

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 556 704**

51 Int. Cl.:

**C12Q 1/68** (2006.01)

**G01N 33/53** (2006.01)

**G01N 33/566** (2006.01)

**C12N 15/113** (2010.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.12.2003 E 10075508 (1)**

97 Fecha y número de publicación de la concesión europea: **16.09.2015 EP 2316966**

54 Título: **Proteína de replicación Ciz1**

30 Prioridad:

**05.12.2002 GB 0228337**

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

**19.01.2016**

73 Titular/es:

**CIZZLE BIOTECHNOLOGY LIMITED (100.0%)  
Innovation Centre York Science Park  
Heslington, York YO10 5DG, GB**

72 Inventor/es:

**COVERLEY, DAWN**

74 Agente/Representante:

**ISERN JARA, Jorge**

**ES 2 556 704 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

## Proteína de replicación Ciz1

5 La presente memoria descriptiva se refiere a un método de detección para la identificación de agentes que modulan la actividad de una proteína de replicación del ADN como diana para la intervención en la terapia contra el cáncer e incluye agentes que modulan dicha actividad.

10 La invención también se refiere al uso de la proteína de replicación del ADN y sus transcritos de ARN en el pronóstico y diagnóstico de una enfermedad proliferativa, por ejemplo, cáncer.

## Introducción

15 El inicio de la replicación del ADN es uno de los principales puntos de control del ciclo celular en los mamíferos, además del punto de acción de muchos productos génicos que cuya regulación está alterada en el cáncer (Hanahan y Weinberg, 2000). El proceso de inicio implica el ensamblaje de proteínas del complejo de prerreplicación, que incluye las proteínas Cdc6, Cdt1 y Mcm del complejo de reconocimiento de origen (ORC), en los orígenes de replicación durante la fase G1 del ciclo celular. A continuación, actúa un segundo grupo de proteínas que facilita la carga de las ADN polimerasas y de sus factores accesorios, incluyendo el PCNA, y la transición a la fase S. El proceso de inicio se encuentra regulado por la proteína quinasa 2 dependiente de ciclina (Cdk2), Cdc7-dbf4 y el inhibidor de Cdt1 llamado geminina (para una revisión véase Bell y Dutta, 2002). En el núcleo de las células en fase S, las horquillas de replicación se agrupan formando cientos de "focos" o fábricas de replicación (Cook, 1999). Las fábricas de replicación se unen aparentemente a un marco estructural dentro del núcleo; sin embargo, la naturaleza de las moléculas que forman la unión y su papel en la actividad de las horquillas de replicación sigue sin conocerse con exactitud.

25 La identificación de las proteínas implicadas en la replicación del ADN eucariótico y el análisis de las rutas básicas que regulan su actividad durante el ciclo celular ha estado dirigida principalmente por la genética de las levaduras. Estas proteínas y rutas generalmente se encuentran conservadas en levaduras y ser humano. Sin embargo, en organismos multicelulares, que se diferencian a lo largo de diversas rutas de desarrollo, se están descubriendo capas adicionales de complejidad. Por ejemplo, en vertebrados, varias proteínas implicadas en la diferenciación neuronal también regulan la transición de fases G1 a S (Ohnuma et al., 2001). Estos incluyen el inhibidor de cdk p21<sup>CIP1/WAF1/SDI1</sup> que está implicado en la diferenciación de los oligodendrocitos tras la detención del crecimiento (Zezula et al., 2001), y en la diferenciación terminal de otros tipos celulares (Parker et al., 1995).

35 El inicio de la replicación del ADN puede reconstituirse *in vitro* con núcleos aislados y extractos citosólicos procedentes de células de mamífero (Krude, 2000; Krude et al., 1997; Laman et al., 2001; Stoeber et al., 1998). Además, utilizando Cdk2 recombinante en complejo con las ciclinas E o A, el ensamblaje del complejo de replicación y la activación de la síntesis del ADN pueden reconstituirse independientemente (Coverley et al., 2002). Los inventores han estudiado la etapa de activación, catalizada *in vitro* por ciclina A-cdk2, y han demostrado que una proteína relativamente no estudiada, la proteína de dedo de cinc de interacción con p21-Cip1 (Ciz1) funciona durante la etapa del proceso de inicio. La Ciz1 humana fue identificada anteriormente utilizando una modificación del cribado de dos híbridos de levadura con ciclina E-p21, y el análisis bioquímico corroboró la existencia de una interacción con p21 (Mitsui et al., 1999). Se propuso, aunque no demostró, un papel potencial en la transcripción, y no se atribuyó ninguna otra función a Ciz1. Más recientemente, se ha aislado el gen Ciz1 a partir de una biblioteca de ADNc derivada de meduloblastoma humano utilizando un modelo de tumorigénesis *in vivo* (Warder y Keherly, 2003). El análisis de los presentes inventores demuestra por primera vez que Ciz1 ejerce una función positiva en el inicio de la replicación del ADN.

40 Se produce una serie de cambios en las proteínas unidas a la cromatina al activarse la síntesis de ADN *in vitro* con ciclina A-cdk2 recombinante. La presente invención se refiere al resultado de que un antígeno relacionado con cdc6, p85, se correlaciona con el inicio de la replicación del ADN y de que se encuentra regulado por la ciclina A-cdk2. La proteína ha sido clonada a partir de una biblioteca embrionaria de ratón y ha sido identificada como Ciz1 de ratón.

55 El análisis *in vitro* ha demostrado que la proteína Ciz1 regula positivamente el inicio de la replicación del ADN y que su actividad se encuentra modulada por la fosforilación de cdk en la treonina 191/2, ligándola a las rutas dependientes de cdk que controlan el inicio. La forma embrionaria de Ciz1 de ratón se somete a corte y empalme alternativo, en contraste con las formas predichas y somáticas. La Ciz1 humana también se somete a corte y empalme alternativo, con variabilidad en los mismos exones que la Ciz1 de ratón. Se ha encontrado que la forma embrionaria recombinante de Ciz1 estimula el inicio de la replicación del ADN de mamífero y que los cánceres pediátricos expresan formas "de tipo embrionario" de Ciz1. Sin limitarse a ninguna teoría en particular, los inventores proponen que el corte y empalme erróneo de Ciz1 produce formas de tipo embrionario de Ciz1 en tiempos inadecuados del desarrollo. Esto estimula la replicación incorrectamente regulada del ADN y contribuye a la formación o avance de linajes de células de cáncer.

65

Se han desarrollado varias técnicas en los últimos años que pretenden eliminar específicamente genes y/o productos génicos. Por ejemplo, la utilización de moléculas de ácidos nucleicos antisentido para unirse y de esta manera bloquear o inactivar moléculas de ARNm diana es un medio efectivo para inhibir la producción de productos génicos.

5 Una técnica mucho más reciente para eliminar específicamente la función génica es a través de la introducción de ARN de doble cadena, también denominado ARN inhibidor (ARNi), en una célula, que tiene como resultado la destrucción del ARNm complementario a la secuencia incluida en la molécula de ARNi. La molécula de ARNi comprende dos cadenas complementarias de ARN (una cadena de sentido y una cadena antisentido) hibridadas entre sí formando una molécula de ARN de doble cadena. La molécula de ARNi típicamente deriva de la secuencia exónica o codificante del gen que debe anularse.

10 Tanto los ácidos nucleicos como las proteínas presentan una estructura de secuencia lineal, definida por su secuencia de bases o de aminoácidos, y también una estructura tridimensional que en parte está determinada por la secuencia lineal y también el ambiente en el que se localizan dichas moléculas. Las moléculas terapéuticas convencionales son moléculas de tamaño reducido, por ejemplo, péptidos, polipéptidos o anticuerpos, que se unen a moléculas diana, produciendo un efecto agonista o antagonista. Se ha puesto de manifiesto que las moléculas de ácidos nucleicos también presentan un potencial con respecto a la provisión de agentes con las propiedades de unión necesarias que podrían presentar utilidad terapéutica. Estas moléculas de ácidos nucleicos típicamente se denominan aptámeros. Los aptámeros son moléculas de ácidos nucleicos pequeñas, habitualmente estabilizadas, que comprenden un dominio de unión para una molécula diana.

15 Los aptámeros pueden comprender por lo menos una base nucleotídica modificada. La expresión "base nucleotídica modificada" comprende nucleótidos con una base y/o azúcar modificado covalentemente. Por ejemplo, entre los nucleótidos modificados se incluyen nucleótidos que presentan azúcares que se unen covalentemente a grupos orgánicos de bajo peso molecular diferentes de un grupo hidroxilo en la posición 3' y diferentes de un grupo fosfato en la posición 5'. De esta manera, los nucleótidos modificados también pueden incluir azúcares 2'-sustituidos, tales como 2'-O-metilo, 2'-O-alquilo, 2'-O-alilo, 2'-S-alquilo, 2'-S-alilo, 2'-fluoro, 2'-halo o 2'-azido-ribosa, análogos carbocíclicos de azúcares, azúcares a-anómericos, azúcares epiméricos tales como arabinosa, xilosas o lixosas, azúcares piranosas, azúcares furanosos y pseudoheptulosas.

20 Los nucleótidos modificados son conocidos de la técnica y comprenden, a título de ejemplo no limitativo, purinas y/o pirimidinas alquiladas, purinas y/o pirimidinas aciladas, u otros heterociclos. Estas clases de pirimidinas y purinas son conocidas de la técnica e incluyen pseudoisocitosina, N<sup>4</sup>, N<sup>4'</sup>-etanocitosina, 8-hidroxi-N<sup>6</sup>metiladenina, 4-acetilcitosina, 5-(carboxihidroximetil)uracilo, 5-fluorouracilo, 5-bromouracilo, 5-carboximetilaminometil-2-tiouracilo, 5-carboximetilaminometil-uracilo, dihidrouracilo, inosina, N<sup>6</sup>-isopentiladenina, 1-metiladenina, 1-metilpseudouracilo, 1-metilguanina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N<sup>6</sup>-metiladenina, 7-metilguanina, 5-metilaminometil-uracilo, 5-metoxi-aminometil-2-tiouracilo, β-D-manosilqueosina, 5-metoxycarbonilmetiluracilo, 5-metoxiuracilo, 2-metiltio-N<sup>6</sup>-isopentiladenina, metil-éster de ácido uracilo-5-oxiacético, pseudouracilo, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, metil-éster de ácido N-uracil-5-oxiacético, ácido uracil-5-oxiacético, queosina, 2-tiocitosina, 5-propiluracilo, 5-propilcitosina, 5-etiluracilo, 5-etilcitosina, 5-butiluracilo, 5-pentiluracilo, 5-pentilcitosina y 2,6-diaminopurina, metilpseudouracilo, 1-metilguanina y 1-metilcitosina;

45 Los aptámeros pueden sintetizarse utilizando nucleótidos convencionales unidos mediante enlaces fosfodiéster, mediante técnicas de estándar síntesis en fase sólida o de solución o en fase solución conocidas de la técnica. Los enlaces entre nucleótidos pueden utilizar moléculas de unión alternativas. Por ejemplo, grupos de unión de fórmula P(O)S, (tioato); P(S)S, (ditioato); P(O)NR'<sup>2</sup>; P(O)R'; P(O)OR<sub>6</sub>; CO; o CONR'<sup>2</sup> en los que R es H (o una sal) o alquilo (C<sub>1-12</sub>) y R<sub>6</sub> es alquilo (C<sub>1-9</sub>) unido a nucleótidos contiguos mediante -O- o -S-.

50 Otras técnicas que pretenden anular específicamente genes y/o producto génicos se centran en modular la función o interferir en la actividad de las moléculas de proteína. Las proteínas pueden ser las dianas de inhibidores químicos, obtenidos de, por ejemplo, bibliotecas existentes de moléculas pequeñas.

55 Pueden cultivarse anticuerpos, preferentemente monoclonales, por ejemplo en ratones o ratas frente a diferentes isoformas de proteína. Los anticuerpos, también conocidos como inmunoglobulinas, son moléculas de proteína que presentan especificidad para moléculas foráneas (antígenos). Las inmunoglobulinas (Ig) son una clase de proteínas relacionadas estructuralmente que consisten de dos parejas de cadenas polipeptídicas, un par de cadenas ligeras (L) (de bajo peso molecular) (K o A) y un par de cadenas pesadas (H) (γ, α, 1, 5 y E), las cuatro unidas mediante enlaces disulfuro. Tanto las cadenas H como las L presentan regiones que contribuyen a la unión de antígenos y que son muy variables de una a otra molécula de Ig. Además, las cadenas H y L contienen regiones que son no variables o constantes.

65 Las cadenas L están constituidas por dos dominios. El dominio carboxi-terminal es esencialmente idéntico en las cadenas L de un tipo dado y se hace referencia al mismo como región "constante" (C). El dominio amino-terminal

varía de una cadena L a otro y contribuye al sitio de unión del anticuerpo. Debido a su variabilidad, se denomina región "variable" (V).

Las cadenas H de las moléculas de Ig son de varias clases,  $\alpha$ ,  $\mu$ ,  $\sigma$ ,  $\delta$ , y  $\gamma$  (de la que existen varias subclases). Una molécula de Ig ensamblada está constituida por una o más unidades de dos cadenas H y L idénticas, deriva su nombre de la cadena H que presenta. De esta manera, existen cinco isotipos de Ig: IgA, IgM, IgD, IgE e IgG (con cuatro subclases basadas en las diferencias entre las cadenas H, es decir, IgG1, IgG2, IgG3 e IgG4). Puede encontrarse más información sobre la estructura del anticuerpo y sus diversas funciones en: Using Antibodies: A laboratory manual, Cold Spring Harbour Laboratory Press.

Los anticuerpos quiméricos son anticuerpos recombinantes en los que la totalidad de las regiones V de un anticuerpo de ratón o rata se combinan con regiones C de anticuerpo humano. Los anticuerpos humanizados son anticuerpos híbridos recombinantes que fusionan las regiones determinantes de complementariedad de una región V de anticuerpo de roedor con las regiones de marco de las regiones V de anticuerpo humano. Las regiones C del anticuerpo humano también se utilizan. Las regiones determinantes de LA complementariedad (CDR) son las regiones en el dominio N-terminal de las cadenas tanto pesada como ligera del anticuerpo en la que se restringe la mayor parte de la variación de la región V. Estas regiones forman bucles en la superficie de la molécula de anticuerpo. Estos bucles proporcionan la superficie de unión entre el anticuerpo y el antígeno.

Los anticuerpos procedentes de animales no humanos provocan una respuesta inmunológica frente al anticuerpo extraño y su eliminación de la circulación. Los anticuerpos tanto quiméricos como humanizados presentan una antigenicidad reducida cuando se inyectan en un sujeto humano debido que existe una cantidad reducida de anticuerpo de roedor (es decir, extraño) dentro del anticuerpo híbrido recombinante, mientras que las regiones de anticuerpo humano no inducen una respuesta inmunológica. Esto resulta en una respuesta inmunológica más débil y en una reducción de la eliminación del anticuerpo. Esto resulta claramente deseable al utilizar anticuerpos terapéuticos en el tratamiento de enfermedades humanas. Los anticuerpos humanizados están diseñados para presentar menos regiones de anticuerpo "extraño" y, por lo tanto, se cree que son menos inmunogénicas que los anticuerpos quiméricos.

Entre otras técnicas para apuntar al nivel de proteínas se incluyen la utilización de péptidos generados aleatoriamente que se unen específicamente a proteínas, y cualquier otra molécula que se una a proteínas o variantes de proteína y modifique la función de las mismas.

La comprensión del proceso de replicación del ADN resulta de especial interés en el campo de la terapia del cáncer. Es conocido que las células de cáncer pueden adquirir resistencia frente a agentes quimioterapéuticos y pueden evitar la detección por parte del sistema inmunológico. Existe una necesidad continuada de identificación de dianas para la terapia del cáncer de manera que puedan identificarse nuevos agentes. El proceso de replicación del ADN representa una diana importante para la intervención farmacológica en la terapia del cáncer. Existe una necesidad de identificar productos génicos que modulen la replicación del ADN y que contribuyan a la formación o avance de linajes de células de cáncer y de desarrollar agentes que afectan a su función.

#### Declaraciones de la invención

De acuerdo con un primer aspecto de la presente invención se proporciona un método diagnóstico *in vitro* para la identificación de un trastorno proliferativo, comprendiendo el método detectar la presencia o la expresión de una variante de corte y empalme de C1z1, en el que el método proporciona una medida cuantitativa de una variante de la proteína C1z1 en una muestra de un sujeto y la variante proteica es una secuencia de proteínas sometida a corte y empalme alternativo humana en la que la secuencia:

(a) QQLQQLQQQLQQQLQQQLLQLQQLLQQSPP, o  
MFSQQQQQLQQQQQLQQQLQQQLQQQLQQQLLQLQQLLQQSPPQA está ausente;

(b) GLDQFAMPPATYDTAGLTMPTATL está ausente;

(c) DSSSQ está ausente;

(d) PQVQQAHSQPPRQVQLQLQKQVQTQTY; o

PQVQQAHSQGPRQVOLQQEAEPLKQVQPQVQVQQAHSQPPRQVQLQLQKQVQTQTY,

o

QVQSQTQPRIPSTDTQVQPKLQKQAQTQTSPEHLVLQKQVQPQLQQEAEQKQVQVQ  
VQVQQAHSQGPRQVQLQQEAEPLKQVQPQVQVQQAHSQPPRQVQLQLQKQVQTQTY

está ausente; o

(e) VEEELCKQ está ausente.

De acuerdo con otro aspecto de la presente invención se proporciona un método diagnóstico *in vitro* para la identificación de un trastorno proliferativo, comprendiendo el método detectar la presencia o la expresión de una

variante de corte y empalme de Ci1z1, en el que el método proporciona una medida cuantitativa de una variante de ARN de Ciz1 en una muestra de un sujeto y en el que la variante de ARN de Ciz1 es una secuencia de ARN sometida a corte y empalme alternativo humana en la que la secuencia:

5 (a)

TGGGGGCTGCGGGGCCGGGCCCATCCGTGGGGGCGACTTGAGCGTTGAGGGCG  
CGCGGGGAGGCGAGCCACCATGTTTCAGCCAGCAGCAGCAGCAGCTCCAGCAACAGC  
AGCAGCAGCTCCAGCAGTTACAGCAGCAGCAGCTCCAGCAGCAGCAATTGCAGCAG  
CAGCAGTTACTGCAGCTCCAGCAGCTGCTCCAGCAGTCCCCACCACAGGCC,

o

10

CAGCAGCTCCAGCAGTTACAGCAGCAGCAGCTCCAGCAGCAGCAATTGCAGCAGCA  
GCAGTTACTGCAGCTCCAGCAGCTGCTCCAGCAGTCCCCACCACA

está ausente;

15 (b)

(b) GGACTGGACCAGTTTGAATGCCACCAGCCACGTATGACACTGCCGGTCTCACCA  
T GCCACAGCAACTG

está ausente;

20

(c) AGGATTCTTCTTCTC está ausente;

(d)

25

(d) CCACAGGTGCAGCCCCAGGCACATTCACAGCCCCAAGGCAGGTGCAGCTGCAG  
CTGCAGAAGCAGGTCCAGACACAGACATATCC,

o

CCACAGGTACAGCCACAGGCACATTCACAGGGCCCAAGGCAGGTGCAGCTGCAGCA  
GGAGGCAGAGCCGCTGAAGCAGGTGCAGCCACAGGTGCAGCCCCAGGCACATTCAC  
AGCCCCAAGGCAGGTGCAGCTGCAGCTGCAGAAGCAGGTCCAGACACAGACATAT,

30

o

CAGGTGCAGTCACAGACTCAGCCGCGGATACCATCCACAGACACCCAGGTGCAGCC  
AAAGCTGCAGAAGCAGGCGCAAACACAGACCTCTCCAGAGCACTTAGTGCTGCAAC  
AGAAGCAGGTGCAGCCACAGCTGCAGCAGGAGGCAGAGCCACAGAAGCAGGTGCA  
GCCACAGGTACAGCCACAGGCACATTCACAGGGCCCAAGGCAGGTGCAGCTGCAGC  
AGGAGGCAGAGCCGCTGAAGCAGGTGCAGCCACAGGTGCAGCCCCAGGCACATTC  
CAGCCCCAAGGCAGGTGCAGCTGCAGCTGCAGAAGCAGGTCCAGACACAGACATA  
T

35 está ausente; o

(e) GTTGAGGAGGAACTCTGCAAGCAG está ausente.

En una forma de realización, el trastorno proliferativo es cáncer.

40

En otra forma de realización, el cáncer es un cáncer pediátrico seleccionado del grupo que consiste en retinoblastoma, neuroblastoma, linfoma de Burkitt, meduloblastoma y tumores de la familia del sarcoma de Ewings.

En otra forma de realización, el cáncer es carcinoma, adenocarcinoma, linfoma o leucemia.

45

En otra forma de realización, el cáncer es cáncer de hígado, de pulmón o de piel.

En otra forma de realización, el método de diagnóstico comprende uno o más de las siguientes etapas:

- (a) poner en contacto la muestra con un agente que se une específicamente a un polipéptido con actividad Ciz1 y medir la unión del agente al polipéptido en la muestra;
- (b) usar la PCR con transcripción inversa o la PCR en tiempo real para controlar la expresión de la isoforma de Ciz1 y medir los niveles de expresión;
- (c) medir la presencia de mutaciones en el ácido nucleico o aminoácido en base a las propiedades de conformación de la molécula alteradas.

En una forma de realización, el agente que se une específicamente a un polipéptido con actividad de Ciz1 es un anticuerpo anti-Ciz1.

El anticuerpo puede ser un anticuerpo monoclonal.

También se describe en el presente documento el uso de secuencia de nucleótidos o de polipéptido de Ciz1, o cualquier fragmento o variante de la misma, como un objetivo para la identificación de agentes que modulan la replicación del ADN.

Tal como se utiliza en el presente documento, el término "fragmento" o "variante" se utiliza para hacer referencia a cualquier secuencia de ácido nucleico o de aminoácidos que deriva de la secuencia de nucleótidos o de aminoácidos de longitud completa de Ciz1 o deriva de una variante de corte y empalme de la misma. En una realización, el fragmento es de longitud suficiente y / o de suficiente homología con la Ciz1 de longitud completa para retener la actividad de replicación de ADN de Ciz1. En una realización alternativa se utilizan fragmentos de Ciz1 inactivos. El término "fragmento" o "variante" se refiere también a transcritos de ARN de Ciz1 descritos en el presente documento y a isoformas de proteínas (o partes de los mismos).

Tal como se utiliza en el presente documento, el término "modular" se utiliza para hacer referencia a aumentar o disminuir la replicación del ADN, por encima y por debajo de los niveles que normalmente se observa en ausencia del agente específico (es decir, cualquier alteración en la actividad de replicación de ADN que está directamente o indirectamente relacionados con el uso del agente). El término "modular" también incluye la referencia a un cambio de organización espacial o temporal de la replicación del ADN.

También se describe en el presente documento un método de detección selectiva para la identificación de agentes que modulan la replicación del ADN en el que el método de detección selectiva comprende el uso de secuencias de nucleótidos o de polipéptidos de Ciz1 o fragmentos o variantes de los mismos.

El método de detección selectiva puede comprender detectar o medir el efecto de un agente en una molécula de ácido nucleico seleccionada de los grupos que consisten en:

- a) una molécula de ácido nucleico que comprende una secuencia de ácido nucleico representada en cualquiera de las Figuras 14, 15, o 21;
- b) una molécula de ácido nucleico que hibrida con la secuencia de ácido nucleico de (a) y que tiene una actividad de Ciz1 o actividad de una variante de la misma;
- c) una molécula de ácido nucleico que tiene una secuencia de ácido nucleico que es degenerada debido al código genético con las secuencias en a) y b) y un agente candidato que se va a analizar;
- d) una molécula de ácido nucleico derivado de la secuencia genómica en el locus Ciz1 o una molécula de ácido nucleico que hibrida con la secuencia genómica.

La molécula de ácido nucleico puede modificarse por delección, sustitución o adición de al menos un residuo de ácido nucleico de la secuencia de ácido nucleico.

Como alternativa, el método de detección selectiva comprende las etapas de:

(i) formar una preparación que comprende una molécula de polipéptido, o un fragmento activo del mismo, codificado por una molécula de ácido nucleico seleccionada del grupo que consiste en:

- a) una molécula de ácido nucleico que comprende una secuencia de ácido nucleico representada en las Figuras 14, 15, o 21;
- b) una molécula de ácido nucleico que hibrida con la secuencia de ácido nucleico de (a) y que tiene una actividad de Ciz1 o actividad de una variante de la misma;
- c) una molécula de ácido nucleico que tiene una secuencia de ácido nucleico que es degenerada debido al código genético con las secuencias en a) y b) y un agente candidato que se va a analizar;
- d) una molécula de ácido nucleico derivado de la secuencia genómica en el locus Ciz1 o una molécula de ácido nucleico que hibrida con la secuencia genómica; y

(ii) detectar o medir el efecto del agente sobre la actividad de dicho polipéptido.

Los ensayos para la detección de la replicación del ADN son conocidos en la técnica. La actividad residente en Ciz1, o fragmentos de péptidos derivados, y el efecto de los potenciales agentes terapéuticos sobre dicha actividad se analizaría *in vitro* o *in vivo*.

5 Los ensayos *in vitro* para la actividad de la proteína Ciz1 comprendería núcleos en fase G1 aislados sincronizados y, extracto en fase S o extracto en fase G1 complementados con quinasas dependientes de ciclina. La inclusión de Ciz1 o fragmentos peptídicos derivados estimula la iniciación de la replicación de ADN en estas circunstancias y se puede monitorizar visualmente (anotando los núcleos que han incorporado los nucleótidos fluorescentes durante las reacciones *in vitro*) o midiendo la incorporación de nucleótidos radiactivos. El ensayo para reactivos terapéuticos que  
10 interfieren con la función de la proteína Ciz1 implicaría buscar la inhibición de la replicación del ADN en estos ensayos. El efecto de los agentes en la localización nuclear de Ciz1, la unión a la cromatina, la estabilidad, la modificación y las interacciones proteína-proteína también podría monitorizarse en estos ensayos.

15 Los ensayos *in vivo* incluirán la creación de modelos celulares y de ratón que sobreexpresan subexpresan Ciz1 o fragmentos derivados, lo que da lugar a la proliferación celular alterada. La preparación de animales transgénicos se conoce generalmente en la técnica y dentro del ámbito de la persona experta. El ensayo para reactivos terapéuticos implicaría análisis del tiempo del ciclo celular, la iniciación de la replicación del ADN y la incidencia de cáncer en presencia y ausencia de fármacos que afectan a la actividad de la proteína Ciz1, o interfieren con la producción de Ciz1 apuntando a Ciz1 y a sus variantes a nivel de ARN.

20 Las condiciones de hibridación pueden ser rigurosas.

25 Las condiciones de hibridación/lavado rigurosas son bien conocidas en la técnica. Por ejemplo, los híbridos de ácido nucleico que son estables después de lavar en 0,1 x SSC, 0,1 % de SDS a 60 °C. En la técnica se sabe bien que las condiciones de hibridación óptimas pueden calcularse si se conoce la secuencia del ácido nucleico. Típicamente, las condiciones de hibridación utilizan 4 - 6 x SSPE (20 x SSPE contiene 175,3 g de NaCl, 88,2 g de NaH<sub>2</sub>PO<sub>4</sub>, H<sub>2</sub>O y 7,4 g de EDTA disuelto en 1 litro y el pH ajustado a 7,4); la solución 10 x Denhardtts (50 x solución Denhardtts contiene 5 g de Ficoll (Tipo 400, Pharmacia), 5 g de polivinilpirrolidona y 5 g de seroalbúmina bovina; 100 µg – 1,0 mg/ml de ADN salmón / arenque sometidos a ultrasonidos; 0,1 – 1,0 % de dodecil sulfato de sodio; opcionalmente  
30 40 – 60 % de formamida desionizada. La temperatura de hibridación variará dependiendo del contenido de GC de la secuencia de ácido nucleico diana, pero será típicamente entre 42– 65 °C.

35 Dicho polipéptido puede modificarse por delección, sustitución o adición de al menos un residuo de aminoácido de la secuencia de polipéptido.

Un modificado o variante, es decir un fragmento de polipéptido y polipéptido de referencia, se puede diferenciar en la secuencia de aminoácidos en una o más sustituciones, adiciones, delecciones, fusiones y truncamientos, que pueden estar presentes en cualquier combinación. Entre las variantes preferidas se encuentran aquellas que varían de un polipéptido de referencia por sustituciones de aminoácidos conservadoras. Dichas sustituciones conservadoras son  
40 aquellas que sustituyen un aminoácido dado por otro aminoácido de características similares. La siguiente lista de aminoácidos no limitante se considera reemplazos conservativos (similares): a) alanina, serina, y treonina; b) ácido glutámico y ácido aspártico; c) asparagina y glutamina d) arginina y lisina; y) isoleucina, leucina, metionina y valina y f) fenilalanina, tirosina y triptófano. Se prefieren las variantes que retienen la misma función biológica y actividad que el polipéptido de referencia del que varía. Como alternativa, las variantes incluyen aquellos con una función biológica alterada, por ejemplo, variantes que actúan como antagonistas, llamadas variantes "dominantes negativas".  
45

50 Como alternativa, o además, las sustituciones no conservadoras pueden dar la actividad biológica deseada, véase Cain SA, Williams DM, Harris V, Monk PN. Selection of novel ligands from a whole-molecule randomly mutated C5a library. Protein Eng. 2001 Mar, 14(3):189 – 93.

Una secuencia polipeptídica funcionalmente equivalente como se describe en el presente documento es una variante en la que uno o más restos de aminoácidos están sustituidos con restos de aminoácidos conservados o no conservados, o una en la que uno o más restos de aminoácidos incluye un grupo sustituyente. Las sustituciones conservadoras son los reemplazos, uno por otro, entre los aminoácidos alifáticos Ala, Val, Leu e Ile; intercambio de  
55 los restos hidroxilo Ser y Thr, intercambio de los restos ácidos Asp y Glu; sustitución entre los restos amida Asn y Gln; intercambio de los restos básicos Lys y Arg; y reemplazos entre los restos aromáticos Phe y Tyr.

60 En el presente documento se describen secuencias de nucleótidos o de polipéptidos que tiene al menos un 50 % de identidad con las secuencias de nucleótidos o polipéptidos descritas en el presente documento, o fragmentos y polipéptidos funcionalmente equivalentes de los mismos. En una realización, las secuencias de nucleótidos o de polipéptidos tienen al menos de un 75 % a 85 % de identidad, más preferiblemente al menos un 90 % de identidad, incluso más preferiblemente al menos un 95 % de identidad, todavía más preferiblemente al menos un 97 % de identidad, y más preferiblemente al menos un 99 % de identidad con las secuencias de nucleótidos y de aminoácidos ilustradas en el presente documento.  
65

En un método preferido de la invención, dicha molécula de ácido nucleico comprende la secuencia de ácido nucleico que codifica la secuencia de aminoácidos de Ciz1 descrita en las figuras 20A y 20B. En un método preferido adicional de la invención, dicha molécula de ácido nucleico consiste en la secuencia de ácido nucleico que codifica la secuencia de aminoácidos de Ciz1 descrita en las figuras 20A y 20B.

5 En un método preferido adicional de la invención, dicha molécula de polipéptido comprende la secuencia de aminoácidos de Ciz1 descrita en las figuras 20A y 20B. En un método preferido adicional de la invención, dicha molécula de polipéptido consiste en la secuencia de aminoácidos de Ciz1 descrita en las figuras 20A y 20B.

10 Dicho polipéptido puede ser expresado por una célula, preferiblemente una célula de mamífero, o animal y dicho método de detección selectiva un método de detección selectiva basado en células.

Dicha célula puede expresar naturalmente el polipéptido Ciz 1. Como alternativa, dicha célula se transfecta con una molécula de ácido nucleico que codifica un polipéptido de Ciz 1 (o una molécula variante del mismo, que se encuentra, por ejemplo, en linajes celulares de cáncer).

15 En el presente documento se describe un agente que puede obtenerse por el método descrito en el presente documento.

20 Dicho agente puede ser un antagonista de la replicación del ADN de Ciz1 medicada. Como alternativa, dicho agente puede ser un antagonista de la replicación del ADN de Ciz1 medicada.

En un método preferido adicional descrito en el presente documento, dicho agente se selecciona del grupo que consiste en: polipéptido; péptido; aptámero; sustancia química; anticuerpo; ácido nucleico; o polipéptido o una sonda nucleotídica.

25 Preferiblemente, el agente comprende una secuencia que es complementaria o de homología suficiente para dar de unión específica a la diana y puede usarse para detectar el nivel de ácido nucleico o proteína con fines de diagnóstico.

30 El agente identificado por el método descrito en el presente documento puede ser un agente terapéutico y se puede utilizar para el tratamiento de la enfermedad.

35 En una realización de la invención, el agente es una molécula de anticuerpo y se une a cualquiera de las secuencias representadas por las figuras 16, 17 o 20.

Preferiblemente, dicho anticuerpo es un anticuerpo monoclonal.

40 En el presente documento se describe un agente que es una molécula de ácido nucleico antisentido que se une a y, de este modo, bloquea o inactiva el ARNm codificado por cualquiera de las secuencias de ácido nucleico de (i) anteriores.

45 En una realización alternativa, dicho agente es una molécula de ARNi y comprende dos cadenas complementarias de ARN (una hebra sentido y una hebra antisentido) hibridadas entre sí para formar una molécula de ARN de doble cadena. Preferiblemente, la molécula de ARNi deriva de la secuencia exónica del gen de Ciz1 o de otro gen solapante.

50 El ARNm no sometido a corte y empalme puede ser el objetivo del ARNi para inhibir la producción de la variante de corte y empalme o el ARNm variante de corte y empalme puede eliminarse sin afectar al ARNm no variante.

Dicho péptido puede ser un péptido modificado.

55 Será evidente para un experto en la técnica que los aminoácidos modificados incluyen, a modo de ejemplo y no a modo de limitación, 4-hidroxiprolina, 5-hidroxilisina, N<sup>6</sup>acetyllysine, N<sup>6</sup>-metilisina, N<sup>6</sup>, N<sup>6</sup>-dimetilisina, N<sup>6</sup>, N<sup>6</sup> N<sup>6</sup>-trimetilisina, ciclohexialanina, D-aminoácidos, ornitina. Otras modificaciones incluyen aminoácidos con un grupo R alquilo C<sub>2</sub>, C<sub>3</sub> o C<sub>4</sub> opcionalmente sustituido con 1, 2 o 3 sustituyentes seleccionados de halo (por ejemplo, F, Br, I), hidroxilo o alcoxi C<sub>1</sub>-C<sub>4</sub>.

60 Como alternativa, dicho péptido se modifica mediante acetilación y / o amidación.

Los polipéptidos o péptidos pueden modificarse por ciclación. La ciclación se conoce en la técnica, (véase Scott et al Chem Biol (2001), 8:801 – 815; Gellerman et al J. Peptide Res (2001), 57: 277 – 291; Dutta et al J. Peptide Res (2000), 8: 398 – 412; Ngoka y Gross J Amer Soc Mass Spec (1999), 10:360 – 363).

También se describe en el presente documento un vector como medio de liberación para, por ejemplo, un antisentido o una molécula de ARNi que inhibe Ciz1 o variantes de la misma y, de este modo, permite el uso como objetivo de células que expresan la proteína diana.

5 Un vector viral puede ser usado como medio de liberación.  
Preferiblemente, los vectores que incluyen el casete de expresión están adaptados para la expresión de genes eucariotas. Típicamente, dicha adaptación incluye, por ejemplo y no a modo de limitación, la provisión de secuencias de control de la transcripción (secuencias promotoras) que median la expresión específica de célula / tejido. Estas secuencias promotoras pueden ser específicas de célula / tejido, inducibles o constitutivas.

10 Los elementos promotores típicamente incluyen también las denominadas caja ATA y secuencias de selección de iniciación de la ARN polimerasa, que funcionan seleccionando un sitio de iniciación de la transcripción. Estas secuencias también se unen polipéptidos que funcionan, entre otras cosas, para facilitar la selección de la iniciación de la transcripción por la ARN polimerasa.

15 Las adaptaciones también incluyen la provisión de marcadores seleccionables y secuencias de replicación autónoma que facilitan el mantenimiento de dicho vector en o la célula eucariota o huésped procarionta. Los vectores que se mantienen autónomamente se conocen como vectores episomales. Otras adaptaciones que facilitan la expresión de genes codificados en el vector incluyen la provisión de secuencias de terminación de la transcripción.

20 Estas adaptaciones son bien conocidos en la técnica. Hay una cantidad significativa de bibliografía publicada con respecto a la construcción del vector de expresión y técnicas de ADN recombinante en general. Véase, por ejemplo, Sambrook et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY y las referencias en el mismo; Marston, F (1987) DNA Cloning Techniques: A Practical Approach Vol III IRL Press, Oxford UK; DNA Cloning: F M Ausubel et al, Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (1994).

25 Según la presente invención, se proporciona un método de diagnóstico para la identificación de trastornos proliferativos que comprende la detección de la presencia o expresión del gen de Ciz1, variantes de corte y empalme de Ciz1 y mutaciones en la secuencia genómica o de proteínas del mismo.

30 El método de diagnóstico de la invención puede proporcionar una medida cuantitativa de ARN de Ciz1 o variantes proteicas en una muestra.

35 También se describe en el presente documento el uso de un agente que modula el ADN de Ciz1 o de la proteína, o variantes del mismo, como un producto farmacéutico.

Dicha sustancia farmacéutica puede comprender un agente descrito en el presente documento en combinación o asociación con un vehículo, excipiente o diluyente farmacéuticamente aceptable.

