

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 391 906**

51 Int. Cl.:  
**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **07291093 .8**  
96 Fecha de presentación: **12.09.2007**  
97 Número de publicación de la solicitud: **2036989**  
97 Fecha de publicación de la solicitud: **18.03.2009**

54 Título: **Polinucleótido apto para un ensayo indicador basado en una sola célula para controlar los patrones de expresión génica con una resolución espacio-temporal elevada**

45 Fecha de publicación de la mención BOPI:  
**30.11.2012**

45 Fecha de la publicación del folleto de la patente:  
**30.11.2012**

73 Titular/es:  
**INSTITUT PASTEUR (100.0%)  
25-28, RUE DU DOCTEUR ROUX  
75724 PARIS CEDEX 15, FR**

72 Inventor/es:  
**MHLANGA, MUSA;  
ENNINGA, JOST;  
SANSONETTI, PHILIPPE y  
NEHRBASS, ULF**

74 Agente/Representante:  
**CURELL AGUILÁ, Mireia**

ES 2 391 906 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

Polinucleótido apto para un ensayo indicador basado en una sola célula para controlar los patrones de expresión génica con una resolución espacio-temporal elevada.

5 La invención se refiere a un ensayo para controlar los patrones de expresión génica, especialmente en una célula viva, recurriendo a un sistema indicador de la transcripción. El ensayo de la invención proporciona unos medios adecuados para una rápida y alta resolución espacio-temporal de dichos patrones. El ensayo de la invención puede utilizarse además para el seguimiento cuantitativo de la transcripción génica.

10 La invención proporciona en particular la posibilidad de realizar el ensayo diseñado en células aisladas. Se proporciona también especialmente la posibilidad de realizar el ensayo diseñado en una célula viva.

15 Así pues, la invención proporciona un ensayo que permite controlar la producción de ARN, en su caso un ARNm específico, en una célula viva.

20 El campo de aplicaciones de la invención comprende el seguimiento de la expresión génica a fin de determinar o controlar la regulación de la homeostasis celular, la actividad celular o la desregulación de la homeostasis o actividad celular u otros procesos celulares.

25 La invención permite en particular controlar la expresión génica en las células que se han sometido a prueba de provocación por episodios extracelulares, tales como el estrés causado por agentes patógenos u organismos, fármacos o productos químicos. Las células para su utilización según la invención comprenden cualquier tipo de células procariotas o eucariotas. En un aspecto concreto de la invención, estas células se pueden obtener de un hospedador, especialmente un hospedador mamífero, en particular un paciente afectado con una afección patológica. También comprende la determinación de las interacciones hospedador/patógeno, en la célula viva.

La invención proporciona así unos medios adecuados para un alto rendimiento y un alto contenido de identificación.

30 Así pues proporciona herramientas para su utilización en estrategias de desarrollo de fármacos, y más en general permite controlar la expresión génica en las células vivas y permite la perfilación de genes con fines de identificación. Las aplicaciones específicas de la invención incluyen el seguimiento de las células de un paciente infectado con organismos o agentes patógenos, tales como bacterias, virus, parásitos. Las células ensayadas como tal puede estar infectadas o no. La invención también proporciona unos medios para el seguimiento de las reacciones inmunitarias en una célula o para el seguimiento del desarrollo o la creación de las células tumorales.

35 Las células vivas están en constante comunicación con su entorno lo que requiere la adaptación de su fisiología a las circunstancias específicas. Estos episodios de comunicación juegan un papel importante (i) para la interferencia entre células individuales de un organismo (por ejemplo durante el desarrollo), o (ii) para la respuesta eficaz a la tensión desde el exterior. El estrés puede ser físico (por ejemplo calor), químico (por ejemplo productos químicos tóxicos) o bioquímico (por ejemplo, bacterias patógenas).

40 En general, las células alteran la expresión de genes específicos para ajustarse durante las situaciones descritas anteriormente. En particular, la expresión génica está muy regulada temporalmente lo que permite una respuesta equilibrada celular al estrés. La alteración de este equilibrio conduce a la enfermedad.

45 En el caso del estrés causado por patógenos, estos factores estresantes han desarrollado estrategias para atacar específicamente y alterar la expresión génica celular del hospedador (Arbibe, L. *et al.* An injected bacterial effector targets chromatin access for transcription factor NF-kappaB to alter transcription of host genes involved in immune responses. *Nat. Immunol.* 8, 47-56 (2007)). Esto puede conducir a una interferencia con la respuesta inmunitaria del hospedador, y al escape del agente patógeno del seguimiento inmunitario del hospedador (por ejemplo *Mycobacterium tuberculosis*) (Monack D.M., Mueller A., Falkow S. Persistent bacterial infections: the interface of the pathogen and the host immune system. *Nat. Rev. Microbiol.* Sep. 2004; 2(9):747-65). Distorsiones similares ocurren durante la desregulación en las enfermedades autoinmunitarias (por ejemplo artritis o lupus) (Kyttaris V.C., *et al.* Immune cells and cytokines in systemic lupus erythematosus: an update. *Curr. Opin. Rheumatol.* Sep. 2005; 17(5):518-22).

50 Por lo tanto, un conocimiento preciso de las respuestas de expresión génica puesta a sitio es fundamental para entender la base molecular de las reacciones celulares al estrés, o las especificidades de las etapas celulares reguladas durante el desarrollo.

55 La época posgenómica ha proporcionado una gran cantidad de información de varios sistemas genómicos, lo que permite el análisis de la complejidad de los procesos biológicos a gran escala, y con un alto rendimiento (Barabási, A.L. y Oltvai, Z.N. Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101-113 (2004)). Los ejemplos de estos son el análisis de proteínas por espectrometría de masas y transcripción y perfilación

de la expresión por chips de proteínas y de ADN (Pepperkok, R. y Ellenberg, J. High-throughput fluorescence microscopy for systems biology. *Nat. Rev. Mol. Cell Biol.* 7, 690-696 (2006)).

La comprensión de las interacciones entre los componentes de un sistema biológico y la forma en la que dan lugar a la función es un objetivo clave cuando se estudia la biología de sistemas. La mayor parte de nuestra información actual sobre la activación de genes corriente abajo en muchas cascadas de transducción de señales procede de los datos de chips o de ensayos con gen indicador de proteínas (Pepperkok, R. y Ellenberg, J. High-throughput fluorescence microscopy for systems biology. *Nat. Rev. Mol. Cell Biol.* 7, 690-696 (2006)). Los métodos con chip pueden proporcionar información de la población o "censo" del comportamiento de millones de células. Sin embargo, cada célula está involucrada de la manera más probable en una fase diferente de la respuesta a la cascada de señalización y lo que se mide es una imagen más global y general. Dentro de este panorama mosaico se encuentra la información sobre cuándo las células específicas se dedican en fases específicas de su respuesta génica. Teóricamente ésta puede estar relacionada temporalmente con cuando se activa la vía y ayuda en la montaje de modelos mecánicos de las modalidades de cómo funcionan dichas cascadas de transducción de señal dinámica. Sin embargo la información temporal precisa de dicha respuesta de la transcripción está enmascarada en el "ruido" o las variaciones estocásticas de los datos del chip.

La exploración de macromoléculas en su entorno natural con alta resolución espacio-temporal ha sido posible mediante la utilización de análisis de detección por la imagen por fluorescencia en las células vivas (Pepperkok, R. y Ellenberg, J. High-throughput fluorescence microscopy for systems biology. *Nat. Rev. Mol. Cell Biol.* 7, 690-696 (2006), Bastiaens, P.I. y Pepperkok, R. Observing proteins in their natural habitat: the living cell. *Trends Biochem. Sci.* 25, 631-637 (2000); Meyer, T. y Teruel, M.N. Fluorescence imaging of signaling networks. *Trends Cell Biol.* 13, 101-106 (2003); Wouters, F.S., Verveer, P.J. y Bastiaens, P.I. Imaging biochemistry inside cells. *Trends Cell Biol.* 11, 203-211 (2001). En principio, se pueden utilizar para explorar las proteínas en su hábitat natural, interrogando a sus interacciones bioquímicas. Sin embargo, esto no se ha extendido fácilmente a la dinámica de detección por la imagen de la expresión génica, por ejemplo mediante la observación de la transcripción del ARN mensajero. El examen de esta actividad en la célula aislada permitiría la relación temporal entre la activación de una cascada de transducción de la señal (los episodios bioquímicos) y una respuesta de transcripción específica que debe relacionarse con precisión.

Aunque esto se ha intentado con ensayos con genes informadores tales como ensayos con proteína fluorescente o luciferasa, lo que se mide es la lectura de la traducción y no de la transcripción de un solo gen en lo que es lo más probable cientos de genes que participan en una respuesta de la transcripción. La cinética de la expresión génica sin embargo no puede deducirse directamente por la aparición de una proteína ya que existe un retraso entre la transcripción de ARNm y la traducción a proteína. Puesto que la traducción se retrasa respecto a la transcripción, la capacidad de extraer relaciones temporales exactas entre los genes puede seguir siendo infructuosa. Como resultado a pesar de las copiosas cantidades de datos, la relación temporal entre la activación de las vías funcionales permanece poco clara. Estas deficiencias se han observado especialmente al estudiar las relaciones temporales entre las vías inmunitarias y las vías inflamatorias (Karin, M. Nuclear factor-kappaB in cancer development and progression. *Nature* 441, 431-436 (2006); Karin, M., Lawrence, T. y Nizet, V. Innate immunity gone awry: linking microbial infections to chronic inflammation and cancer. *Cell* 124, 823-835 (2006)). Hasta la fecha, no existe ninguna etiqueta codificada genéticamente que podría utilizarse para marcar ARNm de una manera similar que GFP y proteínas comparables.

El alto contenido de identificación permite la evaluación de múltiples parámetros bioquímicos y morfológicos en los sistemas celulares, si las lecturas biológicas en el sistema son adecuadas para la recopilación de datos cuantitativos *in vivo*. Al combinar la detección por la imagen de células aisladas con algoritmos de análisis de detección por la imagen, se asignan propiedades cuantitativas a cada uno de los componentes del sistema biológico (Genovesio, A., Belhassine, Z. y Olivo-Marin, J. Adaptive gating in Gaussian Bayesian multi-target tracking. *Image Processing*, 2004. ICIP'04. 2004 International Conference on 1 (2004), Bork, P. y Serrano, L. Towards cellular systems in 4D. *Cell* 121, 507-509 (2005); Genovesio, A., Zhang, B. y Olivo-Marin, J. Interacting multiple model based method to track moving fluorescent biological spots. *Biomedical Imaging: Macro to Nano*, 2004. IEEE International Symposium on, 1239-1242 (2004); Olivo-Marin, J. Extraction of spots in biological images using multiscale products. *Pattern Recognition* 35, 1989-1996 (2002)). Así pues, la naturaleza del sistema dinámico puede modelarse, permitiendo auténtica biología de sistemas.

En la actualidad se puede recopilar información cuantitativa de las interacciones de proteínas en las vías de transducción de señales en las células vivas utilizando métodos basados en detección por la imagen (Starkuviene, V. High-content screening microscopy identifies novel proteins with a putative role in secretory membrane traffic. *Genome Res.* 14, 1948-1956 (2004); Liebel, U. A microscope-based screening platform for large-scale functional protein analysis in intact cells. *FEBS Lett.* 554, 394-398 (2003); Bastiaens, P.I. y Pepperkok, R. Observing proteins in their natural habitat: the living cell. *Trends Biochem. Sci.* 25, 631 -637 (2000)). Sin embargo, la información de la transcripción unicelular *in vivo* en el ARNm que informa de los niveles cuantitativos de actividad de la transcripción en una cascada de transducción de señal, en respuesta a un estímulo dado o en diferentes antecedentes genéticos, no es aún una realidad. La consecución de este objetivo permitirá el modelado preciso de la transcripción en las cascadas de transducción de señales.

La capacidad para seguir y rastrear cada uno de los complejos de mRNA *in vivo* (es decir, en una célula viva) ha experimentado avances importantes. La mayoría de los enfoques sufre sin embargo de una incapacidad para observar conjuntamente las proteínas implicadas en episodios de la cascada de transducción de señales a la vez que el ARNm al que se está haciendo el seguimiento en espacio y en tiempo real. Un avance adicional permitiría estudios resueltos en espacio y tiempo para comprender la relación orquestada entre proteínas de transducción de señales implicadas en la expresión génica, y permitiría la determinación precisa de los sitios de tiempo en los que empieza la expresión génica, la cuantificación de la expresión y el desarrollo de modelos cuantitativos de expresión génica.

El transporte de cada uno de los transcritos recientemente ha sido posible utilizando el sistema informador MS2. Para este enfoque, un plásmido codifica una proteína verde fluorescente (GFP) fusionada a la secuencia de codificación para la proteína de la cápside del fago, MS2, con ARN de fago monocatenario. La proteína de cápside MS2 tiene una gran afinidad por una secuencia específica de ARN con una estructura de tallo-lazo corta (secuencia de unión a MS2) codificada por el fago. El segundo plásmido contiene la secuencia de unión a MS2 polimerizada en 6, 12 ó 24 copias. Ambos plásmidos se cotransfectaron en células o alternativamente estirpes celulares para la expresión inducible de GFP-MS2. Aun cuando, pueden rastrearse transcritos aislados por este método, el control de la velocidad de transcripción es problemático porque la proteína GFP-MS2 fluoresce también en ausencia de transcritos. Por lo tanto, no es adecuado para estudios cuantitativos de perfilación de la transcripción (Bertrand, E. *et al.* Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2, 437-445 (1998); Shav-Tal, Y. *et al.* Dynamics of single mRNPs in nuclei of living cells. *Science* 304, 1797-1800 (2004)).

Las balizas moleculares, las sondas de nucleótidos que fluorescen sólo al hibridar específicamente a secuencias de ARNm complementario, presentan una solución general al problema de prever la expresión génica (Tyagi, S. y Kramer, FR. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* 14, 303-308 (1996)) Conjuntamente con las proteínas fluorescentes y a diferencia de las técnicas de microscopía óptica y análisis de detección por la imagen éstas ofrecen la posibilidad de utilizarse en biología cuantitativa. Sin embargo, todavía se necesitan mejoras para permitir la visualización y el análisis correspondientes de la señal suministrada por estas sondas, sobre todo cuando se contempla la determinación cuantitativa de sondas hibridadas.

Existen varias limitaciones para llevar a cabo ensayos similares de hibridación en células vivas, la principal de ellas, la capacidad de la sonda para lograr emparejarse con las secuencias de ARNm que se encuentran en estructuras secundarias complejas *in vivo*. Además puesto que el ácido nucleico diana no está inmovilizado o fijado antes de la introducción de la sonda, se producen constantemente interacciones dinámicas de ARNm con proteínas y ribosomas, lo que significa que muchas regiones del polímero ARNm están involucradas en interacciones con otros actores celulares.

Previamente, los transcritos de ARNm individuales se rastrearon en el interior de células vivas utilizando un montaje que contenía 96 sitios de unión de baliza que se fusionaron a un elemento de respuesta de tetraciclina modificada juntamente con un promotor mínimo de CMV (D.Y. Vargas *et al.*, Mechanism of mRNA transport in the nucleus, *PNAS* vol. 102, nº 47, 17008-17013 (2005)). El mecanismo de transcripción no reconoce dicho montaje en el interior de las células eucariotas para producir el indicador. Más bien, este sistema indicador requiere la presencia de un factor modificado genéticamente, el transactivador controlado por tetraciclina. Este transactivador consiste en una versión modificada de la molécula represora Tet bacteriana y repeticiones de un dominio de transactivación VP16 mínimo. Este sistema indicador se utiliza para garantizar la transcripción rigurosamente regulada en función de la cantidad de tetraciclina añadida a las células. En conjunto, este sistema indicador no permite el estudio de la expresión génica de genes endógenos en células eucariotas procedentes de su promotor natural, y requiere la presencia de (i) factores modificados genéticamente para activar la transcripción, y de (ii) tetraciclina como inductor de la transcripción. Del Poeta *et al.* (*Infection and Immunity*, abril 1999, págs. 1812-1820) da a conocer un polinucleótido bicatenario que comprende uno de los diversos promotores endógenos de levadura (actina, GAL7, Mfa1) seguido de un gen indicador GFP. Este montaje es adecuado para la perfilación de la expresión génica en diferentes condiciones de crecimiento por detección por fluorescencia del GFP como indicador de la actividad de transcripción de dicho promotor.

La solicitud da a conocer un ensayo que supera los inconvenientes de ensayos conocidos como los descritos anteriormente basados en proteínas indicadoras, por ejemplo, proteínas fluorescentes indicadoras, que introducen un impredecible tiempo de retardo experimental causado por la maduración de la proteína. En este contexto, los inventores han desarrollado un procedimiento y herramientas adecuadas y productos para investigar los cambios de la transcripción en tiempo real en la célula aislada que está basado en indicadores a base de ácidos nucleicos muy sensibles.

La invención proporciona medios adecuados para mejorar el estudio de la biología de sistemas, mediante el desarrollo de una tecnología de detección por la imagen que representa un gran avance en la detección por la imagen de la expresión génica *in vivo*, es decir, en células vivas, especialmente adecuado para la perfilación de genes por separado, es decir, células aisladas.

Los productos de la invención ofrecen una forma para el marcado sistemático de los ARNm en las células vivas, especialmente en la práctica rutinaria.

5 Las etiquetas utilizadas para llevar a cabo la invención son etiquetas a base de ácidos nucleicos, definidas empíricamente y no obstante adecuados para su utilización sistemática para la perfilación de muchos genes.

Los productos de la invención están diseñados para permitir no sólo el avance de la biología de sistemas, sino lo más importante, la investigación básica y clínica.

10 Todos los procesos celulares en células vivas están regulado a través de la expresión fuertemente orquestada de subconjuntos específicos de genes. La desregulación de la expresión génica conduce a una alteración de la homeostasis de la expresión génica y por último a enfermedad. Esto puede controlarse recurriendo a la invención.

15 Es importante destacar que los patógenos o los fármacos a menudo interfieren con las vías de señalización alterando potencialmente el patrón de expresión génica celular. La invención permite el seguimiento de dichas interferencias en la expresión génica celular en la transcripción.

20 Los programas de expresión génica cambian rápidamente en las células vivas provocando enormes desafíos técnicos para su estudio. Una vez realizada, la invención permite el estudio rutinario de la homeostasis celular y, según un aspecto, la identificación de las respuestas de la transcripción asociadas a la infección y la enfermedad, el desarrollo y la respuesta inmunitaria. Especialmente puede ayudar a la identificación de rutas y redes reguladoras de genes que se expresan en las enfermedades, a la respuesta inmunitaria y durante el crecimiento y el desarrollo.

25 Con el fin de alcanzar este objetivo, la invención proporciona polinucleótidos bicatenarios adecuados para llevar a cabo el ensayo de expresión de la perfilación génica cuando se integra en la célula que alberga de forma natural y expresa posiblemente el gen o genes de interés para dicha perfilación, que comprende en su cadena positiva considerada desde su extremo 5' a su extremo 3', (i) un promotor de un gen de interés o varios promotores de varios genes de interés seleccionados de entre varios genes que son (a) endógenos para la célula y están (b) sometidos a perfilación de la expresión génica, en el que dicho promotor es reconocido por el mecanismo de transcripción interna de la célula y, (ii) uno o varios código(s) de barras en el que cada código de barras contiene por lo menos una unidad de código de barras que es un montaje de ADN que comprende repeticiones en tándem de un sitio de unión de reconocimiento de baliza o repeticiones en tándem de por lo menos dos sitios de unión diferentes de reconocimiento de baliza, estando compuesto cada sitio de unión por una secuencia de nucleótidos, que se hibrida específicamente con la parte de lazo de una baliza molecular en la que la parte de lazo tiene de 10 a 55 nucleótidos y en el que cada uno de dicho(s) código(s) de barras está(n) bajo el control de por lo menos uno de dicho(s) promotor(es) para la transcripción.

