



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 338 564**

(51) Int. Cl.:

C12N 15/10 (2006.01)

C07H 21/02 (2006.01)

C07H 21/04 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **04805284 .9**

(96) Fecha de presentación : **22.10.2004**

(97) Número de publicación de la solicitud: **1678300**

(97) Fecha de publicación de la solicitud: **12.07.2006**

(54) Título: **Método de producción de ADNc de longitud completa mediante ligación de adaptador de extremo cohesivo.**

(30) Prioridad: **24.10.2003 FR 03 12458**

(45) Fecha de publicación de la mención BOPI:
10.05.2010

(45) Fecha de la publicación del folleto de la patente:
10.05.2010

(73) Titular/es: **CENTRE NATIONAL DE LA
RECHERCHE SCIENTIFIQUE (CNRS)
3, rue Michel Ange
75016 Paris, FR
INSTITUT NATIONAL DE LA RECHERCHE
AGRONOMIQUE**

(72) Inventor/es: **Clepet, Christian, Yves**

(74) Agente: **Curell Suñol, Marcelino**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Método de producción de ADNc de longitud completa mediante ligación de adaptador de extremo cohesivo.

5 La presente invención tiene por objeto un método de producción de ADNc de longitud completa mediante ligación de un oligonucleótido en el extremo 5' de un ARN gracias a un adaptador bicatenario de extremo cohesivo y una ligasa.

Técnica anterior

10 En biología molecular, el estudio de los ADNc de longitud completa sigue siendo un enfoque ineludible para determinar la estructura de los genes. La realización de librerías de ADNc de calidad constituye un preliminar indispensable en diferentes programas de investigación en genómica. El conocimiento de la secuencia completa de los transcritos permite, por ejemplo, elegir unas sondas de especificidad óptima para la constitución de chips de ADN, o en unos
15 proyectos de análisis funcional mediante "gene silencing". Los constructos de ADNc que contienen la totalidad del marco de lectura permiten, por ejemplo, estudiar la estructura de la proteína, su localización celular, sus eventuales parejas, etc.

20 Un ADNc se denomina de longitud completa cuando cubre la totalidad de la secuencia de un transcrito. Esta terminología se utiliza habitualmente en el caso de los ARN mensajeros (ARNm) eucariotas para los cuales el ADNc de longitud completa contendrá en particular la región directamente corriente abajo del protector en 5'.

25 El aislamiento de los ADNc de longitud completa es una problemática difícil que ha inspirado un gran número de trabajos. El método ideal de producción de ADNc de longitud completa debe ser lo más específico posible del extremo 5' protegido de los ARN mensajeros con el fin de poder contra-seleccionar los ADNc truncados. En la práctica, esta ganancia de especificidad se realiza generalmente en detrimento de la eficacia global del protocolo, lo que puede ralentizar considerablemente la clonación de los ARN mensajeros más raros o inestables, o los más difíciles de convertir en ADN complementarios.

30 En el método de producción de ADNc de longitud completa de Sekine *et al.* (patente EP 0 625 572), después del tratamiento de la preparación de ARN mediante la fosfatasa alcalina, los ARNm se desprotegen mediante la pirofosfatasa ácida del tabaco, y se ligan a un oligonucleótido mediante la T4 ARN ligasa. Los ARN obtenidos se convierten a continuación en ADNc mediante una transcriptasa inversa.

35 En este protocolo, la producción de ADNc de longitud completa está limitada a nivel de la etapa de fijación del adaptador sobre el ARN; en efecto, el experto en la materia sabe que el empalme y la ligación de dos fragmentos de ácido nucleico monocatenarios presentan una eficacia reducida. Además, este procedimiento de etiquetado plantea unos problemas de especificidad debido a que cualquier extremo 3'OH presente en el medio de reacción puede servir de sustrato de ligación en lugar del adaptador, y conducir a la formación de quimeras u otros concatémoros y llegar de
40 esta manera a la desactivación y a la pérdida de los ARNm más difíciles de clonar.

Descripción de la invención

45 De manera sorprendente e inesperada, los inventores han demostrado que el método de producción de ADNc de longitud completa de la presente invención permite resolver estos problemas mediante la utilización de un adaptador bicatenario de extremo cohesivo que crea una estructura bicatenaria a nivel de la unión con el ARN, y de una ligasa para la ligación de un oligonucleótido en el extremo 5' de los ARN. Los inventores han demostrado que dicho método permite aumentar la eficacia de la ligación. La utilización de adaptadores bicatenarios de extremo cohesivo aporta unas mejoras significativas para el posicionamiento del oligonucleótido a nivel del extremo 5' del ARN. La estructura bica-
50 tenaria generada a nivel de la unión con el ARN permite además catalizar la ligación por medio de la T4 ADN ligasa cuya constante de afinidad ($10^{-8} \text{ M} < K_m < 10^{-7} \text{ M}$ para unos polinucleótidos en el seno de estructuras bicatenarias) es significativamente más baja que para la T4 ARN ligasa ($K_m > 1 \text{ mM}$ para unos polinucleótidos).

55 Además de una ganancia de eficacia, este método presenta una especificidad aumentada con relación al protocolo de Sekine *et al.* (patente EP 0 625 572). En este protocolo, cualquier polinucleótido que presenta un extremo 3'OH constituye un sitio aceptor potencial para la T4 ARN ligasa y puede, por lo tanto, ser fijado sobre el extremo 5' fosfato de los ARNm. Además de los problemas de ruido de fondo relacionados con la presencia de los concatémoros y otras quimeras que se pueden generar de esta manera, este tipo de evento puede conducir a la desactivación y a la pérdida de los transcritos más raros. La especificidad de ligación conferida por la T4 ADN ligasa que ejerce su actividad en
60 unos "nicks" en el seno de estructuras bicatenarias, permite evitar este tipo de ruido de fondo.

Por último, se señalará que siendo la actividad de la T4 ARN ligasa más elevada en algunas secuencias que en otras, esto podría conducir a una representación desigual de los transcritos en los bancos de ADNc. Dicha desviación de secuencia no se ha descrito jamás para la T4 ADN ligasa.

65 El objeto de la presente invención es por lo tanto un método de producción de ADNc de longitud completa que comprende una etapa a) de desfosforilación de los fragmentos de ácidos nucleicos no provistos de un protector en 5', una etapa b) de supresión del extremo de la protección del extremo 5' de los ARN protegidos, una etapa c) de

ligación de oligonucleótidos en el extremo 5' de ARN obtenidos en la etapa b) y una etapa d) de síntesis de los ADNc de longitud completa, mediante retro-transcripción de los ARN obtenidos en la etapa c), mediante una transcriptasa inversa, estando dicho método caracterizado porque la etapa c) de ligación del oligonucleótido se lleva a cabo mediante una ligasa y un adaptador bicatenario de extremo cohesivo, comprendiendo dicho adaptador bicatenario de extremo cohesivo una hebra superior que contiene un sitio 3'OH aceptor de ligasa para el acoplamiento con el 5' de los ARN, y una hebra inferior que presenta en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria que puede aparearse con el extremo 5' de los ARN.

La etapa a) se lleva a cabo de manera convencional, conocida por el experto en la materia y puede, entre otros, ser llevada a cabo mediante la fosfatasa alcalina (Maniatis *et al.*, véase Techniques générales de biologie moléculaire).

La etapa b) se lleva a cabo de manera convencional, conocida por el experto en la materia y puede, entre otros, ser llevada a cabo mediante la pirofosfatasa ácida del tabaco.

La etapa d) se lleva a cabo de manera convencional, conocida por el experto en la materia y puede, entre otros, ser llevada a cabo mediante el apareamiento de un cebador en el extremo 3' de un ARN obtenido en la etapa c), la elongación 5'-3' del cebador mediante la transcriptasa inversa para formar la primera hebra de ADNc, la degradación de la hebra de ARN mediante un tratamiento alcalino, y la síntesis de la segunda hebra de ADNc mediante una ADN polimerasa dependiente de ADN. (Maniatis *et al.*, véase Techniques générales de biologie moléculaire).

Mediante la expresión "ARN protegido", se entiende designar de manera preferida según la presente invención unos ARN mensajeros.