40 Preferiblemente, dicha sustancia farmacéutica es para administración oral o tópica o para administración por inyección. En una forma de realización alternativa, la sustancia farmacéutica se administra en forma de aerosol.

45 También se describe en el presente documento el uso de un agente descrito en el presente documento para la fabricación de un medicamento para uso en el tratamiento de enfermedades proliferativas. Preferiblemente, dicha enfermedad proliferativa es cáncer.

50 Preferiblemente, dicho cáncer es un cáncer pediátrico seleccionado de entre el grupo que consiste en retinoblastoma, neuroblastoma, linfoma de Burkett, meduloblastoma y tumores de la familia del sarcoma de Ewing (ESFT).

En otra forma de realización alternativa, el cáncer es carcinoma, adenocarcinoma, linfoma o leucemia.

55 En otra forma de realización alternativa, la enfermedad es cáncer de hígado, de pulmón o de piel o metástasis.

También se describe en el presente documento un método para tratar una enfermedad proliferativa que comprende administrar a un animal, preferiblemente un ser humano, un agente que puede obtenerse por el método descrito en el presente documento.

60 También se describe en el presente documento el uso de un agente descrito en el presente documento para la fabricación de un medicamento para ralentizar la división o el crecimiento celular.

65 En el presente documento se describe el uso de la secuencia de aminoácidos de Ciz1 y la estructura proteica en el diseño racional de fármacos y el uso de secuencias de nucleótidos y aminoácidos de Ciz1 de los mismos o variantes de los mismos para la detección selectiva de bibliotecas químicas de agentes que se unen específicamente a Ciz1.

En el presente documento también se describe un kit que comprende un agente de diagnóstico, pronóstico o terapéutico identificado por el método descrito en el presente documento.

En una realización alternativa se usa un chip de secuenciación basado en una matriz para la detección de Ciz1 alterada.

Descripción detallada

A continuación se describirá una realización de la presente a modo de ejemplo únicamente en referencia a las figuras siguientes:

La Fig. 1 ilustra el efecto de la ciclina A-cdk2 sobre los núcleos en fase G1 tardía. A) El anticuerpo anti-Cdc6 V1 detecta el Cdc6 de ratón y un segundo antígeno en transferencias Western de extracto de células completas 3T3, que migra con una Mr aproximada de 100 kDa (según la movilidad de la proteína Mcm3, se estimó previamente en un valor más próximo a 85 kDa, de manera que el antígeno se denominó p85; los inventores han mantenido el mismo nombre en la presente memoria en aras de la claridad). P85 se encuentra presente tanto en la fracción soluble como en la fracción nuclear insoluble (preparada para condiciones de replicación *in vitro*). B) Inicio de la síntesis del ADN en núcleos en fase G1 "competentes para la replicación" con extracto en fase G1 suplementado con ciclina A-cdk2 recombinante. La barra control muestra la proporción de núcleos ya en fase S (no sombreada), y aquellas que han iniciado la replicación en el extracto de células en fase S (sombreada). C) Tras 15 minutos en condiciones de replicación sin células, se lavaron los núcleos y se aisló nuevamente la fracción de cromatina y se separó mediante SDS-PAGE y se transfirió para identificar Mcm2 y Mcm3. D) Los mismos núcleos transferidos con el anticuerpo V1. El antígeno p85 es más abundante en núcleos expuestos a concentraciones de ciclina A-cdk2 que inducen el inicio. Se utilizó anticuerpo V1 para clonar el gen de p85 a partir de una biblioteca de expresión de embriones de ratón que se identificó como Ciz1.

Fig. 2. Alineación de variantes de Ciz1 de ratón. La secuencia de aminoácidos de Ciz1 de longitud completa predicha ("Completa") es idéntica a un clon de ADNc de tumor mamario de ratón (BC018483), mientras que la Ciz1 embrionaria ("ECiz1", AJ575057) y un clon derivado de melanoma (AK089986) no presentan dos secuencias internas discretas. Además, la primera metionina disponible en ECiz1 se encuentra en la parte central del exón 3 (Met84), que excluye una región rica en poliglutamina del extremo N-terminal. El AK089986 derivado de melanoma puede encontrarse incompleto, ya que finaliza 77 codones antes del extremo C-terminal de todos los demás clones de ratón y humanos. Los asteriscos indican los aminoácidos modificados mediante mutagénesis dirigida a sitio en las construcciones mostradas en D. Los aminoácidos que corresponden a los codones diana de los ARNip se encuentran subrayados. B) La Ciz1 de ratón se encuentra codificada por, como mínimo, 17 exones. Los exones codificantes se muestran en gris, las regiones de corte y empalme alternativo se muestran en negro y las regiones no traducidas, en blanco. Se incluyen dos secuencias del exón 1 alternativas en algunos transcritos de Ciz1 (no representadas), pero todavía no se ha encontrado un sitio de inicio traduccional alternativo cadena arriba de los dos ilustrados en la presente memoria. C) Características de secuencia y dominios putativos en ECiz1. Se muestran la secuencia predicha de localización nuclear (NLS), los sitios putativos de fosforilación de quinasa dependiente de ciclina, los dedos de cinc de tipo C2H2 y un dominio C-terminal con homología respecto a la proteína de matriz nuclear matrina 3 (Nakayasu y Berezney, 1991). Las posiciones de las secuencias ausentes de ECiz1 se indican mediante triángulos. D) ECiz1 y mutantes derivados por truncamiento y puntuales utilizados en los experimentos de replicación de ADN libres de células. Los números entre paréntesis se refieren a las posiciones de aminoácidos en la forma de longitud completa de la Ciz1 de ratón, mostrada en A. Los asteriscos indican los sitios putativos de fosforilación anulados mediante mutagénesis dirigida a sitio.

La Fig. 3. muestra los efectos de la proteína Ciz1 y fragmentos derivados en experimentos de replicación del ADN libre de células e ilustra que ECiz1 estimula el inicio de la replicación del ADN en mamíferos. A) La ECiz1 recombinante estimula el inicio de la replicación del ADN en núcleos en fase G1 "competentes para la replicación" durante la incubación en extracto en fase S. El histograma muestra el número promedio de núcleos que incorporaron nucleótidos biotinilados *in vitro* (en negro), en presencia o en ausencia de ECiz1 ectópico, con desviaciones estándar calculadas a partir de cuatro experimentos independientes. El 17 % de los núcleos que ya se encontraban en fase S al crear la preparación nuclear se muestran en blanco. Las imágenes muestran núcleos en replicación *in vitro*, con o sin ECiz1 1 nM. Todos los núcleos se contratiñeron con yoduro de propidio (rojo). B) La respuesta a ECiz1 recombinante era dependiente de la concentración, con un óptimo brusco en el rango nM. En el presente experimento, y en todos aquellos mostrados en B-1, los resultados se expresan como % del inicio y no como % de replicación. Esto se calcula a partir del número de núcleos con inicio *in vitro* y del número de núcleos que son "competentes" para el inicio *in vitro* (véase la sección de métodos). C) Las treoninas 191/2 están implicadas en la regulación de la actividad de replicación del ADN de Ciz1, ya que el mutante T (191/2) A del sitio cdk de ECiz1 escapaba a la supresión a concentraciones elevadas. D) El mutante T (293) A del sitio de Cdk estimula el inicio con un perfil similar a ECiz1 pero a concentraciones más bajas. E) ECiz1 truncado (Nterm 442) carece de las secuencias C-terminales, pero estimula el inicio *in vitro* en un grado similar a ECiz1 F) Cterm 274 no conservaba ninguna actividad de replicación del ADN en el presente ensayo. G, H, I) El análisis adicional de delección en los dos tercios N-terminales de la proteína ECiz1 mostraba que una región corta situada en 3' respecto al exón 8 resultaba necesaria para la función de Ciz1 al someterla a un ensayo *in vitro*.

Fig. 4 Caracterización de anticuerpos policlonales anti-Ciz1 e identificación de las bandas relacionadas con Ciz1 de 125 kDa. A) Gel de SDS-poliacrilamida teñido con Coomassie que muestra el fragmento de ECiz1 recombinante purificado Nterm442, y transferencias de tipo Western de Nterm442 recombinante utilizando el anticuerpo V1 anti-

Cdc6 y los anticuerpos 1793 y 1794 anti-Ciz1. B) Transferencia de tipo Western de extracto de células 3T3 completas. De las dos bandas detectadas por el anticuerpo 1793 anti-Ciz1, una presentaba la misma movilidad que p85-Ciz1 (100 kDa) reconocida por el anticuerpo Z1 y la otra presentaba una Mr aparente de 125 kDa. El anticuerpo 1794 anti-Ciz1 reconoce únicamente la forma de 125 kDa de Ciz1 (y un segundo antígeno de aproximadamente 80 kDa). C) Inmunoprecipitación a partir de extracto nuclear de células 3T3 utilizando anticuerpo V1 o 1793 anti-Ciz1. Ambos anticuerpos precipitaron p85, que se reconoce por el anticuerpo recíproco en transferencias de tipo Western. P125 precipita por acción del anticuerpo 1793, y en menor grado por el anticuerpo V1, y estos son reconocidos por 1793 en las transferencias de tipo Western. Mcm3 se muestra como control.

Fig. 5 Análisis de inmunofluorescencia de Ciz1 endógena. La Ciz1 reside en focos subnucleares que se solapan con sitios de replicación del ADN. A) Ciz1 endógena (en rojo) en células 3T3 fijadas antes (sin tratar) o después (tratadas con detergente) de la exposición a Triton X100, detectadas con anticuerpo 1793 anti-Ciz1. Los núcleos se contratiñeron con Hoescht 33258 (azul). A título comparativo se muestra Cdc6 (en verde) detectado con anticuerpo monoclonal específico para Cdc6. B) La inclusión de la Ciz1 recombinante bloquea la reactividad del anticuerpo 1793 con los núcleos tratados con detergente. C) La Ciz1 resistente a detergente (en rojo) se encontraba presente en todos los núcleos en poblaciones ciclantes, mientras que el PCNA resistente a detergente (en verde) persistía únicamente en núcleos en fase S. D) Secciones confocales a alto aumento de Ciz1 y PCNA resistentes a detergente, e imagen fusionada que muestra la colocalización de los focos (en amarillo). E) Gráficos de línea de fluorescencias roja y verde de la imagen fusionada en D, en las posiciones indicadas (i e ii). F) Gráfico de correlaciones cruzadas (Rubbi y Milner, 2000; van Steensel et al., 1996) para focos verdes en comparación con los rojos en la totalidad de la imagen fusionada en D, y (gráfico insertado) para la sección señalada tras la búsqueda de valores umbral a los niveles mostrados en Eii. La línea roja en el gráfico insertado en F muestra la pérdida de correlación al rotar 90° la imagen de Ciz1 con respecto a PCNA. La barra es 10  $\mu$ m.

Fig. 6. Interferencia de ARN. La depleción de Ciz1 inhibe la fase S. A) ARNip dirigidos a los transcritos de Ciz1 en cuatro sitios (véase la fig. 2A) se aplicaron individualmente a células 3T3 ciclantes en forma de una sola dosis de 3 nM y se realizó un seguimiento del número de células en los tiempos indicados. Se muestran imágenes de poblaciones celulares a las 16 y 40 horas tras la transfección con ARNip 8 (silueta roja) o células tratadas de forma simulada (silueta azul). B) Proteína Ciz1 detectada con 1793 anti-Ciz1 (en verde) 48 horas después de la exposición a ARNip de Ciz1 (4 y 8) o ARNip de GAPDH de control. C) Niveles de transcritos de Ciz1, GAPDH y  $\alpha$ -actina en células expuestas durante 24 horas a ARNip de Ciz1 (4 y 8). Los números entre paréntesis reflejan la intensidad de las bandas en unidades arbitrarias, y la reducción global de los transcritos de Ciz1 y GAPDH (normalizados frente a la  $\alpha$ -actina) se expresa en porcentaje. D) La proporción de células que incorporaron BrdU en el ADN (en verde) se redujo significativamente en las células con niveles reducidos de Ciz1 48 horas después del tratamiento con ARNip de Ciz1. El histograma muestra los resultados promedio de cuatro experimentos independientes. E) El número de núcleos con Mcm3 resistente a detergente (en verde) se incrementa en poblaciones tratadas con ARNip de Ciz1. F) La proporción de núcleos con PCNA resistente a detergente (en verde) también se incrementa bajo estas condiciones. Todos los núcleos se contratiñeron y se muestran en color falso (rojo).

Fig. 7. Análisis de RT-PCR de la expresión de una variante de corte y empalme alternativo de los exones 3/4 de Ciz1 en células germinales primordiales de ratón y en células madre embrionarias. Los exones 3 y/o 4 se someten a corte y empalme alternativo en estos tipos celulares, pero no en el corazón neonatal. Estos datos son consistentes con la hipótesis de que la Ciz1 de longitud completa es la forma predominante en el tejido somático neonatal, y que las variantes se observan más frecuentemente antes durante el desarrollo y en tejidos de la línea germinal.

Fig. 8. Transfección transitoria de células 3T3 de ratón. A. Se transfectaron construcciones de Ciz1 marcadas con GFP en células NIH3T3, o B. Se microinyectaron en el pronúcleo macho de huevos de ratón fertilizados en la etapa unicelular. A las 24 horas, Ciz1 y ECiz1 se habían localizado en el núcleo formando un patrón moteado subnuclear, mientras que GFP solo se encontraba presente tanto en el núcleo como en el citoplasma. C. Imágenes de alto aumento de núcleos de células 3T3 vivas 24 horas después de la transfección que muestran la organización subnuclear de Ciz1 y ECiz1 marcadas con EGFP y fragmentos derivados en los que se ha eliminado el fragmento C-terminal (equivalente a Cterm274). En ausencia de los dominios C-terminales, GFP-ECiz1 se localiza difusamente en el núcleo 24 horas después de la transfección, mientras que GFP-Ciz1 se agrega formando uno o dos agregados de gran tamaño en el interior del núcleo. D. El dominio 274 C-terminal solo es citoplasmático hasta después de que las células hayan pasado la mitosis (muy probablemente debido a la falta de secuencias de localización nuclear y a la entrada pasiva al núcleo), aunque una vez dentro se une a estructuras nucleares y se condensa con cromosomas. E. Se muestran imágenes representativas de GFP-Ciz1 (verde), BrdU (rojo) y núcleos totales (azul) en una población marcada con BrdU durante las primeras 12 horas tras la transfección. Los histogramas muestran la proporción de células transfectadas (en verde) que han incorporado BrdU en comparación con el número de células no transfectadas (gris) en tres ventanas de marcaje separadas. Durante el periodo de 0 a 22 horas posterior a la transfección, las células en ciclo rápido mostraron un incremento consistente de la fracción marcada con BrdU al ser transfectadas con Ciz1 o ECiz1. Se obtuvieron resultados similares con cultivos densos en los que la mayoría de células había salido del ciclo celular y había entrado en quiescencia. Sin embargo, al exponer las células en ciclo rápido a BrdU durante un pulso corto (de 20 minutos) 22 horas después de la transfección, el número de células en las que se producía síntesis de ADN se redujo en las poblaciones transfectadas con Ciz1 y ECiz1 en comparación con los controles no transfectados y con células transfectadas con GFP solo. Esto indica que, a las 22 horas, la síntesis de ADN había cesado en células que expresaban Ciz1.

Fig. 9. Potencial de proliferación y morfología celular alterados en poblaciones transfectadas. Agrupaciones celulares que aparecen en las poblaciones transfectadas de células 3T3. A. Las células se transfectaron con los dos tercios N-terminales de Ciz1 o ECiz1 (Nterm442) marcados con GFP y se mantuvieron bajo selección con 50  $\mu$ g/ml de G418.

Tras tres semanas bajo selección, los agregados celulares eran visibles con células positivas para GFP en el interior.

Fig. 10 Variantes de corte y empalme de la Ciz1 humana en cánceres pediátricos. Existen siete ADNc de Ciz1 humana en las bases de datos públicas, pero solo uno se deriva de tejido adulto normal (células B) y contiene todos los exones predichos. Los otros seis se derivan de células embrionarias o de cánceres pediátricos. Cinco de ellos se someten a corte y empalme alternativo con variabilidad en los exones 2, 3, 6 y 8 (por ejemplo la ECiz1 de ratón) y también en el exón 4 (tal como las células ES de ratón, las células germinales primordiales y testiculares). El sexto (AF159025) carece de la primera metionina y contiene polimorfismos de un único nucleótido que dan lugar a sustituciones de aminoácidos. Todas las diferencias respecto a la secuencia predicha (AB030835) están marcadas.

Fig. 11. Análisis de la secuencia EST. En cada mapa se incluye una representación esquemática de la proteína Ciz1 como referencia, que muestra las posiciones de los exones alternativamente procesados (en negro), los dominios putativos de interacción con la cromatina (en gris) y los dedos de cinc predichos (líneas verticales negras). Todas las secuencias EST están acompañadas de su número de acceso Genbank con la biblioteca a partir de la que derivaron indicada entre paréntesis. Las secuencias ausentes de las EST de Ciz1 debido al corte y empalme alternativo se muestran en amarillo, los desplazamientos de marco en rojo y las deleciones putativas en gris. Los polimorfismos de un único nucleótido que dan lugar a sustituciones de aminoácidos se indican con puntos negros y algunos de ellos se producen en un sitio de fosforilación cdk de consenso que los presentes inventores han demostrado que resulta importante para la regulación de la actividad de Ciz1 (puntos azules). La posición de la secuencia insertada en la línea celular de carcinoma MGC102 se indica con un triángulo.

- A) EST traducidas procedentes de cánceres pediátricos y de cánceres neurales adultos.
- B) EST traducidas procedentes de diversas células y tejidos no cancerosos
- C) EST traducidas procedentes de leucemias, linfomas y de células hematopoyéticas y linfocíticas normales
- D) EST traducidas procedentes de carcinomas
- E) EST traducidas procedentes de un abanico de otros cánceres
- F) Resumen de las regiones sometidas a corte y empalme alternativo en la Ciz1 humana que muestra secuencias incluidas condicionalmente.

Fig. 12 Expresión de la variante de corte y empalme de Ciz1 en líneas celulares tumorales de sarcoma de Ewings (ESFT) y en líneas celulares de neuroblastoma. A. Muestras de ARN completas de seis líneas celulares de ESFT independientes, dos neuroblastomas y una línea celular control (células HEK293) se sometieron a análisis mediante RT-PCR usando 4 conjuntos de cebadores diferentes. Las líneas celulares ESFT son: 1) A673, 2) RDES, 3) SKES1, 4) SKNMC, 5) TC3, 6) TTC466. Las líneas celulares de neuroblastoma son 1) IMR32 y 2) SKNSH

B) Análisis de productos de PCR de los exones 3/4/5 de Ciz1 en ESFT y neuroblastoma. Los productos de los cebadores h3 y h4 (que comprendían los exones potencialmente variables 4 y 6) se analizaron con mayor detalle. Se purificaron los fragmentos de PCR a partir de geles de agarosa mediante procedimientos estándar, se subclonaron y se secuenciaron para identificar la fuente de las variaciones de tamaño de los fragmentos. Se secuenciaron entre uno y once clones individuales para cada una de las siete líneas celulares y se resumen los resultados en forma de tabla. Ciz1 de las líneas celulares ESFT no presenta exón 4 en 31 % de los transcritos globalmente, y en algunas líneas ESFT este nivel es más próximo a 50 %. DSSSQ más frecuentemente se encuentra ausente en las dos líneas celulares de neuroblastoma sometidas a ensayo en el presente documento.

Fig. 13. Isoformas de Ciz1 en fibroblastos humanos normales (Wi38) y en líneas celulares de cáncer de próstata metastásico (PC3 y LNCAP). A. Ambas líneas celulares de cáncer de próstata contienen un exceso de la variante p125 de Ciz1 de mayor tamaño en la fracción nuclear, en comparación con la línea celular no de cáncer. B. Modelos para la producción de p85 (100) a partir de variantes de p125 mediante procesamiento de las proteínas durante el inicio de la replicación del ADN.

La Fig. 14 Ilustra la secuencia de ARNm de ratón de longitud completa.

La Fig. 15 Ilustra la secuencia de ARNm de ser humano de longitud completa.

La Fig. 16 Ilustra la secuencia de proteínas de ratón de longitud completa.

La Fig. 17 Ilustra la secuencia de proteínas de ser humano de longitud completa.

La Fig. 18. Ilustra las secuencias de proteína sometidas a corte y empalme alternativo. Las secuencias mostradas se encuentran ausentes en las secuencias proteicas sometidas a corte y empalme.

La Fig. 19. Ilustra las secuencias de ARNm sometidas a corte y empalme alternativo. Las secuencias mostradas se encuentran ausentes en las secuencias proteicas sometidas a corte y empalme.

Las Figs. 20A y B. Ilustran las secuencias de unión únicas creadas en las proteínas Ciz1 humanas por exones que faltan. Las secuencias de unión representan los sitios diana principales para agentes terapéuticos identificados por el método de la invención.

Las Figs. 21A a H ilustran las secuencias de unión creadas en el ARNm de la Ciz1 humana.

Identificación de Ciz1. Se ha aprovechado un anticuerpo policlonal (anticuerpo V1) producido la Cdc6 humana recombinante (Coverley et al., 2000; Stoeber et al., 1998; Williams et al., 1998) con el fin de identificar y estudiar un antígeno desconocido cuyo comportamiento se correlaciona con el inicio de la replicación del ADN *in vitro*. El antígeno presenta una Mr aparente de 100 kDa (denominado p85) y es fácilmente detectable en extractos procedentes de células 3T3 (fig. 1A).

La síntesis de ADN puede activarse en experimentos de replicación libres de células utilizando núcleos en fase G1 tardía "competentes para la replicación", en extractos en G1 y ciclina A-cdk2 recombinante. En estas condiciones, los núcleos incorporan nucleótidos marcados en el ADN naciente, de un modo estrictamente dependiente de la concentración de proteína quinasa activa (fig. 1B). En valores superiores e inferiores a la concentración óptima no tiene lugar inicio de la replicación del ADN. Sin embargo, se producen otros sucesos inversamente correlacionados con el inicio (Coverley et al., 2002). En el presente documento, los presentes inventores utilizan la activación de la síntesis del ADN (Fig. 1B) y la fosforilación de Mcm2 (que resulta en una movilidad incrementada, Fig. 1C), con el fin de calibrar los efectos de la ciclina A-cdk2 recombinante en experimentos de replicación libres de células, y correlacionan el comportamiento de p85 con la activación de la síntesis de ADN.

En los núcleos en G1 aislados nuevamente a partir de reacciones que contienen concentraciones inductoras de inicio de ciclina A-cdk2, el antígeno p85 es más prevalente en comparación con los núcleos expuestos a concentraciones más bajas o más altas de quinasa (Fig. 1D). Esto sugiere que p85 se encuentra regulada a cierto nivel por la ciclina A-cdk2 de una manera que es coincidente con la activación de la síntesis del ADN. Ningún otro antígeno se correlaciona de modo tan estrecho con esta etapa en el proceso de inicio libre de células, por lo tanto los presentes inventores utilizaron anticuerpo V1 para clonar el gen para p85 de ratón. Cuando se aplica a una biblioteca de expresión de ADNc derivada de embriones de ratón de 11 días, el anticuerpo V1 seleccionó dos clones que sobrevivieron múltiples rondas de cribado (véanse los métodos). Uno codificaba la Cdc6 de ratón, mientras que la otra codificaba 716 aminoácidos del homólogo murino de la Ciz1 humana (Mitsui et al., 1999). La Ciz1 humana y de ratón de longitud completa presentan una homología global de aproximadamente 70 % a nivel de los aminoácidos, con una homología máxima (>80 %) en las regiones N-terminales y C-terminales. Ciz1 se encuentra conservada entre los vertebrados debido a que existen homólogos en la rata y el fugu, pero no han podido identificarse proteínas con un grado elevado de homología o con una estructura de dominios similar en los eucariotas inferiores, planteando la posibilidad de que Ciz1 evolucionó para llevar a cabo una función especializada en el desarrollo de los vertebrados.

Una publicación anterior sobre la Ciz1 humana (Mitsui et al 1999) ha demostrado que interactúa con la proteína de ciclo celular p21-CIP1, llevando a que se investigase la proposición de que actúa como factor de transcripción, no como factor de replicación del ADN. Un segundo artículo Warder y Keherly, 2003) publicado tras la fecha de prioridad de la presente solicitud de patente sugiere un papel para Ciz1 en la tumorigénesis, pero no demuestra ninguna función en la replicación del ADN ni reconoce la importancia de la expresión de variantes de corte y empalme de Ciz1.

Múltiples isoformas de Ciz1 El marco de lectura abierto de Ciz1 de ratón predicho y un ADNc derivado de una biblioteca de tumor mamario de ratón (BC018483) contiene tres regiones que no se encuentran presentes en el clon embrionario de la presente invención (AJ575057), en lo sucesivo denominado ECiz1 (Fig. 2A). Las regiones variables en ECiz1 aparentemente son el resultado del corte y empalme alternativo de los exones 2/3, 6 y 8 (fig. 2B). El clon de melanoma de ratón AK089986 no presenta dos de las tres regiones de ECiz1 (Fig. 2A), mientras que la tercera codifica un tramo poliglutamina N-terminal que también se encuentra ausente de los clones derivados de meduloblastoma humano. Un cuarto bloque de secuencia derivado de los exones 3/4 se encuentra ausente de los transcritos de Ciz1 derivados de las células ES de ratón y del exón 4 en las células germinales primordiales de ratón (fig. 7). La Ciz1 humana también se procesa alternativamente a nivel de ARN, rindiendo transcritos que excluyen combinaciones de los mismos cuatro bloques de secuencia que Ciz1 de ratón (véase posteriormente). De hecho, todas las variaciones conocidas de los ADNc de Ciz1 de ratón presentan paralelos humanos estrechamente similares, algunos de los cuales son idénticos al nivel de los aminoácidos. Esto sugiere que las diferentes isoformas de Ciz1 presentan un significado funcional. Una quinta región variable (todavía no observada en el ratón) se somete a corte y empalme alternativamente en los transcritos de Ciz1 humana derivados principalmente de carcinomas.

Los datos sugieren que las formas más cortas de Ciz1 (que no presentan los exones alternativamente procesados) son más prevalentes precozmente durante el desarrollo y en linajes celulares que dan lugar a la línea germinal. En el análisis mostrado en la fig. 7, únicamente Ciz1 procedente de corazón neonatal totalmente desarrollado no mostraba ningún procesamiento alternativo, mientras que todos los tipos embrionarios contienen formas sometidas a corte y empalme alternativo. Además, los únicos ADNc de Ciz1 completos en las bases de datos públicos (humanos o de ratón) se derivan de tipos celulares no embrionarios y los únicos derivados de fuentes embrionarias se someten a corte y empalme alternativo. Por lo tanto, la expresión de una variante de corte y empalme de Ciz1 aparentemente se produce diferencialmente en tipos celulares que todavía no se han diferenciado por completo.

Notablemente, los ADNc de Ciz1 procedentes de cánceres pediátricos también se someten a corte y empalme alternativo (ver posteriormente). Esto conduce a la hipótesis de que la expresión de la isoforma de Ciz1 correcta en el punto correcto del desarrollo conduce a una actividad de Ciz1 incorrectamente regulada. Esto podría contribuir a una proliferación y transformación celular no programadas.

ECiz1 estimula la replicación del ADN *in vitro*. Tras la exposición a extracto citosólico procedente de células en fase S, los núcleos en fase G1 tardía inician la replicación del ADN e inician la síntesis de ADN naciente (Krude et al., 1997). Los inventores usaron este ensayo libre de células para someter a ensayo el efecto de ECiz1, y fragmentos recombinantes derivados, sobre la síntesis del ADN (fig. 3). La proteína ECiz1 de longitud completa

consistentemente incrementó el número de núcleos que se replicaban *in vitro*, de 30 % (+/-0,9 %) a 46 % (+/-5,5 %), lo que sugiere que Ciz1 es limitante para el inicio en extractos en fase S (Fig. 3A). Únicamente dos otras clases de proteína (las quinasas dependientes de ciclina, Coverley et al., 2002; Krude et al., 1997; Laman et al., 2001, y la proteína Cdc6, Coverley et al., 2002; Stoeber et al., 1998) se ha encontrado previamente que estimulan el inicio libre de células. De esta manera, ECiz1 es la primera proteína que presentando esta actividad no se conocía ya su participación en el proceso de la replicación. El efecto positivo de ECiz1 recombinante sobre el inicio libre de células apoya que la Ciz1 endógena desempeñe un papel positivo en la replicación del ADN en las células de mamífero.

La estimulación libre de células del inicio es dependiente de la concentración, con actividad máxima en el extracto en fase S en aproximadamente 1 nM (fig. 3B). Este resultado es similar al obtenido en análisis libres de células anteriores con otras proteínas recombinantes (Coverley et al., 2002; Krude et al., 1997), en la que la estimulación del inicio típicamente alcanza un máximo y después cae nuevamente al nivel no estimulado a concentraciones altas. Para ECiz1, el motivo de la caída de actividad a concentraciones altas todavía se desconoce. Sin embargo, los estudios de mutagénesis (ver posteriormente) sugieren que el mecanismo de restricción probablemente se encuentra activo y es específico, y no que se debe a un desequilibrio general de la composición de los complejos de proteínas de orden más alto.

La regulación por disminución de ECiz1 implica las treoninas 191/192. Probablemente Ciz1 es una fosfoproteína *in vivo*, ya que contiene numerosos sitios putativos de fosforilación y muestra una movilidad alterada al tratar extractos de células 3T3 con fosfatasa lambda (no mostrado). La Ciz1 murina contiene dos motivos de unión de RXL-ciclina y cinco sitios putativos de fosforilación de cdk, que se encuentran presentes en todas las variantes conocidas. Cuatro de ellas se encuentran localizadas en el fragmento N-terminal de ECiz1 que contiene actividad de replicación *in vitro* (véase posteriormente) y uno es contiguo al sitio en el que el exón 6 se somete a corte y empalme alternativo para excluir un motivo de secuencia DSSSQ corto (SEC ID N° 1) (Fig. 2A, C). Debido a que este motivo es 100 % idéntico y sometido a corte y empalme alternativo tanto en el ratón como en el ser humano, se consideró que la inclusión condicional podría servir para regular la actividad de Ciz1, identificando esta región de la proteína como potencialmente importante. Por lo tanto, se eligió centrarse en el sitio cdk, que se encuentra cuatro restos cadena arriba y que también se encuentra conservado en el ratón y en el ser humano, mediante la combinación de un enfoque genético y ensayos de replicación libres de células. Partiendo de ECiz1, se cambiaron las treoninas en 191 y 192 por dos alaninas, generando ECiz1T (191/2) A (fig. 2D). Al someterla a ensayo *in vitro* para determinar la actividad de replicación del ADN, ECiz1T (191/2) A estimuló el inicio en núcleos en G1 tardía en un grado similar a ECiz1 (fig. 3C). Sin embargo, al contrario que ECiz1, la estimulación del inicio se mantuvo a través de un abanico amplio de concentraciones que se extendía abrazando por lo menos tres órdenes de magnitud. Por lo tanto, existe un mecanismo que restringe la actividad del exceso de ECiz1 y que funciona en un ambiente libre de células. En una construcción separada, la treonina en la posición 293 también se cambió por alanina, generando ECiz1T (293) A (fig. 2D), aunque esta alteración presentó poco efecto sobre la actividad de ECiz1 según el ensayo *in vitro* (fig. 3D).

Estos resultados demuestran que la regulación negativa de la actividad de ECiz1 implica la treonina 191/2 y probablemente está causada por la fosforilación mediada por quinasa dependiente de ciclina en este sitio. Esto liga la actividad de Ciz1 a las rutas dependientes de cdk que controlan todos los sucesos principales del ciclo celular, incluyendo el inicio de la replicación del ADN.

La mayoría de proteínas del complejo de prerreplicación y muchas proteínas de la horquilla de replicación se encuentran fosforiladas *in vivo*, con frecuencia por quinasas dependientes de ciclina quinasas (Bell y Dutta, 2002; Fujita, 1999). Los datos de los inventores sugieren que la acumulación nuclear de antígeno p85-Ciz1 se encuentra regulada (directa o indirectamente) por ciclina A-cdk2, y demuestra que un sitio de fosforilación de cdk de consenso en la treonina 191/192 participa en el control de la actividad de Ciz1. Al provocar que este sitio no sea fosforilable, se mantiene la actividad de Ciz1 en un abanico más amplio de concentraciones en ensayos libres de células. Por lo tanto, la actividad de Ciz1 normalmente se encuentra regulada por disminución por la modificación en este sitio. Las funciones de los demás sitios de fosforilación de cdk conservados, y el efecto de la inclusión condicional de un motivo de unión de RXL-ciclina en la parte N-terminal procesada alternativamente de Ciz1, todavía no se han determinado. De esta manera, la simple relación negativa entre la actividad de Ciz1 y la fosforilación dependiente de cdk que se ha descubierto en la presente memoria es improbable que sea la historia completa. Sin embargo, el análisis de los presentes inventores hasta el momento relaciona Ciz1 con las rutas dependientes de cdk que controlan todas las transiciones principales del ciclo celular, y por lo tanto concuerda con la conclusión principal de que Ciz1 participa en el inicio de la replicación del ADN.

La actividad de replicación *in vitro* reside en el extremo N-terminal. Ciz1 presenta varias actividades C-terminales que podrían anclar la proteína dentro del núcleo. El dominio de matriz 3 sugiere la interacción con la matriz nuclear y los tres dedos de cinc implican la interacción con ácidos nucleicos. En efecto, datos recientes sugieren que la Ciz1 humana se une a ADN de una manera débilmente específica de secuencia (Warder y Keherley, 2003). Para determinar si los dominios C-terminales resultan importantes para la actividad de replicación de ECiz1 se dividió la proteína en dos fragmentos (fig. 2D). Nterm442 (que contiene el NLS, dos sitios cdk conservados, un dedo de cinc y todos los sitios conocidos en los que se ha observado procesamiento variable) estimula el inicio en un grado similar y a la misma concentración que ECiz1 (fig. 3E). En contraste, la parte C-terminal (Cterm274) no contiene ninguna actividad de replicación residual (fig. 3F). Por lo tanto, el dominio de matriz 3, uno de los sitios de fosforilación de la

quinasa dependiente de ciclina, y dos de los dedos de cinc no resultan necesarios para la actividad de replicación del ADN de ECiz1, al someterlos a ensayo *in vitro*. Sin embargo, debe indicarse que dicho análisis mide la actividad de ECiz1 *in trans* bajo condiciones en las que las consecuencias de la localización incorrecta resulta improbable que se detecten. Por lo tanto, sigue siendo una posibilidad que el dominio de matriz 3 y los dedos de cinc actúen *in vivo* dirigiendo la actividad de Ciz1 a sitios específicos en el núcleo y limitando de esta manera el alcance de la actividad de Ciz1.

Ciz1 endógena. El anticuerpo V1 reconoce Cdc6, así como p85-Ciz1 (fig. 1A), de manera que no resulta adecuado para experimentos de inmunofluorescencia destinados a visualizar la localización subcelular de la Ciz1 endógena. Por lo tanto, los inventores generaron dos nuevos antisueros policlonales de conejo contra el fragmento recombinante de ECiz1 llamado Nterm442, denominados 1793 y 1794 anti-Ciz1. Tal como se esperaba, Nterm442 purificado resulta reconocido por los anticuerpos 1793 y 1794 anti-Ciz1 en filtros western, aunque también resulta reconocido por el anticuerpo V1 (fig. 4A), apoyando la conclusión de que p85 (p100) es, en efecto, Ciz1.

Cuando se aplica a extractos de proteínas derivados de células 3T3 en cultivo, el anticuerpo 1793 anti-Ciz1 reconocía dos antígenos, con Mr de 125 y 100 kDa (fig. 4B), cuyas proporciones relativas varían de preparación a preparación. La banda de 100 kDa comigra con el antígeno sensible a ciclina A que resulta reconocido por el anticuerpo V1 (figs. 1 y 4B), lo que sugiere que ambos anticuerpos reconocen la misma proteína *in vivo*. Los inventores confirmaron mediante inmunoprecipitación que las bandas de p100-Ciz1 reconocidas por los anticuerpos V1 y 1793 son la misma proteína (fig. 4C). El anticuerpo V1 precipitó una banda de 100 kDa que resultó reconocido en transferencias de tipo Western por 1793, y viceversa. Además, en el mismo experimento, 1793, y en menor grado el anticuerpo V1, precipitaron un antígeno de 125 kDa, que se fue reconocido en los filtros western como 1793. Conjuntamente, las observaciones de los presentes inventores demuestran que la banda de 100 kDa en efecto es Ciz1 (previamente conocida como p85) y sugieren que la proteína Ciz1 existe en por lo menos dos formas en las células ciclantes.

Además de la evidencia de inmunoprecipitación proporcionada anteriormente, otras observaciones llevan a la conclusión de que p125 también es una forma de Ciz1. En primer lugar, los dos anticuerpos anti-Ciz1 de la presente invención (1793 y 1794) presentan dicha banda en común. Ambos anticuerpos producen el mismo patrón de tinción nuclear en experimentos de inmunofluorescencia, el cual resulta alterado en células tratadas con ARNip de Ciz1 (véase más adelante). En segundo lugar, las proporciones relativas de p100 y p125 varían entre preparaciones, y por lo tanto podrían ser el resultado de la escisión proteolítica. En tercer lugar, los resultados de los presentes inventores son inesperadamente similares a los de Mitsui et al (1999) cuyo anticuerpo monoclonal anti-Ciz1 humano detectó dos antígenos con Mr aparente de 120 kDa y 95 kDa en células HEK293. Propusieron que la forma de 120 kDa de la proteína Ciz1 humana se procese para producir la forma de 95 kDa y los resultados de la presente invención son consistentes con esta propuesta.