40 El/los sitio(s) de unión de reconocimiento o la secuencia de nucleótidos complementaria puede reconocerse y unirse por una sola sonda de detección, es decir, una sonda molecular (sonda nucleica).

En una forma de realización específica, el número de sitios de unión de reconocimiento pueden ser hasta 500. Especialmente pueden ser más de 100.

45 Por consiguiente, cuando el código de barras comprende 2 o más unidades de unión de reconocimiento, cada unidad de unión es reconocida por una sola sonda de detección.

50 En una forma de realización específica, la unidad de código de barras o por lo menos una de las diversas unidades de código de barras del código de barras varias está constituida por un gran número de sitios de unión de reconocimiento que son todos idénticos entre sí.

55 Según la invención, los transcritos del ácido nucleico que constituye el código de barras expresado bajo el control del promotor o promotores, se detectan con una sonda de detección, especialmente una sonda de detección basada en ácido nucleico (sonda molecular). Dicha sonda de detección reconoce el transcrito de los sitios de reconocimiento del código o códigos de barras.

60 Lo que se describe en la presente memoria con referencia a un promotor, un sitio de unión de reconocimiento, una unidad de código de barras, un código de barras, un gen de interés en el polinucleótido, de manera similar puede aplicarse a varias o a todas estas entidades cuando se asocia en un polinucleótido o cuando está presente en varios polinucleótidos.

Las características del polinucleótido definidas en la presente memoria están posiblemente presentes en el polinucleótido independientemente uno de otro o en cualquier combinación posible entre sí.

65 Se señala especialmente que, excepto cuando se excluye expresamente, cualquiera de estas características, incluyendo las características descritas en los ejemplos, se pueden combinar para proporcionar una forma de realización concreta del polinucleótido definido.

El promotor contenido en el polinucleótido es uno de un gen de interés seleccionado entre los genes que son endógenas para una célula determinada en el que se estudia la perfilación de genes según la invención.

5 Un polinucleótido determinado se diseña de tal manera que en un código de barras, por lo menos 2 de los sitios de unión de reconocimiento están contiguos y/o por lo menos 2 de las unidades de código de barras están contiguos.

10 En otro polinucleótido determinado, en una unidad de código de barras, por lo menos 2 de los sitios de unión de reconocimiento están separados por un separador y/o por lo menos 2 de las unidades de código de barras están separadas por un separador.

15 Estas definiciones de montajes de polinucleótidos específicos abarcan cualquier posibilidad de combinación de sitios de unión de reconocimiento contiguos y/o separados o unidades de códigos de barras o códigos de barras, siempre que el montaje obtenido permita que se detecte la señal de detección y especialmente se conecte cuando los códigos de barras en especial su transcrito, se hibrida con una sonda de detección de ácido nucleico, en condiciones apropiadas.

20 Cuando se utiliza para el ensayo de perfilación génica en una célula, mediante el estudio de la transcripción de ARNm dirigida por el promotor del polinucleótido de la invención, la reacción de hibridación con la sonda de detección se produce con el transcrito de ARN del polinucleótido, o parte del mismo.

25 Para la determinación de la capacidad de la sonda de detección que debe utilizarse en el ensayo de la invención, la capacidad de hibridación puede probarse con el polinucleótido de la invención, proporcionado como molécula monocatenaria. Por conveniencia, si se hace referencia en la presente solicitud a la hibridación con el polinucleótido debe sobreentenderse como que abarca la hibridación con una cadena del polinucleótido, o en el contexto de la realización del ensayo de expresión de perfilación génica de la invención, con su transcrito.

30 El polinucleótido bicatenario es una secuencia de ADN que, en alguna etapa en las aplicaciones de la invención, especialmente para conseguir la hibridación de ensayo, puede estar presente, transcribirse o utilizarse como un ADN monocatenario o polinucleótido de ARN. En la medida en que es técnicamente relevante, las características descritas con respecto al polinucleótido bicatenario se aplican a su ADN monocatenario o a la versión de ARN.

El polinucleótido bicatenario de interés es adecuado y se utiliza en una célula que alberga de forma natural y posiblemente expresa el gen de interés cuyo promotor se utiliza en la montaje polinucleotídico.

35 En una forma de realización concreta de la invención, el polinucleótido bicatenario está diseñado de tal manera que su extremo 5' (en su cadena positiva) comienza esencialmente con una secuencia de promotor de un gen de interés y su extremo 3' termina esencialmente con la secuencia de un sitio de unión de reconocimiento. En consecuencia, dicho polinucleótido específico de la invención está enmarcado por el promotor descrito en la presente memoria y el sitio de unión de reconocimiento y, o bien no contiene más secuencias con 5' y 3' o sólo contiene más nucleótidos 5' y/o 3' que representan la(s) enzima(s) de restricción endonucleasa(s) o una parte de la misma, útil para una etapa de clonación. Entre dicho promotor de partida y el sitio de unión de reconocimiento de terminación, el polinucleótido puede comprender otro(s) sitio(s) de unión de reconocimiento sucesivo(s) y la(s) secuencia(s) promotora(s) de manera que un código de barras hecho de sitios de unión de reconocimiento, posiblemente dispuesto en unidades de código de barras, está necesariamente bajo el control de expresión de un promotor situado corriente arriba. También es posible que el polinucleótido contenga más secuencias 3', tales como, por ejemplo, una cola de poli A y/o una señal de terminación de la transcripción en 3'. El polinucleótido puede comprender también secuencias del gen de interés que son secuencias más distales al promotor, tales como la de una región potenciadora y/o otras secuencias reguladoras de la expresión. La secuencia promotora es, según la definición convencional, una región en el gen que determina el sitio de partida de la transcripción y permite el reconocimiento y la unión del factor o factores de transcripción y, directa o indirectamente, el reconocimiento y la unión de la(s) polimerasa(s) de la célula, lo que permite la ampliación de ARN. El promotor comprende generalmente una secuencia TATA.

55 En una forma de realización concreta, el promotor es el de un gen eucariota.

60 Además, los polinucleótidos o los fragmentos de un polinucleótido de la invención, hechos de un promotor y un código de barras puede contener además secuencias de ADN codificadoras de interés, incluyendo las secuencias para la expresión de una proteína informadora, y/o una secuencia de codificación del gen que proporciona el promotor. Dichos fragmentos de polinucleótidos pueden ser adyacentes entre sí en el polinucleótido o pueden estar separados por otras secuencias de ácido nucleico.

65 Un sitio de unión de reconocimiento como se describe en la presente memoria comprende o consiste en una secuencia de nucleótidos, especialmente una secuencia de ADN, que, en especial cuyo transcrito, es adecuado para el reconocimiento por otro polinucleótido que constituye una sonda de nucleótidos, tal como una secuencia complementaria de nucleótidos capaz de emparejamiento de bases con dicho sitio de unión de reconocimiento, especialmente su transcrito. Cuando se utiliza el polinucleótido de la invención, el transcrito del sitio de unión de

reconocimiento se expresa y es reconocido por la sonda de una manera suficiente para permitir el emparejamiento de bases con la sonda para dar lugar a un cambio mensurable en el transcrito del polinucleótido que puede detectarse.

5 En la invención, el/los sitio(s) de unión de reconocimiento del código de barras es (son) de sitio(s) de unión de baliza, es decir, los sitios que son reconocidos y adecuados para la unión con sondas de baliza. Según esto, el código de barras es un código de barras de baliza reconocido por una o varias sonda(s) de baliza. Las sondas de baliza se describen con más detalles a continuación.

10 Según la invención, la expresión "un gen de interés" se refiere a una secuencia de nucleótidos o montaje, que incluye una unidad de transcripción, que puede transcribirse y traducirse en condiciones ambientales adecuadas especialmente en una célula, para codificar un polipéptido (sea cual sea su longitud). En una forma de realización preferida, un gen de interés es un gen de estructura celular, un gen accesorio o un gen regulador o cualquier otro gen apropiado para la homeostasis, la regulación o el metabolismo de la célula.

15 El gen de interés que proporciona el promotor del polinucleótido de la invención es endógeno en la célula en la que se introduce el polinucleótido para el estudio de transcripción de polinucleótidos y proporciona por lo tanto un promotor natural, es decir, el que alberga de forma natural la capacidad de funcionar como promotor en un gen de una célula. En consecuencia, el promotor se dice que es endógeno a dicha célula. Por esta expresión se entiende también que el promotor es reconocido por el mecanismo de transcripción interna de la célula (especialmente de la célula eucariota), incluyendo el/los factor(es) de transcripción y la(s) polimerasa(s) de la célula, sin necesidad de llevar factores de transcripción externa u otros componentes reguladores de la expresión. En otras palabras, el promotor es un promotor natural de la célula en la que se estudia su actividad de transcripción o procede de dicho promotor como se describe en la presente memoria. Para el propósito de la invención dicho promotor natural puede alterarse y mutarse especialmente por adición, eliminación, inserción de uno o varios nucleótidos, para estudiar el impacto de dicha alteración cuando se utiliza un promotor de variante en el polinucleótido.

20 Ejemplos de genes de interés según la invención son especialmente los genes que participan en una estrategia terapéutica. En una forma de realización concreta, pueden ser un objetivo para un organismo patógeno o agente activado o inhibido, ya sea directa o indirectamente, o un objetivo para activado o inhibido por un fármaco, ya sea directa o indirectamente. Pueden ser, como tales, un fármaco o un candidato a fármaco.

30 Dentro de esta definición de un "gen de interés" y por lo tanto del promotor insertado en el polinucleótido bicatenario de la invención, los genes endógenos de una célula, ya sea una célula procariota o una célula eucariota, útil para llevar a cabo la invención, son los genes que están involucrados, directa o indirectamente, en una respuesta a estímulos extracelulares, especialmente al estrés extracelular, tales como la inducida por organismos o agentes patógenos o por estímulos físicos tales como calor, productos químicos tóxicos, compuestos bioquímicos o condiciones patológicas.

40 En una forma de realización concreta los genes de una célula determinada que están afectados por la alteración de la expresión génica de células o involucrados en la desregulación de la expresión génica pueden ser genes de interés.

45 Ejemplos de genes endógenos de células específicas de interés son los genes implicados en la homeostasis o en el desarrollo de las células concretas estudiadas o de todo el organismo que comprende estas células, incluyendo pero sin limitarse a los genes implicados en el sistema inmunitario, a los genes inducidos o silenciados en condiciones patológicas tales como en estados de tumor o cáncer.

50 Los ejemplos particulares de dichos genes son genes que codifican inmunoglobulinas o fragmentos de unión a antígenos de los mismos, especialmente genes que codifican fragmentos variables de las cadenas pesadas de anticuerpos que tienen capacidad de unión a antígenos.

55 Otros ejemplos incluyen genes que codifican receptores, ya sean receptores celulares o receptores solubles, incluyendo receptores que se encuentran en los linfocitos T.

Ejemplos concretos de genes del sistema inmunitario que pueden dirigirse a la perfilación génica abarcan TGF- $\beta$ , C-fos, ICAM, en particular ICAM-1, CCL20, genes de interferón gamma.

60 Otros genes de interés para la perfilación génica incluyen los genes de interleucina, p. ej., los genes IL10 e IL8.

Otros ejemplos de genes de interés son los genes que codifican antígenos asociados a tumores, concretamente los genes implicados en la leucemia enumerados a continuación:

65 A título de ilustración, se pueden estudiar los siguientes genes indicativos de la adopción del LSC Fate:

**Genes marcadores L-GMP de ratón ("Early"):** *Meis1, HoxA9, HoxA10, MYLK, HoxA5, Stau2*; y para los **Genes Marcadores MLL-AML Humanos:** *Meis1, HoxA9, HoxA10, HoxA5*. En otro ejemplo, los genes muy expresados en las leucemias ALL, MLL y AML como las siguientes pueden caracterizarse:

5                   **Genes            Marcadores            ALL:**        *MME            (CD10),        CD24,        DYRK3,        FOXO1A.*

10                   **Genes Marcadores MLL:** *FLT3, KIAA0428, NKG2D, ADAM-10, PROML-1, KIAAW25, LGALS-1, CCNA-1, DKFZp586O0120,            ITPA,            CDNAag36C04,            KIAA0920,            LMO-2.*

15                   **Genes Marcadores AML:** *GJB-1, BSG, ENSA, CTSD, DF, TFDP-2, DRAP-1, NF2, CDNA20C10, PDE3B, ANPEP, Chrm19clone, Chrm22q11clone, RTN2, CRYAA.*

La expresión "promotor" o "secuencia promotora" designa por lo tanto, según la invención, una secuencia de ADN presente en un montaje que es adecuada para proporcionar, en una célula en la que un gen que comprende dicho promotor es normalmente endógena y natural, la activación de la transcripción de una secuencia de ADN colocada bajo su control en el polinucleótido de la invención, como resultado de unión de reconocimiento del factor o factores de transcripción celular y la(s) polimerasa(s) de dicha célula. El promotor del gen es, naturalmente, capaz de ser reconocido y activado por el mecanismo de la célula en la que se ensaya la transcripción o la expresión después de la transfección de dicha célula con la montaje del polinucleótido.

Una secuencia promotora apropiada para llevar a cabo la invención, puede ser un promotor natural de un gen natural de interés. Alternativamente, puede proceder de un promotor endógeno como se describe en la presente memoria y en consecuencia es un promotor modificado, que tiene especialmente una secuencia mutada, por ejemplo, por eliminación, inserción y/o sustitución de uno o varios, de sus nucleótidos, y cuya actividad de transcripción debe examinarse. Por consiguiente, para la invención, el acceso al promotor endógeno del gen de interés o la información proporcionada por su secuencia de nucleótidos, se requiere permitir su identificación en vista de su preparación para probar la actividad del promotor o para probar la actividad de una contrapartida modificada de dicho promotor, mientras que puede no ser necesario el acceso a toda la secuencia del gen o incluso a su secuencia de codificación. Más en general, la invención proporciona medios para ensayar la supuesta actividad de la transcripción de cualquier fragmento de nucleótidos en una célula que puede ser un promotor experimental. Las secuencias que replican a estas diversas definiciones del promotor pueden prepararse por síntesis (síntesis química o ampliación, p. ej., por RCP) o pueden clonarse según cualquier técnica disponible.

El promotor endógeno es ventajosamente un promotor de un gen eucariota. La célula en la que se estudia la transcripción o expresión es por lo tanto ventajosamente una célula eucariota.

Un promotor tal como un promotor sensible a la doxiciclina como se describe en Vargas D.Y. *et al.*, *PNAS*, 2005 no es ni un promotor natural según la invención, ni un promotor variante tal como se define en la presente memoria.

El "promotor" como se define en la presente memoria consta del fragmento de ADN que tiene la actividad descrita para desencadenar la activación de la transcripción o comprende dicho fragmento de ADN. Preferentemente está desprovisto de toda la secuencia de codificación (es decir, la secuencia que se transcribe como un ARNm prematuro o como un ARNm) del gen de interés, o que sólo comprende menos del 20% de la secuencia de codificación del gen de interés. La secuencia del promotor no es especialmente una secuencia de ADN que se transcribe como un pre-ARNm en una célula que expresa el gen.

Según una forma de realización específica de la invención, la actividad del promotor se prueba sin provocación o interacción de cualquier molécula heteróloga o la molécula pone en contacto con la célula en la que se estudia la transcripción. Por "heterólogo" se entiende una molécula que no se expresa de forma natural en dicha célula. Una molécula que se pone en contacto con la célula se añade para el ensayo.

Según otra forma de realización, la actividad de transcripción del promotor del montaje del polinucleótido bicatenario se prueba bajo provocación o influencia de una molécula que es heteróloga a la célula o se estudia una molécula que se pone en contacto con la célula y cuya capacidad de interacción, ya sea directa o indirecta, con la actividad de transcripción del promotor del gen de interés en el polinucleótido de la invención. Dichas moléculas se proporcionan a título ilustrativo en los ejemplos y especialmente comprenden citocinas, interleucinas o quimiocinas, tales como interleucinas TNF $\alpha$ , IL-1.

En una forma de realización específica de la invención, el promotor de un gen de interés se selecciona de entre el grupo de promotores de genes endógenos tales como los genes de quimioquinas, especialmente promotores del gen de interleucina 8, los promotores de genes de interferón, especialmente el promotor del gen de interferón gamma, o un promotor de un gen de una proteína asociada a tumores.

En una forma de realización concreta de la invención, el promotor de un gen de interés se provoca mediante una molécula presente o se expresa en el interior de las células ensayadas mediante un mecanismo autocrino.

5 Según otra forma de realización, la actividad de transcripción del promotor de la montaje de polinucleótido bicatenario se provoca o está influenciado por un agente, especialmente un agente patógeno que se estudia para su posible, directa o indirecta, la interacción con dicha actividad promotora.

10 Según la descripción, un "código de barras" es un montaje de ADN que contiene por lo menos un sitio de unión de reconocimiento, es decir, un sitio adecuado para ser reconocido y para la hibridación, especialmente por su transcrito, a una sonda molecular, es decir, a una sonda de ácido nucleico, p. ej., una sonda de ADN, incluyendo por ejemplo una sonda del tipo de baliza tal como se explica en la presente memoria. Preferentemente, un código de barras de la invención comprende más de 2 o más de 3 sitios de unión de reconocimiento y/o más preferentemente más de 2 unidades de código de barras o de más de 3 unidades de código de barras, en el que cada unidad comprende o consta de múltiples repeticiones de una o varias (por ejemplo 3 o más) sitios de unión de reconocimiento.

15 En una forma de realización concreta de la invención, la sonda es una baliza y el código de barras es un código de barras de baliza con el/los sitio(s) de unión de la baliza.

20 En una forma de realización cuando varios sitios de unión de reconocimiento o varias unidades de código de barras están presentes en un código de barras, por lo menos algunos de dichos sitios de unión o unidades de código de barras están separadas por un espaciador.

25 En consecuencia, los sitios de unión de reconocimiento y/o las unidades de código de barras pueden estar separados por nucleótidos o secuencias de nucleótidos que intervienen de 1 a 250, especialmente 1, 2, 3, 4, 5, 6, 7, 8 nucleótidos.

30 El código de barras y/o las unidades de códigos de barras de la invención puede estar precedidos o seguidos o enmarcados por una o más sitios de restricción que están situados en su extremos 5' y/o 3'. Aparte de estos tipos diferentes de secuencias, en una forma de realización concreta, el código de barras puede estar desprovisto de otras secuencias funcionales implicadas en la transcripción incluyendo secuencias que provocan o están sometidos a la transcripción.

35 En una forma de realización específica de la invención, el polinucleótido comprende varios códigos de barras que están separados uno de otro por lo menos por un montaje promotor funcional. En consecuencia, cada código de barras se coloca bajo el control de un promotor de la transcripción específico como se define en la presente memoria.