Por el término "ligasa" se entiende designar, según la presente invención, una ADN ligasa, una ARN ligasa o una combinación de ARN ligasa y de ADN ligasa.

De manera preferida, la ligasa utilizada según la presente invención es una ADN ligasa, y más particularmente la T4 ADN ligasa.

Mediante la expresión "adaptador bicatenario de extremo cohesivo" se entiende designar, según la presente invención, una molécula de ácido nucleico bicatenario que comprende una hebra superior que contiene un sitio 3'OH aceptor de ligasa para el acoplamiento con el 5' de los ARN, y una hebra inferior que presenta en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria que puede aparearse con el extremo 5' de los ARN.

De manera opcional, según la presente invención, la hebra inferior del adaptador bicatenario de extremo cohesivo se puede bloquear mediante un grupo amina en 3'.

Mediante la expresión "adaptador bicatenario de extremo cohesivo constituido por una molécula de ácido nucleico" se entiende designar, según la presente invención, un adaptador bicatenario de extremo cohesivo constituido por ADN o por ARN o por una composición mixta de ADN/ARN.

De manera preferida, la hebra superior del adaptador bicatenario está compuesta por 24 desoxirribonucleótidos en 5' seguido de un fragmento de 8 ribonucleótidos en 3'.

De manera preferida, la hebra inferior del adaptador bicatenario comprende en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria de 6 bases. Se debe señalar que dicha extensión oligonucleotídica monocatenaria de secuencia aleatoria puede estar constituida por cualquier tipo de base que puede aparearse con el ARN, y puede comprender menos de 6 bases, o más de 6 bases, sin que se modifique por ello la presente invención.

La utilización de este tipo de adaptador bicatenario de extremo cohesivo y de la T4 ADN ligasa puede tener una aplicación cada vez que un adaptador oligonucleotídico debe ser ligado sobre unos ARN.

En particular, este método es útil para ligar un adaptador sobre el extremo 5' fosfato de un ARN de secuencia desconocida, utilizando unos adaptadores bicatenarios de extremo cohesivo del mismo tipo que los descritos anteriormente.

Si es necesario, se podrá añadir previamente un fosfato sobre el extremo 5' de los ARN, por medio de la actividad polinucleótida quinasa del bacteriófago T4, con el fin de crear un sitio donante para la T4 ADN ligasa, que puede reaccionar con el sitio aceptor 3' hidroxilo del adaptador.

Esta estrategia de etiquetado puede constituir un procedimiento alternativo que permite unas ganancias de eficacia en los experimentos de inmunoprecipitación de ribonucleoproteínas (RIP).

La presente invención tiene asimismo por objeto un método de ligación de un adaptador bicatenario de extremo cohesivo en el extremo 5' de un ARN, que comprende las etapas siguientes:

- a) Obtener un ARN monocatenario que presenta un extremo 5' fosfato

b) Preparar las dos hebras de un adaptador de extremo cohesivo que comprende una hebra superior que contiene un sitio 3'OH aceptor de ligasa para el acoplamiento con el 5' de los ARN, y una hebra inferior de extremo cohesivo que presenta en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria que puede aparearse con el extremo 5' fosfato de los ARN

c) Mezclar dicha hebra superior con dicha hebra inferior de manera que forman un adaptador bicatenario

d) Mezclar dicho adaptador con dicho ARN monocatenario para permitir que el extremo 5' de dicha hebra inferior se hibride con dicha hebra sencilla de ARN

e) Ligar dicho adaptador con dicha hebra sencilla de ARN mediante una ligasa.

Por el término "ARN" se entiende designar, según la presente invención, unos ARN protegidos o unos ARN no protegidos.

Si por el término "ARN" se entienden unos ARN protegidos, entonces la etapa a) es una supresión de la protección del extremo 5' de los ARN protegidos. Dicha supresión de la protección se puede llevar a cabo, por ejemplo, mediante la pirofosfatasa ácida del tabaco. De manera preferida, los ARN protegidos son unos ARN mensajeros.

Si por el término "ARN" se entienden unos ARN no protegidos y provistos de un extremo 5'OH, entonces la etapa a) es una fosforilación del extremo 5'OH de los ARN no protegidos. Dicha fosforilación puede llevarse a cabo, por ejemplo, mediante la T4 polinucleótido quinasa.

Si por el término "ARN" se entienden unos ARN no protegidos y provistos de un extremo 5' fosfato, entonces la etapa a) consiste en aislar dichos ARN.

De manera preferida, la ligasa utilizada en la etapa e) es una ADN ligasa, una ARN ligasa o una combinación de ARN ligasa y de ADN ligasa. De manera más preferida, la ligasa utilizada es una ADN ligasa, y más particularmente la T4 ADN ligasa.

De manera opcional, la hebra inferior del adaptador bicatenario de extremo cohesivo se puede bloquear mediante un grupo amina en 3'.

Mediante la expresión "adaptador bicatenario de extremo cohesivo" se entiende designar, según la presente invención, un adaptador bicatenario de extremo cohesivo constituido por ADN o por ARN o por una combinación mixta ADN/ARN.

De manera preferida, la hebra superior del adaptador bicatenario está compuesta por 24 desoxirribonucleótidos en 5' seguido de un fragmento de 8 ribonucleótidos en 3'.

De manera preferida, la hebra inferior del adaptador bicatenario comprende en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria de 6 bases.

Se debe señalar que dicha extensión oligonucleotídica monocatenaria de secuencia aleatoria puede estar constituida por cualquier tipo de base que puede aparearse con el ARN, y puede comprender menos de 6 bases, o más de 6 bases, sin que se modifique por ello la presente invención.

La presente invención tiene asimismo por objeto un kit para la producción de ADNc de longitud completa, caracterizado porque comprende por lo menos:

- o fosfatasa alcalina y su tampón de reacción
- o pirofosfatasa ácida del tabaco y su tampón de reacción
- o una ligasa, o un cóctel de ligasas, que comprende en particular la T4 ADN ligasa y su tampón de reacción.
- o un adaptador bicatenario de extremo cohesivo, comprendiendo dicho adaptador bicatenario de extremo cohesivo una hebra superior que contiene un sitio 3'OH aceptor de ligasa para el acoplamiento con el 5' de los ARN, y una hebra inferior que presenta en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria que puede aparearse con el extremo 5' de los ARN.
- o una transcriptasa inversa
- o una ADN polimerasa dependiente de ADN
- o unos cebadores oligonucleotídicos para la polimerización del ADN.

Leyenda de las figuras

Figura 1

5 Estrategia de síntesis de ADNc de longitud completa.

Figura 2

10 Análisis de los productos de RACE PCR sobre gel al 2% de agarosa coloreado con bromuro de etidio. 12 µl de producto PCR (1 - 6) amplificado con los cebadores P1 y rbc402.L (1 - 3); P1 (4); rbc402.L (5); rbc64.U y rbc402.L (6); sobre ADNc (3 - 6); sobre ADN genómico (1); control sin ADN (2). L: 100 bp ladder (Life Technologies).

Figura 3

15 Electroferograma de la secuencia de un clon de la RACE PCR basada en el cebador rbc402.L. La flecha hacia la derecha indica la secuencia (complementaria inversa) que corresponde al extremo 5' del transcrito ats1A, fusionada con la secuencia del adaptador P1 (flecha hacia la izquierda).

Figura 4

25 Análisis de los clones de RACE PCR. La secuencia de la región adyacente al adaptador esta presentada para 10 clones. Los fragmentos de ADNc obtenidos corresponden a 3 genes. El número de registro y la secuencia genómica que enmarca el sitio (+1) de iniciación de la transcripción descritos en Genbank están presentados para cada gen. N: número de clones. Cap site: posición relativa del extremo 5' de los fragmentos clonados con relación al +1 anotado en Genbank.

30

Ejemplo1) *Material y métodos*

35 El método de producción de ADNc según la presente invención se representa en la figura 1.

- Los oligonucleótidos

40 Las secuencias están orientadas desde 5' hacia 3'. N corresponde a cualquier tipo de base que puede aparearse con el ARN; V corresponde a uno de los 3 desoxirribonucleótidos dA, dG o dC; rC, rT, rG son unos ribonucleótidos; a indica un grupo NH₂ en el carbono 3'.

