



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 356 910**

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

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04702315 .5**

96 Fecha de presentación : **15.01.2004**

97 Número de publicación de la solicitud: **1583832**

97 Fecha de publicación de la solicitud: **12.10.2005**

54 Título: **Construcciones para la expresión inducible de ARN de interferencia pequeño (ARNs) para el silenciamiento génico seleccionado.**

30 Prioridad: **17.01.2003 EP 03001058**

45 Fecha de publicación de la mención BOPI:
14.04.2011

45 Fecha de la publicación del folleto de la patente:
14.04.2011

73 Titular/es: **Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.**
Hofgartenstrasse 8
80539 Munchen, DE

72 Inventor/es: **Tuschl, Thomas;**
Achsel, Tilmann;
Lührmann, Reinhard y
Dammann, Jutta

74 Agente: **Lehmann Novo, María Isabel**

ES 2 356 910 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

Constructos para la expresión inducible de ARN de interferencia pequeño (ARNsi) para el silenciamiento génico seleccionado.

La invención se refiere a vectores para la expresión inducible de moléculas de ARN en células y organismos eucariotas, particularmente de mamíferos.

Se ha demostrado que los ARN bicatenarios pequeños de aproximadamente 20 a 30 pares de bases dirigen de una forma específica de la secuencia la degradación de ARNm en células de mamíferos (McManus y Sharp 2002). Estos ARN de interferencia pequeños (ARNsi) tienen preferentemente una longitud de 21 nucleótidos (nt), y están emparejados de manera que tienen un tallo de 19 pares de bases y extremos salientes en 3' con dos nt (Elbashir et al. 2001b; Elbashir et al. 2001a; Elbashir et al. 2001c; Elbashir et al. 2002). Tales dúplex de ARNsi se pueden suministrar a células de mamíferos mediante microinyección, transfección o electroporación, y se pueden convertir en una nueva clase de agentes terapéuticos dirigidos contra genes que se han asociado con patogénesis, tales como genes víricos, destruyendo sus ARNm y evitando de ese modo su expresión (Paddison y Hannon 2002; Tuschl y Borkhardt 2002). El ARN bicatenario más largo de 30 pares de bases puede activar la respuesta de interferón, provocando la detención traduccional no específica y la apoptosis; estos efectos no se han observado con los ARNs más cortos (Bitko y Barik 2001; Elbashir et al. 2001b).

Más recientemente, se descubrió una nueva clase de genes que codifican bucles de horquilla de ARNs corto, de alrededor de 25 a 30 pares de bases de longitud, que se procesan en ARN pequeños de 21 a 23 nucleótidos (Lagos-Quintana et al. 2001; Lau et al. 2001; Lee y Ambros 2001; Lagos-Quintana et al. 2002). Esta clase se denominó los microARN. Los microARN funcionan en la misma ruta que los ARNsi, asociándose con las proteínas Argonauta, que son necesarias para guiar el reconocimiento del ARNm diana (Hutvagner y Zamore 2002; Martinez et al. 2002; Mourelatos et al. 2002). Los miARN escinden ARNm diana complementarios en plantas (Llave et al. 2002; Rhoades et al. 2002), pero parece que reprimen la traducción de ARNm en lugar de la escisión del ARNm en animales (Hutvagner y Zamore 2002).

Para experimentos de selección de dianas génicas, hasta hace poco se introdujeron ARNsi en células vía métodos de transferencia génica clásicos, tales como transfección mediada por liposomas, electroporación, o microinyección, que requieren la síntesis química o enzimática de los ARNsi antes de su aplicación; pero los ARNsi también se pueden generar intracelularmente mediante expresión de los ARNsi a partir de ADN plasmídico o constructos retrovíricos, lentivíricos o adenovíricos (Barton y Medzhitov 2002; Brummelkamp et al. 2002a; Brummelkamp et al. 2002b; Devroe y Silver 2002; McManus et al. 2002; Miyagishi y Taira 2002; Xia et al. 2002; Zeng et al. 2002). La transcripción intracelular de moléculas de ARN pequeñas es posible clonando los moldes de ARNsi en unidades de transcripción de ARN polimerasa III (pol III), que normalmente codifican el ARN nuclear pequeño U6 o el resto de ARN H1 de la ARNasa P humana. Se han desarrollado dos enfoques para expresar los ARNsi: (1) las hebras sentido y antisentido que constituyen el dúplex de ARNsi son transcritas por promotores individuales (Lee et al. 2002; Miyagishi y Taira 2002), o (2) los ARNsi son expresados como estructuras de tallo-bucle plegadas hacia atrás que dan lugar a los ARNsi tras el procesamiento intracelular (Brummelkamp et al. 2002b; Paul et al. 2002). Se piensa que la expresión endógena de los ARNsi a partir de moldes de ADN introducidos supera algunas limitaciones del suministro de ARNsi exógeno, en particular la pérdida transitoria de fenotipo.

Los promotores de ARN U6 y H1 son miembros del tipo III de promotores de pol III (Paule y White 2000). Estos promotores son inusuales por cuanto casi todos sus elementos, con la excepción del primer nucleótido transcrito (posición + 1), están localizados en dirección 5' de la región transcrita, de forma que se puede transcribir casi cualquier secuencia insertada más corta de 400 nt. Por lo tanto, son muy adecuados para la expresión de los ARNsi de aproximadamente 21 nt o los tallo-bucles de ARN de aproximadamente 50 nt. El promotor U6 y el promotor H1 tienen un tamaño diferente, pero contienen los mismos elementos de secuencia conservados o sitios de unión a proteínas (Myslinski et al. 2001). El nucleótido en + 1 de los promotores similares a U6 es siempre guanósina, y siempre adenosina para H1. De forma interesante, el cambio de la adenosina en + 1 por U, C o G en las secuencias de tallo-bucle expresadas por H1 no parece afectar al silenciamiento génico, sugiriendo por lo tanto que los promotores H1 pueden ser más flexibles que los promotores U6 para los cambios de secuencia en + 1, o pueden ser capaces de iniciar la transcripción en el primer nucleótido de purina en dirección 3' codificado por el ADN molde (Brummelkamp et al. 2002b). La transcripción del ARN termina cuando pol III encuentra un tramo de 4 ó 5 timidinas después de la incorporación de varios restos de uridina (Myslinski et al. 2001).

Sin embargo, para aplicaciones prácticas, también se deben de considerar el tiempo extra considerable implicado en preparar y amplificar vectores de expresión de ARNsi y la eficiencia de la transfección de plásmidos, con relación a los ARNsi. Adicionalmente, la selección como dianas de genes esenciales provoca la detención en el crecimiento celular o la muerte celular en uno a tres días después del suministro de los ARNsi, haciendo así innecesario el silenciamiento a largo plazo, si no imposible; sin embargo, el desarrollo de sistemas de expresión inducible de ARNsi puede proporcionar una alternativa interesante en tales casos (Ohkawa y Taira 2000). Sin embargo, cuando se seleccionan como dianas proteínas no esenciales, las células estables a las que se les ha reducido la expresión de algún gen (células *knockdown*) pueden ser de gran valor cuando se estudian procesos inducibles tales como la respuesta al daño por irradiación con UV, interacciones hospedante-patógeno, o

diferenciación celular. A fin de superar la limitación de los vectores actualmente disponibles seleccionadores de dianas, se ha explorado la posibilidad de insertar secuencias de unión a proteínas reguladoras en la región promotora de promotores de pol III.

5 Un primer aspecto de la presente invención es un vector recombinante para la expresión inducible de una molécula de ARN monocatenaria o bicatenaria en una célula eucariota, particularmente de mamífero, que comprende al menos una secuencia que codifica la molécula de ARN ligada operativamente a una secuencia de control de la expresión que comprende un promotor de polimerasa III y varios sitios de unión a proteínas reguladoras, y opcionalmente un terminador de la transcripción, en el que al menos un sitio de unión a proteínas reguladoras está localizado en 5', y un sitio de unión a proteínas reguladoras está localizado en 3' con respecto a la caja TATA del promotor de polimerasa III, y en el que los sitios de unión a proteínas reguladoras son sitios de unión para una proteína represora. El promotor de polimerasa III y los sitios de unión a proteínas reguladoras están localizados en 5' con respecto a la secuencia codificante, y el terminador está localizado en 3' con respecto a la secuencia codificante.