La banda de 125 kDa reconocida por el anticuerpo 1793 en células de ratón y humanas se resuelve en tres bandas relacionadas con Ciz1 durante la electroforesis de alta resolución de material derivado de células humanas no transformadas (Wi38, ver posteriormente) y de células de ratón (NIH3T3, no mostrado). Esto puede ser el resultado de la modificación post-traduccional de la proteína Ciz1 o del procesamiento alternativo del transcrito de Ciz1.

Distribución subcelular de Ciz1. Se utilizó 1793 anti-Ciz1 para visualizar la distribución subcelular de la proteína Ciz1 (p85 y p125) en células 3T3 (fig. 5A) y en células HeLa (no mostradas). En ambos tipos celulares, 1793 reaccionó con un antígeno específico nuclear y esto se bloqueó mediante la inclusión del fragmento recombinante Nterm442 (fig. 5B). Al contrario que Cdc6, que se muestra a título comparativo (fig. 5A), Ciz1 es claramente detectable en todas las células 3T3 en esta población ciclante. Por lo tanto, Ciz1 se encuentra presente en el núcleo durante toda la interfase, aunque variaciones menores de cantidad o isoforma no resultarían detectables mediante este método. Tras el tratamiento con detergente, se redujo la tinción nuclear total de Ciz1 en todos los núcleos, lo que sugiere que Ciz1 se encuentra presente en el núcleo tanto en forma de fracción soluble como también unido a estructuras nucleares insolubles.

Cuando proteína soluble se elimina mediante lavado, el antígeno inmovilizado insoluble se resuelve en un patrón moteado subnuclear de puntos a alta magnificación (fig. 5C, D). Las manchas de Ciz1 muestran un intervalo y distribución de tamaños similares a los de los "focos" o "fábricas" de replicación, los sitios en los que tiene lugar la síntesis del ADN en la fase S. Para saber si Ciz1 era coincidente con los sitios de las fábricas de replicación, se compararon la posición de las manchas de Ciz1 con la posición del PCNA, un componente de los complejos de replicación en las células en fase S (fig. 5C). En una sección confocal, los focos de PCNA eran menos abundantes que los focos de Ciz1, aunque prácticamente todos eran coincidentes con Ciz1 (fig. 5D, E, F). Esto resulta particularmente inesperado para los focos en el intervalo de tamaños intermedios. En las imágenes fusionadas, el solapamiento entre las posiciones de los focos de PCNA y de Ciz1 resulta en manchas amarillas, mientras que los focos de Ciz1 restantes que no son coincidentes con PCNA son rojos. Los focos verdes (solo PCNA) se encuentran prácticamente ausentes, lo que sugiere que Ciz1 se encuentra presente en todos los sitios en los que se han formado fábricas de replicación de ADN.

Ciz1 también se encuentra presente en sitios que no contienen PCNA (fig. 5D) y al contrario que PCNA, los focos de Ciz1 persisten durante toda la interfase (fig. 5A). Una interpretación de estas observaciones que Ciz1 marca las posiciones en el núcleo en las que las fábricas de replicación que contienen PCNA pueden formarse en la fase S, pero que no todos estos sitios se usan al mismo tiempo. Todavía no se ha determinado si diferentes focos de Ciz1 se convierten en sitios activos de replicación del ADN en diferentes tiempos en la fase S, o si también se producen otras actividades nucleares en sitios en los que se encuentra unido Ciz1. De hecho, en esta etapa también sigue siendo posible que las variantes de 100 kDa y de 125 kDa de Ciz1 presenten diferentes actividades y que residan en sitios nucleares con diferentes funciones.

Ciz1 es esencial para la proliferación celular. Hasta el momento se ha demostrado que el comportamiento de p85 (p100) -Ciz1 correlaciona con el inicio de la replicación del ADN en ensayos libres de células, que la Ciz1 recombinante estimula la frecuencia del inicio y que Ciz1 reside en los mismos sitios nucleares que la maquinaria de replicación del ADN. Sin embargo, estos datos no muestran que Ciz1 presente una función esencial en las células en proliferación. Con el fin de someter a ensayo esta posibilidad, se utilizó la interferencia de ARN (ARNi) para reducir selectivamente los niveles de transcritos de Ciz1 en células NIH3T3. Se seleccionaron cuatro secuencias diana dentro de Ciz1 (ver la fig. 2A) y se produjeron *in vitro* moléculas de ARN de interferencia pequeño (ip). Cuando se aplican a las células, la totalidad de los cuatro ARNip de Ciz1 restringió el crecimiento (fig. 6A) y provocó una reducción visible del nivel de la proteína Ciz1 tras 48 horas (fig. 6B). El efecto de la reducción de Ciz1 sobre la proliferación se pone de manifiesto entre las 23 y las 40 horas después de la transfección, lo que sugiere que el primer ciclo celular sin ARN de Ciz1 resulta relativamente no afectado. A las 40 horas, los controles y las células tratadas con ARNip de Ciz1 divergieron significativamente, sin proliferación adicional en poblaciones con niveles reducidos de Ciz1. Para verificar la especificidad de la reducción de los niveles de Ciz1, se realizó un seguimiento de los niveles de transcrito a las 24 horas, antes de que la proliferación resultase inhibida significativamente (fig. 6C). En este punto, los transcritos de Ciz1 se habían reducido a 42 % del nivel observado en las células de control tratadas con ARNip de GAPDH. Estos experimentos demuestran que Ciz1 resulta necesario para la proliferación celular y son consistentes con la existencia de una función primaria en la replicación del ADN.

Para someter a ensayo adicionalmente lo anterior, las células se marcaron con un pulso de BrdU 48 horas después del tratamiento de ARNip con el fin de determinar la fracción de células en las que se está produciendo la síntesis de ADN (fig. 6D). Al reducir los niveles de Ciz1, la fracción marcada con BrdU también se reduce, sugiriendo que la síntesis del ADN se encuentra inhibida bajo estas condiciones. Además, las células en la población con nivel reducido de Ciz1 que había incorporado BrdU (aproximadamente 15 % de la población) se encontraban marcadas menos intensamente. Por lo tanto, en algunas células con Ciz1 tratadas con ARNip, la fase S se ralentiza en lugar de encontrarse inhibida por completo, posiblemente debido a una reducción incompleta de los niveles

La inhibición de la síntesis de ADN por parte de los ARNip de Ciz1 podría ser una consecuencia secundaria de una disrupción general de la función nuclear. Por lo tanto, se investigó en mayor detalle un abanico de otras proteínas de replicación cuyos niveles se encuentran regulados de un modo dependiente del ciclo celular, para averiguar si las células con niveles reducidos detienen el ciclo aleatoriamente, o si produce una acumulación en un punto particular del ciclo.

Durante el inicio de la replicación del ADN eucariótico, las proteínas del complejo Mcm se ensamblan en los orígenes de replicación en la etapa G1 tardía de un modo dependiente de Cdc6. Cierta tiempo después, las ADN polimerasas y los factores accesorios de las mismas (incluyendo el PCNA) se unen a la cromatina y se activan los orígenes. Esto se asocia a la exportación nuclear y a la proteólisis de la mayoría de Cdc6 y, a medida que se produce la síntesis del ADN, al desplazamiento gradual del complejo Mcm respecto de la cromatina (Bell y Dutta, 2002). Con el fin de identificar el punto de acción de Ciz1, se utilizó inmunofluorescencia para realizar un seguimiento de Mcm3 y del PCNA. En las células con niveles reducidos de Ciz1 (figs. 6E, F), ambas proteínas eran detectables dentro del núcleo unidas a estructuras nucleares resistentes a detergente. Por lo tanto, dichos factores es improbable que se unan directamente a Ciz1, o que sean dependientes de Ciz1 para su ensamblaje. De hecho, en cuatro experimentos independientes, el número medio de células con Mcm3 unido a cromatina y resistente a detergente se incrementó de 31 % (+/-6 %) a 51 % (+/-5 %) (fig. 6E). Un nivel incrementado de Mcm3 indica que la etapa dependiente de Ciz1 se produce tras el ensamblaje del complejo de prerreplicación (aunque antes de que se complete la fase S). En las mismas poblaciones celulares, también se incrementó la fracción positiva para PCNA, de 32 % (+/- 5 %) a 49 % (+/- 6 %) (fig. 6F), localizando el punto de acción de Ciz1 en un periodo posterior al ensamblaje del PCNA. De esta manera, Ciz1 muy probablemente actúa facilitando la replicación del ADN durante una etapa tardía del proceso de inicio, mientras que su inacción inhibe el avance a través de la fase S, dejando en su sitio a Mcm3 y PCNA.

En conjunto, las investigaciones libres de células o con células de los presentes inventores dibujan un sólido esquema de la función primaria de Ciz1. Sugieren que Ciz1 es un nuevo componente de las fábricas de replicación del ADN, y demuestran que Ciz1 desempeña un papel positivo en el ciclo celular de los mamíferos, actuando en la promoción del inicio de la replicación del ADN.

Tres de las líneas de investigación sugieren que Ciz1 resulta necesario durante una etapa tardía del proceso de inicio después de la formación del complejo de prerreplicación de ciclina A-cdk2 que activan la síntesis de ADN, lo

que implica que Ciz1 actúa durante esta etapa y no durante etapas más tempranas de ensamblaje del complejo de replicación (Coverley et al., 2002). En segundo lugar, los estudios funcionales con núcleos en la etapa G1 tardía demuestran que el ECiz1 recombinante incrementa el número de núcleos que incorporan nucleótidos marcados *in vitro*. Por lo tanto, Ciz1 debe encontrarse activo en una etapa que convierte los núcleos preparados para iniciar la síntesis del ADN en núcleos en síntesis activa de ADN. En tercer lugar, los estudios de interferencia de ARN apuntan a la existencia de una etapa dependiente de Ciz1 posterior a la formación del complejo Mcm y tras el ensamblaje de PCNA sobre el ADN, aunque anterior al desplazamiento de estas proteínas. Estas diferentes líneas de investigación conducen a conclusiones inesperadamente similares acerca del punto de acción de Ciz1, situándolo en las etapas posteriores del proceso de inicio.

ARNip anti-Ciz1 como estrategia terapéutica. El análisis demuestra que Ciz1 resulta esencial para la proliferación celular, y que apuntar a Ciz1 como diana es una estrategia viable para limitar la proliferación. Las formas sometidas a corte y empalme alternativo de Ciz1 que se observan en diversos cánceres (ver posteriormente) implican que Ciz1 podría seleccionarse como diana de un modo selectivo para limitar la proliferación en un subconjunto de células dentro de una población.

A título de ejemplo, lo anterior podría llevarse a cabo dirigiendo los ARNip a la secuencia de unión creada en los transcritos de Ciz1 en el caso de que falte la secuencia C-terminal GTTGAGGAGGAACTCTGCAAGCAG (SEC ID N° 2) en células de carcinoma pulmonar de células pequeñas, o mediante la utilización de proteína Ciz1 que no presenta la secuencia VEEELCKQ (SEC ID N° 3) correspondiente para seleccionar inhibidores químicos específicos.

De acuerdo con lo expuesto anteriormente, la presente invención también proporciona la utilización de secuencias de unión creadas en los transcritos y proteínas de Ciz1 en el caso de que no se encuentren presentes secuencias sometidas a corte y empalme alternativo, a modo de marcador diagnóstico, indicador pronóstico o diana terapéutica.

La forma embrionaria de Ciz1 se localiza en el núcleo. El análisis de RT-PCR de exones potencialmente variables sugiere que las células 3T3 expresan predominantemente Ciz1 de longitud completa, de modo que el trabajo de los presentes inventores de inmunolocalización de la Ciz1 endógena (fig. 5) no refleja necesariamente el comportamiento de ECiz1, el cual no presenta varios bloques de secuencia y posiblemente, por lo tanto, información que se utiliza para localizar la proteína. Con el fin de comparar directamente la localización de ECiz1 y la Ciz1 de longitud completa, se transfectaron construcciones marcadas con GFP mejorada en células 3T3 (fig. 8A) y se microinyectaron en pronúcleos de ratón (fig. 8B). En todos los casos, los Ciz1 y ECiz1 marcados eran exclusivamente nucleares, mientras que una construcción de control que expresaba GFP solo se encontraba presente en el núcleo y en el citoplasma. Tanto GFP-Ciz1 como GFP-ECiz1 eran visibles en células vivas en forma de focos subnucleares, similares a los focos de replicación observados en células fijadas mediante inmunofluorescencia. De esta manera, los tres bloques de secuencia que se encuentran ausentes de ECiz1 aparentemente no contribuyen a la localización nuclear de Ciz1.

Durante el periodo de tres días posterior a la transfección no se observó división celular en las células transfectadas con GFP-Ciz1 y con GFP-ECiz1. Estos datos sugieren que la sobreexpresión de la Ciz1 funcional presenta un efecto inhibitor del ciclo celular (en células que presentan rutas reguladoras intactas).

Coalescencia. Al transfectar células 3T3 con construcciones marcadas con GFP en los que el tercio C-terminal de Ciz1 había sido eliminado, se observaron diferencias entre ECiz1 y Ciz1 de longitud completa (fig. 8C). A las 48 horas, FL Ciz1 N-term (equivalente a 442) se había fusionado formando grandes agregados intranucleares que solo resultaron aparentes en la población transfectada con ECiz1 N-term442 el día 3 o después. . Antes de este tiempo, ECiz1 N-term442 se localizaba en forma de patrón específicamente nuclear pero difuso. De esta manera, la capacidad de coalescencia es diferente en términos cuantificables en el caso de Ciz1 y de ECiz1, y por lo tanto resulta afectada por uno de los tres exones procesados alternativamente (2/3, 6 u 8).

Al igual que las células transfectadas con Ciz1 de longitud completa y ECiz1, no se observó que las células transfectadas con construcciones en las que el tercio C-terminal había sido eliminado se multiplicasen durante el periodo de seguimiento de tres días.

Los dominios C-terminales anclan Ciz1 a las estructuras nucleares. Tal como se ha indicado anteriormente, la diferencia entre Ciz1 y ECiz1 N-term se encuentra enmascarada cuando los dominios C-terminales también se encuentran presentes (fig. 8A). Además, el fragmento C-terminal solo dirige el marcador GFP a la cromatina, formando un patrón irregular que no es tan moteado (focal) como Ciz1 o ECiz1, pero que sigue unido a los cromosomas durante la mitosis (fig. 8D). Esto sugiere que los dominios C-terminales se encuentran implicados en la inmovilización de Ciz1 sobre un marco estructural en el núcleo. Notablemente, las células transfectadas transitoriamente con fragmento C-terminal siguieron dividiéndose, resultando en la dilución gradual de la fluorescencia verde.

La Ciz1 ectópica estimula entrada prematura en la fase S. Se analizaron los sucesos que se producían durante el primer día después de la transfección. La fracción en fase S de células transfectadas (verde) se comparó con la fracción en fase S de células no transfectadas mediante el marcaje con BrdU en diversos intervalos. Durante largas

5 ventanas de marcaje, incluyendo 0 a 22 horas (fig. 8E), 0 a 12 horas y 0 a 7 horas (no mostrado), se estaba produciendo la síntesis de ADN consistentemente en más células transfectadas con Ciz1 y con ECiz1 que en células no transfectadas. Esto sugiere que Ciz1 y ECiz1 presentan un efecto positivo sobre la transición G1-S, promoviendo la entrada no programada en la fase S. Se obtuvieron resultados similares con poblaciones de células 3T3 densamente sembradas en placas antes de la transfección. Esto se llevó a cabo con el fin de minimizar la fracción de la población no transfectada que se encontraba en la fase S como parte del ciclo celular normal. En estas condiciones, la diferencia entre las poblaciones transfectadas y no transfectadas era máxima, demostrando claramente el efecto de la Ciz1 ectópica sobre el inicio de la replicación del ADN.

10 A la inversa, al marcar células con un pulso corto de BrdU administrado a las 22 horas (fig. 8E), o 10 o 12 horas después de la transfección (no mostrado), la fracción marcada se había reducido consistentemente en las poblaciones transfectadas con Ciz1 y con ECiz1. Esto sugiere que la fase S que resulta inducida por Ciz1 o ECiz1 ectópicos es anormal, con una síntesis del ADN lenta o nula que no resulta suficiente para marcar las células durante ventanas cortas de exposición a BrdU.

15 Por lo tanto, Ciz1 y ECiz1 ectópicas presentan dos efectos sobre la fase S en las células en cultivo. Estimulan la replicación del ADN, pero resultando en una síntesis del ADN lenta o nula.

20 Clones con potencial de proliferación alterado. Se realizó asimismo un seguimiento de poblaciones transfectadas de células 3T3 durante un periodo de tres semanas. En células transfectadas con GFP-Nterm442 o el equivalente no procesado alternativamente y mantenido bajo selección con G418, se observaron grandes focos que contenían cientos de células (fig. 9A). Estas agrupaciones contenían un gran número de células que expresaban GFP, demostrando que la sobreexpresión de la parte N-terminal de ECiz1 (en la que reside la actividad de replicación) no es lenta, y sugiriendo que la sobreexpresión conduce a un fenotipo de proliferación alterado, en comparación con las células no transfectadas, incluyendo la pérdida de inhibición por contacto y la incapacidad para formar una monocapa. Este comportamiento alterado dependiente de Ciz1 podría contribuir a la formación de tumores. Una versión truncada de Ciz1 de ratón que no presentaba los dominios de interacción putativos con la cromatina ha sido aislada anteriormente a partir de un melanoma de ratón (fig. 2).

30 Ciz1 humana y cáncer

35 ADNc de Ciz1 en bases de datos públicas. Tal como se ha indicado anteriormente, la Ciz1 humana se procesa alternativamente al nivel del ARN, rindiendo transcritos que no presentan tres de los mismos exones que la Ciz1 embrionaria de ratón. Se han registrado siete ADNc de Ciz1 humanos en las bases de datos públicas (fig. 10), presentadas por Mitsui et al (1999), Warder y Keherly (2003) y proyectos a gran escala de análisis del genoma (proyecto NIH-MGC, proyecto NEDO de secuenciación del ADNc humano). Únicamente uno deriva de tejido adulto normal y éste contiene todos los exones predichos (AB030835). El resto deriva de células embrionarias (AK027287), o notablemente de cuatro tipos diferentes de cáncer pediátrico (meduloblastoma, AF159025, aAF0234161, retinoblastoma, AK023978, neuroblastoma, BC004119 y linfoma de Burkitt, BC021163). La forma embrionaria y las formas derivadas del cáncer no presentan bloques de secuencia de las mismas tres regiones que el clon embrionario de ratón de los presentes inventores y de una cuarta región que corresponde al exón 4. Por lo tanto, los limitados datos sugieren que las formas sometidas a corte y empalme alternativo son más prevalentes en etapas tempranas del desarrollo. Esta correlación no ha sido observada anteriormente en la literatura científica. La presencia de Ciz1 sometida a corte y empalme alternativo en cánceres pediátricos plantea la posibilidad de que el procesamiento incorrecto de Ciz1 esté asociado a una proliferación celular inadecuada.

50 Por ejemplo, uno de los exones variables codifica un motivo de secuencia DSSSQ conservado corto que se encuentra ausente en la ECiz1 de ratón y en un meduloblastoma humano. Este está directamente adyacente al sitio de fosforilación de cdk de consenso que los presentes inventores han demostrado que se encuentra implicado en la regulación de la función de EGiz1. La inclusión condicional de la secuencia DSSSQ podría convertir a Ciz1 en el objeto de la regulación por parte de la familia ATM/ATR de proteína quinasas, las cuales fosforilan las proteínas en las secuencias SQ, limitando de esta manera la función de inicio de Ciz1 en respuesta al daño al ADN.

55 Análisis de los marcadores de secuencia expresada. La presencia de Ciz1 sometida a corte y empalme alternativo en cánceres pediátricos ha motivado un análisis detallado de las EST de Ciz1. Existen 567 marcadores de secuencia expresada (EST) dentro de la Unigene cluster del NCBI nº Hs.23476 (Ciz1 humana). Estas se derivan de un amplio abanico de tejidos y líneas celulares normales y enfermas. Las secuencias se han traducido y mapeado frente a la secuencia de aminoácidos de longitud completa predicha de la Ciz1 humana. Se han registrado alteraciones de secuencias que dan lugar a sustituciones o deleciones de aminoácidos, a desplazamientos de marco o a la terminación prematura de la traducción.

60 También se han observado variantes de corte y empalme alternativo de Ciz1 en estos datos de EST y se incorporan como referencia en la presente memoria. Los cuatro bloques de secuencia que los presentes inventores habían informado anteriormente que se habían sometido a corte y empalme alternativo en la Ciz1 humana y de ratón (fig. 2) han sido observadas en las secuencias de EST, así como la variante previamente no detectada que carece de la secuencia derivada del exón 14 VEEELCKQ. Todos estos bloques de secuencia recurrentemente variantes se

encuentran limitados por sitios de corte y empalme adecuados. Se ha identificado un sexto bloque de secuencia variable en una biblioteca derivada de carcinoma, causado por la inclusión de

**GCCACCCACACCACGAAGAGATGTGTTTGCCACGTTCCAGTGCAGGGGTGGAG**  
**CACAGCCCGGCTTGTTACAGATAT (SEQ ID NO: 4).**

5 Las EST se agrupan según el tipo celular a partir del que se derivaron, con las divisiones primarias presentes entre las células neoplásicas de origen adulto, infantil o embrionario. Las EST de tejido normal de origen embrionario o adulto se incluyen a título comparativo. Los mapas de la proteína Ciz1 derivada de EST se muestran en la fig. 11A-E y los exones sometidos a corte y empalme alternativo se resumen en la fig. 11F.

10 Tres bloques de secuencia en el extremo N-terminal de la Ciz1 humana no se encuentran presentes en transcritos procedentes de meduloblastomas y neuroblastoma (fig. 11A) y ocasionalmente no se encuentran en transcritos de Ciz1 procedentes de otros cánceres. Los presentes inventores también encontraron un corte y empalme alternativo similar en un tercer cáncer pediátrico, el sarcoma de Ewings (véase más adelante). Las secuencias sometidas a corte y empalme alternativo asociadas a cáncer pediátrico proceden de los exones 2/3 (por lo menos dos versiones), del exón 4 y del exón 6.

15 Las variantes del exón 8 en las que no se encuentran presentes una o más copias de una repetición degenerada rica en Q han sido observadas en transcritos derivados de células normales (de origen neural embrionario o adulto) y de diversos cánceres. El corte y empalme alternativo en dicha región podría producir Ciz1 con una actividad inapropiada; por lo tanto, la expresión de variante del exón 8, o la presencia de mutaciones puntuales que influyan sobre el procesamiento en esta región, podrían resultar útiles como marcadores diagnósticos o pronósticos de cáncer. Las repeticiones del exón 8 degeneradas sometidas a corte y empalme alternativo se detallan posteriormente y se resumen en la fig. 11F.

20 En la mitad C-terminal de la proteína Ciz1 humana, dos bloques de secuencia someten a corte y empalme de modo variable. Una de éstas no se encuentra presente en los transcritos derivados de tres de cada cinco bibliotecas de carcinoma pulmonar y de carcinoide pulmonar, y de tres otras bibliotecas de carcinoma (pero muy raramente de transcritos de otros tipos celulares).

25 El segundo bloque de secuencia variante se debe a la inclusión incorrecta de secuencia adicional en transcritos procedentes de la biblioteca del carcinoma epidermoide (MGC102).

30 Dichas secuencias y las secuencias de unión formadas en las proteínas Ciz1, y los transcritos de Ciz1 en el caso de que estos segmentos sean excluidos o incluidos, son dianas potenciales para la inhibición selectiva de la proliferación celular en un amplio abanico de cánceres diferentes. Las secuencias no variantes restantes son dianas potenciales para la inhibición no selectiva de la proliferación celular.

35 Además de variantes de corte y empalme, se ha observado la presencia en algunos cánceres de otros transcritos de Ciz1 no típicos. En los rhabdomyosarcomas, Ciz1 se termina prematuramente, conduciendo a una proteína predicha que no presenta dominios de unión nuclear C-terminales. Esto podría conducir a una replicación incorrecta del ADN y por lo tanto podría ser una diana terapéutica o un marcador en este tipo de cáncer.

40 Varios transcritos contienen mutaciones puntuales que conducen a sustituciones de aminoácidos en sitios putativos de fosforilación de quinasa (cdk) dependiente de ciclina. En la biblioteca MGC12 del carcinoma cervical, lo anterior se produce dos veces. Los presentes inventores han demostrado que se encuentran implicados dos sitios de fosforilación cdk en la limitación de la actividad de Ciz1 (fig. 3C y D), implicando a estas mutaciones en la alteración de la regulación de la proliferación en las células de cáncer. Uno de ellos es el mismo que el mutante derivado del carcinoma mencionado anteriormente (fig. 11E). Los transcritos derivados de cáncer con mutaciones puntuales en Ciz1 también podrían ser dianas para la interferencia de ARN, o presentar un valor como indicadores diagnósticos o pronósticos.

Investigación de la expresión de variantes de Ciz1 en los cánceres pediátricos

55 Se ha investigado la expresión de variantes de Ciz1 en 6 líneas celulares tumorales de la familia del sarcoma de Ewings (ESFT) y en dos líneas celulares de neuroblastoma, utilizando RT-PCR con conjuntos de cebadores que comprenden tres regiones de variabilidad conocida de Ciz1 (fig. 12A). Este análisis ha demostrado que el patrón de expresión variante de Ciz1 es diferente en las células ESFT y en las células del neuroblastoma y en células no transformadas, aunque aparentemente es muy similar en conjuntos de líneas celulares del mismo tumor. Por lo tanto, la expresión variante de Ciz1 podría presentar un potencial pronóstico o diagnóstico para estos cánceres. Algunas variaciones menores dentro de un conjunto de líneas del mismo tipo tumoral podrían presentar valor pronóstico.

60 Mediante la subclonación y secuenciación de transcritos amplificados, los presentes inventores encontraron que la totalidad de las seis líneas ESFT sometidas a ensayo expresaban una forma exón 4 menos de Giz1. Debido a que

Ciz1 resulta esencial para la proliferación celular (ver posteriormente), ofrece una posible vía para la limitación selectiva de las células ESFT. Los transcritos de las dos líneas celulares de neuroblastoma sometidas a ensayo raramente carecen de exón 4 pero frecuentemente carecen de secuencias con el motivo DSSSQ codificado por el exón 6 (fig. 12B).

5 Dicho análisis experimental confirma que los cánceres pediátricos expresan formas de Ciz1 con inclusión variable de los exones 4, 6 y probablemente de los exones 2/3.

10 Dos versiones de la secuencia que comprende el exón 8 y una forma de la secuencia que abarca la secuencia de codificación VEEELCKQ se detectaron en ESFT, neuroblastomas y control, sugiriendo que estas regiones no contribuyen a la desregulación de Ciz1 en dichos cánceres pediátricos.

15 En todos los casos, los productos de RT-PCR de Ciz1 eran más abundantes en reacciones llevadas a cabo con muestras de ARN procedentes de líneas celulares de cáncer, comparado con los controles (Wi38, HEK293, NIH3T3 y osteoblastos humanos primarios). Esto es consistente con que se encuentre incrementada la expresión de las variantes de Ciz1 en los tumores.

#### Análisis de la expresión de la proteína Ciz1 en líneas celulares de cáncer de próstata

20 Análisis de la expresión de la proteína Ciz1 en líneas celulares de cáncer prostático Los fibroblastos pulmonares humanos no transformados normales (y las células NIH3T3 de ratón) expresan dos formas mayores de Ciz1 que se detectan con el anticuerpo policlonal anti-Ciz1 1793 en transferencias de tipo Western (fig. 13A). La banda más grande (aproximadamente 125 kDa) se resuelve en tres bandas diferentes que se encuentran presentes en proporciones iguales en células Wi38, pero en proporciones muy desiguales en las líneas celulares de cáncer prostático PC3 y LNCAP (y en líneas de células ESFT, no mostradas). Los presentes inventores postulan que estas isoformas de proteína se generan mediante la expresión de exones sometidos a corte y empalme de forma variable. Ambas líneas celulares tumorales también contienen más antígeno Ciz1 que las células Wi38, lo que es consistente con la sobreexpresión de Ciz1 en dichas líneas celulares de cáncer.

30 En conjunto, los resultados de los presentes inventores (experimentales y el análisis bioinformático de los datos genómicos) apoyan la conclusión de que Ciz1 se encuentra incorrectamente regulado en un amplio abanico de cánceres humanos. Los presentes inventores han demostrado que la proteína Ciz1 desempeña un papel positivo en el proceso de la replicación del ADN; por lo tanto, una Ciz1 mutante podría contribuir a la transformación celular y no ser una consecuencia de la misma. En el caso de que la alteración de la regulación de Ciz1 sea una etapa común en este proceso, representaría una diana muy atractiva para el desarrollo de agentes terapéuticos.

35 Los inventores también han relacionado asimismo cambios particulares con cánceres específicos, con lo que es una posibilidad real que Ciz1 resulte útil como marcador diagnóstico o pronóstico.

40 Estos incluyen:

- Corte y empalme alternativo en la parte N-terminal de la proteína (que contiene actividad de replicación *in vitro*) en cánceres pediátricos.
- Mutaciones puntuales en sitios de fosforilación de quinasa dependiente de ciclina, las cuales es conocido que se encuentran implicadas en la limitación de la actividad de replicación de Ciz1.
- Expresión no típica y propiedades de unión nuclear de formas de Ciz1-p125 en líneas celulares de carcinoma prostático, posiblemente debido al procesamiento incorrectamente regulado de las repeticiones degeneradas en el exón 8, o en otros exones.
- Exclusión condicional de un motivo discreto (VEEELCKQ) en el extremo C-terminal de Ciz1 (que probablemente participa en la localización de la proteína Ciz1 dentro del núcleo) en el carcinoma pulmonar de células pequeñas y en otros carcinomas.
- Niveles incrementados de proteína y ARN de Ciz1 (detectados mediante transferencia western y mediante RT-PCR) en todas las líneas celulares derivadas de cáncer sometidas a ensayo hasta el momento, en comparación con los niveles en fibroblastos pulmonares embrionarios normales Wi38, de ARN de osteoblastos humanos y en fibroblastomas NIH3T3 de ratón.

55 Las secuencias mostradas en las figuras 14 a 21 resultan útiles para el desarrollo de reactivos terapéuticos, diagnósticos o pronósticos.

60 Materiales y procedimientos

65 Clonación. Se utilizó una biblioteca de expresión de ADNc enriquecida de longitud completa del tramo 5' triplEx de lambda, derivada de embriones de ratón de 11 días (Clontech nº ML5015t) para infectar *E. coli* X11blue según el protocolo recomendado (Clontech). Se levantaron las placas sobre filtros de nitrocelulosa de 0,45 micrómetros previamente sumergidos en IPTG 10 mM (Sigma). Se aplicó anticuerpo V1 purificado mediante afinidad, en

aproximadamente  $3 \times 10^6$  placas a una dilución de 1/1.000 en PBS, leche en polvo desnatada al 10 %, Tween-20 al 0,4 %, tras bloquear durante 30 minutos en ausencia del anticuerpo. Tras dos horas, se lavaron tres veces los filtros con el mismo tampón y las placas reactivas se visualizaron con anticuerpo secundario anticonejo conjugado con peroxidasa de rábano picante (Sigma) y quimioluminiscencia mejorada (ECL, Amersham) siguiendo procedimientos estándar. Se recolectaron 43 placas independientes, pero solo dos cepas de fago sobrevivieron tres rondas adicionales de detección selectiva. Éstas se convirtieron en pTriplex mediante transformación en BM25.8 y se secuenciaron. Una codifica para la Cdc6 de ratón (clon P) y la otra (clon L) para una proteína de ratón desconocida que es homóloga de la Ciz1 humana. Los inventores hacen referencia a esta Ciz1 embrionaria (ECiz1) y se presentó al EMBL bajo el número de acceso AJ575057.

Expresión bacteriana. Se utilizaron construcciones de expresión bacteriana basados en pGEX (Amersham) para producir proteínas ECiz1 para el análisis *in vitro*. Se generó pGEX-ECiz1 mediante la inserción de un fragmento SmaI-XbaI (extremos romos) de 2,3 kb procedente del clon L en el sitio SmaI de pGEX-6P-3. Se generó pGEX-Nterm442 mediante la inserción del fragmento XmaI-XhoI de 1,35 kb en pGEX-6P-3 digerido con XmaI-XhoI y pGEX-Cterm274 mediante la inserción del fragmento XhoI de 0,95 kb en pGEX-6P-3 digerido con XhoI. Se generó pGEX-T (191/2) a partir de pGEX-ECiz1 mediante mutagénesis dirigida a sitio (Stratagene Quikchange) utilizando los cebadores AACCCCTCTTCCGCCGCCCAATCGCAAGA (SEC ID N° 5) y TCTTGCATTGGGGCGCGGAAGAGGGGTT (SEC ID N° 6). Se generó pGEX-T(293)A a partir de pGEX-ECiz1 utilizando los cebadores AAGCAGACACAGCCCCGGATCGGCTGCCT (SEC ID N° 7) y AGGCAGCCGATCCGGGGCCTGTGTCTGCTT (SEC ID N° 8). La integridad y marco de lectura de todos los clones se verificaron mediante secuenciación.

Se produjeron Ciz1 recombinante, fragmentos de Ciz1 y mutantes puntuales en BL21-pLysS (Stratagene) en forma de proteína marcada con glutatión-S-transferasa. Esto se purificó a partir de lisados bacterianos sometidos a ultrasonidos y clarificados mediante la unión a glutatión-sefarosa 4B (Amersham). La proteína recombinante se eluyó mediante escisión del marcador GST utilizando proteasa de precisión (tal como recomienda el fabricante, Amersham) en tampón (Tris-HCl 50 mM, pH 7,0, NaCl 150 mM, DTT 1 mM). Esto proporcionó preparaciones de proteína de entre 0,2 y 2,0 mg/ml. Para las réplicas de ensayo, se prepararon diluciones en serie en Hepes 100 mM, pH 7,8, DTT 1 mM y glicerol al 50 %, de manera que no se añadiese más de 1 ml de solución de proteína a 10 ml de ensayo de replicación, dando las concentraciones mostradas. Consistente con las observaciones anteriores (Mitsui et al., 1999; Warder y Keherly, 2003), la Ciz1 recombinante y el fragmento derivado N-term442 migraron en el SDS-PAGE con un peso molecular anómalamente elevado. Se produjo ciclina A-cdk2 en bacterias tal como se ha descrito anteriormente (Coverley et al., 2002).

Anticuerpos anti-Ciz1. Se cultivó anticuerpo V1 policlonal de conejo (Coverley et al., 2000; Stoeber et al., 1998; Williams et al., 1998) contra un fragmento interno del Cdc6 humano expresado en bacterias, correspondiente a los aminoácidos 145 a 360, y se purificó mediante afinidad siguiendo procedimientos estándares (Harlow y Lane, 1988). Este anticuerpo reacciona fuertemente con p100-Ciz1 endógena y también con el fragmento Nterm442 de ECiz1. La alineación de Nterm442 con Cdc6 aminoácidos 145 a 360 sugiere que el epítipo compartido podría encontrarse en 294-298 o en 304-312 en la Ciz1 de ratón. Se utilizó Nterm442 recombinante para generar dos antiseros policlonales específicos de Ciz1 denominados 1793 y 1794 (Abcam). De forma rutinaria se ha usado 1793 en los experimentos descritos en el presente documento. Su especificidad se verificó mediante inmunoprecipitación recíproca y análisis de transferencia western con anticuerpo V, mediante inclusión de Nterm442 (25 µg/ml en tampón para anticuerpo, BSA 10 mg/ml, SDS al 0,02 %, Triton X100 al 0,1 % en PBS), que bloqueó la reactividad con epítopos endógenos, y mediante reducción mediada por ARNip de Ciz1 que redujo específicamente la tinción nuclear de 1793.

Inmunoprecipitación. Las células 3T3 se lavaron en PBS en crecimiento asíncrono, se aclararon en tampón de extracción (Hepes 20 mM, pH 7,8, acetato potásico 5 mM, cloruro magnésico 0,5 mM) suplementado con cóctel inhibidor de proteasa sin EDTA (Roche) y se recolectaron mediante raspado al igual que los extractos de replicación. Las células se lisaron con Triton X100 al 0,1 % y la fracción de sedimento resistente a detergente se extrajo con NaCl 0,3 M en tampón de extracción. Se utilizaron 5 µl de 1793 o 2 µl de anticuerpo V por cada 100 µl de extracto y se incubaron durante 1 hora a 4 °C. Se extrajeron los complejos de antígeno-anticuerpo con 100 µl de proteína G-sefarosa (Sigma) y las perlas se lavaron cinco veces con Tris 50 mM, pH 7,8, EDTA 1 mM, NP40 al 0,1 %, NaCl 150 mM. Los complejos se sometieron a ebullición en tampón de carga (DTT 100 mM, SDS al 2 %, Tris 60 mM, pH 6,8, azul bromofenol al 0,001 %) y se resolvieron mediante electroforesis en gel de SDS-poliacrilamida al 6,5 %

Inmunofluorescencia. Se cultivaron células sobre cubreobjetos y se fijaron en paraformaldehído al 4 %, con o sin una breve preexposición a Triton X100 al 0,05 % en PBS. Se detectó Ciz1 endógena con suero 1793 diluido 1/2.000 en tampón para anticuerpo siguiendo procedimientos estándar. Se detectó Mcm3 con el anticuerpo monoclonal sc9850 (1/1.000), Cdc6 con el anticuerpo monoclonal sc9964 (1/100) y PCNA con el anticuerpo monoclonal PC100 (1/100, todos de Santa Cruz Biotechnology). El análisis de colocalización de imágenes confocales fluorescentes doblemente teñidas se llevó a cabo tal como se ha descrito (Rubbi y Milner, 2000; van Steensel et al., 1996).

