACGCTGGCTATTTTTTTTTTTTTTTTTTTTTGAGAAGGAGTTTCAATCTTGTGTCGCCAGGCTGGAGTGCATGGCACA
 GCTCAGCTCAGCTCAGGCTCCGCTCTCTGGTTCAAGGCAATCTCTGCTCAGCTCCCGAGTAGCTGGGATATAGGG
 ACCTCCACCAATGCTGGCTAATTTTTTGTATTTTTTAGTAGGATGGGTTTCCACATGT

SEQ ID:28

>STAR28

AGGAGGTTATCTCTGAGCAAAAGGCCAGCTAGTGAATGGAATAAATGCCATGTAAGATCTGTTTACCCTGAGAAGGGCA
 TTTCTAATCTCTCCATAAAAAAGGAGGAGGACCCAGATGAAATAGCTGATATGCTTTCTATACAAGCCATCTAGG
 ACTGGCTTATCAATGACAGGATATTTACCCACTGAAATAGGCTATTAACCAAGTTATGTTAAATGCTGTAGTTAAGGGGG
 TCCCTTCCACATGGACACCCAGGTTATAACAGAAAGGTTCCCAATCTAGACTCCAAGAGAGGGTTCTTAGACTCATG
 CAAGAAAGAACCTGGGCAAGTACATAAAGTGAAGCAAGTTTATAAGAAAGTAAAGAAACAAAAATGGCTACTCCAT
 AAGCAAGTATTTTTCTACTTATATGATTAATAAGAGATGGAATTTATCAATGAGTTTTCTGGGAAAGGGGTGGCAATCCCT
 GGAACAGAGGTTCTCCCACTTTTAGACCATATAGGATATCTTCTGATATGCTCAATGCAATTTGTAACCTGCTATGGCA
 CTGATGGGAGTCTTTTAGCATCTAATGCAATTAATTTAGCATATAATGAGCAGTGAAGATGACAGAGGTCACCTCTG
 TTGCCATATGGTTTTCAGTGGGTTTGGTTGGCTTTTTTTTTTTTAAACCAACCTGTTTTTATTTATTTATTTATTT
 ATTTATTTATTTATTTATTTTATTTTATTTTATTTAGATGGAGTCTGCTCTGTCACCCAGGTTAGAGTGCAGTGGCAACATCT
 GGCTCAGTCAAGCTCTGCTCTCTGGTTTCAAGCCATTTCTGCTGCTCAGCTCCCGAGTAGCTGGGACTACAGGTGCTG
 CCACATACCCGGCTAATTTTTCTATTTTTTCAATTTTTTCAATGAGAGCGGGTTTCCAGCTGTAGCCAGGATGGCT

SEQ ID:29

>STAR29

AGCTTGGACACTTGCATGCCACTTTGGAATGTTGAAGGGCCGCTCTCCCAACCCGCTGGGCACTTTTAAATATGCTCC
 CTCCTGCCAAGGGCCAGAGGAGGGCTGCTGAGGTTGACAGGAGTGCATGCTCTCACAGCAGGGGTTCCCGAGG
 ACCTTTTCTCCCATGGGAGCATAGAAGGACTAGAAGGGCCCTCCAAAGCCAGCTGGGCTGACAGGCAAGGAT

FIG. 20, cont

CGATGCCCTCCCGTACTCAGGTGGCCGTCTCTAAAGGTGTGTGTGTTTCTGTTCGCCAGGGGGTGGCGGATACAGTGG
 AGCATCGTCCCGAAGTGTCTGAGCCCGTGGTAAGTCCCTGGAGGGTGCACGGTCTCCTCCGACTGTCTCCAACACGTCAG
 GCTCAACAGCTGTAGGCACCGCTCGGGGAAGCTCTGGATGAGGCCATGTGGTCAATCCCGCTGGAGTCTTGGCTGGCCCT
 GAAGAGGAGGGGAGGAGGAGGCCAGCCCTCCCTAGCCCAAGGCC TGOGAGGCTGCAAGCCGCCCCACATTC TAGTCC
 AGGCTTGGCTGTGCAAGAAGCAGATTGCTGGCCCTGGCCAGGCTTCCAGCTAGGATGTGGTATGGCAGGGGTGGGGAC
 ATTGAGGGCTGTGTAGCCCCACAACCTCCCAAGTAGGGTGGTGAACAGTAGGCTGGACAAGTGGACC GTTCCCATC
 TGAGATTCAAGAGCCCACTCTCGGAGGTTCAGTGAAGCCGAGATCCTCCACTGCCTCCAGCTGGGCAACAGAGCAAG
 ACTCTGTCTCAAAAAAAGCAGAACGCAACAAAAAACCACCTCTGGCCACTGCTAACCTTGTAAATAAAGTTTTAT
 TGGCACAATAGACACACCCATTCATTTACATACTGCTCGGCTGCTTTTGCATTACCCCTTGAGTAGACGACAGACCAGTGG
 CCA TGGAAAGCCAAAAATA TTTACTGTCTGGCCCTTTACAGAAGTCTGCTCTAGAGGGAGACCCCGGCCATGGGGCAGGAC
 CACTGGGCGTGGGCAGAAAGGGAGCCCTCGGTGCC TCCAGGGCCTAGTGTGGTATCTCAGTGCCTG TTTCTGCA TGGAGC
 ACCAGGGCTCAGGGCAGTACC TGGAGGAGGAGGCTGTGTCCCGCCAGCAGCTGGGACCCAGGAGACC TTAGAGGCTCT
 TAACGAATGGGAGACAAGCAGGACCAAGGCTCCCA TTTGGCTGGCCCTCAG TTTCCCTGGCTGTAAAGTGGGAGGGCAGCT
 GTAAAGTGAACCTGTGAGGCAGAGCTCTGCTCAGCCATTCAGGGGGCCCTCTGCCCACTCCCTG TTTGTCACCCAGT
 GAGGGCCAGGGGTGAGATGTCACCATCAGCCCA TAGGGGTGCTCTCTGGTCCAGGTCGCCCAAGGATG TCCCATCCCC
 CCTGGCTGTGTGGGGACAGCAGAGTCCC TGGGCTGGGAGGGCTCCACAC TGTTTTGTCAAGTGGT TTTCTGAACGT TAA
 ATTTCAAGTGGAAAAATTC TCTTTCCCTTTACTGAAGGAACCTCCAAGGAAGACCTGACTGTCTGAGAAAGTTCCAGCT
 GGTCTGGAGCTGCCCAAGAACCCAGGTAAGTCTCCACGGCCGCGGCCAGGGGTGTGTCTGCGCCAGCCA TGGGCACAG
 CCAGGGTGTGTCTACGGCCGCGCAGGGGTAGGCTCTCCGCGCCCTCCGCTGCTGGCTGGGAGGGCCGTCCTGACACTGC
 AGGCCCGGTTGTCTCCGGGTCAGCTGACT TGTAGTCAACCTGCCCTTGGA TGGTCTTACAGCAAC TCTGSGTGTGGGA
 AGGGCCCTCCTGATTCAGCTCT TGGGAGGGTGGGGAGGGTGGAGCTGCCCTCCCTCCCCACCGCCCTGGCCAGGGTTG
 AACCCCTCTGGGAAGGACTCAGGCCCGGGCTCTGCTGTGCTGTAGCGTGGCCACTCTGCCCTTAGACCAGAGCTGGGCCCT
 TCCCGGCTTAGGACAGCCGGCCAGGACCAAGGGCTCCGAGTGAAC TCAAGGCTGCCGACCTGGAGGCCCTCTCCGGCG
 TCGGGTGTGACTGACAGCCAGGACGGGGGCTGTGTAA TTTGCTGTTTCTCTCCACAGAACCTTTTCGGGAAGATG
 GCTGACATCTGGAGAAGTCAAGAAGTAAAG TCCCGCCCCCAACC

SEQ ID: 30

>STAR30

CCTCCCTGGAGCCTTCAGAAGGAGCATGGCATAGGAGTCTGATTTCAGACGCTGGTCCCCAGAAATGATGGGAGAA TGA
 ATTTCTGTATTTAAGCCACCCAACTGTGGTCTTTGTA TAGCAGCCTCAGGAAAC TAACACACTGCACGTGCCACTA
 TTCCTTTTCCAGTATCTTTCCAGGACTGTCTGGCTTCCCTTTGCTCTGGGCTACACCCA TGCATGGCCCAATCCCCACTTC
 CTAAGAACAACCCCTGACTTAGTCTGT TGGGCTGCTAGAACAATACTATAGACTGGGTGACTTATAAACAACAGAAA
 TTTCAATTCACATCTGGAGGCTGGGAAGTCCAAATA TCGAGGCACCA TCAATTTGGTCTCTGCTCAGGCCCTTCTTA
 GCTCTCAGTGTCTCTTACATGGCAGAAGGGGCAAGGCCAGCTCTCTGGGGTCCCTTTTCAAGGCCACAAA TCCCATTCAT
 TAGGGCTGATGACTTCA TACTAA TCACTCTTAA TGGCCCACTCTTAA TCGCATTTGGGCGTTAGGAT TCAACA TAAA
 TTTGGGGGGACACACATATTCAGACCA TAGCAACCCCAACA TAAAAAACCTTCACTTTAAGGTTCCAAATGGACTGGC
 AGTTAAATCATGTTTATA TTTACATAAAGAAGGAGTAAGTCAACAATA TGTATAAACCGCTGGAGATTTGTTCGGATGGAT
 GTTCAACAAAATGCTGGCCCTTAAAGAGTGAATGGGAAA TGGGAAC TATACATTTCTTCA TACTTTTGGTACTGCCCT
 GCA TGT TAAAAA AAAAAAAGAGCAGAGCATT TTTACAA TCAAGAAAAAACAATGAGGTTATCTTCA TTTCTGGAA
 AAAAAATGGAAA TGAACAGTGGAGTCA CA TCA TGGAAAATGCTTATGGTACAATTCATGTGACATAAAAACAATAGAAATA
 GAGGACCTGTTTATGAC TAAAGCACTGTAAAAA TGAAGGCCCTGGAAGGAGAGATAAAAACCACTCATTTGTTAAGGTAG
 TCAGGTGGCAGGTGATTTCTCTCTTTTAAAAAATTTCCATTTTCA TATA TCGCAGTTTGTGCAATTTAC TAAAACTTTCCG
 TTTGTACACA TCCATAA TAGA TAGATAAATAAGTAGA TAGATGATAGATAAATAGACGGTAGGTAGATAGATAGATAGAT
 ATGAGAAAATAGTCCCTGTACTTGGCCCTGCAGCCA TAACTAGTCAATTCCTCTCTCTGCTCCATTTGCTA TGCCTGATGG
 ACAAGGCAGTCTGTGCCCTCTGGCCCAATTCCAA TGTGCCCTGCTCTCTGGCTGTAGTCCCTTTCCACCCCAATACAA
 TTGCTCCGAGGTCAC TTTAAGTGTGAAGCCCCAGATCAGATGGCTTCTTCTGTGTCCTTACCTTACC CAATTTCTAAT
 ATAACTAAAACACAA TGAAGCTCTAGTAAAA TACCA TGAAGACTTCAGGCCCTCTGTATAACTTCAC TCA TTTAAACCTAAC
 AAGGAAAACCTACCA TGAATCCGAGCCACAGAGCAGCTAAGGAACTCACC AAGGTCAGGCAGCTATTGG TGA TGGAAACAT
 GAGTCAAGCTTCAAGCC TGTGGCTCTAGAA TAGGGTTTCCCAACTCAGCACTGTGGACA TTTTCAGGCTGGATAA TTT
 TCTGTTGGGGGGCTGTCTGTGCTTTGTAGGA TATTAGGAGCA TCTCTGGCCCTTACCCA CTAGACCGCAGCAGCAC TCC
 CATGCCAGTTGTGACAACAAGCAATGCTCCCACTATTGCCAAGTGTCCCTGGGTGGAAA TGCACCC

SEQ ID: 31

>STAR31

CAACCCGCTTGGCCCCAGAGTGC TGGGATACAAAGTGTAAACCACTTCC TGGCTAGATTTAATTTTAAAAAATA
 AGAGAAGTAGGAA TAGTCA TTTAGGAGAGCCCTTAACTGGACAGGGCCAGCACAGGGGTGAGGCTTCCCTTANTTC
 AAGCTCACCTCAAACCCACCCAGGACTGTGTGTACATTTCTCCAATAAAGGAAAGGTTGCTGCCCCCGCTGTGAGTGTCT
 CAGTGGAGGGTAGAGGGCCGTGGCCAGAGTCTTCA TGGACTGCTCA TCAAGAAAGGCTTCA TGACAATCGGCCAGCTGC
 TGTCAATCCCAATTTACTTCCAGCTAGGGAAGGGCCGCTTGC CCAAGTCA CCAAGCCGCAAGTGTCAACCCCTGGGTTG
 GACCCAGAGCTATGA TCTGCCCAGGGTCCAGC TGAGAA TCAAGCCCAAGTCTAGGCGAGGGGCTCACCCTACTGGGAC
 TCCAGTAGCTGTAGTGCATGAGGCA TCA TGGCTGCAGCAGCT TGGACCTGGTCTCACACTGGCTGTCCCTGTGGCCAGGC
 CATCCCTCAA TGGCAGGTCAGGCCCAAGCATGTATCCAGACAA TGCATA TGGGTGGAA TCTCTCTCTGTC CCAAGGCCA
 CTCCTCACTGTCTACCTGAGGAAGCAGGGGCA TGGTGGAA TCTGAAAGCTGCTGTGAGGGTCTCCAGCGAACTTGCAC
 ATGGTCAAGCCCTGCC TCTCTCCCTGAACTAGA TTAGCGGAGAGCAAGAAGGACAT TGAACCAAGCAACCAAGAAATTTG
 GGGAAAGCCCTCA TCA TCCAGGTCAGGCTCACCTCTTTTAAAAATTAATTAATTAATTAATTTT TTAGAGACA
 GAGTCTTACTGTGGCCAGGCTGTAGTGCAGTGGCACAATCA TAGTTCAGTGCAGCCCAACTCCCACTCCAGCTC
 TGGATTAGCTGAGACTACAGGTGCACCACCACCACTCACTAATTTTATTTTGTGAGAGAGGGTTTCCACTCT

FIG. 20, cont

CTGGAGCTCGGCGCAGCGCTGGGAGCGGGCCCTGGAGGCTGGGCTGGGGACCCAAAGTTGGGGGGGGCGCAGGAGGT
GGGCTCAGGGTTCCAGAGAA TCCCAAGAGCTGACCCGCGAGGGCGCGGGCCAGTAGGCACCGGGCCCCGGGTGAC
CTGGCGGACCCGAAGCTGGAGCAGCCACTGCAAAATGCTGGCGTGACCCCAAAATGCTGTGTCC TTTAAATGTTTAAATTAAGA
ATAATTAATAGGTCGGGTGGTAGGCTCAAGCTTAAATCCCAAGCAGCTGGCGAGGCGGAGGAGGA

SEQ ID:39

>STAR39

TCAC TGCAACCTCCACCTCCAGGTTCAAGTGATTCCTCGCC TCGCCCTCCCGAGTAGCTGGGACTACAGGTGCA TGACA
CCGCACTCGGCTAGT TTTGTATTTTAGTAGAGACAGGGTTTCACATAGTTGGCCAGGTGGCTCGAACTCC TGACCTT
GTGATCCGCCCACCTCGGCTCCCAAGTGC TGGGATFACAGAGTGAGCCACTCGCCCTGGCC TGCACCCCTTACTATAT
ATGCTTTGCAATTTCTTTAGATTTGAAGAACCTCAATTAACCTC TAGCACTAATCTTATGTCAAGTAAATGCAATAGCAA
ATACTCCCTGAGCGGGAGAAATAATA TTTGCAAGTCTTCTTGTGAACATATGTTTTCAGTTCTAGGAGCCAGACGCTA
TGAGTGAAGAACCTAGTCACTCGTGAAGATGCTTCAACTTTGTAAGAAACTGCCAAACCTTTATTCATAATGCTGTAT
AAATTTTACATTAACCAATAATGATAGAGTTCCAGTTGCTTCACATCCCTCACAGCATTCTTCTTCTGCTCTTTT
TTCCTTTGGTTATCTAGTGGCATAAGATAAATAGTATCCCTGTGGTTAAATGTAATTCAC TGAAGACTAATAACA
TTTGCAATATCTAAATTAAGGCTTTTAAAGTACTTTCAAGTCTTTGCTCATTTTATAGATATTTGCTTCTTAT
TATTGATTTGAAAGAAATATATTTATA TGCTATATCTGGTTATAAGCCCTTTGTCATATTTTCCAAAACAAATATTGG
TGTCTCTGACTACTTCTTCTGCTCTTTGAAATGACTTGGTGCC TGGCCAAAATCAAATGACCACATACATGGGT
GCATCTCCAGACTACCAATTCGGTTTACTATTGTCCTCTCTCTGTCATTAACACTCTGCTTTGATAA TGGTAAGTTT
TGAGATCAGGTTGTAAGTCTCTAAATTTCTCGGGTTTCAAATATGCTTTGCTTTTAAAAATTTTGTATTTTCAT
TTACATTTTAAATAAATTTGTTAGTGGATTTGATTTGGCATTGCAC TGAACCTGTTGATCAA TTTGGGGAGATTGGACA
TTCTTATATATGATCCCGTGGTCACTCAACTTAAAGAACTTTCTCATCCATAGTAACCTCAA TCTAGGTTTCAATGCTA
CTCGTTTCTGCTCAGTCTGTGCTGAGCCCTTATGCTCTTCAATTTGTCATCCAAATTAACCTCAGCTTTGCATCAATAC
TATTTCTTGGCTTTGGTGGCTGTACTCTCTAA TCCAAATCCACAACTTACC TCCAAATTCAGGGCTTTGCTCATTC
TCCCAAGGAGGAGTGTCTCAGTCTA TCTACTAGTATAATAATTTCTCTGGCTGGTATCAAGGCACTCCCA TTTCCGG
CTCCA TGAGATGCTCAGAGGGCACTGCTCCCGGTGACTGCA TGGTCAAGCTTCTTCA TATCTCTTGGCTCATCACT
AAACTCACTATTTGTACTCTGCTCAGCTATAGGGAGCTACTGTAGTTCTTGAAGACATA TGTCTCTCTCTCTCT
ACATCTGGACCTGAGCACATCTCTTACTGCTGCTTGAACAAATGTGATCCCGGACACACCAATAGCTTAGAAGCCCT
CCCTGATCTCTCAAGGCTGGTTGAGTCTCTCTCTGCTCTCATGACAACTAGTTGGCAATCTCTCTGCTCAGCAGCTGAC
CCATGATCTCTTTGGAGCCAGACACTGAGTCTTTCTACTATGAAATTTCCAGCACTCATACAGAGCCCTGGCATATA
AAGCCCTCCATCAATATGATTAAGTGAATGGATAAATGA AAAAAAGTTATATATATGTACATATA TGTGTATATATGATA
TGTATATATGATATATATGTTGATA TATGCTGTGTATATATATGACATATATATGATCTATG TACATATATG TATAT
GTA TATATATG TGTGTATATATG TGTGTGTATATATATATACAA TGAATAC TATTCAGCTTAA AAGGAGGGA
ATCTGTCAATTAACCAATATGGA TAACTTAGAGGACTCTAAAGGCAAA TACCACATGTTCTCACTCACAAAATCAA
CAAGTGTACTCTTACAGTAGAGAGTAGGA TGA TGGTTACCAAGGGCTGGGGACGGGAGAGGATGGGAAAGCATAGCT
CTCCATCAAGGGTAGAAAGTTTCAATTTAGACAAGAGGAATCAGCTTTAGTGA TCTATTTTCAC

SEQ ID:40

>STAR40

GCTGTGATTCAAACTGTCAGGAGATAAGCCAGCAGATCAAGAAGCACTCCGGGCTCCAGAAGGAGCC TCCAGCCAGC
TTTGAGCATAAGCTGCTGATGAGCAGTGAGTCTGTGAG TAGTGTTCAGGGCAGCA TGTACCA TCAATGCTTGACTTCTA
GOCAGTGTGAGGAGAGGCTGGAGTCAAGTCTCTAGAGAGTGTGAGCAGCTCCAGCC TAGATCTCCAGTCTTATGCGGTGT
GCCCATTCGCTTTGTGCTGCACTCCCTGGCCACACCCAGTAA CAGTCTCTGGGATCTA TGGGAGTAGCTTCTTGTAGTGAG
CTTTCCCTCAAATACTTTGCAACAGGTAGAGAA TTTGGAGTGAAGTTTGTCTCTGTTCTTCA CAATA TGGATA
GCATCTCTTTTGA AATGTTAAAGTAAATACCTCTCTTTTCAGATACTGCTCTCATGCGAACTTGGTATCCTGTPTCCA
TCCAGCCCTTCTATAACCCAGTAACTCTTTT TGA AACCAGTGGGTGAGAAAGACACTGG TCCAGGAACGGGACACAG
GACACTCAGGCTCACCACGGCATCAGACTAAGGCCA AACAAGGACTCTGTATAAAGTACCGGTGGCATGTATATTAGT
GAGATGCAAGCTTCTCTGAGACAGGGAGTCA CACAGACACTTTCTATATAATTTCTAAGTCTTTGAA TGTCAAGTA
GAAAGTCTAACATATAATTTGATGAACAATGTA TATTCATGGATATTTTGGAAACGAAATACCMAAAATGGCAA TAGT
GGTCTTTCTGGATGGAAGACA AACTTTCTTGT TTTAAATATAATTTATTTTATATATTGAGTTGACCACATGACCTT
AAGGATACATATAGACAATAAATGGTTACTACAGTGAAGCAAA TTAACATATCTACCA TCGTACATAGTTACATTTTTT
GTGTGACAGGAACAGCTAAA TCACTATTTAA CAAAAATCTTAAAGCAATACATTTTATAA CTA TAGCCCTCATGA
TGTACATTAGATCTATA

SEQ ID:41

>STAR41

TGCTCTTGTGGCCAGGCTGCAGTGCAATGGCGCTGCTCGGCTCATCGCAACCTCGGCTCCAGATTCAAGTGATCTCT
CTGCTCACCTTCCAAAGTAGCTGGGATACCAGTATG CAGCAACACGCCCCGCTAATTTGTATTTGTAAATAGAGACGGG
GTTTCTCAATGTTGGTCAGGCTGGCTCAAATTTCTGCTCAGTGATCTGCCACCTTTGGCC TCCCAAAAGTCTGGGAT
TACAGGCAAGCCACTGTGCTCCGCTGGGCTGGGCTTTTAAAGGGACTGGAGGTTAGGGGCTGGAAAAATGGGAGAG
TTGATGGTGGGGCAAGGGGATGTAATCACTAGGGTGTACAAACTGCACCTCTGGTTTAGTCAGCTCC TGTGGGGTCT
TGGGAGCACTCTAGTCACTAGCTCCTAGTATACAGGACCCAAAGGAATACTCAAGGGAAACAGCAATTTCTAAGGT
TCAAGTTGTACTACGGAGCAGTACCGAACTACAACTGTGACAGGGCTCTACATGCTTCTGAGGCAATGACACACCA
AGCAGCTACGAGGAGCAGTCAAGAGCAGCCGACCTAGTGTACTGATGCTGTGATGCTGCGAGCTGGGTTCA TTTTCATT
TCTCCCTCCCTTCCCTCA TTAATTTGTAAAGTTATAGGGAACTTTCAACCACTCTGTGTGGA TCCCTGTCACTT
ACGGAGTCTGTCACTTTGGCTGTATGGGCTGTGGCC TCGCGGTCGCAATCTCAGGAGGTGTGAGACCA TGAAGCCCG
AGCTGGACAAGGCTAGAGACACACCCCCGCTCCATCCAA TCAATGTTTCTTGGGCTCTGGTTTCTATGCAAGGCTGCA

FIG. 20, cont

TGTCCTTAGTCCCTGCATGGGAACAGCTCCCTGTTGGTGAGCAGGCCCTGAGGAAGCCCTTGAGCGGGAATGGAGCCTAGGC
 TTAGCGTCCCTGGTAAGAGCTGGAGGGAACAGCCGAGGCCTGTTGCTACTTTTTTTTCCAGAAATGAAATACGTGACTGATG
 TTGGTGTCCCTGAGCGCCACGTTTCCCGCCACAACCAACCGGAACGAGGATGAGGGAACACACTCTCCGTGGACTGCACAC
 G

SEQ ID: 42

>STAR42

AAGGGTAGATCACTAGGGAGGGAGGAGCTATAAAAGAAAGAGGTCACCTCACATCTTACACACTTTTAAACAC
 TTGGTTTTTAAATGTCGGTGTCCCTCATAGCAGTAAGCCCTGTTGGAAGCAGGAGTCTTTCTCATTGACCACCATGACAAAG
 ACCCTATTATGAAACAATAATAGACACACAAAATGTTATCGGATA?TTATGAAAATAGGAATTTTCCCTCCACACCTC
 ATGACCACATCTGGTACATTTGTATGAAATGAAATATACCAATAATTTACCTATGGCTGTATATTAGGCTTTTTCTGTCAGG
 CTATAAAAATATGTAAGGGCCGTACAGTGACTTACGCCCTAGTCCAGAACTTTGGGAGGCCGAGGCCGGTGGATCAC
 CTGAGGTCGGGAGTTCAAACACAGCCTGACCAACATGAGAAACCCCGTCTGCTTAAATAACAAAATTAACGGACAC
 GGTGCCGTATGCCCTGTAATCCAGCTACTCGGGAAGCTGAGGCAGGAGAATCTTTGAAACCCAGGAGGCCGAGGTTGTGGT
 GAGTCGAGATTCGCCATTGCCACTCCAGCCTGGGCAACAGAGCGGAAATTCATCTCAAAAAGAAAGAAATGATGACTG
 TATTTAGAGTAGTATGTGGATTTGAAAATTAATAAGTGTTCGCAACTTACCTTAGGCTTTAACCATTTATGAGGGTGT
 GGTTTC

SEQ ID: 43

>STAR43

CAAAATAGATCTACCAAAAACAGATAATGCTGCCCAATTTTCCAAAAGATAATGGTGAAGTGGGTAGAGAGAAAATGCAT
 CCATCTCCACACCAACCTCTGCTAAATGTCCATGTCACAGTACTGAGACCAGGGGCTTATCCACAGCGGCAGAAATG
 TGCACCAAGCACCTCTGCTCAAAATTTGAGTCTAGGCCCTGCTTATTTGATGGTGTGAAGGCTTGCACCTGGCATGGAAGG
 TCCGTTTGTACTCTCTGCTTTAGCAGTCAAAAGAGCAGGGAGAGCTGCGAGGCCCTCTGCAGCTTCAGATGGATGTGGTCT
 AGCTGTTGGAGGCGCCCTCTGTTGGTCCATTATCTCCAGCCCCCTGGCGGTCTGCTGTTTGGCTTGGCTTGTCTGGCTCTC
 CATGCTTGTGGCTCCAAAATGTCATCATGCTGCCACCCAGGAAGAAATGTCAGGCCCATCTCTTTTATGTCCTTTGGCC
 TATTTTGAATCCCGCTGGGTATATTCCTTAGGTAAGACCCAGAAAGACACAGGAGGTAGTTGCTTTGGGAGAGTTGGACC
 ZATGGGTAAGGTAAZAGACACAGTATCTCTCTTTCAATTTGGTGAGACTGTAGCTCTGGCCGGGACTGAAATCCACA
 CAGCTCACTGGGAAACCTTTATCCAAAACATAGTACATTAACAATCTGGAGAAATGAGGACAGAGAAAGAGGCCCTAG
 ATTTGTACACTCGGGTGTATGCTATAAATAAGAAATGCTTTGGTGGTCACTAGACTGTTCATGTTGALATTTAGTCTTG
 CCTTTCCGGTGGTGTATTAATAATATGATATCTGTTTGGAAATATAGTGGAGCTAGGGTGGGCAATTTTCACTGGCTTT
 TTTGTTAGCTCAGCCCTCTGTTATGGCCAGCCTTGAAGCTCAGTAGCTAAATGAAGGATATCCTCACTCCCTCCAGAG
 AGCCGTCCTCCACAGGCTCAATGACAGTCTCTCAGCACCTTGAATGAGT?TAAAC?TGT?TATTTTAAACATCTCTGG
 TATGAAATGGCTATATTAATTAAGTGAACCTTATGTTGTGAAGAATGATTTGGTCTAAGGCTATATAAATTCAG
 GGACCAGTCTCTGAAGAGTTCAATTTAGCATGAAGTCAAGCTGTGGCAGGTTGGGTGGAGCCAGGGAACAAATGGAGAAGC
 TTTCAATGGCTGG

SEQ ID: 44

>STAR44

TGAGTTGGGGTCTAAGCCAGAAGTTAACTATGCTTTCAATATATCTTGGCAAGTAGAAGTACAGTGTGGGTGTAATTCCT
 CTAGATAGGATAGCTAAGCCAGAGGAAATTAAGTAAATGGAAACCATATGACCGTATGCAATTCATGTGCATATTTATAT
 CAAGAAAGAAACAATAEAGGTCGGGTGAGACCCCTATTTGCTCTGCAAAATGCTACTGTATTTACATGCTGTTTCCGGGAG
 TTTGGATGTCAAAGGATTTCTGCTCGGAATGTAAGCAATGCTCTCTGCTTGAATGAGCTACTCAATTTTGTATTTCTGAC
 TAATAAAGTCAATAAACATAATTCAACTCTGTTGCTGGTCTCTCTTCCATTAATTTATACTTTAGCAAAAAGTATGAAAT
 GTGTGTGTTATGTAACAATTTCTCTAATAATATATAAATGATTTATAGCTTTATTCATAAAGTTTAAAGTGTCTTCTT
 CTATGACTACATTTATTTGTTAACAAGAAATTTCTTTAACTGAAAACCTCAAGGAAGACTATCTGGGTAACCTTTCAAAA
 GAATGTCTCCCTGTATTTGGGATGAAATATATAATTTCTTGTACTGTTTAAACAGCACATAATTTACAAAGCAAGCCAC
 TTTTTCAAAAGCCCTGCTCTCTCTCCATTTCCCTATCTCTGTGATTTGACACCTCCAAACCCCTGAGCCCTGCTCTGCTCTC
 TCTTAAACAGTCTCTACTGATACCTTCTAAGTATTTTTCAGCCCTGCTCTCTCTCTCAATCAATGATGATTCACCTTCCA
 GTTGAATPCTTTATGCTACCTCCCTGCAATTAAGCAGTAATCAGAGAGCTGGTCTCTTAACTCAGGATTCACCTCTCTCT
 CATCTGTTGTTACAGTGCATCAGAAAGATTTTTAAATGATGAACTAGAAATTAATATATAAACAACACATACACACA
 TAAATAATACTTAAATTTTCAATGATGTTCCAAATTAATGAAATATAATAAGGAGCCACTTTATGTTCTGCCCTCAATC
 TTTCAATTCAAACTTATCTCTCTGCCACTATCTCTTTGAACATTTGATTTCCAGCTACTTTAGAAATATAAATAATACATAAT
 ATTCATAGAGCCCTTCTCTGGCTTCTTACACCGTACAAAATACTTCCACATATAACATTTAATCTTTGACAACTTTATTAGG
 CATGCACAATTAATATCTATCTATATACTATATCTATATAATAAATCTATATTTATAGATAAAGAAATAGAGGGTAA
 AAATTTCCAAAATACAAACCTTAGAAGTGTAGCAGTTGGGATTTGAACTAGGCACTCTGCCCTATAGTCTACAGTGG
 CTTTCTGTGCAAAAAGCCTTGCAGTTCCCTAGACTTAACTTTCTCAAAAATCTGCTCTTTTACATGCTCTTCCAAATGT
 CTGGAATAATTTCCAACTCAGTCTAAGTGTGGTACTCATGTTCCACCCACAAGAAATGACTCCATCTGTCCCTCTCTCC
 ATGAAAATTTCTTTGAATCTCAGCACTTGGGAGCCTCAGCCAGGTG

