

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 610 655**

51 Int. Cl.:

C12N 15/85 (2006.01)

C12N 5/10 (2006.01)

C12N 15/11 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.12.2010 PCT/EP2010/070537**

87 Fecha y número de publicación internacional: **30.06.2011 WO11076870**

96 Fecha de presentación y número de la solicitud europea: **22.12.2010 E 10796393 (6)**

97 Fecha y número de publicación de la concesión europea: **19.10.2016 EP 2516653**

54 Título: **Secuencias promotoras inducibles con frío**

30 Prioridad:

22.12.2009 EP 09180432

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

28.04.2017

73 Titular/es:

**SANDOZ AG (100.0%)
Lichtstrasse 35
4056 Basel, CH**

72 Inventor/es:

**ERNST, WOLFGANG;
PONTILLER, JENS;
HESSE, FRIEDEMANN y
THAISUCHAT, HARUTHAI**

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 610 655 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 promotoras inducibles con frío

5 La presente invención se refiere a secuencias reguladoras y promotoras en ácidos nucleicos aislados y casetes de expresión.

Hoy en día, la mayor parte de los biofármacos de proteínas recombinantes son producidos en células de mamífero. El sistema de expresión de Ovario de Hámster Chino (CHO) es la plataforma de producción dominante para la fabricación de tales proteínas complejas. Las células se pueden hacer crecer a densidades celulares elevadas en cultivos en suspensión libres de suero, permitiendo un fácil aumento a escala en biorreactores, son capaces de procesar fielmente las proteínas deseadas y han obtenido la aprobación de las autoridades reguladoras puesto que tienen un largo historial de uso seguro. En la práctica a menudo resulta difícil obtener elevados rendimientos del producto recombinante. Los esfuerzos para optimizar sistemas de expresión de mamífero se han centrado principalmente en el proceso y la mejora de la línea celular, pero otra diana para la mejora del proceso aguas arriba constituye el constructo de vector genético que tiene un impacto significativo. Para la producción, los genes recombinantes se transfectan normalmente en células diana en forma de constructos de ADNc en el contexto de un casete de expresión activo en mamífero para permitir la transcripción del gen heterólogo. El constructo de ADN es reconocido por la maquinaria de transcripción celular en un proceso que implica la actividad de muchos factores de transcripción (TF) que actúan en trans en los elementos reguladores en cis, incluyendo potenciadores, silenciadores, aislantes y promotores. Los ADNc sin intrones normalmente producen niveles bajos de ARNm transcrito en el citoplasma, puesto que el empalme y el transporte al citoplasma están funcionalmente ligados. Con el fin de elevar los niveles de ARNm del gen recombinante, muchos vectores incluyen una secuencia intrónica que se localiza generalmente entre el promotor y el inserto de ADNc.

El promotor del gen es el elemento clave y está involucrado en todos estos niveles de regulación, actuando como factor determinante en la transcripción de genes mediante la integración de las influencias de la secuencia de ADN, la unión al factor de transcripción y las características epigenéticas. Se determina la fuerza y duración de la expresión transgénica que está codificada por el vector plasmídico. Los promotores más comunes utilizados para dirigir la expresión de genes heterólogos en células de mamífero son el promotor temprano inmediato principal de citomegalovirus (CMV) de ser humano y de ratón, utilizados preferentemente con sistemas de células anfitrionas tales como CHO y NS0. Confieren una expresión fuerte a corto plazo y se ha demostrado que son robustos en varios tipos de células, incluyendo células CHO y NS0. Su fuerte actividad se debe a la presencia de múltiples reiteraciones de varios sitios de unión a factores de transcripción dentro del promotor y de la región potenciadora aguas arriba. Otros promotores virales tales como el promotor temprano inmediato de SV40 y el promotor de la larga repetición terminal (LTR) del Virus del Sarcoma de Rous (RSV), también se utilizan con frecuencia en casetes de expresión. Mientras el vector de transferencia está en una fase episomal el constructo no está sujeto a los mismos mecanismos de control que un gen endógeno. Después de la integración satisfactoria en un locus funcional, estos promotores virales pueden representar una elevada expresión constitutiva de la proteína deseada. Sin embargo, esta expresión en exceso comprende una fuerte carga sobre el metabolismo celular y la maquinaria de síntesis de proteínas y por lo tanto puede inducir reacciones de estrés excesivo, tales como la respuesta de la proteína desplegada (UPR) o la respuesta al estrés ER. Estos fenómenos perjudican el correcto procesamiento de las proteínas recombinantes y pueden conducir a la inducción de apoptosis. Otro problema se refiere a la dependencia del ciclo celular de estos promotores, y la actividad transcripcional más alta se observa en la fase S, lo que conduce a una variación extrema de célula a célula en la cantidad de proteína recombinante expresada. Además, estos elementos genéticos virales no son parte de la red de regulación de la célula anfitriona y, a menudo están sujetos a silenciamiento génico. En lo que se refiere a la opinión pública y a las preocupaciones generales de bioseguridad también sería más aceptable el uso de promotores no virales en casetes de expresión en lugar de elementos procedentes de secuencias virales.

Algunas de estas reacciones negativas en principio, podría evitarse si se utilizaran promotores y elementos reguladores específicos del anfitrión. Entre los promotores específicos de la célula mejor conocidos se encuentran los de los genes constitutivos que codifican transcritos celulares transcritos abundantemente, tales como beta-actina, factor de elongación 1 alfa (EF-1 alfa), o ubiquitina. En comparación con los promotores virales, no se sabe suficiente acerca de muchos promotores de mamíferos, ya que la expresión de genes eucarióticos es más compleja y requiere una coordinación precisa de muchos factores diferentes. Aunque la actividad de estos promotores puede ser algo más baja que la del promotor de CMV, se mantiene su integridad funcional dentro de la red celular. Asimismo, confieren una expresión sostenida del transgén y los niveles de expresión a largo plazo son más estables. Un ejemplo satisfactorio representa el promotor del factor-1 de elongación derivado de CHO (CHEF-1) que se ha descrito por dirigir la expresión de alto nivel en células de mamífero (documento WO 98/49289 A1). Otro aspecto importante en relación con el uso de elementos reguladores endógenos para la expresión del transgén es la generación de ARNm estable y esa expresión puede tener lugar por consiguiente en el entorno nativo de la célula anfitriona donde se proporcionan los factores de transcripción que actúan en trans.

Dado que la expresión de genes eucarióticos es controlada por una maquinaria compleja de elementos reguladores

que actúan en cis y en trans, la mayoría de los promotores celulares padecen una falta de caracterización funcional extensiva. Un promotor eucariota normalmente se encuentra localizado inmediatamente aguas arriba de su secuencia transcrita y sirve como punto de inicio de la transcripción. El promotor central rodea inmediatamente el sitio de inicio de transcripción (TSS) que es suficiente para ser reconocido por la maquinaria de transcripción. El 'promotor proximal' comprende la región aguas arriba del promotor central y contiene el TSS y otras características de secuencia requeridas para la regulación transcripcional. Los factores de transcripción actúan de una manera específica de la secuencia mediante la unión a los motivos reguladores de la secuencia promotora y potenciadora activando de este modo la cromatina y las enzimas modificadoras de las histonas que alteran la estructura del nucleosoma y su posición, lo que finalmente permite el inicio de la transcripción. La identificación de un promotor funcional depende principalmente de la presencia de elementos potenciadores asociados aguas arriba o aguas abajo. Como promedio, la secuencia de -300 a -50 nucleótidos de TSS contribuye positivamente a la actividad promotora central y los supuestos elementos negativos a menudo se localizan de 1000 y 500 pares de bases aguas arriba del TSS. Estas características determinan en gran medida el éxito de la identificación de un nuevo promotor.

Las estrategias alternativas para alcanzar una alta expresión de la proteína recombinante se basan en un aumento del número de copias del transgén y/o en la protección del constructo recombinante del silenciamiento génico. Dependiendo de los métodos de transfección aplicados y de la presión selectiva, se generan subclones con mayores niveles de expresión. Un factor central que influye y regula la expresión génica es la estructura dinámica de la cromatina celular. La modificación de las histonas, que sirve como marcador epigenético para eventos celulares, activa las enzimas adicionales y los reguladores transcripcionales que pueden alterar la estructura de la cromatina y la expresión génica. Por ejemplo, la integración aleatoria en la heterocromatina altamente compacta y muy metilada puede conducir a una expresión escasa o nula de la proteína recombinante. Estos elementos cromosómicos que se encuentran cerca del sitio de integración del vector de transferencia pueden dar lugar al silenciamiento génico, también conocido como efecto de posición. Un efecto promotor de la transcripción está mediado por los aislantes que son elementos de barrera que pueden proteger los genes de las influencias de su entorno cromosómico y por lo tanto prevenir la inactivación del gen. Las regiones de unión a la matriz (MAR) o los elementos de apertura de la cromatina (UCOE) son otros motivos dentro de las secuencias no codificantes de los cromosomas eucarióticos que organizan estructuralmente la cromatina del núcleo. Son capaces de unirse específicamente a las proteínas nucleares específicas y juegan un papel importante en la regulación de la expresión génica. Todas las estrategias mencionadas para aumentar los niveles de transcripción de los transgenes pueden verse beneficiadas cuando se utilizan combinadas con promotores específicos de células recientemente identificados. Esto puede reducir significativamente el tiempo y los costes de selección de los clones recombinantes debido al escrutinio de un número más pequeño de transfectantes y a unos períodos de ciclo más cortos necesarios para la amplificación de genes. Estos nuevos elementos reguladores tienen el potencial de aumentar la productividad específica de la célula y mejorar aún más el rendimiento del sistema.

EBI Datenbank Núm. de Acc. AC097705 y el documento WO 99/19481 A2 proporcionan secuencias de *Rattus norvegicus* que comprenden una multitud de genes individuales.

Kumar N. et al., y col, *BMC Biotechnology* 8(42) (2008): 1-13 describen un experimento de electroforesis en gel de dos dimensiones para identificar la expresión diferencial de proteínas en un experimento de cambio de temperatura en células CHO. Se identificaron los genes GANAB, ALDH, DPYSL2, RBM3 y Vimentina como diferencialmente expresados en función de las condiciones de cultivo y la temperatura.

Al-Fageeh M. et al., *Biochem J* 397 (2006): 247-259 se refieren a las respuestas de choque de frío en las células de mamíferos y levaduras.

Kaufmann H. et al, *Biotechnology and Bioengineering* 63(5) (1999): 573-582 describen los efectos de la reducción de la temperatura sobre la fosforilación de proteínas en células CHO.

Zervos E.E. et al., *J. Surgical Res* 135 (2006): 317-322 se refieren a análisis de expresión diferencial en el cáncer.

Por lo tanto, un objetivo de la presente invención es obtener nuevas secuencias promotoras adecuadas para la expresión de genes recombinantes en células de mamífero con altos niveles de expresión y sin los inconvenientes mencionados anteriormente.

Por lo tanto, la presente invención proporciona una molécula de ácido nucleico aislada que comprende i) una secuencia de SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o ii) una secuencia que tiene una identidad de al menos 70% con dicha secuencia del SEQ ID NO: 1 o una secuencia que tiene una identidad de al menos 90% con un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, en donde las secuencias de i) y ii) tienen una actividad promotora, en donde la molécula de ácido nucleico comprende hasta 30.000 nucleótidos. Los aspectos de la invención se definen adicionalmente en las reivindicaciones. La invención ha encontrado sorprendentemente que los fragmentos de la secuencia promotora de la invención de SEQ ID NO: 1 también son activos. Por lo tanto también se proporcionan moléculas de ácido

nucleico que comprenden una secuencia promotora del nucleótido (nt) 579 al nt 800 del SEQ ID NO: 1, o que comprenden una secuencia que tiene una identidad de al menos 90% con dicha secuencia promotora. Estas secuencias de la molécula de ácido nucleico por lo general facilitan la actividad del promotor del gen (también referidas por lo tanto como "secuencias promotoras"). El término "promotor" se refiere a una secuencia de ADN que funciona para dirigir la descripción de un producto génico, p. ej., que codifica una proteína recombinante, que está conectada operablemente al mismo. El promotor puede incluir o no secuencias de control tales como secuencias reguladoras de la transcripción o de la traducción, que participan en la expresión de un producto génico dado. En general, las secuencias reguladoras de la transcripción y la traducción incluyen, pero no se limitan a la secuencia promotora, pueden incluir el elemento de respuesta de ADN para una proteína reguladora de la transcripción, los sitios de unión al ribosoma, las secuencias de inicio y parada de la transcripción y secuencias potenciadoras o activadoras. Por lo general, un promotor también comprende una o más secuencias que se unen a las proteínas reguladoras de la transcripción ("proteína de unión al ADN"). Las proteínas reguladoras de la transcripción en general se unen directamente a tales elementos de respuesta de ADN, sin embargo, en algunos casos la unión a ADN puede ser indirecta por medio de la unión a otra proteína que a su vez se une a, o está unida al elemento de respuesta de ADN.

El SEQ ID NO: 1 de la invención se ha aislado en forma de elementos de secuencia de ADN novedosos para el control de la transcripción de genes en células de ovario de hámster chino (CHO) y otras células anfitrionas de mamífero. Las secuencias descritas pueden funcionar como elementos promotores transcripcionalmente activos y permitir la expresión de los genes expresados de forma transitoria o estable.

La invención describe la identificación y la aplicación de nuevos elementos de secuencia de ADN de mamíferos que pueden reemplazar promotores virales convencionales que actualmente se utilizan en la mayoría de los vectores de expresión eucarióticos. La ventaja principal de la invención es la capacidad de permitir la expresión de genes a un nivel más alto en comparación p. ej., con el promotor de SV40 viral utilizando, elementos reguladores que actúan en cis, autólogos, específicos de células.

Una ventaja de los promotores endógenos es su inclusión y su integridad funcional dentro de la red celular del anfitrión. Por lo tanto, confieren niveles estables de expresión del transgén robustos y estables a largo plazo y muchos de los problemas descritos observados con los actuales casetes de expresión de mamíferos se pueden prevenir. La transcripción tiene lugar en el medio natural de la célula anfitriona que genera un ARNm recombinante estable y por lo tanto se proporcionan correctamente los factores de transcripción que actúan en trans.

Se encontró sorprendentemente que el SEQ ID NO: 1 de la invención y sus secuencias homólogas relacionadas, así como sus fragmentos activos, en particular, la porción de la secuencia promotora del nt 579 al nt 800 del SEQ ID NO: 1 o del nt 579 al nt 910 del SEQ ID NO: 1, tienen una actividad del promotor del gen dependiente de la temperatura. En particular, las moléculas de ácido nucleico de la invención pueden tener una expresión del gen condicionalmente inducible a una temperatura reducida que da como resultado niveles de transcripción del gen aumentados significativamente en comparación con la actividad basal a una temperatura convencional. Semejante actividad dependiente de la temperatura puede ser particularmente ventajosa para controlar la actividad de la célula en ciertas fases durante la expresión de las proteínas recombinantes. Puede ser deseable reducir la expresión en una fase inicial durante un cultivo de células, p. ej., multiplicar las células y en una fase de producción posterior aumentar la expresión del gen recombinante. Esto puede ser facilitado por un cambio de temperatura hacia temperaturas más bajas mediante el uso de los promotores de la invención. Es particularmente ventajoso lograr este aumento de la productividad en condiciones de cultivo de baja temperatura que normalmente reducen el metabolismo celular general, p. ej., de procesos competitivos pero sorprendentemente no la expresión de secuencias codificantes unidas operablemente a las secuencias promotoras de la invención.