Sincronía celular. Se sincronizaron células 3T3 de ratón mediante salida de la quiescencia tal como se ha descrito anteriormente (Coverley et al., 2002). Los núcleos preparados a partir de células recolectadas 17 horas después de

la salida (denominadas "de G1 tardía") se utilizaron en todos los experimentos de replicación libres de células descritos en el presente documento. Esto proporcionó poblaciones que contenían núcleos en fase S, núcleos en G1 tardía competentes para la replicación y núcleos en G1 temprana/G0 no sensibles, en proporciones variables. Los extractos de 3T3 receptores en plena etapa G1 se prepararon a las 15 horas (típicamente contienen aproximadamente 5 % de células en fase S). La serie de experimentos de replicación libres de células descritos en la presente memoria requirió grandes cantidades de extracto estandarizado; por lo tanto, se utilizaron células HeLa debido a que se sincronizan fácilmente en masa. Se prepararon extractos de HeLa en fase S a partir de células que habían salido de la quiescencia durante dos horas procedentes de dos bloqueos secuenciales en fase S inducidos con timidina, tal como se ha descrito (Krude et al., 1997)

Replicación del ADN libre de células. Se llevaron a cabo ensayos de replicación del ADN tal como se ha descrito (Coverley et al., 2002; Krude et al., 1997). Brevemente, se incubaron 10  $\mu$ l de extracto en fase G1 o S media (suplementado con sistema regenerador de energía, nucleótidos y dUTP biotinilado) y  $5 \times 10^4$  núcleos en fase G1 tardía, durante 60 minutos a 37 °C. Las reacciones se suplementaron con lisado baculovírico que contenía ciclina Acck2 (Fig. 1B y C), en el que 0,1  $\mu$ l de lisado presentaba la misma actividad específica que 1 nM de quinasa purificada (Coverley et al., 2002). Todas las proteínas recombinantes se diluyeron en serie en Hepes 100 mM, pH 7,8, DTT 1 mM, glicerol al 50 %, de manera que no se añadiesen más de 1  $\mu$ l por cada 10  $\mu$ l de ensayo de replicación, generando las concentraciones indicadas. Las reacciones se detuvieron con 50  $\mu$ l of 0,5 % de Triton X100 al 0,5 % y se fijaron mediante la adición de 50  $\mu$ l de paraformaldehído al 8 % durante 5 minutos. Todas las transferencias a los núcleos en cubreobjetos se tiñeron con estreptavidina-FITC (Amersham) y se contratiñeron con Toto-3-yoduro (Molecular Probes). Se cuantificó la proporción de núcleos marcados mediante inspección a un aumento de 1.000X y todos los núcleos con focos fluorescentes o marcaje uniforme intenso se consideraron positivos. Las imágenes de los núcleos en replicación *in vitro* se generaron mediante microscopia confocal a un aumento de 600X, de muestras contrateñidas con yoduro de propidio. Para el análisis de las proteínas nucleares, los núcleos se aislaron nuevamente tras la exposición durante 15 minutos a condiciones iniciadoras, mediante dilución de dos veces de las reacciones con PBS frío y centrifugación suave.

Análisis y presentación de los datos. Antes de la utilización en los ensayos de inicio, cada preparación de núcleos en etapa G sincronizados se sometió a ensayo para determinar la proporción de núcleos que ya se encontraba en fase S ("% en S"). Para ello, dichos núcleos se incubaron en un extracto que era incapaz de inducir el inicio de la síntesis de ADN (procedente de células en plena etapa G1 recolectadas 15 horas después de su salida de la quiescencia) pero que permitían eficientemente la síntesis de elongación del ADN a partir de orígenes iniciados *in vivo*. La fracción de núcleos en elongación incorporaba nucleótidos eficientemente durante los ensayos de inicio *in vitro* pero no proporciona ninguna información. Rutinariamente esta fracción se preestableció y restó de los datos crudos. Las poblaciones sincronizadas en las que 20 % o menos de las células se encontraban en fase S se utilizaron para los ensayos de inicio.

Al hacer salir de la quiescencia las células 3T3 utilizando el protocolo utilizado en el presente documento, una proporción no superior al 70 % de la población total entra en la fase S (Coverley et al., 2002). Sin embargo, la frecuencia de replicación máxima observada *in vitro* se aproxima más a 50 %; habitualmente obtenida mediante incubación con ECiz1. Para la población en G1 de núcleos 3T3 utilizada en la presente memoria, el 17 % se encontraba en fase S (% en S) y el número máximo que se replicaba en cualquier ensayo *in vitro* era de 51 % (% en replicación). Por lo tanto, el 34 % de esta población es competente para iniciar la replicación *in vitro* (% de C). De esta manera, para cada punto de datos en las Figs. 3B a F, % en inicio = (% en replicación - % en S) /% de C x 100.

ARNA de interferencia Se utilizó como diana Ciz1 endógena en células NIH3T3 proliferantes utilizando ARNip transcritos *in vitro* (kit Ambion Silencer) dirigido contra cuatro regiones de la Ciz1 de ratón. Las secuencias oligonucleotídicas que se utilizaron para generar los ARNip son AAGCACAGTCACAGGAGCAGACCTGT (SEC ID N° 9) CTC y AATCTGCTCCTGTGACTGTGCCCTGTCTC (SEC ID N° 10) para ARNip 4, AATCTGTAC AAGTTCTACGACCTGTCTC (SEC ID N° 11) y AATCGTAGAACTTGTGACAGACCTGTCTC (SEC ID N° 12) para ARNip 8, AATCGCAAGG ATTCTTCTTCTCCTGTCTC (SEC ID N° 13) y AAAGAAGAAGAA TCCTTGCGACCTGTCTC (SEQ ID NO: 14) para ARNip 9, y AATCTGCAGCAGTTCTTTCCCCTGTCTC (SEQ ID NO: 15) y AAGGAAAGAACTGCTGCAGACCTGTCTC (SEC ID N° 16) for Siria 11 para ARNip 11. Las secuencias diana que se encuentran distribuidas en todo el transcrito de Ciz1 se seleccionaron basándose en predicciones reducidas de estructura secundaria y en localizaciones dentro de los exones que se expresan consistentemente en todas las formas conocidas de Ciz1 (secuencias 4, 8 y 11), con la excepción de uno (ARNip 9) que se sabe que se somete a corte y empalme alternativo. Los controles negativos no se trataron, se trataron falsamente (reactivos de transfección pero no ARNip) y las células se trataron con ARNip de GAPDH (Ambion). Se utilizaron ARNip marcados con Cy3 (Ambion) para estimar la eficiencia de transfección, que se encontró que era superior al 95 %. Los experimentos de interferencia de ARN se llevaron a cabo en un formato de 24 pocillos partiendo de  $2 \times 10^4$  células por pocillo en 500  $\mu$ l de medio (DMEM con glutamax suplementado con FCS al 4 %). Los ARNip se añadieron 12 horas después de la siembra en placa utilizando reactivo oligofectamina para la introducción (Invitrogen). A menos que se indique lo contrario, los ARNip se utilizaron por parejas (a una concentración total de 2 nM en medio), en forma de dos dosis, administrando la segunda dosis en medio fresco 24 horas después de la primera. Los resultados se evaluaron 48 horas después de la primera exposición, mediante recuento del número de células, marcaje en la fase S e inmunotinción. Se llevaron a cabo transferencias de tipo Northern en ARN aislados a partir de células

tratadas durante 24 horas con una única dosis de ARNip, en reacciones que se ampliaron en 5 veces. Se preparó ARN utilizando reactivo Trizol (Invitrogen) y las muestras se sometieron a electroforesis a través de agarosa al 1 %, se transfirieron a membrana de nilón Hybon N+ (Amersham) y se hibridaron secuencialmente a 50 °C con sondas de ADNc utilizando reactivos del kit NorthernMax (Ambion), siguiendo las instrucciones del fabricante. La membrana se aclaró entre cada hibridación utilizando solución de SDS al 0,5 % a 90 °C, se dejó que se enfriase lentamente hasta la temperatura ambiente. Las sondas se marcaron con [<sup>32</sup>P]-dCTP utilizando el sistema de marcaje de ADN de cebadores aleatorios (Gibco BRL,) y se utilizaron en el orden siguiente: i. un fragmento Xmal-XhoI de 1,35 kb derivado a partir de ECiz1. ii. ADNc de  $\alpha$ -actina (Clontech) (Clontech) y iii. ADNc de GAPDH de ratón (RNWAY laboratories). La membrana se lavó dos veces en 2X SSC SDS al 0,2 % durante 30 – 60 minutos cada vez, seguido de un lavado en in 0,2X SSC 0,2 % SDS durante 30 minutos a una temperatura de 55 – 65 °C, dependiendo de la sonda utilizada. Las señales de hibridación se cuantificaron utilizando un aparato de captura de imágenes Amersham Biosciences Typhon 9410 en modo variable y software Image Quant TL (v2002). Las intensidades de banda se expresan en unidades arbitrarias (entre paréntesis) y los resultados para Ciz1 y GAPDH se normalizaron respecto a los de la  $\alpha$ -actina y se expresaron en %.

Marcaje en la fase S. La fracción de núcleos en los que se estaba produciendo síntesis de ADN *in vivo* se siguió mediante la suplementación del medio de cultivo con bromodesoxiuridina 20  $\mu$ M (BrdU, Sigma) durante 20 minutos. La BrdU incorporada se visualizó tras el tratamiento ácido con anticuerpo monoclonal anti-BrdU conjugado con FITC (Alexis Biochemicals) siguiendo las instrucciones del fabricante. Los núcleos se contratiñeron con un aparato Hoescht 33258 y se contaron con aumento alto (1.000X).

#### Ciz1 marcada con proteína fluorescente verde

Se obtuvo ADNc de Ciz1 de ratón de longitud completa del UK HGMP Resource Centre (MGC clon nº 27988) y se verificó por completo la secuencia. Se ligaron en el mismo marco con proteína fluorescente verde mejorada (EGFP) un fragmento Smal-XbaI (de extremos romos) de 2,8 kb de longitud completa de Ciz1 a partir de dicho clon y un fragmento Smal-XbaI (de extremos romos) de 2,3 kb de ECiz1 a partir de pTrip1Ex-clon L, en el sitio Smal de pEGFP-C3 (Clontech). Se utilizó pEGFP-C3 sin inserción a modo de control. Las construcciones se transfectoron en células NIH3T3 utilizando *TransIT-293* (Mirus), siguiendo instrucciones del fabricante o se microinyectaron en el pronúcleo macho de huevos de ratón fertilizados en el estadio unicelular. Las células 3T3 en crecimiento transfectadas con EGFP-Ciz1 de longitud completa o EGFP-ECiz1 se analizaron mediante microscopía fluorescente de células vivas hasta tres días después de la transfección. Se realizó un seguimiento de la síntesis del ADN durante las primeras 24 horas después de la transfección, mediante la inclusión del análogo de nucleótido BrdU en el medio de cultivo celular durante diversos periodos de tiempo tal como se indica en las leyendas de las figuras. Tal como se ha indicado anteriormente, cualquier célula en la que se esté produciendo síntesis de ADN mientras se encuentra expuesta a tinción de BrdU con anticuerpo monoclonal anti-BrdU genera núcleos rojos.

Las células transfectadas con Ciz1 también se mantuvieron bajo selección con 50  $\mu$ g/ml de G418 en medio de cultivo estándar (DMEM Glutamax más suero de feto bovino al 10 %) durante como máximo un mes, rindiendo poblaciones celulares de morfología alterada.

#### Análisis de las secuencias de EST

Se tradujeron marcadores de secuencia expresada (EST) individuales localizadas en el Unigene cluster del NCBI nº Hs.23476 (Ciz1 humana), utilizando GeneJockey y la secuencia de aminoácidos predicha se comparó con la secuencia predicha para la Ciz1 de longitud completa, con el fin de identificar los cambios recurrentes en las células cancerosas. Con el fin de excluir errores reflejo de una secuenciación de ADN de mala calidad, tal como la que se produce al final de análisis largos de secuenciación, únicamente se han incluido en este análisis aquellos cambios situados más de 8 aminoácidos del final de la secuencia ininterrumpida. Los desplazamientos de marco restaurados por una segunda alteración posteriormente en la lectura y los desplazamientos de marco seguidos de un codón de parada únicamente se incluyen en el caso de que a continuación se encuentre una secuencia ininterrumpida. De esta manera, se excluyen de este análisis la mayoría de los errores de secuenciación. Sin embargo, se espera que muchas de las mutaciones puntuales que queden (incluyendo los desplazamientos de marco y las paradas) reflejen errores introducidos durante la secuenciación. Por lo tanto, este análisis pretende descubrir tendencias, ponderando las mutaciones puntuales únicamente en el caso de que aparezcan en más de una ocasión.

De las 567 secuencias, localizadas en el Unigene cluster de Ciz1, se han analizado la mayor parte (todos los cánceres pediátricos, los carcinomas de próstata y de pulmón, las leucemias y los linfomas y un amplio abanico de tejidos no enfermos). Algunos no se mapearon debido a que son lecturas extremadamente cortas o porque proporcionaban secuencias de aminoácidos muy cortas tras realizarse la traducción, y para un número pequeño los presentes inventores detectaron una falta de homología con la secuencia codificante de Ciz1. Se excluyó un número pequeño de EST del análisis debido a múltiples desplazamientos de marco que produjeron tramos de homología en la totalidad de los tres marcos, sin indicaciones de que se utilice el marco de lectura *in vivo*. Todos ellos procedían de material derivado de cáncer, habitualmente de adenocarcinomas.

Análisis mediante RT-PCR de la expresión de isoformas de Ciz1. El ARN se aisló usando reactivo trizol tras los procedimientos recomendados, se trató con DNAsa y se sometió a transcripción inversa usando hexámeros al azar y superscript II, después se amplificó con cebadores específicos de Ciz1: –

- 5 h/m5 CAGTCCCCACCACAGGCC (SEC ID N° 17),  
 h/m2 GGCTTCCTCAGACCCCTCTG (SEC ID N° 18).  
 H/m3 ACACAGACCTCTCCAGAGCACTTAG (SEC ID N° 19)  
 H/m4 ATGGTGACCTTCAGGGAGC (SEC ID N° 20)  
 H4 TCCTTGCGA TGTCTCTGG GCAGG (SEC ID N° 21)  
 10 H3 TCCCTCCTCA ACGGCTCCAT GCTGC (SEC ID N° 22)  
 H6 CG TGGGGGCGAC TTGAGCGTTG AGG (SEC ID N° 23)  
 H1 GATGCCAGGGGT ATGGGGCGCC GGG (SEC ID N° 24)  
 H2 TCCGAGCCCT TCCACTCCTC TCTGG (SEC ID N° 25)

#### 15 Análisis de las isoformas de la proteína Ciz1 en líneas celulares de cáncer

Se cultivaron células en DMEM con FCS al 10 % hasta la subconfluencia, se aclararon en solución salina tamponada con Hepes frío suplementado con el cóctel inhibidor de proteasa sin EDTA (Roche) y después se recolectaron y se suplementaron con Triton X100 al 0,1 %. El material insoluble en detergente (incluyendo los núcleos) se sedimentó mediante centrifugación suave, dando fracciones de sobrenadante (SN) y de sedimento (P). Estas fracciones se sometieron a ebullición en tampón para muestras de SDS-PAGE reductor y las proteínas se resolvieron mediante electroforesis en SDS-PAGE al 8 %. Tras la transferencia a nitrocelulosa, las isoformas de Ciz1 se detectaron con anticuerpo anti-Ciz1 1793). Todos los métodos utilizados en este análisis han sido bien documentados en otras partes.

#### 25 Referencias

- Bell, S. P. y Dutta, A. (2002). DNA replication in eukaryotic cells. *Annu Rev Biochem* 71, 333 – 74.  
 Cook, P. R. (1999). The organization of replication and transcription. *Science* 284, 1790 – 5.  
 30 Corpet, F. (1998). Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Res.* 16, 10881 – 10890.  
 Coverley, D., Laman, H. y Laskey, R. A. (2002). Distinct roles for cyclins E and A during DNA replication complex assembly and activation. *Nat Cell Biol* 4, 523 – 8.  
 Coverley, D., Pelizon, C., Trewick, S. y Laskey, R. A. (2000). Chromatin bound Cdc6 persists in S and G2 phases in human cells, while soluble Cdc6 is destroyed in a cyclin A–cdk2 dependent process. *J. Cell Sci.* 113, 1929 – 1938.  
 35 Fujita, M. (1999). Cell cycle regulation of DNA replication initiation proteins in mammalian cells. *Front Biosci* 4, D816 – 23.  
 Hanahan, D. y Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell* 100, 57 – 70. Harlow, E. y Lane, D. (1988). *Antibodies: A laboratory manual*. New York: Cold Spring Harbour Laboratory Press.  
 Jones, D. L., Alani, R. M. y Munger, K. (1997). The human papillomavirus E7 oncoprotein can uncouple cellular differentiation and proliferation in human keratinocytes by abrogating p21Cip1–mediated inhibition of cdk2. *Genes Dev.* 11, 2101 – 2111.  
 40 Krude, T. (2000). Initiation of human DNA replication *in vitro* using nuclei from cells arrested at an initiation–competent state. *J. Biol. Chem.* 275, 13699 – 13707.  
 Krude, T., Jackman, M., Pines, J. y Laskey, R. A. (1997). Cyclin/Cdk–dependent initiation of DNA replication in a human cell–free system. *Cell* 88, 109 – 119.  
 45 Laman, H., Coverley, D., Krude, T. K., Laskey, R. A. y Jones, N. (2001). Viral cyclin/cdk6 complexes initiate nuclear DNA replication. *Mol. Cell. Biol.* 2, 624 – 635.  
 Mercatante, D. R. y Kole, R. (2002). Control of alternative splicing by antisense oligonucleotides as a potential chemotherapy: effects on gene expression. *Biochim Biophys Acta* 1587, 126 – 32.  
 50 Mitsui, K., Matsumoto, A., Ohtsuka, S., Ohtsubo, M. y Yoshimura, A. (1999). Cloning and characterization of a novel p21cip1/waf1–interacting zinc finger protein, Ciz1. *Biochem. Biophys. Res. Com.* 264, 457 – 464.  
 Nakayasu, H. y Berezney, R. (1991). Nuclear matrins: identification of the major nuclear matrix proteins. *Proc Natl Acad Sci U S A* 88, 10312 – 6.  
 Ohnuma, S., Philpott, A. y Harris, W. A. (2001). Cell cycle and cell fate in the nervous system. *Curr Opin Neurobiol* 11, 66 – 73.  
 55 Parker, S. B., Eichele, G., Zhang, P., Rawls, A., Sands, A. T., Bradley, A., Olson, E. N., Harper, J. W. y Elledge, S. J. (1995). p53–independent expression of p21Cip1 in muscle and other terminally differentiating cells. *Science* 267, 1024 – 7.  
 Rubbi, C. P. y Milner, J. (2000). Non–activated p53 co–localizes with sites of transcription within both the nucleoplasm and the nucleolus. *Oncogene* 19, 85 – 96.  
 60 Sherr, C. J. y Roberts, J. M. (1999). CDK inhibitors: positive and negative regulators of G1–phase progression. *Genes Dev.* 13, 1501 – 1512.  
 Stoeber, K., Mills, A. D., Kubota, Y., Krude, T., Romanowski, P., Marheineke, K., Laskey, R. A. y Williams, G. H. (1998). Cdc6 protein causes premature entry into S phase in a mammalian cell–free system. *EMBO J.* 17, 7219 – 7229.  
 65

- van Steensel, B., van Binnendijk, E. P., Hornsby, C. D., van der Voort, H. T., Krozowski, Z. S., de Kloet, E. R. y van Driel, R. (1996). Partial colocalization of glucocorticoid and mineralocorticoid receptors in discrete compartments in nuclei of rat hippocampus neurons. *J Cell Sci* 109 (Pt 4), 787 – 92.
- 5     Warder, D. E. y Keherly, M. J. (2003). Ciz1, Cip1 interacting zinc finger protein 1 binds the consensus DNA sequence ARYSR(0–2)YYAC. *J Biomed Sci* 10, 406 – 17.
- Williams, G. H., Romanowski, P., Morris, L., Madine, M., Mills, A. D., Stoeber, K., Marr, J., Laskey, R. A. y Coleman, N. (1998). Improved cervical smear assessment using antibodies against proteins that regulate DNA replication. *Proc. Natl. Acad. Sci. USA* 95, 14932 – 14937.
- 10    Zezula, J., Casaccia-Bonnel, P., Ezhevsky, S. A., Osterhout, D. J., Levine, J. M., Dowdy, S. F., Chao, M. V. y Koff, A. (2001). p21cip1 is required for the differentiation of oligodendrocytes independently of cell cycle withdrawal. *EMBO Rep* 2, 27 – 34.

REIVINDICACIONES

1. Un método diagnóstico *in vitro* para la identificación de un trastorno proliferativo, comprendiendo el método detectar la presencia o la expresión de una variante de corte y empalme de Ci1z1, en el que el método proporciona una medida cuantitativa de una variante de la proteína Ciz1 en una muestra de un sujeto y la variante proteica es una secuencia de proteínas sometida a corte y empalme alternativo humana en la que la secuencia:

- (a) QQLQQLQQQLQQQLQQQLLQLQQLLQQSPP, o MFSQQQQQLQQQQQLQQLQQQLQQQLLQLQQLLQQSPPQA está ausente;
- (b) GLDQFAMPPATYDTAGLTMPATL está ausente;
- (c) DSSSQ está ausente;
- (d) PQVQPAHSQPPRQVQLQLQKQVQTQTY;

PQVQPAHSQGPQRQVOLQQEAEPLKQVQPQVQPAHSQPPRQVQLQLQKQ  
VQTQTY,

o

QVQSQTQPRIPSTDTQVQPKLQKQAQTQTSPEHLVLQKQVQPQLQQEAEPPKQVQ  
PQVQPAHSQGPQRQVQLQQEAEPLKQVQPQVQPAHSQPPRQVQLQLQKQVQTQ

Y está ausente; o

- (e) VEEELCKQ está ausente.

2. Un método diagnóstico *in vitro* para la identificación de un trastorno proliferativo, comprendiendo el método detectar la presencia o la expresión de una variante de corte y empalme de Ci1z1, en el que el método proporciona una medida cuantitativa de una variante del ARN de Ciz1 en una muestra de un sujeto y en el que la variante del ARN de Ciz1 es una secuencia de ARN sometida a corte y empalme alternativo humana en la que la secuencia:

- (a) TGGGGGCTGCGGGGCCGGGCCCATCCGTGGGGGCGACTTGAGCGTTGAGG  
GCGCGCGGGGAGGCGAGCCACCATGTTTCAGCCAGCAGCAGCAGCAGCTCCAGCA  
ACAGCAGCAGCAGCTCCAGCAGTTACAGCAGCAGCAGCTCCAGCAGCAGCAATTG  
CAGCAGCAGCAGTTACTGCAGCTCCAGCAGCTGCTCCAGCAGTCCCCACCACAGG  
CC, o

o

CAGCAGCTCCAGCAGTTACAGCAGCAGCAGCTCCAGCAGCAGCAATTGCAGCAGC  
AGCAGTTACTGCAGCTCCAGCAGCTGCTCCAGCAGTCCCCACCACA

está ausente;

- (b)

GGACTGGACCAGTTTGCAATGCCACCAGCCACGTATGACACTGCCGGT

CTCACCAT GCCACAGCAACACTG

está ausente;

- (c) AGGATTCTTCTTCTC está ausente;

- (d)

(d) CCACAGGTGCAGCCCCAGGCACATTCACAGCCCCCAAGGCAGGTGC  
AGCTGCAGCTGCAGAAGCAGGTCCAGACACAGACATATCC, o

o

CCACAGGTACAGCCACAGGCACATTCACAGGGCCCAAGGCAGGTGCAGCT  
GCAGCAGGAGGCAGAGCCGCTGAAGCAGGTGCAGCCACAGGTGCAGCCCCAGGC  
ACATTCACAGCCCCCAAGGCAGGTGCAGCTGCAGCTGCAGAAGCAGGTCCAGACA  
CAGACATAT, o

5

o

CAGGTGCAGTCACAGACTCAGCCGCGGATACCATCCACAGACACCCAGGT  
GCAGCCAAAGCTGCAGAAGCAGGCGCAAACACAGACCTCTCCAGAGCACTTAGTG  
CTGCAACAGAAGCAGGTGCAGCCACAGCTGCAGCAGGAGGCAGAGCCACAGAAG  
CAGGTGCAGCCACAGGTACAGCCACAGGCACATTCACAGGGCCCAAGGCAGGTG  
CAGCTGCAGCAGGAGGCAGAGCCGCTGAAGCAGGTGCAGCCACAGGTGCAGCCC  
CAGGCACATTCACAGCCCCCAAGGCAGGTGCAGCTGCAGCTGCAGAAGCAGGTCC  
AGACACAGACATAT está ausente; o

10 está ausente; o

(e) GTTGAGGAGGAACTCTGCAAGCAG está ausente.

15 3. El método de acuerdo con la reivindicación 1 o la reivindicación 2, en el que el trastorno proliferativo es cáncer.

4. El método de acuerdo con la reivindicación 3, en el que el cáncer es un cáncer pediátrico seleccionado del grupo que consiste en retinoblastoma, neuroblastoma, linfoma de Burkitt, meduloblastoma y tumores de la familia del sarcoma de Ewings.

20 5. El método de acuerdo con la reivindicación 3, en el que el cáncer es carcinoma, adenocarcinoma, linfoma o leucemia.

6. El método de acuerdo con la reivindicación 3, en el que el cáncer es cáncer de hígado, de pulmón o de piel.

25 7. El método de acuerdo con la reivindicación 1 o la reivindicación 2, en el que el método diagnóstico comprende una o más de las etapas siguientes:

30 (a) poner en contacto la muestra con un agente que se une específicamente a un polipéptido con actividad Ciz1 y medir en la muestra la unión del agente con el polipéptido;

(b) usar la PCR con transcripción inversa o la PCR en tiempo real para controlar la expresión de la isoforma de Ciz1 y medir los niveles de expresión;

(c) medir la presencia de mutaciones en el ácido nucleico o aminoácido en base a las propiedades de conformación de la molécula alteradas.

35 8. El método de acuerdo con la reivindicación 7, en el que el agente que se une específicamente a un polipéptido con actividad de Ciz1 es un anticuerpo anti-Ciz1.

9. El método de acuerdo con la reivindicación 8, en el que el anticuerpo es un anticuerpo monoclonal.

Figura 1A -D

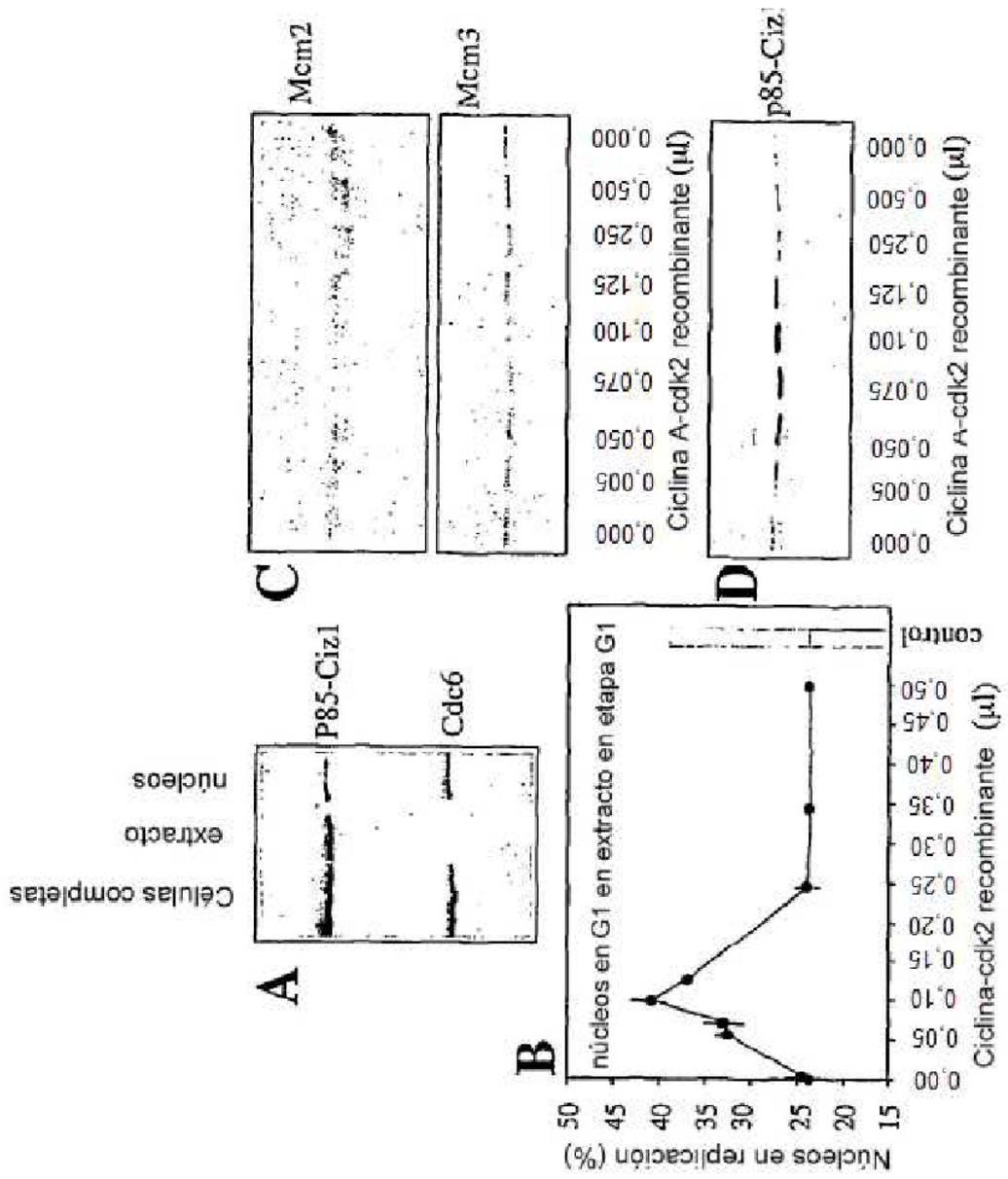


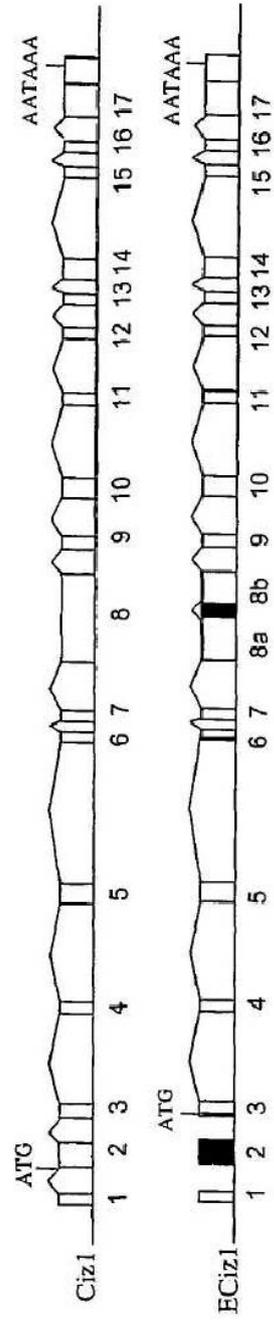
Figura 2A

**A** Isoformas de Ciz1 de ratón

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190	200	210	220
Full	MF	PL	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q
EC1-3	MF	PL	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q
nc1aoms	MF	PL	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q
Full	216	225	235	245	255	265	275	285	295	305	315	325	335	345	355	365	375	385	395	405	415	425	435
EC1-3	216	225	235	245	255	265	275	285	295	305	315	325	335	345	355	365	375	385	395	405	415	425	435
nc1aoms	216	225	235	245	255	265	275	285	295	305	315	325	335	345	355	365	375	385	395	405	415	425	435
Full	434	440	450	460	470	480	490	500	510	520	530	540	550	560	570	580	590	600	610	620	630	640	650
EC1-3	434	440	450	460	470	480	490	500	510	520	530	540	550	560	570	580	590	600	610	620	630	640	650
nc1aoms	434	440	450	460	470	480	490	500	510	520	530	540	550	560	570	580	590	600	610	620	630	640	650
Full	640	650	660	670	680	690	700	710	720	730	740	750	760	770	780	790	800	810	820	830	840	850	860
EC1-3	640	650	660	670	680	690	700	710	720	730	740	750	760	770	780	790	800	810	820	830	840	850	860
nc1aoms	640	650	660	670	680	690	700	710	720	730	740	750	760	770	780	790	800	810	820	830	840	850	860