SEQ ID: 45

>STAR45

CAGCCCCAGCGTGGCTGGACTACTGCTCCCGAGGACTCTGTTCTGCTGCACCCCTGGACTACGGCACAGAGGACCCAG
 CTCCCGCCCGCTGAGCTATGGCACAGAGGACCCAGCTCCCGGACGCTGGACTATGGCACAGAGGACCCAGCCCGCCG
 CTCTCTGGGCTAAGGCACAGTAGGACCCCTGCTCTATGCTGCTCTCTCAGGAGGACCCCTGCGAGGGCCGGCAGCTGGA
 CTAAGCTACTGAAAGGACCCACCCCTGCTAACCCTGGACTAAGGCACTGGAAACTCTTGGCTCCGACAGCCACCGACT
 CTGCAAGAGAACCTCAGCCAGGCTGCTCCCTGGACTGTGGCACAGTAGGGCCACACCCACCCATGGACTCCTGTATT

FIG. 20, cont

GGAGGAAGAGTAGTGAATAATGTCAGGTTTACAACCTGAAAAGTAGCAATCAA TGCCACAAATAGATGGATGTGATGTA
 AAATTATAAA TGA TGA AAAACA TTTATG TGTAA TGCCTAGCCAGAACAGTTACACAAGACAAAGACGTAAAAGAAA TCCACA
 TAGGGAAGGAAGAGGTAAGATTGTTTCGTTTTTTGAAAATAAATCTTAAGATAGAGAAA TCTTAAAGATCCACCAAAA
 ATAAATGOTTATAGCTGATGAAGAAA TCAATAAAAGTAAATAGTTACAAA TCAACATACAAA TATCA TTTATGTTCTAT
 TAACTAATGACAAACTATTACC TGA AAAATAAAGGCAATTC AATTTATAATAGAA TCAAAACAGATATAAATATATAAA
 AGACAGGAGTAAA TTTAA TCAAAAACCTAAAAGATTACA TACTGAAAAC TATAGCACA TTGATGAAAAAATTAATAATGG
 CATAAA TAAA TGGAGAAA CATCTTCAT TGTATGGA TTTCAAAA TTAGTAT TGTAAAAGTGTCAATGCTACC CAAAGCAATC
 TACAGATTAATGCAACCACTA TCAAA TCCAAATGTCA TTTCTCAGAAA TAGAAAAA TACTGCTAAA TTTGTATGGA
 ACCA AAAAGACT TGGACCA CCAAAGCAATCTTGA CAAAAGAACAAAGCTGGAGGCA TCAAGACTACCTGCATCCCAAAC
 TCTATTACAAAGCTATAGGAA TAAAA CAGCATAGCAATGGCA TAAAA CAGACA TGTAAAACAGTACAAAGGGATATAGA
 ACCTGTAAA TAAA TCCGTGTGTCTGTGGTCAA TTGATTTTGTATAAAA TAACTAAAAA TACACAGTGAAGAAA GAAAAAT
 ATTTTCAA TAAATGGTGTAGACAAA CTGACTATCCACATACAGA GAA TAAAA TTTGACTTTTATTTTGCCTTTTATACA
 AGCATCAA TCAAATTAAGATTAAATGTAAA ACTACTACAAGGAAA TATAGAGGAGACTGTATGACA TTTGCC TGAGC
 TATGATTTCTGTAGATTAT TCCAAAAGGCA CAAAAGCAAAA CACAAAATGAGACTGCATAAA ACTTAAA ACTTTTCCA
 CAGGAAAAGAGCAATGATAGAA TTAAGAGAACCCCAA TGGGATAATTTTTAAACCATACATCAGGTAAGGGGCTCA
 TATAATAATATAAAGCAACTCAACTACTCAA AAAA TAAAGAAAAA ACTATGCTTATAAAA AATTAAGCAAGAA TACAGAA
 TAGACATTTCCACA TCA TCAAAAAGGCCAA CCAAGGTACA TGA AAAA TCA TAAACATTCCTAATTA TCAAGAGAGTCAA
 ATCAATGCCACAA TGAAGATACCTCACACA TTTACTAGGGCTA TTA TAAAAAAGATGGAAGA TAAAGTGTGG TGAGG
 ATGTGGAGAAAAGAAACCTGTACACTGTGGTAGGAA TGGAAATAGTACGCCATCTTGGAAAACAGTACCAAGCTTT
 CTCAGAAATTAATAATTTATTTACCTATGATCCAT

SEQ ID: 46

>STAR46

ATGTTTTCTCGCCCTTCGCAATTTCTGCAAA TFCGTTGAA TCAATGCAAGTACTTAGGTTTGCCTCGTCTCCCCCA T
 TACAACCTACTTAGTGGGTTTTCAACCC TAGTCCCTCA TTTTATGATTTATGCTCATTTCTTTGTACACTTCGCTCTG
 CTCCTACTTCCACTCATGGCCCTTGGCTTTGGATATTTGTTTGGTCITTTAATTTTTGCTTCTTACCTCAACACTTT
 ATCTTCCCTCCAGCTCCCGTACCTATCACCAAGGTTGTCATTAACCTTTCATATTTATCC TCAATTA TCCA TGTATTC
 ATTTTCAAATTAAGCTATATTAACAAAATCACAGGTTATAGGAGATATAATTCACA TACCTTAAAAATTCAGGCTTTTAAAG
 TGTACCTTTCAATGTTTTTGGTATA TTTCAAAAAGTTATGCA TTTGATCACCACCACTGATTTCCA TAACTGTTCATAC
 CTCAAAAGAGTCTGTACTCATAGTAGTCAATTCACA TTTCACTCCCTCTGGCTCTGGGCACTCACTGATCTTTGTG
 TCTCTATAGGATTTGCCCTAGTCTAGGTA TTTTATGTAATGGCATATACAACARGTGACTTTTGT TTTGGCTTTTTCAT
 TTAGCAAAATGTTATCAAGGCTCTGCCCTGTGTAGCATCTATTAGCAC TTTCA TTTCTTATA TGTGAA TGA TACTTTA
 TTTGTCATCAGTGTTCATGCTTTTATTTGTCCATCAGTTGATGAACATTTGGCTTTTGGCCACTTTGGGCTATTAAGAAAT
 AATGCTACTGTGAACAAGTGTGTACAAGTTCTCTACAAA TTTTGTGTGGACATATCC TTTCACTTCTCTCAGGTTGATA
 CTCGGGAAATGAA TTTCTGGGTCGTTAGTAGCTATGTTAAACACTTTGAGAAA CTGCTATAA TTTCTCCAGAGCTGTAC
 CATTTTAAATTCGTGTATAGGATTTCCAGTCTTCCACTTCTCACCAGTGTATGGATTTGGGGTATACTTTTAAAAA
 GTGGATTTAGGCTGGCCACAGTGGCTCACACTGTAA TCCAAAC TTAGGAAAGCTCAGTGGGAGGATCACTTGACCT
 AGTAGTTTAGACCCAGCTGGGCAACA TAGGGAGACCTGTCTCTCAAAAATAAATTTAAAA TAAATAGCTGGGGTTG
 TGGCACACACTGTATGCCAGCTACA TGGGAGGCTGAGGTGGAAGGATTCCTTGAGCCAGAAAGTTTGAGGTTGCAGTGA
 GCCATGATGCCAGCACTATCTGTAGCTGGGTGTCAGAGCAAGACTCCGTTTCAGGGAGAAA AAAAAAAGTGGGATGAT
 ATTTTGCACCTTTCTCTTCTTTCTTAAATTTCA TACTCTGGAAA TTTCAATTAATTAGCTGGTACCACCTCAACTCA
 TTGTTTTCATGGCTCAGTAGTAA TATTGCA TAA TAAAA TATACCA TTTCA TCAAAAGT TAGCAGATA TTAGCTGT
 GTGTCAGGCTCCTCTAAGCGTTAAAGAAAACACAAAAACTTTTGCATTTCTAGAGTTTATTTTCCAATGGAGGGG
 GTGGAGGAGGTAAGAAATTTAGGAAATAAATTAATTACATA TATAGCATAGGGTTTCAACCAGTGA TGCAGCTGAAATCGT
 TGGCAGCTTTCTTAGTAGTATAAATACAGTACTAAAGA TGAAT TACTCTAAAATGGTGTACTTAAAATTTACTGAA TAGGT
 ATTACTATTAGTCACTTTGCAGGTGAAAGTGGAAAACCACTCGTAAAATGTA AAA TAGGAAA CAGCTGGTTAATGTT

SEQ ID: 47

>STAR47

ATCATTAGTCAATTAGGGAATGCAAA TGA AAAACACAAGCAGCCACCAATATACACCTACTAGGATGATTTAAAGGAAAA T
 AAGTGTGAAGAAGGACGTAAAGAAAATTTAACCCTGATACATTTGATGGTAGAAA TGGATAAAGTTGCAGCCACTGTGAAAA
 ACAGTCTCAGTGGCTCAGAAGGTTAAATATAGAACCCCTGTGGACCCAGGAACCTACTCTTAGGCCACCCAAAAGAA T
 GAGAACGAAATCAACAGATGTTGTATACTAA TGTTTGTAGCATCACTTTTACAGGAGCCAAAAGGTGGAAATAATCC
 AACCATCAGTGAACAAATGAA TGTAA TAAAGCAAGGTGGTCTGCA TGCATGCTACATCACTCTGTAAAAAAGAAC
 ATCATTTTGA TAGATGATACAACATGGGTGGACATTTAGAACTATATGCTTAGTGAATTAAGCCAGACACAAAAGGAA TAT
 ATTTGATAAATGTTAA TACA TGAAGTGCCTAGAA TAGTCAA TTTCA TACAAGAGAAAAGTGGGATAGGAA TCACTAGGCT
 GGAAA TAGGGGAGGTTGCTATACGCTTATTTGTTGACAAAGTTTCTGTAAGAAA TCA TCAAAA TTTGTTGGTGTAGATG
 GTTGGTTATGCAACCCCTGTGAATATA TGAATGCCA TGGAGTCCACACTTTGGT TAAAAGGTTCAAATGATAAATAATG
 TGTATATA TATTTCCCA CGA TGA AAAACAGCAGCCAAAGCCACA TGGCAGCTTTGTTAGCTGCTTCTTTACCTT
 CAAGAGTGGGCTGAAGCTTGTCAA TCTTTCAAGGTTGCTGAAGACTGATATGATGGAAGTCA TCTGCA TTTGGAAAAGAAAT
 TAAATGGAGAGAGGAGAAA ACTTGAAGATCCACACTACTCACCC TGCAGGGCCAAAGAACTCTGCTCCCA TCTCTTGTCT
 CTGCTCAGTATTTCTGTGACCACTCTCTTTTCAACTGAAGACTTTGTACTGAAGGGGTTCCAGGTTTTTCACTCTG
 GCCCTTGCAGGACTGATCTCTCAACTA

SEQ ID: 48

>STAR48

FIG. 20, cont

ATCATGTAATTGTTTCGAAATTAATTCATTAGATACATTAATGTTTTATGTTACCAATGAAATGATATTATAATATAATAT
 TTTTAAATGGTTCCTACGTTTATAAGAAATTCATTTTCTGTTTACTTTCCCTTCAATCTGAAAAACCTTCGCTGATTTGAT
 TAGTCAATCCCAAAATTTTCCTGGATTTCTATGGGTAATACAAAATCCACACAAATGAGGTTCCAGTTCAGCCAAAGATCA
 CACCACATGACTCCAGCTGGGGCAGCAGATGAGACACCATCTCACAATAACATAAACAACAAACAGAAATCCACAC
 AATGACAAAGTATGCTTTCTTTTCTTCCCTCTTCTATAATAATTTCTTTGCTCTATCTTAACGAACTGGCCAGAAA
 CCCACAGCAATGATAAAATCGAGCAGTCTCAACAGACATCTCAATCCCTTTCTAGCTTTTATAAAAAATACGATTAATGCT
 TTCACATATACATATGGTGGTCTCGATGGTTTTGTATAGATAAGCTTATCAGGTTAAGAAAATTTGCTGCGTTTCTAGT
 TTGGTATAAAGATTTTAAATAAAATGAAATGTTGATTTTATCACTTATTTTCTTACATCTGTAAGGTAATCTGTTG
 TTTTCCCTTTTCAATCTCCATAATGGTGAATGACATTAATAATCTTCTATTTTAAATATTTCTGCAAGCGTGTATA
 GAACCAATGGCTTTATCTGTAATGCTGATGGATTTTGAATAATATGATGGTGGACTTAGTTTCTAAGGGGAATAGAAAT
 TTCTAATAATTTAAAAATTTTGCATGATGTTCTGAAGGACATGGGTGTCTATTTCTATACCACTGCTACTAGAGG
 AGCGACTGAAAGTCCACTGCGGAGGAGGGAGAGGTTGCTCTTCCGTTTCTGGTGTCTGATAGCCATCTCCAGTGGTAGC
 TGCAGTATAAATGCTGACAGTCTGCGACAGTCTGGAAGGAGCAACAACAGTGAATTCAGCAGCAGCAGTATTCGGGGAT
 CCCACGAATGGACAAAGGAAAATAATCTGGAAGCAATGACAAATACTAGCTGTGCTATAGCAGCTGAGATGCTAGTCTC
 ACGTGGCAGCTCAAGGACAGTAGTGAATGGTCCAAATGGCCGCCAGACCTAGAAAATGCACATTTCTCAGCAGCCGCTCCA
 GATGCTGAGCTTGGACAGCTGACGCT

SEQ ID:49

>STAR49

AAACAGAAACCCAAACAAATGGAGTGCATGCTAAAACAGAAAACCCAAACAAATGGAGGTTCTGCTAAACAGAAA
 CCCAAACAAATGGAGTGAAGTGCATAAACCCAGAAAACCCAAACAAATGGAGTGCCTGCTACACAGAAACCCAAACGA
 TGGAGTGCAGTGAATAAACCCAGACACCCAAACAAATGGAGTGCAGTGCATAAACAGAAACCCAAACAAATGGAGTGC
 GTGCTAAACCTCGAAACCTAAACAAATGGAGTGCAGTGCATAAACAGAAATGCAATAACAAATGCTGAGTGCAGTGCATAAC
 CAGAACCCAAACAAATGGAGTGCAGTGCATAAACAGAAACCCAAACAAATGAGAGTGCAGTGCATAAACCCAGAAAACCCAAA
 ACAATGGAAATGAGTGCATAAACCCGAAACCCAAACAAATGGAGTGCAGTGCATAAACAGAAACCCAAACAAATGGAAAT
 GACAATGCTAAACTGGAACCCAAACAAATGGTAACTAAGAGTGAATGCTAAGGCCATACATTTGGTGCACACTCTCAACTAA
 GTGAACTTGCATGAAAGGAGGATTTTCTTCTAAGACAGAGTTTGGTCTGTCCCCAGAGTGGAGTGCAGTGCAGTGCAGT
 GATCTGGCTCAGTGCAGCTCTGCTGCTCCCGGTTTCAGGCCATTTCTGCTGAGCTTCCCTGAGTGGTGGAAATAGAGG
 CACCGCCACACACTGGCTAATTTTGTATTTTAGTAGAGAATGGGTTCCACCAATATAGCAAGGATGGTCTCAATC
 TCTGACTGCTGATCTGCCCACTCAGGCTCCCAAAGTCTGGGATACAGGTTGAGGCCACACCCAGCAAAAAGGA
 GGAATTTTAAAGCAAAATATGGAGGCCATTTGTTGAACTAAGCTCAATGCAATAGGTTCCACAGACCAACCCAAAC
 AAACCAAAATGGAGTGCATGCTAAAATGATGATAATCAAA

SEQ ID:50

>STAR50

CAACCATGTTCCGCAAGAGCGGCTGTTTATAAACATGAAAATGAGGAAAAGGCTAGTAGCTCCATTTGGATTTGGGAAGA
 ATGGCAAGAGAGAGAGCGGCTCAATTTCTAGAAAGCAATCTCACACCTGTTGGTCCACACCATTGAAATGCTCTCACCA
 ATCTCCACACAGAAATGAGTGCATGTTGTCACATCGCTGTCATGTTGAAAGTATGAGTGTGAATGTGCTATATGG
 GAACATAATGTAATGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTA
 GTAATGTAATGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTA
 TGGAGGAAACTATAAACCTGGGCAATTTCTCTCTCAATGAGGAGGCTTTCTGGGCAATAGGGGAGAGCTCAAAAGTTCA
 CAGAGTATGCTGGTGGGAGGCAATGAGGCAAGGGGGAAGTACTGCGAGGAGGGGAGAGGGTCTGACACTTGGGGTTCTA
 ATGGGAAAGAAAGCCCAACATGAAATCCACTTAGCCCAAGCCCTGGGCCCAGCGGTGCGCGCTTCCAAACATACCAAC
 CATTTCCAAAGTGTGGCCGCAAGTTAACCTCTTTAGCCTCAGTTTCCCCACCCTGTAATAATGGCAGAAATGAAACCAAGCT
 TACTTCCCGGAGTGTGAGGATGAAAGAGCTATGTAACGTGATGCACTTAGAAGAAAGGCTAGGGTGTGAGTGGTACT
 CGTCTGGTGGGTTGGAGAAGACATTTAGGCAATGAGGACTGGGAGAGGCTGGCCCAATGGCTTCCACTCAGCAAGGTCA
 GTCTGTTGCTCTGCACTCCAGCCTTCCAGAGAGGACCTTCCCAACAGCACTCCCAAGCTGCCAGTCCACATAGTT
 ACACACATAACATCAATATATGTTCCCATATAGACATTTCACTCAATACCTTCAACATGCAACACGCAATGTCACACA
 CAGTCACTCATTTCTGTGTTGGAGATGGGTGAGGACATTCAAATGGGTGAGGCCAACAGGTTGTAAGATTTCTTTCTAGA
 AAAATGACTCTCTCTCTTTGGCAATCTTCCCAATCCGATGGAGCTACTAGGCTTTTCCCTCATTTCAATGTTTAAATAAAC
 CTTCCCAATGGGCAAAATGGGCTTCTCAAGAAATGGTGGATGCTCCCTGGGTCAGGGGTTGGCAGGGGCAAA
 GCTTCCCTGGAGGAACTGTCAGGCTGATTTCCAGTCCCACTCCAGCTTCCAAACACTCAATCTCCAGGCACTTTCAATC
 TTGGCTTAAATTTGCTCTGTTTCTTTTATTTTATTCGAGAACTGGGTGGAGGCTTTTGGTGTCAATGGGGATGCT
 TTTGAAACCTTCTGCTGCACTGAGGCTGGCTTTCAGTCAACTGGTCCCAATGAAATTTCTTTTATGTTGTTAA
 CAGTAAATTTTGGCAGCTGTTGTCGGTCCAGGTTGGAAAGCAGCTGTTGATGTTGAAATGCTTTGGCTCAGATTTCTT
 GGCAACAGATGCTGCTCTCTCAACTCACCATAAAGAAAGCCAGAAAATGCTGCTGGAGCACCAATGCTGCTGTTATC
 TCTAGTAAATCAGATGGCTTCACTCGGAAGCCCTTCTGSAATGTCAAAGCCAATGAGATAAAGGACATATATATAGTAGCT
 AGGGTGGTCCACTTCTTAGGGGCAATCTCCGGAGTGGTGGAGCAATAGTCCAGGAAGAGAGGAAACTCTGTTTGGAGC
 CAAAGCAATAAAAAACCTTAGCCACAAACCAATGAAATTTGTTTGTGCAAGTCTGAGTCCAGGGAGGGCTTCTGAGGA
 GAGGGCAGCTGGAGCTGTTAGGATTAATGTAAGTGAATGTTGAGATGGGCAAGGGCCCTTAAGAGGTGGGAGCAGCATGAGCAAGGCA
 GAGAGGTTGATATGATAAGGATATGTTGAGGAAAGATTTGGCTGGAACAGATTTACAGAAATAGAAAAATCAACACTA
 TTAATGAGCTCTACTAGCTGCTGACATTTCTAGTCACTGAGATAGGTTTGGTATACAAACAAATAACATCTCTA
 TGACATTTTATGACTAACCAATAATAATAAATAAAGTAAACAAAGCTCAAAACATGCCAGGCACTATTTATTTATTT
 ATTTATTTATTTATTTATTTATTTTGAACAGAGTCTGCTCTGTTGCCAGGCTGGAGTGTAGTGGTCCATCTCGCC
 TCACTG

SEQ ID:51

FIG. 20, cont

>STAR51
 TCACAGGTGACACCAATCCCTGACCAGGCTTTGAGAAGCACTGTACTAGATTGACTTTC TAATGT CAGTC TTTCT
 AGCTCTGTACAGCCATGGTCTCCATAATATCTAGTACAACACACATACAAATATGTGTATACAGTATGAATATAATATA
 AAAATATGTGTTATAATAATAAATAAATAATATAAATAATGCTTTTATACTAGATAAATACTTAATAAGTTGAGTGTTA
 ACTGCTTAAGCACTTACCTGACAGAAACAGTTTTTTTTTATTTGGTGAAATACAATAAATAAATTTATTACAAT
 TTTAAGCATTTTAAGTGTATAGTTTAGTGGAGTTAATAATATCAAAATCTGTGACGCCCTCACCATCATCAGTCTCAT
 AACTCTTTTCATAATGTAAAATAAAGTTTATGCTCATTTAAAAATGACTCCCAAATTTCCCTCCCTCCTCAACCTCTGGAA
 ACTACCATTCATTTTCGCCCTCCGTAGTTTGGCCACTTAAGTACCCTCACATAAGTGAATTTGCTTATTTGCCCTGTT
 TGTGACCGGCTGATTTTCATTTAGTATAATGTCTCAAGTTTATTCACGTTATATAGCATAATGTCATAATTTTCCTCACTT
 TTAAGCTTAGTAAATTTTCACTCGTATGTAATCTCACATTTGCTTAATCCATCTCTCAGTGGACACTTGGATGCTCTC
 TACATTTTAGCTGTGTTGAATACCTGCTGATAAACAATGGGTGATAAATACTCAAGACCTTTTATCAGTTTTTAAAA
 ATATATCTCAGTAGTATGTTAGTGGATTAATGGTAAATTTATTTTAAATTTTGGGAACGTCTCAGCTTTTATTCAC
 ATAGTAGCTATACCAATTCACAAATGGCAATTCCTACCAACAGGCCATAAGGGTCTCAAATCTCCACATAATCCCTGATAC
 TTGTATTTTACAGGTGTTTTTTTTTTTTTTTTTTTTTTTATGGAGCCATGTTAATGGGTGTAAGGTGATATTCATTAAG
 TTTTGTATTTGCAATTTCCCTAAATGATTAGTGAATGTTAAGCATCTCTTCAATGTGCTATTTGGCCATTTGTAATCTCTTAA
 AAAATATAATATACCTCAATTCCTTTGCCATTTTGAATATGTTTATTTTTGTTATTGAGTTTCAATACTTTCTATAATA
 AACTAGGTATTAATCTCTTATCAGACTTAAGATTGCAAAATATCTCTTTCACTCCACAGGTGCTAATCTCTCTGTGTG
 TAAATCTTTTGAATGCTGTGTGCTCCAGAAATGATTCATTCCTGTGGTCTTGGTCTCACTGACTCAAGAATAAAGCTG
 CGGACCTTAGCTGTAGTGTACACTTCTTATAGATGGTGTTCGGAGTTGCTTCCCTCAGATGTCTCCAGAGTTCTTCT
 CTCCAAATGGGTCACTGGTCTGCTGACTTCAGGAAATGAAGCCGAGACCTTCGAGTGGGTTTACAGCTCTTAAAGGTG
 CGGTCTCCAGAGTTGTTGTTCCTCCCTGGTGGTCTGCTGCTTCCAGAAATGAAGCCGAGACCTCCGAGT
 AGTGTACAGCTCAATAAGGTAGTGGGACACAGAGTGAAGTTCAGCAAGATTACTGTGAAGAGCAAAAGAACAAAGCTT
 CCACAGCATAGAAGGACACCCAGCGGGTCTCTGCTGCTGCTCAGGTGGCCAGTTATTTCCCTTATTTGCCCTGCCA
 CATCTCTGATATGGTCCATTTTACAGAGTACTGATTTGGTCCATTTTACAGAGTCTGATTTGGTGCATTTCAAACTCTTA
 CCTAGACACAGAGTCTGATTTGCTGCAATCTTACAGAGTCTGATTTGGTGCATTTACAGTCTCTTAGCTAGATACAGAAAG
 CTGATTTGCTGCGTTTTTACAGAGTCTGATTTGGTGCATTTACAATCTTTAGCTAGACACAGTCTGATTTGGTGGGTTTT
 TACAGAGTCTGATTTGGTGGCTTTTACAGAGTCTGATTTGGTGCATTTACAATCTTTAGCTAGACACAGAGTCTGATTT
 GGTGGGTTZATAATCTCTTAGCTAGACAGAAAGTTTTCCAAGTCCOCCCTGACCCGAGAAAGCCOCCACTGGCTTCACT
 CACTGTTAATCTTTGGACATTTGTCCCCCAAAAATCTCATGTTGAAATGTAACCCCTAATTTGGAACTGAGCCAGACTG
 GATGTGGCTGGGCCATGGGGA

SEQ ID:52
 >STAR52
 CTTATGCCATCTGGCGGTGCCATGTGGAACTTGCTGGAAGAAGCTAAATTTACTGACCACTGTGCTTAGAGCGGGTTCT
 CCAAGGAAAGGCTCTGTAAATCTCGTCCTTTTGAAAATCAGGGGAAACAGCCCTCTTCACTGAGGATTAATTTAAAGAAA
 GGGGAAAATAGGAAAATTCATCGCTGTGGAAGTCCAATTAGATTTCTACATGAACCAATCATATATGTGCACTACATAATCT
 TTAATTTTTTATTTTTAAAAAGGGATAATTAATTTCCAGTGACAAGTTTGGGAAAGGCCAAGCCAAATGGAGTTGA
 ACATTAATGTAGCGTTTATATAGACTTGCAGAGCTCTGTGCAATATCCACCACGTAACAGTGGAGTCTACTCAAGTCTC
 TCTGCGCCCTGGTAAATGTGACTTCCCTTCTTTATTTGCAATGAAATCGCCCTGGAATGGGTGTCAGGTTTTAAAACTCAAAG
 TTTAGCCCTATGTTGTTCAACCAATCAGCATCTTACTTTGACGTGATTTGGCTTCTACTGTAGGTGTCAAATCAATCCAAAT
 TGCACTACTTCTCAGGCCCGGGGAGCTGTGAGTGGCTGTGGCAGCTGGAAGAGAAAGAAATCGGACGGGAAAGAAAT
 AAAAACTACTTTGCTTTCCGAAAGCGAAAGAAAGTATCTTTTCTCATTAATTTTTAAAATAAATTTGATGTAATTTTAC
 CTAATAAAATAAACAATCAATTAACAAAAATAAGCAACTATCAAAAGATTTGTTACTAATTTTCTGTAATGTTACTGTTT
 CAATAAGTAGCCAAAGGAAATATAAACAACAAAAATAAGATGCTGATAATTTATGTCATAAAGACCAATTTAAAACTAA
 AAGTGAACATGGGTTTCTAAAATAAATTTACCGTGTAGCGTAAAAACACTCTTTCAAATCTTGGCCATGCTGAAAGTGC
 TGCACTCTAAGATAAAAAATACCAAGGGGGGGATTTCAAAAGCAATTAATTTGCTTTAAATAATCTGTATTTCTGTCA
 CTTTCCCTTTTTATTTAATTTACCGTGAATCACAGACAGAAATTTACTTGGAGTTTCTGAAAATCTGTGTTGTGACAT
 TCTCATCTTACAGTACCCACACACCCAAAAATAAAAAACAAAGAGAG

SEQ ID:53
 >STAR53
 CCTTGAGGAGATGACGAGTAACTCCGTAGAGAACCTTCCACTCATCCCCACATCCCTGCAGAGTGTCTATCTGTTAT
 GATCTGGTATCCCATCTGTACTGCTGCCCAAAATCATCCCTTCTACAATTTTCTACTGTACAGCATTTGAGGCTGAAC
 GATGAGAGATTTCCATGCTCTTTCTACTCCCTGCCCTGTATAATATCCGGGATCTCTCCCTACCCAGGATGCTGTGGGGT
 CCAAAACCCCAAGTAAAGCCCTGATAATGCGGGCCACACTTTCTCTAGCCTAGGAAATGATAACCCAGCGGAGGAAATCACT
 TGGCAATGAACAGATGTTCACTTGCAGAAACCGTGGAAAGCCGTGTCAGGTCCTGAGATAGGGCAGAAATCCGGAGTGTGCA
 GGTCTGCAGGTACAGGAGGATTTGAGATTTGCGTTGCCAGTGGTGGGAACTCACTGCCACTTATTTCTCTCTCTCTCTTCTG
 CTCACCTCAGGGATACGACACATGCCATGATGAGAAGCAGAACCTGGTGCACCTTCAAGAACATGGCCATGGCTGGCGA
 CCCCCTGCTACAGGTGCAATGCAAGTGAAGCAAGTCTTCAACAGTGAAGTTGAGCGTCAATTTTCTTAGTGTGCC
 AAGAGTTCGATGTTAGCGTTTACGTTGTAATTTCTTACACTGTGTCATTTCTGTAGATACTAACAATTTCAATGATGAGCA
 AGACATACTTAATGCAATTTTGGTTTGTGTAATCCATGCACCTACCTTAGAAAACAAGTATTTCTGGTTACTCTGCAATG
 AACAGCATTTACCTCTCTCTCCAGATGTGACTACTGAGGGCAGTTCTGAGTGTTTAAATTCAGATTTTCTCTCTGCA
 TTTACACACACGACACAAACCAACCAACAAAGTA
 CCAGTATAAGCATCTGCCATCTGCTTTTCCCATTTGCCATGCGTCTGCTGGTCAAGCTCCCTCCTCTGTTTCTCTGGTCA
 TGTACTCCCTCACTCCGATTCCTCTGAGCAGTCACTGACAGTAAATAAACCCTTTGCAAAAGTTCCCAAGTTGTTTCTGCTG
 TGCCATTAATGTGCACACAGCTCTGTGCAGCTGTGTGCAATTTCTTTAGGAAAGATTTAGAAAGTGGAAATGCTGTGTC
 AAAGAGTCAATTTATCAACAAAACACTAATGAGTGGTCTCTGCTGAGGGCTGTTCTAGTGTCTGGAGCGAGCTCAGG

FIG. 20, cont

GAACAAGGCAGACAGGAGTTCC TGACCCCGTTC TAGAGGAGGATGTTCCAGTGTGGGTTTGTGTTGTTGTTCTTC
TAGAGATGTTGCTTTGCTCTGCCAGGCTAGAGTGCAGTGCCATGATCA TAGC