15 El vector puede ser cualquier vector que sea adecuado para la transfección de células eucariotas, por ejemplo un vector de ADN o de ARN. El vector puede ser un plásmido, por ejemplo un plásmido lineal o circular, un cósmido, un vector vírico, por ejemplo un adenovirus, un retrovirus, un virus adenoasociado, un virus de la vacuna, un lentivirus o un cromosoma artificial. El vector puede ser un vector extracromosómico, o un vector que es capaz de integrarse en el genoma de una célula hospedante. Los vectores apropiados son bien conocidos en la técnica, y se describen en Sambrook et al. (1998), *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, y en Ausubel et al. (1998), *Current Protocols in Molecular Biology*, John Wiley & Sons, por ejemplo.

20 La molécula de ARN que está siendo expresada por el vector puede ser cualquier molécula de ARN que tenga una longitud de 15-500 nucleótidos, preferiblemente de 20-400 nucleótidos. Por ejemplo, el ARN puede ser un ARNt, un ARNsn o un microARN. Sin embargo, preferiblemente, el ARN es una molécula de ARN que es capaz de producir la interferencia del ARN, o una molécula que se procesa, por ejemplo mediante mecanismos celulares, para proporcionar una molécula que es capaz de producir la interferencia del ARN. En una realización especialmente preferida, la molécula de ARN es una molécula de ARN monocatenaria que tiene una longitud de 30-100, más preferiblemente 40-80 nucleótidos. La molécula de ARN monocatenaria tiene una porción que es al menos sustancialmente complementaria a un transcrito diana, es decir, un transcrito, particularmente un ARNm, que es expresado en una célula diana. Además, se prefiere que la molécula de ARN monocatenaria sea capaz de formar una estructura de horquilla bicatenaria. La estructura de horquilla tiene preferiblemente un saliente en 3' que puede tener una longitud de 1-5 nucleótidos, más preferiblemente de 1-3 nucleótidos. En una realización adicional preferida, la molécula de ARN es una molécula de ARN bicatenaria que comprende 2 moléculas de ARN monocatenarias que son expresadas individualmente por uno o varios vectores, y que son capaces de formar una molécula de ARN bicatenaria. La molécula de ARN bicatenaria comprende una hebra antisentido, que es al menos sustancialmente complementaria a un transcrito diana, y una hebra sentido, que es al menos sustancialmente complementaria a la hebra antisentido. Cada hebra tiene una longitud de preferiblemente 19-30, más preferiblemente de 19-25 nucleótidos. La molécula de ARN bicatenaria tiene preferiblemente al menos un saliente en 3' que tiene una longitud de 1-5 nucleótidos, más preferiblemente de 1-3 nucleótidos.

35 La secuencia de control de la expresión comprende un promotor de polimerasa III, más particularmente un promotor de polimerasa III que es reconocido en una célula hospedante predeterminada, por ejemplo una célula de mamífero, particularmente una célula humana. Ejemplos apropiados de promotores de polimerasa III son el promotor H1, el promotor U6, un promotor de ARNt, u otros promotores de polimerasa III.

40 La secuencia de control de la expresión comprende varios sitios de unión a proteínas reguladoras, por ejemplo 2 ó 3 sitios de unión a proteínas reguladoras, en la que al menos un sitio de unión a proteínas reguladoras está localizado en 5', y un sitio de unión a proteínas reguladoras está localizado en 3', con respecto a la caja TATA del promotor de polimerasa III. El sitio de unión a proteínas reguladoras permite la unión específica de una proteína reguladora, y tiene preferiblemente una longitud de hasta 25 nucleótidos, más preferiblemente hasta 20 nucleótidos.

45 La proteína reguladora puede ser cualquier proteína que sea capaz de la unión, específica de la secuencia, a una secuencia de ácido nucleico, por ejemplo un represor, un factor de transcripción, un receptor nuclear, etc. Un ejemplo preferido de una proteína reguladora es el represor de tetraciclina, que es capaz de unirse, de forma específica de la secuencia, a una secuencia de ácido nucleico corta. La unión del represor de tetraciclina a su sitio de unión en la secuencia de control de la expresión conduce a una represión de la transcripción. La unión al represor de tetraciclina se puede anular añadiendo tetraciclina o un derivado de la misma, por ejemplo doxiciclina, al medio de cultivo, dando como resultado una inducción de la expresión del ARN. La proteína reguladora que reconoce el sitio de unión es capaz de ser expresada en una célula que se ha transfectado con el vector de la invención. La expresión de la proteína reguladora puede ser constitutiva o regulable. La proteína reguladora puede ser una proteína endógena a la célula transfectada, o una proteína exógena a la célula transfectada, es decir, una proteína codificada por una secuencia de ácido nucleico que se ha insertado en la célula por medios recombinantes. Por ejemplo, el ácido nucleico que codifica la proteína reguladora puede estar localizado en el vector de la invención, en el cual está ligado operativamente a una secuencia de control de la expresión apropiada.

Preferiblemente, la secuencia de control de la expresión comprende además un terminador de la transcripción. El terminador de la transcripción puede ser un terminador de origen natural, por ejemplo el terminador U6 o el H1, o un terminador sintético. Preferiblemente, el terminador comienza con una secuencia de varios nucleótidos T.

5 La secuencia de la molécula de ARN codificada por el vector de la presente invención tiene que tener una complementariedad suficiente con la molécula diana de ácido nucleico, a fin de mediar la interferencia del ARN específico de la diana. Más particularmente, una porción de la molécula de ARN es sustancialmente complementaria al transcrito diana.

10 La reacción de escisión del ARN diana guiada por las moléculas de ARN codificadas por el vector de la presente invención es muy específica de la secuencia. Sin embargo, no todas las posiciones de la molécula de ARN contribuyen por igual al reconocimiento de la diana. Se toleran desemparejamientos, particularmente en el término 3' de la molécula de ARN, más particularmente en los restos 3' con respecto a los primeros 20 nt de la molécula de ARN. Se prefieren especialmente las moléculas de ARN monocatenarias o las hebras antisentido de moléculas de ARN bicatenarias que tienen en el término 5' al menos 15 y preferiblemente al menos 20 nucleótidos que son completamente complementarios a un transcrito diana predeterminado, o tiene sólo un desemparejamiento, y opcionalmente hasta 15 nucleótidos en el término 3', que puede contener uno o varios, por ejemplo 2, 3 o más desemparejamientos.

La Solicitud describe además un método para mediar la interferencia de ARN en una célula o un organismo, que comprende las etapas de:

20 (a) poner en contacto la célula u organismo con el vector de la invención en condiciones en las que se reprimen las modificaciones de ácidos nucleicos específicos de la diana mediadas por la molécula de ARN codificada por el vector, y

25 (b) inducir una modificación de ácidos nucleicos específicos de la diana efectuada por la molécula de ARN codificada por el vector con relación a un ácido nucleico diana que tiene una porción de secuencia sustancialmente complementaria a la molécula de ARN.

30 Preferiblemente, la etapa de puesta en contacto (a) comprende introducir el vector en una célula diana, por ejemplo una célula diana aislada, por ejemplo en un cultivo celular, un microorganismo unicelular o una célula diana o una pluralidad de células diana dentro de un organismo multicelular. Más preferiblemente, la etapa de introducción comprende un suministro mediado por el vehículo, por ejemplo mediante vehículos liposómicos y/o mediante inyección, o mediante suministro por electroporación, precipitación con fosfato de calcio, infección vírica, etc. Otros sistemas de suministro adecuados incluyen Oligofectamine (Invitrogen) y el reactivo de transfección de ARNsi Transit-TKO (Mirus).