Un "promotor" o una "secuencia que tiene actividad promotora" es una región de ADN que facilita la transcripción de un gen particular. La transcripción es el proceso de generación de un ARN del gen o la secuencia codificante del gen durante la expresión génica. La transcripción se puede estimar fácilmente, p. ej., experimentalmente en líneas celulares de expresión de ensayo tales como células CHO, mediante la determinación de la expresión de una secuencia codificante de un gen conectado operablemente al promotor. Las secuencias promotoras pueden ser sometidas a ensayo en análisis de indicadores de luciferasa. El promotor de la presente invención puede actuar modulando la expresión de cualquiera de dichas secuencias codificantes de cualquier gen. Así, la presente invención también se refiere a genes quiméricos, que comprenden cualquier secuencia promotora de la presente invención conectada operablemente a cualquier secuencia codificante que vaya a ser expresada. El promotor puede ser sensible a la temperatura, es decir, la cantidad de transcritos generados y/o de productos génicos expresados se modifica con la temperatura. Los promotores de la invención pueden haber aumentado la actividad de la expresión génica cuando la temperatura se disminuye por debajo de 37°C. En realizaciones preferidas, la expresión difiere en un factor de al menos 1,2, preferiblemente al menos 1,4, particularmente preferiblemente al menos 1,6, entre la expresión a 37°C y a 33°C.

Según se utilizan en la presente memoria, los términos "conectado operablemente" o "situado operablemente" con

respecto a un producto génico o una secuencia codificante, p. ej., en un casete de expresión, representan una secuencia que tiene una actividad promotora de la invención que está en una relación funcional con otro componente de nucleótidos de la molécula de ácido nucleico. Por ejemplo, un promotor o potenciador está conectado operablemente a una secuencia codificante si afecta a la transcripción de la secuencia codificante.

5 Semejante conexión operable puede consistir, por ejemplo, en proporcionar el promotor de la invención en la misma molécula de ADN que la secuencia codificante para un gen (que actúa en cis). De manera particularmente preferida, las secuencias promotoras se encuentran aguas arriba de las secuencias codificantes, es decir, en la posición 5' de la secuencia codificante. El promotor puede estar, p. ej., al menos 1, al menos 10, al menos 30, al menos 50, al menos 100, al menos 150, al menos 200, al menos 250, al menos 300, al menos 400, al menos 500 nucleótidos
10 aguas arriba de la secuencia codificante o directamente adyacente. En combinación con esto, o independientemente de ello, la secuencia promotora puede estar a menos de 5.000, menos de 4.500, menos de 4.000, menos de 3.500, menos de 3.000, menos de 2.500, menos de 2.000, menos de 1.500, menos de 1.000, menos de 750, menos de 500, menos de 250, menos de 100 nucleótidos aguas arriba de la secuencia codificante. Un intervalo preferido es que el fragmento desde el 579 nt al nt 800 del SEQ ID NO: 1 o su homólogo esté de 300 a 1000 nts aguas arriba de una secuencia codificante.
15

Un "fragmento" es una secuencia que comprende una parte de una secuencia completa de la que deriva, que no comprende toda la secuencia de la secuencia completa. Un fragmento de acuerdo con la presente invención puede comprender una secuencia promotora, p. ej., de al menos 100 nt (nucleótidos), al menos 120 nt, al menos 140 nt, al menos 160 nt, al menos 180 nt, al menos 200 nt, al menos 220 nt, o al menos 240 nt, al menos 260 nt, al menos 280 nt, al menos 300 nt, al menos 350 nt o al menos 400 nt, correspondiente a las secuencias de la invención descritas en la presente memoria, en particular del SEQ ID NO: 1 o sus secuencias homólogas.
20

Las posibles aplicaciones para la invención comprenden el desarrollo de líneas celulares altamente expresadas que se basan en promotores de genes no virales, en particular células CHO, para la producción de proteínas recombinantes incluyendo productos biofarmacéuticos.
25

Como se muestra en los ejemplos, las secuencias promotoras activas se han aislado de una biblioteca que codifica ADN genómico de CHO mediante hibridación de colonias utilizando sondas marcadas de genes seleccionados. La secuencia procede generalmente de regiones aguas arriba 5' del respectivo gen CHO. Las secuencias promotoras pueden ser sometidas a ensayo en análisis de indicadores de luciferasa. Tales análisis generalmente comprenden colocar el promotor en una relación funcional, p. ej., aguas arriba 5' como se mencionó anteriormente, con respecto a una secuencia codificante de luciferasa en un vector de expresión y transfectar las células anfitrionas adecuadas, preferiblemente células CHO con este vector. Un promotor activo es capaz de dirigir la expresión de la enzima luciferasa que a su vez puede ser detectada por métodos comunes. Preferiblemente, el promotor de la invención induce la expresión del producto génico al menos 3 veces por encima de un nivel de fondo de secuencias no promotoras. Los dominios funcionales pueden ser identificados y caracterizados mediante el constructo de diferentes mutantes por truncamiento y delección. Las secuencias descritas se caracterizaron sorprendentemente, además, por una capacidad para aumentar los niveles de expresión basales bajo unas condiciones ambientales definidas, tales como una temperatura más baja, como capacidad funcional preferida de las secuencias de la invención. La invención también proporciona vectores y anfitriones que utilizan tales promotores.
30
35
40

Se ha encontrado que no es necesario todo el SEQ ID NO: 1 para la actividad del promotor del gen. De hecho, un fragmento desde el nt 579 al nt 800 del SEQ ID NO: 1 era muy activo. Otros fragmentos del SEQ ID NO: 1 que componen preferiblemente la molécula de ácido nucleico de acuerdo con la presente invención son del nt 1 al nt 1505 (SEQ ID NO: 1 completo), del nt 1 al nt 1334, del nt 1 al nt 1294, del nt 1 al nt 1206, del nt 1 al nt 910, del nt 1 al nt 763, del nt 450 al nt 1167, del nt 450 al nt 1425, del nt 115 al nt 910, del nt 231 al nt 910, del nt 450 al nt 910, del nt 579 al nt 910, del nt 1 al nt 800, del nt 450 al nt 800, del nt 115 al nt 800, del nt 231 al nt 800, del nt 450 al nt 800, y por supuesto, del nt 579 al nt 800 del SEQ ID NO: 1. Los promotores particularmente preferidos de la molécula de ácido nucleico de la invención se ilustran en las Tablas 4 y 6. Cualquier fragmento de la secuencia promotora de la invención del SEQ ID NO: 1 puede comenzar en la posición 1, 115, 231, 344, 450 o la posición 579. Los fragmentos pueden ser de hasta 763, 800, 910, 1167, 1206, 1294, 1334, 1425, 1505 nucleótidos del SEQ ID NO: 1 o cualquier combinación de los mismos. Estas secuencias se denominan posteriormente secuencias promotoras. Se proporcionan otros fragmentos preferidos de acuerdo con la presente invención, con una identidad de 100% o al menos 70% como se describe adicionalmente en la presente memoria, que son del SEQ ID NO: 2 ó 3 y que corresponden a estos fragmentos enumerados anteriormente para el SEQ ID NO 1. Tales fragmentos o posiciones de fragmentos "correspondientes" en una molécula de ácido nucleico se pueden determinar por medio de programas de alineamiento de secuencias tales como Clustal W, utilizando preferiblemente los parámetros por defecto. Semejante alineamiento de secuencias se muestra en la Fig. 4B.
45
50
55
60

La presente invención se refiere además a secuencias homólogas a estas secuencias promotoras de la invención del SEQ ID NO: 1 y sus fragmentos que tienen una identidad de al menos 70% con cualquiera de estas secuencias promotoras. En realizaciones preferidas estas secuencias pueden tener una identidad de al menos 75%, una identidad de al menos 80%, una identidad de al menos 85%, más preferiblemente una identidad de al menos 90%,

una identidad de al menos 95%, una identidad de al menos 98%, una identidad de al menos 99% o una identidad de 100% con dichas secuencias promotoras. Tales secuencias homólogas son, por ejemplo el SEQ ID NO: 2 y 3 (Fig. 4B).

5 Otras secuencias homólogas son secuencias que hibridan con dichas secuencias promotoras en condiciones rigurosas (o sus cadenas complementarias para conservar la relación de la secuencia con la secuencia de los fragmentos del SEQ ID NO: 1: como tal). Las condiciones rigurosas pueden incluir el lavado a aproximadamente 65°C en un tampón que contiene aproximadamente 2 x SSC y aproximadamente SDS al 1% o condiciones equivalentes.

10 Estas secuencias homólogas, secuencias divergentes con identidad de al menos 70%, o secuencias de hibridación, puede relacionarse con cualquiera de las secuencias promotoras de la invención (fragmentos del SEQ ID NO: 1 o el SEQ ID NO: 1 completo). Preferiblemente, la identidad o la hibridación de 90% se mide en comparación con los nucleótidos 579 al nt 800 del SEQ ID NO: 1 o cualquiera de los intervalos de nucleótidos anteriormente mencionados, p. ej., del nt 1 al nt 1505, del nt 1 al nt 1334, del nt 1 al nt 1294, del nt 1 al nt 1206, del nt 1 al nt 910, del nt 1 al nt 763, del nt 450 al nt 1167, del nt 450 al nt 1425, del nt 115 al nt 910, del nt 231 al nt 910, del nt 450 al nt 910, del nt 579 al nt 910, del nt 1 al nt 800, del nt 450 al nt 800, del nt 115 al nt 800, del nt 231 al nt 800, del nt 450 al nt 800, y por supuesto, del nt 579 al nt 800 del SEQ ID NO: 1.

20 La identidad de secuencia se refiere a la identidad de secuencia de ácido nucleico entre dos o más secuencias, cuando se alinean utilizando un programa de alineamiento de secuencias. La referencia para medir la identidad de secuencia es el SEQ ID NO: 1 o sus fragmentos como se mencionó anteriormente. La secuencia puede comprender una desviación o modificación de estas secuencias promotoras de la invención, en particular, una mutación de ácido nucleico, en particular una sustitución, adición o delección de ácido nucleico. Las búsquedas de secuencias se llevan a cabo preferiblemente utilizando el programa BLASTN cuando se evalúa el % de identidad de la secuencia de ácido nucleico dada en relación con secuencias de ácido nucleico de promotores homólogos. BLASTN se puede ejecutar, p. ej., utilizando parámetros por defecto con una penalización de hueco abierto de 11,0 y una penalización por hueco extendido de 1,0 y utilizando la matriz BLOSUM-62.

30 La molécula de ácido nucleico de la presente invención puede ser ADN o ARN o una combinación de los mismos, tales como mezclas o ácidos nucleicos híbridos. En otras realizaciones preferidas, la molécula de ácido nucleico está aislada y/o purificada.

35 La molécula de ácido nucleico puede comprender además elementos reguladores adicionales, p. ej., potenciadores, supresores o aislantes conectados operablemente al promotor de la invención o a una secuencia codificante que está conectada operablemente a secuencias de la invención que tienen la actividad del promotor del gen. La expresión génica se puede modificar aún más por medio de elementos adicionales. Preferiblemente, estos elementos son inducibles, por ejemplo mediante cambios de temperatura o ciertos productos químicos para aumentar el control de la expresión génica. Los potenciadores de la transcripción a menudo se encuentran a una distancia considerable de los promotores diana pertinentes. Otros tipos de potenciadores pueden estar situados en la proximidad de un promotor dado. En realizaciones preferidas las moléculas de ácido nucleico comprenden un elemento de caja TATA, particularmente preferiblemente una porción de secuencia de la secuencia TATAA. También se prefiere una molécula de ácido nucleico de la presente invención que comprende un elemento de unión a NF-KB, particularmente preferiblemente que comprende una porción de secuencia de las secuencias TGGACAGTCC, GGGACGTCC o GGGACAGTCC. También se prefiere la presencia de al menos un elemento de unión Sp1 en la molécula de ácido nucleico de la invención, preferiblemente una porción de secuencia de las secuencias GGGCGT o GGGCGG. En realizaciones aún más preferidas la molécula de ácido nucleico comprende dos o más de tales sitios de unión Sp1. En otras realizaciones preferidas, la molécula de ácido nucleico de la invención comprende una, dos o más secuencias o fragmentos de las mismas que tienen actividad promotora. Mediante la repetición de las secuencias promotoras de la invención es posible un aumento de la actividad.

50 En realizaciones preferidas, las porciones de nucleótidos adyacentes a las secuencias de la invención son ricas en GC con un contenido de G y C de al menos 51%, al menos 52%, o al menos 53%, preferiblemente al menos 54%, preferiblemente al menos 55%, particularmente preferiblemente al menos 56%, al menos 57% o al menos 58%, sobre un área de, p. ej., 10 nucleótidos, 20 nucleótidos, 30 nucleótidos, 40 nucleótidos, 50 nucleótidos, o más. Los nucleótidos adyacentes pueden estar, p. ej., a 50, o a 100, o a 200, o a 300, o a 400 nucleótidos de las secuencias de la invención que tienen actividad promotora de genes. En los casos de las secuencias homólogas que tienen una cierta identidad de secuencia, p. ej., una identidad de al menos 70%, con las secuencias promotoras de la invención de SEQ ID NO: 1 o sus fragmentos, o las secuencias que hibridan, también pueden tener secuencias adicionales que tienen un elevado contenido en GC, que comprende preferiblemente al menos 51%, preferiblemente al menos 52%, al menos 53%, al menos 54%, al menos 55%, al menos 56%, al menos 57%, o al menos 58%, de G y C. Aunque un elevado contenido de GC puede ser importante para la actividad promotora, se prefiere que las modificaciones de las secuencias promotoras de la invención sean sometidas a ensayo con los sistemas de ensayo descritos. En particular, parece que normalmente no es necesario un elemento universalmente requerido dentro de

los promotores para la actividad promotora por lo que hay una alta posibilidad de variar las secuencias de la invención y todavía retener la actividad del promotor.

5 En un segundo y tercer aspecto de la presente invención se proporcionan secuencias promotoras basadas en los SEQ ID NO: 2 y 3. Los SEQ ID NO: 2 y 3 proporcionan secuencias homólogas a las del SEQ ID NO: 1 de la invención con elementos estructurales y/o actividades similares. La descripción anterior y las realizaciones preferidas para el SEQ ID NO: 1 por lo tanto también se aplican para el SEQ ID NO: 2 y el SEQ ID NO: 3. Por ejemplo, la descripción proporciona por lo tanto una molécula de ácido nucleico aislada que comprende una secuencia del SEQ ID NO: 2 o 3 o fragmentos de la misma, teniendo dichos fragmentos una actividad promotora o comprendiendo una
10 secuencia que tiene una identidad de al menos 70% con dicha secuencia del SEQ ID NO: 2 o 3 o dichos fragmentos de la misma, o comprendiendo una secuencia que hibrida con dicha secuencia del SEQ ID NO: 2 o 3 o dichos fragmentos de la misma en condiciones rigurosas, en donde las secuencias tienen actividad promotora.

15 Preferiblemente, el ácido nucleico de la invención se proporciona en una forma adecuada para una fácil manipulación, por ejemplo, con una longitud limitada. La molécula de ácido nucleico puede por lo tanto excluir secuencias genómicas o grandes fragmentos genómicos. En realizaciones de la invención, la molécula de ácido nucleico comprende hasta 30.000 nts (nucleótidos), o preferiblemente hasta 25.000 nts, hasta 20.000 nts, hasta 15.000 nts, hasta 12.500 nts, hasta 10.000 nts, hasta 9.000 nts, hasta 8.000 nts, hasta 7.000 nts, hasta 6.000 nts. En realizaciones preferidas, la molécula de ácido nucleico está flanqueada por sitios de restricción de endonucleasa en su extremo 5' y/o 3'.

