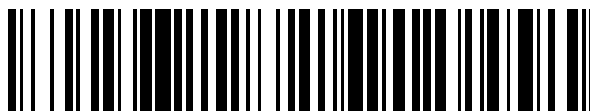


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 768 291**

51 Int. Cl.:

**C12N 15/09** (2006.01)

**C12Q 1/02** (2006.01)

**C12Q 1/68** (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **06.09.2013 PCT/JP2013/074107**

87 Fecha y número de publicación internacional: **21.08.2014 WO14125668**

96 Fecha de presentación y número de la solicitud europea: **06.09.2013 E 13875109 (4)**

97 Fecha y número de publicación de la concesión europea: **06.11.2019 EP 2963113**

54 Título: **Método para aislar una región genómica específica utilizando una molécula que se une específicamente a una secuencia de ADN endógena**

30 Prioridad:

**14.02.2013 JP 2013026310**  
**09.07.2013 JP 2013143282**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**22.06.2020**

73 Titular/es:

**OSAKA UNIVERSITY (100.0%)**  
**1-1 Yamadaoka, Suita-shi**  
**Osaka 565-0871, JP**

72 Inventor/es:

**FUJII, HODAKA y**  
**FUJITA, TOSHITSUGU**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 768 291 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

Método para aislar una región genómica específica utilizando una molécula que se une específicamente a una secuencia de ADN endógena

### Campo de la invención

- 5 La presente invención se refiere a un método para aislar una región genómica específica con el uso de una molécula capaz de unirse específicamente a una secuencia de ADN endógeno y, en particular, a un método para aislar una región genómica específica con el uso de una molécula capaz de unirse específicamente a una secuencia de ADN endógeno, mientras se mantiene la interacción de la región genómica específica y moléculas que interactúan con ella.

### Antecedentes de la técnica

- 10 El análisis bioquímico y biomolecular de los dominios de cromatina es fundamental para comprender los mecanismos moleculares de las funciones del genoma. Sin embargo, la naturaleza bioquímica de los dominios de cromatina es poco conocida. Esto se debe a que los métodos de muestreo que pueden usarse para el análisis bioquímico y biomolecular de los dominios de cromatina son limitados.

- 15 El análisis bioquímico y biomolecular de los dominios de cromatina requiere una muestra que retenga la interacción del ADN genómico y las moléculas que interactúan con el mismo. Los ejemplos conocidos del método para aislar una región genómica específica mientras se mantiene la interacción de la región genómica específica y las moléculas que interactúan con ella incluyen lo siguiente.

#### (i) Inmunoprecipitación de cromatina

- 20 La inmunoprecipitación de la cromatina (en lo sucesivo denominada "ChIP") es un método que implica la inmunoprecipitación de una región genómica específica con un anticuerpo contra una proteína de unión al ADN conocida para aislar la región en cuestión (consúltese la bibliografía no de patentes 1 y 2). Por lo tanto, existe una limitación de que ChIP no se pueda usar sin la información sobre las proteínas de unión al ADN. Además, ChIP tiene dificultades para aislar una única región genómica específica. Esto se debe a que las proteínas de unión al ADN generalmente se unen a una pluralidad de regiones en el ADN genómico y, por lo tanto, el inmunoprecipitado resultante contiene varias regiones genómicas. Además, ChIP no puede aislar una región genómica a la que las proteínas de unión al ADN conocidas no se unan.

#### (ii) Captura de conformación cromosómica (en adelante denominada "3C")

- 30 3C o sus derivados pueden identificar una o varias regiones genómicas que interactúan con una o varias regiones genómicas (véanse las publicaciones no de patente 3 a 5). Sin embargo, es probable que 3C detecte la interacción no fisiológica porque 3C requiere una reacción enzimática con enzimas de restricción o ADN ligasas en presencia de enlaces cruzados, lo cual no es una condición óptima para la reacción enzimática. Además, la presencia de enlaces cruzados provoca una digestión incompleta por las enzimas de restricción y, por lo tanto, las regiones vecinas de la región genómica diana se amplifican inevitablemente por PCR, lo que resulta en un señal de fondo muy alta que dificulta la detección de interacciones no identificadas.

- 35 (iii) Hibridación fluorescente in situ (en adelante denominada "FISH")

FISH, solo o en combinación con inmunofluorescencia, puede detectar la interacción de una región genómica específica con otras regiones genómicas, ARN y proteínas. Sin embargo, este método tiene una baja resolución y no puede detectar ninguna interacción con moléculas no identificadas.

#### (iv) Proteómica de segmentos de cromatina aislados (en lo sucesivo, "PICH")

- 40 PICH es un método para aislar una región genómica específica con el uso de una sonda de ácido nucleico específica, y que se ha demostrado que es capaz de aislar regiones teloméricas que tienen múltiples repeticiones de una secuencia complementaria a la sonda (véase la literatura no de patentes 6). Este método requiere el recocido de una sonda de ácido nucleico y una región genómica diana en presencia de enlaces cruzados, y no se ha demostrado la viabilidad del aislamiento de una región genómica específica de copia única o baja.

- 45 (v) Método de literatura no de patentes 7

- 50 El documento no de patentes 7 describe un intento de aislar una región genómica específica mediante una purificación por afinidad. En este intento, las células utilizadas son células de levadura y las células eucariotas superiores aún no se han probado. Además, el sistema Cre-loxP se utiliza para la escisión de la región genómica específica, pero este procedimiento tiene el riesgo de alterar las estructuras de la cromatina. Además, el tratamiento de reticulación con formaldehído no se puede utilizar y, por lo tanto, las moléculas que interactúan, como las proteínas y el ADN, pueden disociarse de la región genómica específica durante un proceso de purificación.

(vi) iChIP

Este es un método que los presentes inventores desarrollaron previamente (véase el documento patentado 1, y documentos no patentados 8 y 9). Para aislar una región genómica específica que retiene una estructura fisiológica de cromatina, se realizan los siguientes procedimientos.

- 5 (1) Se inserta una secuencia de reconocimiento de una molécula de unión a ADN exógena en la vecindad de una región genómica diana en las células a analizar.
- (2) Las células a analizar están hechas para expresar una molécula de fusión de un dominio de unión al ADN de la molécula de unión al ADN exógena y un marcador reconocible por los anticuerpos existentes u otras proteínas.
- 10 (3) Si fuera necesario, las células se someten a un tratamiento de reticulación con formaldehído, etc. Este procedimiento permite la formación de enlaces cruzados entre la región genómica diana y sus moléculas interactuantes, como proteínas, ARN y ADN. Además, también se forman enlaces cruzados entre la molécula de unión a ADN exógena marcada y la secuencia de reconocimiento insertada.
- (4) Posteriormente, las células se rompen y el ADN reticulado se fragmenta por tratamiento con una o varias desoxirribonucleasas, tal como una enzima de restricción o por ultrasonidos.
- 15 (5) Un complejo que contiene la molécula de unión al ADN exógeno marcada se precipita con un vehículo en el que se inmoviliza un anticuerpo, una proteína u otros que reconocen la marcador.
- (6) Se analizan las moléculas que interactúan con la región genómica diana en el complejo precipitado.

Los problemas de iChIP son los siguientes: se requiere tiempo y esfuerzo para insertar la secuencia de reconocimiento de la molécula de unión al ADN exógeno en la vecindad de la región genómica diana de las células a analizar; y la posibilidad de que la inserción de la secuencia de reconocimiento influya en la interacción de interés no puede eliminarse.

Lista de citas bibliográficas

Documentos de patentes

Documento de patente 1: WO 2011/021684

25 Documentos no de patentes

Documento no de patente 1:

Solomon, M. J., Larsen, P. L., Varshaysky, A., Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* (1988) 53, 937-947

Documento no de patente 2:

30 Solomon, M.J., Varshavsky, A., Formaldehyde-mediated DNA-protein crosslinking: a probe for in vivo chromatin structures. *Proc. Natl. Acad. Sci. USA* (1985) 82, 6470-6474

Documento no de patente 3:

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. Capturing chromosome conformation. *Science* (2002) 295, 1306-1311

35 Documento no de patente 4:

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* (2006) 38, 1348-1354

Documento no de patente 5:

40 Simonis, M., Kooren, J., and de Laat, W. An evaluation of 3C-based methods to capture DNA interactions. *Nat. Methods* (2007) 4, 895-901

Documento no de patente 6:

Dejardin, J., and Kingston, R.E. Purification of proteins associated with specific genomic loci. *Cell* (2009) 136, 175-186

Documento no de patente 7:

45 Griesenbeck, J., Boeger, H., Strattan, J.S., and Kornberg, R. D. Affinity purification of specific chromatin segments

from chromosomal loci in yeast. *Mol. Cell Biol.* (2003) 23, 9275-9282

Documento no de patente 8:

Hoshino, A., and Fujii, H. Insertional chromatin immunoprecipitation: a method for isolating specific genomic regions. *J. Biosci. Bioeng.* (2009) 108, 446-449

5 Documento no de patente 9:

Fujita, H., and Fujii, H. Direct identification of insulator components by insertional chromatin immunoprecipitation. *PLoS One* (2011) 6, e26109.

### Compendio de la invención

Problema técnico

10 Un objeto de la presente invención es proporcionar un método para aislar específicamente cualquier región genómica mientras se mantiene la interacción de la región genómica y las moléculas que interactúan con la misma, caracterizándose el método por el uso de una secuencia de ADN endógeno presente dentro o en las proximidades de una región genómica diana en las células a analizar y sin tener la necesidad de insertar una secuencia de reconocimiento de una molécula de unión a ADN exógena en la vecindad de la región genómica diana en las células a analizar.

Solución al problema

La presente invención incluye lo siguiente para lograr el objeto mencionado anteriormente.

(1) Un método para aislar una región genómica específica mientras se mantiene la interacción de la región genómica específica y las moléculas que interactúan con la misma, el método comprende los siguientes pasos (A) y (B):

20 (A) fragmentar el ADN genómico en un estado en el que la interacción del ADN genómico y las moléculas que interactúan con él se mantiene, y

(B) poner el ADN genómico en contacto con una molécula exógena capaz de unirse a una secuencia de ADN endógeno específica en el ADN genómico.

25 (2) El método según el anterior (1), que comprende además la etapa de (C) realizar un tratamiento para mantener la interacción del ADN genómico y las moléculas que interactúan con el mismo.

(3) El método de acuerdo con los puntos (1) o (2) anteriores, en el que la molécula exógena es una proteína de unión al ADN exógena, un ácido nucleico exógeno o un complejo de proteína-ácido nucleico exógeno.

(4) El método de acuerdo con los puntos (1) o (2) anteriores, en el que la molécula exógena es una proteína de dedo de zinc, una proteína efectora TAL o un complejo de una proteína Cas9 inactiva y un ARN guía (ARNg).

30 (5) El método de acuerdo con alguno cualquiera de los anteriores (1) a (4), en el que el método comprende los siguientes pasos 1 a 3:

paso 1: poner el ADN genómico en las células en contacto con la molécula exógena,

paso 2: fragmentar el ADN genómico en las células en un estado donde se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo, y

35 paso 3: permitir que un fragmento de ADN genómico unido a la molécula exógena forme un complejo con una molécula capaz de unirse específicamente a la molécula exógena, y luego recuperar el complejo.

(6) El método según el anterior (5), en el que la molécula exógena es una molécula de fusión que contiene uno o más tipos de secuencias marcadoras.

40 (7) El método de acuerdo con los puntos (5) o (6) anteriores, en el que la molécula exógena es una molécula de fusión que tiene una señal o varias señales de localización nuclear.

(8) El método de acuerdo con uno cualquiera de los anteriores (5) a (7), en el que, en el paso 1, se introduce un gen o varios que codifican la molécula exógena en las células a analizar para realizar la expresión del gen en las células.

(9) El método de acuerdo con alguno cualquiera de los anteriores (1) a (4), en el que el método comprende los siguientes pasos I a III:

45 paso I: fragmentación del ADN genómico en un estado donde se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo,

paso II: poner el ADN genómico fragmentado en contacto con la molécula exógena, y

paso III: recolectar un fragmento de ADN genómico unido a la molécula exógena.

(10) El método según el anterior (9), en el que, en la etapa II, el ADN genómico fragmentado se pone en contacto con la molécula exógena inmovilizada en un vehículo.

5 (11) El método según el anterior (9), en el que, en la etapa II, la molécula exógena se pone en contacto con el ADN genómico fragmentado y, posteriormente, se inmoviliza en un vehículo.