45

TABLA 1

Oligonucleótidos

50

55

60

65

P1 (SEC ID n° 1)	GGGGACAAGTTTGTACAAAAAAGCtArGrGrCrTrGrGrG
P2 (SEC ID n° 2)	NNNNNNCCCCAGCCTGCa
P3 (SEC ID n° 3)	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTTTTTTTTTT TTTTTTTTTVN
Rbc64.U (SEC ID n° 4)	AATGGCTTCCTCTATGCTCTC
Rbc402.L (SEC ID n° 5)	TGCGGAGAAGGTAGTCAACTTCCT
M13fwd (SEC ID n° 6)	TGTAAAACGACGGCCAGTGA
M13rev (SEC ID n° 7)	GGAAACAGCTATGACCATGAT
Univ221.U (SEC ID n° 8)	TGTAAAACGACGGCCAGTCTTA
Univ221.L (SEC ID n° 9)	CAGGAAACAGCTATGACCATGT
P1per (SEC ID n° 10)	GGGGACAAGTITGTACAAAAAAGCAGGCT
P3pcr (SEC ID n° 11)	GGGGACCACTTTGTACAAGAAAGCTGGGT

ES 2 338 564 T3

- Preparación del adaptador

El adaptador ha sido preparado mezclando en un microtubo 40 μ m de cada uno de los oligonucleótidos P1 y P2, en un tampón que contiene 10 mM de NaCl y 10 mM de Tris-HCl (pH 7,5). El microtubo se calentó a 70°C en un vaso de precipitado que contiene 100 ml de agua, y se dejó volver progresivamente hasta temperatura ambiente durante aproximadamente 20 min. La reacción de hibridación ha sido alicuotada y conservada a -80°C.

- Preparación de los ARN

Los ARN totales han sido purificados mediante el método con Trizol (Invitrogen) a partir de tejidos aéreos vegetativos de *Arabidopsis thaliana*, siguiendo las condiciones recomendadas por el fabricante.

- Fijación del adaptador sobre el sitio de protección de los ARN mensajeros

Desfosforilación

Se han digerido 2 μ g de ARN total por 2 unidades de fosfatasa alcalina (Roche) en 20 μ l de tampón que contenía 50 mM de Tris-HCl (pH 8,5), 5 mM de $MgCl_2$, y 40 unidades de inhibidor de ribonucleasa (ARNsin, Promega) durante 1 h a 37°C. La mezcla de reacción se calentó a 65°C durante 10 min., se completó hasta 100 μ l con H_2O , se extrajo con fenol:cloroformo y se precipitó en presencia de 300 mM de acetato de sodio, 20 μ g de glicógeno, y 2 volúmenes de etanol absoluto, durante 20 min. a -20°C. El ARN se ha recuperado mediante centrifugación a 14.000 rpm durante 20 min., se lavó mediante 200 μ l de etanol a 75% y se secó dejando el tubo abierto sobre la mesa de laboratorio durante 5 min.

Supresión de la protección de los ARN mensajeros

El ARN ha sido digerido a continuación durante 1 h a 37°C, mediante 2 unidades de pirofosfatasa ácida del tabaco (Epicentre) en 20 μ l del tampón recomendado por el fabricante, suplementado por 40 unidades de ARNsin. El volumen de reacción se completó hasta 100 μ l con H_2O , se extrajo con fenol-cloroformo y se precipitó con etanol, tal como se ha descrito anteriormente. El residuo de ARN se lavó mediante etanol a 75% y se secó durante 5 min. sobre la mesa de laboratorio.

Ligación del oligonucleótido

Los ARN liberados de la protección en 5' se recogen en 6,5 μ l de H_2O , se calientan a 65°C durante 5 min., se equilibran a 25°C y se mezclan en un volumen final de 10 μ l con 50 mM de Tris-HCl (pH 7,5), 10 mM de $MgCl_2$, 10 mM de ditioneitol (DTT), 1 mM de ATP, 25 μ g/ml de albúmina de suero de ternera (BSA), 5% de polietilenglicol 8000, 20 unidades de ARNsin, 2 μ M de adaptador bicatenario y 1.000 unidades de T4-ADN ligasa de fuerte concentración (New England Biolabs). La reacción se incubó durante 3 h a 25°C. El volumen de reacción se completó hasta 100 μ l con H_2O , se extrajo con fenol-cloroformo, y se precipitó en presencia de 300 mM de acetato de sodio, 20 μ g de glicógeno y 2 volúmenes de etanol, durante una noche a -20°C. El residuo de ARN se lavó mediante etanol a 75% y se secó durante 5 min. sobre la mesa de laboratorio.

- Transcripción inversa

El residuo de ARN se recogió en 8 μ l de H_2O , se mezcló con el cebador oligonucleotídico P3 (50 pmoles) y los 4 dNTP (10 nmoles de cada uno), en un volumen final de 12 μ l y se calentó hasta 65°C durante 5 min. La disolución se equilibró a 48°C y se completó mediante 40 unidades de ARNsin, 10 mM de DTT, 50 mM de Tris-HCl (pH 8,3), 75 mM de KCl, 3 mM de $MgCl_2$, y 200 unidades de M MLV Superscript III (Invitrogen), en un volumen final de 20 μ l. Después de 50 min. a 48°C, se detuvo la reacción calentando durante 15 min. a 70°C. Los ADN complementarios se congelaron a -20°C.

- PCR

Las PCR se efectuaron utilizando 1,5 unidades de polimerasa AmpliTaq Gold (Roche) en 15 μ l de tampón que contiene 15 mM de Tris-HCl (pH 8,0), 50 mM de KCl, 1,5 mM de $MgCl_2$, 200 μ M de cada dNTP, y 0,2 μ M de cada cebador oligonucleotídico, se han utilizado para los controles genómicos, 60 ng de ADN de *A. thaliana*. Para el cribado del banco de RACE PCR y para la producción de una matriz de secuenciación, se efectuaron las PCR sobre 1 μ l de cada clon conservado a -80°C en 4% de glicerol, con los cebadores (rbc64.U, rbc402.L) o (M13fwd, M13rev). Para el cribado y la producción de una matriz de secuenciación del banco de ADNc de longitud completa, se utilizó el par de cebadores (univ221.U, univ221.L). Salvo que se especifique lo contrario, las PCR se efectuaron en las siguientes condiciones: 8 min. a 94°C y después {94°C durante 10 s, 60°C durante 10 s, 72°C durante 1 min.} durante 30 ciclos y una etapa final de 2 min. a 72°C. Se analizaron las reacciones sobre gel de agarosa a 2%, coloreado con bromuro de etidio.

- Clonación

Para los experimentos de amplificación de los extremos de ADNc (RACE PCR), se ha amplificado 1 μ l de ADNc con el par de cebadores (P1, rbc402.L) antes de ser clonados en el vector de tipo "T/A cloning" pGEM-T (Promega).
 5 Se incubó un μ l del producto de RACE PCR con 20 ng del vector de clonación, durante 1 h a 20°C, en 5 μ l de tampón que contiene 50 mM de Tris-HCl (pH 7,5), 10 mM de MgCl₂, 10 mM de DTT, 1 mM de ATP, 25 μ g/ml de BSA y 200 unidades de T4 ADN ligasa (Biolabs).

Para construir el banco de ADNc de longitud completa, se han sometido 3 μ l de ADNc a una serie de 12 ciclos
 10 de amplificación PCR en las siguientes condiciones {94°C durante 10 s, 66°C durante 20 s, 72°C durante 5 min.}, en un volumen final de 20 μ l con los cebadores P1pcr y P2pcr. Los fragmentos PCR de tamaño superior a 1 kb han sido fraccionados sobre un gel a 1,2% de agarosa Seakem (Tebu) y han sido extraídos del gel mediante electroelución. El ADN eluido ha sido precipitado en presencia de 300 mM de acetato de sodio, 20 μ g de glicógeno y 2 volúmenes de etanol, durante una noche a -20°C. El residuo de ADN se ha lavado mediante etanol a 75%, se ha secado durante 5 min.
 15 en la mesa de laboratorio y se ha recogido en 10 μ l de tampón Tris-HCl pH 7,5 10 mM, EDTA 1 mM. Los ADNc así preparados han sido insertados en el vector pDONR221 (Invitrogen) mediante recombinación *in vitro*, en un volumen final de 20 μ l, a 24°C durante la noche, con la ayuda del kit Gateway y siguiendo las recomendaciones del fabricante (Invitrogen).