25 En una realización adicional de la presente invención, se proporciona un casete de expresión que comprende i) una secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o ii) una secuencia que tiene una identidad de al menos 70% con dicha secuencia del SEQ ID NO: 1 o una secuencia que tiene una identidad de al menos 90% con un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o iii) una secuencia que hibrida con el complemento de dicha secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1 en condiciones rigurosas, en donde las secuencias de i), ii) y iii) tienen una actividad promotora, o una secuencia de un ácido nucleico como se define anteriormente y una secuencia de un producto génico (es decir, una secuencia codificante) en donde la secuencia
30 de la invención se coloca operablemente para la expresión de la secuencia codificante. El casete de expresión puede ser también aislado y/o purificado. Tales casetes de expresión se pueden proporcionar p. ej., en un vector adecuado para la transfección de una célula anfitriona. También, el casete de expresión se puede proporcionar en un genoma modificado de una célula anfitriona. El genoma puede ser modificado, por ejemplo, mediante técnicas recombinantes, en particular procedimientos de activación ("knock-in") o suministro de un cromosoma artificial.

35 Preferiblemente, el casete de expresión de la invención se proporciona en una forma adecuada para una fácil manipulación, p. ej., con una longitud limitada. El casete de expresión puede por lo tanto excluir secuencias genómicas o grandes fragmentos genómicos. En realizaciones preferidas, el casete de expresión comprende hasta 30.000 nts, hasta 25.000 nts, hasta 20.000 nts, hasta 15.000 nts, hasta 12.500 nts, hasta 10.000 nts, hasta 9.000 nts, hasta 8.000 nts, hasta 7.000 nts, hasta 6.000 nts. En realizaciones preferidas, el casete de expresión está flanqueado por sitios de restricción de endonucleasa en su extremo 5' y/o 3'.

45 En otras realizaciones preferidas, el casete de expresión comprende secuencias de intrones que no se traducen y se escinden entre la transcripción y la traducción. Tales secuencias de intrones pueden estar situadas, p. ej., entre una secuencia promotora y un codón de inicio de una secuencia codificante o dentro de la secuencia codificante. Se ha encontrado que tales secuencias de intrones pueden aumentar la expresión de genes debido a una relación mecánica con el procesamiento del transcrito.

50 En un aspecto adicional de la invención, se proporciona un vector que comprende: i) la secuencia del SEQ ID No: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o ii) una secuencia que tiene una identidad de al menos 70% con dicha secuencia del SEQ ID NO: 1 o una secuencia que tiene una identidad de al menos 90% con un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o iii) una secuencia que hibrida con el complemento de dicha secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1 en condiciones rigurosas, en donde las secuencias de i), ii) y iii) tienen una actividad promotora, o una secuencia de una molécula de ácido nucleico como se define más arriba. El vector también puede comprender el casete de expresión mencionado anteriormente. Este puede ser aislado y/o purificado. Preferiblemente, el vector es un vector funcional biológico tal como un vector de expresión.

60 La secuencia de la invención que tiene actividad promotora del gen se coloca preferiblemente flanqueando un sitio de endonucleasa. Esto permite una fácil clonación de las secuencias codificantes en este vector conectado operablemente a las secuencias de la invención con actividad promotora.

En un aspecto adicional de la invención las moléculas de ácido nucleico, los casetes de expresión o los vectores se pueden utilizar para la expresión de un producto génico. Un ejemplo es un método de expresión de un producto

génico, preferiblemente una proteína, que comprende la transfección de una célula aislada o una línea celular en un casete de expresión de acuerdo con la invención y la expresión del producto génico, que opcionalmente comprende además el aislamiento del producto génico expresado. En particular, la invención proporciona una utilización in vitro de una secuencia de un ácido nucleico que comprende i) una molécula del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, ii) una secuencia que tiene una identidad de al menos 70% con dicha secuencia del SEQ ID NO: 1 o una secuencia que tiene una identidad de al menos 90% con un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o iii) una secuencia que hibrida con el complemento de dicha secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1 en condiciones rigurosas, en donde las secuencias de i), ii) y iii) tienen una actividad promotora, para la expresión de un producto génico.

Una célula anfitriona adecuada para la expresión de un producto génico es, por ejemplo, una célula de mamífero, ave, planta o insecto. Preferiblemente, las células son eucarióticas. En realizaciones particularmente preferidas la célula es una célula de un roedor tal como una célula de ratón o rata y preferiblemente una célula de hámster. La célula puede ser de cualquier tejido, preferiblemente una célula de ovario, en particular una célula CHO. Para los sistemas de expresión de células individuales, preferiblemente las células son de un cultivo celular continuo.

En una realización adicional, la invención se refiere a una célula obtenible por un método de este tipo, así como una célula que comprende una molécula de ácido nucleico o un casete de expresión o un vector como se ha definido anteriormente. En realizaciones preferidas, el ácido nucleico, el casete de expresión o el vector se integran de forma estable en el genoma de dicha célula. Alternativamente, también es posible incorporar estos ácidos nucleicos, casetes de expresión o vectores de forma transitoria. Un aspecto adicional de la invención se refiere a una molécula de ácido nucleico que comprende una secuencia complementaria a una de las secuencias de la invención que tienen actividad promotora del gen, en donde la molécula de ácido nucleico comprende hasta 30.000 nucleótidos. Dicha molécula complementaria generalmente se une a la secuencia de la invención. Tales secuencias complementarias pueden ser utilizadas, por ejemplo, para unir las de manera estable a la secuencia de la invención que tiene actividad promotora del gen, p. ej., para controlar la expresión de un gen. Tras la unión estable de dicha secuencia complementaria, la expresión puede potenciarse o suprimirse, en particular suprimirse. La molécula de ácido nucleico con la secuencia complementaria puede ser de cualquier tipo de nucleótido, preferiblemente tipos de nucleótidos que se unen fuertemente a una molécula de ADN con la secuencia de la invención que tiene actividad promotora. Tales tipos de ácidos nucleicos con alta capacidad de unión son, por ejemplo, ARN, LNA (ácidos nucleicos bloqueados) o PNA (ácidos péptido nucleicos).

La presente invención se refiere además a un método in vitro de control de la expresión de un producto génico por un casete de expresión como se define anteriormente, que comprende administrar a una célula con el casete de expresión una molécula de ácido nucleico con una secuencia complementaria como se mencionó anteriormente.

La presente invención se refiere adicionalmente a un método in vitro para el control de la expresión de un producto génico por medio de un casete de expresión de acuerdo con la invención en una célula que comprende la modificación de la temperatura de dicha célula. Por lo general, las células utilizadas para expresar un gen, en particular células de mamífero, se cultivan a 37°C. Se encontró sorprendentemente que las secuencias promotoras de la invención son sensibles a la temperatura, y son adecuadas para aumentar la actividad de la expresión génica cuando la temperatura se hace disminuir por debajo de 37°C. Para modificar la temperatura, la temperatura puede ser regulada, por ejemplo, por debajo de 36°C, por debajo de 35°C, por debajo de 34°C o por debajo de 33°C. Preferiblemente, la temperatura es al menos suficiente para mantener la viabilidad de las células y el metabolismo, p. ej., al menos 10°C, al menos 12°C, al menos 15°C, al menos 17°C, al menos 20°C, al menos 22°C, al menos 24°C, al menos 26°C, al menos 28°C, al menos 30°C.

La presente invención se ilustra adicionalmente mediante las siguientes figuras y ejemplos sin estar limitada por los mismos.

Figuras:

Fig. 1: Tamaño de la UTR 3' flanqueante del gen CHO S100a6 codificado en el clon lambda entre los sitios T7 y del cebador B/D;

Fig. 2

A: Región de la secuencia no traducida 5' del gen CHO S100a6 que comprende 1505 nucleótidos aguas arriba del codón de inicio ATG;

B: Proteína A6 de unión a calcio s100 de *Mus musculus*, 1-1505 pb aguas arriba del codón de inicio ATG;

C: Proteína de unión a calcio A6 S100 de *Rattus norvegicus*, 1-1505 pb aguas arriba del codón de inicio ATG;

Fig. 3: Valores de expresión de luciferasa normalizada de todos los constructos de promotores a 37°C y 33°C

24 horas después de la transfección;

Fig. 4

A: alineamiento de secuencia no traducida 5' del gen S100a6 entre ratón, rata y de hámster chino (CHO).

Se indican NF-kB, dos sitios de unión Sp1 y una secuencia de caja TATA;

B: Alineamientos de secuencia múltiple CLUSTAL W (1.81)

Secuencia 1: CHO (cg) 1.505 pb

Secuencia 2: ratón (mm) 1.505 pb

Secuencia 3: rata (rn) 1.505 pb

Secuencias (1: 2) Alineadas. Puntuación: 76

Secuencias (2:3) Alineadas. Puntuación: 86

Secuencias (1:3) Alineadas. Puntuación: 73

Fig. 5: Actividad luciferasa de 5 constructos diferentes del promotor S100a6 en comparación con dS16, +C (pGL3-P: control positivo) y -C (pGL3-P: control negativo). La actividad Luciferasa se midió 24 horas después de la transfección de muestras a 37°C y 33°C. Todas las actividades relativas de luciferasa se normalizaron a la actividad de luciferasa de Renilla. Se muestra el promedio de dos experimentos independientes con muestras triplicadas para cada constructo;

Fig. 6: Se midió la actividad luciferasa de productos lisados de células CHO 24 horas después de la transfección. Las actividades de luciferasa se normalizaron a los valores de luciferasa de Renilla. Se muestra el promedio de cuatro experimentos independientes, cada uno con muestras por triplicado, para cada constructo.

Ejemplos:

Para la identificación de candidatos promotores se investigaron perfiles de expresión globales y haciendo hincapié en los transcritos muy abundantes que se expresan de forma constitutiva. Estos datos se obtuvieron a partir de resultados de expresión de genes en micromatrices de especies cruzadas de diversos clones CHO que fueron cultivados en diferentes condiciones. Se analizó una colección de aproximadamente 80 muestras diferentes para garantizar un número representativo de experimentos. Los niveles de abundancia de los transcritos se evaluaron de acuerdo con la intensidad media de la señal de cada gen. Después se generó una lista de prioridades de todos los transcritos candidatos utilizando un método de cálculo que permite la predicción de promotores eucarióticos a una escala genómica. La base de datos Orthologous Mammalian Gene Promoter (OMGProm) es una herramienta para la búsqueda de una colección de promotores de genes ortólogos humanos y de ratón conocidos que se obtienen de fuentes que incluyen GenBank, DBTSS, Ensembl, dbEST y RefSeq. La disponibilidad de las secuencias de genes ortólogos permite la comparación entre especies para identificar las regiones centrales que están conservadas entre las especies. Los elementos de secuencia ortólogos entre ser humano y ratón probablemente tendrán una mejor oportunidad para mostrar la misma función también en células CHO ya la regulación transcripcional es un mecanismo conservado evolutivamente. Una vez que se identifica una secuencia candidato se llevan a cabo análisis de delección identificados sobre las regiones promotoras sospechosas en los constructos plasmídicos y se transfectan a células anfitriona cultivadas in vitro. Este método experimental también fue empleado con el fin de caracterizar las secuencias aisladas. Además, la expresión diferencial de genes en células CHO se estudió a temperaturas más bajas (33°C) y se comparó con la temperatura de cultivo convencional (37°C). A partir de estos datos se seleccionaron varios genes con niveles de expresión al menos 1,5 veces mayores que el control. Se analizaron de manera preferente los genes con un nivel de expresión elevado antes de la reducción de la temperatura.

La identificación de la secuencia genómica se logró mediante el escrutinio de una biblioteca de fagos lambda (Stratagene) que codifica fragmentos de ADN genómico de CHO con un tamaño entre 9 kb y 23 kb. Las regiones de ADN que abarcaban aguas arriba y aguas abajo de la señal de inicio de la traducción (TSS) se aislaron utilizando sondas marcadas específicas de genes.

Ejemplo 1: Evaluación in silico de candidatos promotores

Se utilizó la base de datos Orthologous Mammalian Gene Promoters (OMGProm) para buscar los motivos promotores y los sitios supuestos de inicio de transcripción en una región de 2 kb aguas arriba de genes ortólogos humanos y de ratón. Esta comparación entre especies se llevó a cabo para identificar las regiones conservadas con funciones similares y homología de secuencia de ADN con la correspondiente secuencia de CHO del gen S100a6 (proteína A6 de unión a calcio S100, calciclina, Agilent ID: A_51_P281089; genbank: NM_011313; ID Gen: 20200).

Ejemplo 2: Escrutinio de una biblioteca genómica de CHO

Se escrutó una biblioteca genómica de CHO-K1 en vector Lambda FIX II (Stratagene Inc., La Jolla, Calif.) con una sonda de ADNc que codificaba el gen S100a6 de hámster chino. La secuencia de ADNc con un tamaño de 256

pares de bases se amplificó con la transcriptasa inversa Superscript® III (Invitrogen) y el uso de cebadores S100B (5'-ATGGCATGCCCCCTGGATCAG-3') y S100F (5'-CATTGTAGATCAAAGCCAAGG-3'). Se llevó a cabo la reacción de RT utilizando 250 ng de ARN total de CHO en un volumen de 20 µl. Se amplificaron adicionalmente dos microlitros del producto de reacción por medio de una ronda de PCR a 98°C durante 30s para la desnaturalización inicial, 34 ciclos de 98°C durante 10 s, 60°C durante 30 s y 72°C durante 15 s, seguido de una extensión final a 72°C durante 5 min utilizando ADN polimerasa Phusion (Finnzymes) y cebadores de S100B y S100F. El producto de la RT-PCR amplificado se purificó en gel, se secuenció y se marcaron 100 pg con DIG-dUTP en una reacción de marcaje con un volumen final de 50 µl utilizando el kit PCR DIG Probe Synthesis (Roche). Se llevaron a cabo todos los procedimientos de clonación, como se ha descrito (Sambrook y Russel 2001; páginas 1.84-1.97, 1.119-1.122, 5.4-5.17 y 8.1-8.53), las enzimas de restricción y otras enzimas modificadoras se adquirieron de New England Biolabs (NEB, USA) y se utilizaron de acuerdo con las recomendaciones del fabricante.

Para una placa (145 mm) que se utilizó para el escrutinio y la hibridación de placas, se infectaron Escherichia coli XL1-Blue MRA (P2) (600 µl de DO600 = 0,5 células diluidas con MgSO4 10 mM) con 10 µl de 50.000 unidades formadoras de placa durante 20 minutos a 37°C. A continuación, se añadieron 7 ml de agar superior NZY a ~ 50°C, se mezclaron y se extendieron sobre una placa de agar NZY precalentada (37°C). Las placas se hicieron visibles después de 8 horas de incubación a 37°C. Después de la formación de placas, las placas de cultivo se almacenaron al menos 2 horas a 4°C, se transfirieron sobre membrana de nailon (Roche), se desnaturalizaron y se neutralizaron. La etapa de preparación de los cultivos en placa y realización de las transferencias de placa se realizó de acuerdo con el manual de instrucciones para Lambda FIX II Library (Stratagene).