Efectos ventajosos de la invención

De acuerdo con la presente invención, con el uso de una secuencia de ADN endógeno presente dentro o en las proximidades de una región genómica diana en las células a analizar, cualquier región genómica puede aislarse específicamente en un estado donde la interacción de la región genómica y las moléculas interactuando con ella se mantiene. El método de la presente invención para aislar una región genómica específica usa una secuencia de ADN endógeno presente dentro o cerca de una región genómica diana en las células a analizar y no necesita la inserción de una secuencia de reconocimiento de una molécula de unión a ADN exógena y, por lo tanto, puede eliminar la influencia de una secuencia de inserción en la interacción de interés. En consecuencia, la región genómica aislada por el método de la presente invención es muy útil como muestra para el análisis de los mecanismos moleculares de las funciones del genoma.

### Breve descripción de los dibujos

Fig. 1 muestra la primera realización del método de la presente invención para aislar una región genómica específica.

Fig. 2 muestra la segunda realización del método de la presente invención para aislar una región genómica específica.

Fig. 3 muestra el análisis de citometría de flujo de los niveles de expresión de las proteínas indicadas en las células que expresan 3 × FN-Tel-TAL, que es una proteína de fusión de una proteína efectora TAL capaz de unirse a una secuencia telomérica (Telomere-TAL), un 3 × Marcador FLAG y una señal de localización nuclear y en células que expresan una proteína de control (3 × FNLDD).

Fig. 4 muestra el aislamiento exitoso del ADN genómico de la región telomérica usando 3 × FN-Tel-TAL en el análisis de transferencia Southern usando una sonda contra la secuencia telomérica. (A) y (B) muestran la reproducibilidad de los resultados.

Fig. 5 muestra la detección basada en RT-PCR de un ARN de telomerasa contenido en la región telomérica aislada.

Fig. 6 muestra el análisis de inmunotransferencia de los niveles de expresión de las proteínas indicadas en células que expresan una proteína efectora TAL capaz de unirse a la secuencia de ADN de la región promotora del gen Sox2 humano (3 × FN-Sox2-TAL) y en dos tipos de células que expresan diferentes tipos de proteínas de control.

Fig. 7 muestra una vista esquemática de la región promotora del gen Sox2 humano. Las flechas representan los cebadores utilizados para la PCR en tiempo real.

Fig. 8 muestra los resultados de la PCR en tiempo real y demuestra el aislamiento exitoso del ADN genómico de la región promotora del gen Sox2 humano utilizando una proteína efectora TAL capaz de unirse a la secuencia de ADN de la región promotora del gen Sox2 humano (3 × FN-Sox2-TAL).

Fig. 9 muestra el esquema de una realización del método de la presente invención para aislar una región genómica específica (un método para aislar una región genómica específica con el uso de un complejo de una proteína Cas9 inactiva y un gRNA (sistema CRISPR)).

Fig. 10 muestra el análisis de inmunotransferencia de los niveles de expresión de las proteínas indicadas en células que expresan una proteína dCas9 fusionada con un marcador 3 × FLAG (3 × F-dCas9) y en células que expresan una proteína de control (3 × FNLDD).

Fig. 11 muestra los resultados de la PCR en tiempo real y demuestra el aislamiento exitoso del ADN genómico del locus IRF-1 humano usando un gARN correspondiente a una secuencia presente en el locus IRF-1 humano (gRNA-hIRF-1 # 12) y un Proteína dCas9 fusionada con un marcador 3 × FLAG (3 × F-dCas9).

Fig. 12 muestra que una muestra destinada a la identificación basada en espectrometría de masas de las proteínas contenidas en la región aislada de IRF-1 se sometió a electroforesis en un gel SDS-PAGE hasta que la muestra migró aproximadamente 1,0 cm en el gel y, después de tefirse con azul brillante de Coomassie, se dividió en cinco fracciones.

50

**Descripción de las realizaciones**

Método para aislar regiones genómicas específicas

El método de la presente invención para aislar una región genómica específica (en lo sucesivo denominado "el método de la presente invención") comprende los siguientes pasos (A) y (B):

5 (A) fragmentar el ADN genómico en un estado donde se mantiene la interacción del ADN genómico y las moléculas que interactúan con él, y

(B) poner el ADN genómico en contacto con una molécula exógena capaz de unirse a una secuencia de ADN endógeno específica en el ADN genómico. El orden de los pasos (A) y (B) no está limitado, y el paso (A) puede preceder o seguir al paso (B).

10 El método de la presente invención comprende preferiblemente además la siguiente etapa (C):

(C) realizar un tratamiento para mantener la interacción del ADN genómico y las moléculas que interactúan con el mismo.

El paso (C) preferiblemente precede al paso (A).

15 La molécula exógena no está particularmente limitada siempre que sea una molécula exógena capaz de unirse a una secuencia de ADN endógeno específica en el ADN genómico. Por ejemplo, se pueden usar preferiblemente proteínas de unión a ADN exógenas, ácidos nucleicos exógenos, complejos de proteínas-ácido nucleico exógenos, etc. La proteína de unión al ADN exógena es preferiblemente una proteína cuya secuencia de aminoácidos se puede diseñar en asociación con la secuencia de nucleótidos de un ADN diana. Los ejemplos de dicha proteína de unión a ADN exógena incluyen proteínas de dedo de zinc y proteínas efectoras TAL. Los ejemplos del ácido nucleico exógeno incluyen aptámeros de ADN y aptámeros de ARN. Los ejemplos del complejo exógeno de proteína-ácido nucleico incluyen un complejo de una proteína Cas9 inactiva y un ARN guía (gARN).

20 Se sabe que las proteínas de dedo de zinc se pueden modificar para reconocer la secuencia de nucleótidos de un ADN específico (por ejemplo, véase Beerli et al., (2002) *Nature Biotechnol.* 20: 135-141; Pabo y col., (2001) *Ann. Rev. Biochem.* 70: 313-340; Isalan et al., (2001) *Nature Biotechnol.* 19: 656-660; Segal y col., (2001) *Curr. Opin. Biotechnol.* 12: 632-637; y Choo et al., (2000) *Curr. Opin. Struct. Biol.* 10: 411-416). Dicha modificación puede impartir proteínas de dedos de zinc con una nueva especificidad de unión que las proteínas de dedos de zinc nativas no tienen. El método de modificación no está particularmente limitado y, por ejemplo, se puede emplear un método que implique combinar módulos de dedos de zinc que tengan cierta especificidad conocida.

25 La especificidad se puede seleccionar mediante visualización de fagos, sistema de dos híbridos o similares, y dichos métodos de selección se describen en la patente de EE.UU. N° 5.789.538; Pat. de EE.UU. N° 5.925.523; Pat. de EE.UU. N° 6.007.988; Pat. de EE.UU. N° 6.013.453; Pat. de EE.UU. N° 6.410.248; Pat. de EE.UU. N° 6.140.466; Pat. de EE.UU. N° 6.200.759; Pat. de EE.UU. N° 6.242.568; documentos WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197; Patente de Gran Bretaña N° 2.338.237; etc.

30 Las proteínas efectoras TAL son proteínas aisladas de la bacteria *Xanthomonas*, que se conocen como patógenos de plantas. El dominio de unión al ADN de las proteínas efectoras TAL tiene una estructura de repetición en tándem casi completa en la que cada unidad de repetición consta de aproximadamente 34 aminoácidos. Los aminoácidos 12 y 13 de cada unidad de repetición son hipervariables y se denominan diresiduo variable de repetición (RVD). El RVD determina la especificidad de reconocimiento de base de ADN de cada unidad repetida. Para la preparación de una proteína efectora TAL que reconoce una secuencia de nucleótidos específica, las unidades repetidas que tienen cada una un RVD deseado se combinan adecuadamente (véase Moscou y Bogdanove, (2009) *Science* 326: 1501; Boch et al., (2009) *Science* 326: 1509-1512 y Miller et al., (2011) *Nat. Biotechnol.*, 29: 143-148).

35 El ácido nucleico exógeno puede obtenerse, por ejemplo, mediante la purificación por afinidad de una especie molecular capaz de unirse a una secuencia de ADN genómico diana a partir de una biblioteca aleatoria de secuencias de ácido nucleico, seguido del enriquecimiento y la selección mediante el uso de PCR, etc. En dicho ácido nucleico seleccionado, se puede introducir una mutación adecuada para aumentar la afinidad por la secuencia de ADN genómico diana.

40 El complejo exógeno de proteína-ácido nucleico es preferiblemente un complejo de una proteína Cas9 inactiva y un ARNg. La proteína Cas9 es una enzima bacteriana y arqueal que escinde una secuencia específica de un ARNg de manera guiada por ARN (sistema CRISPR). La proteína Cas9 inactiva es un mutante Cas9 que carece de actividad de escisión del ADN pero que es capaz de unirse a una secuencia de ADN complementaria a un ARNg de manera guiada por ARN (Jinek et al., (2012) *Science*, 337: 816-821). Por esta razón, un complejo de Cas9 inactivo y un ARN que contiene una secuencia complementaria a una secuencia de ADN diana es adecuado como molécula exógena utilizada para el método de la presente invención.

45 La secuencia de ADN endógeno específica en el ADN genómico como el sitio de unión de la molécula exógena puede

- 5 seleccionarse a partir de secuencias de ADN que están presentes dentro o en la vecindad de una región genómica a aislar. Es preferible seleccionar una secuencia de ADN que esté presente solo dentro o en la vecindad de la región genómica a aislar y que no esté presente en ninguna otra región del ADN genómico, sino que también se puede seleccionar una secuencia de ADN que esté presente dentro o en la propia vecindad de la región genómica a aislar y que esté presente con poca frecuencia en las otras regiones del ADN genómico.
- 10 En el paso (B), el método para poner el ADN genómico en contacto con la molécula exógena no está particularmente limitado. Tal contacto puede efectuarse en los núcleos de las células. Alternativamente, el ADN genómico en un estado donde se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo puede extraerse de las células antes del contacto. Además, se puede seleccionar una pluralidad de secuencias de ADN endógeno que están presentes dentro o en las proximidades de la región genómica a aislar y ponerlas en contacto con una pluralidad de moléculas exógenas que se unan a las secuencias de ADN correspondientes. Después del contacto con una pluralidad de moléculas exógenas, la implementación de la purificación sucesiva permite un aislamiento más específico de la región genómica diana.
- 15 La fragmentación del ADN genómico en la etapa (A) está destinada a producir un fragmento de ADN genómico de un tamaño apropiado que puede contener una región genómica de interés. Los ejemplos del método para la fragmentación del ADN genómico incluyen la digestión del ADN genómico con una o varias enzimas de restricción, la degradación parcial (escisión parcial) del ADN genómico con una o varias desoxirribonucleasas (endonucleasas) y la escisión física del ADN genómico por ultrasonidos.
- 20 En el caso de que se use una o varias enzimas de restricción, será preferible seleccionar una o varias enzimas de restricción cuyo sitio de escisión no esté dentro de la región genómica para aislarlo y ubicarlo lo más cerca posible de la región genómica de interés.
- 25 En el caso en el que se use una o varias desoxirribonucleasa (endonucleasa), será preferible determinar de antemano las condiciones de reacción para la producción de un fragmento ligeramente mayor que el tamaño de la región genómica a aislar.
- 30 Durante la fragmentación del ADN genómico, es preferible que se mantenga la interacción del ADN genómico y las moléculas que interactúan con el mismo. Por lo tanto, es preferible que, antes de la fragmentación del ADN genómico, se realice un tratamiento para mantener la interacción del ADN genómico y las moléculas que interactúan con él (paso (C)) según sea necesario. El tratamiento para mantener la interacción no está particularmente limitado, pero preferiblemente, la interacción mantenida como resultado del tratamiento puede eliminarse según sea necesario para que las moléculas que interactúan puedan separarse y purificarse para su posterior análisis. Los ejemplos preferibles del tratamiento incluyen la reticulación. Los ejemplos preferibles de un agente usado para la reticulación incluyen formaldehído.
- 35 Primera realización
- 40 El método de la presente invención comprende los pasos 1 a 3 que se muestran a continuación. El método de la presente invención permite el aislamiento de una región genómica específica en un estado en el que se mantiene la interacción de la región genómica específica y las moléculas que interactúan con la misma. Siempre que una región genómica específica pueda aislarse en un estado en el que se mantiene la interacción, el método puede comprender uno o más pasos distintos de los pasos 1 a 3, y los detalles de los pasos no están limitados. En la primera realización, las células a analizar son preferiblemente células viables.
- 45 Paso 1: poner en contacto el ADN genómico en las células con una molécula exógena capaz de unirse a una secuencia de ADN endógeno específica en el ADN genómico.
- Paso 2: fragmentar el ADN genómico en las células en un estado donde se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo.
- Paso 3: permitir que un fragmento de ADN genómico unido a la molécula exógena forme un complejo con una molécula capaz de unirse específicamente a la molécula exógena, y luego recuperar el complejo.
- 50 En lo sucesivo en este documento, el método de la presente invención se describirá en detalle en el orden de los siguientes pasos.
- (1) Paso 1
- El paso 1 es un paso para poner el ADN genómico en las células en contacto con una molécula exógena capaz de unirse a una secuencia de ADN endógeno específica en el ADN genómico.
- En la primera realización del método de la presente invención, la molécula exógena es preferiblemente una molécula de fusión que contiene uno o más tipos de secuencias marcadoras. En el caso de que la molécula exógena contenga

una secuencia o varias secuencias marcadoras, las moléculas capaces de unirse específicamente a la molécula exógena están fácilmente disponibles. La secuencia marcadora no está particularmente limitada y puede seleccionarse según sea apropiado de secuencias marcadoras conocidas. Los ejemplos específicos de la secuencia marcadora incluyen un marcador histidina, un marcador FLAG, un marcador Strep, un péptido de unión a calmodulina, un marcador GST, un marcador MBP, un marcador Halo y un marcador HA.