B

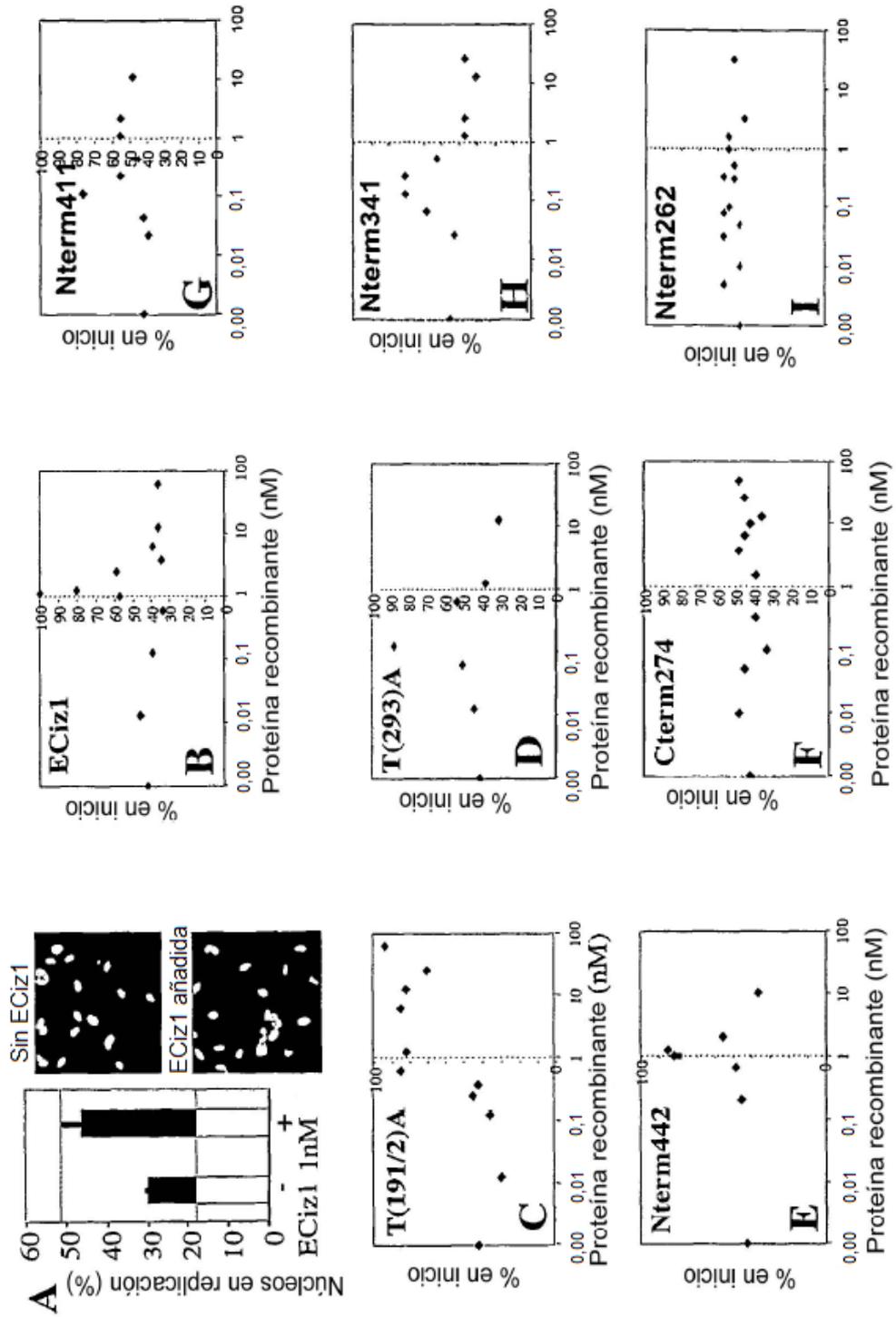
Figura 2B



Uso de exones en la *Ciz1*  
embrionaria de ratón



Figura 3A a I



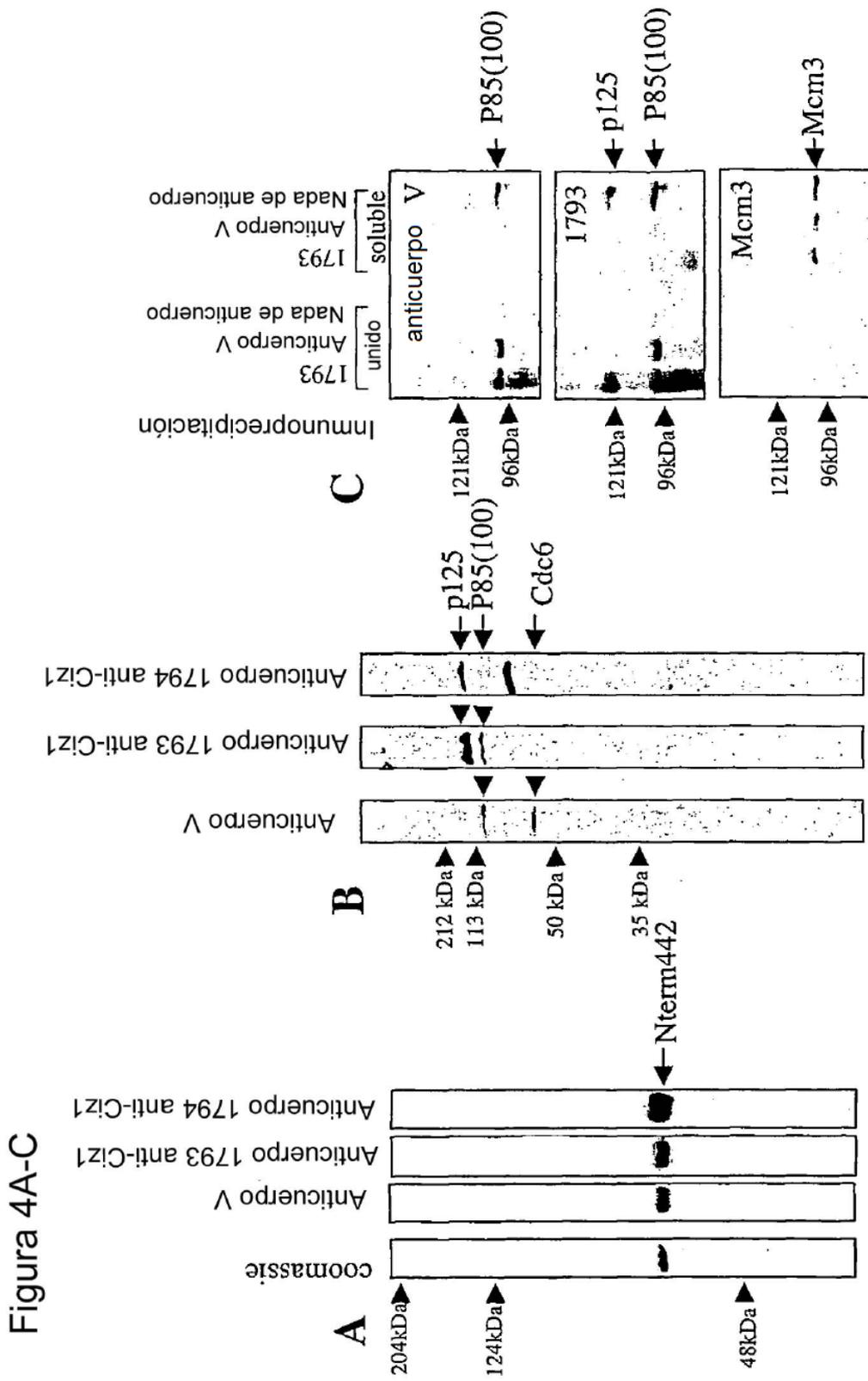


Figura 5A-D

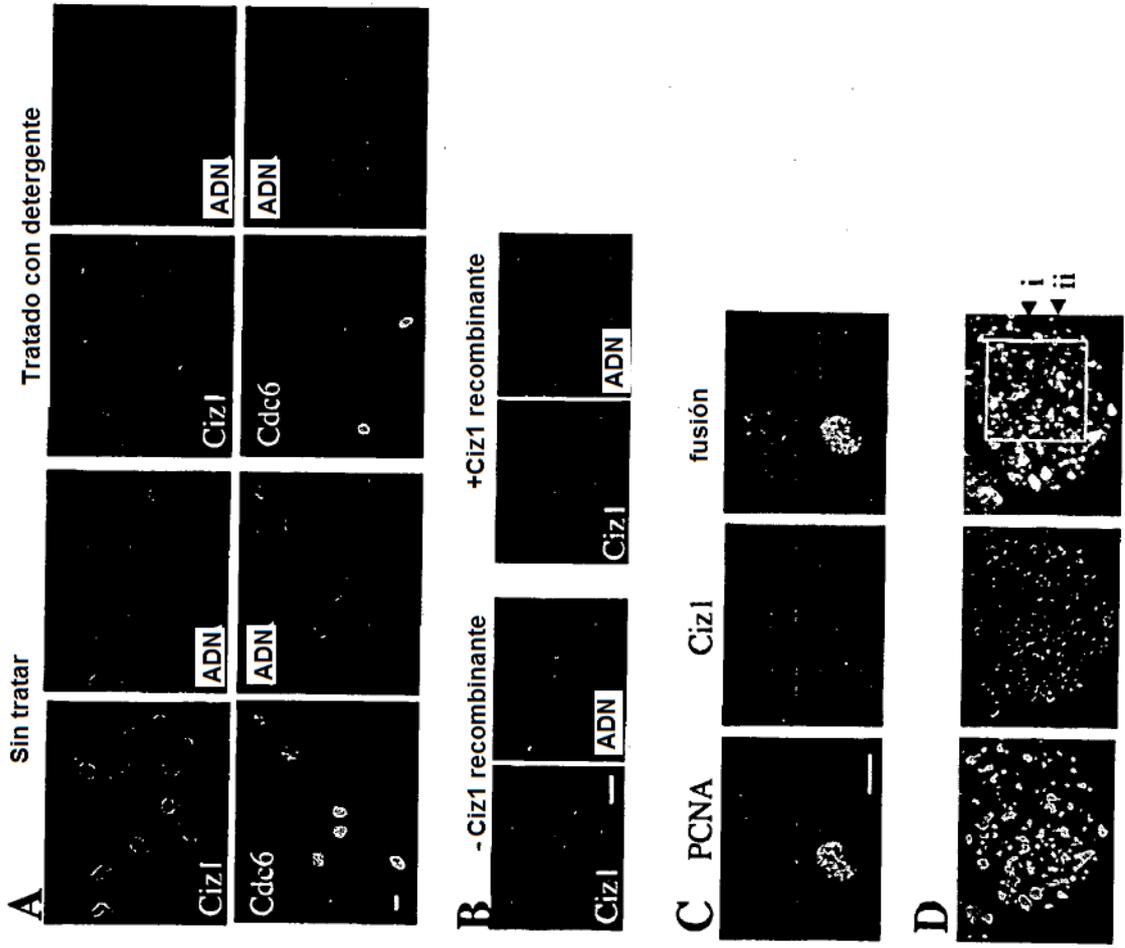


Figura 5E-F

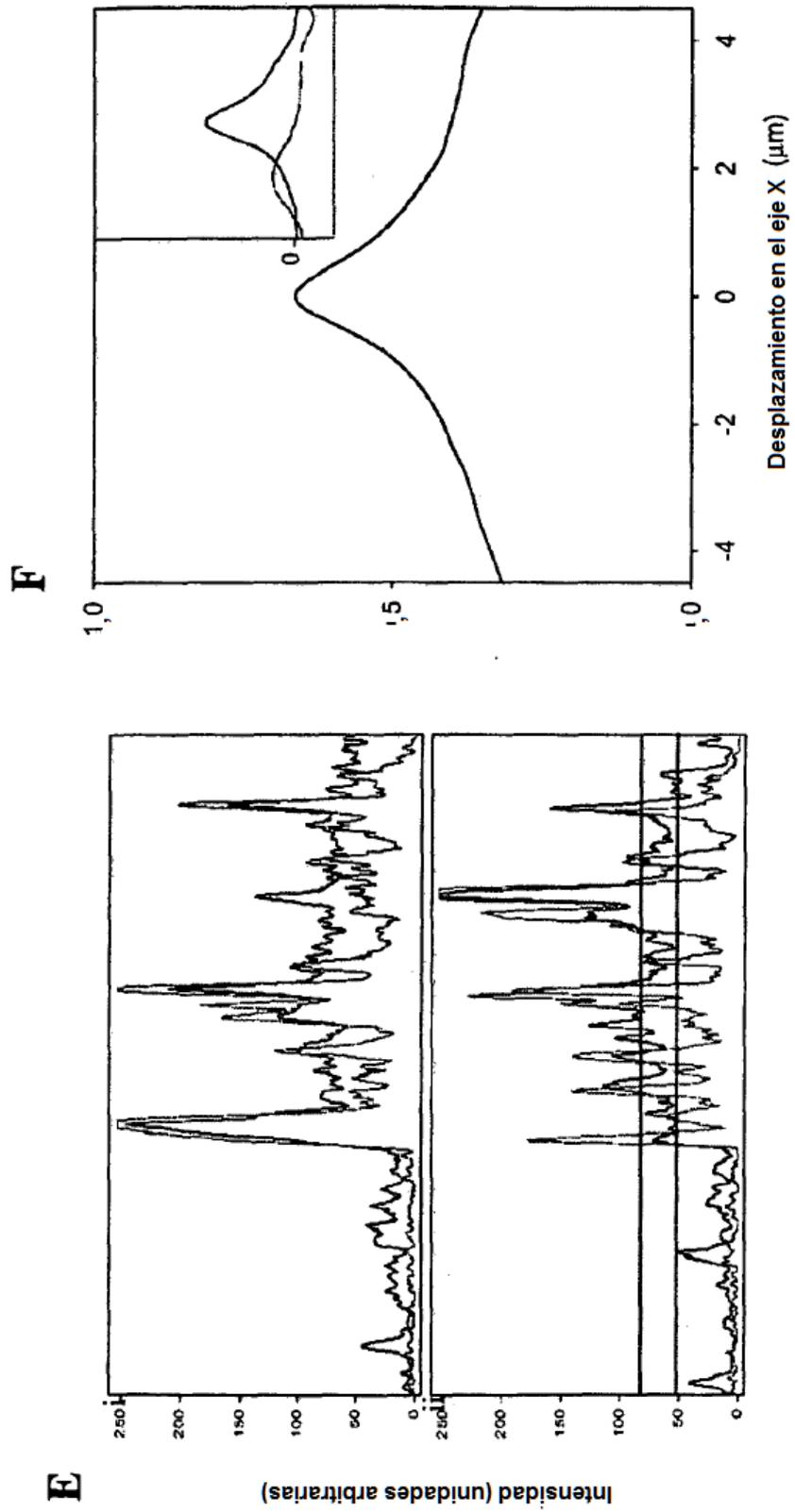


Figura 6A y B

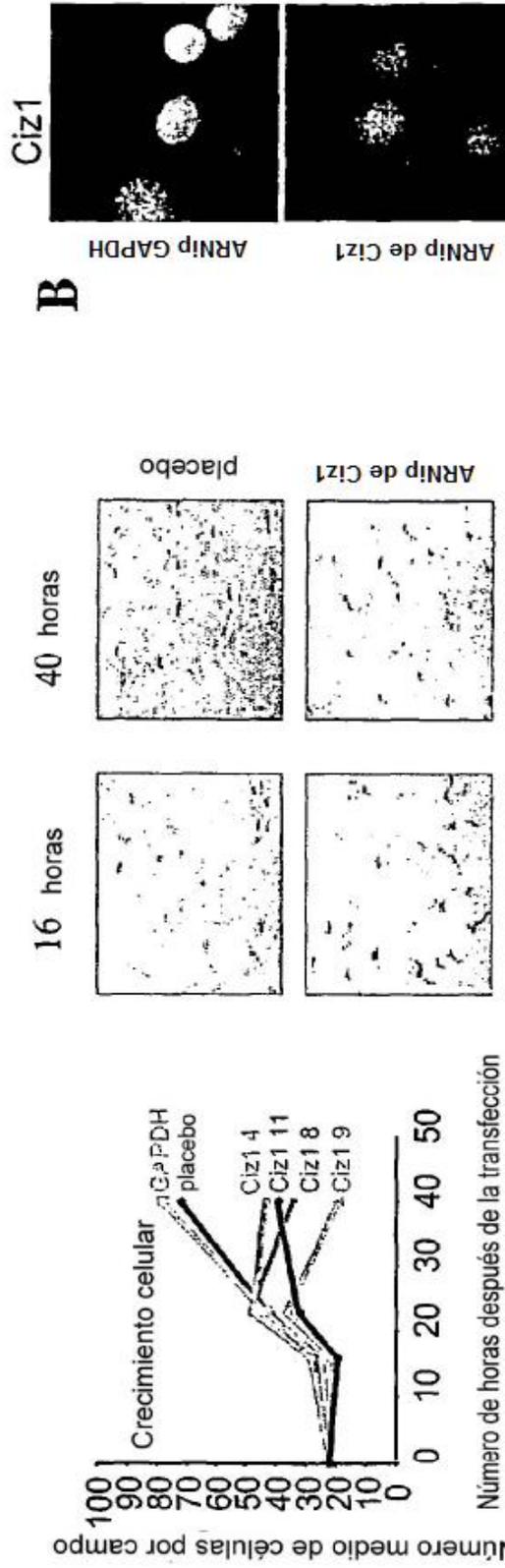


Figura 6C a F

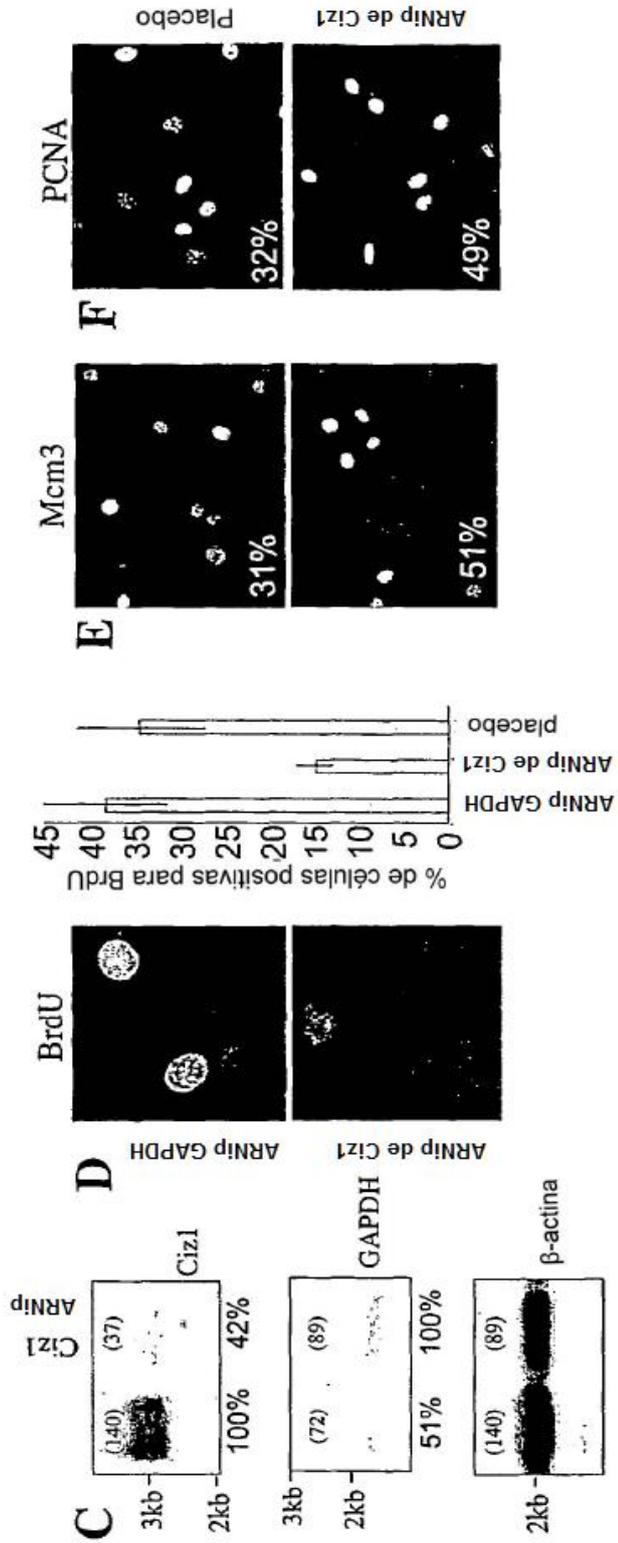
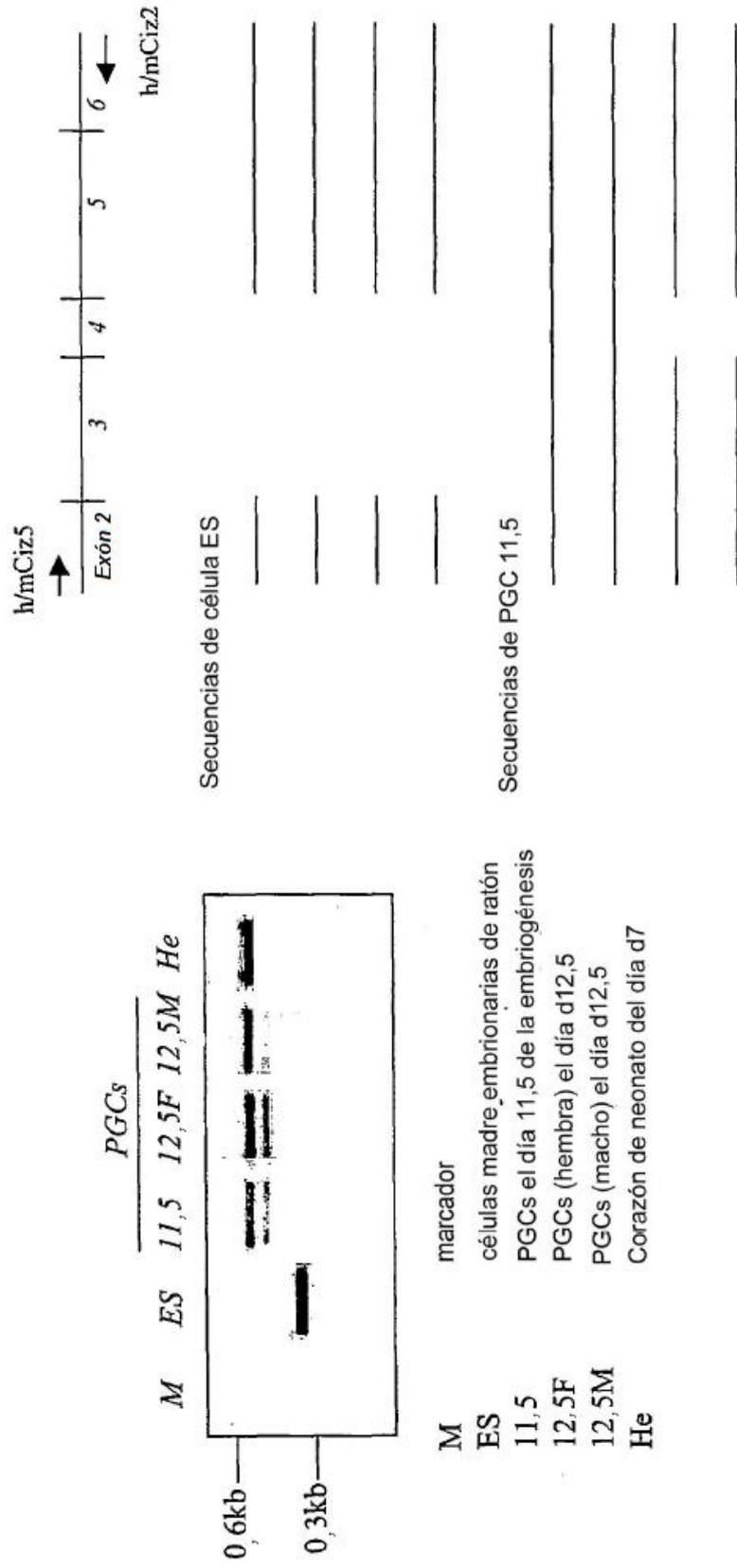


Figura 7



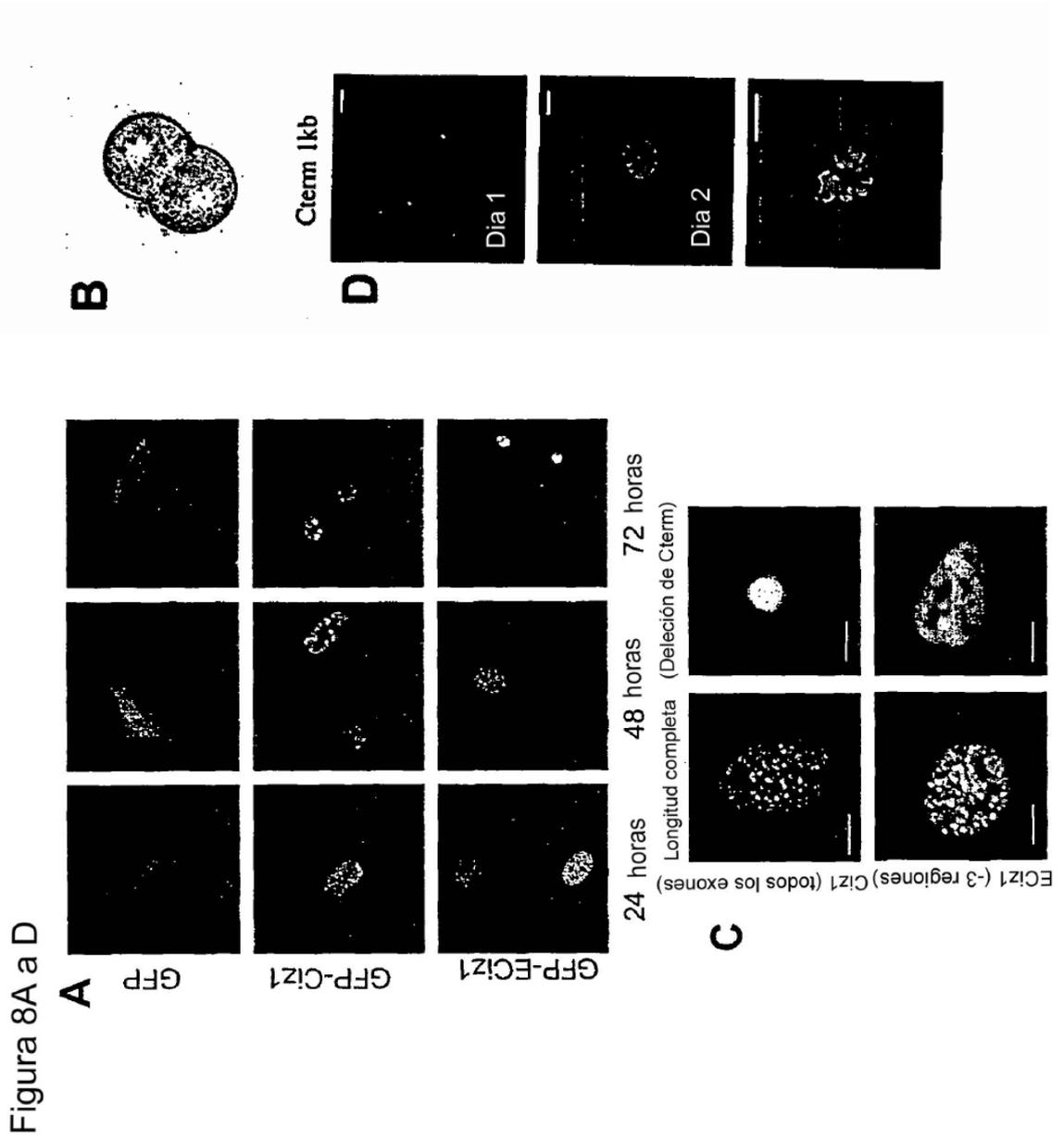
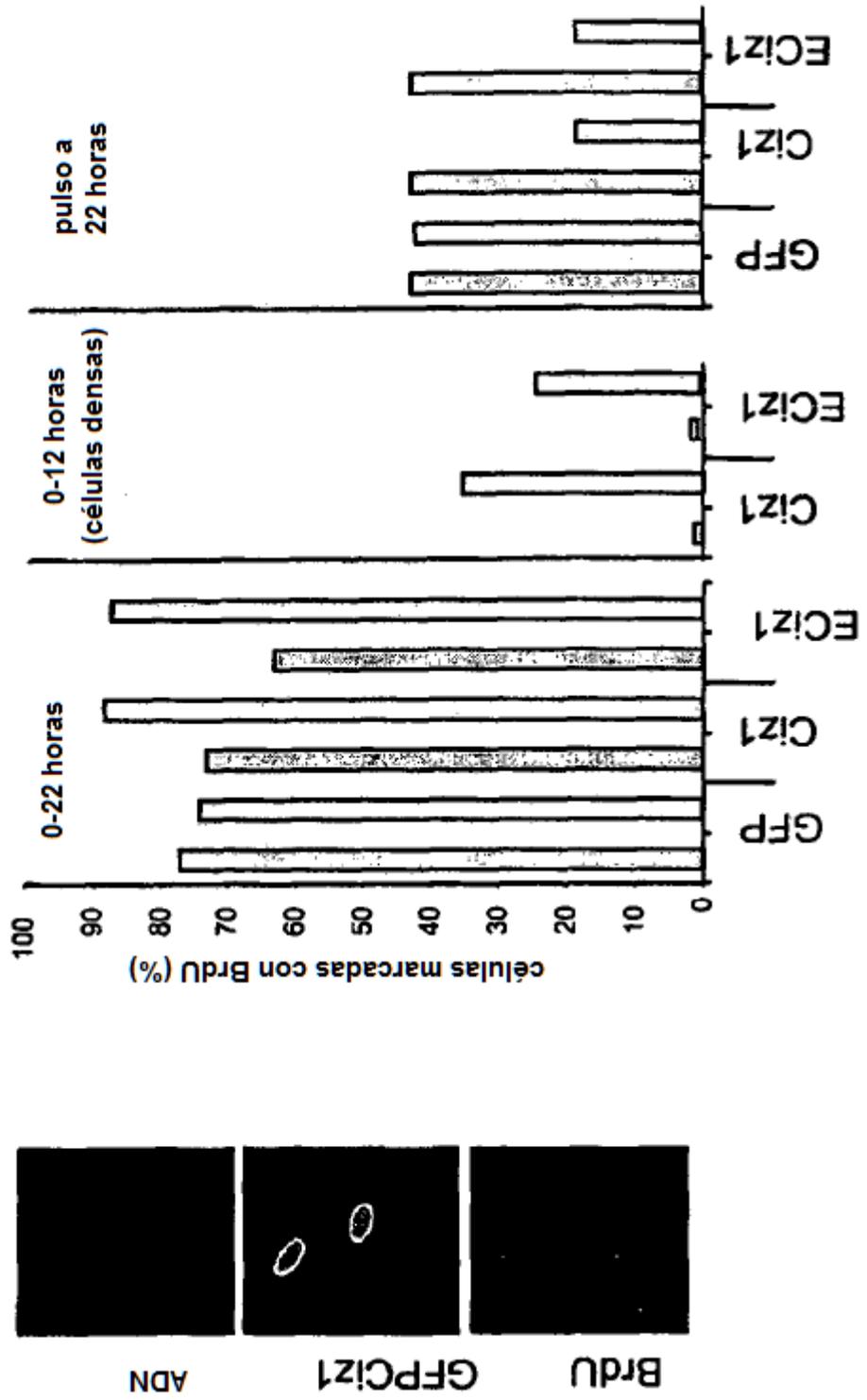
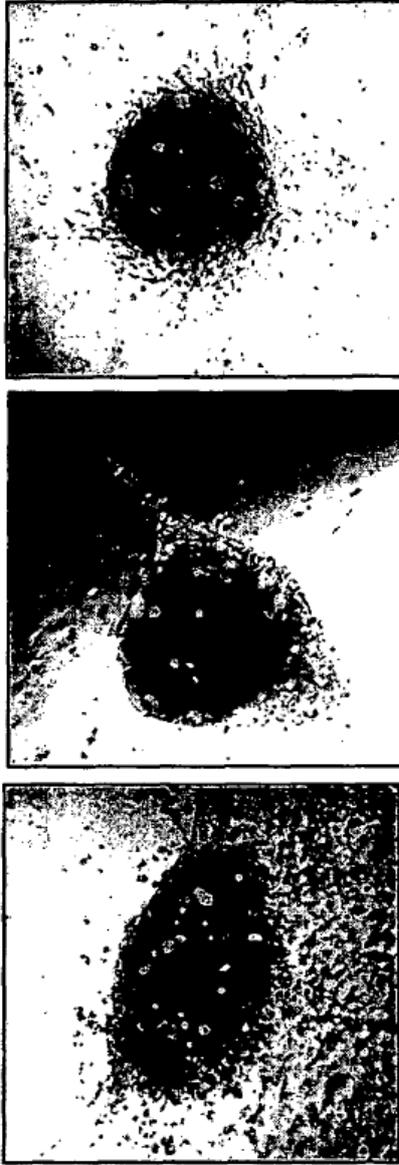


Figura 8E



**Figura 9A**

*ECizJ Nterm442*



*CizJ Nterm* (equivalente a 442)

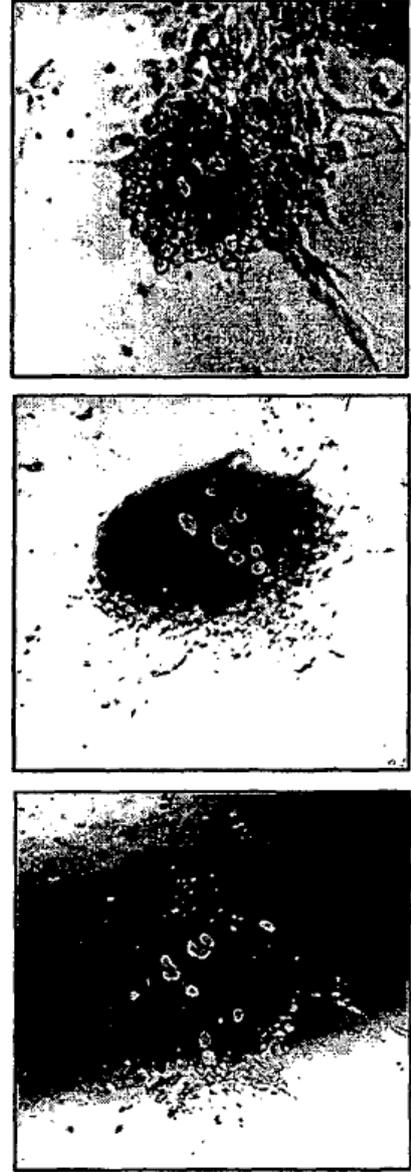


Figura 10

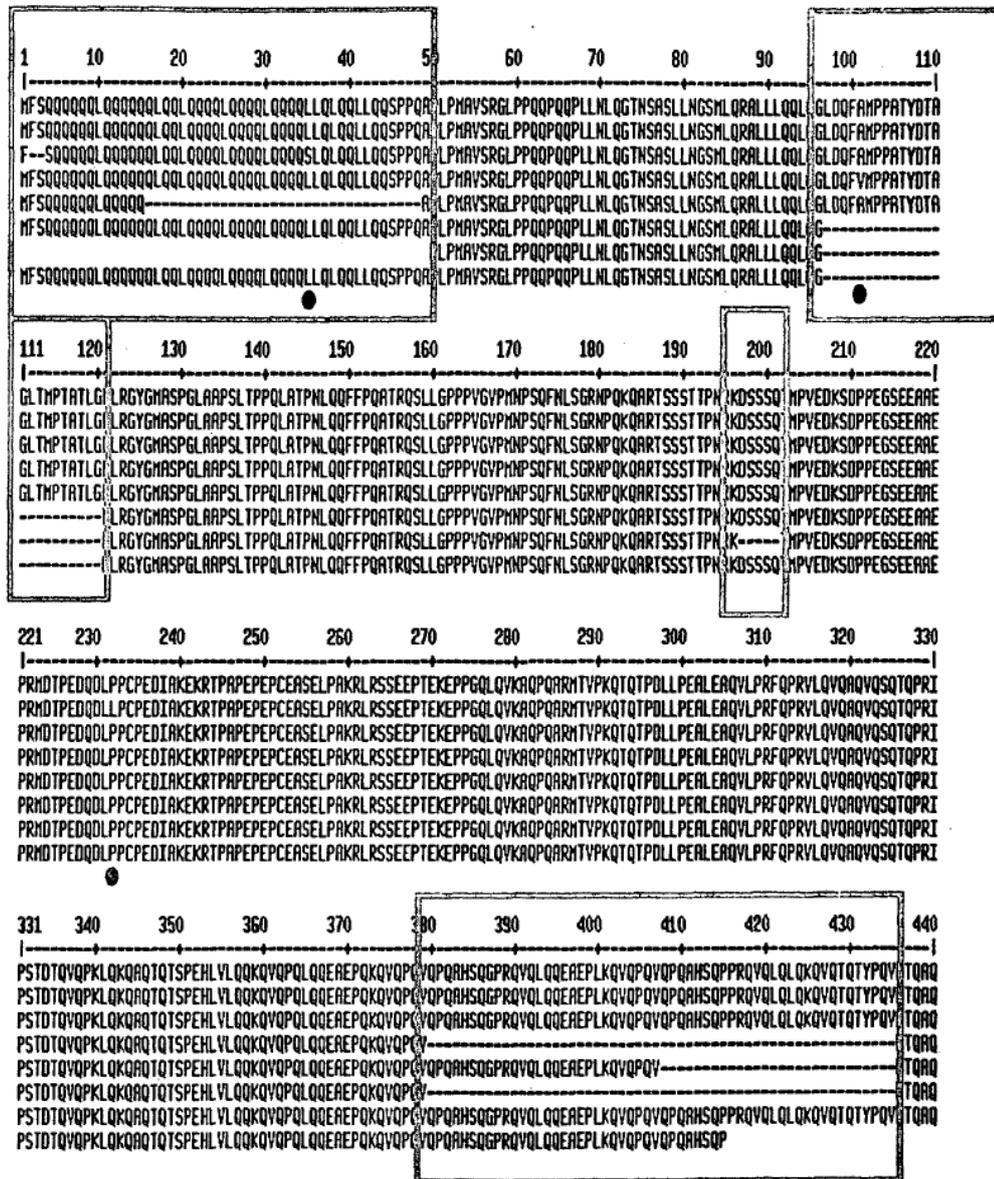


Figura 10 continuación

441 450 460 470 480 490 500 510 520 530 540 550  
 |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|  
 PSVQPQEHPPAQVSVQPPEQTHEQPHTOPQVSLLAPEQTPVYVHVCLEMPDRAVEAGGGMEKTLPEPVGTQVSMEEIQNESACGLDVGECEENRAREMPGVWAGGGS�KY  
 PSVQPQEHPPAQVSVQPPEQTHEQPHTOPQVSLLAPEQTPVYVHVCLEMPDRAVEAGGGMEKTLPEPVGTQVSMEEIQNESACGLDVGECEENRAREMPGVWAGGGS�KY  
 PSVQPQEHPPAQVSVQPPEQTHEQPHTOPQVSLLAPEQTPVYVHVCLEMPDRAVEAGGGMEKTLPEPVGTQVSMEEIQNESACGLDVGECEENRAREMPGVWAGGGS�KY  
 PSVQPQEHPPAQVSVQPPEQTHEQPHTOPQVSLLAPEQTPVYVHVCLEMPDRAVEAGGGMEKTLPEPVGTQVSMEEIQNESACGLDVGECEENRAREMPGVWAGGGS�KY  
 PSVQPQEHPPAQVSVQPPEQTHEQPHTOPQVSLLAPEQTPVYVHVCLEMPDRAVEAGGGMEKTLPEPVGTQVSMEEIQNESACGLDVGECEENRAREMPGVWAGGGS�KY  
 PSVQPQEHPPAQVSVQPPEQTHEQPHTOPQVSLLAPEQTPVYVHVCLEMPDRAVEAGGGMEKTLPEPVGTQVSMEEIQNESACGLDVGECEENRAREMPGVWAGGGS�KY

551 560 570 580 590 600 610 620 630 640 650 660  
 |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|  
 TILQSSDSRAFSTVPLTPVPRPSDSVSTPARTSTPSKQALQFFCYICKASCSSQEFQDHMSEPHQQRLEIQTQHSQACL L SLLPVPRDVLTEDEEPPRRRCHTCQ  
 TILQSSDSRAFSTVPLTPVPRPSDSVSTPARTSTPSKQALQFFCYICKASCSSQEFQDHMSEPHQQRLEIQTQHSQACL L SLLPVPRDVLTEDEEPPRRRCHTCQ

661 670 680 690 700 710 720 730 740 750 760 770  
 |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|  
 LYYMGDLIQHRTQDHKIAKQSLRPFCTVCHRYFKTPRKFVEHVKSQGHKDKAKELKSLEKEIAGQDEDFITVDVAVGCFEGDEEEEEDEDEEETVEEELCKQVRSRD  
 LYYMGDLIQHRTQDHKIAKQSLRPFCTVCHRYFKTPRKFVEHVKSQGHKDKAKELKSLEKEIAGQDEDFITVDVAVGCFEGDEEEEEDEDEEETVEEELCKQVRSRD  
 LYYMGDLIQHRTQDHKIAKQSLRPFCTVCHRYFKTPRKFVEHVKSQGHKDKAKELKSLEKEIAGQDEDFITVDVAVGCFEGDEEEEEDEDEEETVEEELCKQVRSRD  
 LYYMGDLIQHRTQDHKIAKQSLRPFCTVCHRYFKTPRKFVEHVKSQGHKDKAKELKSLEKEIAGQDEDFITVDVAVGCFEGDEEEEEDEDEEETVEEELCKQVRSRD  
 LYYMGDLIQHRTQDHKIAKQSLRPFCTVCHRYFKTPRKFVEHVKSQGHKDKAKELKSLEKEIAGQDEDFITVDVAVGCFEGDEEEEEDEDEEETVEEELCKQVRSRD  
 LYYMGDLIQHRTQDHKIAKQSLRPFCTVCHRYFKTPRKFVEHVKSQGHKDKAKELKSLEKEIAGQDEDFITVDVAVGCFEGDEEEEEDEDEEETVEEELCKQVRSRD