SEQ ID: 54

>STAR54

CCAFAAAAGTGTTC TAACTGCAGAAAA TCCCCCTACAGTCTTACAGTTC AAGAATTTT CAGCATGAAA TGCCTGGTAG
ATTACC TGACTTTT TGGCAAAAA TAAGGCACAGCAGCTCTCCTCC TGACTTGACTTTCTATAGTCTTACTGAAATATA
GTCCTTACGAAATTCATTC TCAAGTGTGCAAGTCTGAAGGACACCCACATTTCTCTTTGCTTTTGTCAATCTTTGTGTT
GTAAGGCAGGATGTTTAAAAGTGAAGTCAATGACTTGC AAAA TGAGAAATTT CAGAGGGCATTTTGTCTCTAGACCA
GTAGCTTAGAGCAGTGTTCACAC TGAGGTTGCTGTAATGTTTCTGCACTTCTACCAATAGTATCATTACCCAGCAACA
GGA TATGA TAGAGGACTTCGAAAACCCAGAAAA TGTTTGGCATA TA TCCAAAGCCCTTTGGGAAA TGGAAAGGAATTC
GGCTCCCA TTTTATA TATGGA TAGATAGAGAC CAAGAAAGACCAAGGCAACTCCATG TGCCTTACATAAATAAAGTACA
AAA TGTAAACATG TAGGAAGTCTAGGCGAAGTTTATGTGAGAAATCTTTACACTAATTTGCAACATTTAAATGCAAGTCT
GAAATTAATGTC AAAATAAGTAAAAA TTTTACAA GTTAAGCAGAGAA TAAACAATGATAGTACAGAGAAATAAGTAGCAAAA
TCTCTCTCAGTATGACTTGGTTGCTTTTCAA TCTC TGAGGACACAGCAGTCTTCCCTTCCAAA TCCACAAGTCAATC
AGTGAGGAGACTCAGCTGAGACTTTGGCTAAATGTTGGGGG TCCCTCTG TGTCTCCAGGCGCAGTGGCTTCAGGCC
GACTCAGCTCGTGGCACACAAC TAAATCTGGGGAGAGCAACCGATGCCAGCATGATGCAGATA TCTCAGGGTATGATCG
GCC

SEQ ID: 55

>STAR55

CCTGAATCATGATCCGCCACC TCAGCTCTCTGAAGTCTGGGATACAGGTGTGAGCCACACACCCAGCCGCAACACA
CTCTGAGCAACCAATG TGTCA TAAAAAGAAA TAAAA TGAAATCAGAAAGTATCTTGAGACAGCAAAAA TGGAAACACA
CATACCAAAA TTTA TGGGCACACAGCAAAAGCAGTTTATAGGAGGGAAGTTTATAGTGA TGAATACCTACCTCAAAAATCATTA
GCC TGA TGGATGACACTACAGTGTATAAAA TGAATGAAAA CCAATGTCGCCCCA TACATA TACAAA TTTTATTTGTT
AATTA AAAA TAAAA TAAAC TTTAAAAAAGAAAGAGACTCAAA TAAACAACCTAACTTTTATACCTCAAGGAAATAGAAG
AGCCAGCTAAGCCCAAAGTTGCAGAGAAGAAAAAATA TGCACAAGAAA TGAACAAGAGACTAGAAAAGCAATGGAAG
AGATCAGCAAAACTA

SEQ ID: 56

>STAR56

ACACAGGAAAGATCGCAATGTTTCAGCAGAGCTTTGAACCGGGGATGACGGTCTCCCTCGTTGCCCGGCAACA TGGTGA
GCAGCCAGCCAGTTATTTCTCTGGCGTAAGCAATACCAGGAAGGAAGTCTTACTGCTCTGCGCGCCGGAGAACAGGTTGTT
CCTGCC TCTGAAC TTGC TCCGCCATA TGAAGCAGATA TAAAGAACTCCAGCCCTGCTCCGCCAAGAAAACGATGGAATA TGA
CTCCTCAAAGAAAGCCGTTGAATA TGGACGGGCAAAAAAGTGA TAGCCGACCGCCCTTA TGGCCCGGGA TGGGGAGTAA
CTTTAG TCAAGCCGTTGCTCCGGGTGTCGGTGCAGCAGTTGCACGTCATTTCTCAGACGAACCGATGACTGGATGGA TGGCC
GCCCGACTGCTCACACTGA TGA TACGGA TGTGCTTCCGTA TACACCA TGTTA TCGGAGAGCTGCCAACG TATGGTTATC
GTCCGGTATG GGGCTGCTTCCAGACAGGCAAGACTTGA TGGTATCCCTGCGATCAATGCCAAACGTTTACCGGATCA
TGGCCAGATA TCCCTG TGTCTTGAGCGAAAACCTGCTGTACCGCA TCGAAAACGGGCACATACAGCCAGAGTGGCCGTA
AAGAAACAA TACGCGA TGTGCTCTGCAGGGTTGAGTTCTGCTGTGATAACGGAGAGACTGCGTGTACAGTTCCGGC
TGGACTGCTGTG

SEQ ID: 57

>STAR57

TCC TCTG TAAATAGGCAAAA TGTATTTAGTTTCCACCACACATGTTCTTTCTGTAGGGCTTGTATGTTGGAAATTTTA
TCAAATTTCAATTAACACTATACCAACAACTGCTAA TTTCTGGAGA TGTGGCAGTGAATAAAAAAGTTATAGTTTCTGA
TTTTGTGGAGCTTGGACTTTAA TGA TGGACAAAACAACACATCTTAAATATATA TTTCA TCAAAA TTA TAGTGGGTGAAT
TATTTATATG TGCAATTTACATG TGTATGTATACA TAAA TGGCCGTTACTGGCTGC ACTGAGAA TGTACACTGTCGCCGAA
CGAGGC TGGCCGTCAGAGAAGGCC TCCAAAGGAGTGGCTTTGAAGCTGAGTGGTGGCTTCCACTGAAAAGGCTTGGAAAG
GGCATTTCCAAAGAAAAGGCTGAGGCCAGCGGAAAAGAGGTTCCAG TCCGCTCTGGGAAACGGAAAGCCACTTCCCTGAAAACG
AAAA TGAGTGTGCTGAAA TAGGACCTGAAAAGGGAGGCGAGGCTGGCAAAAAGCAACCGAGGAGGAGCTCAAAAAGGAGCA
GGCGGGAAGGCCGCTGTGGAGCTTGGAGGAAGCACTTCCGAAGCGCTTCTGAGCGGGTAAGCCCTCGGAGCATGAACTG
CTGAGCAGGTTGTTCCAGAA TTTCTGGGTTTGGTCTCACTGACTTCAAGAA TGAAGAGGACCGCGGACCTCGCGGTG
AGTGTACAGCTCTTAAAGTGGCCGCTGTCGGAGTTTCTCCTTCTGATGTTCCGA TGTGTTACAGATTTCTTCC TCTGCT
GGTTCCTGGTCTCGCTGGCTCAGGAGTGAAGCTGCAGACCTTCCGGTGTAGTGTACAGCTCATAAAAGCAGGTTGGACT
CAAAGACTGAGCAGCAGCAAGA TTTATGCAAGAA TGAAGAACA AAGCTTCCACTCTGGAAGGGGACCCAGCCGGT
TGCAC TGTGCTCCGACGCTGCTTTATCTCTTATCTGGCCCAACCCACATCTGCTGA TGGTAGAGCCGAA TGGT
CTGTTTGCAGCCGCTGA TGGTGGCTTTACAA TCCCTGCGCTAGA TACAAAGGTTCTCCAGTCCCAACAGATTAGCTA
GATAGACTTCCACAAAAGTTCTCCAAGGCCACCAGAGTACG TAGATACAGAGTGTGATGTTGGTCA TCCAAAAC
CTGAGCTAGACAGGGTGA TGAAGTGGTGT TTTACAAAACCTTGGGTAGATACAGACTCAAA TTTGGGTTATTTACAATC
ACTGAGCTAGGCA TAAAGGTTCTCCAGTCCCACAGACTCAGGAGCCAGCTGGCTTCCACCAGTGG

SEQ ID: 58

>STAR58

AAGTTTACCTTAGCCCTAAA TTTATCTATTGTA TGGCA TTTTAGGAAA TATGTA TTAAGGAA TGTCTCTTAGGAGATAA
GGA TAA CATATGCTAAGAAAA TTA TATTGAAA TATTATTA CA TGAAC TAAAA TGTAGAACTGAAAAAAATTTATGTA
CTCCTTCCAGGCTAGGCAGGAGTATCTAGATACCAACTTAAACAAC TCAACTTTAACAAC TCGAACCAACAGATGGCTA

FIG. 20, cont

GTGGGTGACGGCCTGTGCGGGTGACCTGAGAGCCAGGGCTGGCAGGTCAGAGTCAGGAGAGGGA TGCCAGTGGATGCGCCTG
 TGCAGGATCTGCCTAAATCATGGTGGAGGCTGGAGGAA TCCAAAGTGGGCA TGCACCTGCACTCAATTTCTTTATTCATGTGT
 GCGCAATCCCAACAGCAGGGAGCCTGGCCAGGAGGGCCCCGGGAGAGGCACTGATGGGCTGTGTCCATTTAGGAAGGA
 TGGACGGTGTGAGACGGTAAGTCAGAACGGGCTGCCACCTCGGCCGAGAGGGCCCGTGGTGGGTTGGCACCATCTGG
 CCTGGAGAGCTGCTCAGGAGGCTCTTAGGGCTGGGTGACCAGNCTGGGGTACAGTAGCCATGGGAGCAGGTGCTTACC
 TGGGGCTGTCCCTGAGCAGGGGCTGCAATGGGTGCTGTGTGAGCACACTTCTCTATTCACCTGAGTCCCNCTGAGTGT
 GAGACACCCCTTGTTTGCAATGAATCTGAGCATGGAGATGTTAAGTGGCTTGCCAGCCACACAGCAGATGGATGGT
 TAGCTGGGACCTGAGGGCAGGCTCCAGCCGAGGACTTCCCAAGSTTGGGCAACTCTGACACATGACCCCAAGGA
 ACACCCATCTCAGCTCTGGTCAGACACTGCGGAGTTGTGTTGTAACCCACACAGCTGGAGACAGCCACCTTAGCCCAACC
 TTAGCTCTCCCAAGGAACCTGCCCTTCCCTTCAATTTCCCTTACTGCATTGAGGGACACACAGTGGCCAGAAAGGA
 ACATGGCTCAGGACCCAGATGGACTTGCCTCAGCTGACAGCCCTCCTCTCTTTCAGAGTGGCTCTTCCACTGTGAAG
 TTGGGACATCACCAACTCAATCTGCTGGGCCGCTCACAGGTGGCCAGGCAAGGATGGCAGTCACTGGCTGTGGT
 CTGACAGGCTGG

SEQ ID: 62

>STAR62

AGTGTCAAATAGATCTACACAAAACAAGATAATGTCTGCCCAATTTTCCAAAGATAATGTGGTGAAGTGGGTAGAGAGAAA
 TGCAATCACTTCCACCACCAACCTCTGCTAAATGTCCATGTACAGTACTGAGACCCAGGGGCTTATTCACAGCCGGGA
 GAATGTGCACCAAGCAGCTCTTGTCTCAATTTGCACTTAGCCCTGCTATTTGATGGTGTGAAGGCTTGCACC TGGCATG
 GAAAGTCCGTTTGTACTTC TTGCTTTAGCAGTTCAAAGAGCAGGGAGAGCTGCGAGGGGCTCTGCAGCTTCCAGATGGATG
 TGGTCACTGTGTGGAGGGCTCTCTGTGGTCAATATCTCCAGCCGCC TGGGTTGTGCTTTGCTTTGGCTGTCTGG
 CTCTCCATGCTCTGTGGTCCAAAA TGTCATCTAGCTGACCCCAAGAAAGATGTCAGGCCCA TCTCTTTTATGTTGCTT
 TGGCTATTTGATTTCCCGTTGGGTAATTTCCCTAGGTAAGACCCAGAAAGACACAGGAGGTAGTTGCTTTGGAGAGTT
 GGACCTATGGGTATGAGTAAATGACACAGTATCTCTCTTTCAATTTGGTGTGAGACTGTTAGCTCTGGCCCGGACTGAAT
 CCACACAGCTCACTTGGGAAAAC TTTATCCAAAACATAGTCAATGAAACATTTGGAGAA TGGGGACAGAGAGAGGC
 CCTAGATTTGTACATCTGGGTGTTATGTCATATAATAGAAATGCTTTGGTGGTCAACTAGACTTGTTCATGTTGACATTTAG
 TCTTGGCTTTTCCGGTGGTGA TTTAAAAATTA TGTATA TCTTGTGTGGAATA TAGTGGAGCTA TGGTGTGGCAATTTCACT
 GGCCTTTTGTCTAGCTCAGCCCGTCTGTATAGGGCAGCCCTGAAAGCTCAGTAGCTAAAGAGGTA TCC TCACTCCCTC
 CAGAGAGCGGTCCCTCACGGCTCA TTGAGAGTTTGTCA

SEQ ID: 63

>STAR63

CGCTCTGAGCCGCTGGGAACCAATGAGCCCGTCCA TGGAGTTGAGGAAGGGGTTGCCCCACGGGGTGGGCGCCCTCTA
 CACAGCCCGCTTCCCTTCTCTCTCGTTAGCCGCGGGGACAGCCCTCTGGTTCTGCACTCCGCTCTGGGAGCAGGCGCG
 GCTTTGGCGAGGCTTCCCGGGGCTGCGCAGCTCTGCTCCGCTCGCCCGCAGCCCGGCTCCGGAAAGCCCGCAGG
 TCCAGTCCAAAGCCCGAATTCGCCAAGCCAGGGCCCGGGGAGCAATGGAAACAGGGCCGGGACCCAGTCTCCGAGCA
 TGGAGTAACTGCACTTTTGGAAAAGAAAGCGGACCCACCCATCGAGAACCGGGGCGCTTGTTTAGGGACGTTCTCTGG
 CGCTCTACCGAGTGTCCCGGCTCTCTGGGCCCTCCCTCTCCAAAGCCGCCACCCCGACAGCGGCTCCCTGGGGACCTC
 CCTCGGGCTGGCTTTTCCAGCCAAAACAGGGAGGCTCTCCAGGAGCTGCCAGTCCCAAGCAGCAGCCAGGACAGAA
 ACTCCACCTTACTTCCAGGCTTCCAGAGAGGGCGGCTGGACA TGCCTCCGACGGGAGGAGCCCGCTCAGCACCC
 TCCAAAGTGGCAGCAACCCAGAAACCCCGTGGAGGCTCTGAGCAGCCAGGAAGTGGCTGGAAAGCCATAGGCACTCA
 CTCTCTGTAAAGACCAAGGACCGGAGAACACATGCTGACCCCTGCTTTTGCAGAGGGGCGATGCTTTCAGGACAGCCGCT
 CAGCAGGTGTCATCTTATTTACACACTTTGTGTTATATCACTTATTTTGCATTTATGTTCTAAATTAACAATATGACG
 TGGCCAGGGCAGTGGCTCAAGCTCTAA TCCAGCAGCTTTGGGAGCCGAGGCAAGTGTATCACTTGGAGGCGAGGATTC
 GGGACCGGCTTGGGCAACATAGCAAAAACCCATTTGCTAATAAAAAATACAAAAATTAAGCCAGCA TGGTGGCGGGCACTTCG
 AGTCCAGCTACTCCGGAAGCTGAAGCAGGAGAAATCACTTGAACCCAGGAGGCGGAGGTGGCAGTGAAGTATCAAGCCAT
 ACACCTCAGCTGGGCAACAGAAAGAGCTGTCTCAAAAAAAATTAATCCGAGCAGAAATTAATGTTGGTCAAGCCCAAGC
 AGTCCCCCCTCACTGACCCCTCTGCTCCCTACAGCTCCAGGCACTCCCCAGCCCTCCCTGGACAAAGGTAATGGCCAGA
 GGGTGAATAATCCACCAAGGT TAAGCCAGAAAACAAAAGCTCAAAGCTTGGGCACTTCCCTCCGCTCAGACCTTAGAGCAG
 ATTCCTCTCATGACAGCACGATCAGGCTGTG

SEQ ID: 64

>STAR64

AGAGATCTTTAAGGGCTCAAAGACCTGCGGCTCCCTGCCAATAGCTCTGCCATGCTCCAGAGCTTTTCGAGGACCC
 TCCACCA TCGGGCCCAACCC CAGCTGAGCTGGGTGCTGGTCTGAGGCTCTGCTCACTCTCAGCTGAGCATGAGGCTCT
 GCTGTGCTGCTTCCAGCAGCAGGGACAGGGCTGATGAGCCTGGCC TTGCAAGCATCTTCTGTGCGAAATACAATTCAC
 AGCAGAGGATTTAAAAATCCAGTGGAGGTGACAGGAAAGAAAGGAAACCTTCCAGGTA TCAAGAAAGGAGGGGGTGTG
 AAGACAGTATGGGAGGAAAG TCAAGCTGGGGCTCAGCTCTGGGAAG TCCAGCC TGAACAGGAGTCAAGCCCGGCTCCCA
 TGCAGGGAAATGAGGACCGAGCCCTGCA TGTGGCAGGGCTTCCGAGGCTGCCCGCTCTGTGAAACAGGACACCAAGAA
 AGTCTGCTTCCAGCTTGGCAAGTGGCAAGGAACCTCTGGTGGGAAACAAA TCAACAAAACAAAATGTCAGTAAAAAAT
 AGAAACCTCACTCTTCC TTTCTCTTGAACCTTTGAAAAAGCAAA TCCACTGCAGCTCAACAAAAGGCAAGAGAAAACT
 TAAGAAATCCAGAGAAAAGACAGCTTACTTTGCAAAAAGAACATCTAATGCAAGGAGATAATGAAAAATACAGACTCTCA
 AAGGGCTGAAGGAAAAAAAACCGTCCACCTAGAAATCTA TCCCAAACTGTCTCTGAGAGCAAGGGCAAAAACCACTCTT
 CTCAGACAGGCTGGAGAGGTCGCTCAAGCTGTAATCTTAGCACTTTGGGAGGCAAGGTGGGAGGACCGCTTTAAGCCA
 GAAATTTGAGACAGCTGTGGTAAACAATAAGAGCCCACTCTTAAGAAAAGAAA TTAATAAGACAAAGCTTTTTCAGA
 CAACAAAGTCTCTGAGAGCTGGCCTA TCTTGGCTGTCTTGTAAAGAAATGCTGCGAGACACCTCA TTAGGAAAGGACGTA
 ATCTAGAAAGAAAGACAGACATGAGGTACAAATGAGGAGCAATAAAGAGGTACCAATAAAGCAAAACCAAAATACACAT

FIG. 20, cont.

TCACTATACGAAACAATAAAAA TGACTCATTGGGGGGTTAAACACTGTTGAAC TAAAACTCCTGGATAACAGCAGCATGA
AAGGTGGGGTGGTGGTCCCAGGAAAGCAATCAAAGGTCCATGTCCTCATTTGGGAGGAGGGTAGGGAGACTCATGAACTTGA
GGCTCCCTTCAGGCAAGCAGTGCATAAAAAATTAATAATGGGAAACAGATACAGTAGACTGTGATGTACAACTCTCAG
AGCAGTAGAAGGGAGGGTATAAAAAAACTCTGATCCA

SEQ ID: 65

>STAR65

TCGAGACCAGCCTGGCCAACATGGTGAACATTTGTCCTATTACAAATACAAAAATTAGCCAGGTGTAGTGGTGCAATGCCT
GCAGTCCCAGCCATTTGGGAGGCTGAAGTTGGAGAAATCGCTTAAACCTGGGAGGTGGAGGTTGCATTTGAGCCGAGAAGCAC
TCCAGCCTGGATGACGGAGCAAGACTGTCTCAAAAAGAAAAAAGAAAGCAGCAGCAAAATATCCCTGTCTGTAGGGAGG
CTATATAACAACCAACAAGTGAAATGCAATAGACAAATTCAGGTTATGGTAGATACCAATAGTGGGAGATGAACAATGAG
AACACATGGACACAGGGAGGAGAACATCACACACTGGGGCCTCTCGGGGGTGGGGAAA TAGGGGGTGTAGCAATAGGAG
AAATACCTAATGTGATAACAGGTTAGTGGGTGCAGCAAAACCAATGGCAGGTGTATATCTATGTAAACACACCTGCACGT
TCTGCACATGTATCCAGAACTTAAAGTATAATAAAAAAGACATTAAAAAATTATGATATAAAATCCCAATTCAGTGT
TTTAAAAAGAGAAAAACAATTATCTTTTATATAATAGCCGAAAAATAGATGGCCGAAATTAAGCCCTCGTCAATTTTCTAAC
AGAACTTCTGTATAAATGTGATTAATAAAAAATTTAAATATCACTAACACATAGAAGAAAATAAATTTAAACCTTCACAA
AAAAATAAGTACAAATGAATGAAGACAAGGTGTACTTTGAAAAAGAACTGAATAAAATTTCTACATAAAAAAAAATCTGAT
GATATTTGGTGTATCTTACTTTGCTACTAGTTCTCTTTTCTCTGAAAAATTTCTTGGGATGTATTTGGTTTCAT
TAGTAAAAATCTAAGTTCTTTGCAATCTGAACATTTGGAGCTTCATCCATAGCCAGTATGCCCTAACATTAATCTTTGGACA
ACTGTAAAAATTAGAACACTGCCAGACATATTAATGTATGATGTATATCAACACTGGGACACATTTTATACTATCTTTAT
CCAAAATCAAAATGATTCACTGTGGTTTATAAATGTACATGGATATATCTCTACCTAAGCAGATAGTTAGGAGAGTTAGTAA
AAATGAGGTGGAAAAATAGGAGTCACTGTCCCTTCACAGGGAGAGAAATTCGCTTTCTCTAAATAACCTTTGCTTGAAC
AGACTCCAACCCCTCATCTTTTGTCTTTAAATGACCACATTTATTTTAACTTTGATAAACAAACACAGAAAGATATTTGAT
CCATCAACATTCAC

ES 2 344 071 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 04 DIV
- <150> EP 01202581.3
<151> 04-07-2001
- 15 <150> US 60/303.199
<151> 05-07-2001
- 20 <160> 122
- <170> PatentIn versión 3.1
- 25 <210> 1
<211> 749
<212> ADN
<213> Secuencia Artificial
- 30 <220>
<223> secuencia de STAR1
- 35 <400> 1
- | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|------|-----|-----|------|-------|-----|------|-----|-----|------|--------|------|------|------|-----|-------|-------|-----|-----|------|----|-----|-----|
| atg | cgg | tggg | ggc | gcg | ccag | agact | cgt | gg | gat | cct | tggc | ttggat | gttt | ggat | cttt | ct | 60 | | | | | | | |
| g | agt | tg | ccg | cg | gaaa | gac | agg | taca | ttt | ctg | atta | ggc | ctg | tga | gc | ctc | ctgga | 120 | | | | | | |
| g | g | acc | at | ctc | atta | ag | ac | ga | tg | g | tatt | gga | ggg | ag | agt | ca | cagaa | agaac | tgt | gg | cccc | ct | 180 | |
| cc | ct | act | gc | aaa | ac | gga | ag | tg | att | tt | tatt | tta | at | ggg | ag | ttg | ga | at | tg | agg | gg | ct | gc | 240 |
| ag | ga | acc | ag | ct | cc | ct | ct | ctt | tg | gt | tga | aa | ag | ct | ggg | ct | gg | c | ct | ag | ag | g | ttt | 300 |
| ttt | gg | cccc | g | ct | gg | g | ct | gg | c | ag | t | ct | ag | tc | gacc | ct | tt | gt | ag | act | gt | g | ca | 360 |
| ag | ag | ca | act | a | ccc | ct | ata | ca | cc | agg | ct | gg | c | ta | ag | t | g | aaa | ggg | g | ct | ct | gg | 420 |
| gg | aaa | at | ct | g | tg | tc | ct | gg | g | ac | ct | ct | gg | ctt | gt | ct | tc | tc | cc | ct | g | ca | ct | 480 |
| gg | g | t | g | ct | t | at | ct | ct | g | ca | g | a | a | cc | ca | ct | ag | gc | gc | ct | g | ta | g | 540 |
| aa | ca | ca | gc | ca | aaa | ag | cc | ct | ag | ag | at | ca | aaa | ag | ca | tt | ag | ta | tg | gg | ca | gt | tg | 600 |
| ga | at | at | tt | ta | cg | ct | tt | gt | tt | cat | ca | ata | ac | tc | gt | tg | g | ct | tt | tg | ac | ct | gt | 660 |
| ag | ca | ata | agg | tg | aa | t | g | ca | g | tc | ac | ag | cg | ct | ta | aca | aa | ta | tg | aa | at | gt | ta | 720 |
| cc | gg | t | ct | cca | gcc | gg | c | gc | gc | c | agg | ct | ccc | | | | | | | | | | | 749 |
- 60 <210> 2
<211> 883
<212> ADN
<213> Secuencia artificial
- 65 <220>
<223> secuencia de STAR2

ES 2 344 071 T3

<400> 2

	gggtgcttcc tgaattcttc cctgagaagg atggtggccg gtaaggctcc tgtaggtggg	60
5	gtgCGGctcc ccaggccccg gcccgtggtg gtggccgctg cccagcggcc cggcaccccc	120
	atagtccatg gcgcccgagg cagcgtgggg gaggtgagtt agaccaaaga gggctggccc	180
10	ggagttgctc atgggctcca catagctgcc cccacgaag acggggcttc cctgtatgtg	240
	tggggctcca tagctgccgt tgccctgcag gccatgagcg tgcgggtcat agtcgggggt	300
	gccccctgcg cccgcccctg ccgccgtgta gcgcttctgt gggggtggcg ggggtgcgca	360
15	gctgggcagg gacgcagggt aggagcggg gggcagccc taggtaccct gggggggctt	420
	ggagaagggc gggggcgact ggggctcata cgggacgctg ttgaccagcg aatgcataga	480
	gttcagatag ccaccggctc cggggggcac ggggctgcga cttggagact ggcccccgga	540
20	tgacgttagc atgccctgc cttctgatc cttttgtac ttcattcggc gattctggaa	600
	ccagatcttg atctggcgct cagtgaggtt cagcagattg gccatctcca cccggcgcgg	660
	ccggcacagg tagcggttga agtggaaact tttctccagc tccaccagct gcgcgctcgt	720
25	gtaggccgtg cgcgcgcgt tggacgaagc ctgccccggc gggctcttgt cgccagcgca	780
	gctttcgctt gcgaggacag agagaggaag agcggcgta ggggctgccg cggccccgcc	840
	cagcccctga cccagcccgg cccctccttc caccaggccc caa	883

30 <210> 3

<211> 2126

<212> ADN

<213> Secuencia artificial

35

<220>

<223> secuencia de STAR3

40 <400> 3

	atctcgagta ctgaaatagg agtaaacttg aagagcaaat aagatgagcc agaaaacat	60
	gaaaagaaca gggactacca gttgattcca caaggacatt cccaagggtga gaaggccata	120
45	tacctccact acctgaacca attctctgta tgcagattta gcaaggttat aaggtagcaa	180
	aagattagac ccaagaaaat agagaacttc caatccagta aaaatcatag caaatattt	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 gttcgcctcg aacgtagtcc ctggaaattt gatgtccagt	480
55	atagaaaagc agagcagtca aaaaatatag ataaagctga accagatggt gcctgggcaa	540
	tgtagcagc accacactta agatataacc tcaggctgtg gactccctcc ctggggagcg	600
	gtgtgcccg cggcggggcg gctccgcaac tccccggctc tctgccccgc cctcccgctc	660
60	tcctcgggcg gcggcggggg ccgggactgc gccgctcaca gcggcggctc ttctgcgccc	720
	ggcctcggag gcagtgccgg tggcggccat ggcctcctgc gttcggccat gtcagatctt	780
	cgactgagg gtcattctct tgggactggt tagacagtgg gtgcagccca cggaggcgca	840
65	gttgaagcag ggtgggggtg cacctcccc aggaagtcca gtgggtcagg gaactccctc	900

ES 2 344 071 T3

ccctagccaa gggaggccgt gagggactgt gcccggtag agactgtgcc ctgaggaaag 960
 gtgcaactctg gcccagatac tacacttttc ccacggctct 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 ctcctctgct gcgagtagac ccaatccttt 1320
 cctagaggtg gagggggcgg gtaggtggaa gtagaggtag cgcggtatct aggagagaga 1380
 15 aaaagggtcg gaccaatagg tgcccggaag aggcggacc agcggctctg tgattggtat 1440
 tggcagtgga cctcccccg ggggtggtgcc ggaggggggg atgatgggtc gaggggtgtg 1500
 tttatgtgga agcgagatga ccggcaggaa cctgccccaa tgggctgcag agtggttagt 1560
 20 gagtgggtga cagacagacc cgtaggccaa cgggtggcct taagtgtctt tggctcctc 1620
 caatggagca gcggcggggc gggaccgcga ctcgggttta atgagactcc attgggtgt 1680
 aatcagtgtc atgtcggatt catgtcaacg acaacaacag ggggacacaa aatggcggcg 1740
 25 gcttagtcct acccctggcg gcggcggcag cgggtggcga ggcgacggca ctcctccagg 1800
 cggcagccgc agtttctcag gcagcggcag cggccccggc aggcgcggtg gcggtggcgc 1860
 gcagccaggt ctgtcaccca ccccgcgcgt tcccaggggg aggagactgg gcgggagggg 1920
 30 ggaacagacg gggggggatt caggggcttg cgacgccct cccacaggcc tctgcgcgag 1980
 ggtcaccgcg gggccgctcg gggtcaggct gcccctgagc gtgacggtag ggggcggggg 2040
 aaaggggagg agggacaggc cccgcccctc ggcagggcct ctagggcaag ggggcggggc 2100
 35 tcgaggagcg gagggggggc gggcgg 2126

<210> 4

<211> 1625

40 <212> ADN

<213> Secuencia artificial

<220>

45 <223> secuencia de STAR4

<400> 4

gatctgagtc atgttttaag gggaggattc ttttggctgc tgagttgaga ttaggttgag 60
 ggtagtgaag gtaaaggcag tgagaccacg taggggtcat tgcagtaatc caggctggag 120
 atgatggtgg ttcagttgga atagcagtgc atgtgctgta acaacctcag ctgggaagca 180
 55 gtatatgtgg cgttatgacc tcagctggaa cagcaatgca tgtggtggtg taatgacccc 240
 agctgggtag ggtgcatgtg gtgtaacgac ctcagctggg tagcagtgtg tgtgatgtaa 300
 caacctcagc tgggtagcag tgtacttgat aaaatgttgg catactctag atttgttatg 360
 60 agggtagtgc cattaattt ctccacaaat tggttgtcac gtatgagtga aaagaggaag 420
 tgatggaaga cttcagtgtc tttggcctga ataaatagaa gacgtcattt ccagttaatg 480

65

ES 2 344 071 T3

gagacagggg agactaaagg taggggtggg ttcagtagag caggtgttca gttttgaata 540
 tgatgaactc tgagagagga aaaacttttt ctaccttta gtttttgta ctggacttaa 600
 5 gaattaaagt gacataagac agagtaacaa gacaaaaata tgcgaggta tttaatattt 660
 ttacttgagc aggggaatct tcaaaagaaa aatgaagacc caaagaagcc attagggta 720
 aaagctcata tgccttttta agtagaaaat gataaatttt aacaatgtga gaagacaaag 780
 10 gtgtttgagc tgagggcaat aaattgtggg acagtgatta agaaatataat gggggaaatg 840
 aaatgataag ttattttagt agatttattc ttcatatcta ttttgcttc aacttccagt 900
 ctctagtgat aagaatgttc ttctcttctt ggtacagaga gagcaccttt ctcatgggaa 960
 15 attttatgac cttgctgtaa gtagaaaggg gaagatcgat ctctgtttc ccagcatcag 1020
 gatgcaaaaca tttccctcca ttccagttct caaccccatg gctgggcctc atggcattcc 1080
 agcatcgcta tgagtgcacc tttctgcag gctgcctcgg gtagctgggtg cactgctagg 1140
 20 tcagtctatg tgaccaggag ctgggcctct gggcaatgcc agttggcagc ccccatcctt 1200
 ccactgctgg gggcctccta tccagaaggg cttggtgtgc agaacgatgg tgcacatca 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 gagggcctg 1500
 30 ggatggcgct cagtgcctgg tcggctgcaa gaggccagcc ctctgcccatt gaggggagct 1560
 ggcagtgacc aagctgcact gccctggtgg tgcatcttct gcccactct ttccttctaa 1620
 gatcc 1625