La molécula de fusión de la molécula exógena y una secuencia marcadora pueden comprender dos o más tipos de secuencias marcadoras. Los dos o más tipos de secuencias marcadoras pueden estar directamente unidas, o unidas mediante una secuencia de reconocimiento de proteasa y / o una secuencia espaciadora. En el caso de que se emplee dicha estructura, una región genómica unida a la molécula de fusión se puede purificar en dos o más pasos y, por lo tanto, la línea de fondo debido a la unión inespecífica se puede reducir significativamente y la eficiencia del aislamiento específico de regiones genómicas específicas es mejorado significativamente.

La secuencia de reconocimiento de proteasas para unir las secuencias marcadoras se puede seleccionar, según sea apropiado, de las secuencias de reconocimiento de proteasas usadas habitualmente para la eliminación de secuencias marcadoras. Conocida como la proteasa utilizada para la eliminación de secuencias marcadoras, se encuentran, por ejemplo, la proteasa del virus del ataque químico del tabaco (TEV), la proteasa del rinovirus humano (HRV), la enteroquinasa (EK), la trombina (Tb) y el factor Xa.

En el caso en el que las células a analizar sean células eucariotas, la molécula de fusión contiene preferiblemente una señal o varias señales de localización nuclear. La molécula de fusión que tiene una señal o varias señales de localización nuclear puede entrar en el núcleo y entrar en contacto con el ADN genómico. La señal de localización nuclear se puede seleccionar, según corresponda, de las señales de localización nuclear conocidas. Los ejemplos específicos de la señal de localización nuclear incluyen la señal de localización nuclear del antígeno T SV40, la señal de localización nuclear de c-Myc, la señal de localización nuclear de p53 y la señal de localización nuclear de NF-κB p50. Alternativamente, la molécula de fusión puede tener la secuencia de aminoácidos de un motivo permeable a las células (dominio de transducción de proteínas).

En el caso de que la molécula exógena sea una proteína, la molécula exógena se podrá producir, por ejemplo, como una proteína recombinante mediante técnicas recombinantes conocidas, es decir, insertando un ADN que codifique la proteína exógena deseada en un vector de expresión conocido, introduciendo el vector de expresión resultante en células huésped apropiadas para la expresión de la proteína, y purificando la proteína expresada por un método conocido. En el caso en el que la proteína exógena sea una proteína de fusión, los ADN cada uno de los cuales codifican una proteína constituyente diferente de la proteína de fusión se fusionarán adecuadamente, y el ADN de fusión resultante se insertará en un vector de expresión en el curso de la producción de la molécula exógena.

En el caso en el que la molécula exógena sea un ácido nucleico, el ácido nucleico puede producirse, por ejemplo, mediante métodos conocidos como síntesis química, PCR, transcripción in vitro, transcripción inversa, transcripción de genes integrados en células, etc. En el caso en el que el ácido nucleico exógeno sea una molécula de fusión, puede ser producido por una síntesis química conocida.

El método para poner el ADN genómico en las células en contacto con la molécula exógena no está particularmente limitado. Por ejemplo, una proteína exógena o ácido nucleico obtenido mediante técnicas recombinantes conocidas, etc., tal como se describió anteriormente, se introduce por microinyección, electroporación, lipofección, etc. en las células a analizar. Alternativamente, una molécula de fusión con la secuencia de aminoácidos de un motivo permeable a las células (dominio de transducción de proteínas) como la que se mencionó anteriormente puede usarse como la molécula exógena e introducirse directamente en las células a analizar. Alternativamente, aparte de la molécula exógena, se puede preparar un dominio de transducción de proteínas fusionado a una secuencia de aminoácidos capaz de unirse a la molécula exógena y luego introducirlo como una mezcla con la molécula exógena en las células a analizar.

Otra opción es expresar la molécula exógena en las células. Con este fin, las células a analizar se usan como una célula huésped, y se prepara un vector de expresión que puede expresar la molécula exógena en las células y luego introducirse en las células. En el caso en el que el huésped sea una célula procariota, la molécula exógena expresada puede entrar en contacto con el ADN genómico en el citoplasma. En el caso en el que el huésped sea una célula eucariota, se usa una molécula de fusión con una señal o varias señales de localización nuclear como la molécula exógena a expresar, de modo que la molécula exógena expresada pueda entrar en el núcleo desde el citoplasma y ponerse en contacto con el ADN genómico. Alternativamente, un organismo transgénico creado mediante la introducción de una forma expresable de un gen que codifica la molécula exógena también se puede utilizar para lograr el contacto de la molécula exógena con el ADN genómico en las células.

La molécula exógena en contacto con el ADN genómico puede unirse a una secuencia de ADN endógeno diana en el ADN genómico.

## (2) Paso 2

El paso 2 es un paso para fragmentar el ADN genómico en las células en un estado en el que se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo. Es preferible que el paso 2 esté precedido por un



paso para realizar un tratamiento para mantener la interacción del ADN genómico y las moléculas que interactúan con el mismo, aunque dicho paso de tratamiento puede omitirse.

5 Con el ADN genómico, las proteínas que se unen al ADN, como, p. ej. los factores de transcripción, los ácidos nucleicos como el ADN y el ARN, otras moléculas, etc., interactúan según el estado del ciclo celular, la estimulación extracelular, etc. Por lo tanto, es preferible que el momento del tratamiento que mantiene la interacción o fragmenta el ADN genómico se seleccione según sea apropiado para cada propósito analítico particular. Antes de dicho tratamiento, las células pueden ser estimuladas si fuera necesario. Para una fragmentación eficiente, es preferible que la lisis celular, la disrupción celular o la extracción de la fracción de cromatina se realicen de acuerdo con un método conocido antes de la fragmentación.

10 El paso 2 produce un fragmento de ADN al que se une la molécula exógena, y uno o varios fragmentos de ADN al que no se une la molécula exógena.

(3) Paso 3

15 El paso 3 es un paso para permitir que un fragmento de ADN unido a la molécula exógena forme un complejo con una molécula capaz de unirse específicamente a la molécula exógena, y luego recuperar el complejo. Dado que la molécula exógena está unida a una secuencia de ADN endógeno específica presente dentro de la región genómica a aislar o cerca de ésta, se supone que el fragmento de ADN unido a la molécula exógena contiene la región genómica específica deseada. Por lo tanto, la región genómica específica deseada puede aislarse recuperando un complejo formado por la unión de una molécula capaz de unirse específicamente a la molécula exógena.

20 El método utilizado en el paso 3 puede ser cualquier método que permita la formación de un complejo de un fragmento de ADN unido a una molécula exógena que contenga la región genómica específica deseada y una molécula capaz de unirse específicamente a la molécula exógena, y la posterior recuperación del complejo. Por ejemplo, el complejo se puede recuperar después de ser precipitado con un vehículo en el que se inmoviliza un anticuerpo o un péptido capaz de unirse específicamente a la molécula exógena, un ión níquel, etc.

25 Por ejemplo, en el caso en el que la molécula exógena sea una molécula de fusión con dos tipos de secuencias marcadoras unidas mediante una secuencia de reconocimiento de proteasa, son preferibles los siguientes procedimientos. Se hace que el fragmento de ADN obtenido se una a una molécula capaz de unirse específicamente a una primera secuencia marcadora para la formación de un primer complejo y, luego, el primer complejo se trata con una proteasa para la escisión de la secuencia de reconocimiento de proteasa. Se hace que la porción que contiene el ADN genómico separada del primer complejo se una a una molécula capaz de unirse específicamente a una segunda secuencia marcadora para la formación de un segundo complejo. Dicha purificación en dos pasos puede prevenir eventos no deseados, por ejemplo, el enriquecimiento de moléculas como fondo debido a la unión no específica a una molécula capaz de unirse específicamente a la primera secuencia marcadora, y la unión no específica a un vehículo en el que se inmoviliza dicha molécula. Como consecuencia, la eficiencia del aislamiento específico de regiones genómicas específicas se mejora significativamente.

35 En el caso en el que la molécula exógena no sea aquella en la que dos tipos de secuencias marcadoras estén unidas mediante una secuencia de reconocimiento de proteasas, la separación del complejo deseado de una molécula capaz de unirse específicamente a una secuencia marcadora puede reducir la señal de fondo. Específicamente, por ejemplo, en el caso en que se use un marcador FLAG como secuencia marcadora, la adición de un péptido FLAG puede separar el complejo deseado de un anticuerpo anti-FLAG unido al complejo.

40 A continuación en este documento se describirá un ejemplo de la primera realización de la presente invención con referencia a la FIG. 1. Sin embargo, el método de la presente invención no se limita a esta realización.

(1) Una secuencia de ADN endógeno presente dentro o en las proximidades de la región genómica deseada en las células a analizar como se muestra en la FIG. 1A se determina como el objetivo de unión de una molécula exógena.

45 (2) Se diseña una proteína efectora TAL capaz de unirse específicamente a la secuencia de ADN endógeno determinada. Se prepara un vector de expresión para una proteína de fusión de la proteína efectora TAL, una secuencia de señal de localización nuclear (NLS) y una secuencia marcadora que consta de 3 repeticiones de un marcador FLAG (3 × FN-TAL, véase la figura 1B) y se introduce en las células a analizar para que la proteína de fusión pueda expresarse en ellas.

50 (3) Si es necesario, se estimulan las células a analizar, es decir, las células en las que se ha introducido el vector de expresión para 3 × FN-TAL. Si es necesario, para mantener la interacción del ADN genómico con proteínas, ARN, ADN y otras moléculas que interactúan en la región genómica a aislar, las células se someten a un tratamiento de reticulación con un agente de reticulación como el formaldehído. Huelga decir que 3 × FN-TAL está unido a la secuencia de ADN endógeno objetivo (véase la parte superior de la Figura 1C).

55 (4) Las células se lisan y el ADN genómico en un estado en el que se mantiene la interacción se digiere con una o varias desoxirribonucleasas como, p. ej. una enzima de restricción para la fragmentación. Alternativamente, la fragmentación puede realizarse por ultrasonidos. Este paso produce un fragmento de ADN al que se une 3 × FN-TAL, y un fragmento o varios fragmentos de ADN al que no se une 3 × FN-TAL.

(5) El fragmento de ADN al que se une  $3 \times$  FN-TAL forma un complejo con un anticuerpo anti-FLAG (complejo inmunoprecipitado) (véase la parte inferior de la figura 10).

(6) El complejo, incluso después del aislamiento, retiene moléculas que interactúan con la región genómica. Por lo tanto, después de que se elimina la interacción (enlaces cruzados), las moléculas que interactúan tales como proteínas, ARN, ADN y otras moléculas, pueden purificarse e identificarse.

Segunda realización

El método de la presente invención comprende los pasos I a III que se muestran a continuación. El método de la presente invención permite el aislamiento de una región genómica específica en un estado en el que se mantiene la interacción de la región genómica específica y las moléculas que interactúan con la misma. Siempre que una región genómica específica pueda aislarse en un estado en el que se mantiene la interacción, el método puede comprender uno o varios pasos distintos de los pasos I a III, y los detalles de los pasos no están limitados. En la segunda realización, las células a analizar no están limitadas a células viables, y también son preferibles las células en una muestra de tejido, etc.

Paso I: fragmentación del ADN genómico en un estado en el que se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo.

Paso II: poner el ADN genómico fragmentado en contacto con una molécula exógena capaz de unirse a una secuencia de ADN endógeno específica en el ADN genómico.

Paso III: recuperación de un fragmento de ADN genómico unido a la molécula exógena.

En lo sucesivo en este documento, el método de la presente invención se describirá con detalle en el orden de los pasos siguiente.