Después del desalado de las reacciones sobre minicolumnas de sefarsa 4LB (Amersham Pharmacia) se utilizó
 20 1/5 de los productos de ligación (RACE PCR) o de recombinación mediante Gateway (ADNc de longitud completa) para transformar mediante electroporación la cepa *E. coli* DH10B {F-mcrA D(mrr-hsdRMS-mcrBC) F 801acZDM15 DlacX74 deoR recA1 endA1 araD139 D(ara,leu)7697 galU galK 1 - rpsL (StrR) nupG tonA} comercializada por Invitrogen, con un electroporador de marca Life Technologies siguiendo las condiciones preconizadas por el constructor.
 25 Las células transformadas se incubaron durante 45 min. en medio SOC, antes de ser extendidas sobre las cajas de Petri que contienen un medio LB gelosado, suplementado por 50 μ g/ml de carbenicilina, 20 μ g/ml de 5-bromo-4-cloro-3-indolil- β -D-galactósido y de iso-propiltio- β -D-galactósido para la clonación en pGEM-T, o suplementado por 50 μ g/ml de kanamicina para la clonación en pDONR221.

30 - Secuenciación

Los fragmentos de ADNc han sido aislados mediante PCR con dos cebadores localizados en ambos lados del
 sitio de clonación: es decir, (M13fwd, M13rev) para los constructos en pGEM-T y (Univ221.U, Univ221.L) para los
 35 constructos en pDONR221. Las matrices se purificaron utilizando el sistema MinElute de Qiagen, y se secuenciaron en fluorescencia sobre aparatos ABI.

- Análisis de secuencia

Las secuencias han sido analizadas mediante BLAST frente a la base de Genbank, utilizando la versión de abril de
 40 2003 para los productos de RACE PCR, y la versión del 8 de agosto de 2003 para los clones del banco de ADNc de longitud completa. Para el análisis de los clones de ADNc de longitud completa, el locus genómico que corresponde a cada uno de los fragmentos de ADNc se identificó utilizando la base de datos Flagdb22, y una secuencia genómica de 3 kb anclada en posición 1000 sobre la primera base del marco de lectura anotado por la AGI, y que cubre 2.000 bases corriente arriba (es decir, hacia el promotor del gen) y se han aislado 1.000 bases corriente abajo de esta base para
 45 cada uno de los genes. Cada secuencia de 3 kb ha sido utilizada como pregunta para consultar mediante BLAST una base de datos que agrupa las secuencias de los 154 fragmentos clonados por un lado, y una base de datos constituida a partir de la totalidad de las secuencias de ADNc (EST, ARN, ADNc) presentes en Genbank bajo el campo *Arabidopsis thaliana* por otro lado.

Observación: Con el fin de poder utilizar eventualmente el sistema de clonación Gateway patentado por Invitrogen,
 50 los sitios Gateway attB1 y attB2 de recombinación *in vitro* han sido integrados en los adaptadores P1/P2 y P3. Este sistema de clonación ha sido utilizado efectivamente para la construcción del banco de ADNc de longitud completa pero no para la clonación de los productos de RACE PCR que han sido clonados mediante ligación. Se insiste en el hecho de que el método según la presente invención no depende ni del sistema de clonación Gateway ni de la presencia
 55 de las secuencias attB1 y attB2 en los adaptadores utilizados. Otros tipos de secuencia de adaptador, otro tipo de vector diferente de pDONR221 y una clonación "clásica" mediante ligación podrían haber sido utilizados para construir el banco de ADNc de longitud completa. La utilización del sistema Gateway debe ser considerada como una variante de clonación, con sus ventajas y sus inconvenientes, pudiendo "amenizar" la estrategia descrita de producción de ADNc de longitud completa.

60 2) Resultados y discusión

El nuevo método de etiquetado de los ARN y su aplicación para la producción de ADNc de longitud completa
 ha sido validado de manera convergente mediante un experimento de RACE PCR y mediante la construcción y el
 65 análisis de un banco de ADNc. La RACE PCR se ha efectuado sobre un sistema experimental ya documentado: un cebador de PCR (rbc402.L, tabla 1) se selecciona sobre una secuencia homóloga de entre cuatro genes de *Arabidopsis thaliana*; este cebador reconoce una diana idéntica al 100% sobre los genes ats1B, ats2B y ats3B, y presenta una base de diferencia en su parte 5', para el gen ast1A. El cebador está orientado hacia el sitio de iniciación de la transcripción

ES 2 338 564 T3

y se encuentra aproximadamente a 3.000 bases de éste en el ADNc. Unos ARN totales de *Arabidopsis thaliana* han sido purificados recientemente y acoplados en 5' con la etiqueta P1 por medio de la T4 ADN ligasa, según el protocolo enunciado anteriormente. Se han sintetizado unos ADNc a partir de los ARNm acoplados a la etiqueta 5' y de cebadores P3 (tabla 1), y amplificado mediante PCR con los cebadores rbc402.L y P1. El análisis sobre gel de la reacción de RACE PCR hace aparecer una banda compatible con el producto esperado de 350 pares de bases y ausente de los controles de PCR (figura 2). El producto de la RACE PCR ha sido clonado en el vector pGEM-T y se han secuenciado los constructos de 10 clones amplificables por PCR con los cebadores (rbc64.U, rbc402.L). Estos fragmentos de inserciones corresponden a la región 5' de los ADNc ats1A, ats1B y ats3B, fusionados según la configuración esperada con la secuencia de la etiqueta P1; a título de ejemplo, el electroferograma de la secuencia complementaria inversa de un clon que contiene el extremo 5' del transcrito atsA1 se presenta en la figura 3. La figura 4 resume los resultados del análisis de los clones: en particular 3 ADNc "ats1A" corresponden exactamente al sitio de iniciación de transcripción (SIT) descrito en Genbank para este gen, y un fragmento "ats1A" presenta la incorporación de 5 bases suplementarias en dirección del promotor. Los otros 6 constructos corresponden a unos fragmentos de ADNc que tienen un extremo 5' localizado a 1 base o algunas bases corriente abajo de los sitios de iniciación de transcripción descritos en Genbank respectivamente para ats1A, ats3B y ats1B.

La evaluación de la estrategia se buscó después a través de la construcción y el análisis de un banco de ADNc de longitud completa. *Arabidopsis thaliana* constituye un sistema experimental de calidad para validar un método de producción de ADNc de longitud completa, gracias a la extensión y a la calidad de los datos acumulados en este organismo. Se dispone en efecto de la secuencia genómica extensivamente documentada y de más de 220.000 secuencias de ADNc (catalogadas bajo los campos EST, ARN o ADNc), de las cuales una parte importante procede de proyectos de caracterización de ADNc de longitud completa utilizando unas técnicas concurrentes. Un banco de ADNc ha sido construido por lo tanto a partir de los ARN ligados a la etiqueta 5' y retro-transcritos con el cebador P3. Los ADNc han sido amplificados mediante PCR con los cebadores P1pcr y P3pcr (tabla 1), fraccionados sobre gel de agarosa y clonados mediante el sistema Gateway (Invitrogen) en el vector pDONR221. Un total de 154 clones recombinantes han sido secuenciados y sometidos a una serie de análisis bioinformáticos con el fin de evaluar la calidad de los ADNc obtenidos. En ausencia de una "secuencia signatura" conocida que permitiría definir si un fragmento de ADNc está completo o no, la evaluación de los ADNc se ha establecido sobre las consideraciones siguientes: un primer criterio consiste en verificar si la totalidad de la secuencia codificante (ORF) está o no contenida por el fragmento de ADNc; para eso, se hace referencia a las anotaciones de la Arabidopsis Genome Initiative (AGI) (The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796-815 (2000)). En segundo lugar, la posición del extremo 5' de cada ADNc ha sido comparada con las de la totalidad de las secuencias de ADNc relativas a los transcritos considerados y presentes en Genbank (publicación del 13 de agosto de 2003). El resultado de los análisis clon a clon se agrupa en la tabla 2; se presenta en la tabla 3 un balance.