35 <210> 5
 <211> 1571
 <212> ADN
 40 <213> Secuencia Artificial
 <220>
 <223> secuencia de STAR5
 45 <400> 5

agcagagatc ttatttcccg tattcccttg tggcacagca cctcccacgc caaagcaaac 60
 50 caaagcaaaag gagcccttga tgaggagggg ccttcccca acctggtctc ccacaggtcc 120
 tacatacgta cccaccccag acacacagag ctgcttctg ctctcacacc agactgagct 180
 gtgccagac atttccccta gcactaacca actcttcaa aaatacattt ttctctaaaa 240
 55 agaacaagt taaacaagt tgactcattt taagaactgt ttagaagata acctgtgtt 300
 tattaattat gtatttgagc aaattggagg cagaaggta ccaacattgc ctggtgtcca 360
 gccaggagg agagcgtggg ggcattcaga accttctcc aactcctgcc tggcgtgggt 420
 60 tttattcatc tttgtattcc caagaaactt ctcaagtgtc caggagtgtt aggactcag 480
 tacgtgtttg gtagttacat gaatgaatgc ataatgacta agtgagttaa tggatgaagc 540

65

ES 2 344 071 T3

taattgtctc tcccttttgc ttttccagag ctttccaagg tgaaagtgtt ggacactctt 600
 tcttcatctc agatttaate aactaagaat gctgcaaatt gaacaccagt ccacaaaact 660
 5 caggaataca tgaaaagcat tgtgccttat ttttaactaa ctcaaattct atgtcagctc 720
 cccttttatg ctggatggtg gcgctaaatc tcagtgggtt cctcattctg ccagacctgt 780
 gtccagtttg ggggcttcac atagagccac cccatcacag gagaggggag ggtcttgctc 840
 10 ttggttgcca tcaactccacc ctcttgctct cagagctttg atgttcactt tccttttcac 900
 cactcggaag ctctctgcca tgatacattg agacctcaat gttaatgcca attgggggtt 960
 ggggttctca taaactcaga agtccaggaa aatcgcttgc tgcctccac aacactctga 1020
 15 gggcattctg gaatctacc acttacctgg agcctgctgg cctcaactgt tttgaagtct 1080
 gtgtctgggc catgcaggta aatgggagga tgttctgtgg ccataaaaat acccgaagtc 1140
 ccacctaag ttgatgcagg gtcttctgca tttcattgca aaattgttct atcatttcta 1200
 20 tagttttcag cctacagtca ggggccagga ctttgcacc ttggtaaacc tcaatctctt 1260
 ctcttctctg gcttctactc ttttctccct caatcccaa tcaaggccct tgattgtctg 1320
 gaggtaggaa agcctggttc tggtctatga tatagtctac atcatagcct ttgtcatctc 1380
 25 atggattcac tcaacaaccg tgtgtggatg gggccacca atatgtgcca ggagttgagg 1440
 acacgcaggg ttatgatgat gaaatagata aggggcccac actcacggac cctgcaggac 1500
 agtgagctg tggaccagc atgcgagtaa agaccagtg agctcaccag acagatcatt 1560
 30 taaatcaggt g 1571

<210> 6

<211> 1173

35 <212> ADN

<213> Secuencia Artificial

<220>

40 <223> secuencia de STAR6

<400> 6

tgaccacca cagacatccc ctctggcctc ctgagtgggt tcttcagcac agcttcaga 60
 45 gccaaattaa acgttctctc tatgtctata gacaaaaagg gttttgacta aactctgtgt 120
 tttagagagg gagttaaag ctgttaactt tttaggggtg ggcgagagg atgacaaata 180
 50 acaacttctc tgaatgtttt acatttctcc cactgcctc aagaaggttc acaacgaggt 240
 catccatgat aaggagtaag acctcccagc cggactgtcc ctggcccc agaggacct 300
 ccacagagat atgctaactg gacttgaga ctggctcaca ctccagagaa aagcatggag 360
 55 cacgagcgca cagagcaggg ccaaggctcc agggacagaa tgtctaggag ggagattggg 420
 gtgagggtaa tctgatgcaa ttactgtggc agctcaacat tcaaggagg ggaagaaag 480
 aacagctccc tgtcaagtaa gttgtgcagc agagatgta agctccaaa tttgaaactt 540
 60 tggctgctgg aaagttttag ggggagaga taagaagaca taagagactt tgagggttta 600
 ctacacacta gacgctctat gcatttattt atttattatc tcttatttat tactttgtat 660

65

ES 2 344 071 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 ccatagtgga ttgctgcacc tatcaaccg tcatctaggt ttaagcccca	960
10	catgcattag ctatttgtcc tgatgtcttc cctcccctcc ccaaccaga caggccttgg	1020
	tgtgtgatgt tcccctccct gtgtccatgt gttctcactg ttcagctccc acttatgagt	1080
	gagaacgtgt ggtatttggg tttctgttcc tgtgttagtt tgctgaggat gatggcttcc	1140
15	agcttcatcc atgtccctgc aaaggacacg atc	1173

<210> 7

<211> 2101

20 <212> ADN

<213> Secuencia artificial

<220>

25 <223> secuencia de STAR7

<400> 7

	atcatgccag cttagggcag agagtgagac tggacataat aacaataata ataaaaataa	60
	ataaataaaa caattatctg agaggaaaaa tttgattcat aataaagaga ataaaggttt	120
	ttggcgtgtt tgTTTTgttt tcacctaaaga acagctgttc ccctcattgg gttagtttta	180
35	tttgaagca gaaatcatct ccgcatgatt tccaggggtga tggaaaactg aatatgaatc	240
	caccttctgc catctattca cttgtcacat ttaataagac actcatgcct attttagcat	300
	gttttcttcc ctaccaaag agttagtaac atcaagagat taaaataaca caaataagaa	360
40	cattgaagggt attcaaagt tacatacaaa tattaaacac aatattatta taattattcc	420
	tggaaatgac attgcctcta ctctcaagggt aaaggtcatt tttcttgatt taaacttttt	480
	tctcaagttt gaaatctcta agtttcaacc cgtaatctat ttgcaagttt gtgcaaattt	540
45	tagggattga atccatagta attagtgatt tattgtgggtg tagggagaca agtcaaaaaga	600
	atcaggactg ctaggtagat gactaaggaa aggatggttc acgaggtgac ataaagcact	660
	cagaagaaaa aggtcaggaa acggaggaca gaaaaaaacc taagttctgc tgggtgatgc	720
	tgaatTTgtc atcacaaaat ctgcattgtg gaagcttag ctattgagga gattgctcaa	780
50	gtgtagaact gagaacaata ggcagtgaac ccgagagaac atcaagagac tgagagaaaa	840
	tgaaccagac ttccaggtgc tccatgttcc aaccaacatt ttgtattgtc agaaggaatt	900
55	gagaggcaaa aggaaacca ataaaaata aacaggaaa gggcatacat gattaccacc	960
	cTTTTctca ccagctgtc atggaccagc tttctcctag tgctattttc ttggtcactg	1020
	catcactctg ctaacatagt ttcccacta gctctgaggc tgtcccagag ggggaagccag	1080
60	ctgtcatctc cttcttccac actctgttgg aggaaactgt cattagcagc tccctactaa	1140
	acgcatttat gacaaacagg caggagataa ttaactagaa agtgaacaaa ctcaaacttc	1200

65

ES 2 344 071 T3

agagcctctc atttgatga atgcccttgt aaggctctgg gcctatttta atatttataa 1260
 atgtgttatt ttcttctaaa gaaaaccacc aaattgtata agctacagaa tctgcaaac 1320
 5 tgaggctcat ccatgcactc aggatacatt catagcatct ctgagctgga aaatatctta 1380
 aaggctatat atgtcctcca acactgcaag aatctctctg gcagcattct tttaaaatca 1440
 tcatctaaaa gagggaaatc cccagctgtg tttggatttt gctctgtcac ttgtccagtt 1500
 10 tccccatcca taaaagggca acaatatgaa tttcctgata aggtagtgtg taatataaat 1560
 acaaagtgcg tagccacttc cctaagaaaa atatggggtt tctgcttcac agtctagggg 1620
 gaggaaaaaa aagggggggtc agaagtgatt attattatca ttctatattg gaatgttttc 1680
 15 agacataaaa agctcaccac gtcttaggcc agacagatgc attatgaaag ttaagctaag 1740
 tcttcctcat catgagctgc acctatatcc ccattacttc ttctagaact gcataattta 1800
 tttattcttt cttcaaaagt ttgagagagc cattcttgtc ctctaagatt tttttttttt 1860
 20 tttttggaga cagagtctcc gtctgttgcc caggctggag tgcaatggca ctatctcagc 1920
 tcaactgcaac ctctgcctcc cagattcaag tgattctcct gcctcagcct cccgagtagc 1980
 tgggattaca agcacgcacc accacaacca gctaattttt cgtatttttt agtagagacg 2040
 25 aggttttacc atgttggccca ggctggtctt gaactcctga cctcgggtga tccaccacc 2100
 t 2101

<210> 8

30 <211> 1821

<212> ADN

<213> Secuencia artificial

35 <220>

<223> secuencia de STAR8

40 <400> 8

gagatcacct cgaagagagt ctaacgtccg taggaacgct ctcgggttca caaggattga 60
 ccgaacccca ggatacgtcg ctctccatct gaggcttctc ccaaatggcc ctccactatt 120
 45 ccaggcacgt ggggtgtctcc cctaactctc cctgctctcc tgagcccatg ctgcctatca 180
 cccatcggtg caggtccttt ctgaagagct cgggtggatt ctctccatcc cacttccttt 240
 cccaagaaag aagccaccgt tccaagacac ccaatgggac attccccttc cacctccttc 300
 50 tccaagttg cccagggtgtt catcacaggt tagggagaga agccccagg tttcagttac 360
 aaggcatagg acgctggcat gaacacacac acacacacac acacacacac acacacacac 420
 acacgactcg aagaggtagc cacaagggtc attaaacact tgacgactgt tttccaaaaa 480
 55 cgtggatgca gttcatccac gccaaagcca agggtgcaaa gcaaacacgg aatgggtggag 540
 agattccaga ggctcaccaa accctctcag gaatattttc ctgaccctgg gggcagaggt 600
 tggaaacatt gaggacattt cttgggacac acggagaagc tgaccgacca ggcattttcc 660
 60 tttccactgc aatgaccta tggcgggggc atttcacttt cccctgcaaa tcacctatgg 720
 cgaggtagct ccccaagccc ccacccccac ttccgcgaat cggcatggct cggcctctat 780

65

ES 2 344 071 T3

ccgggtgtca ctccaggtag gcttctcaac gctctcggct caaagaagga caatcacagg 840
 tccaagccca aagcccacac ctcttccttt tgttataccc acagaagtta gagaaaacgc 900
 5 cacactttga gacaaattaa gagtccttta ttttaagccgg cggccaaaga gatggctaac 960
 gctcaaaatt ctctgggcc cgaggaagg gcttgactaa cttctatacc ttggtttagg 1020
 aaggggaggg gaactcaaat gcggttaattc tacagaagta aaaacatgca ggaatcaaaa 1080
 10 gaagcaaatg gttatagaga gataaacagt tttaaaaggc aaatggttac aaaaggcaac 1140
 ggtaccaggt gcggggctct aaatccttca tgacacttag atataggtgc tatgctggac 1200
 acgaactcaa ggctttatgt tgttatctct tcgagaaaaa tcctgggaac ttcatgact 1260
 15 gtttgtgcca gtatcttatac 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
 acaaggtag gcatttcact ttcccagaga acgcgcaaac attcaatggg agagaggtcc 1560
 cgagtcgtca aagtcccaga tgtggcgagc ccccgggagg aaaaaccgtg tcttccttag 1620
 25 gatgcccgga acaagagcta ggcttccgga gctaggcagc catctatgtc cgtgagccgg 1680
 cgggagggag accgccggga ggcgaagtgg ggcggggcca tccttcttcc 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 2 344 071 T3

<400> 9

	atgagcccc	aaaaatgatc	ctctggctta	tgacaacctg	atgcagccca	ggaaatgcct	60
5	gcaacatgcc	cactagcagc	tggaacccc	tctgtgagga	agagaacgtt	ttacattaag	120
	aaaccctttg	ttttgcagca	gagactattc	aggtcacaca	tgtgtggcct	ctcagttctt	180
	tgagccattt	gaagttctct	atccttgctg	ggaggtgag	ctctccatgg	aaacctggtc	240
10	cgatagtgag	aggagcagac	cctctggaaa	caccttttta	cacctgacca	aagcagccag	300
	tcatgggcca	gtgatgcaac	aaggtaacc	ggtgcattct	ggcccctcag	aaaagcagcc	360
	cccgggaagg	tcaggaggag	gctgctgact	ccctcttccc	ctgcagccgc	cccaagcaca	420
15	cccaggagcc	ctgcaggttt	gggttcacca	gggtccagca	ggcccacga	tgctgcattt	480
	cttacgagct	cctggaggat	gcagatggtc	ctggtcagag	gctgcattct	gagtatcagg	540
	agccatgggg	caacgtttct	gcgattgagg	aaggggcatt	tctggggtgg	gcagaacaaa	600
20	ggtctttggc	tgagctggag	catccgcctc	catcagtgtt	ttccggcaac	tgactatcc	660
	atcgtcttcc	cttcccacag	ctgaccatgg	ctttggaaaa	tgctctgaaa	ctttcttttc	720
25	agaagagtgt	actcccaact	ccacacttag	gggaagtcaa	gcctacttct	cagaattcag	780
	agaaggcata	aaaaagaatt	catttctaaa	ggccctttag	aagtaacttc	aggtctgaca	840
	gcgccagct	aatttctggt	cgcttccag	gaatcttctg	actgcaaaaa	aaaagcattt	900
30	accacctgaa	cacaaacca	gttacagata	gaaaaacata	gtcatttaaa	tagaatataa	960
	gcatctggcc	tctgcccata	ataatggagt	aacacaaaaa	tctattttca	aaaggaaact	1020
	aaatattatt	gaccaaaaca	tgaatgggga	gacctcaggg	tgatacagct	cttgccctgga	1080
35	tggaatttgt	aatcaagagg	atgagacagg	attgtaactt	gtgccaatgt	gaaagggttt	1140
	gctcaggtat	cattcatttt	gcttaaattgc	atgggtaatt	tccaaagttc	tttgagctg	1200
	aatttcacaa	tttagtgag	gtcctgggta	gcccaccttg	acttatctca	cagtacaatg	1260
40	cagtggcgtg	gctacaatgc	tgggcaagag	aagccaatgt	caacagccca	ggagtggctg	1320
	ggtccttacc	aggctcccag	gcatgcttca	tgggtggccc	tgggctggga	ggaacagcac	1380
	ctttgcctgg	tccatgagta	tctgggtcaa	actctcctgt	ggacacagaa	ggccatggcg	1440
45	acaggcattc	ccaggaaaag	aaaagggcag	cagctgaaat	cgtcaggtgg	agaaggcagt	1500
	catccttgct	cagtcaactc	taatccggct	gcctcctcct	cagcttcagg	gtgaacctct	1560
	cctaagctgt	gtctttggta	tctgatgggc	attaggtgct	ggtgaaaaag	ctggagggtc	1620
50	ctttgggata	ttacagaagc	ccaatctagc	cttgatttca	atatctaggc	actctcacc	1680
	ctgaagttct	acgtttccag	atctctgaaa	acatgggaaa	gcatgtgtgt	gatgtctgag	1740
	gtccccctca	gcctctggtg	tagggttagg	agggctctaa	aggggtggcag	ctccagtgtc	1800
55	ccagtggggc	ctgaagttgg	tcccttccct	tcccagctcc	catccatggt	ttagcccaat	1860
	cccttccgta	cctaagagta	ctgcacatgg	atgctccacg	cagagcctct	gctccactcc	1920
	caggaagtg						1929

60

<210> 10

<211> 1167

<212> ADN

65

<213> Secuencia artificial

<220>

ES 2 344 071 T3

<223> secuencia de STAR10

<220>

5 <221> característica_misc

<222> (452)..(1143)

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

10 <400> 10

	aggtcaggag ttcaagacca gcctggccaa catggtgaaa ccctgtccct acaaaaaata	60
	caaaaattag cggggcgtgg tggggggcgc ctataatccc agctactcag gatgctgaga	120
15	caggagaatt gtttgaaccc gggagggtgga ggttgcagtg aactgagatc gcgccactgc	180
	actccagcct ggtgacagag agagactccg tctcaacaac agacaaacaa acaacaacac	240
20	aacaacaaaa atgtttactg acagctttat tgagataaaa ttcacatgcc ataaaggcca	300
	ccttctacag tatacaattc agtggattta gtatgttcac aaagttgtac gttgttcacc	360
	atctactcca gaacatttac atcacccta aaagaagctc ttagcagtc acttctcatt	420
25	ctccccagcc cctgccaacc acgaatctac tntctgtctc tattctgaat atttcatata	480
	aaggagtccct atcatatggg ccttttactg ctaccttctt tcaacttagca tcatgttttt	540
	aagattcatc cacagtgtag cacgtgtcag ttaattcatt tcatcttatg gctggataat	600
30	gctctattgt atgcatatcc ctcaacttgc ttatccattc atcaactgat tgacatttgg	660
	gttatttcta ctttttgact attatgagta atgctgctat gaacattcct gtaccaatcg	720
	ttacgtggac atatgctttc aattctcctg agtatgtaac tagggttgga gttgctgggt	780
35	catatgttaa ctcaacttgc catttttttg aagaactacc aaatggtttt ccaagtggga	840
	tgcaacactt tacattccca ccagcaagat atgaaggttc caatgtctct acatttttgc	900
	caacacttgt gattttcttt tattttattha tttatttatt tatttttgag atggagtctc	960
40	actctgtcac ccaggctgga gtgcagtggc acaatttcag ctcaactgca tctccacctc	1020
	tcgggctcaa gcgatactcc tgcctcaacc tcccgagtaa ctgggattac aggcgcccac	1080
	caccacacca agctaatttt ttgtattttt agtagagacg gggtttcatc atgtcggcca	1140
45	ggntgtactc gaactctgac ctcaagt	1167

<210> 11

<211> 1377

50 <212> ADN

<213> Secuencia artificial

<220>

55 <223> secuencia de STAR11

60

65

ES 2 344 071 T3

<400> 11

	⁻⁻⁻⁻ ⁻⁻⁻	gattctgggt gggtttgatg atctgagagt cccttgaata aaaagaattc tagaaaagct	60
5		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 tctaaaacct	300
		ttcttccagt taggtgttaa gattttaaca gcctttccta tctgaataaa aactgcacac	360
		aaagtaaact taagagatgt caacaactca tctgtttgtt acaagatgag tctccatgct	420
15		tcatcgctg tggggaatcc tcatcagcgt ctagtggcaa agactcctgt gtgctcaccg	480
		aaacgctccc cttcctccag ggcacacagt cacatggatt tcccatgcac cctggcagct	540
		cagcaggagt ccatgactta agaaggccaa tggactgtgg gtgaagtctg tggacgggga	600
20		agccacatgc gtcacttcca ggcctgggcg tgtgcacacct cactctctt cccctgtggg	660
		tgcagaaggc ggggcagagg gccctgaaac cttggaggtc ggtggagccc aaaatgaagg	720
		agcgtgggcc tctgggtctt catgtaaatt taggtaacac tgaactgtca ggtgaacaag	780
25		aaataaacgt caaatgtatt cagtcgatta gatttgggtga tggttgttac agcggttacc	840
		ctccctcaac ataataaatt ttcaacaac tcataatggc tcactcatgt ataaaatatt	900
		ccatatgaaa tcccgggata acatgcttat tctagctcaa gcttaatcag agtagtccat	960
30		ctgagggagg agatagtaga gggcagcaag gggttgtcac tgaagataac tagccttgct	1020
		aaaagaatgg ttgaagaagt gagctacaga tagggtaaatt ccacatctca gacattctgt	1080
35		gatggtcctg atattatcct aaagtaaaat gtagagttga accattttaa ttagattcta	1140
		gaattctatt aatttataag atgggcattt ccacaaagga ctaaacaag tacaagagga	1200
		ttaaataatc atccacatgg gaggcaccgc cttgcacttt aaaatgatgg agcttatcaa	1260
40		gactggctgt ggatatctgt ccctgggagg gttttttccc ccattttttt cctttttgag	1320
		acatgttctc gctatgttgc ccaggctggc cttgaactcc tgggctcaag tgatcct	1377

<210> 12

45 <211> 1051

<212> ADN

<213> Secuencia artificial

50 <220>

<223> secuencia de STAR12

55

60

65

ES 2 344 071 T3

<400> 12

	atcctgcttc tgggaagaga gtggcctccc ttgtgcaggt gactttggca ggaccagcag	60
5	aaaccaggt 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 ttaaatggag	360
	tccaacttcc ttttctcacc cagacgtcga gaagggtggca ttcaaatgt ttacacttgt	420
15	ttcatctgcc tttttgctaa gtcctgggcc cctacctctt ttcctcact tcacatttgt	480
	cgtttcatcg cacacatatg ctcacttta tatttacata tatataattt ttatatatgg	540
	cttgtgaaat atgccagacg agggatgaaa tagtcctgaa aacagctgga aaattatgca	600
20	acagtgggga gattgggcac atgtacattc tgtactgcaa agttgcacaa cagaccaagt	660
	ttgttataag tgaggctggg tggtttttat ttttctcta ggacaacagc ttgcctggtg	720
	gagtaggcct cctgcagaag gcattttctt aggagcctca acttcccaa gaagaggaga	780
25	gggcgagact ggagttgtgc tggcagcaca gagacaaggg ggcacggcag gactgcagcc	840
	tgacagaggg ctggagaagc ggaggctggc acccagtggc cagcagggcc caggtccaag	900
	tccagcgagg tcgaggtcta gactacagca agccaaggt ccaaggtcag tgagtctaag	960
30	gtccatggtc agtgaggctg agaccaggg tccaatgagg ccaaggtcca gactccagta	1020
	aggccgagat ccaggggtcca gggaggtcaa g	1051

<210> 13

35 <211> 1291

<212> ADN

<213> Secuencia Artificial

40 <220>

<223> secuencia de STAR13

45

50

55

60

65

ES 2 344 071 T3

<400> 13

5 ctgccctgat cccttaatgc ttttggccca gagcaccg ctaagtccaa ccccagaggg 60
 gcctcatccg caaagcctcg ggaagaggac agtgacggag gcggtgccc tgtgagctgc 120
 acggggcaga atgtcctttt ggcgtcatgt tggatgtcca cacatccata tggggtcagt 180
 tctattagga ttccttcggg aagaggtaga gggtaggagg ggttaagcca cgagacgagg 240
 10 catgcagagg ggtggcctgg atgggtctgc actgctgtcc atgcacacgg ggagcgttgc 300
 aaattgtgct tcccagccca tagtgccccc acagaggagc ccgggagtcc ctggtgggagc 360
 tctgtgttcc tgcaaggagc cagtggagat ggccccgtga acttcatcc cccttgccct 420
 15 ggtggggtct ctggcaggtt tatggagccg tacatctttg ggagccgcct ggaccacgac 480
 atcatcgacc tggaacagac agccacgac ctccagctgg ccttgaactt caccgcccac 540
 atggcctacc gcaagggcat catcttgttt ataagccgca accggcagtt ctcgtacctg 600
 20 attgagaaca tggcccgtga ctgtggcgag tacgcccaca ctcgctactt caggggaggc 660
 atgctgacca acgcgcgct cctctttggc cccacggctc gcctgcccga cctcatcatc 720
 ttcctgcaca cgctcaaca catctttgag ccacacgtgg ccgtgagaga cgcagccaag 780
 25 atgaacatcc ccacagtggg catcgtggac accaactgca acccctgcct catcacctac 840
 cctgtaccgg gcaatgacga ctctccgctg gctgtgcacc tctactgag gctcttccag 900
 acggccatca cccgggcca ggagaagcgg cagcaggttg aggtctcta tcgctgagc 960
 30 ggccagaagg agcccgggga ccaggggcca gccaccctc ctggggctga catgagccat 1020
 tccctgtgat gttcactctc ctccaaagc aaaccacagc caagcctgtc tgagctggga 1080
 gtccccttcc ccagccctgg gtcagcggca tcctcagtcg ttgttactta ctcagctgat 1140
 35 gtcacagtgc agacatccac cgttccacca cagaaccagt ggctgagcgg accaacgtt 1200
 ccatgtgctg ttgctctgtg ggaacagag cacagagggt gagcgacatg tgcagaacgg 1260
 ccccttgct gcagttagga cctcagtgcc t 1291

40 <210> 14

<211> 711

<212> ADN

45 <213> Secuencia artificial

<220>

<223> secuencia de STAR14

50 <400> 14

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

60

65

ES 2 344 071 T3

5 **aatcccacc aagagccggc tacaagggtg gtgcaccggc agggcctgat gggcgtccag 240**
 aaggagtctg ccttctgcag acctccccg gttagcaaag acaaactgtc agcccagaag 300
 cctcaggtcc aaggctacga ccacgaccac gcaggagaag gagggcagca caaagccagc 360
 aacccccacc gcccttcaga agcagtacag aataaaagtg ggattaaaag gaacgccagc 420
 accgcaagga ggggaaagcg agtcacgagc gccgtacagg cgcccaggc gtccgactcc 480
 10 **agcagcgacg acggcattga ggaggccatc cagctgtacc aggtgcagaa aacacacaag 540**
 gaggccgacg gggacccgcc ccagagggtc cagctccaag aggaaagagc acctgccctt 600
 cccgcacaca gcacaagcag cgccacaaaa agtgccttgc cagagacca caggaaaaca 660
 15 **cccagcaaga agaagccagt gccaccaag accacggacc ctggtccagg g 711**

<210> 15

<211> 1876

20 <212> ADN

<213> Secuencia artificial

<220>

25 <223> secuencia de STAR16

<400> 15

30 **cagtacatgc agaactgagt ccaaacgaga cggacagcaa acccggcagt gggctcccag 60**
 acattcctgg gggaaagggg tcctaaccac aggcagttaa agtcatctcc tccaaccctc 120
 tatgacacag gctgtgcgct gtcatttaaa agctgagtga aatttaacce ttttccatt 180
 tagaaaaaca aagcgcagct ggctgccagc actcatttaa tttacataa acgtgctctt 240
 35 **tgaggctgaa gcaaatctga ctgattttca atgtgaaaat aaaatgtaa aactgttctt 300**
 ggaattattt ctaaacagaa catcagaatc gtctgaaatc tcagaatcgg ctattttgga 360
 aaaatcggat tcatcaaagc aatcttcggc caacaactgt tagagaacga tgtaacacc 420
 40 **acgcatagga atgttacatt ttctagaatt tgacattttc attgacggaa aattactgta 480**
 tcttgatatat ggaaatacca ctactaaaaa cataatgcta taaatagaat gatgtctttt 540
 gtttccaaag tcaatatact cgagcaatgc aaaaaataa ataaaagtga gatacttcat 600
 45 **ggcaaagctg ccgcaggata aacattgcag ccacaagtgc cccagattt ctcggggcaa 660**
 actggaaaag ggctaacagg caacattttc atgttattct actgagtgca gtaattattt 720
 ttaaaaatat acatgaataa tgaaaaaact gtggtatggt tttaaagaaa tttccataac 780
 50 **ctggtgaaac tcttcacaca gggtaatagg ttcataaagc cttggtcctc tgcaaaacaa 840**
 gcatcaactt gacaatgact aaaagaagca acagcaaac tgtcacgcat ttggagccat 900
 ggcctgggtt gggccggtgt aaagctctcc gccctctgga gcaagtctgg gccccagcgg 960
 55 **ctggcatgtg ggcaactgcag ggccctgggtt gggcagggtg gcagctctcc gtcactctgag 1020**
 cctagtctga ggccctggtg ctggcacgtg ggccctgcag ggcctctact tctcacccca 1080
 gctccacttc cctccctgcc ctctactgggt ctcaacagagc caatgaacac tggggctcaga 1140
 60 **ttcagggccc agcatccact gcagtgggca ctgcccttcc acaaggcctg gctccaggaa 1200**

65

ES 2 344 071 T3

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

20 <210> 16
 <211> 1282
 <212> ADN
 25 <213> Secuencia artificial

<220>
 <223> secuencia de STAR17

30 <400> 16

cgcccacctc ggctttccaa agtgctggga ttacaggcat gagtcactgc gcccatcctg 60
 35 attccaagtc ttagataat aacttaactt tttcgaccaa ttgccaatca ggcaatcttt 120
 gaatctgcct atgacctagg acatccctct ccctacaagt tgccccgcgt ttccagacca 180
 aaccaatgta catcttacat gtattgattg aagttttaca tctccctaaa acatataaaa 240
 40 ccaagctata gtctgaccac ctcaggcacg tgttctcagg acctccctgg ggctatggca 300
 tgggtcctgg tcctcagatt tggctcagaa taaatctctt caaatatttt ccagaatttt 360
 actcttttca tcaccattac ctatcaccca taagtcagag tttccacaa ccccttctc 420
 agattcagta atttgctaga atggccacca aactcaggaa agtattttac ttacaattac 480
 45 caatttatta tgaagaactc aaatcaggaa tagcceaag gaagaggcat agggaaaggt 540
 atggaggaag gggcacaaaag cttccatgcc ctgtgtgcac accaccctct cagcatcttc 600
 atgtgttcac caactcagaa gctcttcaa ctttgtcatt taggggtttt tatggcagtt 660
 50 ccactatgta ggcatggttg ataaatcact ggtcatcggg gatagaactc tgtctccagc 720
 tcctctctct ctcctcccca gaagtcctga ggtggggctg aaagtttcac aaggttagtt 780
 gctctgacaa ccagccccta tcctgaagct attgaggggt ccccaaaaag ttaccttagt 840
 55 atggttgaa gaggcttatt atgaataaca aaagatgctc ctatttttac cactagggag 900
 catatccaag tcttgcggga acaaagcatg ttactggtag caaattcata caggtagata 960
 gcaatctcaa ttcttgcctt ctcagaagaa agaatttgac caagggggca taaggcagag 1020
 60 tgagggacca agataagttt tagagcagga gtgaaagttt attaaaaagt tttaggcagg 1080
 aatgaaagaa agtaaagtac atttggaaga gggccaagtg ggcgacatga gagagtcaaa 1140
 caccatgcc tgtttgatgt ttggcttggg gtcttatatg atgacatgct tctgagggtt 1200
 65 gcatccttct cccctgattc ttccttggg gtgggctgtc cgcatgcaca atggcctgcc 1260
 agcagtaggg aggggccgca tg 1282

ES 2 344 071 T3

<210> 17

<211> 793

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> secuencia de STAR17

10

<400> 17

```

15      atccgagggg aggaggagaa gaggaaggcg agcagggcgc cggagcccga ggtgtctgcg      60
      agaactgttt taaatggttg gcttgaaaat gtcactagtg ctaagtggct tttcggattg      120
      tcttatttat tactttgtca ggtttcctta aggagagggg gtggtggggg tgggggagga      180
      ggtggactgg ggaaacctct gcgtttctcc tcctcggctg cacaggggta gtaggaaacg      240
20      cctcgtgcc acttaacaat ccctctatta gtaaacttac gcggagactc tatgggaagc      300
      cgagaaccag tgtcttcttc cagggcagaa gtcacctgtt gggaacggcc cccgggtccc      360
      cctgctgggc tttcggctc ttctagggcg cctgatttct cctcagccct ccaccagcg      420
25      tccctcaggg acttttcaca cctccccacc cccatttcca ctacagtctc ccagggcaca      480
      gcacttcatt gacagccaca cgagccttct cgttctcttc tcctctgttc cttctctttc      540
      tcttctctc tgttccttct ctttctctgt cataatttcc ttggtgcttt cgccacctta      600
30      acaaaaaag agaaaaaat aaaataaaaa aaaccattc tgagcaaag tattttaaga      660
      tgaatccaag aaagcgacc acatagccct cccaccac ggagtgcgc aagacgcacc      720
      caggctccat cacagggccg agagcagcgc cactctggtc gtacttttg gtcaagagat      780
35      cttgcaaaag agg      793
  