(1) Paso I

El Paso I es un paso para fragmentar el ADN genómico en un estado en el que se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo. En el caso en el que se usen células viables como las células a analizar, es preferible que el paso I sea precedido por un paso para realizar un tratamiento para mantener la interacción del ADN genómico y las moléculas que interactúan con el mismo, pero dicho paso de tratamiento puede ser omitido. En las células viables, se supone que las moléculas que interactúan con el ADN genómico varían según el estado del ciclo celular, la estimulación extracelular, etc. Por lo tanto, es preferible que el momento del tratamiento para mantener la interacción o para fragmentar el ADN genómico se seleccione según sea apropiado para cada propósito analítico particular. Antes de dicho tratamiento, las células pueden ser estimuladas si fuera necesario.

En el caso en el que las células a analizar sean células de una muestra de tejido, etc., se supone que se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo. Por lo tanto, generalmente no hay necesidad de un tratamiento adicional para mantener la interacción.

Cualesquiera que sean las células que analizar, es preferible para una fragmentación eficiente que la lisis celular, la disrupción celular o la extracción de la fracción de cromatina se realice de acuerdo con algún método conocido, antes de la fragmentación.

(2) Paso II

El Paso II es un paso para poner el ADN genómico fragmentado en contacto con una molécula exógena capaz de unirse a una secuencia de ADN endógeno específica en el ADN genómico. La molécula exógena se inmoviliza preferiblemente en un vehículo por adelantado y luego se pone en contacto con el ADN genómico fragmentado preparado en el paso I. Alternativamente, la molécula exógena se puede poner en contacto con el ADN genómico fragmentado y luego inmovilizarse en un vehículo. Con el uso de una molécula exógena inmovilizada en un vehículo, solo un fragmento de ADN que contiene una secuencia de ADN endógeno a la que se une la molécula exógena puede quedar atrapado en el vehículo. La molécula exógena se puede producir mediante el método descrito en la primera realización.

El vehículo sobre el que se inmoviliza la molécula exógena no está particularmente limitado siempre que pueda inmovilizar una proteína o un ácido nucleico. Los ejemplos de vehículos incluyen perlas, perlas magnéticas, discos, barras, tubos, placas de microtitulación, chips de microsensor y microarrays, todos hechos de vidrio, polietileno, polipropileno, acetato de polivinilo, cloruro de polivinilo, metacrilatos, látex, agarosa, celulosa, dextrano, almidón, dextrina, gel de sílice, cerámica porosa, etc. El método para inmovilizar la molécula exógena en el vehículo puede seleccionarse según sea apropiado a partir de métodos conocidos tales como la adsorción física, la unión covalente y la reticulación.

(3) Paso III

El paso III es un paso para la recuperación de un fragmento de ADN genómico unido a la molécula exógena. Dado

que la molécula exógena se une a una secuencia de ADN endógeno específica presente dentro o cerca de la región genómica a aislar, se supone que el fragmento de ADN genómico unido a la molécula exógena contiene la región genómica específica deseada. Por lo tanto, la región genómica específica deseada puede aislarse recuperando el fragmento de ADN genómico unido a la molécula exógena.

5 Los procedimientos concretos en los pasos II y III no están particularmente limitados. Por ejemplo, un vehículo en el que se inmoviliza la molécula exógena se empaqueta en una columna, el fragmento de ADN preparado en el paso I se pasa a través de la columna para que la región de ADN genómico deseada se una a la molécula exógena, la columna se lava y posteriormente la elución se realiza para recuperar la región de ADN genómico deseada. En otro ejemplo, el fragmento de ADN preparado en el paso I se agrega a un tubo que contiene un vehículo en el que se inmoviliza la molécula exógena, la mezcla se somete a agitación u otras operaciones para que la región de ADN genómico deseada entre en contacto con la molécula exógena, el vehículo se lava y se realiza la elución posterior para recuperar la región de ADN genómico deseada.

10 Un ejemplo de la segunda realización de la presente invención se muestra en la FIG. 2. Sin embargo, el método de la presente invención no se limita a esta realización. En lo sucesivo, la segunda realización de la presente invención se describirá con referencia a la FIG. 2.

(1) Una secuencia de ADN endógeno presente dentro o en las proximidades de la región genómica deseada en las células a analizar como se muestra en la FIG. 2A se determina como el objetivo de unión de una molécula exógena.

(2) Se diseña una proteína efectora TAL capaz de unirse específicamente a la secuencia de ADN endógena determinada, y se obtiene una proteína recombinante en la que se marca la proteína efectora TAL. El marcador es, por ejemplo, HaloTag (marca registrada, fabricado por Promega). Este marcador se inmoviliza por unión covalente en cuentas magnéticas HaloLink (nombre comercial, fabricado por Promega) (cuentas conjugadas con TAL, véase la figura 2B). Otro ejemplo de marcador es un marcador GST. Este marcador se inmoviliza por unión por afinidad en Glutathión Sefarosa 4B (fabricado por GE Healthcare).

(3) Si es necesario, se estimulan las células a analizar. Si es necesario, para mantener la interacción del ADN genómico con proteínas, ARN, ADN y otras moléculas que interactúan en la región genómica a aislar, las células se someten a un tratamiento de reticulación con un agente de reticulación como el formaldehído (ver la parte superior de la Figura 2C).

(4) Las células se lisan, y el ADN genómico en un estado en el que se mantiene la interacción se digiere con una o varias desoxirribonucleasas como, por ej. una enzima de restricción para la fragmentación. Alternativamente, la fragmentación puede realizarse por ultrasonidos. Este paso produce un fragmento de ADN que contiene la secuencia de ADN endógeno a la que se unen las perlas conjugadas con TAL, y un fragmento o varios de ADN que no contiene la secuencia de ADN endógeno.

(5) El fragmento de ADN que contiene la secuencia de ADN endógeno a la que se unen las perlas conjugadas con TAL se une a las perlas conjugadas con TAL (véase la parte inferior de la figura 2C).

(6) Las perlas conjugadas con TAL se recuperan y, por lo tanto, se puede aislar la región genómica deseada. La región genómica aislada permanece unida a las moléculas que interactúan con ella. Por lo tanto, después de eliminar la interacción (reticulaciones), las moléculas que interactúan, como, p. ej. proteínas, ARN, ADN y otras moléculas, pueden purificarse e identificarse.

#### Uso de regiones genómicas específicas aisladas

40 La región genómica específica aislada por el método de la presente invención es muy útil como muestra para el análisis de los mecanismos moleculares de las funciones del genoma. Dado que la región genómica específica aislada se encuentra en un estado en el que se mantiene la interacción del ADN genómico y las moléculas que interactúan con ella (proteínas, ADN, ARN, etc.), la identificación de las moléculas que interactúan y la deducción de sus funciones son factibles. Por lo tanto, se espera que el método de la presente invención contribuya en gran medida al análisis de los mecanismos moleculares de las funciones del genoma.

45 La identificación de las moléculas que interactúan en la región genómica específica aislada se puede realizar mediante una combinación apropiada de métodos conocidos sin limitación particular. Esto se puede lograr, por ejemplo, eliminando la interacción, separando y purificando las moléculas que interactúan, y empleando uno o varios métodos de identificación conocidos. En el caso en el que el tratamiento para mantener la interacción sea la reticulación con formaldehído, las reticulaciones se pueden revertir mediante la adición de una solución salina de alta concentración (solución de NaCl 5 M, etc.) y la posterior incubación a una temperatura adecuada durante un período de tiempo adecuado (por ejemplo, a aproximadamente 65°C durante la noche). Los ejemplos del método de identificación conocido para proteínas incluyen espectrometría de masas, inmunotransferencia y ELISA (ensayo inmunoabsorbente ligado a enzimas). Los ejemplos del método de identificación conocido para ADN o ARN incluyen análisis de secuencia de nucleótidos, análisis de microarrays y PCR (reacción en cadena de la polimerasa).

## Ejemplos

En lo sucesivo en este documento, la presente invención se ilustrará con detalle mediante ejemplos, aunque no se limita a los mismos.

### Ejemplo 1

#### 5 Aislamiento de la región telomérica

##### (1) Preparación de células para analizar

Las células que expresan una proteína efectora TAL que reconoce repeticiones teloméricas de TTAGGG se prepararon como sigue. Se sintetizó un ADN que codificaba Telomere-TAL, que es capaz de unirse a un ADN bicatenario de 19 mer que contiene las repeticiones teloméricas de TTAGGG (TAGGGTTAGGGTTAGGGTT: SEQ ID NO: 1). La síntesis de ADN se subcontrató a Life Technologies. Se preparó un vector de retrovirus para la expresión de una proteína de fusión 3 × FN-Tel-TAL (SEQ ID NO: 2), en el que Telomere-TAL se fusiona con un marcador 3 × FLAG y una señal de localización nuclear (NLS) (en adelante en este documento, este vector de expresión se llama 3 × FN-Tel-TAL / pMXs-puro), una línea celular hematopoyética murina Ba / F3 se infectó con 3 × FN-Tel-TAL / pMXs-puro, y solo se seleccionaron células infectadas con éxito en presencia de puromicina. La secuencia de aminoácidos de la proteína 3 × FN-Tel-TAL está representada por la SEQ ID NO: 2.

Las células de control negativo se prepararon como sigue. Un vector de retrovirus para la expresión de una proteína de fusión 3 × FNLDD, que consiste en un marcador 3 × FLAG, un NLS y una proteína bacteriana LexA (Fujita y Fujii, Adv. Biosci. Biotechnol. 3, 626-629 (2012)), se preparó (en adelante en este documento este vector de expresión se llama 3 × FNLDD / pMXs-puro), las células Ba / F3 se infectaron con 3 × FNLDD / pMXs-puro, y solo se infectaron con éxito células infectadas seleccionadas en presencia de puromicina.

Para visualizar la expresión de la proteína 3 × FN-Tel-TAL y la proteína 3 × FNLDD, se realizó una tinción intracelular usando un anticuerpo con marcador anti-FLAG (fabricado por Sigma-Aldrich, anti-FLAG M2 conjugado con isotiocianato de fluoresceína, F4049), seguido de análisis de citometría de flujo. Los resultados se muestran en la fig. 3. En la figura, Ba / F3 (izquierda) muestra los resultados de las células huésped intactas, 3 × FNLDD (centro) muestra los resultados de las células que expresan la proteína 3 × FNLDD, y 3 × FN-Tel-TAL (derecha) muestra los resultados de las células que expresan la proteína 3 × FN-Tel-TAL. Como se muestra en la fig. 3, los niveles de expresión de la proteína 3 × FN-Tel-TAL y la proteína 3 × FNLDD fueron comparables.

##### (2) Aislamiento de la región genómica específica

(a) A una suspensión celular que contenía  $2 \times 10^7$  células en 30 ml de medio RPMI completo suplementado con interleucina 3, se añadieron 810  $\mu$ l de formaldehído al 37%, y las células se incubaron a 37°C durante 5 minutos para realizar el tratamiento de reticulación. A continuación, se añadieron 3 ml de una solución de glicina 1,25 M, y se dejó que se desarrollara una reacción de neutralización a temperatura ambiente durante 10 minutos.

(b) Después de la centrifugación (a 1,3 krpm a 4°C durante 5 minutos), las células se lavaron dos veces con PBS helado, se suspendieron en 10 ml de tampón de lisis celular (Tris 10 mM (pH 8,0), EDTA 1 mM, IGEPAL CA-630 al 0,5%, cóctel inhibidor de proteasa), y la suspensión celular se colocó en hielo durante 10 minutos. Después de la centrifugación (a 2,0 krpm a 4°C durante 8 minutos), se descartó el sobrenadante, se suspendió el sedimento celular en 10 ml de tampón de lisis nuclear (Tris 10 mM (pH 8,0), EDTA 1 mM, NaCl 0,5 M, Triton X-100 al 1%, desoxicolato de sodio al 0,5%, sal de lauroilsarcosina al 0,5%, cóctel inhibidor de la proteasa) y la suspensión celular se colocó en hielo durante 10 minutos con agitación intermitente (intervalo de 2 a 3 minutos) utilizando un mezclador vórtex. Después de la centrifugación (a 2,0 krpm a 4°C durante 8 minutos), se descartó el sobrenadante, se lavó el sedimento celular con PBS helado y se obtuvo una fracción de cromatina.

(c) La fracción de cromatina se suspendió en 800  $\mu$ l de tampón de lisis 3 (Tris 10 mM (pH 8,0), EDTA 1 mM, EGTA 0,5 mM, NaCl 150 mM, desoxicolato de sodio al 0,1%, SDS al 0,1%, cóctel inhibidor de proteasa) y se sometió a ultrasonidos para la fragmentación del ADN genómico reticulado (generador ultrasónico: fabricado por TOMY SEIKO, disruptor ultrasónico UD-201, salida: 3, servicio: 100%, un conjunto de 10 segundos  $\times$  6 con intervalos de 20 segundos se repitió 3 veces a intervalos de 2 minutos (18 veces en total)).