40 En la invención, cada código de barras y/o cada unidad de código de barras comprende repeticiones en tándem de por lo menos una secuencia adecuada para permitir la hibridación de su transcripción con la secuencia de la sonda cuando entra en contacto con dicha sonda en condiciones apropiadas.

45 En una forma de realización específica, cada código de barras y/o cada unidad de código de barras comprende repeticiones de dos sitios de unión de reconocimiento diferentes o repeticiones de más de dos sitios de unión de reconocimiento, especialmente 3 o más, cada sitio de unión de reconocimiento son adecuados cuando se transcribe, para hibridarse con una secuencia de sonda específica.

50 En una forma de realización específica de la invención, el código de barras y/o la unidad de código de barras comprende repeticiones en tándem de por lo menos dos, especialmente 3, 4, 5 o más secuencias diferentes que, cuando se transcribe, son adecuados para la hibridación con secuencias de la sonda diferentes cuando se pone en contacto con dichas sondas en condiciones apropiadas.

55 En otra forma de realización de la invención, las secuencias de los sitios de unión de reconocimiento de un código de barras o de una unidad de código de barras son idénticas pero son apropiadas, cuando se transcriben, para hibridarse con diferentes sondas específicas en las que la diferencia entre los resultados de varias sondas en un cambio mensurable diferente del polinucleótido reconocido e hibridado que da lugar a una molécula de señal diferente, tales como diferentes restos fluorescentes o compuestos emisores de luz.

60 La identificación de fragmentos de nucleótidos apropiados adecuados para constituir sitios de unión de reconocimiento depende del contexto con respecto al ensayo que debe realizarse según la invención. En otra forma de realización dichos fragmentos de nucleótidos pueden estandarizarse.

65 En una forma de realización específica, las secuencias utilizadas para preparar los sitios de unión de reconocimiento no se encuentran en el transcriptoma de la célula en la que se introduce el polinucleótido, y especialmente no se encuentran en los transcriptomas de mamíferos, especialmente en el transcriptoma humano representado en las bases de datos disponibles en el momento de la detección. Especialmente, las secuencias utilizadas para preparar

los sitios de unión de reconocimiento proceden de organismos distintos de mamíferos. Un ejemplo específico de dicho organismo es *Drosophila* que proporciona una fuente para la identificación de las secuencias apropiadas.

Una selección preliminar de las secuencias puede hacerse recurriendo a algoritmos disponibles, tal como "mfold" que proporciona datos sobre secuencias que se considera que albergan los sitios de unión accesibles cuando están comprendidos en secuencias de ARN. La estructura secundaria del ARNm se determina por consiguiente identificando potenciales sitios de hibridación que son accesibles a la unión por sondas de hibridación de ácidos nucleicos. Existe una gran serie de pruebas que sugieren que la formación de heterodúplex está principalmente limitado por la estructura secundaria local y plegamiento de ARN. La composición y longitud de las sondas de ácido nucleico parece ser de menor importancia. Se prueban *in vitro* a continuación sondas experimentales identificadas por el algoritmo por su capacidad para unirse a un ARNm transcrito *in vitro*. Los que pueden formar mejor heterodúplex se conservan y utilizan como sondas de ácido nucleico.

Otra solución para abordar esta cuestión de diseño de sondas se ha centrado en la selección de sondas de ácidos nucleicos por métodos empíricos, utilizando, por ejemplo, matrices de exploración de oligonucleótidos. La utilización de dichas matrices ha reforzado la solidez de métodos empíricos sobre los algoritmos de predicción.

Las comparaciones entre los métodos empíricos para determinar las regiones accesibles a la hibridación, frente a los que determinados teóricamente por el algoritmo de predicción, han resaltado constantemente la fuerza superior de métodos empíricos. De hecho, los oligonucleótidos más apropiados (seleccionados para la hibridación a regiones monocatenarias) no se corresponden con las posiciones previstas como bucles por el algoritmo. Además está lejos de aclararse que las regiones monocatenarias previstas por los algoritmos de predicción sean realmente más accesible a la formación de heterodúplex, que las horquillas.

La siguiente descripción para la preparación de la baliza molecular es una adaptación de Bratu D. (Molecular beacons: Fluorescent probes for detection of endogenous mRNA in living cells. *Methods Mol. Biol.* 319, 1-14 (2006)). Teóricamente, puede seleccionarse cualquier secuencia en un ARN diana como un sitio para la unión de balizas moleculares. Las infinitas posibilidades dan confianza de que dichas regiones se identifiquen fácilmente. Sin embargo, el grado de accesibilidad al objetivo es principalmente una consecuencia de complejas estructuras intramoleculares secundarias y terciarias, que son difíciles de prever y que puede enmascarar muchas de estas regiones. Además, dentro de la célula, los ARNm existen junto con proteínas que ocluyen más partes del ARNm. Aunque las regiones implicadas en la unión de proteínas sólo se pueden identificar por análisis experimental, pueden hacerse intentos razonables para predecir las regiones que no están implicadas en estructuras secundarias herméticas. Hasta ahora, se dispone de varios ensayos *in vitro* y algoritmos teóricos para ayudar a identificar supuestos sitios diana en las secuencias de ARNm, así como las sondas con alta afinidad para la unión (Tyagi, S. & Kramer, F.R. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* 14, 303-308 (1996); Mir, K.U. & Southern, E. M. Determining the influence of structure on hybridization using oligonucleotide arrays. *Nat. Biotechnol.* 17, 788-792 (1999); Matsuo, T. (1998). In situ visualization of messenger RNA for basic fibroblast growth factor in living cells. *Biochim. Biophys. Acta* 1379, 178-84; Sokol, D. L., Zhang, X., Lu, P. & Gewirtz, A. M. (1998). Real time detection of DNA:RNA hybridization in living cells. *Proc. Natl. Acad. Sci. U S A* 95, 11538-43). El algoritmo mfold de plegamiento de ARN se utiliza para predecir la estructura secundaria más termodinámicamente estable junto con un conjunto de estructuras subóptimas (Southern, EM, Milner, N. & Mir, K.U. (1997). Discovering antisense reagents by hybridization of RNA to oligonucleotide arrays. *Ciba Found Symp* 209, 38-44; exposición 44-6; Ho, S. P., Bao, Y., Leshner, T., Malhotra, R., Ma, L. Y., Fluharty, S. J. & Sakai, R. R. (1998). Mapping of RNA accessible sites for antisense experiments with oligonucleotide libraries. *Nat. Biotechnol.* 16, 59-63). Puesto que ninguna de estas estructuras puede considerarse que representa la configuración de origen natural, se analizan los parámetros que describen el conjunto completo. El número de sitios experimentales se reduce empleando de un segundo algoritmo. OligoWalk explora la secuencia de ARN plegado para las regiones a las que pueden unirse oligonucleótidos de varias longitudes. Considerando la composición básica de cada oligonucleótido, y la estructura secundaria predicha del ARN, la producción proporciona información sobre la estabilidad del híbrido esperado, y por lo tanto identifica potenciales regiones diana. Una vez identificadas, las balizas moleculares específicas para esas regiones diseñan y sintetizan y se prueban empíricamente con RNA sintetizado *in vitro* en un espectrofluorómetro.

Determinados polinucleótidos o unidades de código de barras del polinucleótido de la invención comprenden uno de los siguientes sitios de unión, especialmente adecuados para el reconocimiento por sondas de baliza:

5'-TTCTCTTCAAACCTTTTCCGCTTTT-3' o,  
 5'-CGCCAAAACCTATTATCTTAAGTC-3' o;  
 5'-CTCACCTGCTCTTCTCAGACC-3' como sitio de unión 1  
 y 5'-GCTATAGCACTAAGGTAAGACCC-3' como segundo sitio de unión.

Estas secuencias proceden de *Drosophila* y no están presentes en el genoma de los mamíferos.

En una forma de realización específica de la invención, estos fragmentos de nucleótidos se utilizan juntos en una unidad de código de barras. Dicha unidad de código de barras puede contener sitio(s) de unión de reconocimiento y puede estar repetida en el código de barras.

En un polinucleótido específico, cada secuencia de repetición en tándem del código de barras está enmarcada por uno o varios sitios de restricción.

5 En general, la longitud del polinucleótido bicatenario de la invención está determinada por la longitud del código o códigos de barras y la longitud y el número de secuencias de promotor y de las secuencias opcionales adicionales (incluyendo secuencias de clonación o parte de las secuencias de restricción) y se adapta así a la necesidad de las aplicaciones deseadas. La longitud general del montaje de polinucleótido de la invención puede adaptarse al requisito de que dicho polinucleótido se integra de manera estable en una célula hospedadora determinada, especialmente en el genoma de dicha célula hospedadora.

10 A modo de ilustración, el polinucleótido puede ser tal que la unidad de código de barras o cada unidad de código de barras tiene una longitud comprendida en un intervalo de 16 a 200 nucleótidos.

15 También, a modo de ilustración, la longitud total del polinucleótido de la invención puede estar comprendida en un intervalo entre 80 y 30.000 nucleótidos.

A modo de ilustración, la siguiente es una descripción de las estructuras globales que se pueden encontrar en los polinucleótidos de la invención: los polinucleótidos puede comprender o contener, cuando se lee del extremo 5' a 3':

20 - un promotor seguido de un código de barras;

- un promotor seguido de la secuencia de codificación incluyendo el intrón o los intrones del gen que proporciona dicho promotor o seguido de parte o la totalidad de los intrones, seguidos de un código de barras;

25 - un gen completo de interés o una parte de dicho gen incluyendo su promotor seguido por un código de barras;

- un promotor seguido de un ADNc del gen de interés que proporciona el promotor seguido de un código de barras.

30 Estos diversos montajes pueden comprender además, después del promotor o en el extremo del montaje, una molécula indicadora especialmente un ADN indicador. También pueden comprender secuencias adicionales, tales como enzimas de restricción, secuencias de terminación de la transcripción en 3', cola poli A, microARN o ARN no codificador como los descritos en la presente memoria al considerar el polinucleótido de la invención.

35 Se enfatiza que todas las definiciones que se proporcionan en la presente memoria en relación con código de barras, las unidades de código de barras y los sitios de reconocimiento en el polinucleótido, se aplica en particular a los sitios de unión a la baliza, a las unidades de código de barras y a los códigos de barras de baliza, es decir, montajes que son adecuados para el reconocimiento por las sondas de baliza.

40 En una forma de realización específica de la invención, el polinucleótido comprende además, corriente abajo de la secuencia del promotor y corriente arriba o corriente abajo de la secuencia de código de barras, un ADN que codifica una proteína marcadora, en el que dicho ADN de codificación está bajo el control de elementos reguladores de la expresión, incluyendo bajo el control de dicho promotor que controla la transcripción del código de barras.

45 La presencia de una secuencia codificadora para una proteína indicadora o marcadora puede añadir algún nivel de determinación de la funcionalidad del promotor del polinucleótido y, especialmente, puede permitir la detección de la capacidad de expresión del montaje de polinucleótido de la invención. La proteína marcadora puede ser cualquier tipo de marcador conocido utilizado para el estudio de la expresión en una célula, tal como la proteína fluorescente verde (GFP), luciferasa o cloranfenicol acetiltransferasa (CAT), y preferentemente de proteínas marcadoras que no tienen ninguna actividad de expresión endógena.

50 El polinucleótido de la invención puede comprender además una secuencia codificadora de un gen de interés, ya sea en una versión que contiene intrones o como un ADNc, y/o el ADN que codifica una proteína indicadora o marcadora, bajo el control de dicho promotor para la transcripción del código de barras.

55 Como se ha indicado anteriormente, el gen de interés que proporciona el promotor del polinucleótido de la invención comprende genes implicados en la homeostasis de las células, y especialmente los genes de interés en el diseño de una estrategia terapéutica o genes implicados en el mantenimiento de la estructura celular o en la regulación del metabolismo celular, incluyendo los genes implicados en los estados patológicos en un hospedador, o activados como resultado de tal estado, especialmente en un hospedador humano. El gen de interés puede ser especialmente cualquier gen natural de mamífero, especialmente un gen natural humano, cuyas características de expresión o la regulación en una célula tiene que estudiarse y en especial cuya transcripción debe estudiarse. El gen de interés también puede ser una versión mutada de un gen natural, especialmente de un mamífero natural, en particular, un gen humano, cuya presencia e interés para caracterizar la expresión génica celular y especialmente para dicha perfilación en el transcrito debe investigarse. El gen de interés puede ser también un gen cuya expresión se provoca o regula en una célula hospedadora, especialmente en una célula de mamífero o en una célula humana, como

65

resultado de una infección patógena del mamífero u hospedador humano. El gen de interés puede ser especialmente un gen regulado por reacciones inmunológicas o que las regula en un hospedador.

5 A modo de ilustración, un gen de interés puede ser el implicado en la respuesta inmunitaria, en particular en la respuesta inmunitaria innata, tales como las interleucinas incluyendo IL10 o IL8, o moléculas de adhesión celular como ICAM o genes tales como TgF- $\beta$ , C-fos o CCL20.

10 Los genes específicos de interés pueden ser genes que están desregulados en un hospedador, o que son inducidos en un hospedador, especialmente un mamífero, en particular un hospedador humano, cuando dicho hospedador se ve afectado por un estado patológico, incluyendo, por ejemplo patologías que conducen a la proliferación celular descontrolada, especialmente el cáncer o patologías acompañadas de la desregulación del sistema inmunitario del hospedador.

15 Según una forma de realización específica, un gen de interés puede seleccionarse entre los genes que son sensibles a fármacos o a otros factores de tensión externa (incluso los organismos o agentes patógenos) en una célula estudiada.

20 Según una forma de realización específica, el gen de interés puede representar un compuesto experimental o una destino para un tratamiento terapéutico, cuya perfilación de la expresión, ya sea en o fuera y especialmente la perfilación de la transcripción tiene que determinarse después de la administración de este fármaco o factor de estrés.

25 En una forma de realización específica, la perfilación génica puede implicar el cribado de bancos de compuestos para determinar la interacción de los compuestos con los genes de la célula. Por lo tanto, la invención se refiere a la utilización del polinucleótido de la invención para controlar la perfilación génica en una célula, mediante cribado de un banco tal como un banco de ADN, un banco de ARNi, un banco de productos químicos o un banco de organismos patógenos.

30 La perfilación génica según la invención comprende el seguimiento de la expresión génica, mediante la observación de la transcripción del polinucleótido de la invención que comprende el promotor del gen, incluyendo la determinación de las condiciones para activar o provocar la transcripción génica (perfilación en marcha) o para desactivar la transcripción génica o en silencio (sin perfilación). También se refiere a la modulación de la expresión génica, ya sea para aumentar o para disminuir la expresión desde un nivel de partida de la transcripción. Ésta puede estudiarse en uno o varios genes en una célula.

35 La invención también se refiere al polinucleótido definido en la presente solicitud cuando se clona en un vector que permite la propagación en el interior de la célula que debe controlarse y posiblemente permite la inserción estable de dicho polinucleótido en el genoma de la célula.

40 Alternativamente, el polinucleótido puede permanecer como un episoma en la célula.

Los vectores específicos, adecuados para la realización de la invención son especialmente plásmidos, cósmidos, virus o bacs (cromosomas bacterianos artificiales).

45 La invención también se refiere a un conjunto de diferentes polinucleótidos de la invención.

50 La invención también se refiere a una célula, especialmente una célula viva y, en particular, una célula aislada o a una estirpe celular que comprende, especialmente la integración estable (es decir, integrado en el genoma celular) del polinucleótido de la invención. La célula puede ser una célula primaria o una estirpe celular.

55 Una célula o estirpe celular específica está formada por célula(s) eucariota(s) o célula(s) procariotas. Especialmente, es una célula o estirpe celular de mamífero especialmente una célula o estirpe celular humana con la exclusión de la(s) célula(s) o estirpe(s) celular(es) embrionaria(s) humana(s). Alternativamente, las células son de un roedor, especialmente un ratón u otro modelo animal apropiado, o de un pez cebra o de *Drosophila*.

En un aspecto particular de la invención, la célula de mamífero o la estirpe celular de mamífero es una célula o estirpe celular de mamífero no-humano.

60 En una forma de realización específica, la célula es una célula diferenciada o la estirpe celular procede de células diferenciadas.

En otra forma de realización, las células son células desdiferenciadas. Las células desdiferenciadas específicas son las células que han escapado del control de la proliferación *in vivo*.

En otra forma de realización, la célula o la estirpe celular procede de células pluripotentes o de células madre, especialmente de células pluripotentes o células madre humanas pluripotentes, ya sean células pluripotentes o células madre embrionario o adultas con la exclusión de las células o estirpes celulares embrionarias humanas.

5 Los ejemplos específicos de células que pueden utilizarse para realizar la invención abarcan las células implicadas en el sistema inmunitario tales como macrófagos, dendrocitos, monocitos o células epiteliales. Otras células son células cancerosas tales como células en desarrollo en el cáncer de mama.

10 Se ha demostrado que la introducción y especialmente la integración en el genoma, de un polinucleótido de la invención proporcionada en una celda, se puede conseguir de diferentes maneras, incluyendo por inyección en una célula o preferentemente por transfección de la célula con dicho polinucleótido. En el ejemplo proporcionado a continuación se proporciona una descripción de un protocolo de transfección que puede aplicarse a varias células y diversos polinucleótidos de la invención.

15 La invención también se refiere a una célula o estirpe celular definida anteriormente y se ilustra en los ejemplos, la cual comprende, además, una o varias sondas moleculares de detección tal como se definen en la presente solicitud. Como para el polinucleótido de la invención, dicha(s) sonda(s) molecular(es) se introduce(n) en la célula por inyección o preferentemente por transfección.

20 La invención también se refiere a un conjunto de células o a un conjunto de estirpes celulares con la exclusión de las células o estirpes celulares embrionarias humanas, en las que cada conjunto de células o estirpes celulares comprende especialmente integrado en las mismas, uno o varios polinucleótido(s) tal como se define en la presente memoria, teniendo cada célula o estirpe celular una composición en dicho polinucleótido(s) que es diferente de la de la otra célula(s) o estirpe(s) celular(es).

25 Como se describe en la presente memoria, el seguimiento de la expresión génica en una célula, que ha sido modificada mediante la inserción del polinucleótido de la invención requiere medios para detectar la transcripción de dicho polinucleótido. La detección puede realizarse utilizando sondas (sondas moleculares) que pueden detectar cambios mensurables en los transcritos. Los cambios mensurables pueden ser cambios en la configuración del ácido nucleico como resultado de la hibridación de la sonda o los cambios en la emisión de luz (tal como fluorescencia) como resultado de la hibridación de la sonda.

30 Las sondas adecuadas para detectar la expresión del polinucleótido de la invención puede ser sondas lineales o no lineales, tales como las sondas de baliza. Son sondas de ácido nucleico, especialmente sondas de ADN.

35 En una forma de realización específica, la invención por lo tanto se refiere a una baliza molecular y a su utilización como sonda para detectar la transcripción del polinucleótido de la invención. Las balizas moleculares son moléculas de ADN adecuadas para hibridarse con sus secuencias complementarias en las secuencias transcritas procedentes de sitio(s) de unión a la baliza del o de los código(s) de barras de la(s) baliza(s) en los polinucleótidos de la invención, teniendo dicha(s) baliza(s) molecular(es) una estructura de polinucleótido en tallo-lazo y siendo adecuada para visualización cuando se hibrida a su secuencia diana, especialmente de manera reversible.