```

<210> 18

<211> 492

<212> ADN

<213> Secuencia Artificial

45 <220>

<223> secuencia de STAR18

<400> 18

```

50      atctttttgc tctctaaatg tattgatggg ttgtgtttt tttcccact gctaataaat      60
      attacattgc aacattcttc cctcaacttc aaaactgctg aactgaaaca atatgcataa      120
      aagaaaatcc tttgcagaag aaaaaagct attttctccc actgatttg aatggcactt      180
55      gcggatgcag ttcgcaaatc ctattgccta ttcctcatg aacattgtga aatgaaacct      240
      ttggacagtc tgccgcattg cgcagtagac tgcctgcga aggcaagggt atggttccca      300
60      aagcaccag tggtaaatcc taacttatta ttccttaa attccaatgt aacaacgtgg      360
      gccataaag agtttctgaa caaacatgt catcttttg gaaagggtt tttcgtaatt      420
      aatgatgaa tcatgctcat ttcaaatgg aggtccacga tttgtggcca gctgatgcct      480
65      gcaaattatc ct      492
  
```

ES 2 344 071 T3

<210> 19

<211> 1840

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> secuencia de STAR19

10

<400> 19

```

    tcacttcctg atatTTTaca ttcaaggcta gctttatgca tatgcaacct gtgcagttgc      60
15  acagggcttt gtgttcagaa agactagctc ttggTTTaat actctgttgt tgccatcttg      120
    agattcatta taatataatt tttgaatttg tgttttgaac gtgatgtcca atgggacaat      180
    ggaacattca cataacagag gagacaggtc aggtggcagc ctcaattcct tgccaccctt      240
20  ttcacataca gatttgcaa tgcccatga gcacaaaatt tgggggaacc atgatgctaa      300
    gactcaaagc acatataaac atgttacctc tgtgactaaa agaagtggag gtgctgacag      360
    cccccagagg ccacagTTta tgttcaaacc aaaacttgct tagggTgcag aaagaaggca      420
25  atggcagggT ctaagaaaca gcccatacata tccttgttta tTcatgttac gTccctgcat      480
    gaactaatca cttacactga aatattgac agaggaggaa atggaaagat agggcaaccc      540
    atagttcttt tTccttttag tctttcctta tcagtaaacc aaagatagta ttggtaaaat      600
30  gtgtgtgagt taattaatga gttagTTtta ggcagtgttt cactgtttgg ggtaagaaca      660
    aaatatatag gcttgtattg agctattaaa tgtaaattgt ggaatgtcag tgattccaag      720
    tatgaattaa atatccttgt atttgcattt aaaattggca ctgaacaaca aagattaaca      780
35  gTaaaattaa taatgTaaaa gTTtaatttt tacttagaat gacattaaat agcaaataaa      840
    agcaccatga taaatcaaga gagagactgt gGaaagaagg aaaacgTTTT tatttttagta      900
    tatttaatgg gactttcttc ctgatgtttt gTTTTgtttt gagagagagg gatgtggggg      960
40  cagggaggTc tcattttgtt gcccaggctg gacttgaact cctgggctcc agctatcctg      1020
    ccttagcttc ttgagtagct gggactacag gcacacacca cagtgtctga cattttctgg      1080
    atTTTTTTTT tTTTTtatt tTTTTgtga gacaggttct ggctctgtta cTcaggtTgc      1140
45  agTgcagTgg catgatagcg gctcactgca gcctcaacct cctcagctta agctactctc      1200
    ccacttcagc cTcctgagta gccaggacta cagttgtgtg ccaccacacc tgtggctaTt      1260
    ttttTtagag atggggTctc tccacgtTgc cgaggctggT ctccaactcc Tggtctcaag      1320
50  cgaacctcct gactTggcct cccgaagtgc tgggattaca ggctTgagcc actgcatcca      1380
    gcctgtcctc tgtgtTaaac ctactccaat ttgtctttca tctctacata aacggctctt      1440
    ttcaaagTtc ccatagacct cactgtTgtc aatctaataa taaattatct gccttttctt      1500

55  acatggTtca tcagtagcag cattagattg ggctgctcaa tTcttctTgg tatatTTTct      1560
    tcattTggct tctggggcat cacactctct ttgagttact cattcctcat tgatagctTc      1620
60  tTcctagtct tctttactgg tTcttctctc tctccctgac tcttTaatat tgtttttctc      1680
    cccaggcttt agttcttagt cctcttctgt tatctattta cacccaattc tttcagagTc      1740
    tcatccagag tcatgaactt aaacctgttt ctgtgcagat aattcacatt attatatctc      1800
65  cagcccagac tctcccGcaa actgcagact gatcctactg      1840

```

ES 2 344 071 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 gtaaagacca    120
atggaccctt caaacaaaaa cagccaagct tttctgcca aaagatgact gagaagactg    180
ttaaagcaaa aaactctgtt cctgcctcag atgatggcta tccagaaata gaaaaattat    240
20 ttcccttcaa tcctctaggc ttcgagagtt ttgacctgcc tgaagagcac cagattgcac    300
atctcccctt gagtgaagtg cctctcatga tactttgatga ggagagagag cttgaaaagc    360
tgtttcagct gggccccctt tcacctttga agatgccctc tccaccatgg aaatccaatc    420
25 tgttgcagtc tcctttaagc attctgttga ccctggatgt tgaattgcca cctgtttgct    480
ctgacataga tatttaaat tcttagtgct ttagagtttg tgtatatttc tattaataaa    540
gcattatttg tttaacagaa aaaaagatat atacttaaat cctaaaataa aataaccatt    600
30 aaaaggaaaa acaggagtta taactaataa gggacaagag 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 acagcaaatc tctttgtagt agcaaaaggc caaaaggctt atcaacaaga    180
aaattaatac attgtggcac ataatggcat ccttatgcc aataaaatgg atgaaattat    240
55 agttagggtc aaaaggcaag cctccagata atttatatca tataattcca tgtacaacat    300

```

```

tcaacaacia gcaaaactaa acatatacaa atgtcagggg aatgatgaa caaggttaga    360
60 aatgattaa tataaaaaa ctgcacagtg ataacattta atgagaaaaa aagaaggaag    420
ggcttaggga gggacctaca gggaaactcca aagttcatgg taagtactaa atacataatc    480
aaagcactca aatagaaaa tattttagta atgttttagc tagttaatat cttacttaaa    540
65 acaaggctca ggccaggcac ggtggctcac acctgtaatc ccagcacttt gggaggctga    600
ggcgggtt                                     607

```

ES 2 344 071 T3

<210> 22

<211> 1380

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> secuencia de STAR22

10

<400> 22

```

        cccttgatg  ccacccgct  tggcctcca  aagtgctgg  attacaggcg  tgagtcacta  60
15      cgcccgcca  ccctccctgt  atattatttc  taagtatact  attatgtaa  aaaaagtta  120
        aaaatattga  tttaatgaat  tcccagaaac  taggatttta  catgtcacgt  tttcttatta  180
        taaaaataaa  aatcaacaat  aaatatatgg  taaaagtaa  aagaaaaaca  aaaacaaaaa  240
20      gtgaaaaaaa  taaacaacac  tcctgtcaa  aaacaacagt  tgtgataaaa  cttaatgccc  300
        tgaaaattta  gaaacatcct  tctaaagaag  ttctgaataa  aataaggaat  aaaataatca  360
        catagtttg  gtcattggtt  ctgtttatgt  gatggattat  gtttattgat  ttgtgatgt  420
25      tgaacttatc  tcaatagatg  cagacaaggc  cttgataaaa  gtttttaaca  ccttttcagt  480
        ttgaaaactc  tcaatagact  aggtattgat  gaaacatata  tcaaaataat  agaagctatt  540
        tatgataaac  ccatagccaa  tatcatactg  agtgggcaa  agctggaagc  attcccttg  600
30      aaaactggca  caagacaagg  atgccctctc  tcaccactcc  tattaatgt  agtattggaa  660
        gttctggcca  gagcaatcag  gcaggagaaa  gaaaaggat  taaaatagga  agagaggaag  720
        tcaaatgtc  tctgtttgca  gtaaacaatg  ttgtatattt  agaaaacccc  attgtctcat  780
35      cctaaaaact  ccttaagctg  ataaacaact  tcagcaaagt  ctcaggatac  aaaatcaatg  840
        tgcaaaaatc  acaagcattc  ctatacaccg  ataatagaca  gcagagagcc  aaatcatgag  900
        tgaagtccca  ttcacaattg  cttcaaagaa  aataaaatac  ttaggaatac  aactttcacg  960
40      ggacatgaag  gacattttca  aggacaacta  aaaaccactg  ctcaaggaaa  tgagagagga  1020
        cacaagaaa  tggaaaaaca  ttccatgctc  atggaagaat  caatatcatg  aaaatggcca  1080
        tactgcccac  agtaatttat  agattcaatg  ctaaccccat  caagccacca  ttgactttct  1140
45      tcacagaact  agaaaaaac  tattttaaa  ctcatatgta  gtcaaaaaga  gtcggtatag  1200
        ccaagacaat  cctaagcata  aagaacaaag  ctggatgcat  cacgtgact  tcaaaccata  1260
        ctacaaggct  acagtaacca  aacagcatg  gtactggtac  caaacagat  agatagaccg  1320
        atagaacaga  acagaggcct  cggaaataac  accacacatc  tacaaccctt  tgatcttcaa  1380

```

50

<210> 23

<211> 1246

<212> ADN

55 <213> Secuencia Artificial

<220>

<223> secuencia de STAR23

60

65

ES 2 344 071 T3

<400> 23

	atcccctcat ccttcagggc agctgagcag ggcctcgagc agctggggga gcctcactta	60
5	atgctcctgg gagggcagcc agggagcatg gggctctgag gcatggtcca gggctcctgca	120
	ggcggcacgc accatgtgca gccgccccca cctgttgctc tgcctccgcc acctggccat	180
	gggcttcagc agccagccac aaagtctgca gctgctgtac atggacaaga agcccacaag	240
10	cagctagagg accttggtt ccacgtgccc agggagcatg gcccacagcc caaagaccag	300
	tcaggagcag gcaggggctt ctggcaggcc cagctctacc tctgtcttca cacagatggg	360
	agatttctgt tgtgattttg agtgatgtgc ccctttgggtg acatccaaga tagttgctga	420
15	agcaccgctc taacaatgtg tgtgtattct gaaaacgaga acttctttat tctgaaataa	480
	ttgatgcaaa ataaattagt ttggatttga aattctattc atgtaggcat gcacacaaaa	540
	gtccaacatt gcatatgaca caaagaaaag aaaaagcttg cttccttaa atacaatat	600
20	ctgttaacta tatttgcaaa tatatttgaa tacacttcta ttatgttaca tataatatta	660
	tatgtatatg tatatataat atacatata atgttacata taatatactt ctattatggt	720
	acataataa tttatctata agtaaataca taaatataaa gatttgagta gctgtagaac	780
25	attgtcttat gtgttatcag ctactactac aaaaatatct cttccactta tgccagtttg	840
	ccatataaat atgatcttct cattgatggc ccagggcaag agtgcagtggtg gtacttattc	900
	tctgtgagga gggaggagaa aagggaaaca ggagaaagtc acaaaggga aactctgggtg	960
30	ttgccaaaat gtcaagttc acatattccg agacggaaaa tgacatgtcc cacagaagga	1020
	ccctgcccag ctaatgtgtc acagatatct caggaagctt aatgatttt tttaaaagaa	1080
	aagagatggc attgtcactt gtttcttcta gctgaggctg tgggatgatg cagatttctg	1140
35	gaaggcaaag agctcctgct tttccacac cgagggactt tcaggaatga ggccagggtg	1200
	ctgagcacta caccaggaaa tccctggaga gtgttttct tactta	1246

<210> 24

<211> 939

40 <212> ADN

<213> Secuencia artificial

45 <220>

<223> secuencia de STAR24

<400> 24

50	acgaggtcac gagttcgaga ccagcctggc caagatgggtg aagccctgtc tctactaaaa	60
	atacaaāaag tagccgggcg cggtgācggg cgCctgtaat cccagctact caggaggctg	120
	aagcaggaga atctctagaa cccaggaggc ggaggtgcag tgagctgaga ctgccccgct	180

55

60

65

ES 2 344 071 T3

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

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

ataaaaaaat aaaaaaccct gctctaattt gcaaaggctc tatctttcct cccaaccacc 60
 tgaaatttta gtgaaaacgg ggcttcctgt aggaaggagt agctagctat cccggtccgc 120
 40 tacaggttat cagtgcgtga ataccctgac tcctaaggct caggatttga ctgggtcgcc 180
 tcgtccgact gccccgcccc caacgcggac ccacgtcacc gcgcgccagc ctgcggccgt 240
 cctgacctcg cgggatttga gcttcggtgc caacaaacac tcccaccgcy gctgcgtcca 300
 45 ctttacctgc cggcggcgac cagcttctga agaaaagtgt ccaccatggt gtcgaggagc 360
 ttcacctcgc aaatggtagt gccgggtggc acagattccg aagacgacct ctcatgcctt 420
 tttctctcac agccgctgcc tagattggcg ctacttgctt cggccatggt gaagttgaac 480
 50 ctccaaatct aactggcccc gcctccccgc ctgcccggagc tcccgattgg ccgctcccgc 540
 gaagggtgcc tccgattgga agcagtagaa cgtctgtcac cgagcagggc gggggcgggg 600
 aagtcatcgg aggctgaggg cagcggggag gcgaggctct gcgcgggtgg atgtccgcga 660
 55 ccggaaaaat acgcgcaagc caaagctcgg gggctcaata aaaactttta attacatttc 720
 agagacttcg tacagtgcaa cagtgaatat tcaactgtaa tttcacaag agtccatttc 780
 atcaaacggt cagagagtct gccttttcat tccttgctc ctcagtgtc caatcaggtt 840
 60 tccagtctcc cagaggtttc ttttagtttt gattaccgac caaaactcca gtttagggag 900
 aatggaagtc caccgtccca tccccacaa aacatatttc agtcaaacce aatcccagtc 960
 cctaaagaat taggaaagta tgggccaagg gtccttttaa ttatacacac atcacctta 1020
 65 aaactgcgtg tgtgtacgag aaataaagaa aaacacaaga ggggctg 1067

ES 2 344 071 T3

<210> 26
 <211> 540
 <212> ADN
 5 <213> Secuencia Artificial

 <220>
 <223> secuencia de STAR26
 10
 <400> 26

```

ccccctgaca agccccagtg tgtgatgttc cccactctgt gtccatgcat tctcattgtt    60
15 caactcccat ctgtgagtga gaacatgcag tgtttggttt tctgtccttg agatagtttg    120
ctgagaatga tggtttccag cttcatccat gtccttgcaa aggaagtga cttatccttt    180
tttatggctt catagtattc catggcacat atgtgccaca ttttttaat ccagtctatc    240
20 attgatggac atttgggttg gttccaagtc tttgctattg tgaatagcac cacaattaac    300
atatgtgtgc atgtatacat ctttatagta gcatgattta taatccttcg ggtatatacc    360
ctgtaatggg atcgctgggt caaatggtat ttctagtctt agatccttga ggaatcacca    420
25 cactgctttc cacaatggtt gaactaattt acgctcccac cagcagtgta aaagcattcc    480
tatttctcca cgtcctctcc agtatctggt gtttctctgac tttttaatga tcatcattct    540
  
```

<210> 27
 30 <211> 1520
 <212> ADN
 <213> Secuencia Artificial

 35 <220>
 <223> secuencia de STAR27

 40 <400> 27