771 780 790 800 810 820 830 840 850 860 870 880  
 |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|  
 ISREEKGETYSPNTAYGVDFLVPVNGYICRICHKFYHNSGAQLSHCKSLGHFENLQKYKAKNPSPTTRPVSRCAINARNAL TALFTSSGRPPSPNTQDKTPSKV  
 ISREEKGETYSPNTAYGVDFLVPVNGYICRICHKFYHNSGAQLSHCKSLGHFENLQKYKAKNPSPTTRPVSRCAINARNAL TALFTSSGRPPSPNTQDKTPSKV  
 ISREEKGETYSPNTAYGVDFLVPVNGYICRICHKFYHNSGAQLSHCKSLGHFENLQKYKAKNPSPTTRPVSRCAINARNAL TALFTSSGRPPSPNTQDKTPSKV  
 ISREEKGETYSPNTAYGVDFLVPVNGYICRICHKFYHNSGAQLSHCKSLGHFENLQKYKAKNPSPTTRPVSRCAINARNAL TALFTSSGRPPSPNTQDKTPSKV  
 ISREEKGETYSPNTAYGVDFLVPVNGYICRICHKFYHNSGAQLSHCKSLGHFENLQKYKAKNPSPTTRPVSRCAINARNAL TALFTSSGRPPSPNTQDKTPSKV  
 ISREEKGETYSPNTAYGVDFLVPVNGYICRICHKFYHNSGAQLSHCKSLGHFENLQKYKAKNPSPTTRPVSRCAINARNAL TALFTSSGRPPSPNTQDKTPSKV

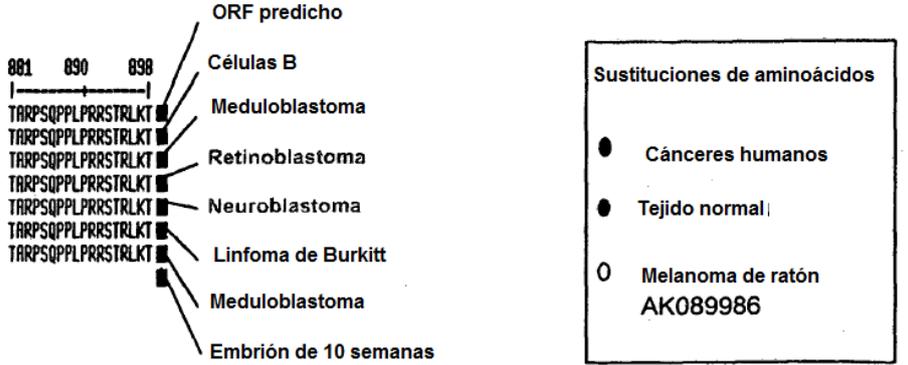
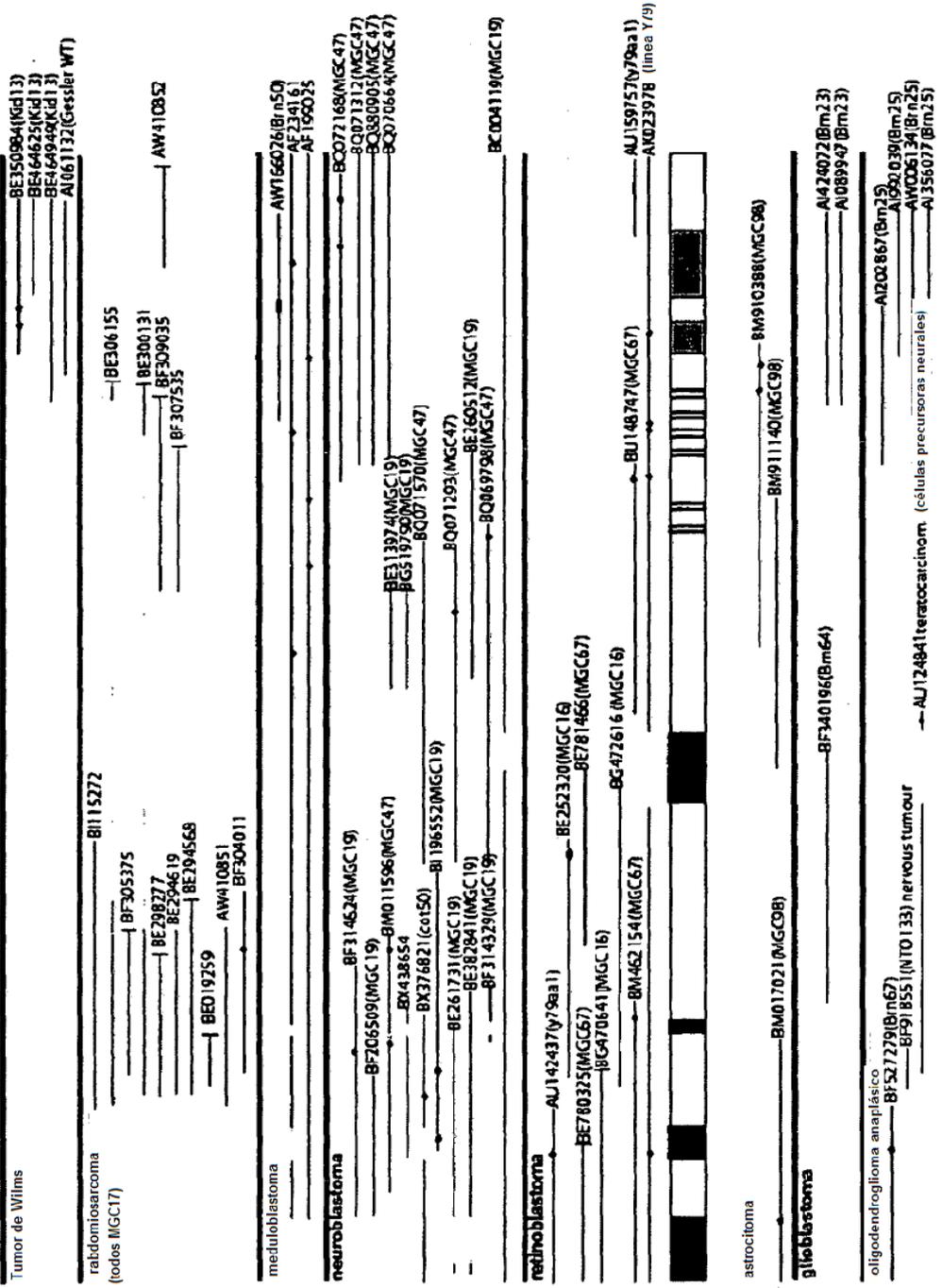
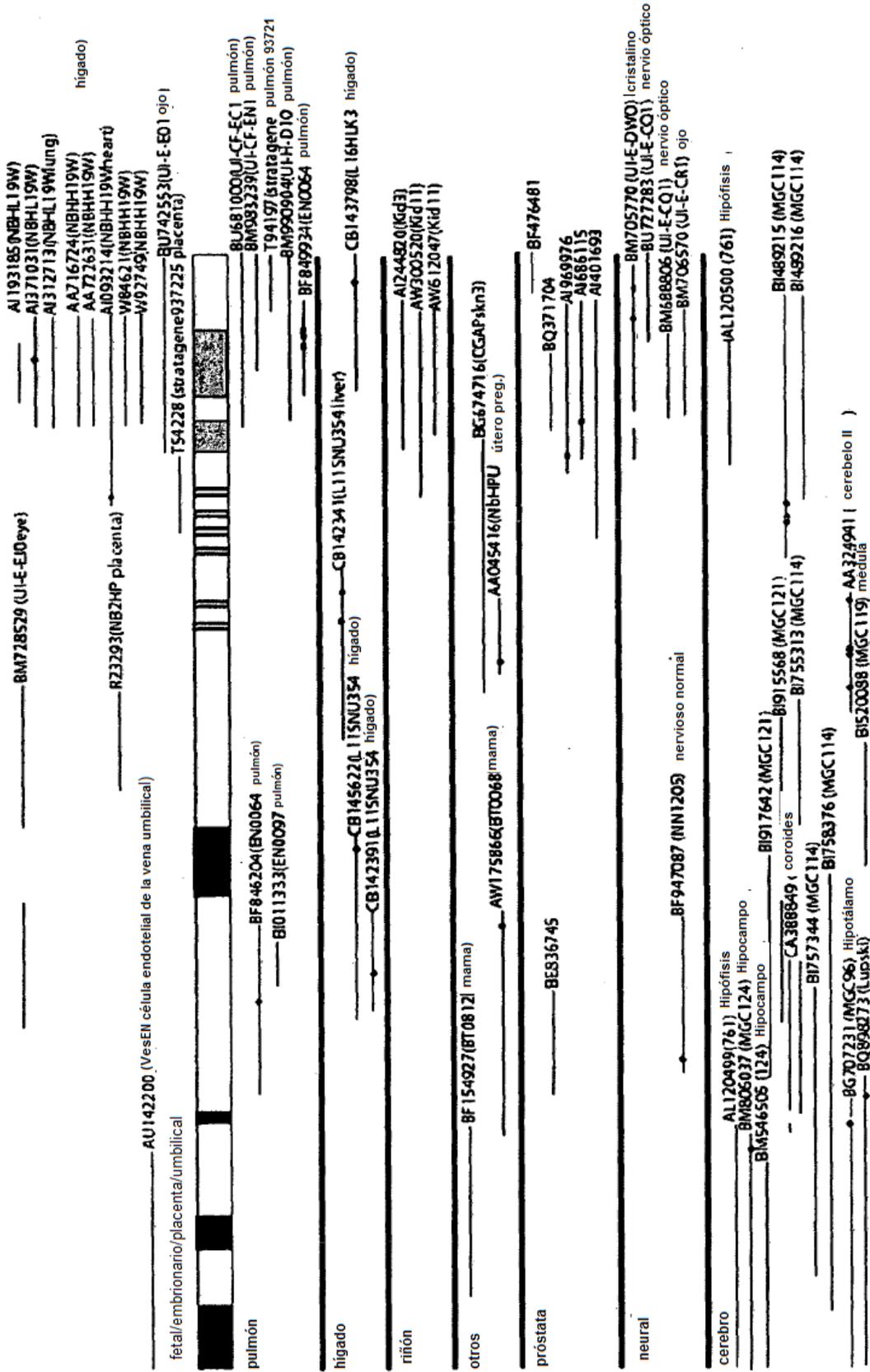


Figura 11a



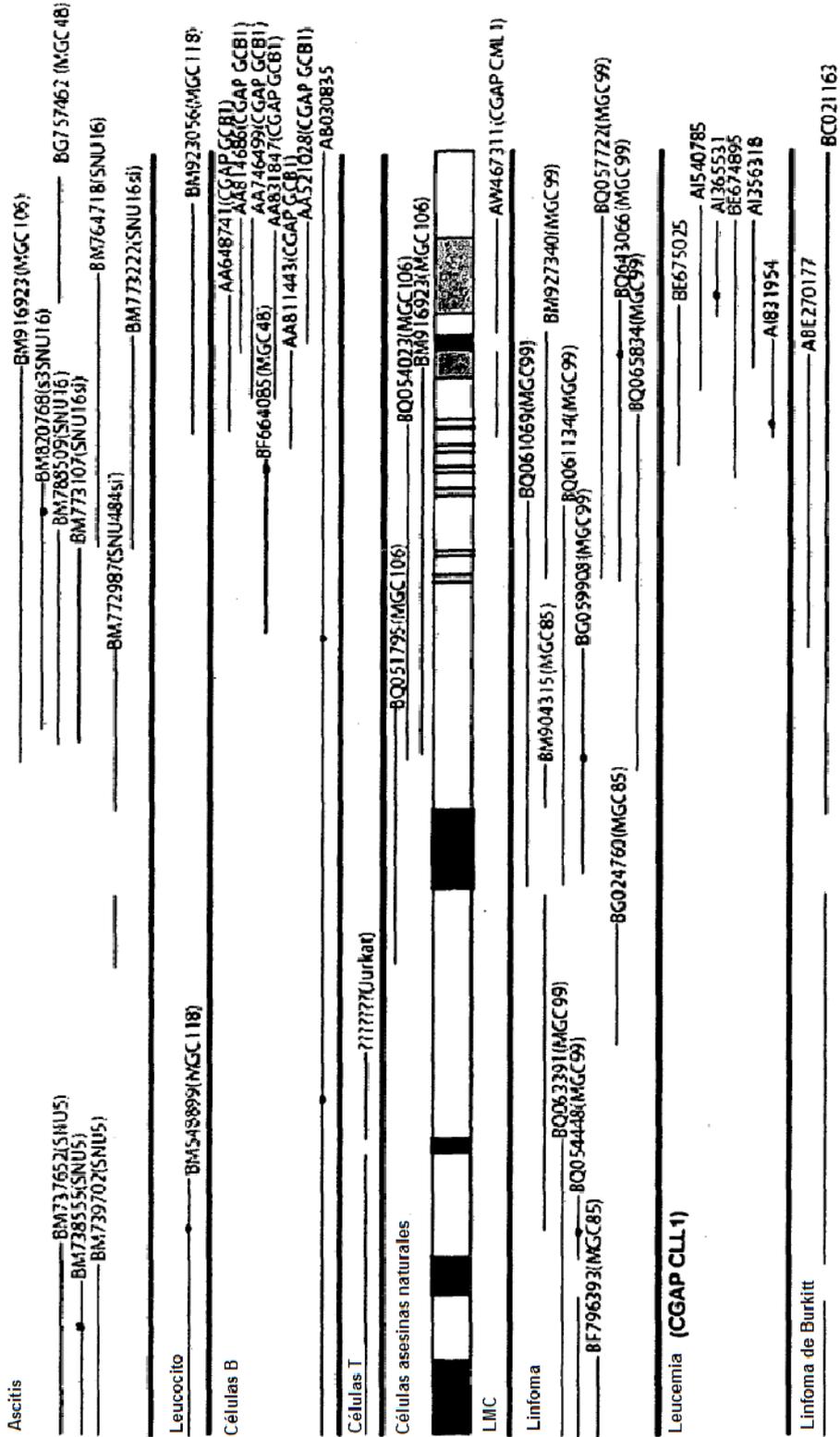
cánceres neurales del adulto      cánceres pediátricos

Figura 11b



No canceroso

Figura 11c



no canceroso

leucemia /linfoma



Figura 11e

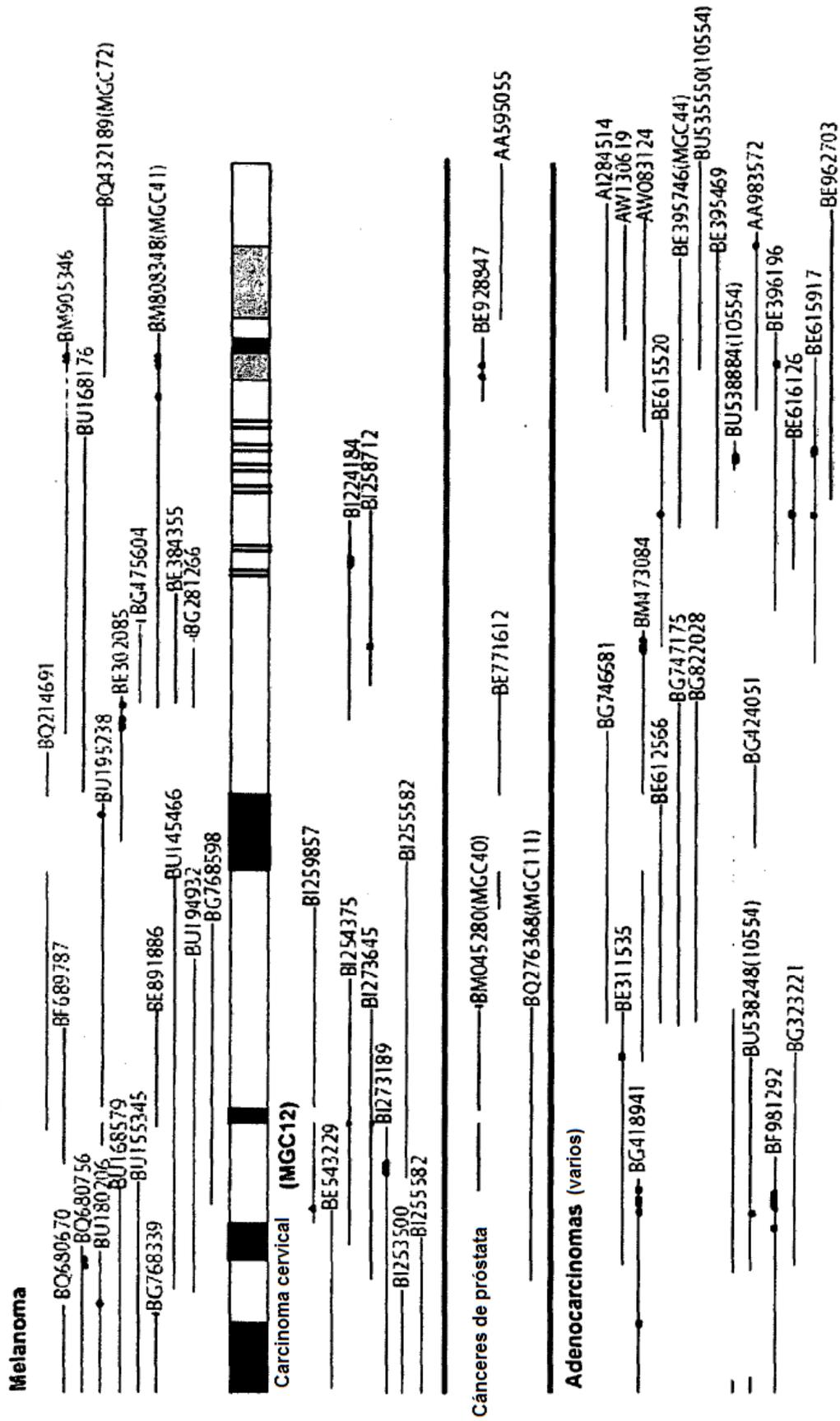


Figura 11F

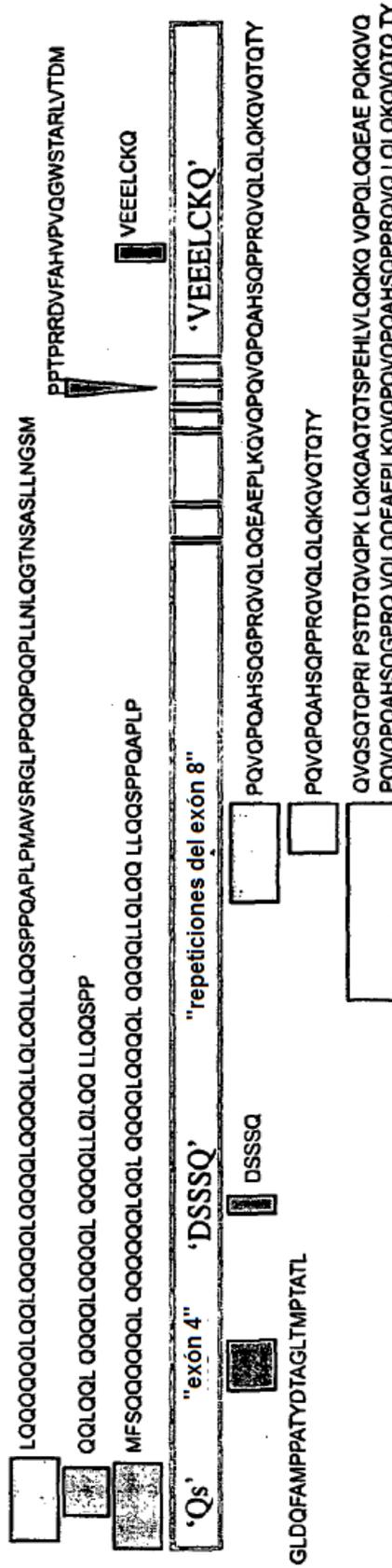




Figura 12B

Ejemplos de productos de PCR

Ewings 6

---



---



---



---



---



---

Neuroblastoma 2

---



---



---



---



---



---

HEK293

---



---



---



---



---



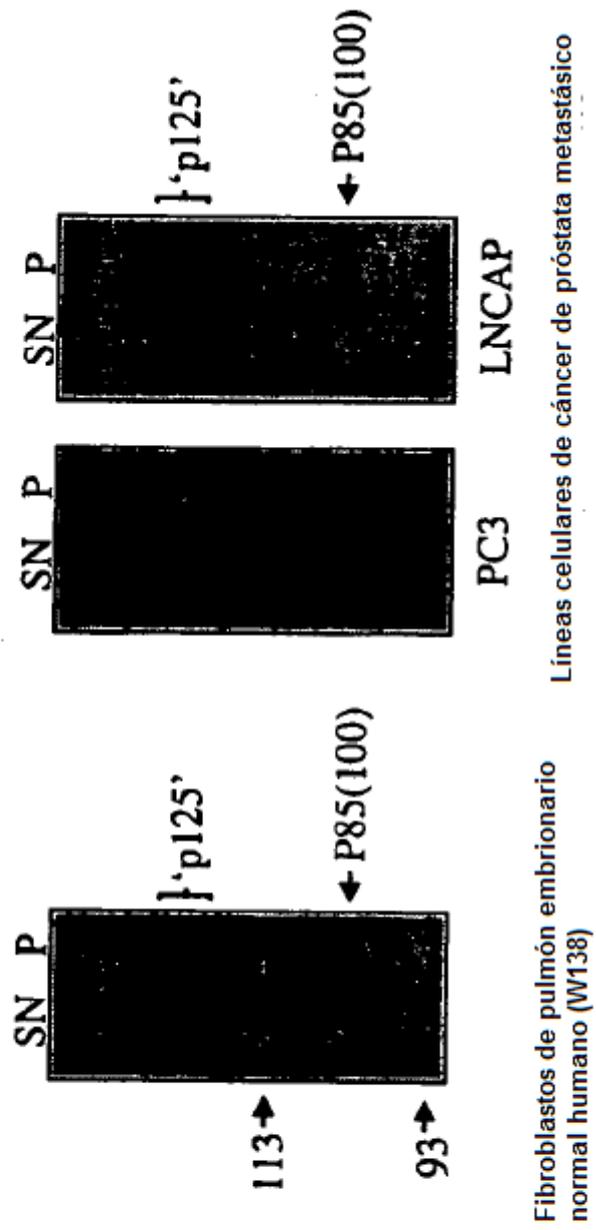
---

Resumen de productos de PCR

	1	2	3	4	5	6	N1	N2	293
'DSSSQ'	0	0	0	0	2	0	1	2	0
'exon4'	1*	0	1*	3	0	3	1	0	0
'FL'	4	1	5	2	2	3	8	3	4
otros	0	0	0	1	0	0	1	0	0

ESFTs                      Neuroblastomas  
DSSSQ 2/26                DSSSQ 3/16  
Exón 4 8/26                Exón 4 1/16

Figura 13A



**Figura 13B**



Corte y empalme de proteínas para excluir las repeticiones del exón 8



Escisión para excluir el extremo N variable

## Figura 14

CATGTTCAAC CCGCAACTCC AGCAGCAGCA ACAGTTGCAG CAGCAGCAGC  
 AACAGTTGCA GCAGCAGCTC CAGCAGCAGC AGCTCCAGCA GCAGCAACAG  
 CAGATACTGC AGCTCCAACA GCTGCTGCAA CAGTCCCCAC CACAGGCCTC  
 CTTGTCCATT CCTGTAGCC GGGGCCTCCC CCAGCAGTCA TCCCCGCAAC  
 AGCTTCTGAG TCTCCAGGGC CTCCACTCGA CCTCCCTGCT CAATGGCCCC  
 ATGCTGCAAA GAGCTTTGCT CCTACAGCAG TTGCAAGGAC TGGACCAAGT  
 TGCAATGCCA CCAGCCACGT ATGACGGTGC CAGCCTCACC ATGCCTACGG  
 CAACACTGGG TAACCTCCGT GCTTTCAATG TGACAGCCCC AAGCCTAGCA  
 GCTCCCAGCC TTACACCACC CCAGATGGTC CCCCCAATC TGCAGCAGTT  
 CTTTCCCCAG GCTACTCGAC AGTCTCTGCT GGGGCCTCCT CTGTTGGGG  
 TCCCAATAAA CCCTTCTCAG CTCAACCACT CAGGGAGGAA CACCCAGAAA  
 CAGGCCAGAA CCCCCTCTTC CACCACCCCC AATCGCAAGG ATTCTTCTTC  
 TCAGACGGTG CCTCTGGAAG ACAGGGAAGA CCCCACAGAG GGGTCTGAGG  
 AAGCCACGGA GCTCCAGATG GACACATGTG AAGACCAAGA TTCACTAGTC  
 GGTCCAGATA GCATGCTGAG TGAGCCCCAA GTGCCTGAGC CTGAGCCCTT  
 TGAGACATTG GAACCACCAG CCAAGAGGTG CAGGAGCTCA GAGGAGTCCA  
 CCGAGAAAGG CCCTACAGGG CAGCCACAAG CAAGGGTCCA GCCTCAGACC  
 CAGATGACAG CACCAAAGCA GACACAGACC CCGGATCGGC TGCCTGAGCC  
 ACCAGAAGTC CAAATGCTGC CGCGTATCCA GCCACAGGCA CTGCAGATCC  
 AGACCCAGCC AAAGCTGCTG AGGCAGGCAC AGACACAGAC CTCTCCAGAG  
 CACTTAGCGC CCCAGCAGGA TCAGGTAGAG CCACAGGTAC CATCACAGCC  
 CCCATGGCAG TTGCAGCCAC GGGAGACAGA CCCACCGAAC CAAGCTCAGG  
 CACAGACCCA GCCTCAGCCC CTCTGGCAGG CGCAGTCACA GAAGCAGGCC  
 CAGACACAGG CACATCCACA GGTACCCACC CAAGCACAGT CACAGGAGCA  
 GACATCAGAG AAGACCCAGG ACCAGCCTCA GACCTGGCCA CAGGGGTCAG  
 TACCCACC ACCAACAAGCG TCAGGTCCAG CCTGTGCCAC GGAACCCAG  
 CTATCCTCTC ACGCTGCAGA AGCTGGGAGT GACCCAGACA AGGCCTTGCC  
 AGAACCAAGTA AGTGCCAGAG GCAGTGAAGACAGGAGCCGG GAGGCTCCG  
 CTGGTGGCCT GGATTTGGGA GAATGTGAAA AGAGAGCGGG AGAGATGCTG  
 GGGATGTGGG GGGCTGGGAG CTCCCTGAAG GTCACCATCC TGCAGAGTAG  
 CAACAGCCGG GCCTTTAACA CCACACCCCT CACATCTGGA CCTCGCCCTG  
 GGGACTCTAC CTCTGCCACC CCTGCCATTG CCAGCACACC CTCCAAGCAA  
 AGCCTCCAGT TCTTCTGCTA CATCTGCAAG GCCAGCAGCA GCAGCCAGCA  
 GGAGTTCCAG GATCACATGT CAGAGGCTCA GCACCAACAG CGGCTTGGGG  
 AAATACAACA CTCGAGCCAG ACCTGCCTGC TGTCCTGCT GCCCATGCCT  
 CGGGACATCC TGGAGAAAGA AGCGGAAGAT CCTCCGCCA AACCGTGGTG  
 CAACACCTGC CAGGTGTA CTGTTGGAGA CTTGATCCAG CACCGTAGGA  
 CACAGGAGCA CAAGGTTGCC AAACAATCCC TGAGGCCCTT CTGCACCATA  
 TGCAACCGTT ACTTCAAGAC CCCTCGAAAG TTTGTGGAGC ACGTGAAGTC  
 CCAGGGACAC AAGGACAAGG CCCAAGAGCT GAAGACACTTGAAAAGGAGA  
 CAGGCAGCCC AGATGAGGAC CACTTCATCA CTGTGGACGC CGTCCGGTTC  
 TTTGAGAGTG GTCAAGAAGA GGACGAGGAT GACGACGAGGAAGAAGAAGA  
 AGAAGGAGAG ATTGAGGCTG AGGAGGAATT CTGCAAGCAG GTGAAGCCGA  
 GAGAAACATC CTCAGAGCAA GGAAGGGCT CTGAGACGTA CAACCCCAAC  
 ACAGCCTATG GTGAGGATTT CCTGGTGCCA GTGATGGGCT ATGTCTGTCA  
 AATCTGTAC AAGTTCTACG ACAGCAACTC AGAATTGCGG CTTTCTCACT  
 GCAAGTCCCT GGCCCACTTT GAGAACCTGC AGAAATACAA AGCCAAGAAC  
 CCAAGCCCTC CTCCTACCCG GCCTGTGAGC CGCAAGTGTG CCATCAACGC  
 CCGCAACGCC CTGACTGCAC TGTTCACTC TAGCCACCAG CCCAGCCCC  
 AGGACACAGT GAAAATGCCC AGCAAGGTGA AGCCTGGATC CCCCAGACTC  
 CCTCCTCCCC TTCGGCGCTC AACACGCCTC AAAACCTGAT AGAGGGAGCT  
 CTGGCCACTC AGCCTGACTA AGGCTCAGTC TGCTAATGCT TCCTAGGTAT  
 CTGTGTAGAA ATGTTCAAGT GGTGGTGT TTTACTCAA ATCCAATAAA  
 GAGTCAGTAG TTTGCCAAAA AAAAAAAAAA AAAAAA

## Figura 15

TGGGGGCTGC GGGGCCGGCC CATCCGTGGG GGGCACTTGA GCGTTGAGGG  
 CGCGCGGGGA GGGCAGCCAC CATGTTTCAGC CAGCAGCAGC AGCAGTCCA  
 GCAACAGCAG CAGCAGCTCC AGCAGTTACA GCAGCAGCAG CTCACAGCAGC  
 AGCAATTGCA GCAGCAGCAG TTA CTGCAGC TCCAGCAGCT GCTCCAGCAG  
 TCCCCACCAC AGGCCCGGTT GCCCATGGCT GTCAGCCGGG GGCTCCCCC  
 GCAGCAGCCA CAGCAGCCGC TTCTGAATCT CCAGGGCACC AACTCAGCCT  
 CCTCCTCAA CGGCTCCATG CTGCAGAGAG CTTTGCTTTT ACAGCAGTTG  
 CAAGGACTGG ACCAGTTTGC AATGCCACCA GCCACGTATG AACTGGCCG  
 TCTCACCATG CCCACAGCAA CACTGGGTAA CCTCCGAGGC TATGGCATGG  
 CATCCCCAGG CCTCGCAGCC CCCAGCCTCA CACCCCCACA ACTGGCCACT  
 CCAAATTTGC AACAGTTCTT TCCCCAGGCC ACTCGCCAGT CCTTGCTGGG  
 ACCTCCTCCT GTTGGGGTCC CCATGAACCC TCCCCAGTTC AACCTTTCAG  
 GACGGAACCC CCAGAAACAG GCCCGGACCT CCTCCTCTAC CACCCCCAAT  
 CGAAAGGATT CTTCTTCTCA GACAATGCCT GTGGAAGACA AGTCAGACCC  
 CCCAGAGGGG TCTGAGGAAG CCGCAGAGCC CCGGATGGAC ACACCAGAAG  
 ACCAAGATTT ACCGCCCTGC CCAGAGGACA TCGCCAAGGA AAAACGCCT  
 CCAGCACCTG AGCCTGAGCC TTGTGAGGGC TCCGAGCTGC CAGCAAAGAG  
 ATTGAGGAGC TCAGAAGAGC CCACAGAGAA GGAACCTCCA GGGCAGTTAC  
 AGGTGAAGGC CCAGCCGAG GCCCGGATGA CAGTACCGAA ACAGACACAG  
 ACACCAGACC TGCTGCCTGA GGCCCTGGAA GCCCAAGTGC TGCCACGATT  
 CCAGCCACGG GTCTGCAGG TCCAGGCCCA GGTGCAGTCA CAGACTCAGC  
 CGCGGATACC ATCCACAGAC ACCCAGGTGC AGCCAAAGCT GCAGAAGCAG  
 GCGCAAACAC AGACCTCTCC AGAGCACTTA GTGCTGCAAC AGAAGCAGT  
 GCAGCCACAG CTGCAGCAG AGGCAGAGCC ACAGAAGCAGGTGCAGCCAC  
 AGGTACAGCC ACAGGCACAT TCACAGGGCC CAAGGCAGGT GCAGCTGCAG  
 CAGGAGGCAG AGCCGCTGAA GCAGGTGCAG CCACAGGTGC AGCCCCAGGC  
 ACATTCACAG CCCCCAAGGC AGGTGCAGCT GCAGCTGCAG AAGCAGGTCC  
 AGACACAGAC ATATCCACAG GTCCACACAC AGGCACAGCC AAGCGTCCAG  
 CCACAGGAGC ATCCTCCAGC GCAGGTGTCA GTACAGCCAC CAGAGCAGAC  
 CCATGAGCAG CCTCACACCC AGCCGAGGT GTCGTTGCTG GCTCCAGAGC  
 AAACACCACT TGTGGTTTCT GTCTGCGGGC TGGAGATGCC ACCTGATGCA  
 GTAGAAGCTG TTGGAGGCAT GGAAAAGACC TTGCCAGAGC CTGTGGGCAC  
 CCAAGTCAGC ATGGAAGAGA TTCAGAATGA GTCGGCCTGT GGCCTAGATG  
 TGGGAGAATG TGA AAAACAGA GCGAGAGAGA TGCCAGGGGTATGGGGCGCC  
 GGGGGCTCCC TGAAGGTAC CATTCTGCAG AGCAGTGACA GCCGGGCTT  
 TAGCACTGTA CCCCTGACAC CTGTCCCCCG CCCCAGTGAC TCCGTCTCCT  
 CCACCCCTGC GGCTACCAGC ACTCCCTCTA AGCAGGCCCT CCAGTTCTTC  
 TGCTACATCT GCAAGGCCAG CTGCTCCAGC CAGCAGGAGT TCCAGGCCA  
 CATGTCCGAG CCTCAGCACC AGCAGCGGCT AGGGGAGATC CAGCACATGA  
 GCCAAGCCTG CCTCCTGTCC CTGCTGCCCG TGCCCCGGGA CGTCTGGAG  
 ACAGAGGATG AGGAGCCTCC ACCAAGGCGC TGGTGCAACA CCTGCCAGCT  
 TACTACATG GGGGACCTGA TCCAACACCG CAGGACACAG GACCACAAGA  
 TTGCCAAACA ATCCTTGCGA CCCTTCTGCA CCGTTTGCAA CCGCTACTTC  
 AAAACCCCTC GCAAGTTTGT GGAGCACGTG AAGTCCCAGG GGCATAAGGA  
 CAAAGCCAAG GAGCTGAAGT CGCTTGAGAA AGAAATTGCT GGCCAAGATG  
 AGGACCACTT CATTACAGTG GACGCTGTGG GTTGCTTCTG GGGTGATGAA  
 GAAGAGGAAG AGGATGATGA GGATGAAGAAGAGATCGAGGTTGAGGAGGA  
 ACTCTGCAAG CAGGTGAGGT CCAGAGATAT ATCCAGAGAG GAGTGGAAGG  
 GCTCGGAGAC CTACAGCCCC AATACTGCAT ATGGTGTGGA CTTCTGGTG  
 CCCGTGATGG GCTATATCTG CCGCATCTGC CACAAGTTCT ATCACAGCAA  
 CTCAGGGGCA CAGCTCTCCC ACTGCAAGTC CCTGGGGCCAC TTTGAGAACC  
 TGCAGAAATA CAAGGCGGCC AAGAACCCCA GCCCACCAC CCGACTGTG  
 AGCCGCGGT GCGCAATCAA CGCCCGGAAC GCTTTGACAG CCCTGTTTAC  
 CTCCAGCGGC CGCCACCCT CCCAGCCCA CACCCAGGAC AAAACACCCA  
 GCAAGGTGAC GGCTCGACCC TCCCAGCCCC CACTACCTCG GCGCTCAACC  
 CGCTCAAAA CCTGATAGAG GGACCTCCCT GTCCCTAGCT TGCCATGGGT  
 CAGATCTGCT AATGCTTTT AGGAGTCTGC CTGGAAGCTT TGACATGGTT  
 CATGTTTTTA CTCAAATCC AATAAAACA GGTAGTTTGG CTGTGCAAAA  
 AAAAAAAAAA AAAAAAAAAA AA

## Figura 16

MFNPQLQQQQ QLLQQQQQLQ QQLQQQQQLQ QQQQILQLQQ LLQQSPPPQAS  
 LSIPVSRGLP QQSSPQQLLS LQGLHSTSL NGPMLQRALL LQQLQGLDQF  
 AMPPATYDGA SLTPTATLG NLRANVTAP SLAAPSLTPP QMVTNPNLQQF  
 FPQATRQSLG GPPVGVVPIN PSQLNHSGRN TQKQARTPSS TTPNRKDSSS  
 QTVPLEDRED PTEGSEATE LQMDTCEDQD SLVGPDSMLS EPQVPEPEPF  
 ETLEPPAKRC RSSEESTEKG PTGQPQARVQ PQTQMTAPKQ TQTPDRLPEP  
 PEVQMLPRIQ PQALQIQTOP KLLRQAQTOT SPEHLAPQQD QVEPQVPSQP  
 PWQLQPRETD PPNQAQAQTQ PQLWQAQSQ KQAQTQAHPQ VPTQAQSQEQ  
 TSEKTQDQPQ TWPQGSVPPP EQASGPACAT EPQLSSHAAE AGSDPKALP  
 EPVSAQSSSED RSREASAGGL DLGECEKRAG EMLGMWGAGS SLKVITILQSS  
 NSRAFNTTTL TSGPRPGDST SATPAIASTP SKQSLQFFCY ICKASSSSQQ  
 EFQDHMSEAQ HQQLGELQEH SSQTCLLSLL PMPRDILEKE AEDPPPKRWC  
 NTCQVYYVGD LIQHRRTQEH KVAKQSLRPF CTICNRYFKT PRKFVEHVKS  
 QGHKDKAQEL KLEKETGSP DEDHFTVDA VQCFESGQEE DEDDDEEEEEE  
 EGEIEAEEEF CKQVKPRETS SEQGKGSETY NPNTAYGEDF LVPVMGYVCQ  
 ICHKFYDSNS ELRLSHCKSL AHFENLQKYK AKNPSPPTTR PVSRRKCAINA  
 RNALTALFTS SHQPSPODTV KMPSKVKPGS PGLPPLRRS TRLKKT