40 Una baliza molecular específica se define como una estructura de polinucleótido en tallo-lazo en la que la parte de lazo del polinucleótido es la secuencia de la sonda adecuada para hibridarse específicamente con un sitio de unión de la baliza del polinucleótido y la parte de tallo se compone de dos brazos formados por secuencias complementarias entre sí, alojando cada una de las secuencias de brazo, unida a su extremo libre, que es contrario a la parte de lazo del polinucleótido, un resto fluorescente o un resto de extinción no fluorescente en el que dichos restos, cuando está unido a dichas secuencias de brazo, están suficientemente próximos unos de otros para producir fluorescencia del resto fluorescente que que va a extinguirse por transferencia de energía de resonancia de fluorescencia, y además dicha parte de lazo del polinucleótido es al menos dos veces más larga en los nucleótidos que cada estructura de polinucleótido de brazo.

45 Según su definición, las balizas moleculares definidas en la presente memoria tienen una secuencia de sonda que es complementaria al sitio o sitios de unión de la baliza de un código de barras de la baliza del polinucleótido o, especialmente, a su transcrito.

50 La expresión "complementaria" significa que la baliza molecular es adecuada para hibridar en condiciones de hibridación definidas, especialmente en condiciones fisiológicas, por emparejamiento de bases con la secuencia complementaria de un sitio de unión de baliza molecular. En el contexto de perfilación génica en una célula según la invención, la hibridación de la sonda de baliza se logra con el transcrito del polinucleótido de la invención, en los sitios de unión de la baliza.

55 Preferentemente, la secuencia de la sonda de la baliza molecular es perfectamente complementaria a los sitios de unión de las balizas moleculares específicas o su transcrito en una célula, y por lo tanto es adecuada para formar un híbrido estable con dicho transcrito en una célula de manera que su hibridación permite un cambio mensurable, p.

ej., un cambio de configuración, que se produce en la baliza molecular, lo que les permite elevar una señal, especialmente para emitir luz fluorescente y especialmente para florecer.

5 Como se conoce en la técnica, la fluorescencia de baliza molecular procede de la capacidad del fluoróforo (resto fluorescente) y del inhibidor (resto atenuador no fluorescente) para alejarse uno de otro, como resultado de la hibridación, permitiendo de este modo la fluorescencia.

10 Las reglas de diseño para la preparación de las balizas moleculares son bien conocidas por los expertos y se puede encontrar sobre todo en el siguiente página web: <http://www.molecular-beacons.org>. Dichas sondas de baliza se ilustran en los ejemplos.

15 En una forma de realización específica, las balizas moleculares son tales que el resto fluorescente y resto atenuador no fluorescente están unidos por enlace covalente a las secuencias de brazo de la estructura de polinucleótido en tallo-lazo.

En una baliza molecular específica, la estructura de polinucleótido de la sonda de baliza tiene de 8 a 100 nucleótidos.

20 Se pueden utilizar muchos restos fluorescentes diferentes para preparar balizas moleculares de la invención junto con un resto atenuador para ilustración. Se proporcionan los siguientes ejemplos. Según formas de realización específicas, el resto fluorescente (fluoróforo) se selecciona en el grupo de Quantum Dots (puntos cuánticos) y derivados, familia Alexafluor de colorantes, FAM, TET o CAL FluorGold 540, HEX o JOE, VIC<sup>B</sup>, CAL Fluor Orange 560<sup>A</sup>; Cy3<sup>C</sup> o NED<sup>B</sup>, Quasar 570<sup>A</sup>, Oyster 556<sup>D</sup>; TMR o CAL Fluor Red 590<sup>A</sup>; ROX o LC rojo 610<sup>E</sup>, CAL Fluor Red 610<sup>A</sup>, rojo Texas, o rojo LC 610<sup>E</sup>, CAL Fluor Red610<sup>A</sup>; LC Rojo 640<sup>E</sup> o CAL Fluor Red 635<sup>A</sup>; Cy5<sup>C</sup> o LC Rojo 670<sup>E</sup>, Quasar 670<sup>A</sup>, Oyster 645<sup>D</sup>, LC rojo 705<sup>E</sup> o Cy5.5<sup>C</sup> o ácido 5-(2'-aminoetil)aminonaftalen-1-sulfónico (EDANS), fluoresceína, antranilamida, cumarina, y quelatos de terbio, y el resto atenuador (extintor) se selecciona en el grupo de DDQ-I<sup>A</sup> (absorción máx. 430 nm), Dabcyl (absorción máx. 475), Eclipse<sup>B</sup> (absorción máx. 530), Iowa Black FQ<sup>C</sup> (absorción máx. 532), BHQ-1<sup>D</sup> (absorción máx. 534 nm), QSY-7<sup>E</sup> (absorción máx. 571), BHQ-2<sup>D</sup> (absorción máx. 580), DDQ-II<sup>A</sup> (absorción máx. 630), Iowa Black RQ<sup>C</sup> (absorción máx. 645), QSY-21<sup>E</sup> (absorción máx. 660), BHQ-3<sup>D</sup> (absorción máx. 670), oro, metales de tierras raras o ácido 4-(4'-dimetilaminofenilazo)benzoico (DABCYL), rodamina, pirenobutirato, eosina, nitrotirosina, etidio y tetrametil-rodamina.

35 Los pares específicos de fluoróforo/atenuador son los que proceden de las listas anteriores por asociación de los respectivos fluoróforo/atenuador marcados por la misma letra mayúscula (A, B, C, D o E) u otros pares seleccionados en el grupo de las siguientes combinaciones: CY5 con BlackHole Quencher 3, CY3 con BlackHole Quencher 2, ácido 5-(2'-aminoetil)aminonaftaleno-1-sulfónico (EDANS) con ácido 4-(4'-dimetilaminofenilazo)benzoico (DABCYL), fluoresceína con rodamina, fluoresceína con pirenobutirato, fluoresceína con eosina, antranilamida con nitrotirosina, cumarina con etidio, y quelatos de terbio con tetrametil-rodamina.

40 Las sondas alternativas son otras sondas a base de ácidos nucleicos. Estas sondas en la hibridación en el sitio de reconocimiento provocan un cambio mensurable tal que puede medirse con diferentes instrumentos. Los ejemplos de sondas lineales de ácido nucleico, sondas de desplazamiento de cadena, sondas Scorpion, sondas adyacentes, sondas cyclicon, HyBeacons, sondas de fijación al surco menor, sondas de baliza molecular (incluyendo sondas de baliza molecular de desplazamiento de longitud de onda), sondas de 5' nucleasa y sondas Amplifluor.

45 Las sondas adyacentes tales como las descritas en Wittwer, C.T., Herrmann, M.G., Moss, A.A. y Rasmussen, R.P. Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechniques* 22, 130-131, 134-138 (1997) son adecuados para la detección de los transcritos de los polinucleótidos según la invención. Estas sondas utilizan dos sondas de hibridación monocatenarias que se unen a sitios próximos en un ácido nucleico diana. Una sonda puede estar marcada con un fluoróforo donante y el otro con un receptor. La distancia entre las dos sondas una vez hibridadas se elige de manera que puede producirse transferencia eficaz de energía fluorescente a partir del donante al receptor. La transferencia de energía debe ser mínima cuando las sondas no se hibridan a su objetivo. Además de los cambios en la señal de fluorescencia, otros cambios mensurables puede utilizarse para indicar la hibridación mediante sondas adyacentes a los sitios cercanos. Así pues, estas sondas no son sólo las sondas a base de FRET.

50 Las sondas de desplazamiento de cadena, incluyendo las sondas Scorpion descritas en Solinas, A. *et al.* Duplex Scorpion primers in SNP analysis and FRET applications. *Nucleic Acids Res.* 29, E96 (2001) son sondas más adecuadas para llevar a cabo la detección del polinucleótido de la invención. Estas sondas utilizan dos sondas de oligonucleótidos complementarias, una sonda marcada con un fluoróforo y la otra marcada con un resto atenuador no fluorescente. Cuando las dos sondas están hibridadas entre sí el fluoróforo y el atenuador están en proximidad estrecha resultante en temple contacto que ocurre y emisión de fluorescencia baja. En presencia de un ácido nucleico diana una de las sondas forma un híbrido sonda blanco resultante más estable en las dos sondas están separadas. Como consecuencia de este desplazamiento, el fluoróforo y el atenuador ya no están en estrecha proximidad y la fluorescencia aumenta. Además de los cambios en la señal de fluorescencia, pueden utilizarse otros

cambios mensurables para indicar los cambios en el desplazamiento de cadena. Por lo tanto, estas sondas no son solo a base de fluoróforo-atenuador.

Es posible utilizar muchos códigos de barras de unión de reconocimiento de ácidos nucleicos distintos, en los que cada uno de ellos se fusiona a una secuencia específica de promotor endógeno, para proporcionar muchos montajes de polinucleótidos para su utilización en una sola célula simultáneamente. En este caso, las sondas de detección se pueden diseñar con moléculas que producen señales distinguibles. Por lo tanto, esto permite la perfilación de múltiples genes en células aisladas en tiempo real. Por ejemplo, esto es posible mediante la utilización de balizas de muchos colores, es decir balizas diferentes con colores diferentes, que se hibridan específicamente a varios sitios de unión de ácidos nucleicos.

En una forma de realización específica, los polinucleótidos de la invención están diseñados de tal manera que, cuando su transcrito se hibrida con una sonda de detección y especialmente con las balizas moleculares, proporcionan un perfil de expresión génica que permite la visualización de más de 1, especialmente hasta 4, o más especialmente hasta 6, y en su caso hasta 15, 20 o hasta 32 transcritos de los polinucleótidos de la invención. Esto puede conseguirse combinando un código de barras, un conjunto de sitios de unión de la baliza, que, cuando se hibridan, proporcionan un espectro con una combinación de colores que pueden descomponerse para proporcionar una huella digital del transcrito del polinucleótido.

A modo de ilustración, cuando se dispone de diferentes colores de fluoróforos, éstos pueden utilizarse combinados para obtener un espectro que tiene más de un color. Si dichos espectros se pueden descomponer en sus diversos componentes cuando se visualizan, esto proporciona posibilidades para permitir la visualización de más transcritos que el número correspondiente a los colores básicos disponibles de los fluoróforos. A fin de conseguir dicho espectro con combinaciones de colores, el código de barras de la baliza que debería dar lugar al espectro multicolor debería comprender sitios de unión de baliza complementarios a las balizas que tienen fluoróforos de diferentes colores.

Como se ilustra en las figuras, las combinaciones de colores fluorescentes de las sondas unidas a cada unidad de código de barras se puede utilizar para caracterizar muchos genes en una sola célula.

Especialmente, cuando una unidad de código de barras comprende 2 sitios de unión de reconocimiento, estando marcado cada uno de los cuales con una etiqueta diferente, especialmente un color diferente, es posible, recurriendo a seis etiquetas diferentes, tales como restos coloreados o fluorescentes, caracterizar hasta 15 genes en una sola célula.

Cuando una unidad de código de barras comprende 3 sitios de unión de reconocimiento, cada uno de los cuales está marcado con una etiqueta diferente, es posible, recurriendo a seis etiquetas diferentes, tales como restos coloreados o fluorescentes, caracterizar hasta 20 genes en una sola célula.

Las sondas de detección definidas de este modo son complementarias a la secuencia del transcrito que se transcribe desde el código de barras o hasta una parte de éste suficiente para caracterizar dicho transcrito. En una forma de realización específica, la complementariedad es tal como dos secuencias de pares de bases perfectamente.

En una forma de realización específica de la invención, la sonda de detección es una baliza molecular y su parte de lazo tiene de 10 a 55 nucleótidos y la estructura de cada brazo de polinucleótido tiene 4 a 16 nucleótidos.

Las balizas moleculares particulares adecuadas para llevar a cabo la invención tienen una de las siguientes secuencias de nucleótidos: 5'-GCUGCAAAAGCGGAAAAGUUUGAAGAGAAGCAGC-3' o 5'-CGACCGACUUAAGAUAAUAGGUUUUGGCGGGUCG-3'.

La invención también proporciona un kit que comprende (i) un polinucleótido como se define en la presente memoria y (ii) células o estirpe celular adecuadas para la integración de dicho polinucleótido y/o una o más sondas moleculares de detección, con la exclusión de células o estirpes celulares embrionarias apropiadas.

En una forma de realización específica, el kit comprende un polinucleótido como se describe en la presente solicitud y una célula o una estirpe celular adecuada para la integración de dicho polinucleótido o un conjunto de dichos polinucleótidos y un conjunto de dichas células o estirpes celulares, con la exclusión de células o estirpes celulares embrionarias apropiadas.

Según otra forma de realización, el kit comprende, además o alternativamente estirpes celulares con polinucleótidos de código de barras integrado corriente abajo de forma estable del gen o promotor de interés.

En una forma de realización específica, el kit comprende además una o varias sondas moleculares de detección adecuada(s) para hibridarse con el sitio de unión de reconocimiento (s) del código de barras (s).

La invención también se refiere a un kit que comprende un polinucleótido como se describe en la presente solicitud y una o varias sonda de detección molecular (s) adecuada para hibridarse con el/los sitio(s) de unión de reconocimiento o el/los código(s) de barras o su transcrito.

5 En una forma de realización específica, el kit comprende un polinucleótido como se describe en la presente solicitud y una o varias sonda(s) de detección molecular(es) adecuada(s) para hibridarse con el/los sitio(s) de unión de reconocimiento del o de los código(s) de barras, o con su transcrito, una célula o una estirpe celular adecuada para la integración de dicho polinucleótido o un conjunto de dichos polinucleótidos y un conjunto de dichas células o estirpes celulares.

10 En una forma de realización específica de la invención, el kit descrito en la presente memoria comprende balizas moleculares que tienen las secuencias de nucleótidos siguientes.

15 5'-GCUGCAAAAGCGGAAAAGUUUGAAGAGAAGCAGC-3' o  
5'-CGACCGACUUAAGAUAAUAGGUUUUGGCGGGUCG-3'.

20 Las sondas de detección molecular están diseñados de tal manera que la visualización de su hibridación con su objetivo se obtiene como resultado de un cambio mensurable, tal como una fluorescencia que se conecta, cuando la sonda de detección molecular se une a su secuencia diana.

25 Opcionalmente, el kit definido comprende además un plásmido con un polinucleótido de código de barras clonado corriente abajo del gen o promotor de interés, y/o sondas (por ejemplo balizas moleculares) capaces de reconocer el polinucleótido de código de barras, y/o reactivos (péptidos, lípidos, productos químicos, etc.) capaces de introducir sondas y plásmidos en las estirpes celulares especialmente por transfección o microinyectarlos en la célula y/o referencias positivas y negativas para cada paso en el procedimiento experimental según se requiera.

La solicitud da a conocer un procedimiento de estudio de la transcripción génica en una célula o en una estirpe celular que comprende las etapas siguientes:

- 30 (i) proporcionar una célula o una estirpe celular que comprende, integrar especialmente de forma estable un polinucleótido como se define en la invención,
- (ii) provocar, silenciar o modular la transcripción del montaje del polinucleótido,
- 35 (iii) detectar un cambio mensurable en la transcripción del polinucleótido.

La etapa que consiste en detectar el cambio mensurable en la transcripción del polinucleótido puede comprender:

- 40 • poner en contacto la célula o estirpe celular en la que se ha introducido el polinucleótido, con una o más sonda(s) de detección tal como se define en la presente memoria;
- detectar la hibridación entre dicha(s) sonda(s) de detección y el transcrito del sitio o sitios de unión de reconocimiento del código de barras, como indicador de la actividad de transcripción del promotor del montaje del polinucleótido.

45 El estudio de la transcripción génica permitida por el procedimiento se consigue como resultado de la activación del promotor en el polinucleótido que da lugar a la transcripción del código de barras molecular, lo que permite la producción de múltiples copias de dicho código de barras molecular que conduce a su unión a sus sondas de detección molecular complementarias específicas.

50 La etapa (ii) del procedimiento puede llevarse a cabo poniendo la célula o estirpe celular en contacto con una molécula o agente de interés, cuyo impacto en la transcripción se estudia.

55 Los procesos de integración especialmente para la integración estable del polinucleótido de la invención en las células son bien conocidos por el experto en la materia y en general incluyen los protocolos de inyección o transfección. Dichos protocolos se ilustran en los ejemplos recurriendo a las células HeLa y polinucleótidos concretos. Técnicas similares se puede repetir para diferentes células y polinucleótidos.

60 Asimismo, puede conseguirse poner en contacto las células con sondas de detección por métodos como los ilustrados en los ejemplos.

65 Cuando un gen de interés y/o una secuencia que codifica una proteína marcadora está también presente en el polinucleótido de la invención, la actividad de transcripción del promotor también permite la transcripción y la expresión de dicho gen y/o proteína marcadora.

De esta manera, la actividad del promotor puede detectarse en la transcripción, dando como resultado un cambio mensurable, especialmente de fluorescencia o de otra señal emisora de luz de la sonda de detección, y posiblemente en la expresión del producto del gen de interés y/o de la proteína marcadora.

5 Se pueden utilizar algunos medios apropiados para la detección de los cambios mensurables, especialmente de fluorescencia, a fin de controlar la actividad de transcripción del promotor del polinucleótido de la invención. Especialmente, la microscopia de alto contenido y técnicas de detección por la imagen de células vivas son adecuados para llevar a cabo la detección.

10 Una ventaja concreta de la invención, es que el procedimiento que se describe permite el estudio de la transcripción génica reflejada por la transcripción de los polinucleótidos que comprenden los promotores de dichos genes en una sola célula viva.

15 El procedimiento dado a conocer está especialmente adaptado para el estudio de la transcripción del polinucleótido en un ensayo en tiempo real, o en un punto final.

20 Otra ventaja del procedimiento dado a conocer es que puede llevarse a cabo para estudiar la transcripción génica en un solo gen, especialmente en una sola célula, o para estudiar la transcripción de genes en una base de múltiples genes, especialmente en una sola célula. Se ha demostrado que la transcripción de múltiples genes reflejada por la transcripción de los polinucleótidos que comprenden sus promotores, se puede estudiar durante un máximo de 35 genes en una sola célula, y especialmente 2, 3, 4, 5, 6 genes se pueden estudiar en un solo ensayo.

25 Según un aspecto dado a conocer, la etapa de provocación, silenciamiento o modulación de la transcripción del polinucleótido se obtiene después de la puesta en contacto de la célula o estirpe celular con un factor exterior.

Dicho factor externo puede ser un banco de compuestos químicos, pequeños ARN de interferencia o un banco de organismos que se detecta en la célula o estirpe celular.

30 También es especialmente apropiado y destinado para la detección cuantitativa de la transcripción de polinucleótidos, por detección cuantitativa de los episodios de unión de las sondas de detección con sus secuencias diana como resultado de la transcripción de los polinucleótidos de la invención.

35 En una forma de realización específica, el procedimiento dado a conocer comprende además la detección de una proteína indicadora expresada codificada por el montaje del polinucleótido y expresada bajo el control del promotor contenido en el montaje del polinucleótido.

40 Los montajes de promotor específico contenidos en la secuencia de polinucleótidos de la invención, y cuya actividad se detecta por el procedimiento de la invención son los promotores de los genes que se han descrito en la presente memoria.