```

cttggccctc acaaagcctg tggccagggg acaattagcg agctgcttat tttgctttgt    60
atccccaatg ctgggcataa tgccctgccat tatgagtaat gccggtagaa gtatgtgttc    120
45 aaggaccaa gttgataaat accaaagaat ccagagaagg gagagaacat tgagtagagg    180
atagtgcag aagagatggg aacttctgac aagagttgtg aagatgtact aggcaggggg    240
aacagcttaa ggagagtcac acaggaccga gctcctgtca agccggctgc catggaggct    300
50 ggggtggggc atggtagctt tcccttcctt ctcaggttca gagtgtcagc cttgaacttc    360
taattcccag aggcatttat tcaatgtttt cttctagggg catacctgcc ctgctgtgga    420
agactttctt ccctgtgggt cgccccagtc cccagatgag acggtttggg tcagggccag    480
55 gtgcaccgtt ggggtgtgtc ttatgtctga tgacagttag ttactcagtc attagtcatt    540
gagggaggtg tggtaaagat ggagatgctg ggtcacatcc ctagagaggt gttccagtat    600
gggcacatgg gagggctgga aggataggtt actgctagac gtagagaagc cacatccttt    660
60 aacaccctgg cttttcccac tgccaagatc cagaaagtc ttgtggtttc gctgctttct    720
cctttttttt tttttttttt tttctgagat ggagtctggc tctgtcgccc aggctggagt    780
  
```

65

ES 2 344 071 T3

gcagtggcac gatttcggct cactgcaagt tccgcctcct aggttcatac cattctccca 840
 cctcagcctc ccgagtagct gggactacag gcgccaccac acccagctaa tttttgtat 900
 5 ttttagtaga gacggcgttt caccatgta gccaggatgg tcttgatccg cctgcctcag 960
 cctcccaaag tgctgggatt acaggcgtga gccaccgcgc ccggcctgct ttcttcttcc 1020
 atgaagcatt cagctgggta aaaagctcag ccaggctggg ctggaactct tgacctcaag 1080
 10 tgatctgcct gcctcagcct cccaaagtgc tgagattaca ggcatgagcc agtccgaatg 1140
 tggctttttt tgttttgttt tgaacaagg tctcactggt gcccaggctg cagtgcagtg 1200
 gcatacctca gctccactgc agcctcgacc tcctgggctc aagcaatcct cccaactgag 1260
 15 cctccccagt agctggggct acaagcgcac gccaccacgc ctggctatct tttttttttt 1320
 tttttttttt gagaaggagt ttcattcttg ttgccaggc tggagtgcaa tggcacagtc 1380
 tcagctcact gcagcctccg cctcctgggt tcaagcgatt ctctgcctc agcctcccgga 1440
 20 gtagctggga ttataggcac ctgccaccat gcctggctaa tttttttgta ttttagtag 1500
 ggatgggggt tcaccatggt 1520

<210> 28

<211> 961

<212> ADN

<213> Secuencia Artificial

<220>

<223> secuencia de STAR28

<400> 28

aggaggttat tcctgagcaa atggccagcc tagtgaactg gataaatgcc catgtaagat 60
 ctgtttacc tgagaagggc atttcctaac tctccctata aaatgccaag tggagcacc 120
 40 cagatgaaat agctgatatg ctttctatac aagccatcta ggactggctt tatcatgacc 180
 aggatattca cccactgaat atggctatta cccaagttat ggtaaatgct gtagttaagg 240
 gggtccttc cacatggaca ccccaggta taaccagaaa gggttcccaa tctagactcc 300
 45 aagagagggg tcttagacct catgcaagaa agaacttggg gcaagtacat aaagtgaaag 360
 caagtttatt aagaaagtaa agaaacaaaa aaatggctac tccataagca aagttatttc 420
 tcacttatat gattaataag agatggatta ttcatgagtt ttctgggaaa ggggtgggca 480
 50 attcctggaa ctgaggggtc ctcccacttt tagaccatat agggtatctt cctgatattg 540
 ccatggcatt tgtaaactgt catggcactg atgggagtggt cttttagcat tctaattgat 600
 tataattagc atataatgag cagtgaggat gaccagaggt cacttctggt gccatattgg 660
 55 tttcagtggg gtttggttgg cttttttttt tttttaacca caacctggtt tttatttatt 720
 tattttattt tttatttatt tatatttttt attttttttt agatggagtc ttgctctgtc 780
 acccaggta gagtgcagtg gcaccatctc ggctcactgc aagctctgcc tccttggttc 840
 60 acgccattct gctgcctcag cctcccagat agctgggact acagggtcct gccaccatac 900
 ccggctaatt ttttctattt ttcagtagag acggggtttc accgtgtag ccaggatggg 960

C

961

<210> 29

<211> 2233

ES 2 344 071 T3

<212> ADN

<213> Secuencia artificial

5 <220>

<223> secuencia de STAR29

<400> 29

10	agcttgaca cttgctgatg ccactttgga tgttgaaggg ccgccctctc ccacaccgct	60
	ggccactttt aaatatgtcc cctctgcca gaagggcccc agaggagggg ctggtgaggg	120
15	tgacaggagt tgactgctct cacagcaggg ggttccggag ggacctttc tccccattgg	180
	gcagcataga aggacctaga agggccccct ccaagcccag ctgggcgtgc agggccagcg	240
20	attcgatgcc tccccctgac tcaggtggcg ctgtcctaaa ggtgtgtgtg ttttctgttc	300
	gccagggggg ggcggataca gtggagcatc gtgcccgaa gttctgagcc cgtggtaagt	360
25	ccctggaggg tgcacggctc cctccgactg tctccatcac gtcaggcctc acagcctgta	420
	ggcaccgctc ggggaagcct ctggatgagg ccatgtggtc atccccctgg agtccctggc	480
30	tggcctgaag aggaggggag gaggaggcca gccctccct agccccaagg cctgcgaggg	540
	tgcaagcccc gccccacatt ctagtccagg cttggctgtg caagaagcag attgcctggc	600
35	cctggccagg cttcccagct aggatgtggt atggcagggg tgggggacat tgaggggctg	660
	ctgtagcccc cacaacctcc ccaggtaggg tgggtaacag taggctggac aagtggacct	720
40	gttcccatct gagattcaag agcccacctc tcggaggttg cagtgaagcc agatccctcc	780
	actgactcc agcctgggca acagagcaag actctgtctc aaaaaaacag aacaacgaca	840
45	acaaaaaacc cacctctggc ccactgccta actttgtaa taaagttta ttggcacata	900
	gacacacca ttcatttaca tactgctgcg gctgcttttg cattaccctt gagtagacga	960
50	cagaccacgt ggccatggaa gccaaaaata tttactgtct ggcctttac agaagtctgc	1020
	tctagagggg gacccccggc catggggcag gaccactggg cgtgggcaga agggaggcct	1080
55	cgggtgcctc acgggcctag ttgggtatct cagtgcctgt ttcttgcag gagcaccagg	1140
	ggtcagggca agtacctgga ggaggcaggc tgttgccgc ccagcactgg gaccacaggag	1200
60	accttgagag gctcttaacg aatgggagac aagcaggacc agggctccca ttggctgggc	1260
	ctcagtttcc ctgcctgtaa gtgagggagg gcagctgta aggtgaactg tgaggcagag	1320
65	cctctgctca gccattgcag gggcggctct gccccactcc tgttgtgac ccagagtgag	1380
	gggcacgggg tgagatgtca ccatcagccc ataggggtgt ctcctggtg ccaggtcccc	1440
70	aagggatgtc ccatcccccc tggctgtgtg gggacagcag agtccctggg gctgggaggg	1500
	ctccacactg ttttgtcagt ggtttttctg aactgttaa tttcagtgga aaattctctt	1560
75	tcccccttta ctgaaggaac ctccaaagga agacctgact gtgtctgaga agttccagct	1620
	ggtgctggac gtcgccaga aagcccagg actgccacgg gcgccggcca ggggtgtgtc	1680

ES 2 344 071 T3

tgcgccagcc atgggcacca gccaggggtg tgtctacgcc ggccaggggt aggtctccgc 1740
 cggcctccgc tgctgcctgg ggagggccgt gcctgacact gcaggcccgg tttgtccgcg 1800
 5 gtcagctgac ttgtagtcac cctgcccttg gatggtcggt acagcaactc tgggtggttg 1860
 ggaaggggcc tcctgattca gcctctgcgg acggtgcgcg aggggtggagc tcccctccct 1920
 ccccaccgcc cctggccagg gtgaaacgcc cctgggaagg actcaggccc gggctctgctg 1980
 10 ttgctgtgag cgtggccacc tctgccctag accagagctg ggccctcccc ggccctaggag 2040
 cagccgggca ggaccacagg gctccgagtg acctcagggc tgcccagcct ggaggccctc 2100
 ctggcgtcgc ggtgtgactg acagcccagg agcgggggct gttgtaattg ctgtttctcc 2160
 15 ttcacacaga accttttcgg gaagatggct gacatcctgg agaagatcaa gaagtaagtc 2220
 ccgcccccca ccc 2233

<210> 30

20 <211> 1851

<212> ADN

<213> Secuencia artificial

25 <220>

<223> secuencia de STAR30

<400> 30

30 cctcccctgg agccttcaga aggagcatgg cataggagtc ttgatttcag acgtctggtc 60
 cccagaatga tgggagaatg aatttctggt atttaagcca cccaacctgt ggtgctttgt 120
 35 tatagcagcc tcaggaaact aacacactgc acgtgcccac tattcccttt tccagtatct 180
 ttcaggactt gctggcttcc tttgttctgg cgtacaccca tgcattggccc cattccccac 240
 ttctataaac aacaacctg acttagtctg tttgggctgc tagaacaaaa tactatagac 300
 tgggtgactt ataaacaaca gaaattcatt tctcacattc tggaggctgg gaagtccaat 360
 40 atcaggacac catcacattt ggtctctgct gagggcccct tcctagctcc tctactgtgtc 420
 cttacatggc agaaggggca aggcagctct ctggggctcc ttttcaaggc cacaaatccc 480
 attcattagg gctgatgact tcatgactta atcacctcct aatggcccca cctcctaatc 540
 45 gcattgggcg ttaggattca acataaattt tggggggaca cacatattca gaccatagca 600
 aacccaaca ataaaaaacc ttcactttaa ggttccaaat ggactggcag ttaaatcatg 660
 ttcataatta cataaaagaa ggagtaagtc aacaaattga taaacgcgtg gagatttggt 720
 50 cggatggatg ttcacaaaa tgctggcctt aaagagtgag atgggaaatg ggaactatta 780
 cattcttctt catactttt ggtactgcct gcattgttaa aaaaaaaaaa aaagagcaca 840
 gagcattttt acaatcagga aaaaaacaat gaggttatct tcattctgga aaaaaatgga 900
 55 aatgaaaca gtggagtcac atcatgaaa atgcttatgg tacaatttca tgtgacataa 960
 aacaatagaa tagaggacct gttttatgac taaagcactg taaaaatgac aggcctggaa 1020
 ggagagatga aaaccactca tttgttaagg tagtcaggtg gcaggtgatt tctcttcttt 1080
 60 tgaaaatttc cattttcatt atatcgcagt ttgtgcattt actaaaactt tcggttgga 1140

65

ES 2 344 071 T3

cacatgcata aatagataga taaataagta gatagatgat agataaatag acggtaggta 1200
 gatagataga tagatatgag aaataagtcc cctgtacttg gccttgacag 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 cttaccaat ttctaattat aactaaaaca caatgaggct ctagtaaat accatgagac 1500
 ttcaggccct ctgtataact tcaactcattt aaacctaaca aggaaaacct accatgaatc 1560
 cgaggcacag agcagctaag gaactcacca aggtcacgca gctattgggtg atggaacct 1620
 15 gagtcaagct tcacagcctg ttggctctag aatagggttt cccaacctca gcactgtgga 1680
 cttttcagg ctggataatt ctctgttggtg gggggctggt ctgtgccttg taggatatta 1740
 ggagcatctc tggcctctac ccactagacg cagcagcact cccatgcccc gttgtgacaa 1800
 20 caagcaatgt ctcccacat tgccaagtgt cccctgggtg gaaatgcacc c 1851

<210> 31
 <211> 1701
 25 <212> ADN
 <213> Secuencia artificial

<220>
 30 <223> secuencia de STAR31

<220>
 <221> característica_misc
 35 <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 gagcccctta 120
 45 actgggacag gggcaggaca ggggtgaggc ttccttant tcaagctcac ctcaaaccca 180
 cccaggactg tgtgtcacat tctccaataa aggaaagggt gctgcccccg cctgtgagtg 240
 ctgcagtgga gggtagaggg ccgtgggcag agtgcttcat ggactgctca tcaagaaagg 300
 50 cttcatgaca atcggcccag ctgctgtcat cccacattct acttccagct aggagaaggc 360
 ggcttgcaca cagtcaccca gccggcaagt gtcaccctg ggttgaccc agagctatga 420
 tcctgcccag ggtccagct gagaatcagg cccacgttct aggcagaggg gctcacctac 480
 55 tgggactcca gtagctgtag tgcatggagg catcatggct gcagcagcct ggacctggtc 540
 tcacactggc tgtccctgtg ggcaggccat cctcaatgcc aggtcaggcc caagcatgta 600
 tcccagacaa tgacaatggg gtggaatcct ctcttgtccc agaagccact cctcactggt 660
 60 ctacctgagg aaggcagggg catggtggaa tcctgaagcc tgctgtgagg gtctccagcg 720
 aacttgcaca tggtcagccc tgcttctcc tccctgaact agattgagcg agagcaagaa 780
 ggacattgaa ccagcaccca aagaattttg gggaacggcc tctcatccag gtcaggctca 840

65

ES 2 344 071 T3

5 cctccttttt aaaatttaat taattaatta attaattttt ttttagagac agagtcttac 900
 tgtgtggccc aggctgtagt gcagtggcac aatcatagtt cactgcagcc tcaaactccc 960
 cacctcagcc tctggattag ctgagactac aggtgcacca ccaccacacc cagctaatat 1020
 ttttattttt gtagagagag ggtttcacca tcttgcccag gctggtctca aactcctggg 1080
 ctcaagtgat cccgccagc tctgaaagcc cccaggctgg cctcagactg tggggttttc 1140
 10 catgcagcca cccgagggcg cccccaagcc agttcatctc ggagtccagg cctggccctg 1200
 ggagacagag tgaaacagc ggtttttatg aacttaactt agagttaaaa agatttctac 1260
 tcgatcactt gtcaagatgc gccctctctg gggagaaggg aacgtgactg gattccctca 1320
 15 ctgttgatc ttgaataaac gctgctgctt catcctgtgg gggccgtggc cctgtccctg 1380
 tgtgggtggg gcctcttcca tttccctgac ttagaaacca cagtccacct agaacagggg 1440
 ttgagaggct tagtcagcac tgggtagcgt tttgactcca ttctcggctt tcttctttt 1500
 20 ctttccagga tttttgtgca gaaatggttc ttttgttgcc gtgttagtcc tccttggag 1560
 gcagctcaga aggcccgta aatgtcgggg gacaggacc ccaggagggg aaccccaggc 1620
 tacgcacttt agggttcgtt ctccaggggg ggcgacctga cccccgnatc cgtcggngcg 1680
 25 cgngngnacn aanntttcc c 1701

<210> 32

<211> 771

30 <212> ADN

<213> Secuencia artificial

<220>

35 <223> secuencia de STAR32

<400> 32

40 gatcacacag ctgtatgtg ggagctagga ttggaacccc agaagtctgg ccccaggttc 60
 atgctctcac cactgcata caatggcctc tcataaatca atccagtata aacattaga 120
 atctgcttta aaaccataga attagtagcg taagtaataa atgcagagac catgcagtga 180
 45 atggcattcc tggaaaaagc ccccagaagg aattttaaat cagctttcgt ctaatcttga 240
 gcagctagtt agcaaatatg agaatacagt tgttcccaga taatgcttta tgtctgacca 300
 tcttaaactg gcgctgtttt tcaaaaactt aaaaacaaaa 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 cttttttctt ctcttctctt ctcttttctt tttttttgag 720
 60 acggagcttt gctctttttg cccaggcttg attgcaatgg cgtgatctca g 771

<210> 33

<211> 1368

65 <212> ADN

<213> Secuencia Artificial

ES 2 344 071 T3

<220>

<223> secuencia de STAR33

5 <400> 33

	gcttttatcc tccattcaca gctagcctgg cccccagagt acccaattct ccctaaaaaa	60
	cggctcatgct gtatagatgt gtgtggcttg gtagtgctaa agtggccaca tacagagctc	120
10	tgacaccaaa cctcaggacc atgttcatgc cttctcactg agttctggct tgttcgtgac	180
	acattatgac attatgatta tgatgacttg tgagagcctc agtcttctat agcactttta	240
	gaatgcttta taaaaacat ggggatgtca ttatattcta acctgttagc acttctgttc	300
15	gtattacca tcacatcca acatcaattc tcatatatgc aggtacctct tgtcacgcgc	360
	gtccatgtaa ggagaccaca aacaggctt tgtttgagca acaaggtttt tatttcacct	420
	gggtgcaggt gggctgagtc tgaaaagaga gtcagtgaag ggagacaggg gtgggtccac	480
20	tttataagat ttgggtaggt agtggaaaat tacaatcaaa gggggttgtt ctctggctgg	540
	ccagggtggg ggtcacaagg tgctcagtgg gagagccttt gagccaggat gagccagaag	600
	gaatttcaca aggtaatgtc atcagttaag gcagggactg gccattttca cttcttttgt	660
25	ggtggaatgt catcagttaa ggcaggaacc ggccattttc acttcttttg tgattcttca	720
	cttgcttcag gccatctgga cgtataggtg caggtcacag tcacagggga taagatggca	780
	atggcatagc ttgggctcag 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 ccaccccaca	1080
	aaagaacaaa caacaaatcc caggacatct tagcgcaaca tttcagtttg catattttac	1140
	atatttactt ttcttacata ttaaaaaact gaaaatttta tgaacacgct aagttagatt	1200
40	ttaaattaag ttgttttta cactgaaaat aatttaatat ttgtgaagaa tactaataca	1260
	ttggtatatt tcattttctt aaaattctga acccctctc ccttatttcc ttttgaccg	1320
	attggtgat tggctatgtg actcatggat ttgccttaag gcaggagg	1368

45 <210> 34

<211> 755

<212> ADN

50 <213> Secuencia Artificial

<220>

<223> secuencia de STAR34

55 <400> 34

	actgggcacc ctcttaggca ggggaatgtg agaactgccg ctgctctggg gctgggcgcc	60
60	atgtcacagc aggagggagg acggtgttac accacgtggg aaggactcag ggtggtcagc	120

65

ES 2 344 071 T3

	caciaagctg ctggtgatga ccaggggctt gtgtcttcac tctgcagccc taacacccag	180
	gctgggttcg ctaggctcca tcctgggggt gcagaccctg agagtgatgc cagtgggagc	240
5	ctcccgcccc tccccttcct cgaaggccca ggggtcaaac agtgtagact cagaggcctg	300
	agggcacatg tttatttagc agacaagggtg gggctccatc agcgggggtgg cctggggagc	360
	agctgcatgg gtggcactgt ggggagggtc tcccagctcc ctcaatggtg ttcgggctgg	420
10	tgcggcagct ggcggcacc cggacagagg tggatatgag ggtgatgggt ggggaaatgg	480
	gaggcaccg agatggggac agcagaataa agacagcagc agtgctgggg ggcaggggga	540
	tgagcaaagg caggcccaag acccccagcc cactgcaccc tggcctccca caagccccct	600
15	cgacgcccgc cagccacact cactgtgcac tcagccgtcg atacactggt ctgttaggga	660
	gaaagtccgt cagaacaggc agctgtgtgt gtgtgtgtgt gtatgagtgt gtgtgtgtga	720
	tccttgactg ccaggctctc tgcactgccc ctggg	755

20

<210> 35

<211> 1193

<212> ADN

25

<213> Secuencia artificial

<220>

<223> secuencia de STAR35

30

<220>

<221> característica_misc

<222> (355)..(1191)

35

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

<220>

40

<221> característica_misc

<222> (312)..(1191)

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

45

<400> 35

	cgacttggtg atgcgggctc ttttttggtt ccatatgaac tttaaagtag tcttttccaa	60
	ttctgtgaag aaagtcattg gtaggttgat ggggatggca ttgaatctgt aaattacctt	120
50	gggcagtatg gccattttca caatgttgat tcttcctatc catgatgatg gaatgttctt	180
	ccattagttt gtatcctctt ttatttcctt gagcagtggt ttgtagtctt ccttgaagag	240
	gtccttcaca tcccttgtaa gttggattcc taggtatttt attctctttg aagcaaattg	300
55	tgaatgggag tncactcacg atttggctct ctgtttgtct gctgggtgta taaanaatgt	360
	ngtgatnttn gtacattgat ttngtatccn tgagacttng ctgaatttgc ttnatcngct	420
	tnngggaacc ttttgggctg aaacnatggg attttctaaa tatacaatca tgtcgtctgc	480
60	aaacagggaa caatttgact tcctcttttc ctaattgaat acactttatc tccttctcct	540
	gcctaattgc cctgggcaaa acttccaaca ctatgntngn aataggagnt ggtgagagag	600
	ggcatccctg ttcttggtgc cagnttttca aagggaatgc ttccagtttt ggccattcā	660
65	gtatgatatg ggctgtgggt nggtgcataa atagctctta tnattttgaa atgtgtccca	720

ES 2 344 071 T3

tcaataccta atttattgaa agtttttagc atgaangcat ngttgaattt ggtcaaaggc 780
 tttttctgca tctatggaaa taatcatgtg gtttttgtct ttggctcntg tttatatgct 840
 5 ggatnacatt tattgatttg tgtatatnga acccagcctn ncatcccagg gatgaagccc 900
 acttgatcca agcttggcgc gcngnctagc tcgaggcagg caaaagtatg caaagcatgc 960
 atctcaatta gtcagcacc atagtcgcc cctacctcg cccatccgcc cctaactcng 1020
 10 nccgttcgcc cattctcgcc catggctgac taatnttttt annatccaag cggngccgcc 1080
 ctgcttganc attcagagtn nagagnnttg gaggccnagc cttgcaaac tccggacngn 1140
 ttctnnggat tgaccccnnt taaatatttg gttttttgtn ttttcannng nga 1193

15 <210> 36
 <211> 1712
 <212> ADN
 20 <213> Secuencia artificial
 <220>
 <223> secuencia de STAR36
 25 <400> 36

gatcccatcc ttagcctcat cgatacctcc tgctcacctg tcagtgcctc tggagtgtgt 60
 30 gtctagccca ggcccatccc ctggaactca ggggactcag gactagtggg catgtacact 120
 tggcctcagg ggactcagga ttagtgagcc ccacatgtac acttggcctc agtggactca 180
 ggactagtga gccccacatg tacacttggc ctcaggggac tcaggattag tgagccccc 240
 35 catgtacact tggcctcagg ggactcagga ttagtgagcc ccacatgtac acttggcctc 300
 aggggactca ggactagtga gccccacatg tacacttggc ctcaggggac tcagaactag 360
 tgagcccccac atgtacactt ggcttcaggg gactcaggat tagtgagccc cacatgtaca 420
 40 cttggacacg tgaaccacat cgatgtgctg cagagctcag ccctctgcag atgaaatgtg 480
 gtcatggcat tccttcacag tggcaccctt cgttccctcc ccacctcacc tcccattctt 540
 gtctgtcttc agcacctgcc atgtccagcc ggcagattcc accgcagcat cttctgcagc 600
 45 acccccagacc acacacctcc ccagcgcctg cttggcctc cagcccagct cccgcctttc 660
 ttcttgggg aagctccctg gacagacacc ccctcctccc agccatggct ttttctgct 720
 ctgccccacg cgggaccctg ccctggatgt gctacaatag acacatcaga tacagtcctt 780
 50 cctcagcagc cggcagacc aggggtggact gctcggggcc tgctgtgag gtcacacagg 840
 tgtcgtaaac ttgccatctc agcaactagt gaatatggg agatgctacc ttcttccgg 900
 ttccttggtg agaggtactg gtggatgtcc tgtgttgccg gccacctttt gtccctggat 960
 55 gccatttatt tttttccaca aatatttccc aggtctcttc tgtgtgcaag gtattagggc 1020
 tgcagcgggg gccaggccac agatctctgt cctgagaaga cttggattct agtgcaggag 1080
 actgaagtgt atcacaccaa tcagtgtaaa ttgttaactg ccacaaggag aaaggccagg 1140
 60 aaggagtggg gcatggtggt gttctagtgt tacaagaaga agccaggag ggcttctctg 1200
 atgaagtggc atctgacctg ggatctggag gaggagaaaa atgtcccaa agagcagaga 1260

65

ES 2 344 071 T3

gcccacccta ggctctgcac caggaggcaa cttgctgggc ttatggaatt cagagggcaa 1320
 gtgataagca gaaagtcctt gggggccaca attaggattt ctgtcttcta aagggcctct 1380
 5 gccctctgct gtgtgacctt gggcaagtta cttcacctct agtgctttgg ttgcctcatc 1440
 tgtaaagtgg tgaggataat gctatcacac tggttgagaa ttgaagtaat tattgctgca 1500
 aagggccttat aagggtgtct aatactagta ctagtaggta cttcatgtgt cttgacaatt 1560
 10 ttaatcatta ttattttgtc atcacctgca ctcttcagg ggactaatgt ccctgctgtt 1620
 ctgtccaaat taaacattgt ttatccctgt gggcatctgg cgagggtggct aggaaagcct 1680
 ggagctgttt cctgttgacg tgccagacta gt 1712

15 <210> 37
 <211> 1321
 <212> ADN
 20 <213> Secuencia artificial
 <220>
 <223> secuencia de STAR37
 25 <400> 37

atctctctct gccaaagcaa cagcggctcc tgccccaacc agactacccc actcagtggg 60
 30 gttacggatg ctgctccagc atcctaacac tgcccagctg gtgcctgcct gtgctcacc 120
 acaaccccca ggccggcctt ccctgcagcc tgggcttggc caccttggcc tgattgagca 180
 ctgaggcctc ctgggcaccc agccccatca ctgcacctgc tgcttcacag cccacccac 240
 35 cggctcaggg gttcttccca gcggcgctga tcatgaagtc aacatgcacg caagtcgtct 300
 caggaaactt tttaatgaaa gtgtcggcca cgggtggtgt taggtggctg agctcagatt 360
 gcagctgcta agacaccagc cacttaccaa gagaaagcca ggctgcttca aaccagggc 420
 40 cggaggcaaa aaagcatcac ttccagccgg ggagtctgga agccacgcct tgtgggaggt 480
 cacactggca tctaggcctt cgctgcact gcagaaggag agccgggtcc cctcctgga 540
 gaacgctgcg ttccccagcc ccacaccggc tttgccacca cacaggctgt tgaggcagga 600
 45 ggcgggtaag acgtagctgt agacccaaag caaccaccag ccctgggacc ctgctgggaga 660
 ggagcacttt tagaacatgg aaaaatgtgg tcatcccatc attagacagc acacatccta 720
 cataaataaa aagtcgtatg ggaagggagg ttggggaggg aataaaaaat tggcacagac 780
 attgatagac tggtttccag tttcaaggta acagatgcac atcatgagac cagaggaggc 840
 50 agagacaagg gctgaattt gcttttctaa gcaacatgtg ttcctgcgca gggctgaatg 900
 gtcgctgaga cagagatgga agccaggaca agggagccca ccgggcccag ataggtacag 960
 55 agagcagagg ctctgttct gtctcgcga cccatgaggg tgacactgct tgtaaattgt 1020
 ggctgtgctc tcccagcaag aaaaaagcac aactaaatcc aactgcaca cagacgcaga 1080
 cagaaagcct tcaagtggct ctgttttctg ctccctgcct tgccaggtcc acaagcagag 1140
 60 aggagtgtca ggcacatggc cccgctgtca ggctccccag tgagctgtag gctcagcagg 1200
 agctgcccac tgacacacag gggacacca ctctgccac cttgggagcg gttgccagac 1260
 agagccgcac tgggtgctgg tgtcatccag ggacccaca cacttcctta aatgtgatcc 1320
 65 t 1321

ES 2 344 071 T3

	<210> 38	
	<211> 1445	
	<212> ADN	
5	<213> Secuencia Artificial	
	<220>	
	<223> secuencia de STAR38	
10	<220>	
	<221> característica_misc	
	<222> (348)..(949)	
15	<223> "N" representa cualquier ácido nucleico en diferentes posiciones	
	<400> 38	
20	gatctatggg agtagcttcc ttagtgagct ttcccttcaa atactttgca accaggtaga	60
	gaattttgga gtgaagggtt tgttcttcgt ttcttcacaa tatggatatg catcttcttt	120
	tgaaaatggt aaagtaaatt acctctcttt tcagatactg tcttcatgcg aacttggtat	180
25	cctgtttcca tcccagcctt ctataaccca gtaacatctt ttttgaacc agtgggtgag	240
	aaagacacct ggtcaggaac gcggaccaca ggacaactca ggctcaccca cggcatcaga	300
	ctaaaggcaa acaaggactc tgtataaagt accggtggca tgtgtatnag tggagatgca	360
30	gcctgtgctc tgcagacagg gagtcacaca gacacttttc tataatttct taagtgcctt	420
	gaatgttcaa gtagaaagtc taacattaaa tttgattgaa caattgtata ttcatggaat	480
	atthtggaaac ggaataccaa aaaatggcaa tagtggttct ttctggatgg aagacaaact	540
35	tttcttgttt aaaataaatt ttattttata ttttgaggt tgaccacatg accttaagga	600
	tacatataga cagtaaactg gttactacag tgaagcaaat taacatatct accatcgtac	660
	atagttacat ttttttgtgt gacaggaaca gctaaaatct acgtatttaa caaaaatcct	720
40	aaagacaata catttttatt aactatagcc ctcatgatgt acattagatc gtgtggttgt	780
	ttcttccgtc cccgccacgc ctctctctg ggatggggat tcattcccta gcagggtgctg	840
	gagaactggc gcccttgtag ggtaggtgcc ccggagcctg aggcgggnac tttaanatca	900
45	gacgcttggg ggccggctgg gaaaaactgg cggaaaatat tataactgna ctctcaatgc	960
	cagctgttgt agaagctcct gggacaagcc gtggaagtcc cctcaggagg ctcccgcat	1020
	gtcctaggtg gctgctccgc ccgccacggt catttccatt gactcacacg cgccgcctgg	1080
50	aggaggaggc tgcgctggac acgccggtgg cgcctttgcc tgggggagcg cagcctggag	1140
	ctctggcggc agcgtgga gcggggcctc ggaggctggg cctggggacc caaggttggg	1200
	cggggcgcag gaggtgggct cagggttctc cagagaatcc ccatgagctg acccgagg	1260
55	cggccgggcc agtaggcacc gggccccgc ggtgacctgc ggaccgaag ctggagcagc	1320
	cactgcaaat gctgcgctga ccccaaatgc tgtgtccttt aaatgtttta attaagaata	1380
	attaataggt ccgggtgtgg aggctcaagc cttaatcccc agcacctggc gaggccgagg	1440
60	agggga	1445
65	<210> 39	
	<211> 2331	
	<212> ADN	

ES 2 344 071 T3

<213> Secuencia artificial

<220>

5 <223> secuencia de STAR39

<400> 39

10	<p>tcactgcaac ctccacctcc caggttcaag tgattctcct gcctcggcct cccgagtagc 60</p> <p>tgggactaca ggtgcatgac accgcacctg gctagttttt gtatttttag tagagacagg 120</p> <p>gtttcactat gttggccagg ttggtctcga actcctgacc ttgtgatccg cccacctcgg 180</p>	
15	<p>cctcccaaag tgctgggatt acagagtgag ccaactgcgc tggcctgcac cccttactat 240</p> <p>tatatgcttt gcattttctt ttagatttga agaacctcat tataaactct agcactaatc 300</p> <p>ttatgtcagt taaatgcata gcaaatatct cctgacgtgg gagaatatat atttgcaagt 360</p>	
20	<p>cttcttgatga acatatgttt tcagttctag ggagccagac gcctatgagt gaaaagccta 420</p> <p>gtcatcgtgg agaagtgcac tcaactttgt aagaaactgc caaaccttta ttcataatgg 480</p> <p>ttgtataaat ttacattac caccaataat gtatgagagt tccagttgct tcacatcctc 540</p>	
25	<p>accagcattt tgttttgtct gtcttttttc ctttggttat tctagtgggc ataagatata 600</p> <p>atagtatccc ttgtggttta atgtaaattc cactgaagac taataacatt tgcatatttc 660</p> <p>taattaataa gcctttttaa gtgacttttc aagtctttgc tcatttttat tagatatttg 720</p>	
30	<p>ccttcttatt attgatttga aagaattata tttatatgct tatattctgg ttataagccc 780</p> <p>tttgtcatta ttttcaaaa caatatttgg ttgtttctgt actactttcc ttgtccttt 840</p> <p>gaattgactt ggtgccttgg ccaaaaatca attgaccaca tacatgtggg tgcatctcca 900</p>	
35	<p>gactaccaca ttccgtttat ctatttgtct ctcttgtgt caataacact ctgtcttgat 960</p> <p>aatggttaagt ttgagatca ggttgtgtaa gtctcctaa tttttcctgg gttttcaata 1020</p> <p>ttgctttgct ttttaaaaat ttgtatattt catttacatt ttaaataaaa cttgttagtg 1080</p>	
40	<p>ggattttgat tggcattgca ctgaactcgt ggatcaattt ggggagattg gacattctta 1140</p> <p>tatatggatc ccgtggtcat caactttaag aactctttct catccattag taactcaatc 1200</p> <p>taggttcaga tgctactcgt tttctgctca gtctgtgtct gagcccctta tgctcttcat 1260</p>	
45	<p>tttgtcatcc aattaacctc agctttgcat caatactatt tcttgctttg tgctctgtta 1320</p> <p>cctctcctct aatcaccaat ccacaactta cctccaaatt cagggcttgt ctcattcttc 1380</p> <p>ccaggaggag tgctgctcag tctatctact tagtattata atttctctgg cttggtatca 1440</p>	
50	<p>aggcactccc atttccggct tccatgagat gtctcagagg gcatgctgcc cgggtgtagct 1500</p> <p>gcatggtcaa gcttcttcat atctcttgcc tcatcactta aactcactat tttgtactcc 1560</p> <p>tgcttcagct ataggagct actgttagtt tcttgaagac atatgctctc tctctctctc 1620</p>	
55	<p>acatctggac ctgagcacat cctgttactg ctgcttgaag caatgtgatc cccaggcaca 1680</p>	

60

65

ES 2 344 071 T3

caccattagc ttagaagcct cccctgattc ttcaaggctg gttgagtccc ttctctgtgc 1740
 tctcatgaca acagttggca attcctcggt gcagcaccta gcccatgatg ctctttggag 1800
 5 gcagagactg agtctttctc actattgaaat 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
 cagggatcc tgtcatttaa cacaatatgg ataaacctag aggactctaa aggcaaatac 2160
 15 cacatgttct cactcacaaa atctaaacaa gttgaaactcc tacaagtaga gagtaggatg 2220
 atggttacca agggctgggg gacgggagag gatggggaaa gcatagctgt ccatcaaagg 2280
 gtagaaagt tcathtagac aagaggaatc agctttagtg 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 aagctgctga tgagcagtga gtgtcttgag 120
 tagtgttcag ggcagcatgt taccattcat gcttgacttc tagccagtgt gacgagaggc 180
 40 tggagtcagg tctctagaga gttgagcagc tccagcctta gatctcccag tcttatgcgg 240
 tgtgccatt cgctttgtgt ctgcagctcc ctggccacac ccagtaacag ttctgggatc 300
 tatgggagta gcttccttag tgagctttcc ctcaaatatc tttgcaacca ggtagagaat 360
 tttggagtga aggttttgtt cttcgtttct tcacaatatg gatatgcatc ttcttttgaa 420
 45 aatgttaaag taaattacct ctcttttcag atactgtctt catgcgaact tggatcctg 480
 tttccatccc agccttctat aaccagtaa catctttttt gaaaccagtg ggtgagaaag 540
 acacctggtc aggaacgcgg accacaggac aactcaggct caccacggc atcagactaa 600
 50 aggcaaacaa ggactctgta taaagtaccg gtggcatgtg tattagtgga gatgcagcct 660
 gtgctctgca gacagggagt cacacagaca cttttctata atttcttaag tgctttgaa 720
 gttcaagtag aaagtctaac attaaatttg attgaacaat tgtatattca tggaaatatt 780
 55 tggaacggaa taccaaaaaa tggcaatagt ggttctttct ggatggaaga caaactttc 840
 ttgtttaaaa taaattttat tttatatatt tgaggttgac cacatgacct taaggataca 900
 tatagacagt aaactggta ctacagtga gcaaattaac atatctacca tcgtacatag 960
 60 ttacattttt ttgtgtgaca ggaacagcta aaatctacgt atttaacaaa aatcctaaag 1020
 acaatacatt tttattaact atagccctca tgatgtacat tagatctcta a 1071

65 <210> 41
 <211> 1135

ES 2 344 071 T3

<212> ADN

<213> Secuencia Artificial

5 <220>

<223> secuencia de STAR41

<400> 41

```

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

```

<210> 42

45 <211> 735

<212> ADN

<213> Secuencia Artificial

50 <220>

<223> secuencia de STAR42

<400> 42

```

55      aagggtgaga tcactagggg gggaggaagg agctataaaa gaaagaggtc actcatcaca      60
      tcttacacac tttttaaacc cttggttttt taatgtccgt gttcctcatt agcagtaagc      120
      cctgtggaag caggagtctt tctcattgac caccatgaca agaccctatt tatgaaacat      180
60      aatagacaca caaatgttta tcggatattt attgaaatat aggaattttt cccctcacac      240
      ctcatgacca cattctggta cattgtatga atgaatatac cataatttta cctatggctg      300

```

65

ES 2 344 071 T3

	tatatttagg tcttttcgtg caggctataa aaatatgtat gggccggtca cagtgactta	360
	cgcccgtagt ccagaaactt 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 ataagtgtg ccaacttacc ttagggttta taccatttat	720
	gagggtgtcg gtttc	735

15 <210> 43
 <211> 1227
 <212> ADN
 20 <213> Secuencia Artificial
 <220>
 <223> secuencia de STAR43
 25 <400> 43

	caaatagatc tacacaaaac aagataatgt ctgccattt ttcaaagat aatgtggtga	60
	agtgggtaga gagaaatgca tccattctcc ccaccaacc tctgctaaat tgtccatgtc	120
30	acagtactga gaccaggggg cttattccca gcgggcagaa tgtgcaccaa gcacctcttg	180
	tctcaatttg cagtctaggc cctgctattt gatggtgtga aggcttgac ctggcatgga	240
35	aggccgttt tgtacttctt gctttagcag ttcaaagagc agggagagct gcgagggcct	300
	ctgcagcttc agatggatgt ggtcagcttg ttggaggcgc cttctgtggt ccattatctc	360
	cagccccct gcggtgttgc tgtttgcttg gcttgtctgg ctctccatgc cttgttgct	420
40	ccaaaatgtc atcatgctgc accccaggaa gaatgtgcag gcccatctct tttatgtgct	480
	ttgggctatt ttgattccc gttgggtata ttccctaggt aagaccaga agacacagga	540
	ggtagtgtc ttgggagagt ttggacctat gggatgagg taatagacac agtatcttct	600
	ctttcatttg gtgagactgt tagctctggc cgcgactga attccacaca gctcacttgg	660
45	gaaaacttta ttccaaaaca tagtcacatt gaacatttg gagaatgagg gacagagaag	720
	aggccctaga tttgtacatc tgggtgttat gtctataaat agaatgctt ggtggtcaac	780
	tagacttgtt catgttgaca tttagtcttg ccttttcggt ggtgatttaa aaattatgta	840
50	tatcttgttt ggaatatagt ggagctatgg tgtggcattt tcatctggct tttgttttag	900
	ctcagcccgt cctgttatgg gcagccttga agctcagtag ctaatgaaga ggtatcctca	960
	ctccctccag agagcgggtcc cctcacggct cattgagagt ttgtcagcac cttgaaatga	1020
55	gtttaaactt gtttattttt aaaacattct tggttatgaa tgtgcctata ttgaattact	1080
	gaacaacctt atggttgtga agaattgatt tgggtctaag gtgtataaat ttcaggacca	1140
	gtgtctctga agagttcatt tagcatgaag tcagcctgtg gcaggttggg tggagccagg	1200
60	gaacaatgga gaagctttca tgggtg	1227

<210> 44
 65 <211> 1586
 <212> ADN
 <213> Secuencia artificial

ES 2 344 071 T3

<220>

<223> secuencia de STAR44

5 <400> 44

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

<210> 45

<211> 1981

55 <212> ADN

<213> Secuencia artificial

<220>

60 <223> secuencia de STAR45

65

ES 2 344 071 T3

<400> 45

	cacgccccag	cgtagccctg	actactgctc	cgaggactc	ctgttctgct	gcaccctgga	60
5	ctacggcacc	agaggacca	gctcccgcg	gcctgagcta	tggcaccaga	ggaccagct	120
	cccggcagcc	tggactatgg	caccagagga	cccagcccc	cgcttctg	gctaaggcac	180
	agtaggaccc	tgctcatcg	tgtactcctg	ctcaggagga	ccctcgcagg	gcggcgact	240
10	ggactaagct	actgaaggag	ccccaccct	gcctaaccct	ggactaaggc	actggagaac	300
	tcttgctccg	cagagccacg	gactcttgca	caagagaacc	tcagcccagc	cgtagccctg	360
	actgtggcac	agtagggccc	acaccacgcc	atggactcct	gtattggagg	aagagtagtg	420
15	ataaatgtcc	aggtttacia	cttgaaaagt	agcaatcaat	gtgccacaat	agatggatgt	480
	gatgtaaaat	tataaatgat	gaaaacatta	tgtgtaattg	cctagccaga	acagttacac	540
	aagacaaaga	cgtaaaagaa	atccacatag	ggaaggaaga	ggtaaagattg	tttctgtttt	600
20	ttgaaaatat	aatcttaaga	tagagaaaat	cttaaagatt	ccacaaaat	aatggttat	660
	agctgatgaa	gaaattcaat	aaagttaata	gttacaaaat	caacatacaa	atatcattat	720
	tgtttctatt	aactaatgac	aaactattac	ctgaaaaata	aaggcaattc	aatttataat	780
25	agaatcaaaa	cagatatata	aatatataaa	agacaggagt	aaatttaatc	aaaaccataa	840
	aagatttaca	tactgaaaac	tatagcacat	tgatgaaaaa	aattaaaatg	gcataaataa	900
	atggagaaaac	atccttcatt	gatggattca	aaaattagta	ttgtaaaagt	gtcaatgcta	960
30	cccaaagcaa	tctacagatt	aatgcaacc	actatcaaat	tccaatgtca	ttcttcacag	1020
	aatagaaaa	attactgcta	aaatttgat	ggaaccacaa	aagacctgga	ccaaccaaag	1080
	caatcttgaa	caaaaagaac	aaagctggag	gcatcagact	acctgactcc	aaactctatt	1140
35	acaaagctat	aggaattaa	acagcatagc	aatggcataa	aacagacat	gtaaaacagt	1200
	acaaagggat	atagaacctg	taaataaatc	cgtagtctg	tggtaaatg	atttttgat	1260
	aaaataacta	aaaatacaca	gtgaagaaag	aaaattattt	tcaataaatg	gtgtagacaa	1320
40	aactgactat	ccacatacag	aagaataaaa	tttgactttt	attttctct	ttatacaagc	1380
	atcaaatcaa	aattaaagt	taaatgtaa	actactacaa	ggaatatag	aaggagactg	1440
	tatgacattg	gcctgagcta	tgattttctg	tagattattc	caaaaggcaa	caaaagcaaa	1500
45	acacacaaat	gagactgcat	aaaacttaaa	acttttccac	aggaaaagaa	gcaatgatag	1560
	aattaagaga	accacaaat	gggataatat	ttttaaacca	tacatcaggt	aaggggctca	1620
	tataataata	tataagcaac	tcaacctact	caaaaataag	aaaaaaacta	tgcttattaa	1680
50	aaaataagca	aagaatcaga	atagacattt	cctacatcat	acaaaaggcc	aaccaggtac	1740
	atgaaaaaat	cataaacatt	cctaattatc	agagaagtg	aatcaatgc	cacaatgaga	1800
	tatcacctca	cacattttac	tagggctatt	ataaaaaag	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

65 <212> ADN

<213> Secuencia artificial

ES 2 344 071 T3

<220>

<223> secuencia de STAR46

5 <400> 46