El método se puede usar para determinar la función de un gen en una célula o un organismo, o incluso para modular la función de un gen en una célula o un organismo, que es capaz de mediar la interferencia del ARN.

35 La célula es preferiblemente una célula eucariota o una estirpe celular, por ejemplo una célula vegetal o una célula animal, tal como una célula de mamífero, por ejemplo una célula embrionaria, una célula madre pluripotente, una célula tumoral, por ejemplo una célula de teratocarcinoma, o una célula infectada con un virus. El organismo es preferiblemente un organismo eucariota, por ejemplo una planta o un animal, tal como un mamífero, particularmente un ser humano.

40 El gen diana al que se dirige la molécula de ARN puede estar asociado con una patología. Por ejemplo, el gen puede ser un gen asociado a un patógeno, por ejemplo un gen vírico, un gen asociado a un tumor, o un gen asociado a una enfermedad autoinmunitaria. El gen diana también puede ser un gen heterólogo, expresado en una célula recombinante o un organismo genéticamente alterado. Determinando o modulando, particularmente inhibiendo, la función de tal gen, se puede obtener información valiosa y beneficios terapéuticos en el campo agrícola o en el campo de la medicina o de la medicina veterinaria.

45 La presente invención también permite una selección muy específica de transcritos en una célula o en un organismo, por ejemplo la selección de isoformas de transcritos individuales o polimorfismos de transcritos.

50 El vector se puede administrar como una composición farmacéutica. La administración se puede llevar a cabo por métodos conocidos, en los que se introduce in vitro o in vivo un ácido nucleico en una célula diana deseada. Las técnicas de transferencia génica usadas habitualmente incluyen fosfato de calcio, DEAE-dextrano, electroporación y microinyección, y métodos víricos (Graham, F.L. y van der Eb, A.J. (1973) *Virology* 52, 456; McCutchan, J.H. y Pagano, J.S. (1968), *J. Natl. Cancer Inst.* 41, 351; Chu, G. et al (1987), *Nucl. Acids Res.* 15, 1311; Fraley, R. et al. (1980), *J. Biol. Chem.* 255, 10431; Capecchi, M.R. (1980), *Cell* 22, 479). Una adición reciente a este arsenal de técnicas para la introducción de ácidos nucleicos en células es el uso de liposomas catiónicos (Felgner, P.L. et al. (1987), *Proc. Natl. Acad. Sci USA* 84, 7413). Las formulaciones de lípidos catiónicos comercialmente disponibles son, por ejemplo, Tfx 50 (Promega) o Lipofectamin2000 (Life Technologies). Un método preferido adicional para la introducción de ácidos nucleicos en un organismo diana, particularmente en un ratón, es

la inyección a alta presión en la vena de la cola (Lewis, D.L. et al. (2002), Nat. Genet. 29, 29; McCaffrey, A.P. et al. (2002), Nature 418, 38-39).

De este modo, la invención también se refiere a una composición farmacéutica que contiene, como agente activo, al menos un vector como se describe anteriormente, y vehículos, diluyentes y/o adyuvantes farmacéuticamente aceptables. La composición se puede usar para aplicaciones de diagnóstico y terapéuticas en medicina humana o en medicina veterinaria.

Para las aplicaciones de diagnóstico o terapéuticas, la composición puede estar en forma de una disolución, por ejemplo una disolución inyectable, una crema, ungüento, comprimido, suspensión o similar. La composición se puede administrar en cualquier forma adecuada, por ejemplo mediante inyección, mediante aplicación oral, tópica, nasal, rectal, etc. El vehículo puede ser cualquier vehículo farmacéutico adecuado. Preferiblemente, se usa un vehículo que es capaz de incrementar la eficacia de las moléculas del vector para entrar en las células diana. Los ejemplos adecuados de tales vehículos son liposomas, particularmente liposomas catiónicos. Un método de administración preferido adicional es la inyección.

Una aplicación preferida adicional del método de ARNi es un análisis funcional de células eucariotas, u organismos no humanos eucariotas, preferiblemente células de mamíferos u organismos mamíferos, y lo más preferible células humanas, por ejemplo estirpes celulares tales como Hela o 293, o roedores, por ejemplo ratas y ratones. Mediante transfección con moléculas del vector que son homólogas a un gen diana predeterminado que codifica una molécula de ARN adecuada, se puede obtener un fenotipo nuligénico (*knockout*) específico en una célula diana, por ejemplo en un cultivo celular o en un organismo diana. El fenotipo nuligénico se puede regular por inducción o represión de la secuencia de control de la expresión en el vector de la invención.

Además, la invención se refiere a una composición farmacéutica para inhibir la expresión de un transcrito diana mediante interferencia del ARN, que comprende como agente activo un vector como se describe anteriormente.

Además, la Solicitud describe un método para la monitorización, prevención o tratamiento de una enfermedad asociada con la sobreexpresión de al menos un gen diana, que comprende administrar a un sujeto que lo necesite un vector como se describe anteriormente.

Todavía una materia objeto adicional de la presente invención es una célula eucariota o un animal transgénico no humano que se transfecta con un vector como se describe anteriormente. La célula puede ser una célula de mamífero, por ejemplo una célula humana. El animal transgénico no humano puede ser un mamífero, por ejemplo un ratón, rata, hámster o mono. La célula o animal se puede transfectar transitoria o establemente. Se prefiere una transfección estable.

Además, la invención se refiere a una población celular clonal derivada de una célula como se describe anteriormente, en la que la población celular clonal consiste sustancialmente en células transfectadas, y está libre de células no transfectadas.

La célula eucariota u organismo no humano eucariota muestra un fenotipo nuligénico específico del gen diana inducible, que comprende una expresión al menos parcialmente deficiente de al menos un ARN génico diana endógeno. Se debería observar que la presente invención permite la transfección simultánea o subsiguiente con varios vectores que codifican moléculas de diferentes secuencias, que están emparentadas con el mismo gen diana o con uno diferente.

Los fenotipos nuligénicos específicos de genes de células u organismos no humanos, particularmente de células humanas o mamíferos no humanos, se pueden usar en procedimientos analíticos, por ejemplo en el análisis funcional y/o fenotípico de procesos fisiológicos complejos tales como el análisis de los perfiles de expresión génica y/o proteomas. Por ejemplo, se pueden preparar los fenotipos nuligénicos de genes humanos en células cultivadas que se supone que son reguladores de los procesos de corte y empalme alternativos. Entre estos genes están particularmente los miembros de la familia del factor de corte y empalme SR, por ejemplo ASF/SF2, SC35, SRp20, SRp40 o SRp55. Además, se puede analizar el efecto de las proteínas SR sobre los perfiles de ARNm de genes predeterminados cortados y empalmados alternativamente, tales como CD44. Preferiblemente, el análisis se lleva a cabo por métodos de alto rendimiento usando matrices a base de oligonucleótidos.

Usando tecnologías nuligénicas a base de ARNi, se puede inhibir la expresión de un gen diana endógeno en una célula diana o un organismo diana. El gen endógeno se puede complementar mediante un ácido nucleico diana exógeno que codifica la proteína diana o una variante o forma mutada de la proteína diana, por ejemplo un gen o un ADNc, que opcionalmente puede estar fusionado a una secuencia de ácido nucleico adicional que codifica un péptido o polipéptido detectable, por ejemplo un marcador de afinidad, particularmente un marcador de afinidad múltiple. Las variantes o formas mutadas del gen diana difieren del gen diana endógeno por cuanto codifican un producto génico que difiere del producto génico endógeno a nivel de aminoácidos por sustituciones, inserciones y/o supresiones de aminoácidos individuales o múltiples aminoácidos. Las variantes o formas mutadas pueden tener la misma actividad biológica que el gen diana endógeno. Por otro lado, el gen diana variante o mutado puede tener también una actividad biológica, que difiere de la actividad biológica del gen diana endógeno, por ejemplo una

actividad parcialmente suprimida, una actividad completamente suprimida, una afinidad potenciada, etc.