45 Cuando el procedimiento dado a conocer se lleva a cabo con objeto de ensayar la actividad de transcripción del promotor de IL8, se provoca dicha actividad por infección de la célula o estirpe celular con bacterias, por ejemplo, con la cepa *S. flexneri*.

En una forma de realización específica del procedimiento dado a conocer, el polinucleótido expresa además un gen indicador seleccionado en el grupo de GFP y luciferasa.

50 Dentro del alcance de la invención se contemplan varias aplicaciones del kit o del polinucleótido de la invención. Según una forma de realización de la invención, pueden utilizarse para el cribado de bancos, incluyendo bancos de ARNi, genotecas, bancos de productos químicos o bancos de organismos patógenos.

55 Según una forma de realización de la invención, pueden utilizarse en un procedimiento de diagnóstico, especialmente como parte de dicho diagnóstico, con objeto de detectar una enfermedad, especialmente una infección tal como una infección vírica o bacteriana en un hospedador, en particular en un hospedador humano.

60 Puede utilizarse también para hacer el seguimiento de reacciones inmunológicas en una sistema celular. Esto puede implicar el ensayar la transcripción de un polinucleótido de la invención cuando la célula ensayada se pone en contacto con los compuestos de una banco para el cribado (incluyendo bancos químicos, bancos de ARNi, genotecas o bancos de patógenos). El resultado del ensayo se detectó mejor utilizando microscopia de alto contenido y alto rendimiento.

Otra aplicación de la invención es el seguimiento de los resultados de un tratamiento terapéutico en un sistema celular.

La invención especialmente se puede utilizar en el proceso de cribado de potenciales compuestos terapéuticos o dianas terapéuticas y especialmente en el proceso de selección de compuestos, interactuando posiblemente con el sistema inmunitario.

- 5 Otra aplicación de la invención es para la investigación de dianas celulares de un compuesto, de un organismo o agente patógeno.

10 La invención se refiere también el seguimiento de las interacciones entre un patógeno y un hospedador, en una célula del hospedador o de una célula derivada de dichas células, especialmente cuando dicha célula se coloca en condiciones de infectarse con dicho patógeno.

15 La solicitud da a conocer también un mamífero transgénico no humano que expresa en el genoma de algunas o todas sus células, un montaje del polinucleótido de la invención, o células transformadas y especialmente transfectadas con dicho polinucleótido de la invención.

20 Mas características de la invención y las propiedades se pondrán de manifiesto a partir de los ejemplos y figuras siguientes: se menciona especialmente que las características descritas en los ejemplos y figuras se pueden aplicar de una manera más general para llevar a cabo la invención y, en consecuencia, excepto donde se indica, no se limitan a los productos específicos ilustradas en los ejemplos.

### Descripción de las figuras

25 Figura 1. A. Principio de balizas moleculares. Las balizas moleculares son sondas de oligonucleótidos en forma de tallo-lazo internamente atenuadas que resultan fluorescentes en la hibridación con su secuencia diana, tales como ARN o ADN.

30 B. Principio de nucleótidos de baliza molecular con código de barras de balizas. Los nucleótidos con código de barras pueden construirse con un promotor de interés que dirige la expresión de una proteína fluorescente indicadora corriente arriba de los nucleótidos de baliza con código de barras. Para los experimentos para controlar la respuesta proinflamatoria se desplegó el primer tipo nucleótidos de baliza con código de barras. Se prepararon dos montajes de baliza con código de barras molecular. IL8WT utilizó el promotor mínimo bien caracterizado de IL8 y se introdujo corriente arriba de la proteína verde fluorescente (GFP) y nucleótidos de baliza con código de barras molecular se introdujeron corriente abajo de GFP con la secuencia 3' no traducida. IL8MUT el segundo montaje de baliza con código de barras era idéntico en todos los aspectos al IL8WT, sin embargo, se introdujeron dos mutaciones en el promotor mínimo haciéndolo insensible a la activación. Estos montajes se han depositado como plásmidos en cepas de *E. coli*, en la CNCM (situada en París, Francia) con el número CNCM I-3.818 (pJoMull8WT) y CNCM I-3819 (pJoMull8Mut) el 4 de septiembre de 2007.

40 Los plásmidos están en la cepa DH5 $\alpha$  de *E. coli* y son resistentes a la kanamicina.

El medio de cultivo para el crecimiento es medio LB con kanamicina y la siembra se realiza con LB-agar-agar. La incubación es a 37°C con agitación constante.

45 El almacenamiento es posible por congelación a -80°C en líquido de suspensión de LB/50% de glicerol con una concentración celular de 10<sup>7</sup>/ml.

50 C. Principio de ensayos basados en células vivas utilizando balizas moleculares y los montajes de balizas con códigos de barras. (De izquierda a derecha) Los montajes de baliza con código de barras IL8WT y IL8MUT se utilizan por separado para crear dos estirpes celulares estables diferentes. En estas estirpes celulares se comprueba la integración estable de cada montaje. El ensayo basado en células se lleva a cabo introduciendo en primer lugar la baliza molecular apropiada por transfección o microinyección en la estirpe celular estable. La célula se incuba a continuación durante unas pocas horas antes de realizar el experimento o ensayo de cribado. Se añaden a las células diferentes factores tales como bacterias, el factor de necrosis tumoral alfa (TNF $\alpha$ ) u otros. La respuesta de la transcripción de las células se controla por microscopia confocal de disco giratorio en condiciones ambientales apropiadas (temperatura y medio). La fluorescencia de la baliza molecular indica la transcripción del ARNm.

60 Las estirpes celulares Hela transfectadas de manera estable con pJoMu I18WT o con pJoMull8Mut se depositaron en la CNCM (París, Francia) el 4 de septiembre de 2007, con los números respectivos CNCM I-3820 (JoMu) y CNCM I-3821 (mutante JoMu).

Cada selección de estirpe celular transfectada de transfectantes estables se llevó a cabo con genética.

65 El medio de cultivo de las estirpes celulares transfectadas es DMEM "Medio Eagle Modificado con Dulbecco", +10% de suero bovino de ternera fetal, + 50 UI/ml de penicilina + estreptomycinina, + L-glutamina 2 mM; pH 7-4, temperatura óptima 37°C, 5% de CO<sub>2</sub> en fase gaseosa.

Las condiciones de cultivo son para células de la monocapa unidas, con un tiempo de duplicación de la población de 16 h, una relación de división óptima de una vez en 3 a 4 días y una vida correspondiente a una estirpe celular de cáncer. El subcultivo se consigue mediante tratamiento sencillo con tripsina similar a la división de las células Hela.

5 El almacenamiento es posible en nitrógeno líquido.

Figura 2

10 Antes de realizar ensayos para seguir la respuesta de la transcripción de las estirpes celulares IL8WT y IL8MUT los inventores han comprobado si el montaje era funcional y sensible a los estímulos normales del promotor IL8. Se sabe que TNF $\alpha$  estimula directamente el promotor de IL8 y los inventores construyeron a propósito el promotor IL8MUT para que fuera refractivo a TNF $\alpha$  que debería dar lugar a que IL8WT conduzca a la expresión de un GFP. Los inventores llevaron a cabo el experimento en cuatro condiciones diferentes para satisfacerse de que ambos nucleótidos de baliza con código de barras funcionaran como se esperaba.

15 Condición 1: IL8WT con inducción de TNF $\alpha$

20 Montaje IL8WT que contiene células estimuladas con TNF- $\alpha$  y controladas durante varias horas tras la adición de TNF $\alpha$ . Las células se tiñeron con DAPI para marcar su presencia en el campo microscópico. Aproximadamente cuatro horas después de la estimulación con TNF $\alpha$  se observó GFP en varias de las células en el campo (fila 1, columnas 1 y 3).

Condición 2: IL8WT sin inducción

25 Montaje IL8WT que contiene células controladas durante varias horas. A menos que se utilicen métodos a los que el promotor de IL8 es sensible para estimular la respuesta proinflamatoria, no se observa fluorescencia en GFP. Se observa un nivel muy bajo de fluorescencia de fondo lo que indica que en algunas células puede estar activo un nivel bajo de estímulo proinflamatorio. (Fila 2, columnas 1 y 3).

30 Condición 3: IL8Mut (mutante NF $\kappa$ B/AP1) con inducción de TNF

Montaje IL8MUT que contiene células estimuladas con TNF $\alpha$  y controladas durante varias horas tras la adición de TNF $\alpha$ . No se observa aumento de fluorescencia y no hay nivel visible de fluorescencia de fondo que se observa en la condición 2. (Fila 3, columnas 1 y 3).

35 Condición 4: IL8Mut (mutante NF $\kappa$ B/AP1), sin inducción

40 Montaje IL8MUT que contiene células controladas durante varias horas. No se observa aumento de fluorescencia durante el período del ensayo. (Fila 4, columnas 1 y 3).

Figura 3

45 Las estirpes celulares IL8WT y IL8MUT sometieron a la prueba de provocación con dos cepas bacterianas diferentes de *Shigella flexneri*. Una cepa, M90T de *S. flexneri* se sabe que provoca una fuerte respuesta proinflamatoria en células humanas. La segunda cepa, OSPF de *S. flexneri* (Arbibe *et al.*, 2007) se utilizó ya que su activación de la respuesta proinflamatoria es ambigua. En todos los casos, las células se incubaron con cualquiera de las cepas de *S. flexneri* durante 30 minutos. Las células se lavaron a continuación cuidadosamente y los medios se intercambiaron con uno que contenía gentomicina 10  $\mu$ M. Las células se incubaron a continuación a 37°C durante 3,5 horas, haciendo seguimiento mientras de la actividad por microscopia confocal. Se realizó el experimento en las siguientes tres condiciones diferentes.

50 Condición 1: IL8WT con M90T de *S. flexneri*

55 Las células incubadas con M90T de *S. flexneri* durante 30 minutos que contienen el montaje IL8WT se controlaron durante varias horas por microscopia confocal de disco Nipkow. Las células se tiñeron con DAPI para marcar su presencia en el campo microscópico. Aproximadamente tres horas después de la incubación se observó GFP en varias de las células en el campo (Fila 1 columnas 1 y 3).

60 Condición 2: IL8MUT con M90T de *S. flexneri*

Las células controladas que contienen el montaje IL8MUT se incubaron con M90T de *S. flexneri* durante 30 minutos. Las células se controlaron a continuación, durante varias horas a través de microscopia confocal de disco Nipkow. No se observó fluorescencia de GFP incluso después de seguimiento durante varias horas lo que indica que el promotor era refractivo a la estimulación por la respuesta inflamatoria. (Fila 2, columnas 1 y 3)

65 Condición 3: IL8WT con OSPF de *S. flexneri*

Las células que contienen el montaje IL8WT incubadas con OSPF de *S. flexneri* durante 30 minutos se controlaron durante varias horas por microscopia confocal de disco de Nipkow. Casi 4 horas después de la incubación se observó un bajo nivel de fluorescencia de GFP en algunas células. El mecanismo por el que las variantes OSPF provocan la respuesta inflamatoria es ambiguo. OSPF se ha construido al producir episodios específicos de remodelación de la cromatina en el hospedador lo que conduce a la estimulación de la respuesta proinflamatoria (fila 3 columnas 1 y 3).

Figura 4

Las estirpes celulares IL8WT y IL8MUT se transfectaron con balizas moleculares complementarias a los nucleótidos de baliza con código de barras codificados en cada montaje. Las células transfectadas se incubaron con las balizas moleculares solas durante por lo menos 2 horas. El medio se intercambiò a continuación, y las dos estirpes celulares se sometieron a la prueba de provocación con la cepa bacteriana M90T de *Shigella flexneri* conocida por provocar una fuerte respuesta proinflamatoria en células humanas. Se añadió M90T de *S. flexneri* a las células, haciendo seguimiento mientras de su actividad de transcripción por microscopia confocal de disco Nipkow.

A. Células IL8WT incubadas con M90T de *S. flexneri*. A los pocos minutos de la adición de las bacterias se observa un rápido aumento de fluorescencia de la baliza molecular. Se observa que la inducción IL8 aumenta rápidamente (C) hasta un nivel que satura el detector CCD después de 20 minutos. Se observan diversos en varios niveles de inducción y se ha seleccionado una célula dentro del campo para la que se observa dicha inducción de su estado "desconectado" para la transcripción.

B. Células IL8MUT incubadas con M90T de *S. flexneri*. Después del seguimiento a las células de más de 50 minutos no se observa ganancia en la intensidad de fluorescencia lo que indica falta de activación del montaje IL8MUT en las células. Aquí el nivel de fluorescencia se muestra en un período similar que en A.

Cinética de inducción de la transcripción de IL8 después de la invasión de M90T de *S. flexneri*. El eje Y indica fluorescencia integrada, mientras que el eje X es el tiempo. El gráfico muestra la inducción rápida de IL8 que es inducida por dos veces en tan sólo 8 minutos. El pico en la intensidad de fluorescencia alcanzado después de 20 minutos es indicador de saturación del detector de CCD y no necesariamente de la detención en la inducción de la vía.

Figura 5

Ilustración de combinación de colores para caracterizar muchos genes.

**Figura 5A:** Esquema de codificación de dos colores. Los puntos diferentes corresponden a diferentes colores de los fluoróforos. Una unidad de baliza con código de barras consta de 2 secuencias tipo I y tipo II. Cada tipo puede estar unido a una baliza molecular o a una sonda de ácido nucleico. Las sondas complementarias al tipo I pueden ser de un color similar o diferente que las de tipo II. Utilización dos colores diferentes unidos a cada unidad de código de barras, como se indica en este esquema, un código de combinación de colores se puede utilizar para caracterizar muchos genes en una sola célula.

**Figura 5B:** Esquema de codificación de tres color. En este esquema, una unidad de baliza con código de barras consta de 3 secuencias de tipo I, tipo II y tipo III. Cada tipo puede estar unido a una baliza molecular o una sonda de ácido nucleico. Las sondas complementarias al tipo I pueden ser de un color similar o diferente que las de tipo II y tipo III. Utilizando tres colores diferentes unidos a cada unidad de código de barras, como se indica en este esquema, se puede utilizar una combinación de colores para caracterizar muchos genes (20 genes) en una sola célula.

### Ejemplos

Para realizar los estudios iniciales de la expresión génica de una sola molécula se eligió examinar la transcripción inducida por patógenos en células hospedadoras (humanas). Un sistema modelo específico es la invasión por *Shigella flexneri* de las células humanas. Un trabajo reciente ha propuesto que *S. flexneri* altera específicamente el estado de la transcripción de células invadidas, activando respuestas inmunitarias oportunistas para la bacteria. El análisis del transcriptoma utilizando chips ha sugerido genes específicos del hospedador que *S. flexneri* dirige durante la invasión de las células hospedadoras, induciendo un patrón muy específico de respuesta de la transcripción. Los inventores han desarrollado herramientas moleculares para seguir esta respuesta de la transcripción en tiempo real. En combinación con los procedimientos existentes basados en detección por la imagen para controlar los cambios bioquímicos, de proteínas y morfológicos, los inventores han construido una plataforma que ofrece una forma revolucionaria de llevar a cabo la biología experimental. Los experimentos se realizan en células aisladas que se dirigen a un único gen hospedador y su respuesta a la invasión bacteriana se investiga en modo de alto rendimiento. Además, la plataforma puede utilizarse para la detección de factores químicos, mutantes de *S. flexneri*, o mutantes de hospedador (utilizando RNAi), que pueden alterar la respuesta génica del hospedador favoreciendo al hospedador o a la bacteria.

Biología molecular

Se diseñaron dos secuencias complementarias de oligonucleótidos de 81 nt, que contienen dos sitios de unión de baliza molecular con una secuencia intermedia de 8 nt. La secuencia 1 (cadena +ve) era 5'-TTCTCTTCAAACCTTTTCCGCTTTT-3' y la secuencia 2 (cadena +ve) era 5'-CGCCAAAACCTATTATCTTAAGTC-3'. La cadena positiva completa de la secuencia del oligonucleótido fue 5'-ACGCGTCGACTTCTTCAAACCTTTTCCGCTTTTAGAGAGAGCGCCAAAACCTATTATCTTAAGTCCTCGAGGGATCCGCG-3', que incluía en el extremo 5' un sitio de reconocimiento de la enzima de restricción Sall, y en el extremo 3' y sitios de reconocimiento de enzimas de restricción XhoI, BamHI, respectivamente. Las cadenas de oligonucleótidos positivas y negativas se disolvieron en Tris-HCl 20 mM (pH 8,0) y MgCl<sub>2</sub> 2,5 mM a temperatura ambiente antes de calentarse a 95°C y se enfrió en hielo. El oligonucleótido ahora bicatenario se digirió con Bam HI y Sal I y se clonó en un plásmido pUC19 (Sambrook, J. *et al.*, Molecular Cloning: A Laboratory Manual 2ª edición, Nueva York: Cold Spring Harbor Laboratory 18, 58 (1989)) (que contiene G418 y marcadores de selección de ampicilina) se abrió con Bam HI y Sal I para crear pJOMU. pJOMU se digirió a continuación en dos productos de la digestión por separado, Hind III con Xho I y RI Ecor con Sal I; los fragmentos de cada producto de la digestión se purificaron en gel y ambos se clonaron en pUC19 abierto con Hind III y Eco RI para crear pJOMU2. Las dos últimas etapas se repitieron varias veces dando como resultado la polimerización de Secuencia 1 y 2 creando 64 sitios de unión de baliza molecular en el plásmido pJOMU64. Utilizando cebadores directos e inversos de RCP y el gen IL8 mínimo humano como plantilla (Hitmann *et al.* 1999), se ampliaron tres variantes del promotor de IL8 y se clonaron posteriormente en un plásmido pGFP 5' de una secuencia de codificación de nucleótidos de la proteína verde fluorescente creando por lo tanto tres variantes de pIL8GFP (una variante natural y dos variantes mutantes), los cebadores para esta amplificación fueron cebador directo de IL8: 5'-CACTGAATTAATGAAAGTGTGATGACTCAGGTTTGCC-3' y cebador inverso de IL8 5'-TCAGTGGCTAGCGAAGCTTGTGTGCTCTGCTGTCT-3'. Las fusiones IL8-GFP con una variante natural y dos mutantes, (promotor de IL8 insertado corriente arriba de GFP) se escindieron de pIL8GFP con las enzimas AseI y NotI y se clonaron 5' (corriente arriba) de los 64 sitios de unión de balizas moleculares en pJOMU64 para crear pJOMU-IL8wt, pJOMUIL8mut1 y pJOMUIL8 mut2.

Diseño y síntesis de baliza molecular

Se sintetizaron balizas moleculares en un sintetizador 394 DNA de Applied Biosystems (Foster City, CA) utilizando 2'-O-metil-ribonucleótido β-cianoetil fosforamiditas. Las secuencias de las balizas moleculares eran para MB 1: Cy5-5'-GCUGC AAAAGCGGAAAAGUUUGAAGAGAA GCAGC-3'-BHQ3 complementaria de la secuencia 1 y para MB 2: CY5-5'-CGACC GACUUAAGUAUAGGUUUUGGCG GGUCG-3'-BHQ3 complementaria de la secuencia 2. Para ambas balizas moleculares la secuencia subrayada representaba los extremos complementarios de la baliza molecular. Para la síntesis de las balizas moleculares que tenían un atenuador en su extremo 3', se utilizó una columna de vidrio con poro controlado que contiene Black Hole Quencher 3 (BHQ3). Cada baliza molecular se purificó por cromatografía líquida de alta presión (HPLC) a través de una columna C-18 de fase inversa. Los reactivos de síntesis de ADN se adquirieron en Glen Research (Sterling, VA) y Biosearch Technologies (Novato, CA) y los fluoróforos activados se adquirieron en Molecular Probes (Eugene, OR). Un protocolo detallado para la síntesis de balizas moleculares está disponible en <http://www.molecular-beacons.org>.