Antes de la hibridación, el ADN se fijó a la membrana de nailon por cocción a 80°C durante 2 horas. La prehibridación (39°C durante 1 hora) y la hibridación (39°C durante la noche) se llevaron a cabo en la misma solución, tampón convencional (5x SSC, 0,1% (p/v) N-lauroilsarcosina, SDS al 0,02% (p/v) y reactivo de bloqueo al 1%). Se utilizaron diez mililitros de tampón convencional que contenía 10 µl de sonda producto de PCR marcado con DIG para la hibridación de una membrana de nailon. Las membranas se lavaron dos veces con tampón de lavado de baja rigurosidad (2x SSC, SDS al 0,1%) a temperatura ambiente, seguido de tampón de lavado de alta rigurosidad (0,5 x SSC, SDS al 0,1%) a 65°C. Después del lavado y el bloqueo, la solución de anticuerpo (Anti-Digoxigenina-AP (Roche) diluida 1:10.000 en solución de bloqueo) se añadió a la membrana para permitir que el anticuerpo se uniera a la sonda marcada con DIG. Las membranas se lavaron, se equilibraron, se cubrieron con CDP-Star (NEB) y se expusieron a Lumi-Imager (Roche) para la detección con un análisis quimioluminiscente. Todas las etapas después de realizar las transferencias de placa se realizaron de acuerdo con el manual de aplicación de DIG (Roche).

El clon verdadero positivo se confirmó mediante una segunda y una tercera etapa de escrutinio. Un área cuadrada de agar correspondiente a la ubicación del clon positivo del primer escrutinio se recogió y se utilizó para un segundo escrutinio. A continuación, los clones positivos fueron recogidos de nuevo utilizando puntas de pipeta de 1 ml. En la última etapa de escrutinio, se utilizó también una sonda que codificaba la segunda secuencia del exón del gen S100a6. Tenía un tamaño de 135 pares de bases y se sintetizaba mediante el uso de los cebadores S2exB (5'-ATGGCATGCCCCCTGGATCAG-3') y S2exF (5'-GCCAATGGTGAGCTCCTTCTG-3'). La purificación del ADN de lambda del clon positivo se realizó con un kit Lambda (Qiagen Inc) y luego se utilizó el ADN en reacciones de PCR con T3, T7 y varios cebadores específicos de genes S100a6 (para más detalles véase la Tabla 1) para verificar la presencia y la ubicación del gen. El ADN de lambda se secuenció por medio de GATC Biotech.

Tabla 1: Secuencia de los cebadores usados para confirmar la presencia y la ubicación de S100a6 en el vector lambda

Cebador	Secuencia (5'-3')	Dirección/ubicación del cebador
A	CTCCTTTGGCTCTTCGCTGTC	efector/exón 1
B	ATGGCATGCCCCCTGGATCAG	efector/exón 2
C	CCTTCTTGTTGCCATCTTCC	efector/exón 2
D	CTGAGATTGCAAGGCTGATGG	efector/exón 3
E	GCCAATGGTGAGCTCCTTCTG	inverso/exón 2
F	CATTGTAGATCAAAGCCAAGG	inverso/exón 3

Ejemplo 3: Constructos de promotores

De acuerdo con la información de la secuencia del ADN de lambda, se amplificaron dos fragmentos solapantes de la región no codificante 5' del gen S100a6 por medio de PCR. Los fragmentos se produjeron mediante el uso de 1 de los 2 cebadores efectoras: -1425B (5'-GATGATggtaccGGAAGAAGCGAGT-TAGACAG-3') y -1167B (5'-GATGATggtaccTACGGTGTGAAGCAGCAGTG-3') junto con el cebador inverso -450F (5'-GATGATaagcttAGACC -CCAGTGTAGACCACC-3'). Las PCR se realizaron en un volumen de reacción de 50 µl que contenía 1x tampón

ES 2 610 655 T3

Phusion HF, 0,2 mM de cada dNTP, 0,5 μ M de cada cebador, 1 unidad de ADN polimerasa Phusion (Finnzymes) y 1 ng de ADN de lambda como molde. Se utilizó el termociclador T3 en las siguientes condiciones de ciclación de PCR: desnaturalización inicial a 98°C durante 2 min, 30 ciclos de 98°C durante 10 s, 60°C durante 15 s y 72°C durante 45s seguidos por una extensión final a 72°C durante 7 min. Los productos de PCR amplificados se clonaron en pGL3-basic (Promega) en los sitios KpnI-HindIII, creando pGL3-S3 y pGL3-S4. Los constructos resultantes se secuenciaron para verificar la secuencia de inserto correcta en el vector.

Con el fin de localizar el promotor central de S100a6 más precisamente, se generó un conjunto de constructos de delección 5', que abarcaba la región de ~ 1,5 kpb aguas arriba del gen S100a6, por medio de PCR con el uso del cebador inverso (-1) F junto con 1 de los 11 cebadores sentido enumerados en la Tabla 2. El número de nucleótidos en el nombre de los cebadores es relativo al codón de inicio ATG.

Tabla 2: Secuencia de los cebadores utilizados para producir los constructos de vectores indicadores con delecciones seriadas 5' de la secuencia promotora de S100a6.

Cebador	Secuencia (5'-3')
Cebador efector	
(-1505)B	GATGATggtaccGCATGCTGGCTGGGCTGGG
(-1343)B	GATGATggtaccTGAGACAGGGTTTTATATAGCC
(-1294)B	GATGATggtaccGGATCAACCTACTGAGCTATAT
(-1206)B	GATGATggtaccCACAAAGTATTTACACTGAGATTC
(-910)B	GATGATggtaccATGCCGGAGTCACGAGTCAC
(-763) B	GATGATggtaccAGAGGCGTGGAAAAGTGGAGG
(-683)B	GATGATggtaccACTCCTTGGGCGGGCCTC
(-663)B	GATGATggtaccGGATGCTAGCCGCTATAAGG
(-534)B	GATGATggtaccAGGTCCGCTCCTGGGCTGG
(-322)B	GATGATggtaccTAGGGCGGCTCCCCGAGT
(-256)B	GATGATggtaccTCGCAGTGTGTGGTCCTGTC
Cebador inverso	
(-1) F	GATGATaagcttAATTGACCACTGGGCTAGAAG

Veinticinco microlitros de mezcla de PCR contenían 1x tampón núm. 1, 0,2 mM de cada dNTP, 1 mM de MgCl₂, 0,4 μ M de cada cebador, 1 unidad de ADN polimerasa de Hifi KOD (Novagen) y 1 ng de ADN de lambda como molde. Los ciclos fueron como sigue: 95°C durante 2 min, 30 ciclos de 98°C durante 15 s, 58°C durante 5 segundos y 72°C durante 20 s, extensión final a 72°C durante 5 min. Los 11 productos de PCR se clonaron en pGL3-basic en los sitios KpnI-HindIII y se secuenciaron. Los constructos resultantes se designaron de pGL3-dSp1 a pGL3-dSp11 acuerdo con el tamaño de la delección partiendo del constructo completo derivado con el cebador (-1505) B al más pequeño obtenido con el cebador (-256) B.

Con un segundo conjunto de constructos con delección 3', se determinó la secuencia promotora mínima que se requiere absolutamente para una actividad completa. En experimentos anteriores, se demostró que región de la secuencia aguas arriba 5' más allá de la posición del nucleótido -910 con respecto al codón de inicio de la traducción ATG no es esencial para contribuir positivamente a la fuerza del promotor. Se generaron varios constructos por medio de PCR con el uso del cebador efector (-910) B junto con 1 de los 7 cebadores inversos que se enumeran en la Tabla 3.

Tabla 3: Secuencia de los cebadores utilizados para generar los constructos de vectores indicadores con delecciones seriadas 3'.

Cebador	Secuencia (5'-3')
Cebador inverso	
S-HindIII (-115) -F	GATGATaagcttTCAGGGATGTAAGAACGGAAGC
S-HindIII (-231) -F	GATGATaagcttCTTCCTGACAGGACCACACAC
S-HindIII (-344) -F	GATGATaagcttGCTTGCCTGGCACAACCAAGC
S-HindIII (-461) -F	GATGATaagcttTAGACCACCCGCGGAACCCG
S-HindIII (-579) -F	GATGATaagcttGGCAGGTAGACAGCGAAGAGC
S-HindIII (-651) -F	GATGATaagcttGCGGCTAGCATCCGGGAGG
Cebador efector	
(-910)B	GATGATggtaccATGCCGGAGTCACGAGTCAC

Ejemplo 4: Cultivo de células y análisis de transfección transitoria

5 Se cultivaron células CHO con deficiencia de dihidrofolato-reductasa (ATCC; CRL-9096) en medio de Eagle modificado por Dulbecco (DMEM) y medio F12 de Ham mezclados en una proporción de 1:1 (Invitrogen, Carlsbad, USA) con un suplemento de L-glutamina 4 mM, Pluronic al 0,1% (p/v), soja-peptona al 0,25% (p/v), 13,60 mg/l de hipoxantina y 3,88 mg/l de timidina, a 37°C en CO₂ al 7%. Las células se dividieron rutinariamente 1:6 dos veces a la semana y se utilizaron para análisis de transfección transitoria.

10 Los plásmidos construidos se transfectaron o co-transfectaron junto con el vector indicador de luciferasa de Renilla pRL-SV40 (Promega) en células CHO dhfr- mediante el uso de Nucleofector de Amaxa (Lonza). Las células se transfirieron a un tubo de 15 ml a la cantidad de 4×10^6 células por una muestra de nucleofección y se centrifugaron a 1000 rpm durante 10 min. Después de que el sobrenadante se eliminara por completo, las células se resuspendieron con la mezcla nucleofección que contenía 10 µg de ADN plasmídico, 1 µg de pRL-SV40 en 100 µl de solución de NucleofectorTM. La muestra se transfirió a una cubeta y luego se sometió a Nucleofector para la electroporación. Después de la transfección, las células se transfirieron a medio de cultivo pre-calentado en placas de 6 pocillos y se incubaron a 37°C o cuando fue necesario, las células se mezclaron bien antes de la división 1:1 en dos placas de 12 pocillos y se incubaron a 33°C y 37°C para la investigación del cambio de temperatura. Las actividades luciferasa se midieron en un luminómetro Biotek Synergy 2 (soporte lógico Gen5) utilizando el Dual-Luciferase Reporter Assay System (Progema) de acuerdo con las instrucciones del fabricante preferentemente después de 24 horas y, opcionalmente, en los puntos temporales de 48 ó 72 horas después de la transfección. Para normalizar la eficacia de transfección, se expresó la actividad del promotor de cada constructo como la razón de la actividad luciferasa de luciérnaga con respecto a la actividad luciferasa de Renilla. Se utilizaron pGL3-Basic y pGL3-Promoter como vectores positivos y negativos, respectivamente.

Ejemplo 5: Resultados

30 La secuencia activa promotora se identificó por hibridación del ADN del fago con una sonda marcada con DIG específica para la secuencia codificante del gen S100a6 de CHO. La secuencia de ADN se enriqueció y se purificó a partir de un solo clon del fago después de tres rondas sucesivas de escrutinio en placa. El tamaño del fragmento genómico insertado se determinó por medio de diferentes etapas de amplificación por PCR utilizando el cebador D específico para el exón-3 y el cebador B específico para el exón-2 (véase la Tabla 1) junto con el cebador del vector de T7 (5'-TAATACGACTCACTATAGGG-3') específico para la unión al promotor T7 dentro del sitio de clonación múltiple del ADN de lambda. Los resultados revelaron la presencia de aproximadamente 6 kb de la secuencia no traducida 3' del gen S100a6 de CHO (Fig. 1). Las dos bandas específicas están marcadas por flechas. Las reacciones de amplificación utilizando cebadores de exón antisentido combinados con el cebador de promotor de T3 no proporcionaron un producto específico, presumiblemente debido a la región flanqueante genómica más grande en este sitio.

Ejemplo 6: Cartografía y disección del promotor

45 Puesto que el promotor central y las secuencias adicionales que contribuyen positivamente a la actividad del promotor están generalmente ubicados a menos de 500 pares de bases aguas arriba de la señal de inicio de la traducción, se investigaron esta región y un fragmento sustancialmente mayor de la secuencia no traducida 5' y varios mutantes de la misma. Las secuencias completas seleccionadas (SEQ ID NO: 1-3) que son analizadas se muestran en la Fig. 2.

El gen S100a6 consiste en tres exones, de los cuales el exón-1 no está traducido. En el exón-2 el codón de inicio de la traducción ATG está situado en el nucleótido 25 y la traducción se termina en la posición 132 del exón-3 cerca del final de la secuencia del ARNm. Los exones están intercalados con dos intrones, el intrón-1 que tiene un tamaño de 550 nucleótidos y el intrón-2 que consiste en 287 nucleótidos. La organización estructural del gen de CHO es similar a la de ratón y ambos, el exón-2 y el exón-3 son idénticos en tamaño entre hámster y ratón. La homología de secuencia de la región codificante entre el ratón y CHO representa el 94%, mientras que los intrones sólo muestran una homología promedio de 77%.

Se pronostican dos sitios de inicio de la transcripción dentro de la secuencia en la Fig. 2 por medio del programa Neural Network Promoter Prediction (http://www.fruitfly.org/seq_tools/), uno de los cuales está ubicado en el exón-1 en la posición 620 aguas arriba del codón de inicio ATG y el segundo dentro del intrón-1 en la posición 420. Además, la secuencia de la Fig. 2 comprende una región que incorpora una alta frecuencia de nucleótidos CG como se identifica por medio del programa EMBOSS cpplot (<http://www.ebi.ac.uk/emboss/>) proporcionado por el European Bioinformatics Institute (EBI) parte del European Molecular Biology Laboratory (EMBL). Los parámetros que se utilizaron para la búsqueda de la isla CpG se fijaron a un contenido de GC de más de 50% y una razón de CpG observada/esperada que es mayor de 60%. Se identificó una supuesta región isla de 170 nucleótidos que alcanza desde los nucleótidos 577 hasta 747 en relación con la señal de inicio ATG que se solapa con la secuencia del exón-1. Hay tres motivos de unión al factor de transcripción SP1 pronosticados presentes en la secuencia de unión, dos dentro de la región de la isla CpG y el exón-1 en la posición 671-676 y 690-695 y un tercer sitio dentro del intrón-1 en los nucleótidos 315-320, aguas arriba de la codón de inicio ATG.

A partir de la secuencia completa, se diseñaron los mutantes por delección mediante la eliminación sucesiva de nucleótidos 5' y 3' con el fin de identificar la actividad promotora de la transcripción de las regiones flanqueantes del gen. Además, se determinaron la ubicación y el tamaño de la región promotora mínima y su fuerza de expresión en proporción a toda la secuencia. Los diferentes constructos del promotor se describen en la Tabla 4. Los caracteres 3d y 5d en el nombre de los constructos indican el sitio donde se realizó la delección. Los números detrás denotan los nucleótidos que fueron truncados de la secuencia completa. En la siguiente columna se muestran el tamaño y la posición de los nucleótidos que representan la región para cada constructo. En la columna a la derecha se representa de la longitud de cada construcción por delección en la tabla.