La complementación se puede lograr coexpresando el polipéptido codificado por el ácido nucleico exógeno, por ejemplo una proteína de fusión que comprende la proteína diana y el marcador de afinidad y la molécula de ARN para anular el gen endógeno en la célula diana. Esta coexpresión se puede lograr usando un vector de expresión adecuado que expresa tanto el polipéptido codificado por el ácido nucleico exógeno, por ejemplo la proteína diana modificada con el marcador, como la molécula de ARN, o, como alternativa, usando una combinación de vectores de expresión. Las proteínas y complejos proteicos que se sintetizan de novo en la célula diana contendrán el producto génico exógeno, por ejemplo la proteína de fusión modificada. A fin de evitar la supresión de la expresión del producto génico exógeno por la molécula de ARNi, la secuencia nucleotídica que codifica el ácido nucleico exógeno se puede alterar a nivel del ADN (provocando o no mutaciones a nivel de los aminoácidos) en la parte de la secuencia que es homóloga a la molécula de ARN. Como alternativa, el gen diana endógeno se puede complementar mediante secuencias nucleotídicas correspondientes procedentes de otras especies, por ejemplo de ratón.

Las aplicaciones preferidas para la célula u organismo de la invención es el análisis de perfiles de expresión génica y/o proteomas. En una realización especialmente preferida, se lleva a cabo un análisis de una variante o forma mutada de una o varias proteínas diana, en el que dicha variante o formas mutadas se reintroducen en la célula u organismo mediante un ácido nucleico diana exógeno como se describe anteriormente. La combinación de la anulación de un gen exógeno y el rescate usando una diana exógena mutada, por ejemplo parcialmente suprimida, tiene ventajas en comparación con el uso de una célula nuligénica. Además, este método es particularmente adecuado para identificar dominios funcionales de la proteína diana. En una realización preferida adicional, se lleva a cabo una comparación, por ejemplo, de los perfiles de expresión génica y/o proteomas y/o características fenotípicas de al menos dos células u organismos. Estos organismos se seleccionan de:

- (i) una célula de control u organismo de control sin la inhibición del gen diana,
- (ii) una célula u organismo con inhibición del gen diana, y
- (iii) una célula u organismo con inhibición del gen diana más complementación del gen diana mediante un ácido nucleico diana exógeno.

El método y la célula también se pueden usar en un procedimiento para identificar y/o caracterizar agentes farmacológicos, por ejemplo identificar nuevos agentes farmacológicos a partir de una colección de sustancias de ensayo y/o mecanismos caracterizadores de acción y/o efectos secundarios de agentes farmacológicos conocidos.

De este modo, la Solicitud también describe un sistema para identificar y/o caracterizar agentes farmacológicos que actúan sobre al menos una proteína diana, que comprende:

- (a) una célula eucariota o un organismo no humano eucariota capaz de expresar al menos un gen diana endógeno que codifica dicha proteína diana,
- (b) al menos un vector como se describe anteriormente que codifica una molécula de ARN capaz de inhibir la expresión de dicho al menos un gen diana endógeno por ARNi, y
- (c) una sustancia de ensayo o una colección de sustancias de ensayo en la que se han de identificar y/o caracterizar propiedades farmacológicas de dicha sustancia de ensayo o dicha colección.

Además, el sistema como se describe anteriormente comprende preferiblemente:

- (d) al menos un ácido nucleico diana exógeno que codifica la proteína diana o una variante o forma mutada de la proteína diana, en el que dicho ácido nucleico diana exógeno difiere del gen diana endógeno a nivel del ácido nucleico de forma que la expresión del ácido nucleico diana exógeno está sustancialmente menos inhibida por la molécula de ARN que la expresión del gen diana endógeno.

Adicionalmente, el método de complementación nuligénico del ARN se puede usar con fines preparativos, por ejemplo para la purificación mediante afinidad de proteínas o complejos proteicos procedentes de células eucariotas, particularmente células de mamíferos, y más particularmente células humanas. En esta realización de la invención, el ácido nucleico diana exógeno codifica preferiblemente una proteína diana que se fusiona a un marcador de afinidad.

Se puede emplear el método preparativo para la purificación de complejos proteicos de peso molecular elevado que tienen preferiblemente una masa de ≥ 150 kD, y más preferiblemente de ≥ 500 kD, y que opcionalmente pueden contener ácidos nucleicos tales como ARN. Los ejemplos específicos son el complejo proteico heterotrimérico que consiste en las proteínas de 20 kD, 60 kD y 90 kD de la partícula U4/U6 snRNP, el factor de corte y empalme SF3b de la 17S U2 snRNP, que consiste en 5 proteínas que tienen pesos moleculares de 14, 49, 120, 145 y 155 kD, y la partícula 25S U4/U6/U5 tri-snRNP, que contiene las moléculas de ARNs U4, U5 y U6 y

alrededor de 30 proteínas, que tienen un peso molecular de alrededor de 1,7 MD.

Adicionalmente, la presente invención se explica con más detalle en las siguientes Figuras y Ejemplos.

Figuras

5 Figura 1. Secuencias de la secuencia del promotor H1 sin modificar, que incluye la secuencia de horquilla de GL2 y las secuencias terminadoras para polimerasa III. Las secuencias están unidas perfectamente entre sí en el plásmido.

Figura 2. Ilustración de la posición de los sitios de unión a tetO insertados en la región del promotor H1. La transcripción comenzaría inmediatamente en dirección 3' del nucleótido más hacia el extremo de 3' mostrado. Preferiblemente, el primer nucleótido transcrito codifica un nucleótido de purina.

10 Figura 3. Selección como diana de luciferasa GL2 con vectores de expresión de horquilla dirigidos contra GL2.

Figura 4. Selección como diana de luciferasa GL2 con vectores de expresión de horquilla dirigidos contra GL2. Los constructos de horquilla del promotor H1 se transfectaron un día antes de la transfección del plásmido informador, para permitir la expresión del ARN de horquilla antes de la expresión de los plásmidos informadores.

15 Figura 5. Secuencia de ácidos nucleicos del plásmido pH1-tetO4 usado para el silenciamiento inducible de luciferasa GL2.

Ejemplos

1. Materiales y Métodos

1.1 Construcción de plásmido

20 Para la construcción plasmídica de constructos del promotor H1, se prepararon oligonucleótidos sintéticos, y se hibridaron y ligaron entre sí usando T4 ARN ligasa. El producto de ligación se purificó entonces en gel y se ligó en el vector secuenciador pBluescript SK(-) (Stratagene) que estaba restringido con XhoI y BamHI, y se purificó en gel para eliminar el fragmento cortado del sitio de clonación múltiple.

25 Para generar el inserto de H1-tetO4, se sintetizaron las siguientes hebras oligonucleotídicas que forman el inserto de orientación sentido,

5'-TCGAAGATCTAATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAAT,
5'-pCCCTATCAGTGATAGAGACTTATAAGTCCCTATCAGTGATAGAGAATTCATTTCGAAGTATTCCG
CGTACGTTTCGACG, y

30 5'-pTACGCGGAATACTTCGAAATTTTTCTCGAGCTTCCTTCGGGAAGCTCTCCATATTTTTTG
se combinaron con los siguientes oligonucleótidos formando el inserto de orientación antisentido,

5'-TATGGTGATTTCCAGAACACATAGCGACATGCAAATATTAGATCT,
5'-pTACGCGGAATACTTCGAAATGAATTCTCTATCACTGATAGGGAAGCTTATAAGTCTCTATCACTGA
TAGGGATTTACGTT, y

35 5'-pGATCCAAAAAATATGGAGAGCTTCCCGAAGGAAGCTCGAGAAAAAATTTTGAAGTATTCCGCG
TACGTCGAAACG.