(d) Después de la centrifugación (a 1,3 krpm a 4°C durante 10 minutos), se recogió el sobrenadante (800  $\mu$ L). A 160  $\mu$ L (equivalente a  $4 \times 10^6$  células) del sobrenadante, 240  $\mu$ L de tampón de lisis 3, 50  $\mu$ L de tampón de lisis 3 suplementado con 10% de Triton X-100 y 50  $\mu$ L de una solución inhibidora de proteasa 10  $\times$  fueron añadidos a un volumen total de 500  $\mu$ L. Esto se aclaró anteriormente con Dynabeads conjugadas con IgG de control, para eliminar de ese modo los componentes unidos de forma no específica. Posteriormente, se realizó la inmunoprecipitación usando Dynabeads conjugadas con un anticuerpo anti-FLAG M2 (fabricado por Sigma-Aldrich, anti-FLAG M2, 811804).

(e) Después del lavado, el complejo inmunoprecipitado se suspendió en 300  $\mu$ l de TE. Después de la reversión de la reticulación, se realizó el tratamiento con RNasa A y Proteinasa K, seguido de tratamiento con fenol / cloroformo

para dar un ADN purificado.

(f) El ADN obtenido se usó para el análisis de transferencia Southern. Para la detección del ADN telomérico, se utilizó el kit Telo TAGGG Telomerase Length Assay (12209136001, fabricado por Roche).

5 Los resultados del análisis de transferencia Southern se muestran en las Figs. 4 (A) y (B). (A) y (B) muestran la reproducibilidad de los resultados. Como se desprende de las Figs. 4 (A) y (B), en el análisis de transferencia Southern de las dos muestras obtenidas por inmunoprecipitación, el ADN telomérico se detectó solo en el complejo inmunoprecipitado que contenía 3 × FN-Tel-TAL. Estos resultados muestran que la región telomérica se puede recuperar por inmunoprecipitación usando 3 × FN-Tel-TAL.

(3) Identificación de proteínas contenidas en la región telomérica aislada

10 (g) Se realizaron los mismos procedimientos que en los anteriores (a) a (d) para dar un inmunoprecipitado. Para la inmunoprecipitación, se usaron 300 µL de Dynabeads y 30 µL de un anticuerpo. Al inmunoprecipitado obtenido, se añadieron 200 µL de una solución de péptido FLAG 3x (fabricada por Sigma-Aldrich, F4799, 0,5 mg / ml), y esto se incubó a 37°C durante 20 minutos para realizar la elución de las perlas. El eluato se precipitó con isopropanol, se suspendió en 40 µl de tampón de muestra 2x SDS, y se incubó a 100°C durante 30 minutos para lograr tanto la  
15 degeneración de proteínas como la reversión de la reticulación.

(h) Posteriormente, se realizó la electroforesis en un gel SDS-PAGE con gradiente de 4-20% hasta que la muestra migró aproximadamente 1,5 cm dentro del gel. Después de teñir con azul brillante de Coomassie (CBB), la muestra en el gel se dividió en cinco fracciones y se sometió a análisis de espectrometría de masas.

20 Las proteínas detectadas más abundantemente en el caso de la inmunoprecipitación usando 3 × FN-Tel-TAL que usando 3 × FNLDD incluían muchas de las siguientes proteínas: proteínas de unión a telómeros conocidas, proteínas que se sabe que interactúan con otras proteínas de unión a telómeros, y proteínas que genéticamente demostraron estar involucradas en las funciones de los telómeros (ver Tabla 1). Estos resultados demuestran que, según el método de la presente invención, las proteínas que se unen específicamente a la secuencia de una región genómica específica pueden aislarse como un complejo de cromatina e identificarse por espectrometría de masas.

25 Tabla 1

Proteínas de unión a telómeros de mamíferos	PML, RPA, CDK1, PPAR1, PCBP1
Proteínas de unión a telómeros en levaduras u otros organismos	IMP4
Proteínas que interactúan con proteínas de unión a telómeros (proteína de unión telómero asociada)	ADN polimerasa α (Cdc13p), ARMC6 (TRF2), CTBP1 (FoxP2-POT1), exportin-5 (TERT), GNL3L (TRF1), exportin-1 (TERT), 14-3-3 (TERT)
Proteínas que se localizan en la heterocromatina	BEND3
Proteínas cuyas mutaciones afectan a la función de los telómeros	ADN polimerasa α, HAT1, Nup133, CDK7, DPOE1, PRDX1, TYSY, glutamato-cisteína ligasa, glutaredoxina, SMRC1

(4) Identificación del ARN contenido en la región telomérica aislada

30 (i) Se realizaron los mismos procedimientos que en los anteriores (a) a (e) para dar un inmunoprecipitado. Todos los tampones que utilizar se complementaron con RNasin Plus (Promega). Después del tratamiento con Proteinasa K, la purificación de ARN se realizó con Isogen II (NIPPON GENE CO., LTD.).

(j) Posteriormente, la RT-PCR se realizó utilizando TITANIUM One-Step RT-PCR (Clontech) en un intento de detectar una ARN telomerasa. Los cebadores utilizados para la RT-PCR fueron 5'-CCGGCGCCCCGCGGCTGACAGAG-3' (SEQ ID NO: 3) (para transcripción inversa), 5'-GCTGTGGGTTCTGGTCTTTTGTTC-3' (SEQ ID NO: 4) y 5'-GCGGCAGCGGAGTCC-3' (SEQ ID NO: 5).

35 Los resultados de la RT-PCR se muestran en la FIG. 5. Como se desprende de la fig. 5, el ARN de la telomerasa se detectó solo en el caso de la inmunoprecipitación usando 3 × FN-Tel-TAL. Estos resultados demuestran que, según el método de la presente invención, los ARN que se unen específicamente a la secuencia de una región genómica específica pueden aislarse como un complejo de cromatina y cuantificarse mediante RT-PCR.

Ejemplo 2: Aislamiento de la región promotora del gen Sox2

40 (1) Preparación de células para analizar

Las células que expresan una proteína efectora TAL que reconoce una secuencia de nucleótidos (TTATCCCTGACA, SEQ ID NO: 6) presente en la región promotora del gen Sox2 humano se prepararon como sigue. Se adquirió un ADN

que codificaba la proteína efectora TAL que reconocía la secuencia anterior (SEQ ID NO: 6) de Addgene (Addgene N°. 35388). Un vector plasmídico para la expresión de una proteína de fusión 3 × FN-Sox2-TAL, en el que la proteína efectora TAL se fusiona con un marcador 3 × FLAG y una señal de localización nuclear (NLS), se preparó y transfectó en células 293T de riñón embrionario humano para preparar células que expresaban 3 × FN-Sox2-TAL.

- 5 Las células de control negativo se prepararon como sigue. Un ADN que codificaba una proteína efectora TAL que reconocía una secuencia de nucleótidos (TCTTACTTATAAC, SEQ ID NO: 7) presente en la región promotora del gen KLF4 humano (Addgene N°. 35389) y un ADN que codificaba una proteína efectora TAL que reconocía una secuencia de nucleótidos ( ACCACTCACTATA, SEQ ID NO: 8) presentes en la región promotora de citomegalovirus (Addgene N°. 27970) se compraron en Addgene, y se realizó el mismo procedimiento que anteriormente para preparar células que expresaban 3 × FN-KLF4-TAL y células que expresaban 3 × FN-CMV-TAL8.

La expresión de estas proteínas efectoras TAL se confirmó mediante transferencia Western usando un anticuerpo con marcador anti-FLAG (fabricado por Sigma-Aldrich, anti-FLAG M2, F1804). Los resultados se muestran en la fig. 6. Como se muestra en la fig. 6, los niveles de expresión de estas proteínas efectoras TAL fueron comparables.

(2) Aislamiento de la región genómica específica

- 15 (a) A una suspensión celular que contenía  $1,5 \times 10^7$  células en 30 ml de medio DMEM completo, se añadieron 810 µl de formaldehído al 37%, y las células se incubaron a 37°C durante 5 minutos para el tratamiento de reticulación. A continuación, se añadieron 3 ml de una solución de glicina 1,25 M, y se dejó que se desarrollara una reacción de neutralización a temperatura ambiente durante 10 minutos.

- 20 (b) a (d) Se realizaron los mismos procedimientos que en los anteriores (b) a (d) del Ejemplo 1 (2) para la inmunoprecipitación de la región promotora Sox2 a analizar.

- 25 (e) Después del lavado, se añadieron 40 µL de una solución de péptido FLAG 3x (fabricada por Sigma-Aldrich, F4799, 0,5 mg / ml) al inmunoprecipitado obtenido, y esto se incubó a 37°C durante 30 minutos para su elución de las cuentas. El inmunoprecipitado eluido se suspendió en 60 µL de TE y se sometió a tratamiento con RNasa A y Proteínasa K, seguido de reversión de la reticulación. Después de esto, la purificación de ADN se realizó con la columna CHIP DNA Clean & Concentrator-Capped (fabricada por ZYMO RESEARCH, D5205).

- (f) La PCR en tiempo real se realizó utilizando el ADN obtenido. Los cebadores utilizados para la PCR en tiempo real fueron 5'-ATTGGTGCCTAGAAACCCATTTATT-3' (SEQ ID NO: 9) y 5'-CTGCCTTGACAACCTCCTGATACTTT-3' (SEQ ID NO: 10). Estos cebadores se diseñaron para flanquear la secuencia de reconocimiento 3 × FN-Sox2-TAL (TTATTCCCTGACA, SEQ ID NO: 6), que está presente en la región promotora del gen Sox2 humano (ver Figura 7).

- 30 Los resultados de la PCR en tiempo real se muestran en la FIG. 8. Como se desprende de la fig. 8, en el caso de la inmunoprecipitación usando 3 × FN-Sox2-TAL, la región promotora de Sox2 se enriqueció aproximadamente de 2,5 a 3 veces en comparación con los controles negativos. Estos resultados muestran que la región promotora Sox2 puede recuperarse mediante inmunoprecipitación usando 3x FN-Sox2-TAL. Los resultados anteriores indican que el método de la presente invención es aplicable incluso en los casos en los que el análisis de interés es una región genómica que está presente en solo una o dos copias por célula.

Ejemplo 3: Aislamiento del locus IRF-1 utilizando el complejo inactivo de proteína Cas9-ARN (sistema CRISPR)

El esquema de este ejemplo se muestra en la FIG. 9.

(1) Preparación del sistema CRISPR para analizar

- 40 Se adquirió un plásmido que codifica un mutante Cas9 D10A que llevaba la señal de localización nuclear del antígeno T SV40 de Addgene (Addgene N°. 41816). En este plásmido, se introdujo la mutación H840A para preparar un mutante Cas9 enzimáticamente inactivo (dCas9). La secuencia de codificación dCas9 se insertó en p3 × FLAG-CMV-7.1 (Sigma-Aldrich) para preparar un vector para la expresión de una proteína dCas9 fusionada con un marcador 3 × FLAG (en adelante en este documento, este vector de expresión se llama 3 × F-dCas9 / pCMV-7.1). 3 × F-dCas9 / pCMV-7.1 se transfectó a células 293T de riñón embrionario humano para la expresión transitoria de 3 × F-dCas9. Además, 3 × FNLDD / pCMV-7.1 se transfectó a células 293T para la expresión transitoria de 3 × FNLDD, y las células resultantes se usaron para comparar con las células que expresaban 3 × F-dCas9.

- 50 De cada tipo de células, se preparó un extracto nuclear y se sometió a transferencia Western usando un anticuerpo con marcador anti-FLAG (fabricado por Sigma-Aldrich, anti-FLAG M2, F1804) para confirmar la expresión de 3 × F-dCas9 y 3 × FNLDD. Los resultados se muestran en la fig. 10. En la figura, (-) representa células huésped intactas. La imagen de la tinción de CBB que se muestra en el panel inferior indica que se cargaron las mismas cantidades de muestras en todos los carriles. Los resultados de la transferencia western muestran que los niveles de expresión de 3 × F-dCas9 y 3 × FNLDD fueron comparables.

A continuación, de acuerdo con Mali et al., Science (2013) 339: 823-826, un vector de expresión para un gRNA correspondiente a la secuencia 5'-CCGGGGGCGCTGGGCTGTCCCGG-3' (SEQ ID NO: 11), que está presente en el

locus IRF-1 de ser humano se construyó de la siguiente manera. Las siguientes dos oligonucleótidos:

5' -TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCGGGGGCGCTGGGCT

GTCC-3' (SEQ ID NO: 12) y

5' -GACTAGCCTTATTTAACTTGCTATTTCTAGCTCTAAAACGGACAGCCCAGCGCCC

CCGC-3' (SEQ ID NO: 13)

se hibridaron, y una polimerasa Phusion (New England BioLabs Inc.) y se sintetizó un fragmento de ADN bicatenario de 100 pb. Este fragmento se purificó por electroforesis en agarosa y se ligó a un vector de clonación de ARNg (Addgene N°. 41824) usando el Kit de ensamblaje Gibson (New England BioLabs Inc.) para preparar un vector para la expresión del ARNg (en adelante en este documento, este vector de expresión se denomina gRNA-hIRF -1 N°. 12).