Tabla 4: Constructos del promotor de la región aguas arriba 5' de S100a6

Núm.	Nombre	nt	Longitud
p3	5d388.3d449	450-1167	718
p4	5d80.3d449	450-1425	976
Sps1	5d0	1-1505	1505
Sps2	5d162	1-1334	1343
Sps3	5d211	1-1294	1294
Sps4	5d299	1-1206	1206
Sps5	5d595	1-910	910
Sps6	5d742	1-763	763
Sps7	5d822	1-683	683
Sps8	5d842	1-663	663
Sps9	5d971	1-534	534
Sps10	5d1183	1-322	322
Sps11	5d1249	1-256	256
dS13	5d595.3d114	115-910	796
dS14	5d595.3d230	231-910	680
dS15	5d595.3d343	344-910	567
dS12	5d595.3d460	450-910	461
dS16	5d595.3d578	579-910	332
dS17	5d595.3d650	651-910	260

Ejemplo 7: Análisis de indicadores

Todos los diferentes constructos del promotor se insertaron en el vector básico menos el promotor pGL3-B (básico) (Promega) entre los sitios KpnI-HindIII inmediatamente aguas arriba del gen indicador de luciferasa de luciérnaga. Como vector de control se utilizó el plásmido básico pGL3-B vacío para determinar los niveles de fondo en este análisis. Como control positivo se utilizó el vector pGL3-promotor (promotor) (Promega) que contenía el promotor del Virus de Simios 40 (SV40) y la misma estructura del vector. Los análisis de indicadores cuantitativos se realizaron en experimentos de co-transfección con el fin de normalizar las variaciones experimentales como la eficacia de transfección o las diferencias en el número de células entre las muestras. Como segundo plásmido, el plásmido pRL-SV40 (Promega) que codifica el gen de luciferasa de Renilla bajo el control del promotor de SV-40 fue co-transfectado junto con todos los constructos del promotor en estudio y los plásmidos de control apropiados. Dado que el nivel de expresión del plásmido de renilla generalmente es mayor en comparación con la luciferasa de luciérnaga, sólo se usó una décima parte de la concentración del vector. Para la medición de la señal de luminiscencia de ambas proteínas indicadoras se utilizó el análisis Dual-Glo que permite unas mediciones consecutivas de ambas proteínas de la misma muestra celular.

Después de la co-transfección, se separaron las células y se transfirieron a placas de 12 pocillos para someter a ensayo la expresión del gen indicador a diferentes temperaturas. Se incubó un conjunto de placas a 37°C y el otro a 33°C. A las 24 horas de la transfección, se prepararon alícuotas de cada muestra y se analizaron cuantitativamente. El cálculo de la actividad promotora se realizó dividiendo la señal de luciferasa de luciérnaga por la de renilla y después eso los valores se normalizaron a las correspondientes muestras de control positivo de pGL3-P que se fijaron a 100%. Una visión general de la actividad de todos los constructos del promotor se muestra en la Fig. 3.

En la primera serie de constructos los mutantes por truncamiento 5' fueron diseñados para estudiar el efecto de una pérdida sucesiva de elementos reguladores esenciales como la caja TATA y los sitios de unión al factor de transcripción SP1 que están presentes en los constructos completos. La ausencia de motivos de caja TATA pronosticados no influyó significativamente en la actividad mientras que la pérdida de uno de los tres sitios de unión a SP1 redujo drásticamente la función promotora a casi los niveles de fondo. El constructo Sps6 que codificaba los nucleótidos 1 a 763 aguas arriba de la señal de ATG mostró 80-90% de fuerza de expresión del gen del vector de control. Por el contrario, el constructo Sps7 que codificaba los nucleótidos 1 a 683, pero carecía del sitio de unión a Sp1 esencial, perdió completamente la funcionalidad del promotor. Los otros mutantes por truncamiento Sps8 a Sps11 que codificaban los nucleótidos 663 a 256 del extremo 3' de la secuencia completa tampoco mostraron actividad funcional. El constructo Sps5 que incluye 910 nucleótidos del extremo 3' muestra un valor de la expresión que es considerablemente más alto que el control positivo y igual o incluso mayor que los constructos más largos que se extienden hacia el nucleótido 1505. Por lo tanto la posición 910 fue seleccionada como punto de partida para el diseño de los mutantes por delección 3' que incluye Sps12 a Sps17 con una longitud variable entre 796 y 260 nucleótidos. Todos los constructos excepto Sps17 y Sps15 mostraron una actividad promotora al menos comparable con o hasta 1,5 veces el valor del control. La caída de la actividad observada en el constructo más corto Sps17 puede ser debida a la ausencia de la caja TATA supuesta en la posición 324 u otra región crítica. El constructo dS16 que comprende 332 nucleótidos con 595 nucleótidos eliminados en el sitio 5' y 578 eliminados en el extremo 3' de Sps1 fue identificado como la región promotora mínima. La secuencia activa del promotor se encuentra localizada en el exón-1 que también contiene la supuesta isla CpG de 170 nucleótidos. Su fuerza de expresión se encuentra entre las tres primeras de todos los constructos comparados.

Los análisis de promotores en condiciones de hipotermia revelan un factor de inducción aún más alto para todas las variantes sometidas a ensayo en comparación con el control (Tabla 5). En general, la reducción de temperatura causó un aumento en la expresión de 200% como promedio. A 33°C, tres de los constructos por delección 5', Sps4, Sps5 y Sps6, muestran una mejora sustancial en la expresión del gen indicador de más de 2 veces en comparación con las muestras correspondientes del cultivo a 37°C. Esto corresponde a más de 300% de la fuerza de expresión del vector de control de SV40 no sensible a la temperatura cuando se mide en condiciones de cambio de temperatura. Asimismo, los restantes constructos por truncamiento 3', Sps1 a Sps3 que codifican todos tres sitios de unión al factor de transcripción SP1, pueden activarse constantemente >2 veces después de un cambio a 33°C. Fuera de estos seis constructos, Sps4 y Sps5 muestran la respuesta más fuerte al cambio de temperatura.

Todos los sitios SP1 se encuentran dentro de los primeros setecientos nucleótidos de la región no traducida, dos de los cuales están presentes en el exón-1 y un en el primer intrón (Tabla 5). La respuesta estimuladora de la expresión a la temperatura más baja no es tan pronunciada para los constructos Sps13, Sps14, p3 y p4. Los dos últimos constructos han perdido un sitio SP1, pero todavía muestran un factor de inducción que oscila entre 1,57 y 1,81 veces su valor a 37°C. Se observó una reducción sucesiva en la respuesta a la temperatura con la disminución de la longitud de la secuencia promotora. Los constructos Sps12, Sps16, y Sps17 oscilan entre los niveles de 1,46 a 1,74 veces sus valores de control, con la tasa más baja de expresión mejorada por la temperatura observada para el constructo Sps17 que codifica sólo 260 nucleótidos de la secuencia promotora y carece completamente del intrón-1 y un tercio del primer exón.

Tabla 5. Aumento de la expresión inducido por la temperatura

Núm.	te-1*	te-2*	MEDIA	DT	SP1**	nt
p3	1,74	1,89	1,81	0,11	3-2--	450-1167
p4	1,60	1,53	1,57	0,05	3-2--	450-1425
Sps1	2,22	2,18	2,20	0,02	3-2-1	1-1505
Sps2	2,23	2,25	2,24	0,01	3-2-1	1-1334
Sps3	2,10	2,24	2,17	0,10	3-2-1	1-1294
Sps4	2,70	2,29	2,49	0,29	3-2-1	1-1206
Sps5	2,26	3,36	2,81	0,78	3-2-1	1-910
Sps6	2,25	2,48	2,37	0,16	3-2-1	1-763
dS13	1,87	2,10	1,98	0,16	3-2-1	115-910
dS14	1,70	2,57	2,14	0,61	3-2-1	231-910
dS15	1,78	1,87	1,83	0,07	3-2--	344-910
dS12	1,57	1,48	1,52	0,07	3-2--	450-910
dS16	1,56	1,91	1,74	0,25	3-2--	579-910
dS17	1,44	1,47	1,46	0,02	3-2--	651-910

*: Los efectos de la temperatura (et) se analizaron en dos series de mediciones independientes (1 y 2). Los números muestran los valores de multiplicidad de cambio a 33°C en comparación con 37°C, con respecto a la muestra correspondiente del vector de control pGL3-P.

** : Esta columna muestra el número de sitios de unión al factor de transcripción SP1 presentes en el respectivo constructo. Los SP1 sitios se producen en las posiciones 671-676, 690-695 y 315-320. Los nucleótidos codificados por constructos individuales se muestran en la columna de la derecha.

Ejemplo 8: Secuencia promotora acortada activa y mutantes de la misma

5 En otros experimentos se generaron otros constructos diferentes (Tabla 6).

Tabla 6. Descripción de otros constructos de promotores

Núm.	Nombre	Nucleótido	Longitud
16-ΔSp1	5d595. 3d578, Δ784-847	579-658, 721-910	268
16-rev	5d595. 3d578 *	579-910	332
dS18	5D705. 3d578	579-800	222
18-rev	5D705. 3d578 *	579-800	222
18-doble	5D705. 3d588, 5d714. 3d578	589-800, 579-791	425

* Los fragmentos fueron clonados en orientación inversa en el vector indicador

10 Para generar los constructos de promotores inversos y dobles, mínimos, se utilizaron los cebadores enumerados en la Tabla 7. Se generó un constructo orientado en sentido inverso de dS16 (16-rev) por clonación del fragmento amplificado con el uso de los cebadores *HindIII* (-910) y *B KpnI* (-579) F en el vector pGL3-Básico. El constructo de promotor mínimo se produjo mediante la eliminación de 110 pb aguas arriba de la secuencia promotora de dS16, produciendo el constructo dS18. El inserto para el constructo dS18 se amplificó por medio de PCR utilizando los cebadores *KpnI* (-800) y *B (-579) M*. El constructo 18-rev se generó por medio de PCR con el uso de los cebadores *HindIII* (-800) y *B KpnI* (-579) F y se clonó en el vector en orientación inversa. Para investigar la actividad promotora de la secuencia promotora central en dos repeticiones en tándem, se generó el constructo pGL3-18doble. Inicialmente, se clonaron dos fragmentos de este promotor. La primera secuencia se amplificó con *KpnI* (-800) y *B XhoI* (-589) F y el segundo fragmento con *XhoI* (-791) y *B (-579) M*. Los productos de PCR resultantes se cortaron con las enzimas correspondientes y posteriormente se ligaron juntos en el vector pGL3-Básico. Todos los fragmentos fueron clonados en pGL3-Básico en los sitios *KpnI-HindIII* y se secuenciaron adicionalmente para verificar la autenticidad de los insertos.

15

20

Tabla 7. Secuencia de los cebadores utilizados para producir los constructos de promotores inversos y dobles, mínimos

Designación del cebador	Secuencia del cebador (5'-3')
<i>HindIII</i> (-910) B	GATGATAAGCTTATGCCGGAGTCACGAGTCAC
<i>KpnI</i> (-579) F	GATGATGGTACCGGCAGGTAGACAGCGAAGAGC
<i>KpnI</i> (-800) B	GATGATGGTACCCCTCATGCCACTCCCAATCC
<i>HindIII</i> (-800) B	GATGATAAGCTTCCTCATGCCACTCCCAATCC
<i>XhoI</i> (-791) B	GATGATCTCGAGACTCCCAATCCGGGACAGTC
<i>XhoI</i> (-589) F	GATGATCTCGAGCAGCGAAGAGCCAAAGGAGTG

5 La entidad funcional más pequeña dS16 (579 a 910) del ejemplo 6 fue acortada aún más mediante la eliminación de otras 110 pb del extremo 5', generando un constructo de promotor truncado (dS18) que contenía los nucleótidos 579 a 800 de la secuencia de Sps1 completa. De este modo se mantiene un sitio de unión al supuesto factor de transcripción NF-κB que se encuentra en la posición 780 a 771 de la secuencia promotora. El motivo de consenso NF-κB se conserva entre el gen S100a6 de hámster chino, ratón y rata (Fig. 4). El constructo dS18 se compone de 10 222 nucleótidos, incluyendo un sitio NF-κB, dos sitios de unión a Sp1 y una caja TATA. El sitio de unión a NF-κappa B en dS18 es necesario para la función completa del promotor puesto que la carencia de este sitio en Sps6 (1 a 763) da como resultado la reducción de la actividad del promotor (disminución de 65%) en comparación con Sps5 (1 a 910). La actividad del constructo dS18 es similar a la del promotor dS16 (Fig. 5).

15 Además, se generó un mutante por delección de ds16 (16-ΔSp1) para verificar la necesidad de dos sitios de unión al factor de transcripción Sp1. La falta solamente de 64 pares de bases, que incluye dos supuestos sitios de unión a Sp1 en el constructo 16-ΔSp1 (579-658.suprimido.721-910) redujo drásticamente la función del promotor a aproximadamente 20% (Fig. 5). Esta observación es relevante puesto que el sitio Sp1, un hexanucleótido de GC, es suficiente para la actividad del promotor de SV40 y con frecuencia aparece en múltiples copias. Una repetición en 20 tándem de dos cajas GC estimula la transcripción significativamente de manera más eficiente que un único elemento, y la distancia del sitio o los sitios Sp1 con respecto a la caja TATA puede ser crítica.

25 Se ha demostrado que algunos promotores de mamífero muestran una actividad de expresión cuando se utilizan en la orientación antisentido, principalmente como consecuencia de la función inversa de la caja GC. Así, la función de las dos versiones del promotor central S100a6 fue investigada mediante la orientación inversa de dS16 y dS18, dando como resultado dos constructos 16-rev y 18-rev. En los análisis de indicadores cuantitativos se detectó una actividad de aproximadamente 20% para estas dos variantes en comparación con la secuencia completa (Fig. 5).

30 Se sabe que las reiteraciones de elementos promotores centrales permiten una mayor actividad transcripcional en comparación con las versiones de copia única. Se generó el constructo 18 doble para investigar la actividad de dos secuencias promotoras centrales repetidas. Los análisis de indicadores de esta repetición en tándem del promotor dieron lugar a una mejora de la funcionalidad del promotor ya que su actividad era 1,4 veces más eficaz en comparación con el constructo de copia única dS18 (Fig. 5).

35 El aumento inducido por la temperatura en la expresión a 33°C también se observó para los dos nuevos constructos dS18 y 18-doble y se comparó con la de ds16 (Tabla 8).

Tabla 8. Aumento de la expresión inducido por la temperatura

Núm.	te-1 *	te-2 *	MEDIA	DT	nucleótido
dS16	2,50	2,19	2,35	0,15	579-910
dS18	2,43	1,95	2,18	0,24	579-800
18-doble	2,05	2,10	2,04	0,03	589-800, 579-791

* Los efectos de la temperatura (et) se analizaron en dos series de mediciones independientes (1 y 2). Los números muestran los valores de multiplicidad de cambio a 33°C en comparación con 37°C, con relación a la muestra correspondiente del vector de control pGL3-P.

40 Ejemplo 9: Inducción por disminución de la temperatura

Las variantes del promotor S100a6 Sps5 y dS16 se investigaron bajo diferentes condiciones de hipotermia a temperaturas de 33°C, 31,5°C, 30°C, 28,5°C y a la temperatura de referencia de 37°C. Los productos lisados celulares se midieron para determinar la expresión de la luciferasa 24 horas después de la transfección. El resultado

de este análisis se muestra en la Fig. 6.

El ensayo de la expresión de genes a temperaturas más bajas se basó en los descubrimientos generales de que la producción de proteínas recombinantes en células CHO se puede mejorar mediante la reducción de la temperatura.

5 Como se observa en varios experimentos de cambio de temperatura, el punto temporal en el que se mide la actividad del gen indicador es crítica, ya que esto tiene un gran impacto en la razón entre la muestra a 37°C y una muestra a una temperatura más baja. Con el tiempo, los cultivos de baja temperatura muestran un aumento continuo de expresión, presumiblemente debido a un metabolismo retardado y a una acumulación intracelular de la proteína indicadora. Por lo tanto, se calculó la razón de cada constructo a una temperatura dada para el vector de control (pGL3-P) en condiciones idénticas de acuerdo con: expresión del clon (n) a la temperatura (k) / expresión de pGL3-P a la temperatura (k). Estos valores específicos se muestran en la Tabla 9.