El sitio XhoI ya no existe con la ligación del inserto del promotor H1; el sitio BamHI permanece intacto. Para insertar una nueva secuencia de horquilla, la secuencia de ARN de horquilla de GL2 se puede cortar digiriendo el plásmido con EcoRI y XhoI y ligando los oligonucleótidos sintéticos, prehibridados entre sí en ese sitio.

1.2 Transfección y análisis de la expresión reducida

40 Por pocillo de una placa de 24 pocillos, se transfectaron 1 µg de plásmido de luciferasa de luciérnaga pGL2-SV40, 0,2 µg de plásmido de luciferasa de pensamiento de mar pRL-TK, y 1 µg de plásmido de horquilla del promotor H1 usando Lipofectamine 2000 (Invitrogen) según el protocolo del fabricante. Se usó doxiciclina a una concentración de 0,2 µg/ml en el medio. El medio que contiene el reactivo de transfección se sustituyó con medio reciente 8 h después de la transfección. Las células se recolectaron mediante tripsinación, y se lisaron según las

instrucciones del kit de luciferasa dual (Promega). Después, se monitorizó la actividad de luciferasa. Si las células se cebaron en primer lugar con plásmidos de horquilla de H1, se usaron 2 μ g de ADN plasmídico para la transfección. El medio que contiene el reactivo de transfección se sustituyó con medio reciente 8 h después de la transfección, y la transfección se repitió con la mezcla de plásmidos informador y H1 como se describe.

5 2. Resultados

Se construyó un total de 5 promotores derivados de H1 para conducir la transcripción de una estructura de ARN de tallo-bucle dirigida contra la luciferasa de luciérnaga GL2. Todos contienen el mismo elemento de secuencia proximal, PSE, y el mismo terminador de la transcripción. El comienzo de la transcripción supuesto es el primer nucleótido del tallo-bucle como se enumera más abajo. La terminación supuesta es alrededor de dos a cuatro nucleótidos en la primera carrera de 6 timidinas de la porción terminadora.

La región alrededor de la caja TATA es variable: en H1, deriva del promotor H1 "apropiado". H1-tetO1, H1-tetO2 y H1-tetO3 contienen cada uno un sitio de unión para el represor de tetraciclina tetR. H1-tetO4 contiene dos sitios de unión a tetR.

Para comprobar el efecto de estos constructos in vivo, se cotransfectaron transitoriamente los plásmidos que contienen el casete de H1 con dos plásmidos informadores que codifican la luciferasa GL2 y la luciferasa de Renilla. El plásmido que codifica la luciferasa GL2 muestra una homología de secuencia con el tallo-bucle expresado; su expresión debería ser suprimida por el tallo-bucle. El plásmido que codifica la luciferasa de pensamiento de mar (Renilla) (pRL-TK), por el contrario, no muestra homología de secuencia, y no se ve afectado. Por lo tanto, se calculó la relación de la actividad de luciferasa GL2 con respecto a la actividad de luciferasa de Renilla (las dos enzimas se pueden distinguir puesto que usan sustratos diferentes). La relación obtenida en presencia de un vector vacío en lugar del casete de H1 se ajustó arbitrariamente a 1. De este modo, una relación menor que 1 indica expresión reducida de la luciferasa GL2, y por tanto supresión del gen. En este estudio, se usaron células HeLa "T-Rex" (Invitrogen) que expresan constitutivamente la proteína tetR. Se anticipa que los constructos de H1-tetO son completamente activos sólo en presencia del análogo de tetraciclina doxiciclina (dox) cuando la proteína tetR no se une a su elemento sensible en el ADN. Obsérvese que las actividades de luciferasa varían ligeramente en presencia o ausencia de doxiciclina. Por lo tanto, los controles respectivos se ajustaron independientemente a 1. Las actividades de luciferasa se evaluaron dos días después de la transfección, para permitir la expresión completa del tallo-bucle represor. Después de tres días, las actividades absolutas estaban enormemente disminuidas debido a la naturaleza transitoria de la transfección, pero las relaciones son en gran parte las mismas.

Como se muestra en la figura 3, la cotransfección del casete de H1 reduce la actividad de GL2 hasta 40% del control, independientemente de la presencia o ausencia de doxiciclina. Se obtienen valores similares para todos los constructos de H1-tetO en presencia de doxiciclina, es decir, en ausencia de la unión a tetR. En ausencia de doxiciclina, sin embargo, tetR se puede unir al promotor H1, y los constructos son menos activos: H1-tetO1 reduce la actividad de GL2 hasta 0,82, H1-tetO2 hasta 0,62, y H1-tetO3 hasta 0,58. De este modo, el represor de tet parece suprimir la transcripción del tallo-bucle de forma más eficaz cuando se une en dirección 5' de la caja TATA, en comparación con la dirección 3'. El efecto más drástico se observó para el constructo que contiene dos sitios de unión a tetR: en presencia de doxiciclina, este casete reprime la expresión de la luciferasa GL2, así como los otros constructos. En ausencia de doxiciclina, sin embargo, la actividad de luciferasa GL2 aparentemente incluso aumenta, hasta 1,09. El incremento está probablemente dentro del error experimental y de este modo no es significativo, pero la actividad de GL2 ciertamente no disminuye. Por lo tanto, el promotor H1 que contiene dos sitios de unión a tetR es inactivo en ausencia de doxiciclina, y de este modo es adecuado para la supresión regulada de un gen diana.

En este experimento, ninguno de los constructos redujo la actividad de luciferasa GL2 a menos de 0,4, dejando así un poso significativo de luciferasa expresada. Se cree que este poso no resulta de una incapacidad general del tallo-bucle para suprimir la luciferasa. En su lugar, se cree que la transcripción del tallo-bucle y la maduración en un complejo de RISC funcional toma más tiempo que la expresión de luciferasa activa, dejando así un poso significativo de luciferasa expresada en ausencia de tallo-bucles silenciadores. Adicionalmente, la relación de plásmidos introducidos en células individuales puede diferir, es decir, una célula puede obtener más del plásmido H1, y otra, menos. Aquellas células que no obtuvieron suficiente plásmido represor pueden haber contribuido a la actividad del poso. Para apoyar estas nociones, la expresión prolongada (hasta tres días después de la transfección) condujo a un "nivel de poso" reducido de la actividad de GL2: en este experimento, 0,26 a 0,33 (datos no mostrados).

De forma más importante, la transfección del casete H1 solamente, seguido de la cotransfección de las dos luciferasas más el casete H1 un día después, incrementa el efecto de la expresión del tallo-bucle (figura 4): en este conjunto de experimentos, las actividades de GL2 en presencia de doxiciclina oscilaron entre 20 y 30% del control. Como antes, el constructo H1-tetO4 tiene, en ausencia de doxiciclina, el mismo efecto que H1 sin los sitios de unión a tetR, mientras que, en presencia de doxiciclina, no es significativamente activo (0,94).

Como horquilla silenciadora, se seleccionó un tallo de 23 pares de bases, con lo que el tallo-bucle de la

secuencia UUGG se conectó al extremo 3' de la hebra antisentido complementaria al ARNm diana de GL2. La señal terminadora se situó de manera que pudiese estar presente un saliente 3' de oligoU de 2 a 4 nucleótidos en la horquilla transcrita. La secuencia del bucle también se puede intercambiar con una secuencia de seis nucleótidos de una enzima de restricción, que permite la linealización del plásmido antes de la secuenciación. La secuenciación de los ARN de horquilla es algunas veces incómoda debido a que la estructura de horquilla estable, a través de ella, puede evitar la lectura de la polimerasa secuenciadora. En el silenciamiento génico mediado por horquillas, se pueden usar tallos de 19 pares de bases o más (hasta 29 pares de bases).