### Figura 17

MF SQQQQQQQ QQQQQQQQ  
 LQRALLLQQQLQ GL DQFAMP PATYDTAGLT MPTATLGNLR GYGMASPLA APSLTPPQLATPN LQQFFPQ ATRQSLGPP PVGVPM  
 NPSQ FNLSGRNPQK QARTSSSTPNRK DSSSQTM PVEDKSDPPE GSEEAAPRM DTPEDQDLPP CPEDIAKEKRTPA PEPEPCE ASEL  
 PAKRLR SSEEPTKEP PGQLQVKAQP QARMTVPKQTQP DLLPEAL EAQVLPFRQP RVLVQAQVQ SQTQPRIPST DTQVQPQLQK  
 QAQTQTSPEH LVLQQKQVQP QLQQAEPQK QVQPQVQPQAHSQGRQ VQLQQAEPKQV QPQVQPQAHS QPPRQVQLQL QKQV  
 QTQYYP QVHT QAQPSVQPEHPPAQV SVQPPEQTHE QPHTQPQVSL LAPEQTPVVV HVC GLEMPDPAVEAGGMEK TLPEPVGTTQ  
 V SMEEQNES A CGLDVGECEN RAREMPGVWGAGGSLKVITL QSSDSRAFST VPLTPVPRPS DSVSSTPAAT STPSKQALQFFCYICKA  
 SCS SQQEFQDHMS EPQHQQRLGE IQHMSQACLL SLLPVPRDVLETEDEEPPPR RWCNTCQLY YMGDLIQHRT QDHKIAKQSL RPF  
 CTVCNRYFKTPRKFVEH VKSQGHKDKA KELKSLEKEI AGQDEDFHIT VDAVGCFEGDEEEDEDE EEEIEVEEELC KQVRSRDISR E  
 EWKGETYS PNTAYGVDFLVPVMGYICRI CHKFYHSNSG AQLSHCKSLG HFENLQKYKA AKNPSPTTRPVSRRCAINAR NALTALFTS  
 S GRPPSQPNTQ DKTPSKVTAR PSQPPLPRRSTRLKT



## Figura 19

Desde los exones 2/3 (al menos dos versiones)

TGGGGCTGC GGGGCCGGCC CATCCGTGGG GGCGACTTGA GCGTTGAGGGCGCGGGGA GCGGAGCCAC  
 CATGTTCAGC CAGCAGCAGC AGCAGCTCCAGCAACAGCAG CAGCAGCTCC AGCAGTTACA GCAGCAGCAG  
 CTCAGCAGCAGCAATTGCA GCAGCAGCAG TTACTGCAGC TCCAGCAGCT GCTCCAGCAGCTCCCCACCAC  
 AGGCC

CAGCAG CTCCAGCAGT TACAGCAGCA GCAGCTCCAG CAGCAGCAATTGCAGCAGCA GCAGTTACTG CAGC  
 TCCAGC AGCTGCTCCA GCAGTCCCCACCACA

Exón 4

GGACTGGAC CAGTTTGCAA TGCCACCAGC CAGGTATGAC ACTGCCGGTCTCACCATGCC CACAGCAACA CTG

Desde el exón 6

AGGATTCITCTTCTC

Desde el exón 8 (al menos tres versiones)

CCACAGGTGC AGCCCCAGGC ACATTCACAG CCCCCAAGGC AGGTGCAGCTGCAGCTGCAG AAGCAGGTCC  
 AGACACAGAC ATATCC

CCACAGGTAC AGCCACAGGC ACATTCACAG GGCCCAAGGC AGGTGCAGCTGCAGCAGGAG GCAGAGCCGC  
 TGAAGCAGGT GCAGCCACAG GTGCAGCCCCAGGCACAITC ACAGCCCCCA AGGCAGGTGC AGCTGCAGCT  
 GCAGAGCAGGTCCAGACAC AGACATAT

CAGGTGCAGT CACAGACTCA GCCGGGATA CCATCCACAG ACACCCAGGTGCAGCCAAAG CTGCAGAAGC  
 AGGGCAAAC ACAGACCTCT CCAGAGCATTAGTGTGCA ACAGAAGCAG GTGCAGCCAC AGCTGCAGCA  
 GGAGGCAGGCCACAGAAGC AGGTGCAGCC ACAGGTACAG CCACAGGCAC ATTCACAGGGCCCAAGGCAG  
 GTGCAGCTGC AGCAGGAGGC AGAGCCGCTG AAGCAGGTGCAGCCACAGGT GCAGCCCCAG GCACATTAC  
 AGCCCCCAAG GCAGGTGCAGCTGCAGCTGC AGAAGCAGGT CCAGACACAG ACATAT

Desde el exón 14

GTTGAGGAGGA ACTCTGCAAGCAG

La secuencia siguiente se inserta en los transcritos de Ciz1 en una biblioteca de carcinoma (desde el intrón 12 de Ciz1)

GCCACCCACACCAGAGAGATGTGTTTGCCCCACGTTCCAGTGCAGGGGTGGAGCAGCCCGGCTTGTACAGATAT

# Figura 20A

Parte de los exones 2/3 ausente

```

MF SQQQQQLQQQ QQ APLPM AVSRGLPPQQ FQOPLLNLQG TNSASLLNGS MLQRALLLQOQ GL DOFAMP PATYDTAGLT MPTATLGNLR GYGMASPLA APSLTPQLATPN LQQFFPQ
ATRQSLGGPP PVGYPMNPSQ FNLSGRNPQK QARTSSSTPNRK DSSSQTM PVEDKSDPPE GSEEAEPFRM DTPEDQDLPP CPEDIAKEKRTPA PEPECE ASELPAKRLR SSEEPTKEP PGQL
QVKAQP QARMTVPKQTPT DLLPEAL EAQVLPFRFP RVLQVAQVQ SQTQPRIPST DTQVQPKLQKQAOQTSPH LVLQKQVQP QLQEAEPQK QVQVQVQAHSQGPRQ VOLQQEA
EPLKQV QPVQVQVQAHS QPRRQVQLQ QKQVQTQTP QVHT QAQSVQVQEHPPAQV SVQPEQTHE QPHTQVQVSL LAPEQTPVV HVC GLEMPDVAEAGGMEK TLPVPVGTQV SME
EIQNESA CGLDVGECEN RAREMPVWVWAGGSLKVTIL QSSDSRAFST VPLTPVPRPS DSVSSTPAAT STPSKALQFFCYICKASCSS SQEFQDHMS EPQHQRIGE IQHMSQALL SLLPVPDR
VLETEDEEPPR RWCNTCQLYY MGDLIQHRT QDHKIAKQSL RPFCTVCNRYFKTRKRVFVH VKSQGHKDKA KELKLEKEI AGQDEDHFT VDAVGCPEGDEEEDEDE EIEIEVEELC K
QVRSRDISR EEWKGSSETYS PNTAYGVDFLVPVMGYICRI CHKFYHSNSG AQLSHCKSLG HFENLQKYA AKNPSPTTRPVSRRCAINAR NALTALFTSS GRPPSQNTQ DKTPSKVSTAR PSQPP
LPRRSTRLKT

```

Exón 4 ausente

```

MF SQQQQQLQQQ QQQQLQQQ QLLQLQQLLQQSPPQ APLPM AVSRGLPPQQ FQOPLLNLQG TNSASLLNGS MLQRALLLQOQNLG GYGMASPLA APSLTPQLA
TPN LQQFFPQ ATRQSLGGPP PVGYPMNPSQ FNLSGRNPQK QARTSSSTPNRK DSSSQTM PVEDKSDPPE GSEEAEPFRM DTPEDQDLPP CPEDIAKEKRTPA PEPECE ASELPAKRLR SSE
EPTKEP PGQLQVKAQP QARMTVPKQTPT DLLPEAL EAQVLPFRFP RVLQVAQVQ SQTQPRIPST DTQVQPKLQKQAOQTSPH LVLQKQVQP QLQEAEPQK QVQVQVQAHSQ
GPRQ VOLQQEAEPKQV QPVQVQVQAHS QPRRQVQLQ QKQVQTQTP QVHT QAQSVQVQEHPPAQV SVQPEQTHE QPHTQVQVSL LAPEQTPVV HVC GLEMPDVAEAGGMEK T
LPEVGTQV SMEIQNESA CGLDVGECEN RAREMPVWVWAGGSLKVTIL QSSDSRAFST VPLTPVPRPS DSVSSTPAAT STPSKALQFFCYICKASCSS SQEFQDHMS EPQHQRIGE IQHM
SQACLL SLLPVPDRVLETEDEEPPR RWCNTCQLYY MGDLIQHRT QDHKIAKQSL RPFCTVCNRYFKTRKRVFVH VKSQGHKDKA KELKLEKEI AGQDEDHFT VDAVGCPEGDEEEDE
DEDE EIEIEVEELC KQVRSRDISR EEWKGSSETYS PNTAYGVDFLVPVMGYICRI CHKFYHSNSG AQLSHCKSLG HFENLQKYA AKNPSPTTRPVSRRCAINAR NALTALFTSS GRPPSQNTQ
DKTPSKVSTAR PSQPP LPRRSTRLKT

```

Parte del exón 6 ausente

```

MF SQQQQQLQQQ QQQQLQQQ QLLQLQQLLQQSPPQ APLPM AVSRGLPPQQ FQOPLLNLQG TNSASLLNGS MLQRALLLQOQ GL DOFAMP PATYDTAGLT MPTAT
LGNLR GYGMASPLA APSLTPQLATPN LQQFFPQ ATRQSLGGPP PVGYPMNPSQ FNLSGRNPQK QARTSSSTPNRK DSSSQTM PVEDKSDPPE GSEEAEPFRM DTPEDQDLPP CPEDIAKEKRTPA
PEPECE ASELPAKRLR SSEEPTKEP PGQLQVKAQP QARMTVPKQTPT DLLPEAL EAQVLPFRFP RVLQVAQVQ SQTQPRIPST DTQVQPKLQKQAOQTSPH LVLQKQVQP QLQ
EAEPQK QVQVQVQAHSQGPRQ VOLQQEAEPKQV QPVQVQVQAHS QPRRQVQLQ QKQVQTQTP QVHT QAQSVQVQEHPPAQV SVQPEQTHE QPHTQVQVSL LAPEQTPVV HVC
GLEMPDVAEAGGMEK TLPVPVGTQV SMEIQNESA CGLDVGECEN RAREMPVWVWAGGSLKVTIL QSSDSRAFST VPLTPVPRPS DSVSSTPAAT STPSKALQFFCYICKASCSS SQEF
QDHMS EPQHQRIGE IQHMSQALL SLLPVPDRVLETEDEEPPR RWCNTCQLYY MGDLIQHRT QDHKIAKQSL RPFCTVCNRYFKTRKRVFVH VKSQGHKDKA KELKLEKEI AGQDE
DHFT VDAVGCPEGDEEEDEDE EIEIEVEELC KQVRSRDISR EEWKGSSETYS PNTAYGVDFLVPVMGYICRI CHKFYHSNSG AQLSHCKSLG HFENLQKYA AKNPSPTTRPVSRRCAIN
AR NALTALFTSS GRPPSQNTQ DKTPSKVSTAR PSQPP LPRRSTRLKT

```

Variante 1 menos el exón 8

```

MF SQQQQQLQQQ QQQQLQQQ QLLQLQQLLQQSPPQ APLPM AVSRGLPPQQ FQOPLLNLQG TNSASLLNGS MLQRALLLQOQ GL DOFAMP PATYDTAGLT MPT
ATLGNLR GYGMASPLA APSLTPQLATPN LQQFFPQ ATRQSLGGPP PVGYPMNPSQ FNLSGRNPQK QARTSSSTPNRK DSSSQTM PVEDKSDPPE GSEEAEPFRM DTPEDQDLPP CPE
DIAKEKRTPA PEPECE ASELPAKRLR SSEEPTKEP PGQLQVKAQP QARMTVPKQTPT DLLPEAL EAQVLPFRFP RVLQVAQVQ SQTQPRIPST DTQVQPKLQKQAOQTSPH LVL
QQKQVQP QLQQEAEPQK QVQVQVQAHSQGPRQ VOLQQEAEPKQV QPVQVQVQAHS QPRRQVQLQ QKQVQTQTP QVHT QAQSVQVQEHPPAQV SVQPEQTHE QPHTQVQVSL LAPEQTPVV HVC GLEMPDVAEAGG
MEK TLPVPVGTQV SMEIQNESA CGLDVGECEN RAREMPVWVWAGGSLKVTIL QSSDSRAFST VPLTPVPRPS DSVSSTPAAT STPSKALQFFCYICKASCSS SQEFQDHMS EPQHQR
IGE IQHMSQALL SLLPVPDRVLETEDEEPPR RWCNTCQLYY MGDLIQHRT QDHKIAKQSL RPFCTVCNRYFKTRKRVFVH VKSQGHKDKA KELKLEKEI AGQDEDHFT VDAVGC
PEGDEEEDEDE EIEIEVEELC KQVRSRDISR EEWKGSSETYS PNTAYGVDFLVPVMGYICRI CHKFYHSNSG AQLSHCKSLG HFENLQKYA AKNPSPTTRPVSRRCAINAR NALTAL
FTSS GRPPSQNTQ DKTPSKVSTAR PSQPP LPRRSTRLKT

```

**Figura 20B**

Variante 2 menos el exón 8

MF SQQQQLQQQ QQQQLQQQ QLLQLQQLLQQSPPO APLPM AVSRGLPPQQ PQPPLNLQOQ TNSASLLNGS MLQRALLLQOQO GL DQFAMP PATYDTAGLT  
MPTATLGNLR GYGMA SPGLA APSLTPPOLATPN LQQFFPO ATRQSLGPP PVGVPMNPSQ FNLSGRNPQK QARTSSSTPNRK DSSSQTM PVEDKSDPPE GSEEAAPRM DTPED  
QDLPP CPEDIAKEKRTPA PEPECE ASELPAKRLR SSEEPEKEP PGQLQVKAQP QARMTVPKQTQTP DLLPEAL EAQVLPFRFQ RVLQVQAQVQ SQTPRIPST DTQVQPKLQK  
QAQTQTSPEH LVLQKQVQP QLQQAEPQK QVQ P QVHT QAQSVQPEHPAQV SVQPPEQTHE QPHTQPVSL LAPEQTPV VV HVC GLEMPDVAEAGGMEK TLPEPVGTO  
V SMEEIQNESA CGLDVGECEN RAREMPGVWAGGSLKVTIL QSSDSRAFAST VPLTPVPRPS DSVSSTPAAT SPSKQALQFFCYCKASC SQQEFQDHMS EPQHQRRLGE IQHMSQ  
ACLL SLLPVRDVLTEDEEPPRR RWCNTCQLYY MGDLIQHRT QDHKIAKQSL RPFCTVCNRYFKTRKFVEH VKSQGHDKA KELKLEKEI AGQDEDFHIT VDAVGCFEDEE  
EEDEDEE EIEIEVEELC KQVRSRDIR EEWKGETYS PNTAYGVDFLVPVMGYICRI CHKFYHNSG AQLSHCKSLG HFENLQYKA AKNPSPTTRPVRRRCAINAR NALTALFTSS  
GRPPSQPNTQ DKTPSKVTAR PSQPPLPRRSTRLKT

Variante 3 menos el exón 8

MF SQQQQLQQQ QQQQLQQQ QLLQLQQLLQQSPPO APLPM AVSRGLPPQQ PQPPLNLQOQ TNSASLLNGS MLQRALLLQOQO GL DQFAMP PATYDTAGLT  
MPTATLGNLR GYGMA SPGLA APSLTPPOLATPN LQQFFPO ATRQSLGPP PVGVPMNPSQ FNLSGRNPQK QARTSSSTPNRK DSSSQTM PVEDKSDPPE GSEEAAPRM DTPED  
QDLPP CPEDIAKEKRTPA PEPECE ASELPAKRLR SSEEPEKEP PGQLQVKAQP QARMTVPKQTQTP DLLPEAL EAQVLPFRFQ RVLQVQAQVQ SQTPRIPST DTQVQPKLQK  
QPPEQTHE QPHTQPVSL LAPEQTPV VV HVC GLEMPDVAEAGGMEK TLPEPVGTO V SMEEIQNESA CGLDVGECEN RAREMPGVWAGGSLKVTIL QSSDSRAFAST VPLTPV  
RPS DSVSSTPAAT STPSKQALQFFCYCKASC SQQEFQDHMS EPQHQRRLGE IQHMSQACLL SLLPVRDVLTEDEEPPRR RWCNTCQLYY MGDLIQHRT QDHKIAKQSL RPFCTV  
CNRIFYKTRKFVEH VKSQGHDKA KELKLEKEI AGQDEDFHIT VDAVGCFEDEEEDDEE EIEIEVEELC KQVRSRDIR EEWKGETYS PNTAYGVDFLVPVMGYICRI  
CHKFYHNSG AQLSHCKSLG HFENLQYKA AKNPSPTTRPVRRRCAINAR NALTALFTSS GRPPSQPNTQ DKTPSKVTAR PSQPPLPRRSTRLKT

Variante menos el exón 14

MF SQQQQLQQQ QQQQLQQQ QLLQLQQLLQQSPPO APLPM AVSRGLPPQQ PQPPLNLQOQ TNSASLLNGS MLQRALLLQOQO GL DQFAMP PATYDTAGLT  
MPTATLGNLR GYGMA SPGLA APSLTPPOLATPN LQQFFPO ATRQSLGPP PVGVPMNPSQ FNLSGRNPQK QARTSSSTPNRK DSSSQTM PVEDKSDPPE GSEEAAPRM DTPEDQ  
DLPP CPEDIAKEKRTPA PEPECE ASELPAKRLR SSEEPEKEP PGQLQVKAQP QARMTVPKQTQTP DLLPEAL EAQVLPFRFQ RVLQVQAQVQ SQTPRIPST DTQVQPKLQK  
QAQTQTSPEH LVLQKQVQP QLQQAEPQK QVQVQVQAHSQPRQ VOLQQAEPKQV QVQVQVQAHS QPRQVQLQ QKQVQTQTYV QVHT QAQSVQPEHPAQV SVQ  
PPEQTHE QPHTQPVSL LAPEQTPV VV HVC GLEMPDVAEAGGMEK TLPEPVGTO V SMEEIQNESA CGLDVGECEN RAREMPGVWAGGSLKVTIL QSSDSRAFAST VPLTPVPR  
S DSVSSTPAAT STPSKQALQFFCYCKASC SQQEFQDHMS EPQHQRRLGE IQHMSQACLL SLLPVRDVLTEDEEPPRR RWCNTCQLYY MGDLIQHRT QDHKIAKQSL RPFCTV  
CNRIFYKTRKFVEH VKSQGHDKA KELKLEKEI AGQDEDFHIT VDAVGCFEDEEEDDEE EIEIEVEELC KQVRSRDIR EEWKGETYS PNTAYGVDFLVPVMGYICRI CHKFYHNSG  
AQLSHCKSLG HFENLQYKA AKNPSPTTRPVRRRCAINAR NALTALFTSS GRPPSQPNTQ DKTPSKVTAR PSQPPLPRRSTRLKT

También se deben proteger los transcritos que carecen de combinaciones de los exones variables. Por ejemplo:

Variante menos el exón 4 y parte del exón 6

MF SQQQQLQQQ QQQQLQQQ QLLQLQQLLQQSPPO APLPM AVSRGLPPQQ PQPPLNLQOQ TNSASLLNGS MLQRALLLQOQOGLNLR GYGMA SPGLA APS  
LTPPOLATPN LQQFFPO ATRQSLGPP PVGVPMNPSQ FNLSGRNPQK QARTSSSTPNRK DSSSQTM PVEDKSDPPE GSEEAAPRM DTPEDQDLPP CPEDIAKEKRTPA PEPECE  
ASELPAKRLR SSEEPEKEP PGQLQVKAQP QARMTVPKQTQTP DLLPEAL EAQVLPFRFQ RVLQVQAQVQ SQTPRIPST DTQVQPKLQQAQQTQTSPEH LVLQKQVQP QLQK  
EAEPQ QVQVQVQAHSQPRQ VOLQQAEPKQV QVQVQVQAHS QPRQVQLQ QKQVQTQTYV QVHT QAQSVQPEHPAQV SVQPPEQTHE QPHTQPVSL LAPEQ  
PVV HVC GLEMPDVAEAGGMEK TLPEPVGTO V SMEEIQNESA CGLDVGECEN RAREMPGVWAGGSLKVTIL QSSDSRAFAST VPLTPVPRPS DSVSSTPAAT STPSKQALQ  
FCYICKASC SQQEFQDHMS EPQHQRRLGE IQHMSQACLL SLLPVRDVLTEDEEPPRR RWCNTCQLYY MGDLIQHRT QDHKIAKQSL RPFCTVCNRYFKTRKFVEH VKSQ  
HKDKA KELKLEKEI AGQDEDFHIT VDAVGCFEDEEEDDEE EIEIEVEELC KQVRSRDIR EEWKGETYS PNTAYGVDFLVPVMGYICRI CHKFYHNSG AQLSHCKSLG  
HFENLQYKA AKNPSPTTRPVRRRCAINAR NALTALFTSS GRPPSQPNTQ DKTPSKVTAR PSQPPLPRRSTRLKT

Figura 21A

Parte de los exones 2/3 ausente

TGGGGGCTGC GGGGCCGGCC CATCCGTGGG GGCAGCTTGA GCGTTGAGGG  
CGCGCGGGGA GCGGAGCCAC CATGTTTCAGC CAGCAGCAGC AGCAGCTCCA  
GCAACAGCAG GGCCCGTT GCCCATGGCT GTCAGCCGGG GGCTCCCCC  
GCAGCAGCCA CAGCAGCCGC TTCTGAATCT CCAGGGCACC AACTCAGCCT  
CCCTCCTCAA CGGCTCCATG CTGCAGAGAG CTTTGCTTTT ACAGCAGTTG  
CAAGGACTGG ACCAGTTTGC AATGCCACCA GCCACGTATG AACTGCCGG  
TCTCACCATG CCCACAGCAA CACTGGGTAA CCTCCGAGGC TATGGCATGG  
CATCCOAGG CCTCGCAGCC CCCAGCCTCA CACCCCACT ACTGCCCCT  
CCAAATTTGC AACAGTTCTT TCCCCAGGCC ACTCGCCAGT CCTTGCTGGG  
ACCTCCTCCT GTTGGGGTCC CCATGAACCC TTCCAGTTT AACCTTTCAG  
GACGGAACCC CCAGAAACAG GCCCGGACCT CCTCCTCTAC CACCCCAAT  
CGAAAGGATT CTCTCTCA GACAATGCCT GTGGAAGACA AGTCAGACCC  
CCCAGAGGGG TCTGAGGAAG CCGCAGAGCC CCGGATGGAC ACACCAGAAG  
ACCAAGATT ACCGCCCTGC CCAGAGGACA TCGCCAAAGGA AAAACGCACT  
CCAGCACCTG AGCCTGAGCC TTGTGAGGCG TCCGAGCTGC CAGCAAAGAG  
ATTGAGGAGC TCAGAAGAGC CCACAGAGAA GGAACCTCCA GGGCAGTTAC  
AGGTGAAGGC CCAGCCGAG GCCCGGATGA CAGTACCGAA ACAGACACAG  
ACACCAGACC TGCTGCCTGA GGCCCTGAA GCCCAAGTGC TGCCACGATT  
CCAGCCACGG GTCTGCAGG TCCAGGCCA GGTGCAGTCA CAGACTCAGC  
CGCGGATACC ATCCACAGAC ACCCAGGTGC AGCCAAAGCT GCAGAAGCAG  
GCCAAACAC AGACTCTCC AGAGCACTTA GTGCTGCAAC AGAAGCAGGT  
GCAGCCACAG CTGCAGCAGG AGGCAGAGCC ACAGAAGCAG GTGCAGCCAC  
AGGTACAGCC ACAAGGCACAT TCACAGGGCC CAAGGCAGGT GCAGCTGCAG  
CAGGAGGCAG AGCCGCTGAA GCAGGTGCAG CCACAGGTGC AGCCCCAGGC  
ACATTCACAG CCCCCAAGGC AGGTGCAGCT GCAGCTGCAG AAGCAGGTCC  
AGACACAGAC ATATCCACAG GTCCACACAC AGGCACAGCC AAGCCTCAG  
CCACAGGAGC ATCCTCCAGC GCAGGTGTCA GTACAGCCAC CAGAGCAGAC  
CCATGAGCAG CCTCACACCC AGCCGCAGGT GTCGTTGCTG GCTCCAGAGC  
AAACACCAGT TGTGGTTCAT GTCTGCGGGC TGGAGATGCC ACCTGATGCA  
GTAGAAGCTG GTGGAGGCAT GGAAGAGACC TTGCCAGAGC CTGTGGGCAC  
CAAAGTCAGC ATGGAAGAGA TTCAGAATGA GTCGGCCTGT GGCCTAGATG  
TGGGAGAATG TGAAAACAGA GCGAGAGAGA TGCCAGGGGT ATGGGGCGCC  
GGGGGCTCCC TGAAGGTAC CATTCTGCAG AGCAGTGACA GCCGGCCCTT  
TAGCACTGTA CCCCTGACAC CTGTCCCCCG CCCCAGTGAC TCCGTCTCCT  
CCACCCCTGC GGCTACCAGC ACTCCCTCTA AGCAGGCCCT CCAGTTCTTC  
TGCTACATCT GCAAGGCCAG CTGCTCCAGC CAGCAGGAGT TCCAGGACCA  
CATGTCCGAG CCTCAGCACC AGCAGCGGCT AGGGGAGATC CAGCACATGA  
GCCAAGCCTG CCTCCTGTCC CTGCTGCCCG TGCCCCGGGA CGTCTGGAG  
ACAGAGGATG AGGAGCCTCC ACCAAGGGCG TGGTGCAACA CCTGCCAGCT  
CTACTACATG GGGGACCTGA TCCAAACCCG CAGGACACAG GACCACAAGA  
TTGCCAAACA ATCCTTGCGA CCCTTCTGCA CCGTTTGCAA CCGTACTTC  
AAAACCCCTC GCAAGTTTGT GGAGCACGTG AAGTCCCAGG GGCATAAGGA  
CAAAGCCAAG GAGCTGAAGT CGCTTGAGAA AGAAATTGCT GGCCAAGATG  
AGGACCACTT CATTACAGTG GACGCTGTGG GTTGCTTCGA GGGTGATGAA  
GAAGAGGAAG AGGATGATGA GGATGAAGAA GAGATCGAGG TTGAGGAGGA  
ACTCTGCAAG CAGGTGAGGT CCAGAGATAT ATCCAGAGAG GAGTGGAAAG  
GCTCGGAGAC CTACAGCCCC AATACTGCAT ATGGTGTGGA CTCTCTGGTG  
CCCGTGATGG GCTATATCTG CCGCATCTGC CACAAGTTCT ATCAAGCAA  
CTCAGGGGCA CAGCTCTCCC ACTGCAAGTC CCTGGGCCAC TTTGAGAACC  
TGCAGAAATA CAAGGCGGCC AAGAACCCCA GCCCACCAC CCGACCTGTG  
AGCCGCGGT GCGCAATCAA CGCCCGGAA GCTTTGACAG CCCTGTTTAC  
CTCCAGCGC CGCCACCCT CCCAGCCCA CACCCAGGAC AAAACACCCA  
GCAAGGTGAC GGCTCGACCC TCCCAGCCCC CACTACCTCG GCGTCAACC  
CGCTCAAAA CCTGATAGAG GGACCTCCCT GTCCCTGGCC TGCTGGGT  
CAGATCTGCT AATGCTTTTT AGGAGTCTGC CTGGAACCTT TGACATGGT  
CATGTTTTTA CTCAAAATCC AATAAAACA GGTAGTTTGG CTGTGCAAAA  
AAAAAAAAAA AAAAAAAAAA AA

Figura 21B

Exón 4 ausente

TGGGGGCTGC GGGGCCGGCC CATCCGTGGG GGCGACTTGA GCGTTGAGGG  
CGCGCGGGGA GGGGAGCCAC CATGTTTCAGC CAGCAGCAGC AGCAGCTCCA  
GCAACAGCAG CAGCAGCTCC AGCAGTTACA GCAGCAGCAG CTCCAGCAG  
AGCAATTGCA GCAGCAGCAG TTA CTGCAGC TCCAGCAGCT GCTCCAGCAG  
TCCCCACCAC AGGCCCCGTT GCCATGGCT GTCAGCCGGG GGCTCCCCC  
GCAGCAGCCA CAGCAGCCG TTCTGAATCT CCAGGGACC AACTCAGCCT  
CCCTCTCAA CGGCTCCATG CTGCAGAGAG CTTTGCTTTT ACAGCAGTTG  
CAAGGTAACC TCCGAGGCTA TGGCATGGCA TCCCCAGGCC TCGCAGCCCC  
CAGCCTCACA CCCCCACAAC TGGCCACTCC AAATTTGCAA CAGTTCTTTT  
CCCAGGCCAC TCGCCAGTCC TTGCTGGGAC CTCCTCTGT TGGGGTCCCC  
ATGAACCCTT CCCAGTTCAA CCTTTCAGGA CGGAACCCCC AGAAACAGGC  
CCGGACCTCC TCCTTACCA CCCCCAATCG AAAGGATTCT TCTTCTCAGA  
CAATGCCTGT GGAAGACAAG TCAGACCCCC CAGAGGGGTC TGAGGAAGCC  
GCAGAGCCCC GGATGGACAC ACCAGAAGAC CAAGATTAC CGCCCTGCC  
AGAGGACATC GCCAAGGAAA AACGCACTCC AGCACCTGAG CCTGAGCCTT  
GTGAGGGCTC CGAGCTGCCA GCAAAGAGAT TGAGGAGCTC AGAAGAGCCC  
ACAGAGAAGG AACCTCCAGG GCAGTTACAG GTGAAGGCC AGCCGACGG  
CCGGATGACA GTACCGAAAC AGACACAGAC ACCAGACCTG CTGCTGAGG  
CCCTGGAAGC CCAAGTGCTG CCACGATTCC AGCCACGGGT CCTGCAGGT  
CAGGCCAGG TGCAGTACA GACTCAGCCG CGGATACCA CCACAGACAC  
CCAGGTGACG CCAAAGCTGC AGAAGCAGGC GCAAACACAG ACCTCTCCAG  
AGCACTTAGT GCTGCAACAG AAGCAGGTGC AGCCACAGCT GCAGCAGGAG  
GCAGAGCCAC AGAAGCAGGT GCAGCCACAG GTACAGCCAC AGGCACATTC  
ACAGGGCCCA AGGCAGGTGC AGCTGCAGCA GGAGGCAGAG CCGCTGAAGC  
AGGTGCAGCC ACAGGTGACG CCCCAGGCAC ATTCACAGCC CCAAAGGCA  
GTGCAGCTGC AGCTGCAGAA GCAGGTCCAG ACACAGACAT ATCCACAGGT  
CCACACACAG GCACAGCCAA GCGTCCAGCC ACAGGAGCAT CCTCCAGGC  
AGGTGTCACT ACAGCCACCA GAGCAGACCC ATGAGCAGCC TCACACCCAG  
CCGACAGGTG CGTTGCTGGC TCCAGAGCAA ACACCAAGTTG TGGTTTATGT  
CTGCGGGCTG GAGATGCCAC CTGATGCAGT AGAAGCTGGT GGAGGCATGG  
AAAAGACCTT GCCAGAGCCT GTGGGCACCC AAGTCAGCAT GGAAGAGATT  
GAGAATGAGT CGGCCTGTGG CCTAGATGTG GGAGAATGTG AAAACAGAGC  
GAGAGAGATG CCAGGGGTAT GGGGCGCCGG GGGCTCCCTG AAGGTACCA  
TTCTGCAGAG CAGTGACAGC CCGGCCTTTA GCACTGTACC CCTGCACCT  
GTCCCCCGCC CCAAGTACTC CGTCTCTCC ACCCCTGGG CTACCAGCAC  
TCCCTCTAAG CAGGCCCTCC AGTTCTTCTG CTACATCTGC AAGGCCAGCT  
GCTCCAGCCA GCAGGAGTTC CAGGACCACA TGTCGGAGCC TCAGCACCAG  
CAGCGGCTAG GGGAGATCCA GCACATGAGC CAAGCCTGCC TCCTGTCCCT  
GCTGCCCGTG CCCCAGGACG TCCTGGAGAC AGAGGATGAG GAGCCTCCAC  
CAAGGCGCTG GTGCAACACC TGCCAGCTCT ACTACATGGG GGACCTGATC  
CAACACCGCA GGACACAGGA CCACAAGATT GCCAAACAAT CCTTGGCAGC  
CTTCTGCACC GTTTGCAACC GCTACTTCAA AACCCCTCGC AAGTTTGTGG  
AGCACGTGAA GTCCCAGGGG CATAAGGACA AAGCCAAGGA GCTGAAGTCC  
CTTGAGAAAAG AAATTGCTGG CCAAGATGAG GACCACTTCA TTACAGTGA  
CGCTGTGGGT TGCTTCGAGG GTGATGAAGA AGAGGAAGAG GATGATGAGG  
ATGAAGAAGA GATCGAGGTT GAGGAGGAAC TCTGCAAGCA GGTGAGGTCC  
AGAGATATAT CCAGAGAGGA GTGGAAGGGC TCGGAGACCT ACAGCCCCAA  
TACTGCATAT GGTGTGGACT TCCTGGTGCC CGTGATGGGC TATATCTGCC  
GCATCTGCCA CAAGTTCTAT CACAGCAACT CAGGGGCACA GCTCTCCCAC  
TGCAAGTCCC TGGGCCACTT TGAGAACCTG CAGAAATACA AGGCGGCCAA  
GAACCCAGC CCCACCACC GACCTGTGAG CCGCCGGTGC GCAATCAACG  
CCCGAACGC TTTGACAGCC CTGTTCACTT CCAGCGGCCG CCCACCCTCC  
CAGCCCAACA CCCAGGACAA AACACCCAGC AAGGTGACGG CTCGACCCTC  
CCAGCCCCCA CTACTCGGC GCTCAACCCG CCTCAAAAAC TGATAGAGGG  
ACCTCCCTGT CCTGGCCCTG CCTGGGTTCA GATCTGTAA TGCTTTTATG  
GAGTCTGCCT GGAAACTTTG ACATGGTTCA TGTTTTTACT CAAAATCCAA  
TAAACAAGG TAGTTTGGCT GTGCAAAAAA AAAAAAAAAA AAAAAAAAAA