(2) Aislamiento del locus IRF-1 humano

(a) Los 3 × F-dCas9 / pCMV-7.1 y gRNA-hIRF-1 N°. 12 preparados anteriormente se transfectaron en células 293T de riñón embrionario humano (3 × 10<sup>6</sup> células) con el uso de Lipofectamina 2000 (Invitrogen). El día después de la transfección, las células fueron sembradas de nuevo. Al día siguiente, se añadieron 810 µl de formaldehído al 37% a las células en 30 ml de medio completo DMEM, y las células se incubaron a 37°C durante 5 minutos para el tratamiento de reticulación. A continuación, se añadieron 3 ml de una solución de glicina 1,25 M, y se dejó que se desarrollara una reacción de neutralización a temperatura ambiente durante 10 minutos.

(b) a (d) Se realizaron los mismos procedimientos que en los anteriores (b) a (d) del Ejemplo 1 (2) para la inmunoprecipitación de la región IRF-1 a analizar.

(e) Después del lavado, el complejo inmunoprecipitado se suspendió en 285 µL de TE. Después de la adición de 12 µl de NaCl 5 M, se continuó la incubación a 65°C durante la noche para la reversión de la reticulación. Posteriormente, se realizó el tratamiento con RNasa A y Proteinasa K, seguido de un tratamiento con fenol / cloroformo para dar un ADN purificado.

(f) La PCR en tiempo real se realizó utilizando el ADN obtenido. Los cebadores utilizados para la PCR en tiempo real fueron 5'-CGCCTGCGTTTCGGGAGATATAC-3' (SEQ ID NO: 14) y 5'-CTGTCTCTCACTCCGCCTTGT-3' (SEQ ID NO: 15). Estos cebadores fueron diseñados para recocer las proximidades de la secuencia de ADN objetivo del locus IRF-1 humano. Además, los cebadores de SEQ ID NO: 9 y 10 se usaron para la cuantificación del locus Sox2, que es una región no relacionada.

Los resultados de la PCR en tiempo real se muestran en la FIG. 11. Como se desprende de la fig. 11, el locus IRF-1 apenas se enriqueció en el control negativo y, por el contrario, cuando se usaban 3 × F-dCas9 y gRNA-hIRF-1 N°. 12, el porcentaje de enriquecimiento en comparación con la entrada fue tan alto como 7,7%. El locus Sox2 no relacionado no se enriqueció en absoluto. Estos resultados muestran que el locus diana puede recuperarse de manera específica y eficiente mediante inmunoprecipitación usando 3 × F-dCas9 y el ARNg. Los resultados anteriores indican que el método de la presente invención es aplicable incluso en casos en los que el analito de interés es una región genómica que está presente en solo una o dos copias por célula.

(3) Identificación de proteínas contenidas en la región aislada de IRF-1

(g) Se realizaron los mismos procedimientos que en los anteriores (a) a (d) para dar un inmunoprecipitado. Para la inmunoprecipitación, se usaron 300 µL de Dynabeads y 30 µL de un anticuerpo. Al inmunoprecipitado obtenido, se añadieron 200 µL de una solución de péptido 3 × FLAG (fabricada por Sigma-Aldrich, F4799, 0,5 mg / ml), y esto se incubó a 37°C durante 20 minutos para elución de las perlas. El eluato se precipitó con isopropanol, se suspendió en 40 µl de tampón de muestra 2x SDS, y se incubó a 100°C durante 30 minutos para lograr tanto la degeneración de proteínas como la reversión de la reticulación.

(h) Posteriormente, la electroforesis se realizó en un gel SDS-PAGE con gradiente de 4-20% hasta que la muestra migró aproximadamente 1,0 cm dentro del gel. Después de teñir con CBB, la muestra en el gel se dividió en cinco fracciones y se sometió a análisis de espectrometría de masas (véase la figura 12).

Las proteínas detectadas incluían muchas de las siguientes proteínas: componentes proteicos conocidos de la cromatina, tales como histonas, proteínas de unión a histonas, proteínas de enzimas metabólicas y helicasas de ARN (ver Tabla 2). Estos resultados demuestran que, según el método de la presente invención, las proteínas que se unen a la secuencia de una región genómica específica pueden aislarse como un complejo de cromatina e identificarse por espectrometría de masas.





ES 2 768 291 T3

<400> 2

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp  
 1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Leu Ala Ala Gly Arg Ala Thr Met Gly  
 20 25 30

Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr Gly Gly Met  
 35 40 45

Ala Pro Lys Lys Lys Arg Lys Val Asp Gly Gly Val Asp Leu Arg Thr  
 50 55 60

Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg  
 65 70 75 80

Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr  
 85 90 95

His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr

ES 2 768 291 T3

		100						105						110			
Val	Ala	Val	Lys	Tyr	Gln	Asp	Met	Ile	Ala	Ala	Leu	Pro	Glu	Ala	Thr		
		115					120					125					
His	Glu	Ala	Ile	Val	Gly	Val	Gly	Lys	Gln	Trp	Ser	Gly	Ala	Arg	Ala		
	130					135					140						
Leu	Glu	Ala	Leu	Leu	Thr	Val	Ala	Gly	Glu	Leu	Arg	Gly	Pro	Pro	Leu		
145					150					155					160		
Gln	Leu	Asp	Thr	Gly	Gln	Leu	Leu	Lys	Ile	Ala	Lys	Arg	Gly	Gly	Val		
				165					170					175			
Thr	Ala	Val	Glu	Ala	Val	His	Ala	Trp	Arg	Asn	Ala	Leu	Thr	Gly	Ala		
			180					185						190			
Pro	Leu	Asn	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile		
		195					200					205					
Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu		
	210					215					220						
Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser		
225					230					235					240		
Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro		
				245					250					255			
Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile		
			260					265					270				
Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu		
		275					280						285				
Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln	Val	Val		
	290					295					300						
Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln		
305					310					315					320		
Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln		
				325					330					335			
Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr		
			340					345						350			

ES 2 768 291 T3

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro  
 355 360 365

Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu  
 370 375 380

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
 385 390 395 400

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln  
 405 410 415

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
 420 425 430

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly  
 435 440 445

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln  
 450 455 460

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn  
 465 470 475 480

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu  
 485 490 495

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser  
 500 505 510

Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro  
 515 520 525

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile  
 530 535 540

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu  
 545 550 555 560

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val  
 565 570 575

Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln  
 580 585 590

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln  
 595 600 605

ES 2 768 291 T3

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr  
610 615 620

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro  
625 630 635 640

Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu  
645 650 655

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
660 665 670

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln  
675 680 685

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
690 695 700

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly  
705 710 715 720

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln  
725 730 735

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly  
740 745 750

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu  
755 760 765

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser  
770 775 780

Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser  
785 790 795 800

Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala  
805 810 815

Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly  
820 825 830

Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro  
835 840 845

Glu Arg Thr Ser His Arg Val Ala Asp His Ala Gln Val Val Arg Val  
850 855 860

ES 2 768 291 T3

Leu Gly Phe Phe Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp  
865 870 875 880

Ala Met Thr Gln Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe  
885 890 895

Arg Arg Val Gly Val Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro  
900 905 910

Pro Ala Ser Gln Arg Trp Asp Arg Ile Leu Gln Gly Ser Arg Leu Ile  
915 920 925

His Arg  
930

- 5 <210> 3  
<211> 23  
<212> ADN  
<213> Secuencia artificial
- 10 <220>  
<223> Secuencia de cebador  
  
<400> 3  
ccggcgcccc gcggtgaca gag 23  
  
<210> 4  
<211> 24  
<212> ADN  
<213> Secuencia artificial
- 15 <220>  
<223> Secuencia de cebador  
  
<400> 4  
gctgtgggtt ctggtctttt gttc 24  
  
<210> 5  
<211> 19  
<212> ADN  
<213> Secuencia artificial
- 20 <220>  
<223> Secuencia de cebador  
  
<400> 5  
gcggcagcgg agtcctaag 19  
  
<210> 6  
<211> 13  
<212> ADN  
<213> Homo sapiens
- 25 <220>  
<223> Secuencia de cebador  
  
<400> 6  
ttattccctg aca 13  
  
<210> 7  
<211> 13  
<212> ADN  
<213> Homo sapiens
- 30 <220>  
<223> Secuencia de cebador  
  
<400> 6  
ttattccctg aca 13  
  
<210> 7  
<211> 13  
<212> ADN  
<213> Homo sapiens
- 35 <220>  
<223> Secuencia de cebador  
  
<400> 6  
ttattccctg aca 13  
  
<210> 7  
<211> 13  
<212> ADN  
<213> Homo sapiens

ES 2 768 291 T3

	<400> 7 tcttacttat aac	13	
5	<210> 8 <211> 13 <212> ADN <213> Homo sapiens		
	<400> 8 accactcact ata	13	
10	<210> 9 <211> 25 <212> ADN <213> Secuencia artificial		
	<220> <223> Secuencia de cebador		
15	<400> 9 attggtcgct agaaacccat ttatt	25	
20	<210> 10 <211> 25 <212> ADN <213> Secuencia artificial		
	<220> <223> Secuencia de cebador		
25	<400> 10 ctgccttgac aactcctgat acttt	25	
	<210> 11 <211> 23 <212> ADN <213> Homo sapiens		
30	<400> 11 ccgggggcgc tgggctgtcc cgg	23	
	<210> 12 <211> 60 <212> ADN <213> Secuencia artificial		
35	<220> <223> Oligonucleótido		
40	<400> 12 tttcttgct ttatatatct tgtgaaagg acgaaacacc gcgggggcgc tgggctgtcc	60	
	<210> 13 <211> 60 <212> ADN <213> Secuencia artificial		
	<220> <223> Oligonucleótido		
45	<400> 13 gactagcctt attttaactt gctatttcta gctctaaaac ggacagccca ggcgggggcgc	60	
50	<210> 14 <211> 22 <212> ADN <213> Secuencia artificial		
	<220>		

<223> Secuencia de cebador  
<400> 14  
cgctgcggtt cgggagatat ac  
22  
5 <210> 15  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Secuencia de cebador  
10 <400> 15  
ctgtcctctc actccgcctt gt  
22

**REIVINDICACIONES**

1. Un método para aislar una región genómica específica mientras se mantiene la interacción de la región genómica específica y las moléculas que interactúan con ella, comprendiendo el método los siguientes pasos 1 a 3:
- 5 Paso 1: poner en contacto el ADN genómico con una molécula exógena capaz de unirse a una secuencia específica de ADN endógeno en el ADN genómico,
- Paso 2: fragmentar el ADN genómico en un estado en el que se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo, y
- 10 Paso 3: permitir que un fragmento de ADN genómico unido a la molécula exógena forme un complejo con una molécula capaz de unirse específicamente a la molécula exógena, y luego recuperar el complejo,
- en donde la fragmentación del ADN genómico se realiza por la digestión del ADN genómico con una o varias enzimas de restricción, la degradación parcial (escisión parcial) del ADN genómico con una o varias desoxirribonucleasas (endonucleasas) o la escisión física del ADN genómico por ultrasonidos, y
- en donde la molécula exógena es una proteína de unión al ADN exógena o un complejo de proteína-ácido nucleico exógeno, y
- 15 en donde la molécula exógena es una molécula de fusión que contiene uno o más tipos de secuencias marcadoras.
2. El método según la reivindicación 1, que comprende además la etapa de realizar un tratamiento para mantener la interacción del ADN genómico y las moléculas que interactúan con el mismo.
3. El método de acuerdo con la reivindicación 1 o 2, en el que la molécula exógena es una proteína de dedo de zinc, una proteína efectora de TAL o un complejo de una proteína Cas9 inactiva y un ARN guía (gARN).
- 20 4. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que la molécula exógena es una molécula de fusión que tiene una señal o varias señales de localización nuclear.
5. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en el que, en el paso 1, un gen o varios genes que codifican la molécula exógena se introducen en las células para analizar la expresión del gen en las células.
- 25 6. Un método para aislar una región genómica específica mientras se mantiene la interacción de la región genómica específica y las moléculas que interactúan con la misma, comprendiendo el método los siguientes pasos 1 a 3:
- Paso 1: fragmentar el ADN genómico en un estado en el que se mantiene la interacción del ADN genómico y las moléculas que interactúan con el mismo.
- Paso 2: poner el ADN genómico fragmentado en contacto con una molécula exógena capaz de unirse a una secuencia de ADN endógeno específica en el ADN genómico, y
- 30 Paso 3: permitir que un fragmento de ADN genómico unido a la molécula exógena forme un complejo con una molécula capaz de unirse específicamente a la molécula exógena, y luego recuperar el complejo,
- en el que la fragmentación del ADN genómico se realiza por la digestión del ADN genómico con una o varias enzimas de restricción, la degradación parcial (escisión parcial) del ADN genómico con una o varias desoxirribonucleasas (endonucleasas) o la escisión física del ADN genómico por ultrasonidos, y
- 35 en el que la molécula exógena es una proteína de unión al ADN exógena o un complejo de proteína-ácido nucleico exógeno, y
- en el que la molécula exógena es una molécula de fusión que contiene uno o más tipos de secuencias marcadoras.
7. El método de acuerdo con la reivindicación 6, que comprende además la etapa de realizar un tratamiento para mantener la interacción del ADN genómico y las moléculas que interactúan con el mismo.
- 40 8. El método de acuerdo con la reivindicación 6 o 7, en el que la molécula exógena es una proteína de dedo de zinc, una proteína efectora de TAL o un complejo de una proteína Cas9 inactiva y un ARN guía (gARN).
9. El método de acuerdo con cualquiera de las reivindicaciones 6 a 8, en el que, en el paso 2, el ADN genómico fragmentado se pone en contacto con la molécula exógena inmovilizada en un vehículo.
- 45 10. El método de acuerdo con una cualquiera de las reivindicaciones 6 a 8, en el que, en el paso 2, la molécula exógena se pone en contacto con el ADN genómico fragmentado y, posteriormente, se inmoviliza en un vehículo.