Ensayo *in vitro* de enlace baliza molecular-código de barras

La relación señal a fondo de todas las balizas moleculares se midió con un espectrofluorómetro. En primer lugar, se determinó la fluorescencia de referencia de la solución. Posteriormente, la fluorescencia de una solución de 200 µl que contiene balizas moleculares de 30 nM en MgCl<sub>2</sub> 5 mM y Tris-HCl 20 mM (pH 8,0) se determinó a 25°C con un espectrofluorómetro QuantaMaster (Photon Technology International, South Brunswick, NJ) utilizando máxima excitación y longitudes de onda de emisión. Se añadió un exceso molar doble del ARNm transcrito *in vitro* de la diana pJOMU-IL8 y se controló el aumento de fluorescencia hasta que se alcanzó un nivel estable. El aumento de fluorescencia sobre la señal de la baliza molecular sola (sin la adición de ARNm diana) se calculó para determinar la relación señal a fondo y la eficiencia de atenuación de la baliza molecular.

Cultivo celular

Todos los reactivos de cultivos de células se adquirieron en Gibco-Invitrogen. Las células eucariotas se cultivaron en DMEM enriquecido con 10% v/v de suero de ternera fetal (FBS), 50 ui/ml de penicilina, 50 µg/ml de estreptomycin y L-glutamina 2 mM a 37°C, 5% de CO<sub>2</sub>. Toda la microscopía fluorescente de células vivas se realizó en tampón EM (NaCl 120 mM, KCl 7 mM, CaCl<sub>2</sub> 1,8 mM, MgCl<sub>2</sub> 0,8 mM, glucosa 5 mM, Hepes 25 mM a pH 7,3).

Creación de estirpes celulares estables

Se crearon estirpes celulares que expresaban de manera estable los montajes de baliza molecular con 'código de barras' (véase el apartado sobre biología molecular) mediante protocolos estándares (Ref. Maniatis *et al.* 1989). En resumen, 2,5x10<sup>5</sup> células HeLa se transfectaron con montajes de baliza con código de barras (pJOMU-IL8wt y pJOMUIL8mut1), utilizando el reactivo de transfección Eugene (Roche, Suiza) según los protocolos del fabricante. 24 horas después de la transfección, se añadió genética (G418) (Invitrogen, Carlsbad, CA) a las células

a una concentración final de 400 µg/ml. A continuación se mantuvieron durante dos semanas para eliminar las células que no contenían una integración estable de la montaje promotor de baliza molecular con código de barras. Posteriormente, las células se dividieron y diluyeron para permitir el aislamiento de clones de células individuales con plásmidos integrados de forma estable. Después de la selección continua con 400 µg/ml de G418 durante cuatro semanas más, las estirpes celulares obtenidas a partir de clones individuales continuaron en presencia de G418, y se utilizaron para los ensayos funcionales.

#### Cultivo bacteriano

M90T (Sansonetti P.J. *et al.*, Infection and Immunity Mar 1982, págs. 852-860) es la cepa natural invasora de *S. flexneri*. M90Tafal es la cepa natural que expresa la adhesina, afal. BS176afal es un mutante no invasor de M90T curado del plásmido con virulencia y expresa la adhesión, afal. En el caso de ospF, se trataron bacterias con poli-L-lisina antes de la invasión para aumentar la infección bacteriana. Todas las cepas bacterianas utilizadas en este estudio se cultivaron a 37°C en caldo tripsínico de caseína de soja (TCSB) enriquecido con 100 µg/ml de ampicilina.

#### Inyección y transfección de balizas moleculares

Se disolvieron balizas moleculares a una concentración de 2,5 ng/µl, y una solución de aprox 0,1 a 1 fentolitros se microinyectó en cada célula utilizando un aparato de microinyección FemtoJet (Eppendorf). Con objeto de introducir las balizas moleculares en las células vivas por transfección, las células se cultivaron a un 30 por ciento de confluencia y se lavaron con suero exento de Opti-MEM1 (Invitrogen, Carlsbad, CA). El reactivo de transfección Oligofectamina (Invitrogen, Carlsbad, CA) se incubó durante cinco minutos en medio exento de suero (1 µl de reactivo añadido a 9 l de Opti-MEM1) antes de mezclar con balizas moleculares (MB1 y MB2) (1 ng/µl en Opti-MEM1). Las balizas moleculares y el reactivo de transfección se incubaron durante 20 minutos a 25°C para formar complejos entre ellos. Después de diluir estos complejos con 200 µl de medio exento de suero se añadieron a las células. Las células se incubaron durante 3 horas en presencia de estos complejos. Por último, las células se lavaron con medio EM y se detectaron por la imagen.

#### Infección bacteriana

Las células con los indicadores promotor de baliza con código de barras integrados de forma estable se sembraron a una densidad de  $2 \times 10^5$  células por pocillo en placas de cultivo MATTEK de 35 nm con fondo de vidrio (Mattek Cooperación, PA). 24 h antes de la infección, los cultivos de *S. flexneri* requeridos se inocularon en TCSB y se cultivaron durante la noche a 37°C. Durante la noche se inocularon cultivos bacterianos a una dilución 1/100 en TCSB y se cultivaron a una densidad óptica a 600 nm ( $DO_{600}$ ) de ~0,3. Antes de la infección, las bacterias se lavaron dos veces con PBS (solución salina tamponada con fosfato) y se volvieron a poner en suspensión en EM. La baliza molecular inyectada o células transfectadas se lavaron dos veces con PBS y se mantuvieron en EM. Por último, las células se infectaron directamente con bacterias a una MDI (multiplicidad de infección) de 10, y se mantuvieron a 37°C utilizando una cámara de calentamiento. La invasión bacteriana fue seguida de microscopía multidimensional de lapso de tiempo.

#### Microscopía y detección por la imagen de células vivas

La invasión bacteriana se siguió utilizando un microscopio confocal de disco Perkin Elmer Nipkow (Perkin Elmer, UK) con un objetivo de 40x. Para detección por la imagen de células vivas, se utilizaron placas de cultivo MATTEK con fondo de vidrio (Mattek Cooperación, PA) que posee un cubreobjetos de 0,17 mm con una capa de material conductor en su fondo para permitir el calentamiento controlado. Estas placas se recubrieron con gelatina antes de colocar las células en placas. La temperatura de las placas de cultivo MATTEK y el objetivo del microscopio se mantuvieron a 37°C. Durante la microinyección y la detección por la imagen, las células se mantuvieron en un medio EM. La microscopía se realizó en múltiples dimensiones utilizando un microscopio ZEISS (Zeiss, Alemania) equipado con láser de Ar y He/Ne conectado a un disco Nipkow de Perkin Elmer.

#### Análisis de datos

Todos los análisis de datos y la cuantificación se realizó utilizando el programa informático gratuito ImageJ (descarga gratis en <http://rsb.info.nih.gov/ij/>). Para la cuantificación, las imágenes se situaron en el umbral por encima del fondo citosólico de las células que contienen baliza molecular. A continuación, se midió la densidad integrada de la fluorescencia dentro del núcleo celular durante el transcurso de la invasión bacteriana. Los resultados de esta cuantificación se representaron utilizando Excel (Microsoft, Seattle, WA).

#### Referencias

1. Arbibe, L. et al. An injected bacterial effector targets chromatin access for transcription factor NF-kappaB to alter transcription of host genes involved in immune responses. Nat Immunol 8, 47-56 (2007).

2. Monack DM, Mueller A, Falkow S. Persistent bacterial infections: the interface of the pathogen and the host immune system. *Nat Rev Microbiol.* 2004 Sep;2(9):747-65.
- 5 3. Kytтары VC, Juang YT, Tsokos GC. Immune cells and cytokines in systemic lupus erythematosus: an update. *Curr Opin Rheumatol.* 2005 Sep;17(5):518-22
4. Barabasi, A.L. & Oltvai, Z.N. Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5, 101-113 (2004).
- 10 5. Pepperkok, R. & Ellenberg, J. High-throughput fluorescence microscopy for systems biology. *Nat Rev Mol Cell Biol* 7, 690-696 (2006).
6. Bastiaens, P.I. & Pepperkok, R. Observing proteins in their natural habitat: the living cell. *Trends Biochem. Sci.* 25, 631-637 (2000).
- 15 7. Meyer, T. & Teruel, M.N. Fluorescence imaging of signaling networks. *Trends Cell Biol.* 13, 101-106 (2003).
8. Wouters, F.S., Verveer, P.J. & Bastiaens, P.I. Imaging biochemistry inside cells. *Trends Cell Biol.* 11, 203-211 (2001).
- 20 9. Karin, M. Nuclear factor-kappaB in cancer development and progression. *Nature* 441, 431-436 (2006).
10. Karin, M., Lawrence, T. & Nizet, V. Innate immunity gone awry: linking microbial infections to chronic inflammation and cancer. *Cell* 124, 823-835 (2006).
- 25 11. Genovesio, A., Belhassine, Z. & Olivo-Marin, J. Adaptive gating in Gaussian Bayesian multi-target tracking. *Image Processing, 2004. ICIP'04. 2004 International Conference on* 1 (2004).
12. Bork, P. & Serrano, L. Towards cellular systems in 4D. *Cell* 121, 507-509 (2005).
- 30 13. Genovesio, A., Zhang, B. & Olivo-Marin, J. Interacting multiple model based method to track moving fluorescent biological spots. *Biomedical Imaging: Macro to Nano, 2004. IEEE International Symposium on*, 1239-1242 (2004).
14. Olivo-Marin, J. Extraction of spots in biological images using multiscale products. *Pattern Recognition* 35, 1989-1996 (2002).
- 35 15. Starkuviene, V. High-content screening microscopy identifies novel proteins with a putative role in secretory membrane traffic. *Genome Res.* 14, 1948-1956 (2004).
16. Liebel, U. A microscope-based screening platform for large-scale functional protein analysis in intact cells. *FEBS Lett.* 554, 394-398 (2003).
- 40 17. Shav-Tal, Y. et al. Dynamics of single mRNPs in nuclei of living cells. *Science* 304, 1797-1800 (2004).
- 45 18. Mhlanga, M.M., Vargas, D.Y., Fung, C.W., Kramer, F.R. & Tyagi, S. tRNA-linked molecular beacons for imaging mRNAs in the cytoplasm of living cells. *Nucleic Acids Res* 33, 1902-1912 (2005).
19. Bratu, D.P. Molecular beacons: Fluorescent probes for detection of endogenous mRNAs in living cells. *Methods Mol Biol* 319, 1-14 (2006).
- 50 20. Tyagi, S. & Kramer, F.R. Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* 14, 303-308 (1996).
21. Bertrand, E. et al. Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2, 437-445 (1998).
- 55 22. Shav-Tal, Y. et al. Dynamics of single mRNPs in nuclei of living cells. *Science* 304, 1797-1800 (2004).
23. Vargas et al
- 60 24. Bratu, D. P., Cha B.J., Mhlanga, M.M., Kramer, F.R., and Tyagi, S. (2003). Visualizing the distribution and transport of mRNAs in living cells. *Proc Natl Acad Sci USA*
25. Matsuo, T. (1998). In situ visualization of messenger RNA for basic fibroblast growth factor in living cells. *Biochim Biophys Acta* 1379, 178-84.
- 65

26. Sokol, D. L., Zhang, X., Lu, P. & Gewirtz, A. M. (1998). Real time detection of DNA:RNA hybridization in living cells. *Proc Natl Acad Sci U S A* 95, 11538-43.
- 5 27. Southern, E. M., Milner, N. & Mir, K. U. (1997). Discovering antisense reagents by hybridization of RNA to oligonucleotide arrays. *Ciba Found Symp* 209, 38-44; discussion 44-6.
28. Ho, S. P., Bao, Y., Leshner, T., Malhotra, R., Ma, L. Y., Fluharty, S. J. & Sakai, R. R. (1998). Mapping of RNA accessible sites for antisense experiments with oligonucleotide libraries. *Nat Biotechnol* 16, 59-63
- 10 29. Mir, K.U. & Southern, E.M. Determining the influence of structure on hybridization using oligonucleotide arrays. *Nat Biotechnol* 17, 788-792 (1999).
30. Milner, N., Mir, K.U. & Southern, E.M. Selecting effective antisense reagents on combinatorial oligonucleotide arrays. *Nat Biotechnol* 15, 537-541 (1997).
- 15 31. Southern, E.M., Milner, N. & Mir, K.U. Discovering antisense reagents by hybridization of RNA to oligonucleotide arrays. *Ciba Found Symp* 209, 38-44; discussion 44-36 (1997).
32. Sohail, M. et al. Antisense oligonucleotides selected by hybridisation to scanning arrays are effective reagents in vivo. *Nucleic Acids Res* 29, 2041-2051 (2001).
- 20 33. Ooms, M., Verhoef, K., Southern, E., Huthoff, H. & Berkhout, B. Probing alternative foldings of the HIV-1 leader RNA by antisense oligonucleotide scanning arrays. *Nucleic Acids Res* 32, 819-827 (2004).
- 25 34. Nguyen, H.K. & Southern, E.M. Minimising the secondary structure of DNA targets by incorporation of a modified deoxynucleoside: implications for nucleic acid analysis by hybridisation. *Nucleic Acids Res* 28, 3904-3909 (2000).
35. Sohail, M., Doran, G., Riedemann, J., Macaulay, V. & Southern, E.M. A simple and cost-effective method for producing small interfering RNAs with high efficacy. *Nucleic Acids Res* 31, e38 (2003).
- 30 36. Wittwer, C.T., Herrmann, M.G., Moss, A.A. & Rasmussen, R.P. Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechniques* 22, 130-131, 134-138 (1997).
- 35 37. Li, Q., Luan, G., Guo, Q. & Liang, J. A new class of homogeneous nucleic acid probes based on specific displacement hybridization. *Nucleic Acids Res* 30, E5 (2002).
38. Solinas, A. et al. Duplex Scorpion primers in SNP analysis and FRET applications. *Nucleic Acids Res* 29, E96 (2001).
- 40 39. Sambrook, J., Fritsch, E. & Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd. New York: Cold Spring Harbor Laboratory 18, 58 (1989).
- 45 40. Sansonetti P.J. et al - Involvement of a plasmid in the invasive ability of *Shigella flexneri* - *Infection and Immunity*, Mar 1982, p. 852-860.

**Listado de secuencias**

5 <110> INSTITUT PASTEUR

<120> Polinucleótido apto para un ensayo indicador basado en una sola célula para controlar los patrones de expresión génica con una resolución espacio-temporal elevada

10 <130> B7203 -AD/CAL

<160> 14

<170> PatentIn version 3.3

15 <210> 1  
<211> 24  
<212> ADN  
<213> Artificial

20 <220>  
<223> oligonucleótido

<400> 1

25 ttctctcaa actttccgc ttt 24

<210> 2  
<211> 24  
<212> ADN  
<213> Artificial

30 <220>  
<223> oligonucleótido

<400> 2

35 cgccaaaacc tattatctta agtc 24

<210> 3  
<211> 30  
<212> ADN  
<213> Artificial

40 <220>  
<223> oligonucleótido

<400> 3

45 gcgcaaaagc ggaaaaggaa gagaagcagc 30

<210> 4  
<211> 25  
<212> ADN  
<213> Artificial

50 <220>  
<223> oligonucleótido

<400> 4

55 cgaccgacaa gaaaaggggc gggcg 25

<210> 5  
<211> 24  
<212> ADN  
<213> Artificial

60 <220>  
<223> oligonucleótido

65

# ES 2 391 906 T3

	<400> 5 ttcttctcaa actttccgc tttt	24
5	<210> 6 <211> 24 <212> ADN <213> Artificial	
10	<220> <223> oligonucleótido	
	<400> 6 cgccaaaacc tattatctta agtc	24
15	<210> 7 <211> 28 <212> ADN <213> Artificial	
20	<220> <223> oligonucleótido	
	<400> 7 acgcgctgcac ttcttctcaa actttcc	28
25	<210> 8 <211> 53 <212> ADN <213> Artificial	
30	<220> <223> oligonucleótido	
	<400> 8 gcttttagag agagcgccaa aacctattat ctaagtcct cgagggatcc gcg	53
35	<210> 9 <211> 38 <212> ADN <213> Artificial	
40	<220> <223> oligonucleótido	
	<400> 9 cactgaatta atgaaagtgt gatgactcag gtttgccc	38
45	<210> 10 <211> 35 <212> ADN <213> Artificial	
50	<220> <223> oligonucleótido	
	<400> 10 tcagtggtta gcgaagcttg tgtgctctgc tgtct	35
55	<210> 11 <211> 30 <212> ADN <213> Artificial	
60	<220> <223> oligonucleótido	
65	<220> <223> oligonucleótido	

# ES 2 391 906 T3

	<400> 11 gcgcaaaaagc ggaaaaggaa gagaagcagc	30
5	<210> 12 <211> 25 <212> ADN <213> Artificial	
10	<220> <223> oligonucleótido	
	<400> 12 cgaccgacaa gaaaaggggc gggcg	25
15	<210> 13 <211> 21 <212> ADN <213> Artificial	
20	<220> <223> oligonucleótido	
25	<400> 13 ctcacctgct cttctcagac c	21
	<210> 14 <211> 23 <212> ADN <213> Artificial	
30	<220> <223> oligonucleótido	
35	<400> 14 gctatagcac taaggaaga ccc	23

## REIVINDICACIONES

- 5 1. Polinucleótido bicatenario apto para llevar a cabo el ensayo de expresión de la perfilación génica cuando se integra en la célula que aloja de manera natural y expresa posiblemente el gen o genes de interés para dicha perfilación, que comprende en su cadena positiva considerada desde su extremo 5' a su extremo 3', (i) un promotor de un gen de interés o varios promotores de varios genes de interés seleccionados de entre varios genes que son (a) endógenos a la célula y están (b) sometidos a perfilación de la expresión génica, en el que dicho promotor es reconocido por el mecanismo de transcripción interna de la célula y, (ii) uno o varios códigos de barras en el que cada código de barras contiene por lo menos una unidad de código de barras que es un montaje de ADN que comprende repeticiones en tándem de un sitio de unión de reconocimiento de baliza o repeticiones en tándem de por lo menos dos sitios de unión de reconocimiento diferentes de baliza, estando compuesto cada sitio de unión por una secuencia de nucleótidos que se hibrida específicamente con la parte de lazo de una baliza molecular en la que la parte de lazo presenta de 10 a 55 nucleótidos y en el que cada uno de dicho(s) código(s) de barras está(n) bajo el control de por lo menos uno de dicho(s) promotor(es) para la transcripción.
- 10 2. Polinucleótido según la reivindicación 1, en el que en un código de barras, por lo menos 2 de los sitios de unión de reconocimiento son contiguos y/o por lo menos 2 de las unidades de código de barras son contiguos.
- 20 3. Polinucleótido según cualquiera de las reivindicaciones 1 a 2 en el que en un código de barras, por lo menos 2 de los sitios de unión de reconocimiento están separados por un separador y/o por lo menos 2 de las unidades de código de barras están separadas por un separador.
- 25 4. Polinucleótido según cualquiera de las reivindicaciones 1 a 3, en el que el número de sitios de unión de reconocimiento en un código de barras es hasta 500.
- 30 5. Polinucleótido según cualquiera de las reivindicaciones 1 a 4, en el que cada unidad de código de barras comprende por lo menos 3 sitios de unión de reconocimiento diferentes y en el que cada código de barras comprende o consiste en 3 unidades de código de barras.
- 35 6. Polinucleótido según cualquiera de las reivindicaciones 1 a 5, que comprende además corriente abajo de un promotor, controlado por dicho promotor, un ADN que codifica una proteína marcadora, en el que dicho ADN de codificación está dispuesto bajo el control de los elementos reguladores de la expresión, inclusive, bajo el control de dicho promotor.
- 40 7. Polinucleótido según cualquiera de las reivindicaciones 1 a 6, que comprende además una secuencia codificadora de un gen de interés, bajo el control de dicho promotor para la transcripción y la expresión.
- 45 8. Polinucleótido según cualquiera de las reivindicaciones 1 a 7, en el que el promotor de un gen de interés se selecciona de entre el grupo de promotores de genes implicados en la respuesta inmunitaria, especialmente en la respuesta inmunitaria innata, tales como promotores de genes de quimiocina o de citocinas, especialmente promotores de genes de interleucina, o promotores de genes de moléculas de adhesión celular tales como genes ICAM, promotores de gen de interferón, o se selecciona de entre promotores de genes que codifican proteínas asociadas a tumores.
- 50 9. Polinucleótido según cualquiera de las reivindicaciones 1 a 8, en el que las secuencias de los sitios de unión de reconocimiento son secuencias no mamíferas, especialmente no humanas.
10. Polinucleótido según cualquiera de las reivindicaciones 1 a 9, en el que un sitio de unión de reconocimiento presenta una de las siguientes secuencias:  
5' -TTCTCTTCAAAC TTTTCCGCTTTT -3' o,  
5' -CGCCAAAACCTATTATCTTAAGTC -3' o,  
5' -CTCACCTGCTCTTCTCAGACC -3'  
y 5' -GCTATAGCACTAAGGTAAGACCC -3'.
- 55 11. Polinucleótido según cualquiera de las reivindicaciones 1 a 10, en el que cada secuencia de repetición en tándem está enmarcada por uno o varios sitios de restricción.
- 60 12. Polinucleótido según cualquiera de las reivindicaciones 1 a 11, en el que el código de barras contiene una o varias secuencias de nucleótidos o varias repeticiones de las mismas seleccionadas de entre la secuencia de terminación de la transcripción y la secuencia de poli A.
- 65 13. Vector que comprende un polinucleótido según cualquiera de las reivindicaciones 1 a 12, clonado en el mismo.
14. Vector según la reivindicación 13, que es un plásmido, un cósmido, un virus o un bac.