TABLA 2

Clon	locus ID	GB 5'	ATG	nota
a192B10	AT5G44730	ausente	-143	(6)
a189D5	AT5G19550	-72	-167	
a189E2	AT5G25752	-34	-61	
a209D1	AT2G35450 y AT2G35440	-19	-8	(1)
a189E12	AT4G17050	-16	-88	
a209E4		+55	-26	
a188F11		+60	-21	
a209D10	AT2G42240 y AT2G42245	-15	-72	(2)
a209A8	AT1G06040	-14	-244	
a189F12	AT4G01310	-10	-65	
a189F9	AT1G71190	-5	-50	
a209B4	AT1G29930	-5	-71	
a189A4		+1	-66	
a189E11		+1	-66	
a189A11		+1	-66	
a188G8		+1	-66	
a192C3		+1	-66	
a192C4		+1	-66	
a209E5		+1	-66	
a189C7		+1	-66	

ES 2 338 564 T3

	Clon	locus ID	GB 5'	ATG	nota
5	a189D4		+1	-66	
	a192B7		+2	-65	
	a188B7		+5	-62	
	a063B1	AT1G09590	-3	-15	
10	a197bG2	AT1G06680	-2	-63	
	a209C5		+4	-58	
	a197bB6		+4	-58	
	a189D11		+41	-21	
	a209D2	AT1G43800	+1	-51	
15	a209A7	AT3G17410	+1	-163	
	a189C1	AT4G05320	+1	-68	
	a189B5	AT5G57850	+1	-39	
20	a197bH5	AT5G57930	+1	-86	
	a197bG1	AT1G19570	+1	-42	
	a192C11	AT3G12930	+1	-56	
	a209D12	AT3G44860	+1	-56	
25	a192A11		+1	-56	
	a063D1		+1	-62	
	a192C6		+2	-62	
	a189E9		+2	-62	
30	a209E2	AT5G22340	+1	-36	
	a209C8	AT1G15430	+2	-136	
	a189D12	AT1G78630	+2	-40	
	a189B10	AT3G54890	+3	-66	
35	a197bH1		+16	-53	
	a209D5		+3	-166	
	a209D8		+3	-10	
40	a189C6	AT1G44575	+3	-43	
	a209A1		+7	-39	
	a192A9		+3	-57	
	a209D6		+4	-56	
45	a209B9	AT3G08940	+3	-39	
	a189D8		+3	-39	
	a189F4		+27	-15	
	a209A9	AT4G12730	+4	-64	
50	a189D6	AT4G20260	+4	-134	
	a209B3	AT2G34420	+5	-51	
	a197bE1		+5	-50	
	a189E10		+6	-50	
55	a197bG5		+9	-46	
	a197bA2		+10	-45	
	a197bG3	AT4G34215	+6	-110	
	a209B10	AT3G26740	+7	-60	
60	a209B2	AT4G14450	+9	irrelevante	(3)
	a063A1	AT3G15353	+9	-71	
	a192A6	AT4G37040	+9	-27	
65	a189E4	AT1G17530	+10	-82	

ES 2 338 564 T3

	Clon	locus ID	GB 5'	ATG	nota
5	a192B9	AT1G15820	+10	-86	
	a209E3		+25	-71	
	a189F11		+32	-64	
	a188H4	AT4G00430	+13	-74	
	a209E6		+26	-61	
10	a189F10	AT2G33800	+13	-66	
	a209C4	AT2G42070	+18	-84	
	a192A8	AT3G14210	+22	-42	
	a209A12		+32	-32	
15	a209B5	AT1G78620	+27	-10	
	a197bF3	AT5G54270	+28	-67	
	a209B12		+31	-64	
	a197bH2		+32	-63	
20	a18909	AT2G19080	+29	-32	
	a197bF1	AT1G20340	+31	-64	
	a209B1	AT2G36830	+33	-60	
25	a209C9	AT5G10450	+33	-23	
	a197bH3	AT1G73230	+38	-86	
	a209C3	AT1G48440	+38	-36	
	a209D7	AT4G25130	+39	-27	(4)
30	a209C1	AT3G26520	+41	-62	
	a209E8	AT4G10340	+43	-52	
	a189C11		+43	-52	
	a209A3		+43	-52	
35	a209A10	AT5G04170	+43	-43	
	a189B12	AT1G28200	+47	-20	
	a197bF5	AT2G33040	+48	-86	
	a189D10	AT5G65430	+49	-39	
40	a192A7	AT1G44810	+54	-25	
	a197bH4	AT1G61520	+54	-88	
	a192D6		+54	-88	
	a197bA1		+54	-88	
45	a189F1		+56	-86	
	a192D5		+59	-83	
	a188A9		+64	-78	
	a192D8		+64	-78	
50	a209D3	AT1G22300	+55	-45	
	a065aE8	AT4G25050	+56	-54	
	a197bG4	AT1G23820	+56	-68	
55	a209B11	AT4G05180	+56	-18	
	a189D3	AT5G03240	+61	-31	
	a209C10	AT1G18080	+61	-63	
	a189E8		+63	-61	
60	a209A5	AT5G01020	+69	-106	
	a065aF7	AT2G39390	+73	-42	
	a063D5	AT1G72020	+77	-16	
65	a197bH6	AT5G48020	+84	-84	

ES 2 338 564 T3

Clon	locus ID	GB 5'	ATG	nota
a209A11	AT1G04170	+85	-116	
a209C11	AT3G26650	+87	-186	
a197bF4		+87	-186	
a209D11		+87	-186	
a192A5		+88	-185	
a188E8		+108	-165	
a189B8	AT5G01530	+112	-35	
a192C5		+120	-27	
a189F7		+139	-8	
a19B11	AT3G26070	+139	-92	
a209A4	AT3G16370	+147	-35	
a192A12	AT4G39090	+148	-9	
a209E1	AT2G20230	+152	-70	
a197bA3	AT2G43750	+152	-79	
a209C7	AT3G47470	+159	-412	
a209D9		+159	-415	
a197b F2	AT2G45440	+209	-246	
a065aBl	AT1G14320	+243	-114	
a209C2	AT1G60950	+466	-51	
a209C6	AT5G64250	+202	-190	(5)
a209B6	AT5G03880	+27	+2	
a192D11	AT3G53750		+7	
a209C12	AT2G14750		+30	
a192D12	AT2G20420		+46	
a192A10	AT5G53480 y AT5G53485		irrelevante	(3)
a192C7	AT5G16715		irrelevante	(3)
a192C12	AT5G02870		>+100	
a189E3	AT1G73940		>+100	
a189D1	AT4G21190		>+100	
a192C10	AT1G56340		>+100	
a209B7	AT5G19540		>+100	
a189C12	AT2G34510		>+100	
a209E12	AT1G62750		>+100	
a188H7			>+100	
a192B5	AT1G49240		>+100	
a192D7	AT1G55490		>+100	
a192D9	AT1G65930		>+100	
a192B12	AT3G61440		>+100	
a192D10	AT2G16400		>+100	
a209E9	AT1G07230		>+100	

- Descripción de la tabla 2: Se indican los clones (referenciados en la 1ª columna) y su región correspondiente en el genoma de *Arabidopsis thaliana* (Locus ID en la 2ª columna). La columna Genbank (tercera columna) proporciona la posición relativa del extremo 5' de los fragmentos clonados con relación a las secuencias publicadas; para cada gen, se ha establecido un "+1" de referencia a partir del ADNc de Genbank que remonta lo más lejos en 5'; las distancias se miden en número de bases de ADNc; los valores negativos indican un fragmento clonado que se extiende más allá del +1 en dirección del promotor y recíprocamente, sabiendo que un 5' idéntico a la secuencia de referencia se traduce por el valor +1 y que el valor -1 indica una ganancia de una base