Fig. 1

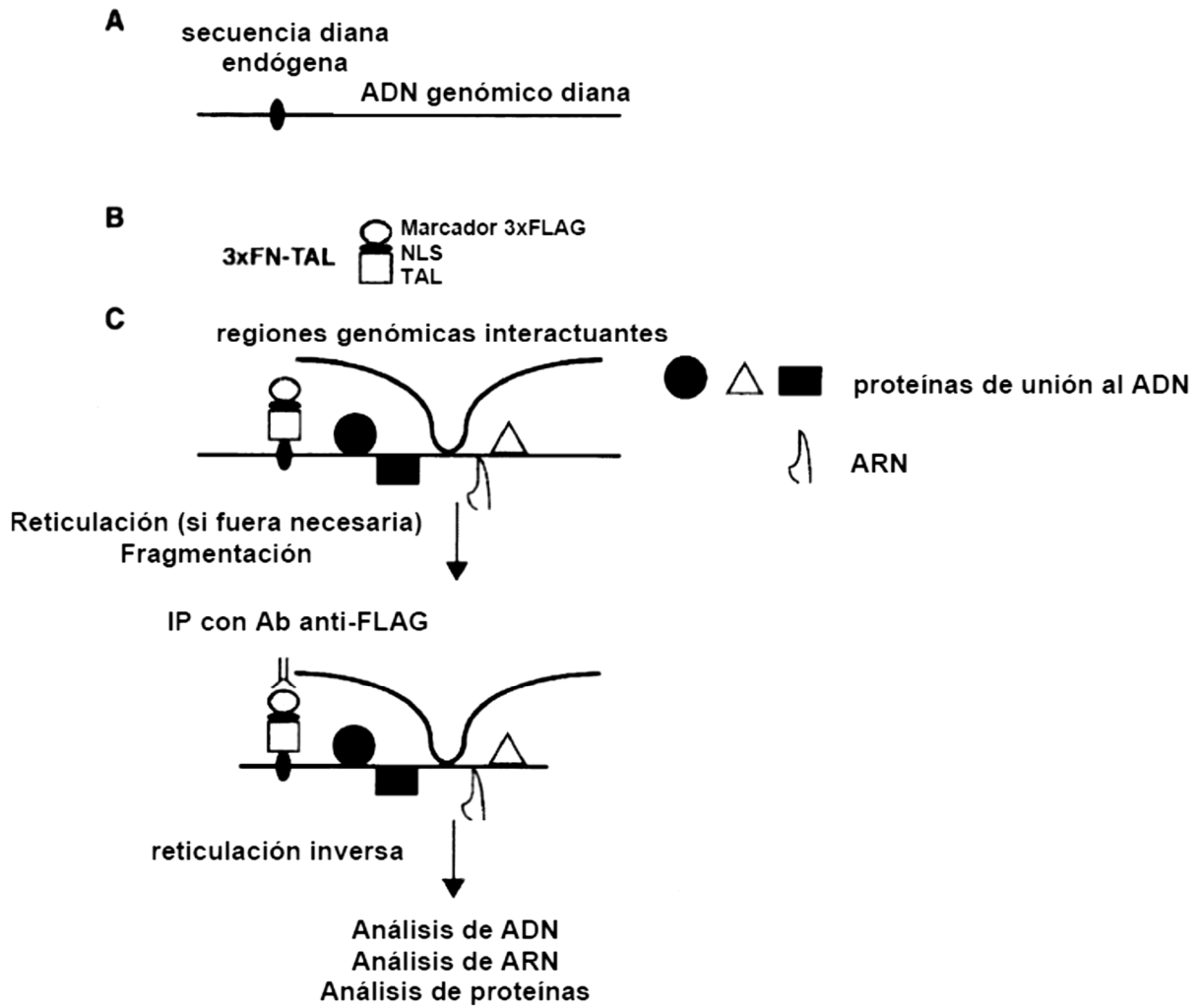


Fig. 2

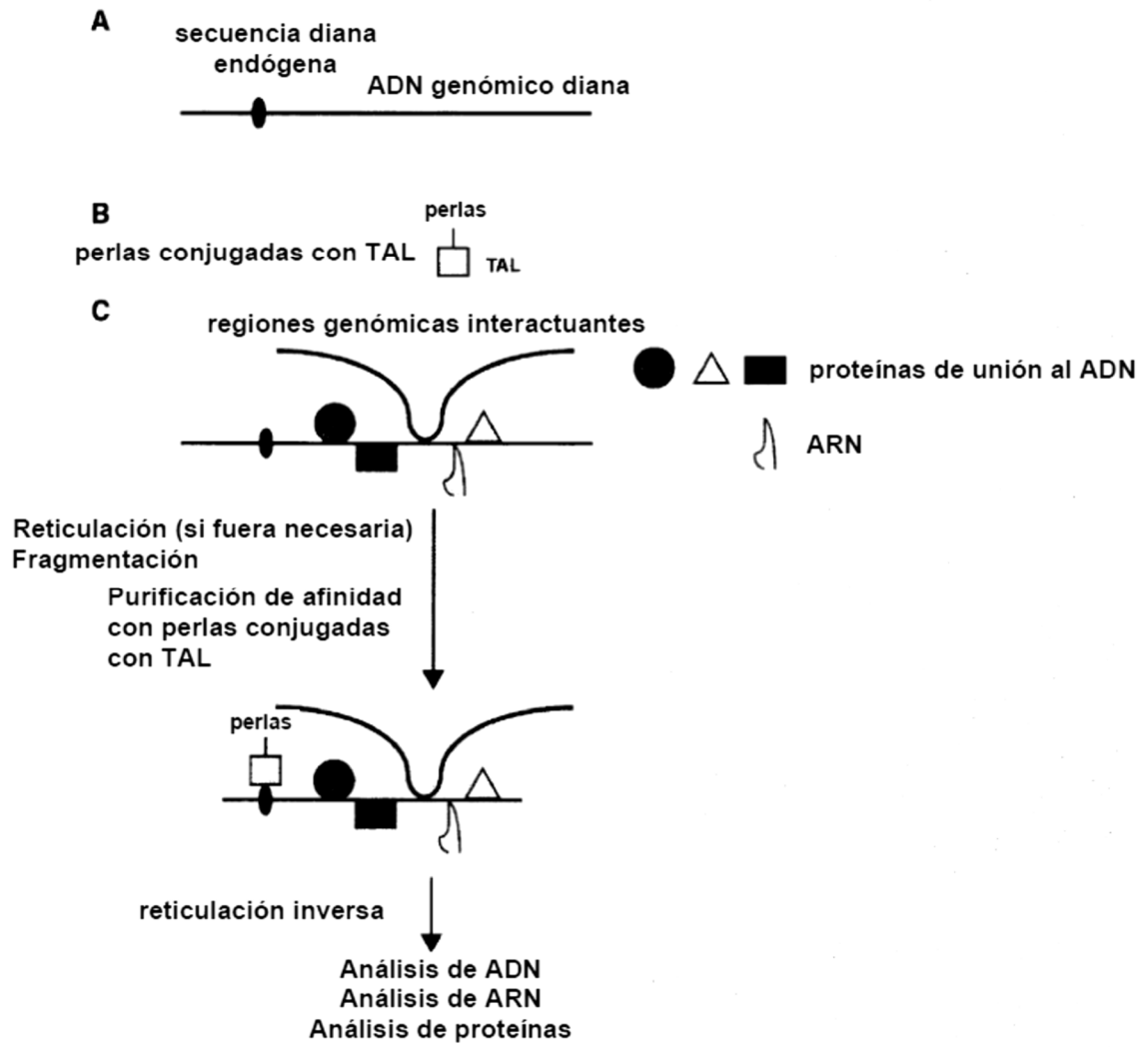


Fig. 3

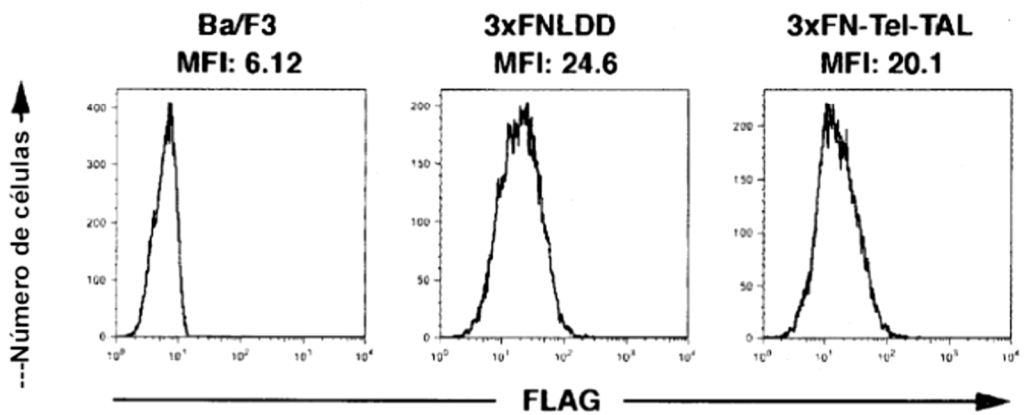


Fig. 4

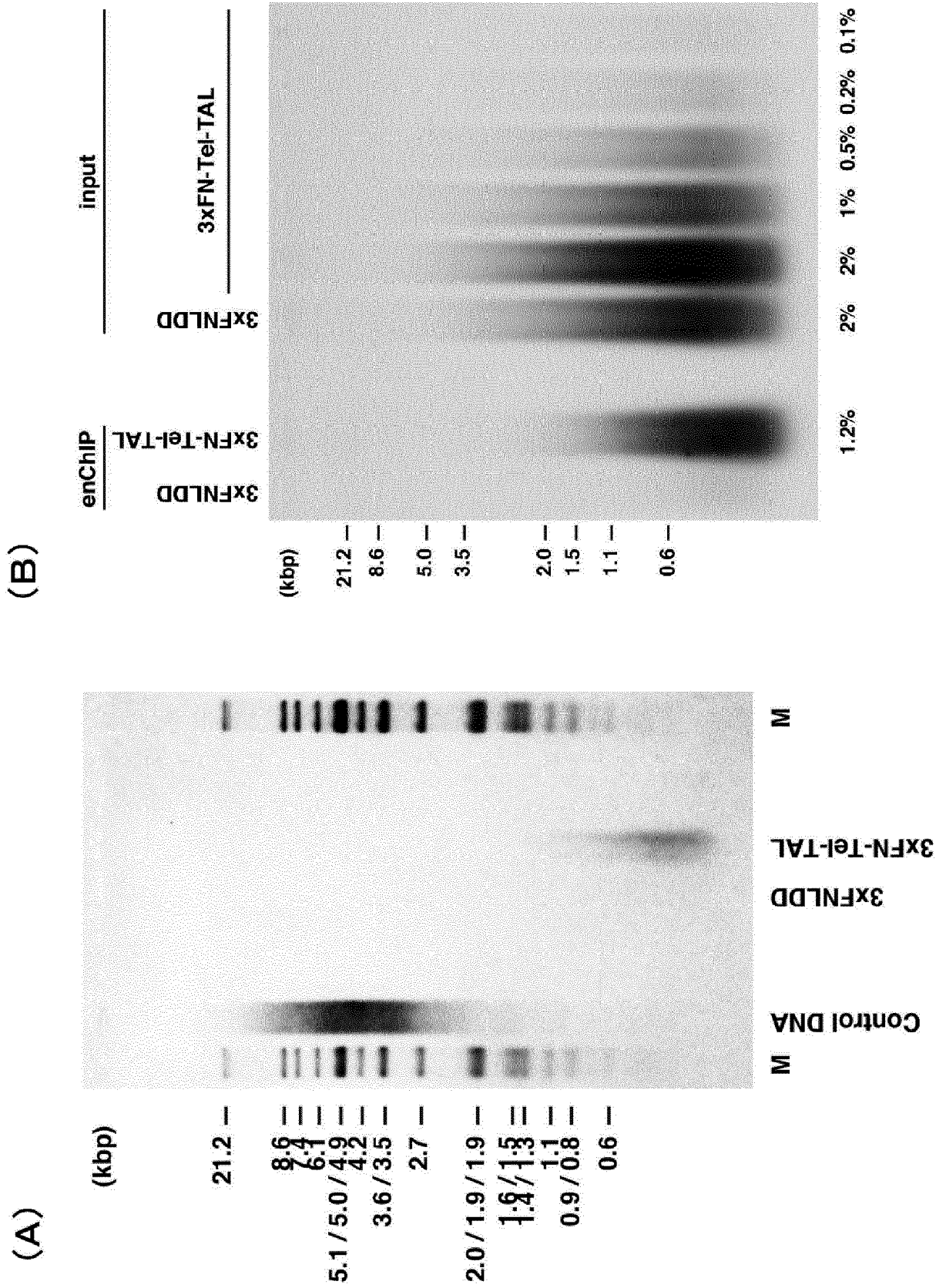


Fig. 5

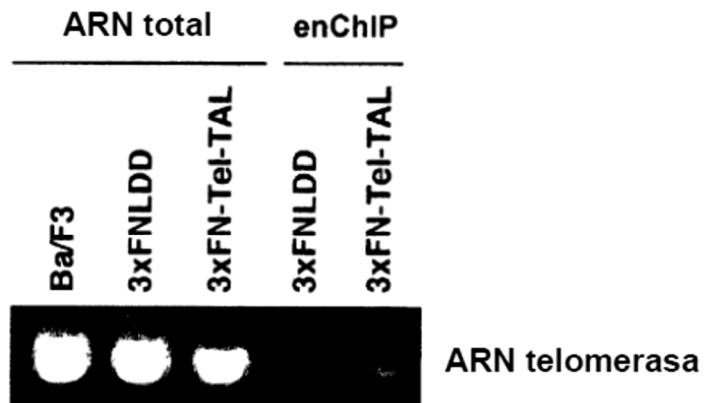


Fig. 6

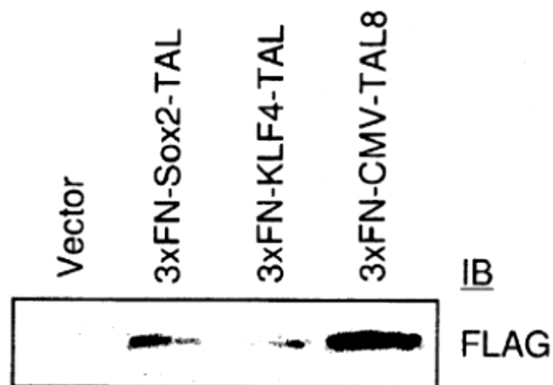