```

attgtttttc tcgcccttct gcattttctg caaattctgt tgaatcattg cagttactta      60
ggtttgcttc gtctcccca ttacaaacta cttactgggt ttttcaacc tagttccctc      120
10 atttttatga tttatgctca tttctttgta cacttcgtct tgctccatct cccaactcat      180
ggccccctggc tttggattat tgttttggtc ttttattttt tgccttcttc tacctcaaca      240
cttatcttcc tctcccagtc tccgggtacc tatcaccaag gttgtcatta acctttcata      300
15 ttattcctca ttatccatgt attcatttgc aaataagcgt atattaacaa aatcacaggt      360
ttatggagat ataattcaca taccttaaaa ttcaggcttt taaagtgtac ctttcatgtg      420
gtttttggta tattcacaaa gtatgcatg gatcaccacc atctgattcc ataacatggt      480
20 caatacctca aaaagaagtc tgtactcatt agtagtcatt tcacattcac cactccctct      540
ggctctgggc agtcaactgat ctttgtgtct ctatggattt gcctagtcta ggtattttta      600
tgtaaatggc atcatacaac atgtgacctt ttgtttggct tttttcattt agcaaaatgt      660
25 tatcaaggtc tgtccctggt gtagcatgta ttagcacttc atttcttata tgctgaatga      720
tatactttat ttgtccatca gttgttcatg ctttatttgt ccatcagttg atgaacattt      780
gctgttttgc cactttgggc tattaagaat aatgctactg tgaacaagtg tgtacaagtt      840
30 cctctacaaa tttttgtgtg gacatatcct ttcagttctc tcaggtgat atctgggaat      900
tgaattgctg ggtcgtgtag tagctatggt aaacactttg agaaactgct ataatgttct      960
ccagagctgt accattttaa attctgtgta tgaggattcc acgttctcca cttcctcacc      1020
35 agtgtatgga tttgggggta tactttttaa 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 cagtgagcca 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 tacttaaatt actggaatag gtattactat tagtcacttt      1800
gcaggtgaaa gtggaacac catcgtaaaa tgtaaaatag gaaacagctg gttaatggt      1859

```

60

<210> 47

<211> 1082

<212> ADN

65

<213> Secuencia artificial

ES 2 344 071 T3

<220>

<223> secuencia de STAR47

5 <400> 47

	atcattagtc attagggaaa tgcaaatgaa aaacacaagc agccaccaat atacacctac	60
	taggatgatt taaaggaaaa taagtgtgaa gaaggacgta aagaaattgt aaccctgata	120
10	cattgatggt agaaatggat aaagttgcag cactgtgaa aaacagtctg cagtggctca	180
	gaaggttaaa tatagaacct ctggtggacc caggaactct actcttaggc accccaaaga	240
	atagagaaca gaaatcaaac agatgtttgt atactaatgt ttgtagcatc acttttcaca	300
15	ggagccaaaa ggtggaaata atccaacat cagtgaacaa atgaatgtaa taaaagcaag	360
	gtggtctgca tgcaatgcta catcatccat ctgtaaaaa cgaacatcat tttgatagat	420
	gatacaacat ggggtggacat tgagaacatt atgcttagtg aaataagcca gacacaaaag	480
20	gaatatattg tataattgta attaccatgaa 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

<210> 48

40 <211> 1242

<212> ADN

<213> Secuencia artificial

45 <220>

<223> secuencia de STAR48

50

55

60

65

ES 2 344 071 T3

<400> 48

5 atcatgtatt tgttttctga attaaitctt agatacatta atgttttatg ttaccatgaa 60
 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 ccatctcaca aaaacacata 300
 aacaaacaaa cagaaactcc acacaatgac aacgtatgtg ctttcttttt ttcttctctc 360
 ttctataata tttctttgtc ctatcttaac tgaactggcc agaaacccca ggacaatgat 420
 15 aaatacgagc agtgtaaca gacatctcat tccctttcct agcttttata aaaataacga 480
 ttatgcttca acattacata tgggtgtgtc gatggttttg ttatagataa gcttatcagg 540
 ttaagaaatt tgtctgcggt tcctagtttg gtataaagat tttaatataa atgaatgttg 600
 20 tattttatca tcttattttt ttctacatc tgctaaggta atcctgtgtt ttcccctttt 660
 caatctccta atgtggtgaa tgacattaaa ataccttcta ttgttaaaat attcttgcaa 720
 cgctgtatag aaccaatgcc tttattctgt attgctgatg gatttttgaa aaatgtgatg 780
 25 gtggacttag ttttctaagg ggaatagaat ttctaataa tttaaaatat tttgcatgta 840
 tgttctgaag gacattgggtg tgtcatttct ataccatctg gctactagag gagccgactg 900
 aaagtcacac tgccggagga ggggagaggt gctcttccgt ttctggtgtc tgtagccatc 960
 30 tccagtggta gctgcagtga taataatgct gcagtgccga cagttctgga aggagcaaca 1020
 acagtgattt cagcagcagc agtattgcgg gatccccacg atggagcaag ggaataaatt 1080
 ctggaagcaa tgacaatatc agctgtggct atagcagctg agatgtgagt tctcacgggtg 1140
 35 gcagcttcaa ggacagtagt gatgggtcaa tggcgcccag acctagaaat gcacatttcc 1200
 tcagcaccgg ctccagatgc tgagcttggga 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
 agggctctgc taaaccagaa acccaaaaca atgggagtgga agtgctaaaa ccagaaaccc 120
 55 aaaacaatgg gagtgcctcg ctacaccaga aacccaaaac gatgggagtg acgtgataaa 180
 accagacacc caaaacaatg ggagtgacgt gctaaaccag aaacccaaaa caatgggagt 240
 gacgtgctaa aacctggaaa cctaaaacaa tgcgagtgag gtgctaacac cagaatccat 300
 60 aacaatgtga gtgacgtgct aaaccagaac ccaaaacaat gggagtgacg tgctaaaaca 360
 ggaacccaaa acaatgagag tgacgtgcta aaccagaaac ccaaaacaat gggaatgacg 420
 tgctaaaacc ggaacccaaa acaatgggag tgatgtgcta aaccagaaac ccaaaacaat 480

65

ES 2 344 071 T3

5 gggaatgaca tgctaaaact ggaacccaaa acaatggtaa ctaagagtga tgctaaggcc 540
 ctacattttg gtcacactct caactaagtg agaacttgac tgaaaaggag gatttttttt 600
 tctaagacag agttttggtc tgtcccccag agtggagtgc agtggcatga tctcggtca 660
 ctgcaagctc tgctcccgg gttcaggcca ttctctgcc tcagcctct gagtagctgg 720
 gaatacaggc acccgcacc acacttggct aattttttgt attttagta gagatgggg 780
 10 ttcacatat tagcaaggat ggtctcaatc tcctgacctc gtgatctgcc cacctcaggc 840
 tcccaaagtg ctgggattac aggtgtgagc caccacacc agcaaaaagg aggaattttt 900
 aaagcaaat tatgggaggc cattgttttg aactaagctc atgcaatagg tcccaacaga 960
 15 ccaaaccaa ccaaaccaa atggagtac tcatgctaaa thtagcataa tcaaa 1015

<210> 50

<211> 2355

20 <212> ADN

<213> Secuencia Artificial

<220>

25 <223> secuencia de STAR50

<400> 50

30 caaccatcgt tccgcaagag cggcttgttt attaaacatg aaatgagggg aaagcctagt 60
 agctccattg gattgggaag aatggcaaag agagacaggc gtcattttct agaaagcaat 120
 cttcacacct gttggctctc acccattgaa tgtcctcacc caatctcaa cacagaaatg 180
 35 agtgactgtg tgtgcacatg cgtgtgcatg tgtgaaagta tgagtgtgaa tgtgtctata 240
 tgggaacata tatgtgattg tatgtgtgta actatgtgtg actggcagcg tggggagtgc 300
 tggttgaggt gtggtgtgat gtgagtatgc atgagtggct gtgtgtatga ctgtggcggg 360
 40 aggcggaagg ggagaagcag caggctcagg tgtcgcagga gaggctggga ggaaactata 420
 aacctgggca atttctctct catcagcgag ctttcttg gcaatagggg cagagctcaa 480
 agttcacaga gatagtgcct gggaggcatg aggcaaggcg gaagtactgc gaggaggggc 540
 45 agagggctctg acacttgagg ggttctaata ggaaaggaaa gaccacact gaattccact 600
 tagccccaga ccctgggcc agcgggtgcc gcttccaacc ataccaacca tttccaagtg 660
 ttgccggcag aagttaacct ctcttagcct cagtttccc acctgtaaaa tggcagaagt 720
 50 aaccaagctt accttcccgg cagtgtgtga ggatgaaaag agctatgtac gtgatgact 780
 tagaagaagg tctagggtgt gagtggact cgtctggtgg gtgtggagaa gacattctag 840
 gcaatgagga ctggggagag cctggcccat ggcttccact cagcaaggtc agtctcttgt 900
 55 cctctgact cccagcctc cagagaggac cttccaacc agcactcccc acgctgccag 960
 tcacacatag ttacacacat acaatcacat atatgttccc atatagacac attcacactc 1020
 ataccttcac acatgcacac gcatgtgcac acacagtcac tcatttctgt gttggagatt 1080
 60 gggtgaggac attcaatggg tgaggaccaa cagggtgtgaa gattgcttc tagaaaatga 1140
 ctctgtctc tctttgcat tcttccaat ccgatggagc tactaggctt ttccctcatt 1200

65

ES 2 344 071 T3

5 tcatgtttaa taaaccttcc caatggcgaa atgggctttc tcaagaagtg gtgagtgtcc 1260
 catccctgcg gtggggacag ggggtggcagc ggacaagcct gcctggaggg aactgtcagg 1320
 ctgattccca gtccaactcc agcttccaac acctcatcct ccaggcagtc ttcattcttg 1380
 gctctaattt cgctcttggt ttctttttta tttttatcga gaactgggtg gagagctttt 1440
 ggtgtcattg gggattgctt tgaaacctt ctctgcctca cactgggagc tggcttgagt 1500
 10 caactggctt 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 ctagggtggt ccacttctta 1800
 ggggccatct ccggaggtgg tgagcactaa gtgccaggaa gagaggaaac tctgttttgg 1860
 20 agccaaagca taaaaaac ttagccacaa accactgaac atttgtttg tgcaggttct 1920
 gagtccaggg agggcttctg aggagagggg cagctggagc tggtaggagt tatgtgagat 1980
 ggagcaaggg ccctttaaga ggtgggagca gcatgagcaa aggcagagag gtgtaatgt 2040
 25 ataaggtatg tcatgggaaa gagtttggtc ggaacagagt ttacagaata gaaaaattca 2100
 acactattaa ttgagcctct actacgtgct cgacattggt ctagtactg agataggttt 2160
 ggtatacaaa acaaaatcca tcctctatgg acattttagt gactaacaac aatataaata 2220
 30 ataaaagtga acaaagctc aaaacatgcc aggcactatt atttatttat ttatttattt 2280
 atttatttat tttttgaaac agagtctcgc tctgttggcc aggctggagt gtagtgggtg 2340
 gatctcggct cactg 2355

35 <210> 51
 <211> 2289
 <212> ADN
 40 <213> Secuencia Artificial
 <220>
 <223> secuencia de STAR51
 45 <400> 51

50 tcacaggtga caccaatccc ctgaccacgc tttgagaagc actgtactag attgactttc 60
 taatgtcagt cttcattttc tagctctggt acagccatgg tctccatatt atctagtaca 120
 acacacatac aatatgtgtg gatacagtat gaatataata taaaaatag tgttataata 180
 taaatataat attaaaatat gtctttatac tagataataa tacttaataa cgttgagtgt 240
 55 ttaactgctc taagcacttt acctgcagga aacagttttt tttttatttt ggtgaaatac 300
 aactaacata aattttatta caattttaag catttttaag tgtatagttt agtggagtta 360
 atatattcaa aatgttgtgc agccgtcacc atcatcagtc ttcataactc ttttcatatt 420
 60 gtaaaattaa aagtttatgc tcatttaaaa atgactccca atttcccccc tctcaacct 480
 ctggaaacta ccattctatt ttctgcctcc gtagttttgc cactctaag tacctcatat 540

65

ES 2 344 071 T3

aagtggaatt tgtcttattt gcctgtttgt gaccggctga tttcatttag tataatgtcc 600
 tcaagtttta ttcacgttat atagcatatg tcataatttt cttcactttt aagcttgagt 660
 5 aatatttcat cgtatgtatc tcacattttg cttatccatt catctctcag tggacacttg 720
 agttgcttct acatttttagc tgttgatgaat actgctgcta tgaacatggg tgtataaata 780
 tctcaagacc tttttatcag ttttttaaaa tatatactca gtagtagttt agctggatta 840
 10 tatggtaatt ttatttttaa tttttgagga actgctctac ctttttattc aatagtagct 900
 ataccaattg acaattggca ttcctaccaa cagggcataa gggttctcaa ttcctccat 960
 attccctgat acttgttatt ttcaggtgtt tttttttttt tttttttttt atgggagcca 1020
 15 tgtaaatggg tgtaaggtga tatttcatta tagttttgat ttgcatttcc ctaatgatta 1080
 gtgatgttaa gcatctcttc atgtgcctat tggccatttg tatactctct ttaaaaatat 1140
 atatatactc attcctttgc ccatttttga attatgttta ttttttgta ttgagtttca 1200
 20 atacttttct atataaccta ggtattaatc ctttatcaga ctttaagattt gcaaatattc 1260
 tctttcattc cacaggttgc taattctctc tgttggaat atcttttgat gctgttgtgt 1320
 ccagaattga ttcattcctg tgggttcttg gtctcactga cttcaagaat aaagctgcgg 1380
 25 accctagtgg tgagtgttac acttcttata gatggtgttt ccggagtttg ttccttcaga 1440
 tgtgtccaga gtttcttctt tccaatgggt tcatggtctt gctgacttca ggaatgaagc 1500
 cgcagacctt cgcagtgagg tttacagctc ttaaagggtg cgtgtccaga gttgtttgtt 1560
 30 ccccctggtg ggttcgtggt cttgctgact tcaggaatga agccgcagac cctcgcagtg 1620
 agtgttacag ctcataaagg tagtgcgac acagagtgag ctgcagcaag atttactgtg 1680
 aagagcaaaa gaacaaagct tccacagcat agaaggacac cccagcgggt tctgctgct 1740
 35 ggctcaggtg gccagttatt attcccttat ttgccctgcc cacatcctgc tgattggtcc 1800
 attttacaga gtactgattg gtccatttta cagagtgctg attggtgcat ttacaatcct 1860
 ttagctagac acagagtgct gattgctgca ttcttacaga gtgctgattg gtgcatttac 1920
 40 agtcctttag ctagatacag aacgctgatt gctgcgtttt ttacagagtg ctgattggtg 1980
 catttacaat cctttagcta gacacagtgc tgattggtg gtttttacag agtgctgatt 2040
 ggtgcgtctt tacagagtgc tgattggtgc atttacaatc ctttagctag acacagagtg 2100
 ctgattggtg cgtttataat cctctagcta gacagaaaag tttccaagt ccccacctga 2160
 45 ccgagaagcc cactggctt cacctctcac tgttatactt tggacatttg tcccccaaa 2220
 atctcatgtt gaaatgtaac ccctaagtgt ggaactgagg ccagactgga tgtggctggg 2280
 ccatgggga 2289

50

<210> 52

<211> 1184

<212> ADN

55

<213> Secuencia Artificial

<220>

60

<223> secuencia de STAR52

65

ES 2 344 071 T3

<400> 52

5 cttatgccat ctggcgggctg catgtggaac ttcgctgaag aagctaaatt tactgacat 60
 ctgtgcctag agcgggtttc tccaaggaaa ggctctgtaa atctcgtcct tttgaaatct 120
 aggggaaaac agcctccttc actgaggatt aatttaaaga aagggggaaa taggaaaatt 180
 ccatgcgttg gaagtcatt tagatttcta catgaacat catatatgtg cactacataa 240
 10 ttcctatttt tttattttta aaaaagggat aatttatatt ccagtgacaa gtttgggaaa 300
 ggccaaggca agcaattgag ttgaacatta tgtagcgtt atatagacct tgcagacgtc 360
 tgtgcaatat ccaccactga acacgtgagg tcgtactcaa gtctctctgg cccctggtaa 420
 15 tgtgactccc ttcctttatt tgcattgaat gcctggattg ggtgtcagg ttttaaaccg 480
 tcaaggttta cgcctattgt tgtcaaccaa tcagcatcct actttgacgt gattggcttc 540
 tactgtagg gtcaatcatc caaatttgc atactactcc tcaggccgcc gggagcctgt 600
 20 cagtcggctg tggcagctgg aagagaagga atcggacgga gaagaatgaa aatcacttt 660
 gctttcgsaa agcgaagaa aagtattctt ttcctcatta tttttaaata aatttgattg 720
 tatatttacc taataaaata aacattcaat taacaaaaa taagcaacta tcaaagattt 780
 25 gtttactaat tttcgtaatg tttactgttt caataagtag ccaaaggaat attaaaacac 840
 aaaaatatga atgctgataa ttttatgtca taaagacct tttaaaacta aaagtgaaca 900
 tggggtttct aaataaaatt accgtggtag cgtaaaaaca ctgctttcaa tacttgggca 960
 30 tgcctgaaagt 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

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

65

ES 2 344 071 T3

tcttgcctca gcctcagggg tacgacacat gcccatgatg agaagcagaa cgtggtgacc 540
 tttcacgaac atgggcatgg ctgctgaccc ctcgtcatca ggtgcatagc aagtgaagc 600
 5 aagtgttcac aacagtgaag agttgagcgt ctttttctt agtgtgcaa gagttcgatg 660
 ttagcgttta cgttgatatt tcttacctg tgcattctg ttagatacta acattttcat 720
 tgatgagcaa gacatactta atgcatatt tggtttgggt atccatgac ctaccttaga 780
 10 aaacaagtat tgcggttac ctctgcatgg aacagcatta cctcctctc tccccagatg 840
 tgactactga gggcagttct gagtgtttaa tttcagatt tttcctctg atttacacac 900
 acacgcacac aaaccacacc acacacacac acacacacac acacacacac acacacacac 960
 15 acacaccaag taccagtata agcatctgcc atctgctttt cccattgcca tgcgtcctgg 1020
 tcaagctccc ctactctgt ttcttggtca gcatgtactc cctcatccg attcccctgt 1080
 agcagtcact gacagttaat aaaccttgc aaacgttccc cagttggtg ctcgtgcat 1140
 20 tattgtgac acagctctgt gcacgtgtgt gcatatttct ttaggaaaga ttcttagaag 1200
 tggaaattgct gtgtcaaagg agtcatttat tcaacaaac actaatgagt gcgtcctcgt 1260
 gctgagcgt gttctaggtg ctggagcgac gtcagggaac aaggcagaca ggagttcctg 1320
 25 acccccgtc tagaggagga tgtttccagt tgttgggtt tgtttggtt tttcttctag 1380
 agatggtggt cttgctctgt ccaggctaga gtgcagtggc atgatcatag c 1431

<210> 54

30 <211> 975

<212> ADN

<213> Secuencia artificial

35 <220>

<223> secuencia de STAR54

<400> 54

40 ccataaaagt gtttctaacc tgcagaaaaa tccccctaca gtcttacagt tcaagaattt 60
 tcagcatgaa atgctggta gattacctga cttttttg caaaaataag gcacagcagc 120
 tctctctga ctctgacttt ctatagctct tactgaatta tagtcctac tgaattcatt 180
 45 cttcagtgtt gcagctgaa ggacacccac attttctctt tgcctttgct aattctttgt 240
 gttgtaaggg caggatgttt aaaagttgaa gtcattgact tgcaaaatga gaaatttcag 300
 agggcatttt gttctctaga ccatgtagct tagagcagtg ttcacctga ggttgcctg 360
 50 aatgtttctg cagttcttac caatagtatc atttaccag caacaggata tgatagagga 420
 cttcgaaaac cccagaaaat gttttgcat atatccaaag ccttttggga aatgaaagg 480
 aattgcgggc tcccattttt atatatggat agatagagac caagaaagac caaggcaact 540
 55 ccatgtgctt tacattaata aagtacaaaa tgttaacatg taggaagtct aggcgaagtt 600
 tatgtgagaa ttctttacac taattttgca acattttaat gcaagtctga aattatgtca 660
 aaataagtaa aaatttttac aagttaagca gagaataaca atgattagtc agagaaataa 720
 60 gtagcaaaat cttcttctca gtattgactt ggttgctttt caatctctga ggacacagca 780
 gtcttcgctt ccaaatccac aagtcacatc agtgaggaga ctacagctgag actttggcta 840
 atgttggggg gtcctcctg tgtctcccca ggcgcagtga gcctgcaggc cgacctcact 900
 65 cgtggcacac aactaaatct ggggagaagc aaccgatgc cagcatgatg cagatatctc 960
 aggtatgat cggcc 975

ES 2 344 071 T3

<210> 55

<211> 501

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> secuencia de STAR55

10

<400> 55

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

30 <210> 56

<211> 741

<212> ADN

<213> Secuencia artificial

35

<220>

<223> secuencia de STAR56

40 <400> 56

	acacaggaaa agatcgcaat tgttcagcag agctttgaac cggggatgac ggtctccctc	60
45	gttgcccggc aacatgggtgt agcagccagc cagttatttc tctggcgtaa gcaataccag	120
	gaaggaagtc ttactgctgt cgccgccgga gaacagggtg ttcctgcctc tgaacttgct	180
	gccgccatga agcagattaa agaactccag cgctgctcg gcaagaaaac gatggaaaat	240
50	gaactcctca aagaagccgt tgaatatgga cgggcaaaaa agtggatagc gcacgcgcc	300
	ttattgcccg gggatgggga gtaagcttag tcagccgttg tctccgggtg tcgctgctgc	360
	agttgcacgt cattctcaga cgaaccgatg actggatgga tggccgccgc agtcgtcaca	420
55	ctgatgatac ggatgtgctt ctccgtatac accatgttat cggagagctg ccaacgtatg	480
	gttatcgctg ggtatgggcg ctgcttcgca gacaggcaga acttgatggt atgcctgcga	540
	tcaatgccaa acgtgtttac cggatcatgc gccagaatgc gctgttgctt gāgcgaaaac	600
60	ctgctgtacc gccatcgaaa cgggcacata caggcagagt ggccgtgaaa gaaagcaatc	660
	agcgatggtg ctctgacggg ttcgagttct gctgtgataa cggagagaga ctgctgttca	720
65	cgttcgcgct ggactgctgt g	741

<210> 57

ES 2 344 071 T3

<211> 1365

<212> ADN

<213> Secuencia Artificial

5

<220>

<223> secuencia de STAR57

10 <400> 57

	tccttctgta aataggcaaa atgtatttta gttccacca cacatgttct tttctgtagg	60
15	gcttgatgt tggaaatatt atccaattat tcaattaaca ctataccaac aatctgctaa	120
	ttctggagat gtggcagtga ataaaaagt tatagtttct gattttgtgg agcttgact	180
	ttaatgatgg acaaaacaac acattcttaa atatatattt catcaaaatt atagtgggtg	240
	aattatttat atgtgcattt acatgtgtat gtatacataa atgggcggtt actggctgca	300
20	ctgagaatgt acacgtggcg cgaacgaggc tgggcggtca gagaaggcct cccaaggagg	360
	tggctttgaa gctgagtggg gcttccacgt gaaaaggctg gaaagggcat tccaagaaaa	420
	ggctgaggcc agcgggaaag aggttccagt gcgctctggg aacggaaagc gcacctgcct	480
25	gaaacgaaaa tgagtgtgct gaaataggac gctagaaagg gaggcagagg ctggcaaaag	540
	cgaccgagga ggagctcaaa ggagcgagcg ggggaagccg ctgtggagcc tggaggaagc	600
	acttcggaag cgcttctgag cgggtaaggc cgctgggagc atgaactgct gacgaggtgt	660
30	gtccagaatt cgtgggttct tgggtctact gacttcaaga atgaagagg acccgggacc	720
	ctcgcggtga gtgttacagc tcttaagggtg gcgctctggg agtttgttcc ttctgatgtt	780
	cggatgtgtt cagagtttct tccttctggt gggttcgtgg tctcgtggc tcaggagtga	840
35	agctgcagac cttcgcggtg agtgttacag ctcataaaag cagggtggac tcaaagagtg	900
	agcagcagca agatttattg caaagaatga aagaacaaag cttccacact gtggaagggg	960
	acccagcgg gttgccactg ctggctccgc agcctgcttt tatttcttta tctggcccca	1020
40	cccacatcct gctgattggt agagccgaat ggtctgtttt gacggcgctg attggtgcgt	1080
	ttacaatccc tgcgctagat acaaaggttc tccacgtccc caccagatta gctagataga	1140
	gtctccacac aaaggttctc caaggcccca ccagagtagc tagatacaga gtgttgattg	1200
45	gtgcattcac aaacctgag ctagacacag ggtgatgact ggtgtgttta caaaccttgc	1260
	ggtagataca gagtatcaat tggcgtattt acaatcactg agctaggcat aaaggttctc	1320
	caggtcccca ccagactcag gagcccagct ggcttcccc agtgg	1365

50

<210> 58

<211> 1401

<212> ADN

55 <213> Secuencia artificial

<220>

<223> secuencia de STAR58

60

65

ES 2 344 071 T3

<400> 58

	aagtttacct tagccctaaa ttatttcatt gtgattggca ttttaggaaa tatgtattaa	60
5	ggaatgtctc ttaggagata aggataacat atgtctaaga aaattatatt gaaatattat	120
	tacatgaact aaaatgtag aactgaaaa aaattattgt aactcctcc agcgtaggca	180
	ggagtatcta gataccaact ttaacaactc aactttaaca acttcgaacc aaccagatgg	240
10	ctaggagatt cacctattta gcatgatatc ttttattgat aaaaaatat aaaacttcca	300
	ttaaattttt aagctactac aatcctatta aattttaact taccagtgtt ctcaatgcta	360
	cataatttaa aatcattgaa atcttctgat ttaactcct cagtcttgaa atctacttat	420
15	ttttagttac atatatatcc aatctactgc cgctagtaga agaagcttg aatttgagaa	480
	aaaaatcaga cgtttgtat attctcatat tctaataatt attttttaa tgagtttctg	540
	caatgcatca agcagtggca aaacaggaga aaaattaaaa ttggttgaaa agatatgtgt	600
20	gccaacaat cccttgaaat ttgatgaagt gactaatcct gagttattgt ttcaaatgtg	660
	tacctgttta tacaagggtg tcaccttga aatctcaaca ttaaatgaaa tttataagc	720
	aatttgttgt aacatgatta ttataaaatt ctgatataac attttttatt acctgtttag	780
25	agtttaaaga gagaaaagga gttaagaata attacatctt cattagcatt gtccgggtgc	840
	aaaaacttct aacactatct tcaaatctt ttctccattg ccttctgaac ataccactt	900
	gggtatctca ttagcactgc aaattcaaca ttttcgattg ctaatttttc tcctaataa	960
30	tttatttgtt ttctcagctt tagccaatgt ttcactattg accatttct caagtatagt	1020
	gacgcttcaa tgacctcag agagctgtt cagtcctcc tggactactt gcatgcttcc	1080
	aacaaaatga agcactctt atgtcagtca ctcaataaa tggaaatggg cccatttact	1140
35	aggaatgta acagaataaa aagatagacg tgacaccagt tgcttcagtc catctccatt	1200
	tacttgctta aggctggcc atatttctca cagttgatat ggcgcagggc acatgtttaa	1260
	atggctgttc ttgtaggatg gtttgactgt tggattcctc atcttccctc tccttaggaa	1320
40	ggaaggttac agtagtactg ttggctcctg gaatatagat tcataaagaa ctaatggagt	1380
	atcatctccc actgctcttg t	1401

<210> 59

45 <211> 866

<212> ADN

<213> Secuencia artificial

50 <220>

<223> secuencia de STAR59

<400> 59

	gagatcacgc cactgcactc cagcctgggg gacagagcaa gactccatct cagaaacaaa	60
	caaacacaca aagccagtca aggtgtttaa ttcgacggtg tcaggctcag gtctcttgac	120
60	aggatacatc cagcaccggg gggaaacgtc gatgggtggg gtggaatcta ttttgtggcc	180
	tcaaggagg gtttgagagg tagtcccga agcgggtgat gcctaaggaa gccctccgc	240

65

ES 2 344 071 T3

ccaagaagcg atattcattt ctagcctgta gccacccaag agggagaatc gggctcgcca 300
cagaccccac aacccccaac ccaccccacc cccacccctc ccacctcgtg aaatgggctc 360
5 tcgctccgtc aggctctagt cacaccgtgt ggttttggaa cctccagcgt gtgtgctggtg 420
gttgctggtt ggggtggggc cggctgtgga cagaggagg gataaagcgg cgggtgcccg 480
cgggtgcccg ggacgtgggg cgtggggcgt ggggtggggtg gccagagcct tgggaactcg 540
10 tcgctgctcg ggacgtctcc cctcctggtc ccctctctga cctacgctcc acatcttcgc 600
cgttcagtgg ggaccttggt ggtggaagtc accatccctt tggactttag ccgacgaagg 660
ccgggtccc aagagtctcc ccggaggcgg ggccttgggc aggtcaciaa ggatgctgac 720
15 ggtgacggtt ggtgacggtg atgtacttcg gaggcctcgg gccaatgcag aggtatccat 780
ttgacctcgg tgggacaggt cagctttgctg ggtcccgtg cgtccttcca gagactcatc 840
cagcgttagc aagcatggtc ccgagg 866

20 <210> 60

<211> 2067

<212> ADN

25 <213> Secuencia artificial

<220>

<223> secuencia de STAR60

30

<220>

<221> característica_misc

<222> (92)..(1777)

35 <223> Todos los "N" en diferentes posiciones representan cualquier ácido nucleico

40

45

50

55

60

65

ES 2 344 071 T3

<400> 60

	agcagtgag aactggggaa gaagaagagt ccctacacca cttaatactc aaaagtactc	60
5	gcaaaaaata acaccctca ccagggtggca tnattactct ccttcattga gaaaattagg	120
	aaactggact tcgtagaagc taattgcttt atccagagcc acctgcatac aaacctgcag	180
	cgccacctgc atacaaacct gtcagccgac cccaaagccc tcagtcgcac caagcctctg	240
10	ctgcacaccc tcgtgccttc aactggccg ttccccaagc ctggggcata ctntcccagct	300
	ctgagaaaatg tattcatcct tcaaagccct gctcatgtgt cctnntcaac aggaaaatct	360
	cccatgagat gctctgctat ccccatctct cctgccccat agcttaggca nacttctgtg	420
15	gtgggtgagc ctgggctgtg ctgtgatgtg ttcgcctgcn atgtntgttc ttccccacaa	480
	tgatggggcc ctgaattctc tatctctagc acctgtgtc agtaaaggct tgggaaacca	540
	ggctcaaagc ctggcccaga tgccacctt tccagggtgc ttccgggggc caccaaccag	600
20	agtgacagct tctctccac caggaactct tgcagcccca cccctgagca cctgcacccc	660
	attacccatc tttgtttctc cgtgtgatcg tattattaca gaattatata ctgtattctt	720
	aatacagtat ataattgtat aattattctt aatacagtat ataattatac aatacaaaa	780
25	tatgtgttaa tggaccgttt atgttactgg taaagcttta agtcaacagt gggacattag	840
	ttaggttttt ggcgaagtca aaagttatat gtgcattttc aacttcttga ggggtcggta	900
30	cntctnacc ccatgttgtt caanggtcaa ctgtctacac atatcatagc taattcacta	960
	cagaaatgtt agcttgtgtc actagtatct ccccttctca taagcttaac acacatacct	1020
	tgagagagct cttggccatc tctactaatg actgaagttt ttatttatta tagatgtcat	1080
35	aataggcata aaactacatt acatcattcg agtgccaatt ttgccacctt gacctcttt	1140
	tgcaaaacac caacgtcagt acacatatga agaggaaact gcccgagaac tgaagttcct	1200
	gagaccagga gctgcaggcg ttagatagaa tatggtgacg agagttacga ggatgacgag	1260
40	agtaaatact tcatactcag tacgtgccaa gcaactgctat aagcgtctg tatgtgtgaa	1320
	gtcatttaat cctcacagca tcccacggtg taattatfff cattatcccc atgaggggaa	1380
	agaaactcag aacggttcaa cacatatgag agaagtcgca gccggtcagt gagagagcag	1440
45	gttcccgtcc aagcagtcag accccgagtg cacactctcg acccctgtcc agcagactca	1500
	ctcgtcataa ggcggggagt gntctgtttc agccagatgc tttatgcatc tcagagtacc	1560
	caaaccatga aagaatgagg cagtattcan gagcagatgg ngctgggagc taaggctggg	1620
50	cttcagaata gctggaaagc tcaagtnatg ggacctgcaa gaaaaatcca ttgttngat	1680
	aaatagccaa agtccctagg ctgtaagggg aagggtgtgccc aggtgcaagt ggagctctaa	1740
	tgtaaaatcg cacctgagtc tcctggtctt atgagtnctg ggtgtacccc agtgaaaggt	1800
55	cctgctgcca ccaagtgggc catggttcag ctgtgtaagt gctgagcggc agccggaccg	1860
	cttctcttaa cttcacctcc aaaggcacag tgcacctggt tcctccagca ctcagctgcy	1920
	aggcccctag ccagggtccc ggccccggc ccccgccagc tgetccagct tccttcccca	1980
60	cagcattcag gatggctctg gttcatgtag acctttgttt tcagtctgtg ctccgaggtc	2040
	actggcagca ctagccccgg ctctgt	2067

<210> 61

65 <211> 1470

<212> ADN

<213> Secuencia Artificial

ES 2 344 071 T3

<220>
 <223> secuencia de STAR61

5 <220>
 <221> característica_misc
 <222> (130)..(976)
 <223> "N" representa cualquier ácido nucleico en diferentes posiciones

10 <400> 61