Clon	locus ID	GB 5'	ATG	nota
hacia el promotor. La columna ATG (cuarta columna) proporciona la posición relativa del 5' clonado con relación al principio del ORF propuesto por la AGI. En negrita se indican los clones de ADNc que cubren la totalidad del ORF y cuyo extremo 5' está localizado corriente arriba o a menos de 10 bases corriente abajo, del +1 de referencia. La columna "nota" (quinta columna) explica ciertas líneas de la tabla que corresponden a los clones: (1) ADNc que cubre 2 loci, el 5' está a 8 bases corriente arriba del ORF AT2G35450 y a más de 530 bases corriente arriba del ORF AT2G35440. (2) ADNc que cubre 2 loci, el 5' está a 72 bases corriente arriba del ORF AT2G42240 y a 661 bases corriente arriba del ORF AT2G42245. (3) Predicciones AGI incompatibles con los datos experimentales. (4) El clon Genbank X97326 muy extendido y no empalmado ha sido excluido del análisis. (5) Este ADNc comprende casi la totalidad de la secuencia codificante predicha, con la excepción de los primeros codones que pertenecen a un primer intrón alternativo cubierto por el constructo. (6) Ningún ADNc había sido descrito antes para este locus.				

TABLA 3

clones de ADNc analizados	154
genes	103
Nuevo	1
Posición relativa del extremo 5' del clon de ADNc	
a) en corriente arriba del +1	12
Posición relativa al extremo 5' del clon de ADNc	
b) identical al +1	20
c) a menos de 10 bases corriente abajo del +1	31
d) en corriente arriba del ATG	134 (87%)
Total (a+b+c)	51 (33%)

Los fragmentos de ADNc contenidos por los 154 constructos han podido ser asignados a 103 loci distintos en el genoma de *A. Thaliana*. Con la excepción de 5 clones, los ADNc obtenidos son globalmente compatibles con las unidades de transcripción y las secuencias codificantes anotadas por la AGI. Un total de 134 ADNc globalmente compatibles con las anotaciones AGI comprenden asimismo el ATG de iniciación; el porcentaje de clones de ADNc que cubren un ORF completo se eleva por lo tanto a ~87% en este experimento. La posición relativa del extremo 5' de los ADNc ha sido comparada con las secuencias de ADNc descritas en Genbank. Se tomará como sitio de iniciación de transcripción de referencia ("++1"), la posición que corresponde a la secuencia de ADNc (de Genbank) que remonta como máximo hacia el 5'. Doce clones de ADNc corresponden a una ganancia de algunas bases hasta varias decenas de bases en dirección del promotor. En particular, el ADNc a189D5 presenta la incorporación de 78 bases suplementarias en 5' con relación a los ADNc ya descritos para este gen. Veinte clones tienen unos extremos 5' idénticos y 31 clones tienen unos 5' localizados a menos de 10 bases corriente abajo de los +1 disponibles en Genbank para los genes correspondientes. Se señalará por último que a pesar de la extensión muy modesta de este inventario de transcritos, se ha caracterizado un clon (secuencia a192B10) que cubre la totalidad de un "gen huérfano" (locus AT5G44730). La obtención de este transcrito que todavía no estaba representado entre las 221.000 secuencias de ADNc disponibles de Genbank para *A. thaliana*, puede ser considerada asimismo como indicador de la buena sensibilidad del método.

Estos resultados permiten concluir que la T4 ADN ligasa se puede utilizar para catalizar la fijación de un oligonucleótido sobre el ARN. Un adaptador parcialmente bicatenario de extremo cohesivo P1/P2 o de un tipo similar a P1/P2, permite obtener un acoplamiento eficaz con unos fragmentos de ARN de secuencias desconocidas. La estrategia de producción de ADNc de longitud completa basada en particular en este procedimiento de etiquetado del ARN se valida de manera convergente mediante RACE PCR y mediante el análisis de un banco de ADNc. La calidad general de los clones obtenidos por los dos enfoques experimentales demuestra la buena sensibilidad de este método para caracterizar el extremo 5' de los transcritos y para la producción de ADNc de longitud completa. En lo referente a la especificidad del método frente al sitio de adición de la protección, los fragmentos de ADNc "más cortos" que se obtuvieron en los dos tipos de experimentos podrían corresponder a unos sitios alternativos de iniciación de transcripción o a una desactivación incompleta mediante la fosfatasa alcalina de ARN mensajeros parcialmente degradados.

REIVINDICACIONES

1. Método de producción de ADNc de longitud completa que comprende una etapa a) de desfosforilación de los fragmentos de ácidos nucleicos no provistos de un protector en 5', una etapa b) de supresión de la protección del extremo 5' de los ARN protegidos, una etapa c) de ligación de oligonucleótidos en el extremo 5' de ARN obtenidos en la etapa b) y una etapa d) de síntesis de los ADNc de longitud completa, mediante retro-transcripción de los ARN obtenidos en la etapa c), mediante una transcriptasa inversa, estando dicho método **caracterizado** porque la etapa c) de ligación del oligonucleótido se lleva a cabo mediante una ligasa y un adaptador bicatenario de extremo cohesivo, comprendiendo dicho adaptador bicatenario de extremo cohesivo una hebra superior que contiene un sitio 3'OH aceptor de ligasa para el acoplamiento con el 5' de los ARN, y una hebra inferior que presenta en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria que puede aparearse con el extremo 5' de los ARN.
2. Método según la reivindicación 1, **caracterizado** porque los ARN protegidos son unos ARN mensajeros.
3. Método según la reivindicación 1, **caracterizado** porque la ligasa utilizada en la etapa c) es una ADN ligasa, una ARN ligasa o una combinación de ARN ligasa y de ADN ligasa.
4. Método según la reivindicación 3, **caracterizado** porque la ligasa utilizada es una ADN ligasa.
5. Método según las reivindicaciones 3 y 4, **caracterizado** porque la ligasa utilizada es la T4 ADN ligasa.
6. Método según la reivindicación 1, **caracterizado** porque la hebra inferior del adaptador bicatenario de extremo cohesivo se puede bloquear mediante un grupo amina en 3'.
7. Método según la reivindicación 1, **caracterizado** porque las hebras superior e inferior del adaptador bicatenario de extremo cohesivo pueden estar constituidas por ADN o por ARN o ser de composición mixta ADN/ARN.
8. Método según la reivindicación 1, **caracterizado** porque la hebra superior del adaptador bicatenario está compuesta por 24 desoxirribonucleótidos en 5' seguido de un fragmento de 8 ribonucleótidos en 3'.
9. Método según la reivindicación 1, **caracterizado** porque la hebra inferior del adaptador bicatenario comprende en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria de 6 bases.
10. Método de ligación de un adaptador oligonucleotídico en el extremo 5' de un ARN, que comprende las etapas siguientes:
 - a) Obtener un ARN monocatenario que presenta un extremo 5' fosfato
 - b) Preparar las dos hebras de un adaptador de extremo cohesivo que comprende una hebra superior que contiene un sitio 3'OH aceptor de ligasa para el acoplamiento con el 5' de los ARN, y una hebra inferior de extremo cohesivo que presenta en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria que puede aparearse con el extremo 5' fosfato de los ARN
 - c) Mezclar dicha hebra superior con dicha hebra inferior de manera que forman un adaptador bicatenario
 - d) Mezclar dicho adaptador con dicho ARN monocatenario para permitir que el extremo 5' de dicha hebra inferior se hibride con dicha hebra sencilla de ARN
 - e) Ligar dicho adaptador con dicha hebra sencilla de ARN mediante una ligasa.
11. Método según la reivindicación 10, **caracterizado** porque el ARN monocatenario es un ARN protegido.
12. Método según la reivindicación 11, **caracterizado** porque el ARN protegido es un ARN mensajero.
13. Método según la reivindicación 10, **caracterizado** porque el ARN monocatenario es un ARN no protegido.
14. Método según la reivindicación 10, **caracterizado** porque la ligasa utilizada en la etapa e) es una ADN ligasa, una ARN ligasa, o una combinación de ARN ligasa y de ADN ligasa.
15. Método según la reivindicación 10, **caracterizado** porque la ligasa utilizada es una ADN ligasa.
16. Método según la reivindicación 10, **caracterizado** porque la ligasa utilizada es la T4 ADN ligasa.
17. Método según la reivindicación 10, **caracterizado** porque la hebra inferior del adaptador bicatenario de extremo cohesivo se puede bloquear mediante un grupo amina en 3'.