15. Célula viva o estirpe celular, excluyendo una célula o estirpe celular embrionaria humana, que comprende, integrada especialmente de manera estable, uno o varios polinucleótidos según cualquiera de las reivindicaciones 1 a 12.
- 5 16. Célula viva o estirpe celular según la reivindicación 15, que es una célula eucariota, seleccionada especialmente de entre células de mamíferos, especialmente humanos, y se selecciona de entre células diferenciadas, células pluripotentes y células madre, o es una estirpe celular derivada de dichas células, especialmente una estirpe celular de mamífero, en particular una estirpe celular humana.
- 10 17. Célula o estirpe celular según la reivindicación 15 o 16, que resulta de la transfección de una célula o estirpe celular con un polinucleótido según cualquiera de las reivindicaciones 1 a 12.
- 15 18. Conjunto de células o estirpes celulares excluyendo una célula o estirpe celular embrionaria humana en el que cada una de las células comprende, integrados especialmente de manera estable en las mismas, uno o varios polinucleótidos según cualquiera de las reivindicaciones 1 a 12, presentando cada célula o estirpe celular un contenido en dicho(s) polinucleótido(s), diferente de las otras células o estirpes celulares.
- 20 19. Conjunto de polinucleótidos diferentes según cualquiera de las reivindicaciones 1 a 12.
- 25 20. Kit que comprende (i) un polinucleótido según cualquiera de las reivindicaciones 1 a 12, o un conjunto de polinucleótidos diferentes según la reivindicación 18, y (ii) una célula o una estirpe celular o un conjunto de células o estirpes celulares excluyendo una célula o estirpe celular embrionaria humana, adaptadas para la integración de dicho polinucleótido o conjunto de polinucleótidos, y/o una o varias sondas de detección molecular.
- 30 21. Kit que comprende un conjunto de células o estirpes celulares según la reivindicación 18 y reactivos apropiados para mantener las células en condiciones adecuadas para ensayar la actividad de transcripción de los polinucleótidos integrados en las mismas.
- 35 22. Kit según cualquiera de las reivindicaciones 20 o 21, que comprende además una o varias sondas de detección que pueden hibridar con el/los sitio(s) de unión de reconocimiento del (de los) código(s) de barras, resultando dicha(s) sonda(s) de detección adecuadas para la visualización o la medición cuando se hibridan con su secuencia diana.
- 40 23. Kit según la reivindicación 22, en el que la(s) sonda(s) de detección es (son) una o varias baliza(s) molecular(es) que pueden hibridar con el/los sitio(s) de unión de reconocimiento de baliza del (de los) código(s) de barras de baliza, presentando dicha(s) baliza(s) molecular(es) una estructura de polinucleótido en tallo-lazo y que es adecuada para la visualización o la medición cuando se hibrida a su secuencia diana, especialmente de manera reversible.
- 45 24. Kit según la reivindicación 22 o 23, en el que la visualización de la hibridación de la(s) sonda(s) de detección con su diana se obtiene como resultado de la fluorescencia, que se activa cuando la sonda de detección se une a su secuencia diana.
- 50 25. Kit según la reivindicación 23 o 24, en el que la baliza molecular es una estructura de polinucleótido en tallo-lazo en el que la parte de lazo del polinucleótido es la secuencia de la sonda adecuada para hibridarse específicamente con un sitio de unión de la baliza y la parte de tallo se compone de dos brazos formados por secuencias complementarias entre sí, alojando cada una de las secuencias de brazo, unida a su extremo libre que es contrario a la parte de lazo del polinucleótido, uno de un resto fluorescente o de un resto de atenuación no fluorescente en el que dichos restos, cuando están unidos a dichas secuencias de brazo, están suficientemente próximos unos de otros para producir la fluorescencia del resto fluorescente que debe atenuarse por transferencia de energía de resonancia fluorescente, y además dicha parte de lazo del polinucleótido es por lo menos dos veces más larga en los nucleótidos que cada estructura de polinucleótido de brazo.
- 55 26. Kit según la reivindicación 25, en el que el resto fluorescente y el resto atenuador no fluorescente están unidos por enlace covalente a las secuencias de brazo en la estructura de polinucleótido en tallo-lazo.
- 60 27. Kit según cualquiera de las reivindicaciones 24 a 26, en el que el resto fluorescente (fluoróforo) se selecciona de entre el grupo de puntos cuánticos y derivados, familia Alexafluor de colorantes, FAM, TET o CAL FluorGold 540, HEX o JOE, VIC<sup>B</sup>, CAL Fluor Orange 560<sup>A</sup>; Cy3<sup>C</sup> o NED<sup>B</sup>, Quasar 570<sup>A</sup>, Oyster 556<sup>D</sup>; TMR o CAL Fluor Red 590<sup>A</sup>; ROX o LC rojo 610<sup>E</sup>, CAL Fluor Red 610<sup>A</sup>, rojo Texas, o rojo LC 610<sup>E</sup>, CAL Fluor Red 610<sup>A</sup>; LC Rojo 640<sup>F</sup> o CAL Fluor Red 635<sup>A</sup>; Cy5<sup>C</sup> o LC Rojo 670<sup>E</sup>, Quasar 670<sup>A</sup>, Oyster 645<sup>D</sup>, LC rojo 705<sup>E</sup> o Cy5.5<sup>C</sup> o ácido 5-(2'-aminoetil)aminonaftalen-1-sulfónico (EDANS), fluoresceína, antranilamida, cumarina, y quelatos de terbio.
- 65 28. Kit según la reivindicación 27, en el que el atenuador se selecciona en el grupo de DDQ-I<sup>A</sup> (absorción máx. 430 µm), Dabcyl (absorción máx. 475), Eclipse<sup>B</sup> (absorción máx. 530), Iowa Black FQ<sup>C</sup> (absorción máx. 532), BHQ-

1<sup>D</sup> (absorción máx. 534), QSY-7<sup>E</sup> (absorción máx. 571), BHQ-2<sup>D</sup> (absorción máx. 580), DDQ-II<sup>A</sup> (absorción máx. 630), Iowa Black RQ<sup>C</sup> (absorción máx. 645), QSY-21<sup>E</sup> (absorción máx. 660), BHQ-3<sup>D</sup> (absorción máx. 670), oro, metales de tierras raras o ácido 4-(4'-dimetilaminofenilazo)benzoico (DABCYL), rodamina, pirenobutirato, eosina, nitrotirosina, etidio y tetrametil-rodamina.

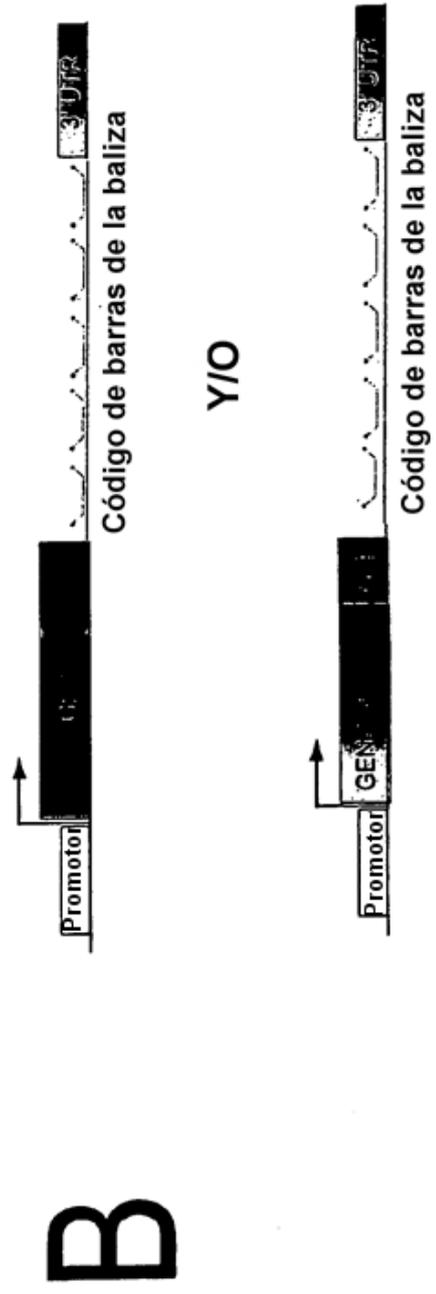
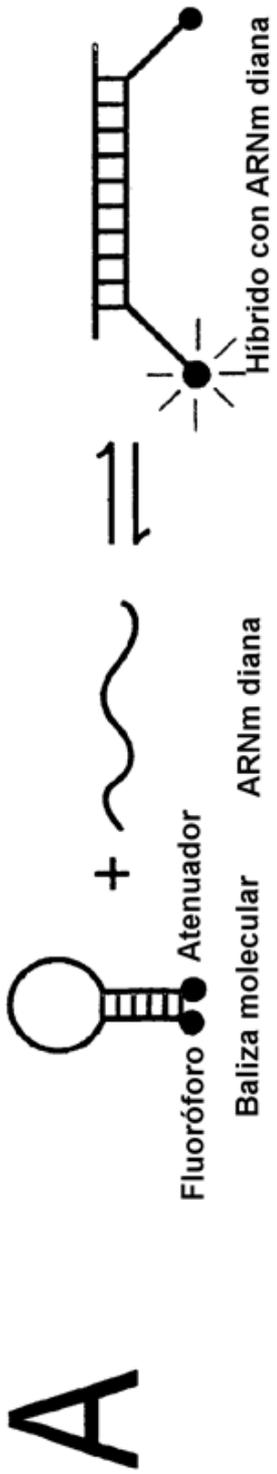
5 29. Kit según cualquiera de las reivindicaciones 23 a 28, en el que la parte de lazo de la baliza molecular presenta de 10 a 55 nucleótidos y cada estructura de polinucleótido del brazo presenta de 4 a 16 nucleótidos.

10 30. Kit según cualquiera de las reivindicaciones 23 a 29, en el que las balizas moleculares presentan una de las siguientes secuencias de nucleótidos:

5'-GCUGC AAAAGCGGAAAAGUUUGAAGAGAA GCAGC-3' o  
5'-CGACC GACUUAAGAUAAUAGGUUUUGGCG GGUCG-3'

15 31. Kit según cualquiera de las reivindicaciones 22 a 30, que comprende además un plásmido con polinucleótido de código de barras clonado corriente abajo del gen o promotor de interés, y/o sondas que pueden reconocer el polinucleótido de código de barras, y/o reactivos tales como péptidos, lípidos, productos químicos que pueden introducir sondas y plásmidos en las estirpes celulares o microinyectarlos dentro de la célula, y/o referencias positivas y negativas para cada etapa en el procedimiento experimental según se necesite.

20 32. Célula o estirpe celular según cualquiera de las reivindicaciones 15 a 17, que comprende además una o varias sondas de detección moleculares como se define en cualquiera de las reivindicaciones 22 a 30.



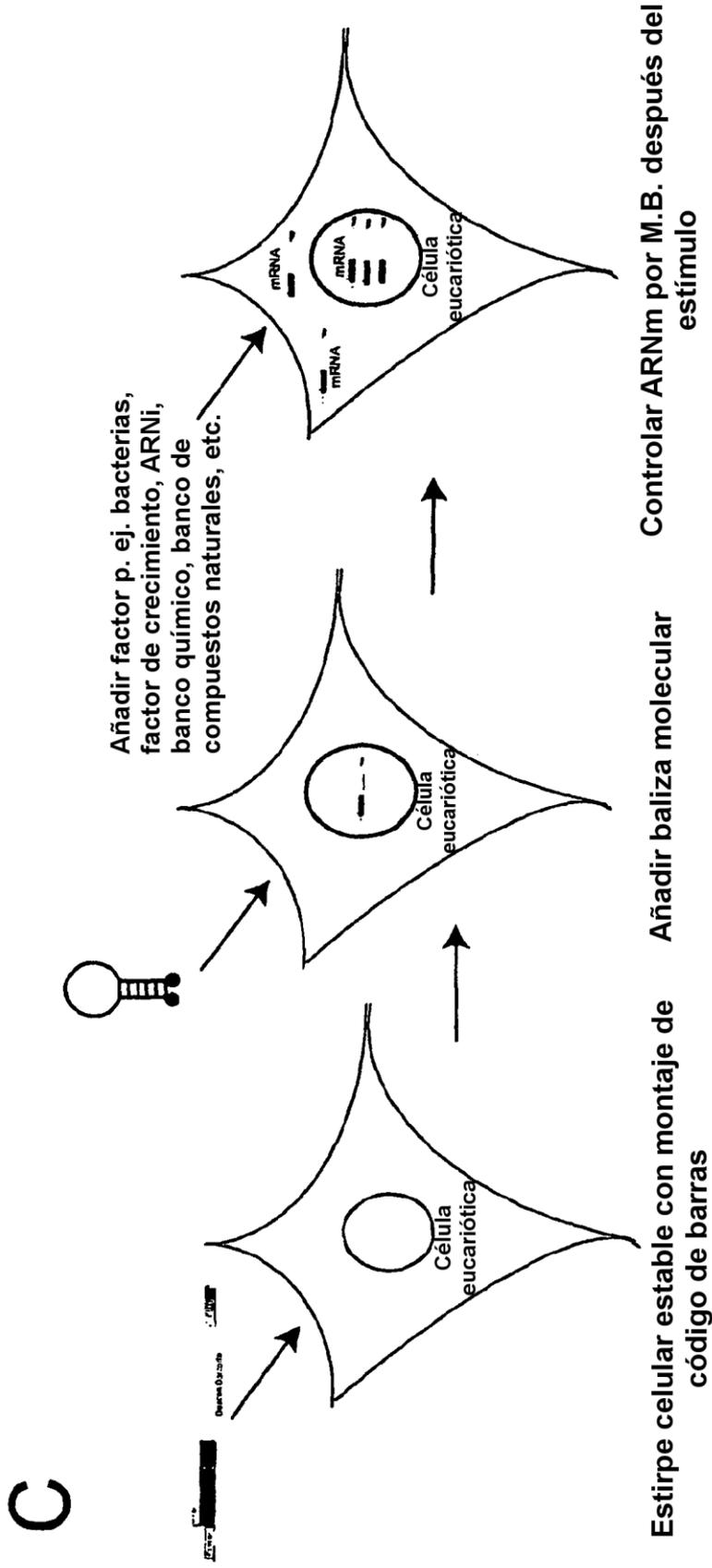
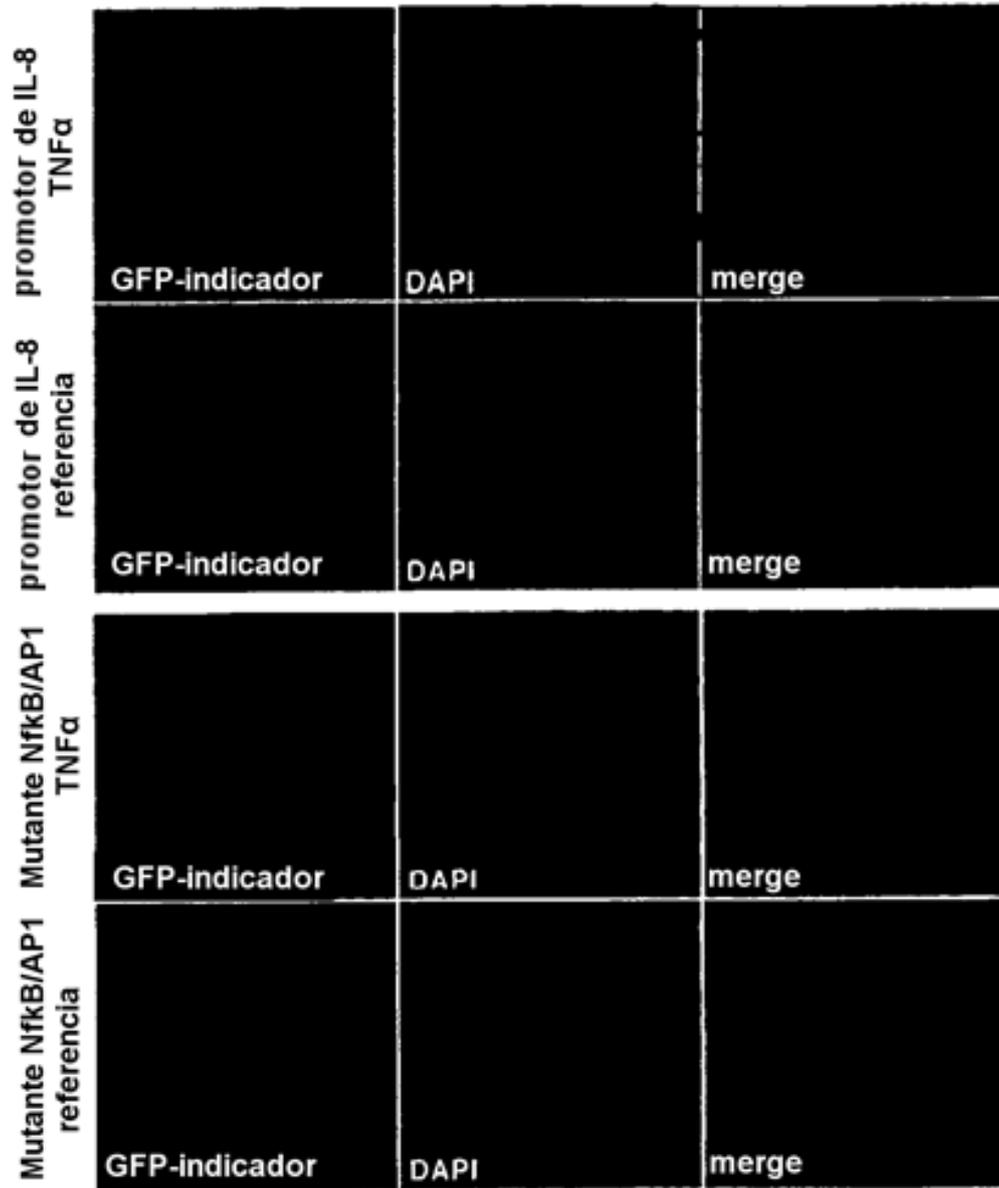