Figura 21C

Transcrito menos el exón 6

TGGGGGCTGC GGGGCCGGCC CATCCGTGGG GCGCACTTGA GCGTTGAGGG  
 CGCGCGGGGA GGCGAGCCAC CATGTTTCAGC CAGCAGCAGC AGCAGCTCCA  
 GCAACAGCAG CAGCAGCTCC AGCAGTTACA GCAGCAGCAG CTCCAGCAGC  
 AGCAATTGCA GCAGCAGCAG TTA CTGCAGC TCCAGCAGCT GCTCCAGCAG  
 TCCCCACCAC AGGCCCGTT GCCCATGGCT GTCAGCCGGG GGCTCCCCC  
 GCAGCAGCCA CAGCAGCCGC TTCTGAATCT CCAGGGCACC AACTCAGCCT  
 CCCTCCTCAA CGGCTCCATG CTGCAGAGAG CTTTGCTTTT ACAGCAGTTG  
 CAAGGACTGG ACCAGTTTGC AATGCCACCA GCCACGTATG AACTGCCC  
 TCTACCATG CCCACAGCAA CACTGGGTAA CCTCCGAGGCT ATGGCATGG  
 CATCCCCAGG CCTCGCAGCC CCCAGCCTCA CACCCCCACA ACTGGCCACT  
 CCAAATTTGC AACAGTTCTT TCCCCAGGCC ACTCGCCAGT CTTTGCTGGG  
 ACCTCCTCCT GTTGGGGTCC CCATGAACCC TTCCCAGTTC AACCTTTCAG  
 GACGGAACCC CCAGAAACAG GCCCGGACCT CCTCCTCTAC CACCCCAAT  
 CGAAAGACAA TGCCTGTGGA AGACAAGTCA GACCCCCAG AGGGGTCTGA  
 GGAAGCCGCA GAGCCCGGA TGGACACACC AGAAGACCAA GATTTACCGC  
 CCTGCCCAGA GGACATCGCC AAGGAAAAAC GCACTCCAGC ACCTGAGCCT  
 GAGCCTTGTG AGGCGTCCGA GCTGCCAGCA AAGAGATTGA GGAGCTCAGA  
 AGAGCCACA GAGAAGGAAC CTCCAGGGCA GTTACAGGTG AAGGCCAGC  
 CGCAGGCCCG GATGACAGTA CCGAAACAGA CACAGACACC AGACCTGCTG  
 CCTGAGGCCC TGGAAGCCCA AGTGCTGCCA CGATTCCAGC CACGGGTCTC  
 GCAGGTCCAG GCCCAGGTGC AGTCACAGAC TCAGCCCGGG ATACCATCCA  
 CAGACACCCA GGTGCAGCCA AAGCTGCAGA AGCAGGCCA AACACAGACC  
 TCTCCAGAGC ACTTAGTGCT GCAACAGAAG CAGGTGCAGC CACAGCTGCA  
 GCAGGAGGCA GAGCCACAGA AGCAGGTGCA GCCACAGGTA CAGCCACAGG  
 CACATTACA GGGCCCAAGG CAGGTGCAGC TGCAGCAGGA GGCAGAGCCC  
 CTGAAGCAGG TGCAGCCACA GGTGCAGCCC CAGGCACATT CACAGCCCC  
 AAGGCAGGTG CAGCTGCAGC TGCAGAAGCA GGTCCAGACA CAGACATATC  
 CACAGGTCCA CACACAGGCA CAGCCAAGG TCCAGCCACA GGAGCTCCT  
 CCAGCGCAGG TGTCAGTACA GCCACCAGAG CAGACCCATG AGCAGCCTCA  
 CACCCAGCCG CAGGTGTCGT TGCTGGCTCC AGAGCAAACA CCAGTTGTGG  
 TTCATGTCTG CGGGCTGGAG ATGCCACCTG ATGCAGTAGA AGCTGGTGA  
 GGCATGGAAA AGACCTTGCC AGAGCCTGTG GGCACCCAAG TCAGCATGGA  
 AGAGATTGAG AATGAGTCGG CCTGTGGCCT AGATGTGGGA GAATGTGAAA  
 ACAGAGCGAG AGAGATGCCA GGGGTATGGG GCGCCGGGGG CTCCTGAA  
 GTCACCATTC TGCAGAGCAG TGACAGCCGG GCCTTTAGCA CTGTACCCCT  
 GACACCTGTC CCCCAGCCCA GTGACTCCGT CTCCTCCACC CCTGCGGTA  
 CCAGCACTCC CTCTAAGCAG GCCCTCCAGT TCTTCTGCTA CATCTGCAAG  
 GCCAGCTGCT CCAGCCAGCA GGAGTTCCAG GACCACATGT CGGAGCCTCA  
 GCACCAGCAG CGGCTAGGGG AGATCCAGCA CATGAGCCAA GCCTGCCCTC  
 TGTCCTGCT GCCCGTGCCC CGGGACGCTC TGGAGACAGA GGATGAGGAG  
 CCTCCACCAA GGCCTGGTG CAACACCTGC CAGCTCTACT ACATGGGGGA  
 CCTGATCCAA CACCGCAGGA CACAGGACCA CAAGATTGCC AAACAATCCT  
 TGCGACCCCTT CTGCACCGTT TGCAACCGCT ACTTCAAAA CCCTCGCAAG  
 TTTGTGGAGC ACGTGAAGTC CCAGGGGCAT AAGGACAAA GCAAGGAGCT  
 GAAGTCGCTT GAGAAAGAAA TTGCTGGCCA AGATGAGGAC CACTTCATTA  
 CAGTGGACGC TGTGGGTTGC TTCGAGGGTG ATGAAGAAGA GGAAGAGGAT  
 GATGAGGATG AAGAAGAGAT CGAGGTTGAG GAGGAACTCT GCAAGCAGGT  
 GAGGTCCAGA GATATATCCA GAGAGGAGTG GAAGGGCTCG GAGACCTACA  
 GCCCAATAC TGCATATGGT GTGGACTTCC TGGTGCCCGT GATGGGCTAT  
 ATCTGCCGCA TCTGCCACAA GTTCTATCAC AGCAACTCAG GGGCACAGCT  
 CTCCCACTGC AAGTCCCTGG GCCACTTTGA GAACCTGCAG AATACAAGG  
 CGGCCAAGAA CCCCAGCCCC ACCACCCGAC CTGTGAGCCG CCGGTGCGCA  
 ATCAACGCCC GGAACGCTTT GACAGCCCTG TTCACCTCCA GCGGCCGCC  
 ACCCTCCAG CCCAACACCC AGGACAAAA ACCCAGCAAG GTGACGGCTC  
 GACCTCCCA GCCCCACTA CCTCGGCGCT CAACCCGCT CAACAACTGA  
 TAGAGGGACC TCCCTGTCCC TGGCCTGCCT GGGTCCAGAT CTGCTAATGC  
 TTTTATAGGAG TCTGCCTGGA AACTTTGACA TGGTTCATGT TTTTACTCAA  
 AATCCAATAA AACAAGGTAG TTTGGCTGTG CAAAAAAAAAAAAAAAAAAAA  
 AAAAAA

Figura 21D

Variante 1 menos el exón 8

TGGGGGCTGC GGGGCCGGCC CATCCGTGGG GCGGACTTGA GCGTTGAGGG  
CGCGCGGGGA GCGGAGCCAC CATGTTTCAGC CAGCAGCAGC AGCAGCTCCA  
GCAACAGCAG CAGCAGCTCC AGCAGTTACA GCAGCAGCAG CTCCAGCAGC  
AGCAATTGCA GCAGCAGCAG TTA CTGCAGC TCCAGCAGCT GCTCCAGCAG  
TCCCCACCAC AGGCCCGTT GCCCATGGCT GTCAGCCGGG GGCTCCCCC  
GCAGCAGCCA CAGCAGCCGC TTCTGAATCT CCAGGGCACC AACTCAGCCT  
CCCTCCTCAA CGGCTCCATG CTGCAGAGAG CTTTGCTTTT ACAGCAGTTG  
CAAGGACTGG ACCAGTTTGC AATGCCACCA GCCACGTATG ACACTGCCGG  
TCTCACCATG CCCACAGCAA CACTGGGTAA CCTCCGAGGC TATGGCATGG  
CATCCCCAGG CCTCGCAGCC CCCAGCCTCA CACCCCCACA ACTGGCCACT  
CCAATTTGC AACAGTTCTT TCCCCAGGCC ACTCGCCAGT CCTTGCTGGG  
ACCTCCTCCT GTTGGGGTCC CCATGAACCC TTCCAGTTT ACAGCAGTTG  
GACGGAACCC CCAGAAACAG GCCCGGACCT CCTCCTTAC CACCCCAAT  
CGAAAGGATT CTTCTTCTCA GACAATGCCT GTGAAGACA AGTCAGACCC  
CCAGAGGGG TCTGAGGAAG CCGCAGAGCC CCGGATGGAC ACACCAGAA  
ACCAAGATTT ACCGCCCTGC CCAGAGGACA TCGCCAAGGA AAAACGCCT  
CCAGCACCTG AGCCTGAGCC TTGTGAGGCG TCCGAGCTGC CAGCAAAGAG  
ATTGAGGAGC TCAGAAGAGC CCACAGAGAA GGAACCTCCA GGCGAGTTAC  
AGGTGAAGGC CCAGCCGAG GCCCGGATGA CAGTACCGAA ACAGACACAG  
ACACCAGACC TGCTGCCTGA GGCCCTGGAA GCCCAAGTGC TGCCACGATT  
CCAGCCACGG GTCCTGCAGG TCCAGGCCA GGTGCAGTCA CAGACTCAGC  
CGCGGATACC ATCCACAGAC ACCCAGGTGC AGCCAAAGCT GCAGAAGCAG  
GGCAAACAC AGACCTCTCC AGAGCACTTA GTGCTGCAAC AGAAGCAGGT  
GCAGCCACAG CTGCAGCAGG AGGCAGAGCC ACAGAAGCAG GTGCAGCCAC  
AGGTACAGCC ACAGGCACAT TCACAGGGCC CAAGGCAGGT GCAGCTGCAG  
CAGGAGGCAG AGCCGCTGAA GCAGGTGCAG ACAG GTCCACACAC AGGCA  
CAGCC AAGCGTCCAG  
CCACAGGAGC ATCCTCCAGC GCAGGTGTCA GTACAGCCAC CAGAGCAGAC  
CCATGAGCAG CCTCACACCC AGCCGAGGT GTCGTTGCTG GCTCCAGAGC  
AAACACCACT GTGTGTTTCAT GTCTGCGGGC TGGAGATGCC ACCTGATGCA  
GTAGAAGTGT GTGGAGGCAT GGAAAAGACC TTGCCAGAGC CTGTGGGCAC  
CCAAGTCAAG ATGGAAGAGA TTCAGAATGA GTCGGCTGT GGCTAGATG  
TGGGAGAATG TGAAAACAGA GCGAGAGAGA TGCCAGGGGT ATGGGGCGCC  
GGGGCTCCC TGAAGGTCAC CATTCTGCAG AGCAGTGACA GCCCGGCTT  
TAGCACTGTA CCCTGACAC CTGTCCCCG CCCCAGTGA TCCGCTCCT  
CCACCCCTGC GGCTACCAGC ACTCCCTTA AGCAGGCCCT CCAGTTCTT  
TGCTACATCT GCAAGGCCAG CTGCTCCAGC CAGCAGGAGT TCCAGGACCA  
CATGTCGGAG CCTCAGCACC AGCAGCGGCT AGGGGAGATC CAGCACATGA  
GCCAAGCCTG CCTCCTGTCC CTGCTGCCCG TGCCCCGGA CGTCTGGAG  
ACAGAGGATG AGGAGCCTCC ACCAAGGGCG TGGTGCAACA CCTGCCAGCT  
CTACTACATG GGGGACCTGA TCCAACCCG CAGGACACAG GACCACAAGA  
TTGCCAAACA ATCCTTGCGA CCCTTCTGCA CCGTTTGCAA CCGCTACTTC  
AAAACCCCTC GCAAGTTTGT GGAGCACGTG AAGTCCAGG GGCATAAGGA  
CAAAGCCAAG GAGCTGAAGT CGCTTGAGAA AGAAATGCT GGCCAAGATG  
AGGACCACTT CATTACAGTG GACGCTGTGG GTTGCTTGA GGGTATGAA  
GAAGAGGAAG AGGATGATGA GGATGAAGAA GAGATCGAGG TTGAGGAGGA  
ACTCTGCAAG CAGGTGAGGT CCAGAGATAT ATCCAGAGAG GAGTGGAAAG  
GCTCGGAGAC CTACAGCCCC AATACTGCAT ATGGTGTGGA CTTCTGGTG  
CCCGTATGG GCTATATCTG CCGCATCTGC CACAAGTTCT ATCACAGCAA  
CTCAGGGGCA CAGCTCTCC ACTGCAAGTC CCTGGGCCAC TTTGAGAACC  
TGCAGAAATA CAAGGCGGCC AAGAACCCA GCCCAACCAC CCGACCTGTG  
AGCCGCCGGT GCGCAATCAA CGCCCGAAC GCTTTGACAG CCCTGTTTAC  
CTCCAGCGGC CGCCACCCT CCCAGCCAA CACCCAGGAC AAAACACCCA  
GCAAGGTGAC GGCTCGACCC TCCAGCCCC CACTACCTCG GCGCTAACCC  
CGCCTCAAAA CTTGATAGAG GGACCTCCT GTCCCTGGCC TGCTGGGT  
CAGATCTGCT AATGCTTTT AGGAGTCTGC CTGGAACCT TGACATGGTT  
CATGTTTTTA CTCAAAATCC AATAAAACAA GGTAGTTTGG CTGTGCAAAA  
AAAAAAAAAAAAAAAAAAAA AA

**Figura 21E**

Variante 2 menos exón 8

TGGGGGCTGC GGGGCCGGCC CATCCGTGGG GGGCACTTGA GCGTTGAGGG  
 CCGCGGGGA GCGAGCCAC CATGTTGAGC CAGCAGCAGC AGCAGCTCCA  
 GCAACAGCAG CAGCAGCTCC AGCAGTTACA GCAGCAGCAG CTCCAGCAGC  
 AGCAATTGCA GCAGCAGCAG TTAAGTGCAG TCCAGCAGCT GCTCCAGCAG  
 TCCCCACCAC AGGCCCGTT GCCATGGCT GTCAGCCGGG GGCTCCCCC  
 GCAGCAGCA CAGCAGCCGC TTCTGAATCT CCAGGGCACC AACTCAGCCT  
 CCTCCTCAA CGGCTCCATG CTGCAGAGAG CTTTGCTTTT ACAGCAGTTG  
 CAAGGACTGG ACCAGTTTGC AATGCCACCA GCCACGTATG AACTGCCGG  
 TCTCACCATG CCCACAGCAA CACTGGGTAA CCTCCGAGGC TATGGCATGG  
 CATCCCCAGG CCTCGCAGCC CCCAGCCTCA CACCCACACA ACTGGCCACT  
 CCAAATTTGC AACAGTTCTT TCCCCAGGCC ACTCGCCAGT CCTTGCTGGG  
 ACCTCCTCT GTTGGGGTCC CCATGAACCC TTCCCAAGTT AACCTTTCAG  
 GACGGAACCC CCAGAAACAG GCCCGGACCT CCTCCTCTAC CACCCCAAT  
 CGAAAGGATT CTCTTCTCA GACAATGCCT GTGGAAGACA AGTCAGACCC  
 CCCAGAGGGG TCTGAGGAAG CCGCAGAGCC CCGGATGGAC ACACAGAAG  
 ACCAAGATTT ACCGCCCTGC CCAGAGGACA TCGCAAGGA AAAACGCACT  
 CCAGCACATG AGCCTGAGCC TTGTGAGGCG TCCGAGTGC CAGCAAAGAG  
 ATTGAGGAGC TCAGAAGAGC CCACAGAGAA GGAACCTCCA GGGCAGTTAC  
 AGGTGAAGGC CCAGCCGAG GCCCGGATGA CAGTACCGAA ACAGACACAG  
 ACACCAGACC TGCTGCCTGA GGCCCTGGAA GCCCAAGTGC TGCCACGATT  
 CCAGCCACGG GTCTGCAGG TCCAGGCCA GGTGCAGTCA CAGACTCAGC  
 CGGGATACC ATCCACAGAC ACCCAGGTGC AGCCAAAGCT GCAGAAGCAG  
 GCGCAAACAC AGACCTCTCC AGAGCACTTA GTGCTGCAAC AGAAGCAGCT  
 GCAGCCACAG CTGCAGCAGG AGGCAGAGCC ACAGAAGCAG GTGCAGCCAC  
 AGGTCCACAC ACAAGCACAG CCAAGCGTCC AGCCACAGGA GCATCCTCCA  
 GCGCAGGTGT CAGTACAGCC ACCAGAGCAG ACCCATGAGC AGCCTCACAC  
 CCAGCCGAG GTGTGCTGCG TGCCCTCAGA GCAAACACCA GTTGTGGTTC  
 ATGTCTGCGG GCTGGAGATG CCACCTGATG CAGTAGAAGC TGGTGGAGC  
 ATGGAAGA CTTGCCAGA GCCTGTGGGC ACCCAAGTCA GCATGGAAGA  
 GATTCAGA AT GAGTCCGCT GTGGCTAGA TGTGGGAGAA TGTGAAACA  
 GAGCGAGAGA GATGCCAGGG GTATGGGGCG CCGGGGGTCC CTGAAGGTC  
 ACCATTCTGC AGAGCAGTGA CAGCCGGGCC TTTAGCACTG TACCCTGAC  
 ACCTGTCCC CGCCCCAGTG ACTCCGTCTC CTCCACCCCT GCGGTACCA  
 GCACTCCCTC TAAGCAGGCC CTCCAGTTCT TCTGCTAC AT CTGCAAGGCC  
 AGCTGTCCA GCCAGCAGGA GTTCCAGGAC CACATGTCGG AGCCTCAGCA  
 CCAGCAGCG CTAGGGGAGA TCCAGCACAT GAGCCAAGCC TGCCTCCTGT  
 CCTGTGTC CGTGCCCGG GACGTCTGG AGACAGAGGA TGAGGAGCCT  
 CCACCAAGGC GCTGGTGCAA CACCTGCCAG CTCTACTACA TGGGGGACCT  
 GATCCAACAC CGCAGGACAC AGGACCACAA GATTGCCAAA CAATCCTTGC  
 GACCCTTCTG CACCGTTTGC AACCGCTACT TCAAAAACCC TCGCAAGTTT  
 GTGGAGCAG TGAAGTCCA GGGCATAAG GACAAAGCCA AGGAGCTGAA  
 GTCGCTTGA AAAGAAATTG CTGGCCAAGA TGAGGACCAC TTCATTACAG  
 TGGACGCTGT GGGTTGCTTC GAGGGTATG AAGAAGAGGA AGAGGATGAT  
 GAGGATGAAG AAGAGATCGA GGTTGAGGAG GAACTCTGCA AGCAGGTGAG  
 GTCCAGAGAT ATATCCAGAG AGGAGTGAA GGGCTCGGAG ACCTACAGCC  
 CCAATACTGC ATATGGTGTG GACTTCCTGG TGCCCGTAT GGGCTATATC  
 TGCCGATCT GCCACAAGTT CTATCACAGC AACTCAGGGG CACAGCTCTC  
 CCACTGCAAG TCCCTGGGCC ACTTTGAGAA CCTGCAGAAA TACAAGCCG  
 CCAAGAACC CAGCCACC ACCCGACCTG TGAGCCCGCG GTGCGCAATC  
 AACGCCGGA ACGCTTGAC AGCCCTGTT ACCTCCAGCG GCCGCCACC  
 CTCCCAGCCC AACACCCAGG ACAAACACC CAGCAAGGTG ACGGCTCGAC  
 CCTCCAGCC CCCACTACCT CGGCGCTCAA CCGCCTCAA AACCTGATG  
 AGGGACCTCC CTGTCCCTGG CCTGCCTGG TCCAGATCTG CTAATGCTTT  
 TTAGGAGTCT GCCTGAAAC TTTGACATGG TTCATGTTTT TACTCAAAAT  
 CCAATAAAC AAGGTAGTTT GGCTGTGCAA AAAAAAAAAA AAAAAAAAAA

Figura 21F

Variante 3 menos exón 8

TGGGGGCTGC GGGGCGGGCC CATCCGTGGG GCGACTTGA GCGTTGAGGG  
CGCGCGGGGA GCGGAGCCAC CATGTTTCAGC CAGCAGCAGC AGCAGCTCCA  
GCAACAGCAG CAGCAGCTCC AGCAGTTACA GCAAGCAGCAG CTCCAGCAGC  
AGCAATTGCA GCAGCAGCAG TTA CTG CAGC TCCAGCAGCT GCTCCAGCAG  
TCCCCACCAC AGGCCCCGTT GCCCATGGCT GTCAGCCGGG GGCTCCCCC  
GCAGCAGCCA CAGCAGCCGC TTCTGAATCT CCAGGGCACC AACTCAGCCT  
CCCTCCTCAA CGGCTCCATG CTGCAGAGAG CTTTGCTTTT ACAGCAGTTG  
CAAGGACTGG ACCAGTTTGC AATGCCACCA GCCACGTATG ACACTGCCGG  
TCTCACCATG CCCACAGCAA CACTGGGTAA CCTCCGAGGC TATGGCATGG  
CATCCCCAGG CCTCGCAGCC CCCAGCCTCA CACCCCCACA ACTGGCCACT  
CCAAATTTGC AACAGTTCTT TCCCCAGGCC ACTCGCCAGT CCTTGCTGGG  
ACCTCCTCCT GTTGGGGTCC CCATGAACCC TTCCCAGTTC AACCTTTCAG  
GACGGAACCC CCAGAAAACAG GCCCGGACCT CCTCCTCTAC CACCCCAAT  
CGAAAGGATT CTCTTCTCA GACAATGCCT GTGGAAGACA AGTCAGACCC  
CCCAGAGGGG TCTGAGGAAG CCGCAGAGCC CCGGATGGAC ACACCAGAAG  
ACCAAGATT ACCGCCCTGC CCAGAGGACA TCGCCAAGGA AAAACGCACT  
CCAGCACCTG AGCCTGAGCC TTGTGAGGGC TCCGAGCTGC CAGCAAAGAG  
ATTGAGGAGC TCAGAAGAGC CCACAGAGAA GGAACCTCCA GGGCAGTTAC  
AGGTGAAGGC CCAGCCGCAG GCCCGGATGA CAGTACCGAA ACAGACACAG  
ACACCAGACC TGCTGCCTGA GGCCCTGGAA GCCCAAGTGC TGCCACGATT  
CCAGCCACGG GTCTGCAGG TCCAGGCCTC CACAGGTCCA CACACAGGCA  
CAGCCAAGCG TCCAGCCACA GGAGCATCCT CCAGCCGAGG TGTCAGTACA  
GCCACCAGAG CAGACCCATG AGCAGCCTCA CACCCAGCCG CAGGTGTCGT  
TGCTGGCTCC AGAGCAAACA CCAGTTGTGG TTCATGTCTG CGGGCTGGAG  
ATGCCACCTG ATGCAGTAGA AGCTGGTGGG GGCATGGAAA AGACCTTGCC  
AGAGCCTGTG GGCACCCAAG TCAGCATGGA AGAGATTAG AATGAGTCGG  
CCTGTGGCCT AGATGTGGGA GAATGTGAAA ACAGAGCCAG AGAGATCCA  
GGGTATGGG GCGCCGGGGG CTCCCTGAAG GTCACCATTG TGCAGAGCAG  
TGACAGCCGG GCCTTTAGCA CTGTACCCCT GACACCTGTC CCCCGCCCA  
GTGACTCCGT CTCCTCCACC CCTGCGGCTA CCAGCACTCC CTCTAAGCAG  
GCCCTCAGT TCTTCTGCTA CATCTGCAAG GCCAGCTGT CCAGCCAGCA  
GGAGTTCCAG GACCACATGT CGGAGCCTCA GCACCAGCAG CGGCTAGGGG  
AGATCCAGCA CATGAGCCAA GCCTGCCTCC TGTCCCTGCT GCCCGTGCC  
CGGGACGTCC TGGAGACAGA GGATGAGGAG CCTCCACCAA GGCGCTGGTG  
CAACACCTGC CAGCTCTACT ACATGGGGGA CCTGATCCA CACCCGAGGA  
CACAGGACCA CAAGATTGCC AAACAATCCT TGCGACCCCT CTGCACCCTT  
TGCAACCGCT ACTTCAAAAC CCCTCGCAAG TTTGTGGAGC ACGTGAAGTC  
CCAGGGCAT AAGGACAAAAG CCAAGGAGCT GAAGTCGCTT GAGAAAGAAA  
TTGCTGGCCA AGATGAGGAC CACTTCATTA CAGTGGACGC TGTGGGTTGC  
TTCGAGGGTG ATGAAGAAGA GGAAGAGGAT GATGAGGATG AAGAAGAGAT  
CGAGGTTGAG GAGGAACTCT GCAAGCAGGT GAGGTCCAGA GATATATCCA  
GAGAGGAGTG GAAGGGCTCG GAGACCTACA GCCCAATAC TGCAATATGGT  
GTGGACTTCC TGGTGCCCGT GATGGGCTAT ATCTGCCGCA TCTGCCACAA  
GTTCTATCAC AGCAACTCAG GGGCACAGCT CTCCCACTGC AAGTCCCTGG  
GCCACTTGA GAACCTGCAG AAATACAAGG CGGCCAAGAA CCCCAGCCCC  
ACCACCCGAC CTGTGAGCCG CCGGTGCGCA ATCAACGCC GGAACGCTTT  
GACAGCCCTG TTCACCTCA GCGGCCGCC ACCCTCCAG CCCAACACCC  
AGGACAAAAC ACCCAGCAA GGTGACGGCTC GACCCTCCA GCCCCACTA  
CCTCGGCGCT CAACCCGCT CAAAACCTGA TAGAGGGACC TCCCTGTCCC  
TGGCCTGCCT GGGTCCAGAT CTGCTAATGC TTTTATAGGAG TCTGCCTGGA  
AACTTTGACA TGGTTCATGT TTTTACTCAA AATCCAATAA AACAAGGTAG  
TTTGGCTGTG CAAAAAAAAA AAAAAAAAAA AAAAAA

## Figura 21G

Transcrito menos exón 14

TGGGGGCTGC GGGGCCGGCC CATCCGTGGG GGGCACTTGA GCGTTGAGGG  
 CGCGCGGGA GCGGAGCCAC CATGTTACG CAGCAGCAGC AGCAGCTCCA  
 GCAACAGCAG CAGCAGCTCC AGCAGTTACA GCAGCAGCAG CTCCAGCAG  
 AGCAATTGCA GCAGCAGCAG TTA CTGCAGC TCCAGCAGCT GCTCCAGCAG  
 TCCCCACCAC AGGCCCCGTT GCCATGGCT GTCAGCCGGG GGCTCCCCC  
 GCAGCAGCCA CAGCAGCCGC TTCTGAATCT CCAGGGCACC AACTCAGCCT  
 CCCTCCTCAA CGGCTCCATG CTGCAGAGAG CTTTGCTTTT ACAGCAGTTG  
 CAAGGACTGG ACCAGTTTGC AATGCCACCA GCCACGTATG AACTGCCCCG  
 TCTCACCATG CCCACAGCAA CACTGGGTA A CCTCCGAGGC TATGGCATGG  
 CATCCCCAGG CCTCGCAGCC CCCAGCCTCA CACCCCCACA ACTGGCCACT  
 CCAAATTTGC AACAGTTCTT TCCCCAGGCC ACTCGCCAGT CCTTGCTGGG  
 ACCTCCTCCT GTTGGGGTCC CCATGAACCC TTCCAGTTT AACCTTTCAG  
 GACGGAACCC CCAGAAACAG GCCCGGACCT CCTCCTCTAC CACCCCAAT  
 CGAAAGGATT CTTCTTCTCA GACAATGCCT GTGGAAGACA AGTCAGACCC  
 CCCAGAGGGG TCTGAGGAAG CCGCAGAGCC CCGGATGGAC ACACCAGAAG  
 ACCAAGATTT ACCGCCCTGC CCAGAGGACA TCGCAAGGA AAAACGCACT  
 CCAGCACCTG AGCCTGAGCC TTGTGAGGCG TCCGAGCTGC CAGCAAAGAG  
 ATTGAGGAGC TCAGAAGAGC CCACAGAGAA GGAACCTCCA GGGCAGTTAC  
 AGGTGAAGGC CCAGCCGAG GCGCGGATGA CAGTACCGAA ACAGACACAG  
 ACACCAGACC TGCTGCCTGA GGCCCTGGAA GCCCAAGTGC TGCCACGATT  
 CCAGCCACGG GTCCTGCAGG TCCAGGCCCA GGTGCAGTCA CAGACTCAGC  
 CGCGATACC ATCCACAGAC ACCCAGGTGC AGCCAAAGCT GCAGAAGCAG  
 GCGCAAACAC AGACCTTCC AGAGCACTTA GTGCTGCAAC AGAAGCAGGT  
 GCAGCCACAG CTGCAGCAGG AGGCAGAGCC ACAGAAGCAG GTGCAGCCAC  
 AGGTACAGCC ACAGGCACAT TCACAGGGCC CAAGGCAGGT GCAGCTGCAG  
 CAGGAGGCAG AGCCOCTGAA GCAAGTGCAG CCACAGGTGC AGCCCCAGGC  
 ACATTCACAG CCCCCAAGGC AGGTGCAGCT GCAGCTGCAG AAGCAAGTCC  
 AGACACAGAC ATATCCACAG GTCCACACAC AGGCACAGCC AAGCGTCCAG  
 CCACAGGAGC ATCCTCCAGC GCAGGTGTCA GTACAGCCAC CAGAGCAGAC  
 CCATGAGCAG CCTCACACCC AGCCGAGGT GTCGTTGCTG GCTCCAGAGC  
 AAACACCACT TGTGGTTCAT GTCTGCGGC TGGAGATGCC ACCTGATGCA  
 GTAGAAGCTG GTGGAGGCAT GGAAAAGACC TTGCCAGAGC CTGTGGGCAC  
 CCAAGTCAGC ATGGAAGAGA TTCAGAATGA GTCCGCTGT GGCCTAGATG  
 TGGGAGAATG TGAAAACAGA GCGAGAGAGA TGCCAGGGGT ATGGGGCGCC  
 GGGGGCTCCC TGAAGGTAC CATTCTGCAG AGCAGTGACA GCCGGGCTT  
 TAGCACTGTA CCCTGACAC CTGTCCCCC CCCCAGTGAC TCCGTCTCCT  
 CCACCCCTGC GGCTACCAGC ACTCCCTCTA AGCAGGGCCT CCAGTTCTTC  
 TGCTACATCT GCAAGGCCAG CTGCTCCAGC CAGCAGGAT TCCAGGACCA  
 CATGTCCGAG CCTCAGCACC AGCAGCGGCT AGGGGAGATC CAGCATGA  
 GCCAAGCCTG CCTCCTGTCC CTGCTGCCC TGCCCCGGGA CGTCTGGAG  
 ACAGAGGATG AGGAGCCTCC ACCAAGGGCG TGGTGAACA CCTGCCAGCT  
 CTA CTACATG GGGGACCTGA TCCAACCCG CAGGACACAG GACCACAAGA  
 TTGCCAAACA ATCCTTGCGA CCCTTCTGCA CCGTTTGCAA CCGCTACTTC  
 AAAACCCCTC GCAAGTTTGT GGAGCACGTG AAGTCCCAGG GGCATAAGGA  
 CAAAGCCAAG GAGCTGAAGT CGCTTGAGAA AGAAATTGCT GGCCAAGATG  
 AGGACCACTT CATTACAGTG GACGCTGTGG GTTGCTTCGA GGGTGATGAA  
 GAAGAGGAAG AGGATGATGA GGATGAAGAA GAGATCGAGG TGAGGTCCAG  
 AGATATATCC AGAGAGGAGT GGAAGGGCTC GGAGACCTAC AGCCCCAATA  
 CTGCATATGG TGTGGACTTC CTGGTGCCC TGATGGGCTA TATCTGCCG  
 ATCTGCCACA AGTTCTATCA CAGCAACTCA GGGGCACAG TCTCCCCTG  
 CAAGTCCCTG GGCCACTTTG AGAACCTGCA GAAATACAAG GCGGCCAAGA  
 ACCCCAGCCC CACCACCCGA CCTGTGAGCC GCCGGTGGC AATCAACGCC  
 CGGAACGCTT TGACAGCCCT GTTACCTCC AGCGGCCGCC CACCCTCCA  
 GCCCAACACC CAGGACAAAA CACCAGCAA GGTGACGGCT CGACCCTCCC  
 AGCCCCACT ACCTCGGCGC TCAACCCGCC TCAAAACCTG ATAGAGGGA  
 CTCCCTGTCC CTGGCCTGCC TGGGTCCAGA TCTGCTAATG CTTTTTAGGA  
 GTCTGCCTGG AAATTTGAC ATGGTTCATG TTTTACTCA AAATCCAATA  
 AAACAAGGTA GTTTGGCTGT GCAAAAAAAAA AAAAAAAAAA AAAAAA

Figura 21H

También se deben proteger los transcritos que carecen de las combinaciones de los exones variables. Por ejemplo:

Variante menos exón 14 y parte del exón 6

TGGGGGCTGC GGGGCCGGCC CATCCGTGGG GGCGACTTGA GCGTTGAGGG  
CGCGCGGGGA GGCGAGCCAC CATGTTTACG CAGCAGCAGC AGCAGCTCCA  
GCAACAGCAG CAGCAGCTCC AGCAGTTACA GCAGCAGCAG CTCCAGCAGC  
AGCAATTGCA GCAGCAGCAG TTA CTG CAGC TCCAGCAGCT GCTCCAGCAG  
TCCCACCAC AGGCCCGTT GCCCATGGCT GTCAGCCGGG GGTC CCCCC  
GCAGCAGCCA CAGCAGCCGC TTCTGAATCT CCAGGGCACC AACTCAGCCT  
CCCTCCTCAA CGCTCCATG CTGCAGAGAG CTTTGCTTTT ACAGCAGTTG  
CAAGGACTGG ACCAGTTTGC AATGCCACCA GCCACGTATG AACTG CCGG  
TCTCACCATG CCCACAGCAA CACTGGGTAA CCTCCGAGGC TATGGCATGG  
CATCCCCAGG CCTCGCAGCC CCCAGCCTCA CACCCCCACA ACTGGCCACT  
CCAAATTTGC AACAGTTCTT TCCCCAGGCC ACTCGCCAGT CCTTGCTGGG  
ACCTCCTCCT GTTGGGGTCC CCATGAACCC TTCCCAAGTC AACCTTTAG  
GACGGAACCC CCAGAAACAG GCCCGGACCT CCTCCTCTAC CACCCCCAAT  
CGAAAGACAA TGCTGTGGA AGACAAGTCA GACCCCCCAG AGGGGTCTGA  
GGAAGCCGCA GAGCCCCGGA TGGACACACC AGAAGACCAA GATTTACCGC  
CCTGCCCAGA GGACATCGCC AAGGAAAAAC GCACTCCAGC ACCTGAGCCT  
GAGCCTTGTG AGGCGTCCGA GCTGCCAGCA AAGAGATTGA GGAGCTCAGA  
AGAGCCCA CA GAGAAGGAAC CTCCAGGGCA GTTACAGGTG AAGGCCCAGC  
CGCAGGCCG GATGACAGTA CCGAAACAGA CACAGACACC AGACTGCTG  
CCTGAGGCCG TGGAAAGCCCA AGTGCTGCCA CGATTCCAGC CACGGGTCT  
GCAGGTCCAG GCCCAGGTGC AGTACAGAC TCAGCCCGCG ATACCATCCA  
CAGACACCCA GGTGCAGCCA AAGCTGCAGA AGCAGGCCA AACACAGACC  
TCTCCAGAGC ACTTAGTGCT GCAACAGAAG CAGGTGCAGC CACAGCTGCA  
GCAGGAGGCA GAGCCACAGA AGCAGGTGCA GCCACAGGTA CAGCCACAGG  
CACATTCACA GGGCCCAAGG CAGGTGCAGC TGCAGCAGGA GGCAGAGCCG  
CTGAAGCAGG TGCAGCCACA GGTGCAGCCC CAGGCACATT CACAGCCCT  
AAGGCAGGTG CAGCTGCAGC TGCAGAAGCA GGTCCAGACA CAGACATATC  
CACAGGTCCA CACACAGGCA CAGCCAAGCG TCCAGCCACA GGAGCATCCT  
CCAGCGCAGG TGTCAGTACA GCCACCAAG AGCAGCCATG AGCAGCCTCA  
CACCCAGCCG CAGGTGTCGT TGCTGGCTCC AGAGCAAACA CCAGTTGTGG  
TTCATGTCTG CCGGCTGGAG ATGCCACCTG ATGCAGTAGA AGCTGGTGG  
GGCATGAAA AGACCTTGCC AGAGCCTGTG GGCACCCAAG TCAGCATGGA  
AGAGATTCAG AATGAGTCCG CCTGTGGCCT AGATGTGGGA GAATGTGAAA  
ACAGAGCGAG AGAGATGCCA GGGGTATGGG GCGCCGGGGG CTCCCTGAAG  
GTCACCATTC TGCAGAGCAG TGACAGCCGG GCCTTTAGCA CTGTACCCCT  
GACACCTGTC CCCCAGCCCA GTGACTCCGT CTCCTCCACC CCTGGCGGTA  
CCAGCACTCC CTCTAAGCAG GCCCTCCAGT TCTTCTGCTA CATCTGCAAG  
GCCAGCTGCT CCAGCCAGCA GGAGTTCCAG GACCACATGT CGGAGCCTCA  
GCACCAGCAG CCGCTAGGGG AGATCCAGCA CATGAGCCAA GCCTGCCTCC  
TGTCCCTGCT GCCCGTGGCC CGGGACGTCC TGGAGACAGA GGATGAGGAG  
CCTCCACCAA GCGCTGGTG CAACACCTGC CAGCTCTACT ACATGGGGGA  
CCTGATCCAA CACCGCAGGA CACAGGACCA CAAGATTGCC AAACAATCCT  
TGCGACCTT CTGCACCGTT TGCAACCGCT ACTTCAAAAAC CCCTCGCAAG  
TTTGTGGAGC ACGTGAAGTC CCAAGGGCAT AAGGACAAAAG CCAAGGAGCT  
GAAGTCGCTT GAGAAAGAAA TTGCTGGCCA AGATGAGGAC CACTTCATTA  
CAGTGGACGC TGTGGGTTGC TTCGAGGGTG ATGAAGAAGA GGAAGAGGAT  
GATGAGGATG AAGAAGAGAT CGAGGTGAGG TCCAGAGATA TATCCAGAGA  
GGAGTGGAAAG GGCTCGGAGA CCTACAGCCC CAATACTGCA TATGGTGTGG  
ACTTCTGGT GCCCGTATG GGCTATATCT GCCGCATCTG CCACAAGTTC  
TATCAGCA ACTCAGGGGC ACAGCTCTCC CACTGCAAGT CCCTGGGCCA  
CTTTGAGAAC CTGCAGAAAT ACAAGCCGGC CAAGAACCCC AGCCCCACCA  
CCCCACCTGT GAGCCGCGG TGCGAATCA ACGCCCGGAA CGCTTTGACA  
GCCCTGTTCA CTCCAGCGG CCGCCACCC TCCCAGCCCA ACACCCAAGGA  
CAAAACACCC AGCAAGGTGA CCGCTCGACC CTCCCAGCCC CCACTACCTC  
GGCGCTCAAC CCGCCTCAA ACCTGATAGA GGGACCTCCC TGTCCCTGGC  
CTGCTGGGT CCAGATCTGC TAATGCTTTT TAGGAGTCTG CCTGGAAACT  
TTGACATGGT TCATGTTTTT ACTCAAAATC CAATAAAAACA AGGTAGTTTG  
GCTGTGCAAA AAAAAAAAAA AAAAAAAAAA AAA