3. Sumario

En conclusión, el casete H1 descrito aquí es capaz de reducir la expresión de un gen diana, y la variante H1-tetO4 permite el control completo de la represión mediante adición de doxiciclina al medio. Para demostrar que el constructo puede suprimir completamente un gen diana dado, se prefiere establecer permanentemente una estirpe celular que sea insensible a la transfección de luciferasa GL2, o que no exprese una proteína endógena. La proteína (endógena) se puede encender y apagar por la doxiciclina cuando se usa la variante H1-tetO4. Eso en particular prepararía el terreno para apagar genes esenciales de una manera fiable y reproducible.

El casete de expresión, cuando se integra en el contexto de una estructura principal plasmídica seleccionable, se puede integrar de forma estable en células de mamíferos. Las células de mamífero que también expresan de forma estable la proteína tetR se pueden inducir entonces para expresar el ARN de horquilla. Los genes esenciales se pueden apagar entonces en células en cualquier momento dado con la adición de doxiciclina al medio de cultivo tisular. Se pueden obtener poblaciones celulares homogéneas con expresión génica reducida, superando los problemas de la transfección no homogénea. Estas estirpes celulares serán de gran valor para los análisis de tipo de perfil proteómico o génico. También se puede usar un sistema de promotor/horquilla inducible en animales transgénicos para generar expresiones génicas reducidas (en lugar de expresiones génicas anuladas) alimentando tetraciclina o doxiciclina. Esto puede ser útil para generar modelos de animales de enfermedad que están provocadas por la reducción de la expresión génica endógena.

Referencias

Barton G.M., y Medzhitov R. 2002. Retroviral delivery of small interfering RNA into primary cells. *Proc. Natl. Acad. Sci. USA* 99: 14943-14945.

Bitko V., y Batik S. 2001. Phenotypic silencing of cytoplasmic genes using sequence-specific double-stranded short interfering RNA and its application in the reverse genetics of wild type negative-strand RNA viruses. *BMC Microbiol.* 1: 34.

Brummelkamp T., Bernards R., y Agami R. 2002a. Stable suppression of tumorigenicity by virus-mediated RNA interference. *Cancer Cell* 2: 243.

Brummelkamp T.R., Bernards R., y Agami R. 2002b. A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296: 550-553.

Devroe E., y Silver P.A. 2002. Retrovirus-delivered siRNA. *BMC Biotechnol* 2: 15.

Elbashir S.M., Lendeckel W., y Tuschl T. 2001a. RNA interference is mediated by 21 and 22 nt RNAs. *Genes & Dev.* 15: 188-200.

Elbashir S.M., Harborth J., Lendeckel W., Yalcin A., Weber K., y Tuschl T. 2001b. Duplexes of 21-nucleotide RNAs mediate RNA interference in mammalian cell culture. *Nature* 411: 494-498.

Elbashir S.M., Martinez J., Patkaniowska A., Lendeckel W., y Tuschl T. 2001c. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.* 20: 6877-6888.

Elbashir S.M., Harborth J., Weber K., y Tuschl T. 2002. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* 26: 199-213.

Hutvagner G., y Zamore P.D. 2002. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 1: 1.

Lagos-Quintana M., Rauhut R., Lendeckel W., y Tuschl T. 2001. Identification of novel genes coding for small expressed RNAs. *Science* 294: 853-858.

Lagos-Quintana M., Rauhut R., Yalcin A., Meyer J., Lendeckel W., y Tuschl T. 2002. Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* 12: 735-739.

Lau N.C., Lim L.P., Weinstein E.G., y Bartel D.P. 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294: 858-862.

Lee N.S., Dohjima T., Bauer G., Li H., Li M.J., Ehsani A., Salvaterra P., y Rossi J. 2002. Expression of small

- interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat. Biotech.* 20: 500-505.
- Lee R.C., y Ambros V. 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294: 862-864.
- 5 Llave C., Xie Z., Kasschau K.D., y Carrington J.C. 2002. Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* 297: 2053-2056.
- Martinez J., Patkaniowska A., Urlaub H., Lührmann R., y Tuschl T. 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* 110: 563-574.
- McManus M.T., y Sharp P.A. 2002. Gene silencing in mammals by small interfering RNAs. *Nat. Rev. Genet.* 3: 737-747.
- 10 McManus M.T., Petersen C.P., Haines B.B., Chen J., y Sharp P.A. 2002. Gene silencing using micro-RNA designed hairpins. *RNA* 8: 842-850.
- Miyagishi M., y Taira K. 2002. U6 promoter driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat. Biotech.* 20: 497-500.
- 15 Mourelatos Z., Dostie J., Paushkin S., Sharma A., Charroux B., Abel L., Rappsilber J., Mann M., y Dreyfuss G. 2002. miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes & Dev.* 16: 720-728.
- Myslinski E., Amé J.-C., Krol A., y Carbon P. 2001. An unusually compact external promoter for RNA polymerase III transcription of the human H1 RNA gene. *Nucleic Acids Res.* 29: 2502-2509.
- 20 Ohkawa J., y Taira K. 2000. Control of the functional activity of an antisense RNA by a tetracycline-responsive derivative of the human U6 snRNA promoter. *Hum. Gene Ther.* 11: 577-585.
- Paddison P.J., y Hannon G.J. 2002. RNA interference: the new somatic cell genetics? *Cancer Cell* 2: 17-23.
- Paul C.P., Good P.D., Winer I., y Engelke D.R. 2002. Effective expression of small interfering RNA in human cells. *Nat. Biotech.* 20: 505-508.
- Paule M.R., y White R.J. 2000. Transcription by RNA polymerase I and III. *Nucleic Acids Res.* 28: 1283-1298.
- 25 Rhoades M., Reinhart B., Lim L., Burge C., Bartel B., y Bartel D. 2002. Prediction of plant microRNA targets. *Cell* 110: 513.
- Tuschl T., y Borkhardt A. 2002. Small interfering RNAs - a revolutionary tool for analysis of gene function and gene therapy. *Mol. Intervent.* 2: 42-51.
- 30 Xia H., Mao Q., Paulson H.L., y Davidson B.L. 2002. siRNA-mediated gene silencing in vitro and in vivo. *Nat. Biotech.* 20: 1006-1010.
- Zeng Y., Wagner E.J., y Cullen B.R. 2002. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol. Cell* 9: 1327-1333.