Fig. 7

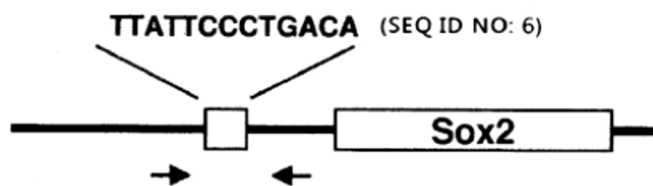


Fig. 8

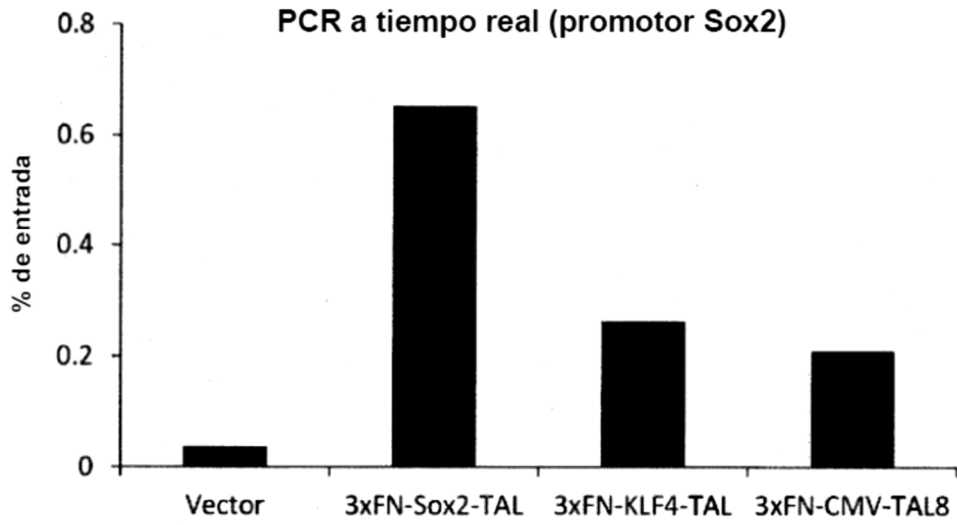


Fig. 9

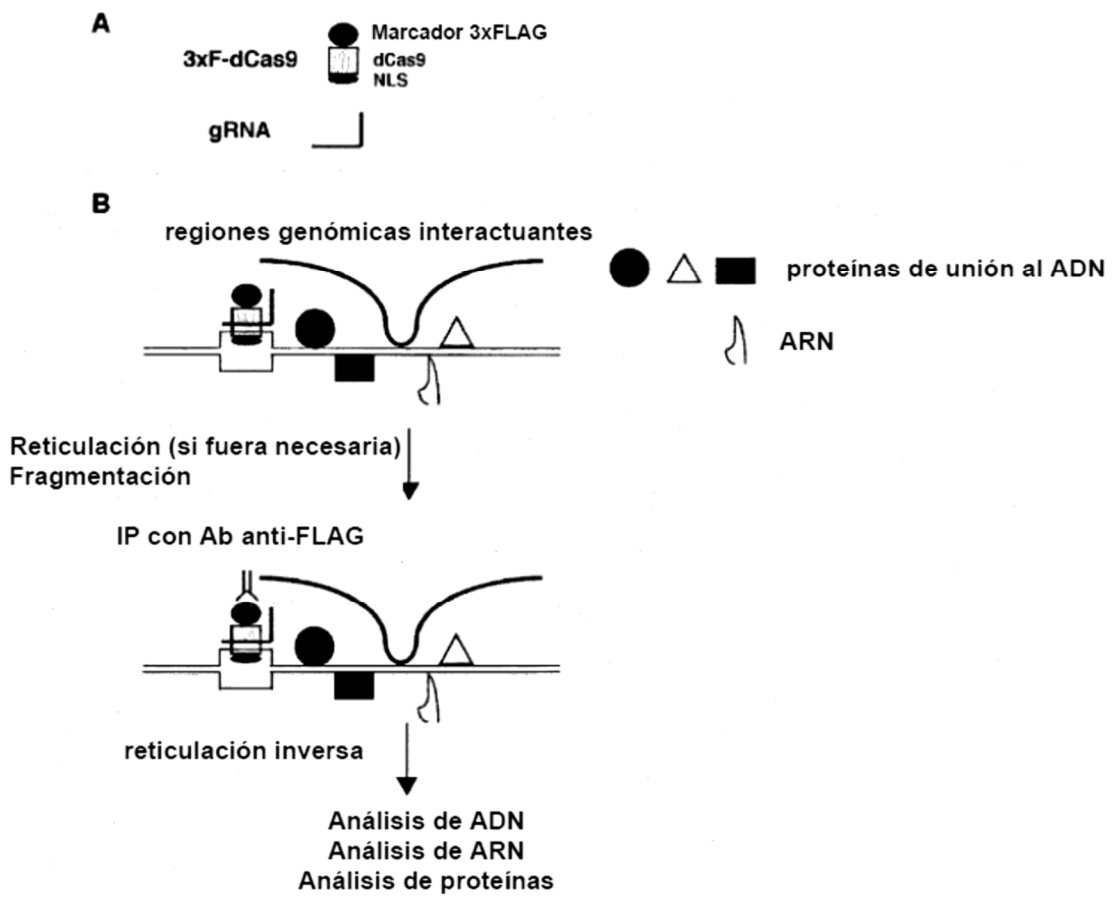


Fig. 10

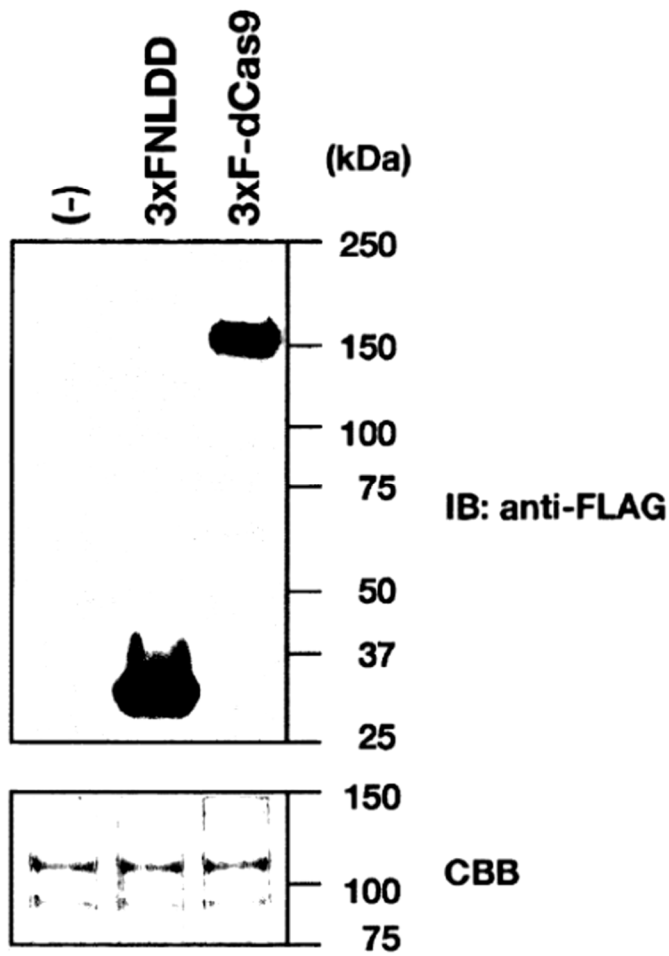


Fig. 11

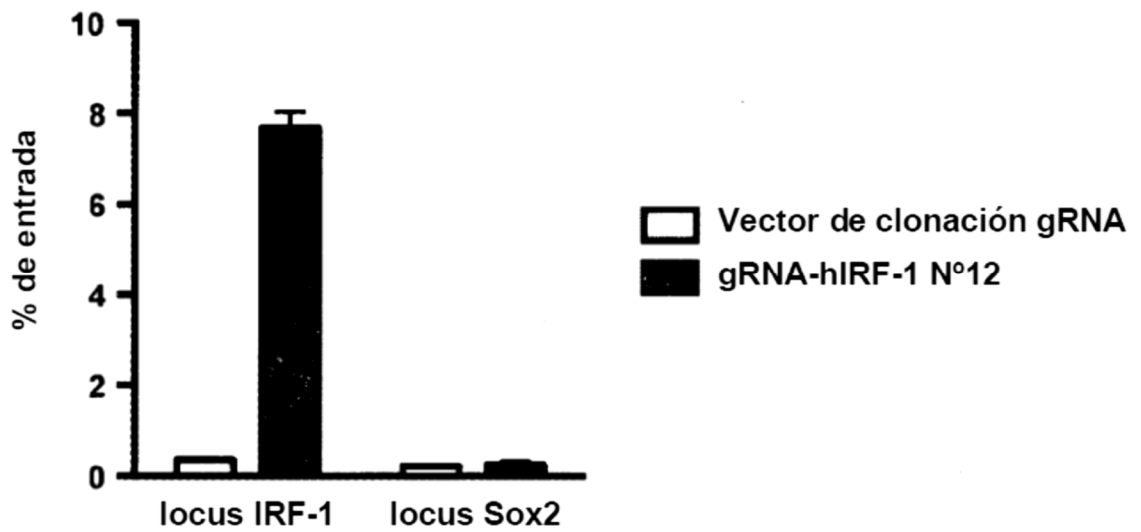


Fig. 12