```

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 gcaccccagg cacggccagc 240
ctggggtatg gacccggggc cntctgttct agagcaggaa ggtatggtga ggacctcaaa 300
aggacagcca ctggagagct ccaggcagag gnacttgaga ggcctgggg ccatcctgtc 360

tcttttctgg gtctgtgtgc tctgggcctg ggccttcct ctgctcccc gggcttggag 420
agggctggcc ttgcctcgtg caaaggacca ctctagactg gtaccaagtc tggcccatgg 480
cctcctgtgg gtgcaggcct gtgcgggtga cctgagagcc agggctggca ggtcagagtc 540
aggagagggga tggcagtgga tgccctgtgc aggatctgcc taatcatggt gaggctggag 600
gaatccaaag tgggcatgca ctctgcactc atttctttat tcatgtgtgc ccatcccaac 660
aagcagggag cctggccagg agggcccctg ggagaaggca ctgatgggct gtgttccatt 720
taggaaggat ggacggttgt gagacgggta agtcagaacg ggctgcccac ctcggccgag 780
agggccccgt ggtgggttgg caccatctgg gcctggagag ctgctcagga ggctctctag 840
ggctgggtga ccaggnctgg ggtacagtag ccatgggagc aggtgcttac ctggggctgt 900
ccctgagcag gggctgcatt ggggtctctg tgagcacaca cttctctatt cacctgagtc 960
ccnctgagtg atgagnacac ccttgttttg cagatgaatc tgagcatgga gatgttaagt 1020
ggcttgcttg agccacacag cagatggatg gtgtagctgg gacctgaggg caggcagtc 1080
cagcccgagg acttcccaag gttgtggcaa actctgacag catgaccca ggaacaccc 1140
atctcagctc tggtcagaca ctgaggagtt gtgttgtaac ccacacagct ggagacagcc 1200
accctagccc cacccttatc ctctcccaaa ggaacctgcc ctttccctc attttcctc 1260
tactgcattg agggaccaca cagtgtggca gaaggaacat gggttcagga cccagatgga 1320
cttgcttcac agtcagccc tcctgtctc ttgcagagtg cgtcttcac tgtgaagttg 1380
ggacagtcac accaactcaa tactgctggg cccgtcacac ggtgggcagg caacggatgg 1440
cagtactggt ctgtgggtct gcagaggtgg 1470

```

<210> 62
 <211> 1011

60 <212> ADN
 <213> Secuencia artificial

<220>
 65 <223> secuencia de STAR62

ES 2 344 071 T3

<400> 62

	agtg tcaaat agatctacac aaaacaagat aatgtctgcc catttttcca aagataatgt	60
5	gg tgaagtgg gtagagagaa atgcatccat tctccccacc caacctctgc taaattgtcc	120
	atgtcacagt actgagacca gggggcttat tcccagcggg cagaatgtgc accaagcacc	180
	tcttgtctca atttgcagtc taggccctgc tatttgatgg tgtgaaggct tgcacctggc	240
10	atggaaggtc cg tttgtac ttcttgcttt agcagttcaa agagcagggg gagctgcgag	300
	ggcctctgca gcttcagatg gatgtgg tca gcttggtgga ggcgccttct gtggtccatt	360
	atctccagcc cccctgcggt gttgctg ttt gcttggttg tctggctctc catgccttgt	420
15	tggctccaaa atgtcatcat gctgcacccc aggaagaatg tgcaggccca tctcttttat	480
	gtgctttggg ctattttgat tccccgttg gtatattccc taggtaagac ccagaagaca	540
	caggaggtag ttgctttggg agagtttggga cctatgggta tgaggtaata gacacagtat	600
20		
	cttctctttc atttggtgag actg ttagct ctggccgcgg actgaattcc acacagctca	660
	cttgggaaaa ctttattcca aaacatagtc acattgaaca ttgtggagaa tgagggacag	720
25	agaagaggcc ctagatttgt acatctgggt gttatgtcta taaatagaat gctttggtgg	780
	tcaactagac ttgttcatgt tgacatttag tcttgcc ttt tccggtggtga tttaaaaatt	840
	atgtatatct tg tttggaat atagtg gagc tatggtgtgg cattttcatc tggctttttg	900
30	tttagctcag cccgtcctgt tatgggcagc cttgaagctc agtagctaat gaagaggat	960
	cctcactccc tccagagagc ggtcccc tca cggctcattg agagtttgtc a	1011

35 <210> 63

<211> 1410

<212> ADN

<213> Secuencia Artificial

40

<220>

<223> secuencia de STAR63

45

50

55

60

65

ES 2 344 071 T3

<400> 63

	gcgtctgagc cgctgggaac ccatgagccc cgtccatgga gttgaggaag ggggttcgcc	60
5	ccacgggggtg ggcgccctct acacagcgcg cttcctcttc tctcgttagc gccgcgggac	120
	cagcctctgg ttctgcacct cgcgctctgg gagcagcgcc cggctttggc gagcgcttcc	180
	ccggggctgc ccagcctctg ctccgctcgc cccgccaggc ccggctccgc gaagccccc	240
10	gggtccagtc caaggccccg attccccaaag gccaggggccc cggggcagca ttggaacagg	300
	gcgcggacgc cagtcctccg agcatggagt aactgcagct tttgagaaaa gaaagcggac	360
	cccccccat cgagaacgcg ggccttggt tagggacggt cctgggccgt cacggagtgt	420
15	cgccggctcc tcggcccctc cctcctcaa gccccacc cgcacagcgg cctccctggg	480
	gacctcccct cgggctgcgc tttcagccca aacacagggg ggtcttccag gagcctgccc	540
	agtccccaca gcagcccaga gacccccact cccacctgta cctgccaagc cttcagagag	600
20	ggcggcctgg acatgccccg cacgggagga gccccgctc agcaccctg caagtggcag	660
	caaccagaa caccctgag aggcctctga gcagcccagg aagtggctgg aagacgcata	720
	ggcagctcac tcctctgtaa gagcaaggac cggagaacac atgctgacct ctgcttttgc	780
25	agaggggca tgcttcagga caggcgcgct cagcaggtgt ccatcttatt tcacacctt	840
	gtgtttatat catcttatt tgcattttat gtctaattaa caatatgcag ctggccaggc	900
	gcagtggctc aagcctctaa tcccagcact ttgggaggcc gaggcaggtg tatcacttga	960
30	gggcaggagt tcgggaccgg cctgggcaac atagcaaac cccattgcta ataaaaatac	1020
	aaaaattagc cagccatggt ggcgggcacc tgcaagccca gctactccgg aagctgaagc	1080
	aggagaaatca cttgaaccca ggaggcggag gtggcagtg gctatcaagc cattacactc	1140
35	cagcctgggc aacagagaaa gactgtctca aaaaaaatt aatacgcagc agaatattat	1200
	gtggtcagcc caagcagtc cccccactca gccctctgtc cctacagctc caggcactcc	1260
	cccagcccct cccctggaca agaggtaatg cccagagggt gaaaatccac caaggttaag	1320
40	 	
	ccagaaacaa aaagctcaaa gcttcggcat ctccctccgc tcagaccctt agagcagatt	1380
	cctctcatcg acagcacgat caggctgtgg	1410

45 <210> 64

<211> 1414

<212> ADN

50 <213> Secuencia Artificial

<220>

<223> secuencia de STAR64

55

60

65

ES 2 344 071 T3

<400> 64

	agagatcttt taagggctca aaagaccctg cggctcccct gccaatagct ctgccatcgt	60
5	ccccagagct ttcgaggacc ctccaccatc ggcgccaacc ccagctgagc tgggtgctcg	120
	tctgcaggcc tctgtccat ctcagcctga 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 tgccttcag cctggcaaag tggcaaggaa cctctgggtg ggaaaacaaa	540
	tcaacaaca aattgtcagt aaaaaacaga aacctcacac tttcctttct cttgacctct	600
20	tgaaaaagc aaatccactg cagctcacca aaggcaaaga gaaaacctta agaataccca	660
	gagagaaaag acacgttact tgcaaaagaa catctaagc agggagataa tgaaaataca	720
	gactcttcaa agggctgaag gaaaaaaccc gtccacctag aattctatcc ccaaactgtc	780
25	atctgagagc aagggcaaaa caaacgcttt ctcagacagg ctggacgagg tcgctcacgc	840
	ctgtaatcct agcactttgg gaggccaagg tgggaggacc gctttaagcc agaagtttga	900
	gaccagtgtg ggtaacataa tgagaccca 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 taaaatgac tcatttgggg ggttaaaca ctgttgaact aaaatcctgg	1200
	ataacagcag catgaaaggt ggggtggtgg tcccaggaaa gcattcaaag gtccatgtct	1260
	catttgggag gagggtaggg agactcatga acttgaggct cccttcaggc aagcacagt	1320
40	caaaaaatt ataataatgg gaaacagata cagtagactg tgatgtacaa ctctcagagc	1380
	agtagaaggg aggtataaa acaaatctga tcca	1414

<210> 65

45 <211> 1310

<212> ADN

<213> Secuencia artificial

50 <220>

<223> secuencia de STAR65

55

60

65

ES 2 344 071 T3

<400> 65

t**cgagaccag cctggccaac atggtgaaac attgtctcta ttacaaatac aaaaattagc** 60
 5 c**aggtgtagt ggtgcatgcc tgcagtccca gccatttggg aggctgaagt tggagaatcg** 120
 ct**taaacctg ggaggtggag gttgcattga gccgagaagc actccagcct ggatgacgga** 180
 g**caagactgt ctcaaaaaga aaaaaaaaaag aagcagcagc aaatatccct gtccctgatgg** 240
 10 ag**gctatata acaaccaaac aagtgaatgc ataagacaat ttcaaggta tggtagatac** 300
 ca**taagtggg agatgaacaa tgagaacaca tggacacagg gaggagaaca tcacacactg** 360
 g**ggcctctcg gggggtgggg aaataggggg tgatagcatt aggagaaata cctaattgctg** 420
 15 at**aacagggtt agtgggtgca gcaaaccacc atggcacgtg tataatctatg taacacacct** 480
 g**cacgttctg cacatgtatc ccagaactta aagtataata aaaaaagaca ttaaaaaatt** 540
 at**gatataaa atcccaattc aagttgtttt aaaaagagaa aacaattatc tttatataat** 600
 20 ag**cggaaaat atagatggcg gaattaaagc ctcgtcatat tttctaacag aactttctga** 660
 ta**aaacttgat taaataaaaa ttttaaatat cactaaacac atagaagaaa taaatttaaa** 720
 ct**ttcacaaa aaataaagta caatgaatga agacaagggtg tacttgaaaa aagaactgaa** 780
 25 ta**aatattct acatataaaa aaaatctgat gatattgtgg tgattcttta ctttgctact** 840
 ag**tttctctt tttttcttct gaaaaatttc ttgggatgta tttggtttca ttagtaaaat** 900
 t**ctaagttc tttgcaatct gaacattgga gcttcatcca tagccagtat gccctaacat** 960
 30 ta**tcttttga caactgtaaa attagaacac tgccagacat atttaatgta tgatgtatat** 1020
 ca**actggtg acacatttta tactatcttt attccaaaat caaatgattc actgtggttt** 1080
 at**aatgtac atggatatat ctctacctaa gcagatagtt aggagagtta gtaaaaatga** 1140
 35 gg**tggaaaat aggagtcact gtcccttcac agggagagaa ttctgctttt ctcctaatat** 1200
 ac**cttttgct tgaacagact ccaaccctc atcttttgtc ctttaaatga ccacatttat** 1260
 tt**taactttg ataaacaaca cagaaagata tttgatccat caacattcac** 1310

40 <210> 66

<211> 2500

<212> ADN

<213> Secuencia artificial

45

<220>

<223> secuencia de T2F (STAR66F)

50 <400> 66

g**caggttga tgggtctgac ccctcctcgg gttggttcc tgtctccagg tggacgtcct** 60
 g**tactccagg gtctgcaagc ctaaaaggag ggaccaggga cccaccacag acccgctgga** 120
 55 c**ccaagggc cagggagcga ttctggccct ggcgggtgac ctggcctacc agaccctccc** 180
 g**ctcagggcc ctggatgtgg acagcggccc cctggaaaac gtgtatgaga gcatccggga** 240
 g**ctggggggac cctgctggca ggagcagcac gtgctgggct gggacgcccc ctgcttccag** 300

60

65

ES 2 344 071 T3

	ctgccccagc ctagggaggg gctggagacc cctccctgcc tccctgccct gaacactcaa	360
	ggacctgtgc tccttcctcc agagtgaggc cegtcccccg ccccgccccg cctcacagct	420
5	gacagcgcca gtcccaggtc cccgggctgc cagcccgtga ggtccgtgag gtcctggccg	480
	ctctgacagc cgcggcctcc cccgggctcca gagaaggccc gcgtctaaat aaagcgccag	540
	cgcaggatga aagcggccag cctcgcagcc tgctcttctt gaaagctggg cgggttgggg	600
10	cggggggctt ctctggaagg cttggagctg tcccctctgg ccttggggga ctggctgccc	660
	ccggggcgcc cgggcctagc cgaggcggtg ctcctgccgg ccagactctc ggtcagtgcg	720
	ggcacggggt cccagccact cctagggggc agcgcagccg gcaggggtggc cgccccggg	780
15	tgggacttgg accctggact ccacgggagg gctccgccac ccagcctggt gttacataag	840
	gggtggtgga ggtgggcagt cgagcgttaa agagtaacct gctgccggga agcccccaa	900
	gcaatcgcgg ccccttcccc ggctctggca gctctgcgag cgcgcccggt gggaaacggg	960
20	cctccccggc ggggcgcgcg ggcgcgcgag gtgggcggag gcctcggagc tgtgccgggc	1020
	cgggcctccc tccctaggcc agcgcgggag cgaccgggag ggggcgggcc cggggcgggg	1080
	cctcgaagcg ctggccggcg ggagcgcggc cggccgggcc cgcccgcctg cgggtgtggac	1140
25	gccgcgcggc caatgcgcgc gccgggacgg gacgggacgg ggcggggcgg ggcgggacga	1200
	gacggggcgg ggcggggcgg gccgggcagc ctccgggcgg cgcggcgcgg gcggcggccg	1260
	gatccagggc gggggtcggc ggcccggcca gccggccccg gccggggcc gcgtcctgag	1320
30	agtcagccct cgccgctgca gcctcggcgc cggccggcc gcccatggag cgccccccgc	1380
	cccgcgccgc cggccgggac cccagtgcgc tgcgggccga ggcgccgtgg ctgcgcgcgg	1440
	agggtcggg gccgcgcgc cgcgccgtga cgggtccac gccgccgag gtaccgggcg	1500
35	ccggtgggcg ggggcgccga ccaagtctt ctcgctgcaa agatggcgtc agtgctgcc	1560
	aaacttcggg cccccggggg cggggcagcg gggagggcgg ccgcgtcggc ccgcgcgtgt	1620
	ccgtgggtcc cgccggggct ggcgccggcg gccggggagc ccttccccgc gcgccgggt	1680
40	gggggcgggg ccgggggcgg ggccgcgccg tccacaccgg ccgcagccgg ttttcgaggc	1740
	gggcgccgag cggatcccg gcggagggtg agggaccccc ctccccggc caccgcctcc	1800
	gctgagtctg cccccctccc atccgcagg ctcttccgtg ggcggcggct tcgcgggctt	1860
45	ggagttcgcg cggccgcagg agtcggagcc gcgggcctcg gacctggggg cccccggac	1920
	gtggacgggg gcggcggcgg ggccccggac tccgctggcg cacatccccg tcccagcgca	1980
	gaggtgagcg ggaggcccgg tgccctcggga ctcggtgtgc gcaggggcgg tgggtgggt	2040
	gcggagacac cggccccgac ggaggccagg tcagggcccc aggtttgtaa ttaccagcca	2100
50	cccccaagct cttcagccct ggaggagctg agcagaaatg atcgatgact gggagtccct	2160
	acacctccct ccaccgcagt tcctcggggc tagagctcag aaccgggagc ggggtggctgt	2220
	gcgtctctgt gcagaagagg ctgcgcggtc ggcattggggc gactgtccag gaatccctgg	2280
55	ggctcctgac cgccacctcc caaccctgc caggccggac acctcggctt ggctgccagg	2340
	gcaggggcgg gccctggcct ggctcgtctg ggcctgggga gctgcccgct ctccagccc	2400
60	agtctcccc tggctgctgc cggctgctgg ccactccac ctcccaggcc tggcgtgagg	2460
	cccacagctg ctggtgcaca accctggtta atgtgtgatg	2500

65 <210> 67
 <211> 2500
 <212> ADN

ES 2 344 071 T3

	gaggactccc tattcattaa atcaactcaa gacggaccaa aaacctaataat 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 ggcctataa ggaacttaaa caaatttaca aacaaaaaac aaacaacact	1920
	attacgagtt ggaaaaggac atgaatcgac acttttcaaa agaagacata catgtggcta	1980
	acaagcatat gaaaaaaatg ctcaacatta ctaatcatta gagaaatgca aatcaaaacc	2040
15	acaatgagat accatctcaa ccagctctgaa tggctgttat taaaaaatc agaaaaaac	2100
	agatgctggc aaggttgtgg 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 aagacatggt	2340
	atcaacctaa atgccatta acagtaaact ggataaggaa aatatggtac atatacactg	2400
25	tggaatacta tgcagtcata aaaagaatga gataatgttc attgcagcaa catggatgga	2460
	actggagacc attatccttg ggaactaac aaagcaacag	2500

<210> 68

30 <211> 13

<212> ADN

<213> Secuencia Artificial

35 <220>

<223> secuencia loxP

<400> 68

40

ataacttctg ata

13

<210> 69

45 <211> 13

<212> ADN

<213> Secuencia artificial

50 <220>

<223> secuencia FRT

<400> 69

55

gaagttccta tac

13

<210> 70

60 <211> 631

<212> ADN

<213> Secuencia artificial

65 <220>

<223> elementos STAR en STAR3 directo

ES 2 344 071 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 aaaaataggg gtatcacgag      60
      gccctttcgt cttcactcga gcggccagct tggatctcga gtactgaaat aggagtaaat      120
      ctgaagagca aataagatga gccagaaaac catgaaaaga acagggacta ccagttgatt      180
15      ccacaaggac attcccaagg tgagaaggcc atatacctcc actacctgaa ccaattctct      240
      gtatgcagat ttagcaaggt tataaggtag caaaagatta gacccaagaa 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 gagagaggtt ctaaatacat tggctcctca taggcaaacc gcagttcact      540
25      ccgaacgtag tccctggaaa tttgatgtcc agnatagaaa agcanagcag ncnhnnnnta      600
      tanatnnngn tganccanat gntnctggn c      631
  
```

<210> 71

30 <211> 774

<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 cgcgccaaagc ttggatcccc ccccgcccc tccgcectcg agccccgccc      60
      cttgccttag aggccctgcc gaggggcggg gcctgtccct cctccccctt cccccgcccc      120
50      ctaccgtcac gctcaggggc agcctgaccc cgagcggccc cgcggtgacc ctgcgcaga      180
      ggctgtggg aggggcgtcg caagcccctg aatccccccc cgtctgttcc cccctcccgc      240
      ccagtctcct ccccctggga acgcgcgggg tgggtgacag acctggctgc gcgccaccgc      300
55      caccgcgcct gccgggggcg ctgccgctgc ctgagaaact gcggtgccg cctggaggag      360
      gtgccgtcgc ctccgccacc gctgccgcc cgcagggg taggagctaa gccgccgcca      420
      ttttgtgtcc ccctgttgtt gtcgttgaca tgaatccgac atgacactga ttacagccca      480
60      atggagtctc attaaaccgg agtcgcggtc ccgccccgcc gctgctccat tggaggagac      540
      caaagacact taaggccacc cgttggccta cgggtctgtc tgtcaccac tactaacca      600
      ctctgcagcc cattggggca ggttcctgcc ggtcatntcg cttccaataa acacaccct      660
65      tcgaccccat nattcccccc cttcgggaac cccccggg ggaggggtcc actggncaat      720
      accaattnaa nagaaccgct ngggtccgcc tntttncggg cncctattg gggt      774
  
```

ES 2 344 071 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
 aaragcagtg catgtgctgt aacaacctca gctgggaagc agtatatgtg gcgttatgac 180
 25 ctcagctgga acagcaatgc atgtggtggt gtaatgacct cagctgggta gggtgcatgt 240
 gatggaacaa cctcagctgg gtagcagtggt acttgataaa atgttggcat actctacatt 300
 tgttatgagg gtagtgccat taaatttctc cacaaattgg ttgtcacgta tgagtgaaaa 360
 30 gaggaagtga tgggaagactt cagtgcctttt ggcctgaata aatagaagac gtcattttca 420
 gtaatggaga caggggaagac taangnaggg tggattcagt agagcaggtg ttcagttttg 480
 aatatgatga actctgagag aggaaaaact ttttctacct cttagttttt gngnctggac 540
 35 ttaanattaa aggacatang acngaganca gaccaaatnt gcgangtttt tatatnttac 600
 ttgcnagggg aatttncaag aaaaagaaga cccaananc attggtcaaa actatntgcc 660
 ttttaaaaaa aaganaatta caatgganan anaagtgttg nctnggcaaa aattggg 717

 40 <210> 73
 <211> 541
 <212> ADN
 45 <213> Secuencia artificial

 <220>
 <223> elementos STAR en STAR4 inverso

 50 <220>
 <221> característica_misc
 <222> (6)..(533)
 55 <223> "N" representa cualquier ácido nucleico en diferentes posiciones

 60

 65

ES 2 344 071 T3

<400> 73

```

    ggattngagc tagcggcgcg ccaagcttgg atcttagaag gacagagtgg ggcatggaaa      60
    tgcaccacca gggcagtgca gcttggtcac tgccagctcc nctcatgggc agagggctgg      120
5    cctcttgagc ccgaccaggc actgagcgcc atcccagggc cctcgccagc cctcagcagg      180
    gccaggacac acaagccttt gacttcctcc tgtcactgct gctgccattc ctgtttttgtg      240
    gtcatcactc ctccctgtc ctcagactgc ccagactca aggatgtcct gtgggtggcat      300
10   cagaccatat gccctgaan aggagtgagt tgggtgtttt tgccgcgccc anagagctgc      360

    tgtcccctga aagatgcaag tgggaatgat gatgntcacc atcntctgac accaagcctt      420
    ttggatagag gccccaacag tgaggatggg gctgcactgc attgccaagg caactctgtn      480
    ntgactgcta cangacantic ccaggacctg ngaagnncta tanatntgat gcnaggcacc      540
20   t                                                                                   541

<210> 74
<211> 794
25 <212> ADN
    <213> Secuencia artificial

<220>
30 <223> elementos STAR en STAR6 directo

<220>
<221> característica_misc
35 <222> (374)..(792)
    <223> "N" representa cualquier ácido nucleico en diferentes posiciones

<400> 74
40   ccaccacaga catcccctct ggcctcctga gtggtttctt cagcacagct tccagagcca      60
    aattaaacgt tcaactctatg tctatagaca aaaagggttt tgactaaact ctgtgtttta      120
    gagagggagt taaatgctgt taacttttta ggggtgggcg agaggaatga caaataacaa      180
45   cttgtctgaa tgttttacat ttctccccac tgcctcaaga aggttcacaa cgaggtcadc      240
    catgataagg agtaagacct cccagccgga ctgtccctcg gccccagag gacactccac      300
    agagatatgc taactggact tggagactgg ctcacactcc agagaaaagc atggagcacg      360
50   agcgcacaga gcangggcca aggtcccagg gacnfaatgt ctaggagggg gattgggggtg      420
    aggtantct gatgcaatta ctgngcagct caacattcaa gggaggggaa gaaagaaach      480
    gtcctgtaa gtaagttgtn cancagagat ggtaagctcc aaatttnaac tttggctgct      540
55   ggaagtttn ngggccnana naanaaacan aanatttga ggttanacc cactaaccn      600
    tatnantant tattaatacc cctaattana ccttggatan ccttaaaata tcntntnaaa      660
    cggaaccctc nttcccnitt nnaaatnnaa aaggccattn ngnncnagta aaaatctnnn      720
60   ttaagnnttg ggccnaaca aacntnttcc nagacacntt tttntccng gnattntaa      780
    tttatttcta ancc                                                                                   794

<210> 75
65 <211> 379
    <212> ADN

```

ES 2 344 071 T3

<213> Secuencia Artificial

<220>

5 <223> elementos STAR en STAR6 inverso

<220>

<221> característica_misc

10 <222> (287)..(379)

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

<400> 75

```

15      atcgtgtcct ttccagggac atggatgaag ctggaagcca tcatcctcag caaactaaca      60
      caggaacaga aaaccaata ccacatgttc tactcataa gtgggagctg aacagtgaga      120
      acacatggac acagggaggg gaacatcaca caccaaggcc tgtctggtgt ggggagggga      180
20      gggagagcat caggacaaat agctaatagca tgtggggctt aaacctagat gacgggttga      240
      taggtgcagc aatccactat ggacacatat acctatgtaa caaccnacc ttnttgacat      300
      gtatcccaga acttaaagga aaataaaaat taaaaaaat tncctggaa taaaaaagag      360
25      tgtggacttt ggtgagatn                                     379
  
```

<210> 76

<211> 398

30 <212> ADN

<213> Secuencia artificial

<220>

35 <223> elementos STAR en STAR8 directo

<220>

40 <221> característica_misc

<222> (98)..(395)

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

45 <400> 76

```

      ggatcacctc gaagagagtc taacgtccgt aggaacgctc tcgggttcac aaggattgac      60
      cgaaccccag gatacgtcgc tctccatctg aggcttgntc caaatggccc tccactattc      120
50      caggcacgtg ggtgtctccc ctaactctcc ctgctctcct gagcccatgc tgccatcac      180
      ccatcgggtc aggtcctttc tgaanagctc ggggtgattc tctccatccc acttcctttc      240
      ccaagaaaga agccaccgtt ccaagacacc caatgggaca ttccnttcc acctccttnt      300
55      cnaaagtng cccaggtggt cntaacaggt tagggagaga ancccccagg tttnagttnc      360
      aaggcatagg acgtggctt gaacacacac acacnctc                                     398
  
```

<210> 77

60 <211> 309

<212> ADN

<213> Secuencia artificial

65 <220>

<223> elementos STAR en STAR8 inverso

ES 2 344 071 T3

<220>
 <221> característica_misc
 <222> (72)..(72)
 5 <223> "N" representa cualquier ácido nucleico

<400> 77

10	ggatcccgac tctgcaccgc aaactctacg ggcacctgca ggacggcggc ctccctgccgc	60
	ttggacgccca gncaggagct ccccggcagc agcagagcag aaagaaggat ggccccgccc	120
	cacttcgcct cccggcggtc tccctcccgc cggctcacgg acatagatgg ctgcctagct	180
15	ccggaagcct agctcttggt ccgggcattc taaggaagac acggtttttc ctcccggggc	240
	 ctcaccacat ctgggacttt gacgactcgg acctctctcc attgaatggt tgcgcgttct	300
20	ctgggaaag	309

<210> 78
 <211> 606
 25 <212> ADN
 <213> Secuencia artificial

<220>
 30 <223> elementos STAR en STAR18 directo

<220>
 <221> característica_misc
 35 <222> (524)..(524)
 <223> "N" representa cualquier ácido nucleico

<400> 78

40	tggatcctgc cgctcgcgtc ttagtgtttc tccctcaaga ctttccttct gttttgttgt	60
	cttgtgcagt attttacagc cctcttgtg tttttcttta tttctcgtac acacacgcag	120
45	ttttaagggg gatgtgtgta taattaaag gacccttggc ccatactttc ctaattcttt	180
	agggactggg attgggtttg actgaaatat gttttggtgg ggatgggacg gtggacttcc	240
	attctcccta aactggagtt ttggtcggta atcaaaacta aaagaaacct ctgggagact	300
50	ggaaacctga ttggagcact gaggaacaag ggaatgaaaa ggcagactct ctgaacgttt	360
	gatgaaatgg actcttgtga aaattaacag tgaatattca ctggtgcact gtacgaagtc	420
	tctgaaatgt aattaaagt tttattgag cccccgagct ttggcttgcg cgratttttc	480
55	cggctcgcgga catcccacg cgcagagcct cgcctccccg ctgncctcag ctccgatgac	540
	ttcccgcgcc ccgccctgct cggtgacaga cgttctactg cttccaatcg gaggcaccct	600
	tcgcgg	606

60 <210> 79
 <211> 465
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <223> elementos STAR en STAR18 inverso

ES 2 344 071 T3

<400> 79

	tggatcctgc cgctcgcgtc ttagtgtttc tccctcaaga ctttccttct gttttgttgt	60
5	cttgtgcagt attttacagc ccctcttggtg tttttcttta tttctcgta acacacgcag	120
	ttttaagggt gatgtgtgta taattaaag gacccttggc ccatacttcc ctaattcttt	180
	agggactggg attgggtttg actgaaatat gttttggtgg ggatgggacg gtggacttcc	240
10	attctcccta aactggagtt ttggtcggta atcaaaacta aaagaaacct ctgggagact	300
	ggaaacctga ttggagcact gaggaacaag ggaatgaaa ggcagactct ctgaacgttt	360
	gatgaaatgg actcttggtga aaattaacag tgaatattca ctgttgcaact gtacgaagtc	420
15	tctgaaatgt aattaaagt ttttattgag cccccgagct ttggc	465

<210> 80

<211> 48

20 <212> ADN

<213> Secuencia artificial

<220>

25 <223> oligonucleótido C65

<400> 80

	aacaagcttg atatcagatc tgctagcttg gttcgagctga tacttccc	48
--	---	----

30 <210> 81

<211> 54

35 <212> ADN

<213> Secuencia artificial

<220>

40 <223> oligonucleótido C66

<400> 81

	aaactcgagc ggcccgcaat tcgtcgactt taccactccc tatcagtgat agag	54
--	---	----

45 <210> 82

<211> 39

50 <212> ADN

<213> Secuencia artificial

<220>

55 <223> oligonucleótido C67

<400> 82

	aaaccgctgc atggaagacg ccaaaaacat aaagaaagg	39
--	--	----

60 <210> 83

<211> 32

65 <212> ADN

<213> Secuencia artificial

ES 2 344 071 T3

	<220>		
	<223> oligonucleótido C68		
5	<400> 83		
	tatggatcct agaattacac ggegatcttt cc		32
10	<210> 84		
	<211> 33		
	<212> ADN		
	<213> Secuencia artificial		
15	<220>		
	<223> oligonucleótido C81		
20	<400> 84		
	aaaccatggc cgagtacaag cccacgggtgc gcc		33
25	<210> 85		
	<211> 35		
	<212> ADN		
	<213> Secuencia Artificial		
30	<220>		
	<223> oligonucleótido C82		
35	<400> 85		
	aaatctagat caggcaccgg gcttgcgggt catgc		35
40	<210> 86		
	<211> 22		
	<212> ADN		
	<213> Secuencia artificial		
45	<220>		
	<223> oligonucleótido C85		
50	<400> 86		
	cattccccg aaaagtgcca cc		22
55	<210> 87		
	<211> 24		
	<212> ADN		
	<213> Secuencia Artificial		
60	<220>		
	<223> oligonucleótido D30		
65	<400> 87		
	tcactgctag cgagtggtaa actc		24

ES 2 344 071 T3

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

ES 2 344 071 T3

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

ES 2 344 071 T3

<210> 97
<211> 20
<212> ADN
5 <213> Secuencia Artificial

<220>
<223> oligonucleótido E12
10
<400> 97

taggggggat ccaaagtgc 20

15
<210> 98
<211> 20
<212> ADN
20 <213> Secuencia artificial

<220>
<223> oligonucleótido E13
25
<400> 98

cctaaaagaa gatcttagc 20

30
<210> 99
<211> 20
<212> ADN
35 <213> Secuencia Artificial

<220>
<223> oligonucleótido E14
40
<400> 99

aagtgtgga tccacttgg 20

45
<210> 100
<211> 20
<212> ADN
50 <213> Secuencia artificial

<220>
<223> oligonucleótido E15
55
<400> 100

tttgaagatc taccaaatgg 20

60
<210> 101
<211> 20
<212> ADN
65 <213> Secuencia Artificial

ES 2 344 071 T3

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

ES 2 344 071 T3

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

ES 2 344 071 T3

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

ES 2 344 071 T3

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

ES 2 344 071 T3

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