REIVINDICACIONES

- 5 1.- Un vector recombinante para la expresión regulable de una molécula de ARN monocatenaria o bicatenaria en una célula eucariota, que comprende al menos una secuencia que codifica la molécula de ARN ligada operativamente a una secuencia de control de la expresión que comprende un promotor de ARN polimerasa III y varios sitios de unión a proteína represora, en el que al menos un sitio de unión a proteína represora está situado 5', y un sitio de unión a proteína represora está situado 3' con respecto a la caja TATA del promotor de ARN polimerasa III.
- 10 2.- El vector de la reivindicación 1, en el que la molécula de ARN codificada es capaz de provocar la interferencia del ARN, o se procesa para proporcionar una molécula de ARN capaz de provocar la interferencia del ARN.
- 3.- El vector de las reivindicaciones 1 ó 2, en el que la molécula de ARN es monocatenaria.
- 4.- El vector de la reivindicación 3, en el que la molécula de ARN tiene una longitud de 30-100 nucleótidos.
- 5.- El vector de la reivindicación 3 ó 4, en el que la molécula de ARN tiene una porción que es al menos sustancialmente complementaria a un transcrito diana.
- 15 6.- El vector de una cualquiera de las reivindicaciones 3-5, en el que la molécula de ARN es capaz de formar una estructura de horquilla.
- 7.- El vector de la reivindicación 6, en el que la estructura de horquilla tiene un saliente en 3'.
- 8.- El vector de las reivindicaciones 1 ó 2, en el que la molécula de ARN es bicatenaria.
- 20 9.- El vector de una cualquiera de las reivindicaciones 1-8, en el que el promotor de polimerasa III es un promotor H1.
- 10.- El vector de una cualquiera de las reivindicaciones 1-9, en el que los sitios de unión a proteína reguladora son sitios de unión para el represor de tetraciclina.
- 11.- El vector de una cualquiera de las reivindicaciones 1-10, en el que la secuencia de control de la expresión comprende además un terminador de la transcripción.
- 25 12.- El vector de una cualquiera de las reivindicaciones 1-11, para la inhibición de la expresión de un gen diana in vitro.
- 13.- El vector de una cualquiera de las reivindicaciones 1-11, para la inhibición de la expresión de un gen diana in vivo.
- 30 14.- Una composición farmacéutica que comprende al menos una molécula de ARN vector de una cualquiera de las reivindicaciones 1-13 como ingrediente activo, y vehículos, diluyentes y/o excipientes farmacéuticamente aceptables.
- 15.- La composición de la reivindicación 14, que es una formulación liposómica o lipídica catiónica.
- 16.- La composición de las reivindicaciones 14 ó 15, para aplicaciones de diagnóstico.
- 35 17.- La composición de la reivindicación 16, para monitorizar enfermedades asociadas con la sobreexpresión de al menos un gen diana.
- 18.- La composición de las reivindicaciones 14 ó 15, para aplicaciones terapéuticas.
- 19.- La composición de la reivindicación 18, para la prevención o tratamiento de enfermedades asociadas con la sobreexpresión de al menos un gen diana.
- 40 20.- La composición de las reivindicaciones 17 ó 19, en la que las enfermedades se seleccionan de enfermedades tumorales, enfermedades inflamatorias, enfermedades infecciosas, por ejemplo infecciones víricas, enfermedades degenerativas y enfermedades autoinmunitarias.
- 21.- Una célula eucariota que se transfecta con un vector de una cualquiera de las reivindicaciones 1-13.
- 22.- La célula de la reivindicación 21, que es una célula de mamífero.
- 45 23.- La célula de la reivindicación 22, que es una célula humana.

- 24.- La célula de una cualquiera de las reivindicaciones 21-22, que se transfecta de forma estable.
- 21-24. 25.- Una población celular clonal derivada de una célula de una cualquiera de las reivindicaciones
- 26.- La población celular de la reivindicación 25, que consiste en células transfectadas.
- 5 27.- Un animal transgénico no humano que se transfecta con un vector de una cualquiera de las reivindicaciones 1-13.
- 28.- El animal no humano de la reivindicación 27, que es un mamífero.

Fig 1

promotor H1	5' AATATTGCATGTCGCTAIGTGTTCIGGGAAATCACCATAAACGTGAAATGCTTTGGATTG
horquilla de GL2	GGAATCTTATAAGTTCTGTATGAGACCACTCGAATTC
terminador	5' ATTTCGAAGTATCCGCGTACGTTTCGACGTACGCGGAATACTTCGAAA
	5' TTTTCTCGAGCTTCCCTCGGGAAGCTCTCCATATTTTTGGATCC

Fig 2

H1 PSE ... TGICTTGGATTGGGAATCTTATAAGTTCTGTATGAGACCACTCGAATC ... GL2 ...
tetO1 PSE ... TCCCTATGACCTGATGACCTTATAAGTTCTGTATGAGACCACTCGAATC ... GL2 ...
tetO2 PSE ... TGICTTGGATTGGGAATCTTATAARCCCTATGACCTGATGACCTTATAAGTTCTGTATGAGACCACTCGAATC ... GL2 ...
tetO3 PSE ... TGICTTGGATTGGGAATCTTATAAGTRCCCTATGACCTGATGACCTTATAAGTTCTGTATGAGACCACTCGAATC ... GL2 ...
tetO4 PSE ... TCCCTATGACCTGATGACCTTATAAGTTCTGTATGAGACCACTCGAATC ... GL2 ...
 (verde: caja TATA; rojo: sitio de unión a tetR).