Tabla 9. Expresión específica de temperatura

Temperatura (°C)	Núm.	FC *	Núm.	FC *
37	Sps5	1,4	dS16	1,3
33	Sps5	3,0	dS16	2,2
31,5	Sps5	3,1	dS16	2,4
30	Sps5	4,4	dS16	3,3
28,5	Sps5	4,7	dS16	3,4

* Las veces del cambio (FC) indican la razón de cada constructo a una temperatura dada con respecto al vector de control (pGL3-P) en condiciones idénticas

15 Los resultados indican la dependencia de la temperatura de ambos constructos que aumenta cuando disminuye la temperatura (en comparación con el promotor de SV40). Este efecto de la temperatura es más pronunciado en el constructo Sps5 que comprende los nucleótidos 1-910 que en dS16 que representa una porción mucho más pequeña de sólo 332 nucleótidos.

20

LISTADO DE SECUENCIAS

	<110> Sandoz AG	
	<120> Secuencias Promotoras	
5	<130> r57905	
	<150> EP 09180432.8	
	<151> 2009-12-22	
10	<160> 40	
	<170> PatentIn Versión 3.5	
15	<210> 1	
	<211> 1505	
	<212> ADN	
	<213> Cricetulus griseus	
20	<400> 1	
	gcatgctggc tgggctggc tccattgtgt gcacattaat ttgtaagctg ctctaaagat	60
	gaacttccag gcagtgagct ggaagaagcg agttagacag aaatttattg ttggtggggg	120
	atggtgtctg aaatccttta gactgtgtcc ctcccccttt tttgagacag ggttttatat	180
	agcccagggt ggctcagaat tctgcctcgt gggatcaacc tactgagcta tatccccaaag	240
	tcttaaaacta gtgaggtcaa accaccctat cagagggggt gcctaagatc atcgaaaaac	300
	acaagtatctt aactgagat tcataaacagt agcaaaatta cgggtgtgaag cagcagtgaa	360
	aataatctta tgattggggg acaccacaac atgagaatct gtgtccaagg gtcatagaat	420
	taggaagggt gagaactatt agccaatcta gtagaccact aggggcttcc cctccttccc	480
	tggagctgac cttgccacca gagggcgaca gcatcagtga ggttcccact ccccctcaca	540
	ttgatgctga ctttagggac acattgtgct ctgtctggca gatggcccag cacacatgcc	600
	ggagtcacga gtcacgtgcc ataagggcaa actgaagtat ggaaattagg gaaaactcga	660
	tgtctctggt ttgtgctggt ctcccagacc agggtcacta ggctccctca tgccactccc	720
	aatccgggac agtcctggca gcagaggcgt ggaaaactga gggggttgtt ggggtgtgtt	780
	ttgctagcct caggcgcccg gtggggctcg gggcgggccc gcactccttg ggcgggctc	840
	ccggatgcta gccgctataa ggccagccgg actgcgacac agtccatccc ctcgaccact	900
	cctttggctc ttogctgtct acctgcctac ggtgcgggtga gctcttgctg gggcagttct	960
	aggccctct aaggctcggct cctgggctgg ggggtcaagc cacttcctgc cacatccagc	1020
	ccctacgggt tccgcggtg gtctacactg gggcttaaat ctgccgagca cggggtggtg	1080
	ggggtggggg tgggggtggg gtggggaggt aagaggggga ggtagggaga gccaaaggttc	1140
	agcttggttg tgccaggcaa gcccgaggc taaggcatcc ttataggcg gctccccgag	1200
	tctgcttttc tggggtgcag gagggctcgc cctgggtgtg tcattgtcgt cgcagtgtgt	1260
	ggtcctgtca ggaagtgcc tggagcagcc tccatctctt cctctgctca gtcatattcc	1320

ES 2 610 655 T3

ccagctctct tggaatccct ggagatcagt gttcagacac cccaaagccg cttccgttct 1380
 tacatccctg accctagttg ccctgggctg cctgcacctg tgttggctaa ggctagctgg 1440
 ttcagacagg cagcactgac tagcccctct ctgtcaaaca gcttcttcta gcccagtggt 1500
 caatt 1505

<210> 2
 <211> 1505
 <212> ADN
 <213> Mus musculus

5

<400> 2
 tcaaaaacttt tcacttgaga tgagtaactg agaatgctcg ctgggcttgg ctctactatg 60
 tgcacattca tgcataagct gctctaagga tgaaaatcca ggcagtgagt gggaggagggc 120
 gagtgatggt gctagaaatt tatggttggg gggggatgct gtttaaaatc ttttagactg 180
 tgttctttcc ccatgcccc tttgatataa ggctcaaaat tctacctttt gggatcaacg 240
 tcttttagcta tatatcaaga ttttagaccag tggttctcaa cctgtggggt gagaccctt 300
 tcacaggaat tgcctaagac catctgaaaa cacagatata cacattaaga ttcataacag 360
 tagcaaaatt atagctatga agtagaaacg agaataactt tatgattggg ggaccaccac 420
 aacttgaaga acagtattaa agggccgcag cattagtcag gttgagtgag aaccatttct 480
 tagaggatgt ggtagacaga ctgcttcccc tccctcactt ggggaccttg cactagagg 540
 gcaacagcat cagtgtggtt cccagtcccc ctcacactga tgctaacttt aaggacactg 600
 ctctcgggct ggcagaaggt tcagcacaca agccagagtt tggagtcacg tgccagaagg 660
 gaaaactaaa cacggaatta gagaaaactt gatgcctctg gcttgactg gtctcctttg 720
 ggcccgttag ggcccgctaa actccctcat tccgctccta atcctggaca gtccaggcaa 780
 caggggcgtg gaaagttgag ggggctggga tgttcgttt ccttgctca ggcgctgggt 840
 ggggtcgggg cgtgccagca ctccctgggc ggacctcacg gatgctggcc actataaggc 900
 cggccagact gcgacacatt ccatcccctc gaccactcct ttggcgcttc gctgtcgacc 960
 gtgcggtgag ctctcgctgg ggggtccctc tagggtcttt ctgctcctgg gcaagggtt 1020
 aagccacttc cttccccctg cagcctctgc aggctcagtg ggtggaatgc attgggatcc 1080
 aagttttcgg agcccaggga ggcagggaga gccatgatta ggtgggttgt atcaggcaaa 1140
 cccagaggct aatgcatccc tatggggcgg caacctgagt ctgctcttct ggggtgcagg 1200
 agggtttgcc ctgggtgtgt catcgtccca gtgtgtggct atgtcaggag gtgcccaggg 1260
 gcactctcca ttctcttctt tgctcagtca tatgccctag ttctcttga agccctggag 1320
 gacagtgtc acagatccca aagccccttc cattcttata tccctcacct aagttgcccc 1380
 ggctgccacc tgtgttggct tgagactgac tgcctcaggc aggggggtgg tgagagaact 1440
 ctctgctatc agcagcactg actagcccct ctgtcaaaca gcttcttcta gcccagtgat 1500
 cagtc 1505

10

<210> 3
 <211> 1505
 <212> ADN

15

ES 2 610 655 T3

<213> Rattus norvegicus

<400> 3

tgaaaacttt ttacttgaga tgagtaagta actaacgatg cttcctgggc ttggctccac	60
tgtgtgcaca ttaaggcata agctgctctg aggatgaaat tccaggcagt gagtgggagg	120
aggcaagtga tgttgtaga aatttatggt tggtggggga tgctgtttaa aatcttttag	180
actgtgttcc ctttctgtct cccttttgag acatgggtct tatatagggt ggctcaaaat	240
tctacctctt gggatcagcc tatctcatca agatthagcc cagtgggtct caacctgtgg	300
agacccttt cacaggaatt gcctgagacc atctgaaaac acagtattta tgtcacgatt	360
cataacagta gcaaaaatat agttatgaag cagcaacgaa aatcacttta tggttggagc	420
gtcaccacaa catgaagaat gtattaatcc gcagtattag agaggtcgag aaccactatc	480
ttagaggatg cggtagactg actgcttccc ctctcgcttg gagttgacct tgccactaga	540
gggcaacagc atcagattt ttcccagtc ccctcacact gattcgaact ttaaggacac	600
tgatctctgg ctggtagagg gttcagcaca cataccagag ttacgagtca cgtgccagaa	660
gggcaaactg aacacggaat tagagggaa ccatggtctc cggcttgcac tggctctctc	720
ttgcactaga atccttcatc ctgctcccag tccgggacgt ccaggcaaca agggcgtgga	780
aagtgagggg gctgggaggt gtgtttgctt tgctcaggc gctgggtggg gttggggcgt	840
gccagcactc cctgggcggg cctcaccgat gctggccact ataaggccag ccagactgcg	900
acacagtcca tcccctcgac cactcttttg gcgcttcatt gtcgacgtgt ggtgagctct	960
cactggggcg tccctctaag atctgtccac tcctgggtcta ggggttaagc ctttctgcc	1020
ctagtcagcc tctgcgggct ccatgggtgg aatgcactgg gatccaagtt ttcggagccc	1080
agggaaagcag ggagagccat ggtctgctgg gctgtaccag acaaaccctg aggctaaggc	1140
atccccatgg gcggcaacct gaatctgctt ttctgggggtg caggagggtt tgccctgggt	1200
gtgtcatcct cgtcccagtg tgtggccctg tcaggaggtt tccaggggca gcctccattc	1260
tcttcttgc tcagtcatat gctccagttc tcttggaaag cctggaggac agtgttcaca	1320
gaccccaaag ccccttccat tcttagactc ctcaoctcag tggccctggc tgctacctgt	1380
gttggcttga ggctagctgc ttcaggcagg tagtctcctg gctcagggga tggtagagg	1440
actctctgct accagtagca ctgaatagcc cctctgtcaa acagcttcta gccagtggt	1500
cagtc	1505

5

<210> 4

<211> 21

<212> ADN

<213> Secuencia Artificial

10

<220>

<223> cebador

15

<400> 4

atggcatgcc ccctggatca g

21

<210> 5

ES 2 610 655 T3

<211> 21
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> cebador
 <400> 5
 10 cattgtagat caagccaag g 21
 <210> 6
 <211> 21
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> cebador
 20 <400> 6
 atggcatgcc ccctggatca g 21
 <210> 7
 <211> 21
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 30 <223> cebador
 <400> 7
 35 gccaatggtg agctccttct g 21
 <210> 8
 <211> 21
 <212> ADN
 40 <213> Secuencia Artificial
 <220>
 <223> cebador
 45 <400> 8
 ctctttggc tcttcgctgt c 21
 <210> 9
 <211> 21
 <212> ADN
 50 <213> Secuencia Artificial
 <220>
 55 <223> cebador
 <400> 9
 atggcatgcc ccctggatca g 21
 60 <210> 10
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

ES 2 610 655 T3

<220>

<223> cebador

5 <400> 10
gccatcttcc ccttctgtg 20

<210> 11
<211> 21
10 <212> ADN
<213> Secuencia Artificial

<220>

15 <223> cebador

<400> 11
ctgagattgc aaggctgatg g 21

20 <210> 12
<211> 21
<212> ADN
<213> Secuencia Artificial

25 <220>

<223> cebador

30 <400> 12
gccaatggtg agctccttct g 21

<210> 13
<211> 21
35 <212> ADN
<213> Secuencia Artificial

<220>

40 <223> cebador

<400> 13
cattgtagat caagccaag g 21

45 <210> 14
<211> 32
<212> ADN
<213> Secuencia Artificial

<220>

50 <223> cebador

<400> 14
gatgatgta ccggaagaag cgagtagac ag 32

55 <210> 15
<211> 32
<212> ADN
<213> Secuencia Artificial

60 <220>

<223> cebador

ES 2 610 655 T3

<400> 15
gatgataagc ttagacccca gtgtagacca cc 32

5 <210> 16
<211> 31
<212> ADN
<213> Secuencia Artificial

10 <220>

<223> cebador

<400> 16
gatgatgga cgcgatgctg gctgggctgg g 31

15 <210> 17
<211> 34
<212> ADN
<213> Secuencia Artificial

20 <220>

<223> cebador

25 <400> 17
gatgatgga cctgagacag ggtttatat agcc 34

30 <210> 18
<211> 34
<212> ADN
<213> Secuencia Artificial

35 <220>

<223> cebador

<400> 18
gatgatgga cggatcaac ctactgagct atat 34

40 <210> 19
<211> 35
<212> ADN
<213> Secuencia Artificial

45 <220>

<223> cebador

50 <400> 19
gatgatgga cccacaagta ttactactga gattc 35

55 <210> 20
<211> 32
<212> ADN
<213> Secuencia Artificial

<220>

<223> cebador

60 <400> 20
gatgatgga ccatgccgga gtcacgagtc ac 32

<210> 21

ES 2 610 655 T3

<211> 32
 <212> ADN
 <213> Secuencia Artificial

5 <220>

 <223> cebador

 <400> 21
 10 gatgatggta ccagaggcgt ggaaaactga gg 32

 <210> 22
 <211> 30
 <212> ADN
 15 <213> Secuencia Artificial

 <220>

 <223> cebador

20

 <400> 22
 gatgatggta ccactccttg ggcgggcctc 30

 <210> 23
 <211> 32
 <212> ADN
 25 <213> Secuencia Artificial

 <220>

30

 <223> cebador

 <400> 23
 35 gatgatggta ccggatgcta gccgtataa gg 32

 <210> 24
 <211> 31
 <212> ADN
 40 <213> Secuencia Artificial

 <220>

 <223> cebador

45

 <400> 24
 gatgatggta ccaggtcggc tctgggctg g 31

 <210> 25
 <211> 30
 <212> ADN
 50 <213> Secuencia Artificial

 <220>

55

 <223> cebador

 <400> 25
 gatgatggta cctagggcgg ctccccgagt 30

60

 <210> 26
 <211> 32
 <212> ADN
 <213> Secuencia Artificial

ES 2 610 655 T3

<220>
<223> cebador
5 <400> 26
gatgatggta cctcgcagtg tgtggctctg tc 32
<210> 27
<211> 33
10 <212> ADN
<213> Secuencia Artificial
<220>
15 <223> cebador
<400> 27
gatgataagc ttaattgacc actgggctag aag 33
20 <210> 28
<211> 34
<212> ADN
<213> Secuencia Artificial
25 <220>
<223> cebador
30 <400> 28
gatgataagc tttcaggat gtaagaacgg aagc 34
<210> 29
<211> 33
35 <212> ADN
<213> Secuencia Artificial
<220>
40 <223> cebador
<400> 29
gatgataagc ttctcctga caggaccaca cac 33
<210> 30
<211> 33
45 <212> ADN
<213> Secuencia Artificial
<220>
50 <223> cebador
<400> 30
55 gatgataagc ttgcttgct gcacaacca agc 33
<210> 31
<211> 32
<212> ADN
60 <213> Secuencia Artificial
<220>
<223> cebador