ES 2 338 564 T3

18. Método según la reivindicación 10, **caracterizado** porque las hebras superior e inferior del adaptador bicatenario de extremo cohesivo pueden estar constituidas por ADN o por ARN o ser de composición mixta ADN/ARN.

19. Método según la reivindicación 10, **caracterizado** porque la hebra superior del adaptador bicatenario de extremo cohesivo está compuesta por 24 desoxirribonucleótidos en 5' seguido de un fragmento de 8 ribonucleótidos en 3'.

20. Método según la reivindicación 10, **caracterizado** porque la hebra inferior del adaptador bicatenario de extremo cohesivo comprende en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria de 6 bases.

21. Kit para la producción de ADNc de longitud completa, **caracterizado** porque comprende por lo menos:

- fosfatasa alcalina y su tampón de reacción
- pirofosfatasa ácida del tabaco y su tampón de reacción
- una ligasa, o un cóctel de ligasas, que comprende en particular la T4 ADN ligasa y su tampón de reacción
- un adaptador bicatenario de extremo cohesivo, comprendiendo dicho adaptador bicatenario de extremo cohesivo una hebra superior que contiene un sitio 3'OH aceptor de ligasa para el acoplamiento con el 5' de los ARN, y una hebra inferior que presenta en 5' una extensión oligonucleotídica monocatenaria de secuencia aleatoria que puede aparearse con el extremo 5' de los ARN.
- una transcriptasa inversa
- una ADN polimerasa dependiente de ADN
- unos cebadores oligonucleotídicos para la polimerización del ADN.