Fig 3

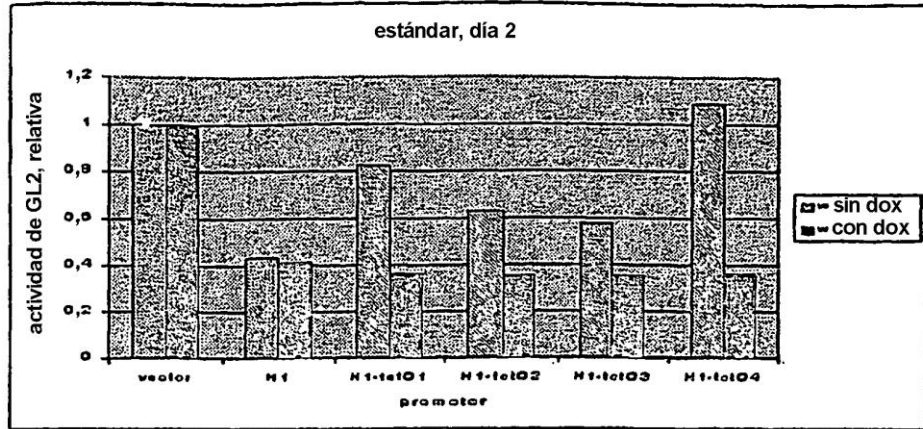


Fig 4

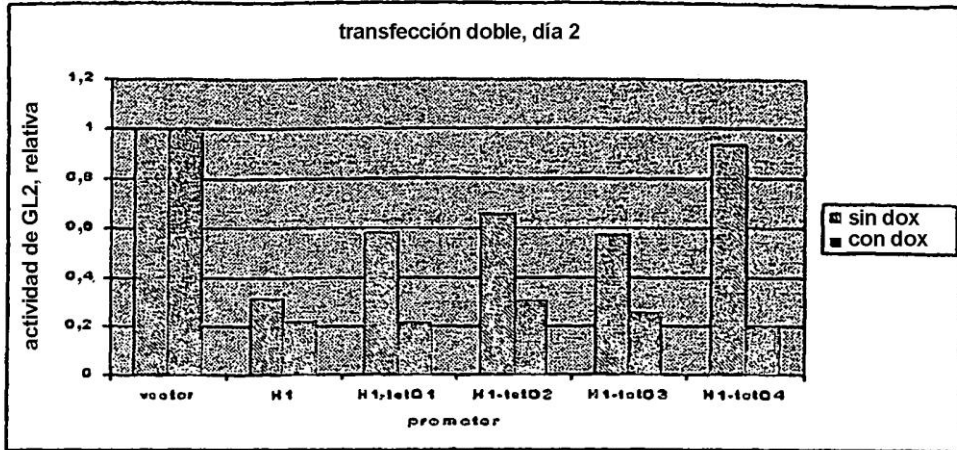


Fig 5

LOCUS	pH1-TetO4, derivado de pBluescript SK(-)					
DEFINICIÓN						
FUENTE	T. Achsel, J. Meyer, I. Tuschl, MPI f. biophys. Chem, Goettingen, Alemania					
ORGANISMO	desconocido					
CARACTERÍSTICAS	Localización/Cualificadores					
característica misc	749..752 / nota = "" / nombre estándar = "caja TATA"					
característica misc	orden (728..746,756..774) / nota = "" / nombre estándar = "sitio de tetO2"					
característica misc	679..778 / nota = "" / nombre estándar = "promotor H1"					
característica misc	orden (779..826,826..828) / nota = "" / nombre estándar = "ARN de horquilla de GL2"					
característica misc	668..673 / nota = "" / nombre estándar = "sitio de XhoI"					
terminador	828..868 / nota = "" / nombre estándar = "terminador"					
RECuento DE BASES	749 a	784 c	753 g	822 t		
ORIGEN						
1	CACCTGACGC	GCCTGTAGC	GGCGCATTAA	CGCGGGCGGG	TGTGGTGGTT	ACGCGCAGCG
61	TGACCGCTAC	ACTTGCCAGC	GCCCTAGCGC	CCGCTCCTTT	CGCTTTCTTC	CCITCCTTTC
121	TCGCCACGTT	CGCCGGCTTT	CCCCGTCAAG	CICTAAATCG	GGGGCTCCCT	TTAGGGTTCC
181	GATTTAGTGC	TTTACGGCAC	CTCGACCCCA	AAAAACTTGA	TTAGGGTGAT	GGTTCACGTA
241	GTTGGCCATC	GCCCTGATAG	ACGGTITTTT	GCCCTTTGAC	GTTGGAGTCC	ACGTICTTTA
301	ATAGTGGACT	CTTGTTCCAA	ACTGGAACAA	CACTCAACCC	TATCTCGGTG	TATCTTTTGG
361	ATTTATAAGG	GATTTTGCCG	ATTTGCGCCT	ATTGGTTAAA	AAATGAGCTG	ATTTAACAAA
421	AAATTTAACGC	GAATTTTAAAC	AAAAIATTTAA	CGCTTACAAT	TTCCATTCGC	CATTGAGGCT
481	GCGCAACTGT	TGGGAAGGGC	GATCGGTGCG	GGCCTCTTCG	CTATTACGCC	AGCTGGCGAA
541	AGGGGGATGI	GCTGCAAGGC	GATTAAGTTG	GGTAACGCCA	GGGTTTCCCG	AGTCACGACG
601	TTGTAAAACG	ACGGCCAGTG	AATTGTAATA	CGACTCACTA	TAGGGCGAAT	TGGGTACCCG
661	GCCCCCCTC	GAAGATCTAA	TATTTGCAATG	TCGCTATGTG	TTCTGGGAAA	TCACCAJAAA
721	CGTGAATATC	CTATCAGTGA	TAGAGACTTA	TAAGTTCCCT	ATCAGTGATA	GAGAATTCAT
781	TTGGAAGTAT	TCGCGTACG	TTTCGACGTA	CGCGGAATAC	TTGGAATTTT	TTTCTCGAGC
841	TTCTTCGGG	AAGCTCTCCA	TATTTTITGG	ATCCACTAGT	TCTAGAGCGG	CCGCCACCGC
901	GGTGGAGCTC	CAGCTTTTGT	TCCCTTTAGT	GAGGGTAAAT	TTGAGCTTGG	GGGTAATCAT
961	GGTCATAGCT	GTTTCCTGTG	TGAAATTGTT	ATCCGCTCAC	AATTCACAC	AACATACGAG
1021	CCGGAAGCAT	AAAGTGTTAA	GCCTGGGGTG	CCTAATGAGT	GAGCTAACTC	ACATTAATTT
1081	CGTTGCGCTC	ACTGCCCGCT	TCCAGTCCGG	GAAACCTGTC	GTGCCAGCTG	CATTAATGAA
1141	TCGGCCAACG	CGCGGGGAGA	GGCGGTTTGC	GIATTTGGCG	CTCTTCCGCT	TCCTCGGTCA
1201	CTGACTCGCT	GCCTCCGTC	GTTCCGCTGC	GGCGAGCGGT	ATCAGCTCAC	TCAAAGGCGG
1261	TAATACGGTT	ATCCACAGAA	TCAGGGGATA	ACGCARGAAA	GAACATGTGA	GCAAAAGGCC
1321	AGCAAAAAGGC	CAGGAACCGT	AAAAAGGCCG	CGTTGCTGGC	GTTTTCCAT	AGGCTCCGCC
1381	CCCCTGACGA	GCATCACAAA	AATCGACGCT	CAAGTCAGAG	GTGGCGAAAC	CCGACACGGC
1441	TATAAAGATA	CCAGGCGTTT	CCCCCTGGAA	GCTCCCTCGT	GCGCTCTCCT	GTTCCGACCC
1501	TGCCGCTTAC	CGGATACCTG	TCCGCTTTC	TCCCTTCGGG	AAGCGTGGCG	CTTCTCATA
1561	GCTCAGCTG	TAGGTATCTC	AGTTCGGTGT	AGGTTCGTTG	CTCCAAGCTG	GGCTGTGTGC
1621	ACGAACCCCC	CGTTCAGCCC	GACCGCTGCG	CCTTATCCGG	TAACATCTGT	CTTGAGTCCA
1681	ACCCGGTAAAG	ACACGACTTA	TCGCCACTGG	CAGCAGCCAC	TGGTAACAGG	ATTAGCAGAG
1741	CGAGGTATGT	AGGCGGTGCT	ACAGAGTTCT	TGAAGTGGTG	GCCTAACTAC	GGCTACACTA
1801	GAAGGACAGT	ATTTGGTATC	TGCGCTCTGC	TGAAGCCAGT	TACCTTCGGA	AAAAGAGTTG
1861	GTAGCTCTTG	ATCCGGCAAA	CAAACCACCG	CTGGTAGCGG	TGGTTTTTTT	GTTTGCAGC
1921	AGCAGATTAC	GCGCAGAAAA	AAAGGATCTC	AAGAAGATCC	TTTGATCTTT	TTTACGGGGT
1981	CTGACGCTCA	GTGGAACGAA	AATCAGCTTT	AAGGGATTTT	GGTCATGAGA	TTATCAAAAA
2041	GGATCTTAC	CTAGATCCTT	TIAAATTTAA	AATGAAGTTT	TAAATCAATC	TAAAGTATAT
2101	ATGAGTAAAC	TTGGTCTGAC	AGTTACCAAT	GCTTAATCAG	TGAGGCACCT	ATCTCAGCGA
2161	TCTGTCTATT	TCGTTTATCC	AATGTTGCC	GAATCCCGT	CGTGTAGATA	ACTACGATAC
2221	GGGAGGGCTT	ACCATCTGGC	CCCAGTGTG	CAATGATACC	GCGAGACCCA	CGCTCACCGG
2281	CTCCAGATT	ATCAGCAATA	AACCAGCCAG	CCGGAAGGGC	CGAGCGCAGA	AGTGGTCTTG
2341	CAACTTTATC	CGCCTCCATC	CAGTCTATTA	ATTGTTGCCG	GGAAGCTAGA	GTAAGTATGT
2401	CGCCAGTTAA	TAGTTTGGCG	AACGTTGTTG	CCATTGCTAC	AGGCATCGTG	GTGTCAGCTT

Fig 6

```
2461 CGTCGTTTGG TATGGCTTCA TTCAGCTCCG GTTCCCAACG ATCAAGGCGA GTTACATGAT
2521 CCCCCATGTT GTGCAAAAAA GCGGTTAGCT CCTTCGGTCC TCCGATCGTT GTCAGAAGTA
2581 AGTTGGCCGC AGTGTTATCA CTCATGGTTA TGGCAGCACT GCATAATTCT CTTACTGTCA
2641 TGCCATCCGT AAGATGCTTI TCTGTGACTG GTGAGTACTC AACCAAGTCA TTCTGAGAAT
2701 AGTGTATGCG GCGACCGAGT TGCTCTTGCC CGGCGTCAAT ACGGGATAAT ACCGCGCCAC
2761 ATAGCAGAAC TTTAAAAGTG CTCATCATTG GAAAACGTTT TTCGGGGCGA AAACCTCTCAA
2821 GGATCTTACC GCTGTTGAGA TCCAGTICGA TGTAACCCAC TCGTGCACCC AACTGATCTT
2881 CAGCATCTTT TACTTTTACC AGCGTTTCTG GGTGAGCAAA AACAGGAAGG CAAAATGCCG
2941 CAAAAAAGGG AATAAGGGCG ACACGGAAAT GTTGAATACT CATACTCTTC CTTTTTCAAT
3001 ATTATTGAAG CATTATCAG GGTATTGTC TCAIGAGCGG ATACATATTT GAATGTATTT
3061 AGAAAAATAA ACAAATAGGG GTTCCGCGCA CATTCCCCG AAAAGTGC
```