ES 2 610 655 T3

<400> 31
gatgataagc ttagaccac ccgcggaacc cg 32

5 <210> 32
<211> 33
<212> ADN
<213> Secuencia Artificial

10 <220>

<223> cebador

<400> 32
gatgataagc ttggcaggta gacagcgaag agc 33

15 <210> 33
<211> 31
<212> ADN
<213> Secuencia Artificial

20 <220>

<223> cebador

25 <400> 33
gatgataagc ttgcgctag catccgggag g 31

30 <210> 34
<211> 32
<212> ADN
<213> Secuencia Artificial

35 <220>

<223> cebador

<400> 34
gatgatgta ccatgccgga gtcacgagtc ac 32

40 <210> 35
<211> 32
<212> ADN
<213> Secuencia Artificial

45 <220>

<223> cebador

50 <400> 35
gatgataagc ttatgccgga gtcacgagtc ac 32

55 <210> 36
<211> 33
<212> ADN
<213> Secuencia Artificial

<220>

<223> cebador

60 <400> 36
gatgatgta ccggcaggta gacagcgaag agc 33

<210> 37

ES 2 610 655 T3

<211> 32
<212> ADN
<213> Secuencia Artificial

5 <220>

<223> cebador

<400> 37
10 gatgatggta cccctcatgc cactccaat cc 32

<210> 38
<211> 32
<212> ADN
15 <213> Secuencia Artificial

<220>

<223> cebador

20 <400> 38
gatgataagc ttccctcatgc cactccaat cc 32

<210> 39
<211> 32
<212> ADN
25 <213> Secuencia Artificial

<220>

30 <223> cebador

<400> 39
35 gatgatctcg agactcccaa tccgggacag tc 32

<210> 40
<211> 33
<212> ADN
40 <213> Secuencia Artificial

<220>

<223> cebador

45 <400> 40
gatgatctcg agcagcgaag agccaaagga gtg 33

50

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada que comprende
- 5 i) una secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o
- ii) una secuencia que tiene una identidad de al menos 70% con dicha secuencia del SEQ ID NO: 1 o una secuencia que tiene una identidad de al menos 90% con un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, en donde las secuencias de i) y ii) tienen una actividad promotora, en donde la molécula de ácido nucleico comprende hasta 30.000 nucleótidos.
- 10 2. La molécula de ácido nucleico de acuerdo con la reivindicación 1, **caracterizada porque** el fragmento comprende del nt 1 al nt 1334, del nt 1 al nt 1294, del nt 1 al nt 1206, del nt 1 al nt 910, del nt 1 al nt 763, del nt 450 al nt 1167, del nt 450 al nt 1425, del nt 115 al nt 910, del nt 231 al nt 910, del nt 450 al nt 910, del nt 579 al nt 910, del nt 1 al nt 800, del nt 450 al 800, del nt 115 al nt 800, del nt 231 al nt 800, del nt 450 al nt 800 o del nt 579 al nt 800 del SEQ ID NO: 1, preferiblemente en donde el fragmento comprende del nt 579 al nt 800 del SEQ ID NO: 1.
- 15 3. La molécula de ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 1 ó 2, **caracterizada porque** la molécula de ácido nucleico comprende hasta 25.000 nucleótidos.
- 20 4. La molécula de ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en donde la actividad del promotor es sensible a la temperatura.
5. La molécula de ácido nucleico de la reivindicación 4, **caracterizada porque** la actividad del promotor se determina en un análisis de indicador de luciferasa.
- 25 6. La molécula de ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 1 a 5, **caracterizada porque** la molécula de ácido nucleico ii) comprende una secuencia que tiene una identidad de secuencia de al menos 75%, preferiblemente al menos 80%, más preferiblemente al menos 85%, incluso más preferiblemente al menos 90%, particularmente preferiblemente al menos 95%, con el SEQ ID NO: 1.
- 30 7. Un casete de expresión que comprende
- i) una secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o
- 35 ii) una secuencia que tiene una identidad de al menos 70% con dicha secuencia del SEQ ID NO: 1 o una secuencia que tiene una identidad de al menos 90% con un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o
- iii) una secuencia que hibrida con el complemento de dicha secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1 en condiciones rigurosas,
- 40 en donde las secuencias de i), ii) y iii) tienen una actividad promotora, o una secuencia de un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 1 a 6 y una secuencia codificante, en donde la secuencia está situada operativamente para la expresión de la secuencia codificante.
- 45 8. Un vector que comprende una secuencia promotora que comprende
- i) la secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o
- ii) una secuencia que tiene una identidad de al menos 70% con dicha secuencia del SEQ ID NO: 1 o una secuencia que tiene una identidad de al menos 90% con un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1, o
- 50 iii) una secuencia que hibrida con el complemento de dicha secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1 en condiciones rigurosas, en donde las secuencias de i), ii) y iii) tienen una actividad promotora,
- o una secuencia de un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 1 a 6 o el casete de expresión de la reivindicación 7.
- 55 9. El vector de la reivindicación 8, **caracterizado porque** dicha secuencia está situada flanqueando un sitio de endonucleasa.
- 60 10. El uso in vitro de una secuencia de una molécula de ácido nucleico que comprende
- i) una secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1,
- ii) una secuencia que tiene una identidad de al menos 70% con dicha secuencia del SEQ ID NO: 1 o una secuencia que tiene una identidad de al menos 90% con un fragmento de la misma que comprende del nt 579

al nt 800 del SEQ ID NO: 1, o

iii) una secuencia que hibrida con el complemento de dicha secuencia del SEQ ID NO: 1 o un fragmento de la misma que comprende del nt 579 al nt 800 del SEQ ID NO: 1 en condiciones rigurosas, en donde las secuencias de i), ii) y iii) tienen una actividad promotora,

5 o una molécula de ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 1 a 6 para la expresión de un producto génico.

10 11. Un método para expresar un producto génico, preferiblemente una proteína, que comprende la transfección de una célula aislada o una línea celular con un casete de expresión de acuerdo con la reivindicación 7 y la expresión del producto génico, que comprende opcionalmente además la etapa de aislamiento del producto génico expresado.

15 12. El método de la reivindicación 11, **caracterizado porque** la célula es una célula de mamífero o aviar, preferiblemente una célula de hámster, y/o una célula de ovario, preferiblemente una célula CHO.

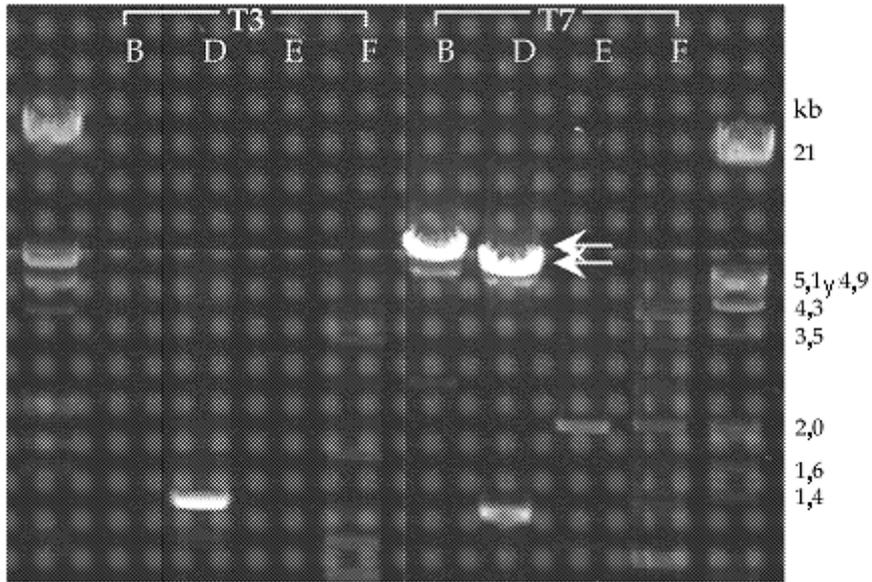
13. Una célula que comprende un casete de expresión de acuerdo con la reivindicación 7 o un vector de acuerdo con la reivindicación 8 ó 9, preferiblemente en donde el casete de expresión o vector se integra de forma estable en el genoma de dicha célula.

20 14. Una molécula de ácido nucleico que comprende una secuencia complementaria a una secuencia de una molécula de ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en donde la molécula de ácido nucleico comprende hasta 30.000 nucleótidos.

25 15. Un método in vitro de control de la expresión de un producto génico por un casete de expresión de acuerdo con la reivindicación 7 en una célula, que comprende la administración a la célula de una molécula de ácido nucleico con una secuencia complementaria de acuerdo con la reivindicación 14 y/o que comprende la modificación de la temperatura de dicha célula.

30

Fig. 1:



ES 2 610 655 T3

Fig. 2 A:

SEQ ID NO 1

GCATGCTGGCTGGGCTGGGCTCCATTGTGTGCACATTAATTTGTAAGCTGCTCTAAAGATGA
ACTTCCAGGCAGTGAGCTGGAAGAAGCGAGTTAGACAGAAATTTATTGTTGGTGGGGGATGG
TGTCTGAAATCCTTTAGACTGTGTCCCTCCCCCTTTTTTGAGACAGGGTTTTATATAGCCCA
GGTTGGCTCAGAATTCTGCCTCGTGGGATCAACCTACTGAGCTATATCCCCAAGTCTTAAAC
TAGTGAGGTCAAACCACCCTATCAGAGGGGTTGCCAAGATCATCGGAAAACACAAGTATTT
ACACTGAGATTCATAACAGTAGCAAAATTACGGTGTGAAGCAGCAGTGAAAATAATTTTATG
ATTGGGGGACACCACAACATGAGAATCTGTGTCCAAGGGTCATAGAATTAGGAAGGTTGAGA
ACTATTAGCCAATCTAGTAGACCACTAGGGGCTTCCCCCTCCTTCCCTGGAGCTGACCTTGCC
ACCAGAGGGGCGACAGCATCAGTGAGGTTCCCCTCCCCCTCACATTGATGCTGACTTTAGGG
ACACATTGTGCTCTGTCTGGCAGATGGCCCAGCACACATGCCGGAGTCACGAGTCACGTGCC
ATAAGGGCAAACCTGAAGTATGGAAATTAGGGAAAACCTCGATGTCTCTGGTTTTGTGCTGGTCT
CCCAGACCAGGGTCACTAGGCTCCCTCATGCCACTCCCAATCCGGGACAGTCCTGGCAGCAG
AGGCGTGGAAAACCTGAGGGGGTTGTTGGGGTGTGTTTTGCTAGCCTCAGGCGCCGGGTGGGG
CTCGGGGCGGGCCGGCACTCCTTGGGCGGGCCTCCCGGATGCTAGCCGCTATAAGGCCAGCC
GGACTGCGACACAGTCCATCCCCCTCGACCACTCCTTTGGCTCTTCGCTGTCTACCTGCCTAC
GGTGCGGTGAGCTCTTGCTGGGGCAGTTCCTAGGTCCCTCTAAGGTCGGCTCCTGGGCTGGGG
GGTCAAGCCACTTCCTGCCACATCCAGCCCCTACGGGTTCCGCGGGTGGTCTACACTGGGGT
CTAAATCTGCCGAGCACGGGGTGGTGGGGTGGGGTGGGGTGGGGTGGGGAGGTAAGAGG
GGGAGGTAGGGAGAGCCAAGGTTAGCTTGGTTGTGCCAGGCAAGCCCGGAGGCTAAGGCAT
CCTTATAGGGCGGCTCCCCGAGTCTGCTTTTCTGGGGTGCAGGAGGGTTCGCCCTGGGTGTG
TCATTGTCGTGCGAGTGTGTGGTCTGTGAGGAAGTGCCCTGGAGCAGCCTCCATCTCTTCC
TCTGCTCAGTCATATTCCCCAGCTCTCTTGAATCCCTGGAGATCAGTGTTTACAGACACCCCA
AAGCCGCTTCCGTTCTTACATCCCTGACCCTAGTTGCCCTGGGCTGCCTGCACCTGTGTTGG
CTAAGGCTAGCTGGTTCAGACAGGCAGCACTGACTAGCCCCTCTCTGTCAAACAGCTTCTTC
TAGCCCAGTGGTCAATT

Fig. 2 B:

SEQ ID NO 2

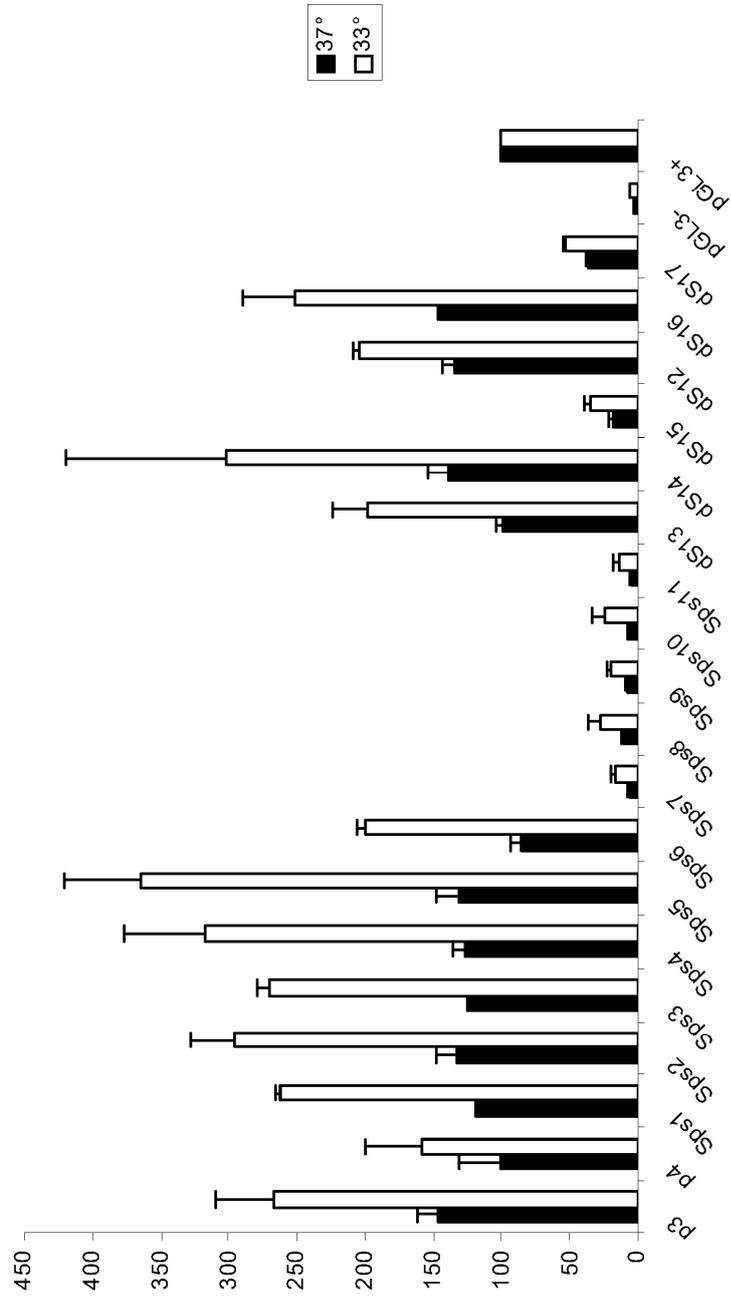
TCAAAACTTTTCACTTGAGATGAGTAACGAGAATGCTCGCTGGGCTTGGCTCTACTATGTGCACATTCATGCAT
AAGCTGCTCTAAGGATGAAAATCCAGGCAGTGAGTGGGAGGAGGCGAGTGATGTTGCTAGAAAATTTATGGTTGGT
GGGGGATGCTGTTAAAATCTTTTAGACTGTGTCTTTCCCCATGCCCCCTTTGATACAAGGCTCAAAAATCTAC
CTTTTGGGATCAACGTCTTTAGCTATATATCAAGATTTAGACCAGTGGTTCTCAACCTGTGGGTTGAGACCCCTT
TCACAGGAATTGCCTAAGACCATCTGAAAACACAGATATACACATTAAGATTCATAACAGTAGCAAAAATTTATAGC
TATGAAGTAGAAACGAGAATAAAGTTATGATTTGGGGACCACCACAACCTTGAAGAACAGTATTAAAGGGCCGAG
CATTAGTCAGGTTGAGTGAGAACCATTTCTTAGAGGATGTGGTAGACAGACTGCTTCCCTCCCTCACTTGGGGA
CCTTGGCCACTAGAGGCCAACAGCATCAGCTGTGGTTCCAGTCCCCCTCACACTCATGCTAAGCTTAAAGGACACTG
CTCTCGGGCTGGCAGAAGGTTCCAGCACACAAGCCAGAGTTTCGAGTCACGTGCCAGAAAGGAAAACTAAACACGG
AATTAGAGAAAACCTTGATGCCTCTGGCTTGCCTGGTCTCCTTTGGGCCCGTTAGGGCCCGCTAAACTCCCTCAT
TCCGCTCCTAATCCTGGACAGTCCAGGCAACAGGGCGTGAAAGTTGAGGGGGCTGGGATGTTGTTTTGCCTTG
CCTCAGGCGCTGGTGGGGTGGGGCGTCCAGCACTCCCTGGGCGGACCTCACGGATGCTGGCCACTATAAGGC
CGGCCAGACTGGACACATTCATCCCTCGACCCTCCTTTGGCGCTTCCGCTGTGACCGTGCCTGGTGGCTCTC
GCTGGGGGCTCCCTTAGGGTCTTTCTGCTCCTGGGCAAGGGGTTAAGCCACTTCCCTCCCCCGTACGCTCTGC
AGGCTCAGTGGGTGGAATGCATTGGGATCCAAGTTTTCCGAGCCAGGGAGGCAGGGAGAGCCATGATTAGGTGG
GTTGTATCAGGCAAAACCAGAGGTAATGCATCCCTATGGGGCGCAACCTGAGTCTGCTCTCTGGGGTGCAGG
AGGGTTTGCCTGGGTGTGTGCATCGTCCAGTGTGTGGCTATGTCAGGAGGTGCCAGGGGCACTCTCCATTTCT
TTCTTGCTCAGTCATATGCCCTAGTTCTCTTGAAGCCCTGGAGGACAGTGTCTCACAGATCCCAAAGCCCTTC
CATCTTATATCCCTCACCTAAGTTGCCCGGCTGCCACCTGTGTGGCTTGAGACTGACTGCCTCAGGCAGGGG
GGTGGTGAGAGAAGCTCTCTGCTATCAGCAGCACTGACTAGCCCTCTGTCAAACAGCTTCTTCTAGCCCAAGT
CAGTC