Figura 1C



**Figura 2**

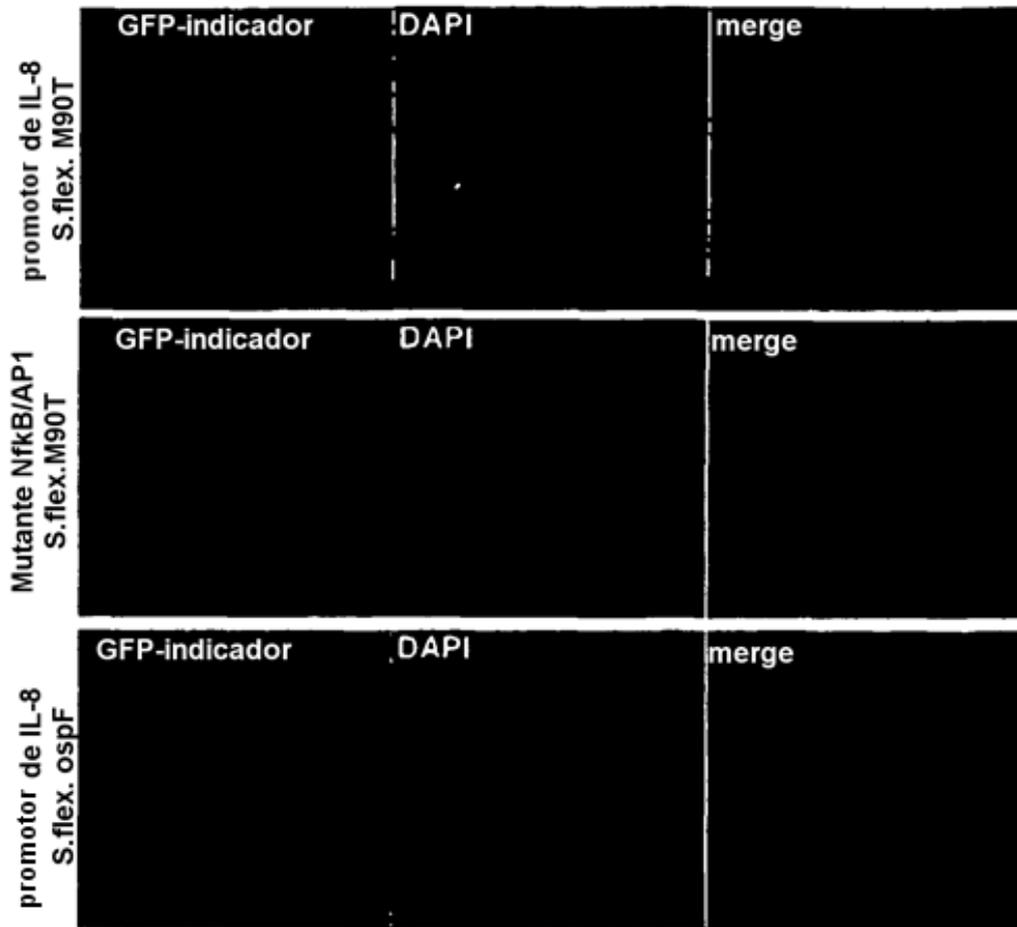
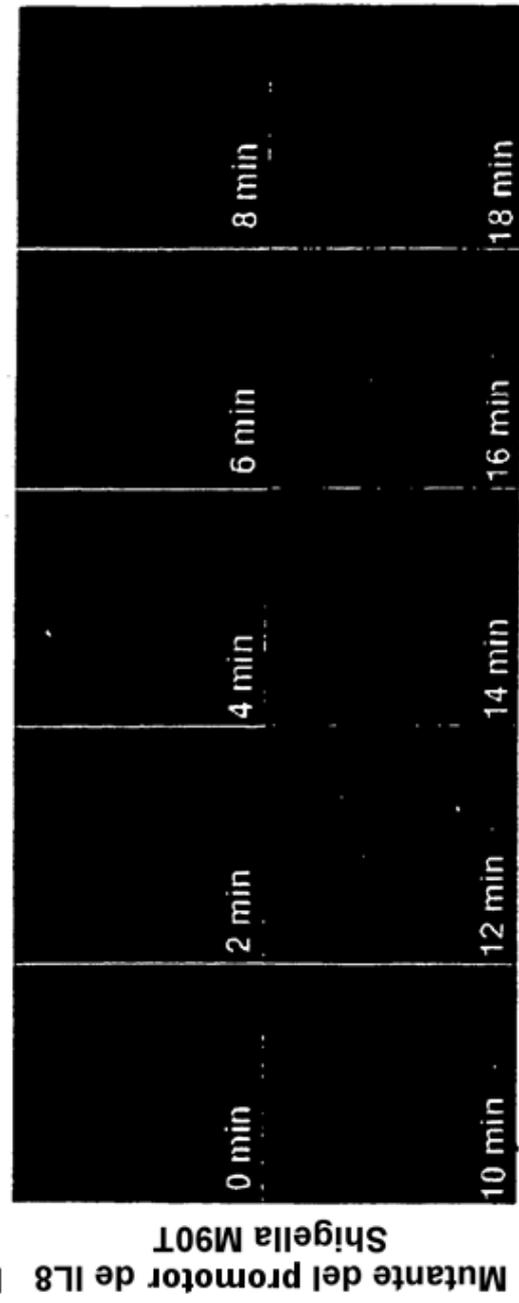
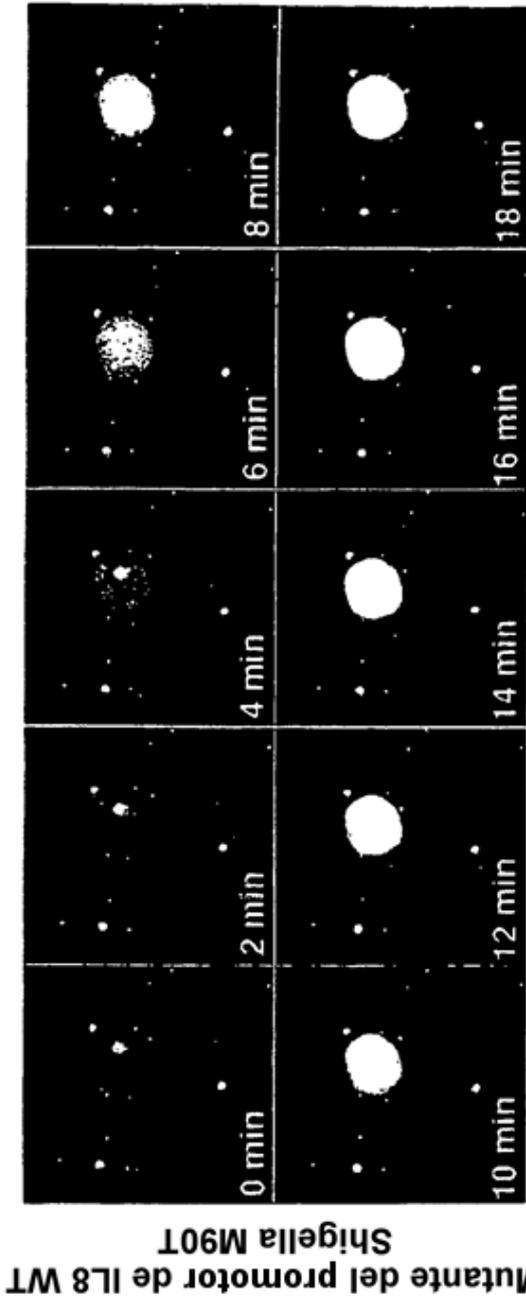


Figura 3



Inducción de la transcripción de Interleucina 8 tras la invasión de Shigella

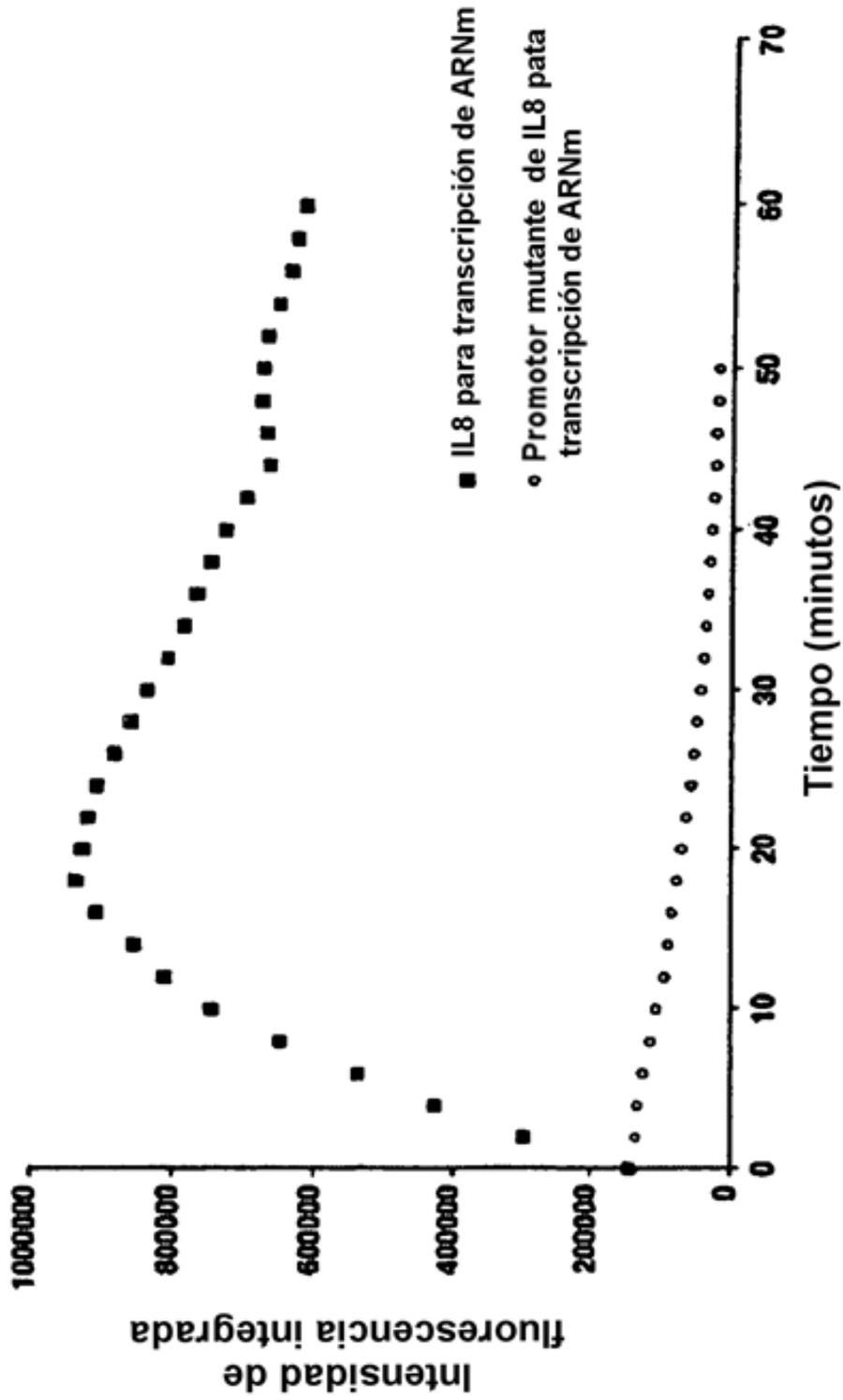


FIGURA 4C

**Esquema de codificación de dos colores**

Gen	Color 1	Color 2	Color 4	Color 5	Color 6
Gen A	●	●			
Gen B	●		●		
Gen C	●		●		
Gen D	●			●	
Gen E	●				●
Gen F		●	●		
Gen G		●	●		
Gen H		●		●	
Gen I		●			●
Gen J			●	●	
Gen K			●	●	
Gen L			●		●
Gen M			●	●	
Gen N			●		●
Gen O				●	●

**Figura 5A(1)**

Utilización de combinaciones de colores para caracterizar genes múltiples en una sola célula

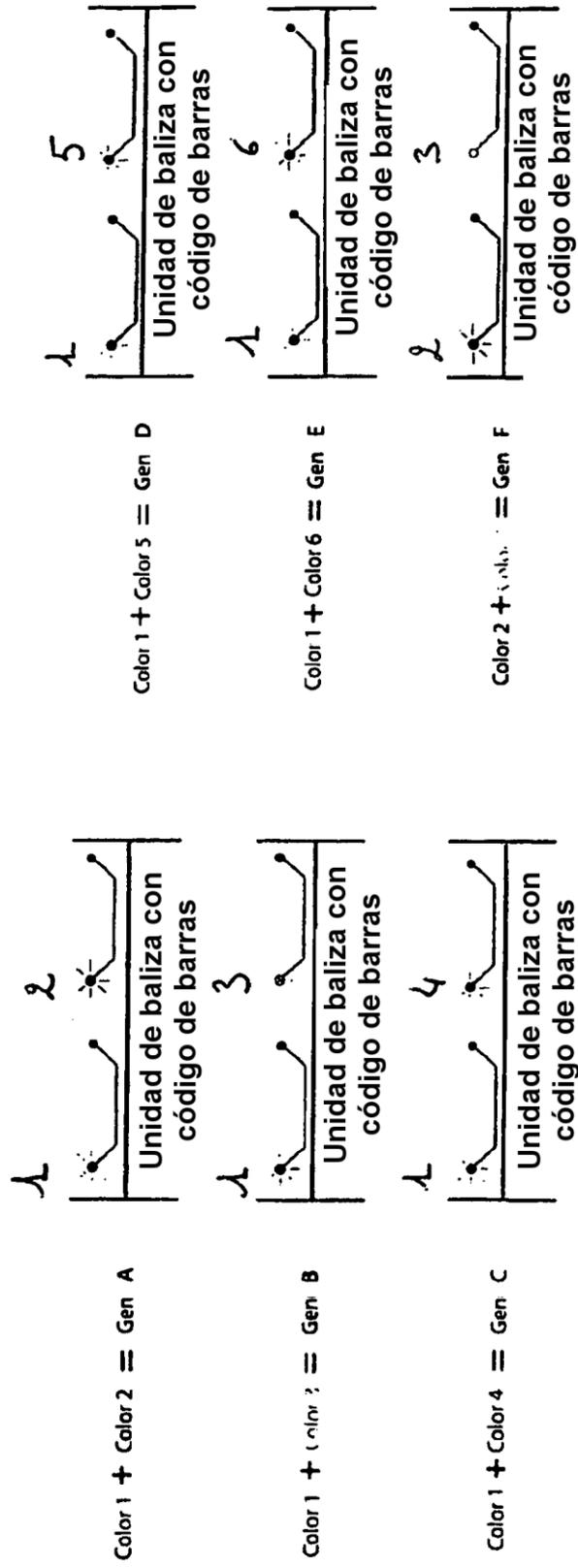


Figura 5A(2)

Esquema de codificación de tres colores

Gen	Color 1	Color 2	Color 4	Color 5	Color 6
Gen A	●	●	⊙		
Gen B	●	●		●	
Gen C	●	●		●	
Gen D	●	●			●
Gen E	●		●	●	
Gen F	●		●	●	
Gen G	●		●		●
Gen H	●		●	●	
Gen I	●		●		●
Gen J	●			●	●
Gen K		●	●	●	
Gen L		●	●	●	
Gen M		●	●		●
Gen N		●	●	●	
Gen O		●	●		●
Gen P		●		●	●
Gen Q			●	●	●
Gen R			●	●	●
Gen S			●	●	●
Gen T			●	●	●

FIGURA 5B

Utilización de combinaciones de colores para caracterizar genes múltiples en una sola célula

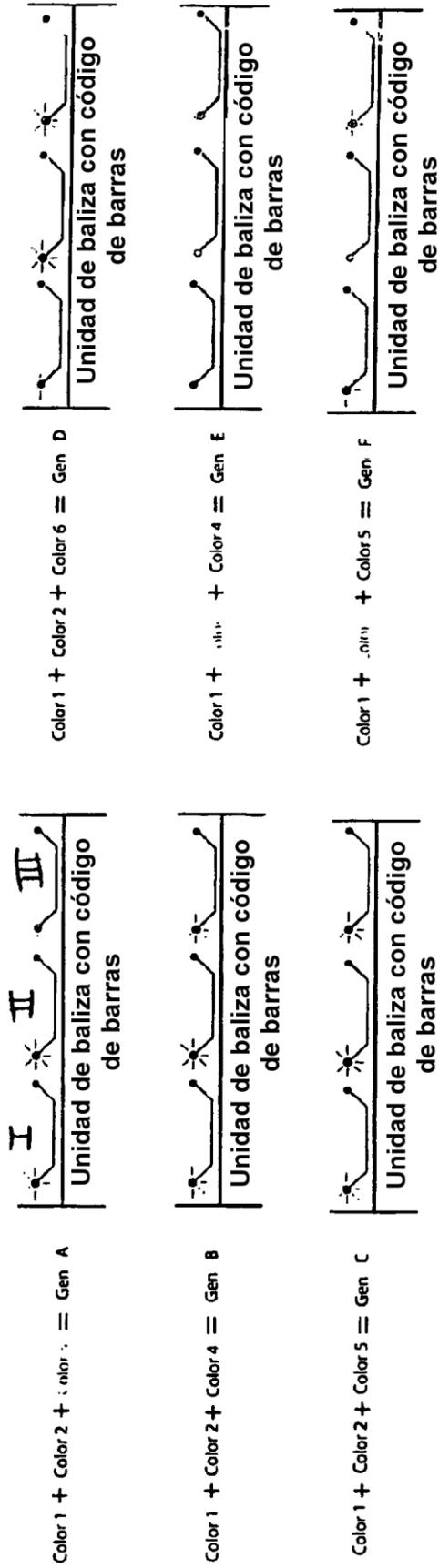


Figura 5c