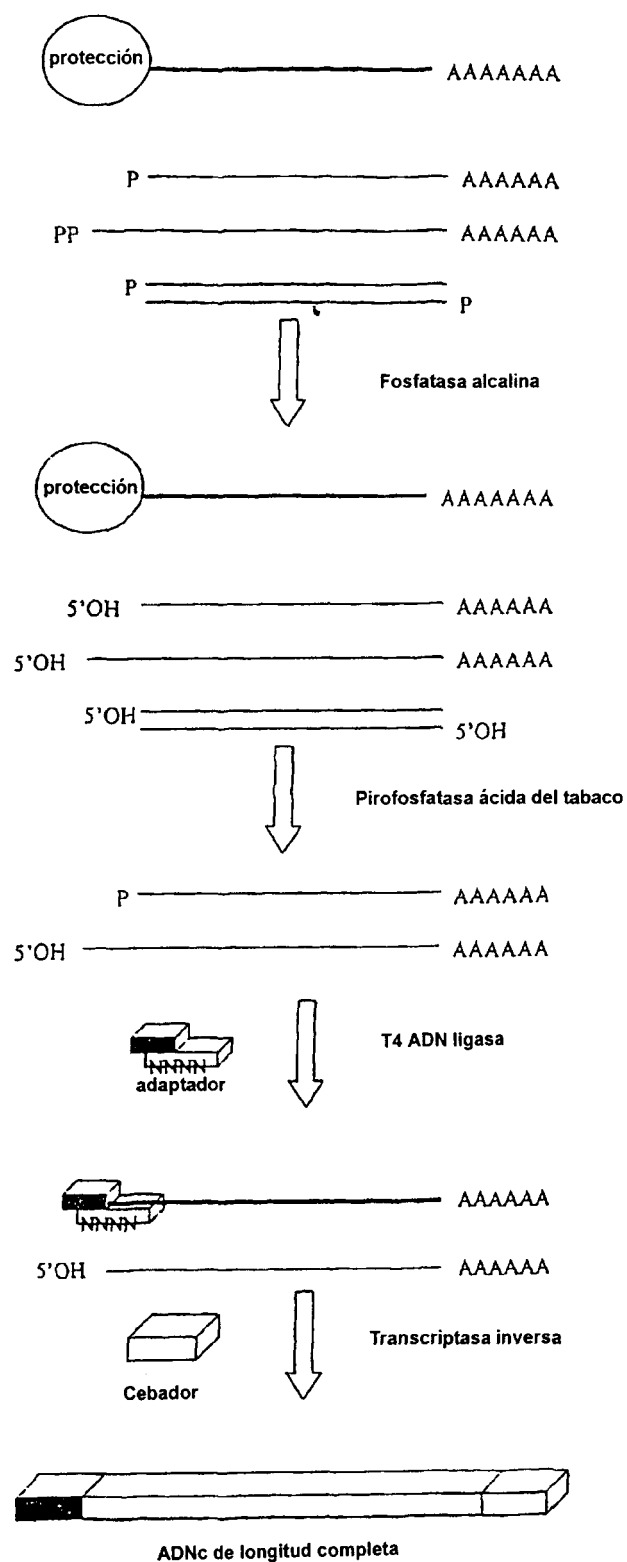


Figura 1

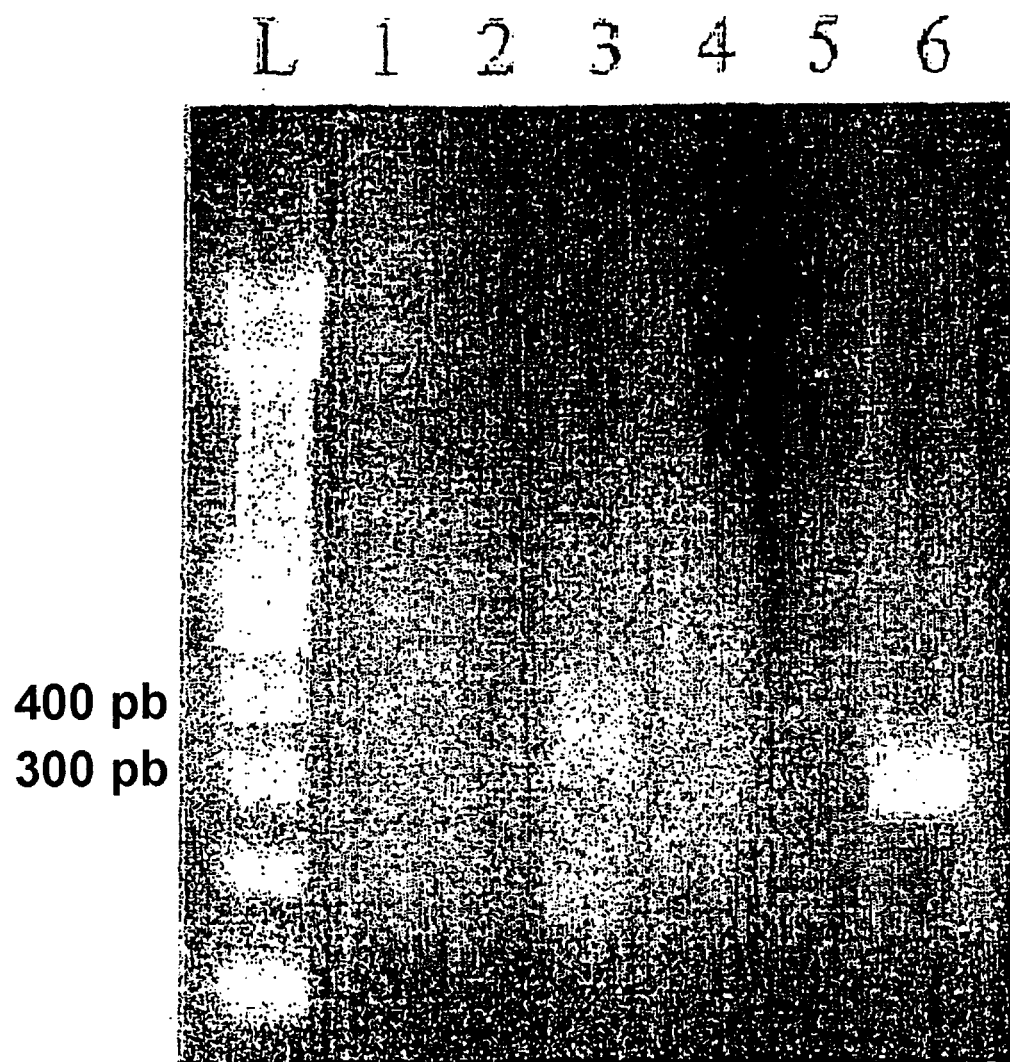


Figura 2

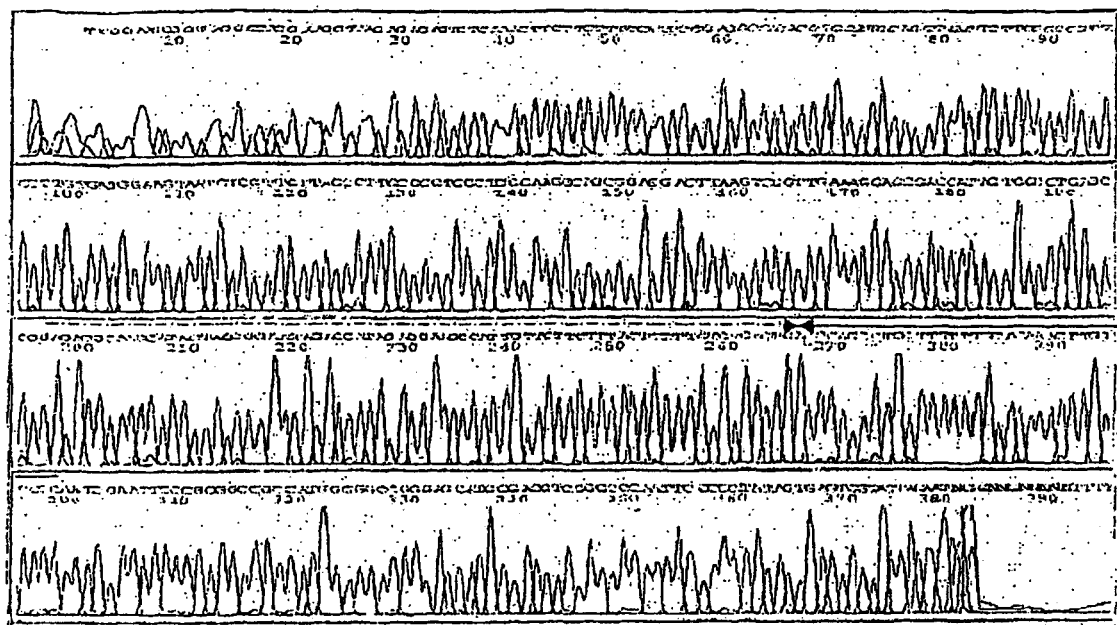


Figura 3

N	Gen	secuencia terminal 5'	Sitio de prot.
	ats1A (X13611)	--GACCAAACCTCAGTCACACAAAGAGTAAAGAA---	
		>+1	
1		AAACCTCAGTCACACAAAGAGTAAAGAA---	-5
3		TCAGTCACACAAAGAGTAAAGAA---	+1
1		CAGTCACACAAAGAGTAAAGAA---	+2
2		AGTCACACAAAGAGTAAAGAA---	+3
	ats3B (X14564)	--CCAGTAGGAAAACAAGTCAGTAAGTAAACGAG---	
		>+1	
1		AAGTCAGTAAGTAAACGAG---	+6
1		GTCAGTAAGTAAACGAG---	+8
	ats1B (X14564)	--CCAATAGAAAAACAATTAAGCAAAAGAAGAAGA---	
		>+1	
1		ATTAAGCAAAAGAAGAAGA---	+7

Figura 4

ES 2 338 564 T3

LISTA DE SECUENCIAS

<110> CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)

5 <120> Método de producción de ADNc de longitud completa mediante ligación de adaptador de extremo cohesivo

<130> D21714

10 <160> 11

<170> PatentIn version 3.2

15 <210> 1

<211> 32

<212> ADN

20 <213> Secuencia artificial

<220>

<221> fuente

25 <222> (1)..(32)

<223> /nota= "Descripción de la secuencia artificial: hebra superior de un adaptador bicatenario a extremo cohesivo (P1)"

<220>

30 <221> misc_feature

<222> (25)..(32)

<223> ribonucleótidos

35 <400> 1

ggggacaagt ttgtacaaaa aagcaggctg gg

32

40 <210> 2

<211> 16

<212> ADN

45 <213> Secuencia artificial

<220>

<221> fuente

50 <222> (1)..(16)

<223> /nota= "Descripción de la secuencia artificial: hebra inferior de un adaptador bicatenario a extremo cohesivo (P2)"

55 <220>

<221> misc feature

<222> (1)..(6)

<223> n corresponde a cualquier tipo de base que puede aparearse con el ARN

60

<220>

<221> misc_feature

<222> (16)..(16)

65 <223> Grupo amina sobre el carbono 3'

ES 2 338 564 T3

<400> 2

nnnnnnccca gcctgc 16

5 <210> 3
 <211> 50
 <212> ADN
 10 <213> Secuencia artificial

<220>
 <221> fuente
 15 <222> (1)..(50)
 <223> /nota= "Descripción de la secuencia artificial: cebador de retro-transcripción P3"

<220>
 20 <221> misc_feature
 <222> (50)..(50)
 <223> n corresponde a cualquier tipo de base que puede aparearse con el ARN

25 <400> 3

ggggaccact ttgtacaaga aagctgggtt tttttttt tttttttvn 50

30 <210> 4
 <211> 21
 <212> ADN
 35 <213> Secuencia artificial

<220>
 <221> fuente
 40 <222> (1)..(21)
 <223> /nota= "Descripción de la secuencia artificial: cebador de PCR Rbc64.U"

<400> 4

45 aatggcttc tctatgctct c 21

<210> 5
 50 <211> 24
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <221> fuente
 <222> (1)..(24)
 <223> /nota= "Descripción de la secuencia artificial: cebador de PCR Rbc402.L"

60 <400> 5

tgcggagaag gtagtcaact tcct 24

65 <210> 6
 <211> 20

ES 2 338 564 T3

	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<221> fuente	
	<222> (1)..(20)	
	<223> /nota= "Descripción de la secuencia artificial: cebador de PCR M1 3fwd"	
10	<400> 6	
	tgtaaaacga cggccagtga	20
15	<210> 7	
	<211> 21	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<221> fuente	
25	<222> (1)..(21)	
	<223> /nota= "Descripción de la secuencia artificial: cebador de PCR M13rev"	
30	<400> 7	
	ggaaacagct atgacatga t	21
35	<210> 8	
	<211> 22	
	<212> ADN	
	<213> Secuencia artificial	
40	<220>	
	<221> fuente	
	<222> (1)..(22)	
45	<223> /nota= "Descripción de la secuencia artificial: cebador de PCR Univ221.U"	
	<400> 8	
50	tgtaaaacga cggccagtct ta	22
	<210> 9	
	<211> 22	
55	<212> ADN	
	<213> Secuencia artificial	
	<220>	
60	<221> fuente	
	<222> (1)..(22)	
	<223> /nota= "Descripción de la secuencia artificial: cebador de PCR Univ221.L"	
65	<400> 9	
	caggaaacag ctatgacat gt	22

ES 2 338 564 T3

<210> 10
<211> 29
<212> ADN
5 <213> Secuencia artificial

<220>
<221> fuente
10 <222> (1)..(29)
<223> /nota= "Descripción de la secuencia artificial: cebador de PCR Piper"

<400> 10
15
 ggggacaagt ttgtacaaaa aagcaggct 29

<210> 11
20 <211> 29
<212> ADN
<213> Secuencia artificial

25 <220>
<221> fuente
<222> (1)..(29)
30 <223> /nota= "Descripción de la secuencia artificial: cebador de PCR P3pcr"

<400> 11
35 ggggaccact ttgtacaaga aagctgggt 29

40

45

50

55

60

65