Fig. 2 C:

SEQ ID NO 3

TGAAAACCTTTTACTTGAGATGAGTAAGTAACTAACGATGCTTCTGGGCTTGGCTCCACTGTGTGCACATTAAG
GCATAAGCTGCTCTGAGGATGAAATTCAGGCAGTGAGTGGGAGGAGGCAAGTGATGTTGTTAGAAAATTTATGGT
TGGTGGGGGATGCTGTTAAAATCTTTTAGACTGTGTTCCTTCTGTCTCCCTTTTGGAGACATGGGTCTTATAT
AGGTTGGCTCAAAAATCTACCTCTTGGGATCAGCCTATCTCATCAAGATTTAGCCCAGTGGTGTCAACCTGTGG
AGACCCCTTTCACAGGAATTGCCTGAGACCATCTGAAAACACAGTATTTATGTCACGATTCATAACAGTAGCAAA
AATATAGTTATGAAGCAGCAACGAAAATCACTTTATGGTTGGAGCGTCACCACAACATGAAGAATGTATTAATCC
GCAGTATTAGAGAGGTCGAGAACCATACTTAGAGGATGCCGTAGACTGACTGCTTCCCTCTCGCTTGGAGTT
GACCTTGGCCACTAGAGGGCAACAGCATCAGTATTGTTCCAGTCCCTCACACTGATTCGAACTTTAAGGACAC
TGATCTCTGGCTGGTAGAGGGTTCAGCACACATACCAGAGTTACGAGTCACGTGCCAGAAAGGGCAAACTGAACAC
GGAATTAGAGGGAACCTCGATGTCTCCGGCTTGCCTGGTCTTCTTGCCTAGAAATCCTTCATCTCTGCTCCAG
TCCGGGACGTCAGGCAACAAGGGCGTGAAAGTGAGGGGGCTGGGAGGTGTGTTTGCCTTGCCTCAGGCGCTGG
GTGGGGTGGGGCGTCCAGCACTCCCTGGGCGGGCTCACCGATGCTGGCCACTATAAGGCCAGCCAGACTGCG
ACACAGTCCATCCCTCGACCCTCTTTGGCGCTTCACTTGTGACGTGGTGGTGGTCTCACTGGGGGCTCCCT
CTAAGATCTGTCCACTCCTGGTCTAGGGTTAAGCCTTTCTGCTTAGTCAAGCTTCTGCGGGCTCCATGGGTGG
AATGCACTGGGATCCAAGTTTTTCGGAGCCCAGGGAAGCAGGGAGGCCATGGTCTGTGGGCTGTACCAGACAAA
CCCCGAGGCTAAGGCATCCCCATGGGCGGCAACCTGAATCTGCTTTTCTGGGGTGCAGGAGGGTTTGCCTGGGT
GTGTCATCTCTGCTCCAGTGTGTGGCCCTGTGAGGAGTTCCAGGGGCAAGCCTCCATTTCTTCTCTGCTCAGT
CATATGCTCCAGTTCTCTTGGAAAGCCCTGGAGGACAGTGTTCACAGACCCCAAAGCCCTTCCATTTCTTAGACTC
CTCACCTCAGTGGCCCTGGCTGCTACCTGTGTGGCTTGGGCTAGCTGCTTCAAGCAGGTAGTCTCTGGCTCA
GGGATGGTGGAGGACTCTCTGTACCAGTAGCACTGAATAGCCCTCTGTCAAACAGCTTCTTCTAGCCCAAGT
CAGTC

Fig. 3:



ES 2 610 655 T3

```

mm      GTCAGGTTGAGTGAGAACCATT-TCTTAGAGGAIGTGGTAGAC---AGACTGCTTCCCCT
rn      GAGAGGTCGAG----AACCACTATCTTAGAGGATGCGGTAGAC---TGACTGCTTCCCCT
cg      GGAAGGTTGAG----AACTATT-----AGCCAATCTAGTAGACCACTAGGGGCTTCCCCT
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      CCCTCACTTGGG---GACCTTGCCACTAGAGGGCAACAGCATCAGTGTGGTTCCCAGTCC
rn      C--TCGCTTGGAGTTGACCTTGCCACTAGAGGGCAACAGCATCAGTATTGTTCCCAGTCC
cg      CCTTC-CCTGGAGCTGACCTTGCCACCAGAGGGCGACAGCATCAGTGTGGTTCCCAGTCC
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      CCCTCACACTGATGCTAACTTTAAGGACAC--TGCTCTCGGGCTGGCAGAAGGTTTCAGCA
rn      CCCTCACACTGATTCGAACTTTAAGGACAC--TGATCTCTGGCTGGTAGAGGTTTCAGCA
cg      CCCTCACATTGATGCTGACTTTTAGGGACACATTGTGCTCTGTCTGGCAGATGGCCACGCA
***** *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      CACAAGCCAGAGTTTCGAGTCACGTGCCAGAAAGGAAAACATAAA-CACGGAA-TTAGAGA
rn      CACATACCAGAGTTACGAGTCACGTGCCAGAAAGGCAAACATAAA-CACGGAA-TTAGAGG
cg      CACATGCCGAGTTCACGAGTCACGTGCCATAAGGGCAAACATAAGTATGGAAATTAGGGA
***** *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      AAACFTGATGCTCTGGCTTGCACCTGGTCTCCTTTGGGCCCGTTAGGGCCCGCTAAACTC
rn      GAACFCGATGCTCCTGGCTTGCACCTGGTCTCCTTTG-----CACTAGAATC
cg      AAACFCGATGCTCCTGGCTTGCACCTGGTCTCCTAGACC-----AGGCTACAGGCTC
***** *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      CCTCATTCGGCTCCTAATCCGGGACAGTCCAGGCAACAGGGGCGTGAAAGTTGAGGGGG
rn      CTTTCATCTGCTCCAGTCCGGGAC-GTCCAGGCAACAAGGGCGTGAAAGT-GAGGGGG
cg      CCTCATGCCACTCCCAATCCGGGACAGTCCAGGCAACAGGGGCGTGAAACTGAGGGGG
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      CTGG-GATGTTTCGTTTGCCCTTCGCTCAGGCGCTGGGTGGGG-TCGGGGCTGCCAGCACT
rn      CTGG-GAGGTGTGTTTGCCCTTCGCTCAGGCGCTGGGTGGGG-TTCGGGGCTGCCAGCACT
cg      TTGTTGGGGTGTGTTTTCGCTCAGGCGCTGGGTGGGG-TCGGGGCTGCCAGCACT
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      CCTTGGGCGGACCTCACGGATGCTGGCCAGTATAAGGCCGGCCAGACTGCGACACATCC
rn      CCTTGGGCGGSCCTCACCGATGCTGGCCAGTATAAGGCCAGCCAGACTGCGACACAGTCC
cg      CCTTGGGCGGSCCTCCCGGATGCTAGCCGATAAGGCCAGCCAGACTGCGACACAGTCC
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      ATCCCTCGACCACTCCTTTGGCGCTTCGCTGTCGACC-----GTGCGGTGAGCTCT
rn      ATCCCTCGACCACTCCTTTGGCGCTTCATGTCGAC-----GTGTGGTGAGCTCT
cg      ATCCCTCGACCACTCCTTTGGCTCTTCGCTGTCCTACCTGCCTACGGTGCCTGAGCTCT
***** *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      CGTGGGGG-----GTCCCTCTAGGGTCTTTCGCTCCTGGGCAAGGGG-TTAAGCC
rn      CACTGGGGC-----GTCCCTCTAAGATCTGTCACCTCCTGGTCTAGGGG-TTAAGCC
cg      TGCTGGGGCAGTTCTAGGTCCTCTAAGGTCG---GCTCCTGGGCTGGGGGTTCAAGCC
***** *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      ACTTCCTTCCCCCGTACGCTCTGCAGGCTCAGTGGGTGGAATGCATTGGGATCCAAGTT
rn      -TTTCCTGCCTAGTACGCTCTGCAGGCTCCTAGGGTGGAAATGCATTGGGATCCAAGTT
cg      ACTTCCTGCACATCCAGCCCTACGGGTTCGCGGGTGTCTACACTGGGGTCTAAATC
***** *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      TTCGGAGCCAGGG-----AG
rn      TTCGGAGCCAGGG-----AA
cg      TCGGAGCAGCGGGTGGTGGGGTGGGGTGGGGTGGGGTGGGGAGGTAAAGAGGGGAG
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

mm      GCAGGGAGAGCCATGAT-TAGGTGGGTGTATCAGGCAAAACCCAGAGGCTAATGCATCCC
rn      GCAGGGAGAGCCATGGT-CTGCTGGGCTGTACCAGACAAACCCGAGGCTAAGGCATCCC
cg      GTAGGGAGAGCCAGGTTACGTTGGTTGTGCCAGGCAAGCCGGAGGCTAAGGCATCCT
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

```

ES 2 610 655 T3

```

mm      TATGGGGCGGCAACCTGAGTCTGCTCTTCTGGGGTGCAGGAGGGTTTGCCCTGGGTGTGT
rn      CATGGG-CGGCAACCTGAATCTGCTTTTCTGGGGTGCAGGAGGGTTTGCCCTGGGTGTGT
cg      TATAGGGCGGCTCCCCGAGTCTGCTTTTCTGGGGTGCAGGAGGGTTTGCCCTGGGTGTGT
          ** ** ***** ** ** ***** ***** ***** ***** *****
mm      CAT---CGTCCAGTGTGTGGCTATGTCAGGAGGTGCCAGGGGCACTCTCCATTCTCTT
rn      CATCCTCGTCCCAGTGTGTGGCCCTGTCAGGAGGTTTCCAGGGGCAGCCTCCATTCTCTT
cg      CATTGTCGTCGAGTGTGTGGTCCCTGTCAGGAAGTGCCCTGGAGCAGCCTCCATCTCTT
          ***  **** ***** ***** ***** ** ** ** * ***** *

mm      CCTTGCTCAGTCATATGCCCTAGTTCTCTGGAAGCCCTGGAGGACAGTGCACAGATC
rn      CCTTGCTCAGTCATATGCTCCAGTTCTCTGGAAGCCCTGGAGGACAGTGTTCACAGACC
cg      CTCTGCTCAGTCATATTCCCAGCTCTCTGGAATCCCTGGAGATCAGTGTTCAGACACC
          *  ***** * * ** ***** ***** ***** * * *

mm      CCAAAGCCCCTTCCATTCTTATATCCCTCACCTAAGTTGCCCCGG-CTGCC---ACCTGT
rn      CCAAAGCCCCTTCCATTCTTAGACTCCTCACCTCAGTGGCCCTGG-CTGCT---ACCTGT
cg      CCAAAGCCGCTTCCGTTCTTACATCCCTGACCCTAGTTGCCCTGGGCTGCCCTGCACCTGT
          ***** ***** ***** * *** ** * * * * * * * * * *

mm      GTTGGCTTGAGACTGACTGCCTCAGGCAGG-----GGGGTGGTGAGAGA
rn      GTTGGCTTGAGGCTAGCTGCTTCAGGCAGGTAGTCTCCTGGCTCAGGGGATGGTGAGAGG
cg      GTTGGCT-AAGGCTAGCTGGTTCAGACAGG-----
          ***** ** ** * * * * * *

mm      ACTCTCTGCTATCAGCAGCACTGACTAGCCCCTCT--GTCAAACAGCTTCTTCTAGCCCA
rn      ACTCTCTGCTACCAGTAGCACTGAATAGCCCCTCT--GTCAAACAGCTTCT---AGCCCA
cg      -----CAGCACTGACTAGCCCCTCTCTGTCAAACAGCTTCTTCTAGCCCA
          ***** ***** ***** ***** *****

mm      GTGATCAGTC
rn      GTGGTCAGTC
cg      GTGGTCAATT
          *** ** *

```

Fig. 5:

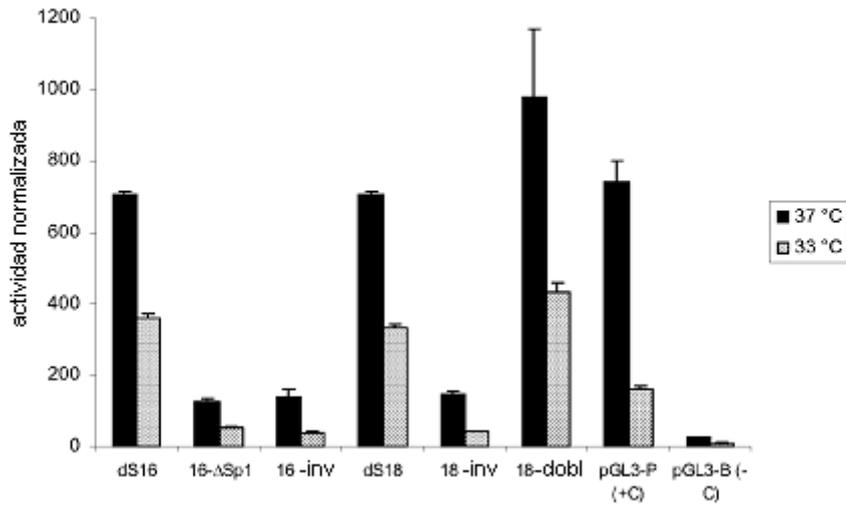


Fig. 6:

