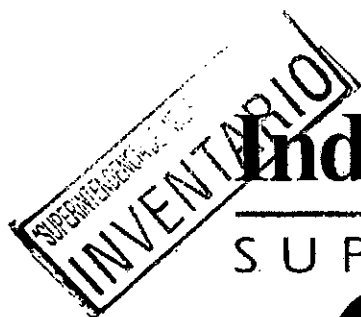




SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-050300-00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 220

Industria y Comercio

SUPERINTENDENCIA

PLT

Q P

SOLICITUD

PATENTE DE INVENCION

11-50300

21. EXPEDIENTE N°

54. TÍTULO VECTORES DE TERAPIA DE GEN Y DESAMINASAS
CITOSINA.

51. CLASIFICACIÓN INTERNACIONAL C12N 15/52

71. SOLICITANTE TOCAGEN INC.

DOMICILIO San Diego, California, Estados Unidos de América

74. APODERADO DILIA MARIA RODRIGUEZ D'ALEMAN

22. BOGOTÁ, D.C.

Pl, C
26-04-2011



No. 11-050300-00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eje: 378 FASENACIONAL
Act. 411 PRESENTACION Folios: 220

**FORMULARIO ÚNICO DE SOLICITUD DE PATENTE
PCT
ENTRADA EN FASE NACIONAL**

SOLICITUD DE:

1

☒

Patente de Invención

☐

Patente de Modelo de Utilidad

☒

Capítulo I.

☐

Capítulo II.

2

**SOLICITANTE
(71)**

Nombre: TOCAGEN INC.

Dirección: 3030 Bunker Hill Street, Suite 230,
San Diego, California 92109, Estados Unidos de
América.

Nacionalidad o domicilio: San Diego,
California,, Estados Unidos de América.

Lugar de Constitución: San Diego,
California, Estados Unidos de América

Teléfono:

Fax:

E-mail

IDENTIFICACIÓN

C.C. ☐

NIT ☐

C.E. ☐

Otro ☐

Número

3

**REPRESENTANTE
O APODERADO
(74)**

Nombre: DILIA MARÍA RODRÍGUEZ
D'ALEMAN

Dirección: Carrera 11 No. 86-53, piso 6

Teléfono: 6 181088

Fax: 6350824

E-mail: clarkemodet@clarkemodet.com.co

IDENTIFICACIÓN

C.C. ☒

NIT ☐

C.E. ☐

Otro ☐

Número 41.654.882
TP 20824 CSJ

4

**INVENTOR (ES)
(72)**

Nombre: VER HOJA ANEXA

Dirección:

Nacionalidad o domicilio:

5 **PCT (86)**

Solicitud Internacional No. PCT/US2009/058510

Fecha 26 de septiembre de 2009

Publicación Internacional No. WO 2010/045002 A2

Fecha 22 de abril de 2010

6 **TÍTULO (54)** VECTORES DE TERAPIA DE GEN Y DESAMINASAS CITOSINA.

7 **Clasificación Internacional (51)** C12N 15/52

8 **Prioridad**

SI ☒ NO ☐

(33) País de Origen

Estados Unidos de América

Estados Unidos de América

Estados Unidos de América

(32) Fecha

26 de septiembre de 2008

08 de diciembre de 2008

13 de junio de 2009

(31) Número de Solicitud

61/100,666

61/120,618

61/186,823

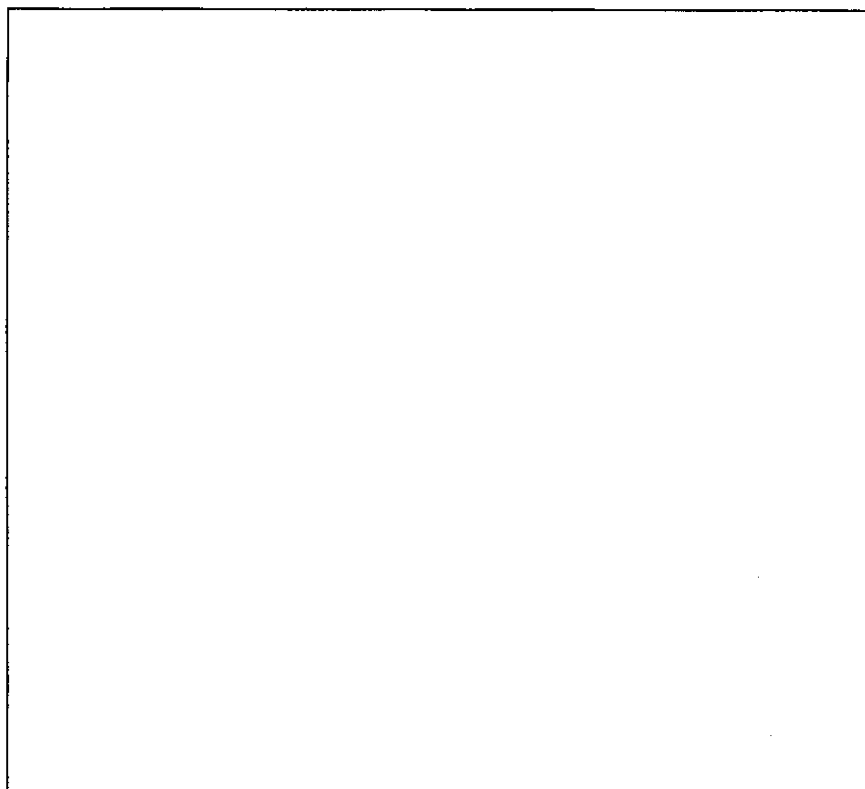
9 **Comprobante de pago No.** 11-35996// 11-36016//11-36017//11-36018//11-36186// 11-36296 //
11-36316 // 11-27069// **Fecha** 15 de abril de 2011

10

ANEXOS

- ☐ Comprobante de pago de la tasa de presentación de la solicitud
- ☐ Documento que acredite la existencia y representación legal cuando el solicitante sea persona jurídica
- ☐ Poderes si fuere el caso
- ☐ Documento de cesión del inventor al solicitante o a su causante
- ☐ Declaraciones
- ☐ Copia de la solicitud internacional. (Resumen, descripción, reivindicaciones, dibujos)
- ☐ Traducción de la solicitud internacional al castellano. (Resumen, descripción, reivindicaciones, dibujos)
- ☐ Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen Preliminar (IPEA).
- ☐ Traducción del examen preliminar internacional efectuado por la IPEA
- ☐ De ser el caso copia de contrato de acceso.
- ☐ De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas
- ☐ De ser el caso, certificado de deposito del material biológico
- ☐ Arte final 12 X 12 .
- ☐ De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- ☐ De ser el caso, modificaciones efectuadas a la solicitud internacional
- ☐ Otros

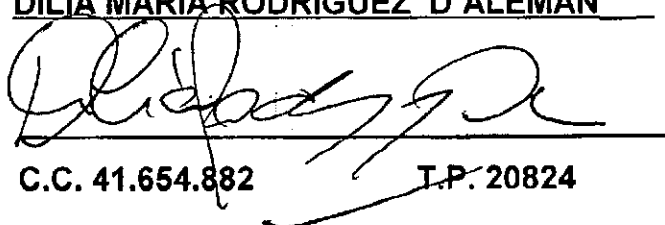
11 FIGURA CARACTERÍSTICA



12 Solicito la concesión de la patente

NOMBRE : DILIA MARÍA RODRÍGUEZ D'ALEMAN

FIRMA:


C.C. 41.654.882 T.P. 20824

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -

 Industria y Comercio SUPERINTENDENCIA	RECIBO DE CAJA	No. 11 - 0035996
	Bogotá D.C., Abril 15 de 2011 - 18:03:59	

RECIBIDO DE : CLARKE MODET

NI 830.008.643

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	431342063	15/04/2011	108.185.000.00

*** Conceptos Pagados ***

CANT. RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1 50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE INVENCIÓN	571.000.00	571.000.00
			\$571.000.00

SON: **QUINIENTOS SETENTA Y UN MIL PESOS MONEDA CORRIENTE***

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-050300-00000-0000

Fecha: 2011-04-25 15:43:45
 Tra. 11 PATENTE IYII
 Act. 411 PRESENTACION

Dep. 2020 DIR. NUEVAS CR
 Eve: 378 FASE NACIONAL I
 Folios: 220

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Abril 15 de 2011 - 18:06:29

3

2-2866

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

-/-


Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 0036016

Bogotá D.C., Abril 15 de 2011 - 18:03:59

RECIBIDO DE : CLARKE MODET

NI 830.008.643

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	431342063	15/04/2011	108.185.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	500.000.00	500.000.00
				\$500.000.00

=====

SON: **QUINIENTOS MIL PESOS MONEDA CORRIENTE**

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-050300- -00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 220

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Abril 15 de 2011 - 18:06:29

u
p-2866

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -



RECIBO DE CAJA

No. 11 - 0036017

Bogotá D.C., Abril 15 de 2011 - 18:03:59

RECIBIDO DE : CLARKE MODET

NI 830.008.643

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	431342063	15/04/2011	108.185.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	VR. UNDITARIO	VR. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	500.000.00	500.000.00
				\$500.000.00

SON: **QUINIENTOS MIL PESOS MONEDA CORRIENTE***

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-050300- -00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 220

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Abril 15 de 2011 - 18:06:20

5

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 0036018

Bogotá D.C., Abril 15 de 2011 - 18:03:59

RECIBIDO DE : CLARKE MODET

NI 830.008.643

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	431342063	15/04/2011	108.185.000.00

*** Conceptos Pagados ***

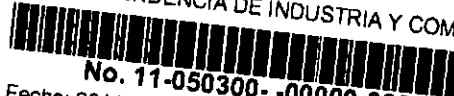
CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1 50005-01-09	OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	500.000.00	500.000.00
				\$500.000.00

SON: **QUINIENTOS MIL PESOS MONEDA CORRIENTE**

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-050300-00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eje: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 220

6

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Abril 15 de 2011 - 18:06:29

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -



RECIBO DE CAJA

No. 11 - 0036186

Bogotá D.C., Abril 15 de 2011 - 18:03:59

RECIBIDO DE : CLARKE MODET

NI 830.008.643

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	431342063	15/04/2011	108.185.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	VL. UNDITARIO	VL. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	50.000.00	50.000.00
				<u>\$50.000.00</u>

SON: **CINCUENTA MIL PESOS MONEDA CORRIENTE***

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-050300-00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
 Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
 Act. 411 PRESENTACION Folios: 220

X

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Abril 15 de 2011 - 18:06:34

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -



RECIBO DE CAJA

No. 11 - 0036296

Bogotá D.C., Abril 15 de 2011 - 18:03:59

RECIBIDO DE : CLARKE MODET

NI 830.008.643

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	431342063	15/04/2011	108.185.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	10.000.00	10.000.00
				\$10.000.00

SON: **DIEZ MIL PESOS MONEDA CORRIENTE**

Responsable _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-050300-00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
 Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
 Act. 411 PRESENTACION Folios: 220

4

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C. - Colombia

Web: www.sic.gov.co e-mai: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Abril 15 de 2011 - 18:05:38

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

-/-

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 0036316

Bogotá D.C., Abril 15 de 2011 - 18:03:59

RECIBIDO DE : CLARKE MODET

NI 830.008.643

***** Soporte del Pago *****

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	431342063	15/04/2011	108.185.000.00

***** Conceptos Pagados *****

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	5.000.00	5.000.00
				\$5.000.00

SON: **CINCO MIL PESOS MONEDA CORRIENTE**

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-050300- -00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 220

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mai: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Abril 15 de 2011 - 18:08:38

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -



RECIBO DE CAJA

No. 11 - 0027069

Bogotá D.C., Marzo 25 de 2011 - 07:06:52

RECIBIDO DE : CLARKE MODET

NI 830.008.643

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	325077546	23/03/2011	95.510.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	VR.UNDITARIO	VR.CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	1.000.00	1.000.00
				\$1.000.00

SON: **MIL PESOS MONEDA CORRIENTE***

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-050300- -00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
 Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
 Act. 411 PRESENTACION Folios: 220

10

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Marzo 25 de 2011 - 07:07:26

VECTORES DE TERAPIA DE GEN Y DESAMINASAS CITOSINA**REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS**

5 [0001] Esta solicitud reivindica la prioridad de la Solicitud Provisional Estadounidense Serie No. 61/100,666, presentada en Septiembre 26, 2008, la Solicitud Provisional Estadounidense Serie No. 61/120,618, presentada en Diciembre 8, 2008 y la Solicitud Provisional Estadounidense Serie No. 61/186,823, 10 presentada en Junio 13, 2009 cuyas descripciones se incorporan aquí como referencia.

CAMPO TÉCNICO

15 [0002] Esta descripción se relaciona con desaminasas citosina modificadas (CD). La descripción se relaciona adicionalmente con células que expresan tales CD modificados y métodos para utilizar tales CD modificados en el tratamiento de enfermedades y trastornos.

20

ANTECEDENTE

[0003] La levadura, o desaminasa citosina bacteriana se convierte en el profármaco antibiótico inocuo 5-FC en el agente 25 quimioterapéutico citotóxico 5-fluorouracilo (5-FU). No se sabe que los humanos (y los mamíferos en general) posean un gen de ocurrencia natural que codifica una enzima con actividad desaminasa citosina significativa. La levadura y la desaminasa

citosina bacteriana tienen gran reconocimiento en el tratamiento de cánceres utilizando suministro de gen y vectores víricos para el suministro de la enzima seguido por tratamiento con 5-FC, que luego se convierte de la enzima a un fármaco citotóxico (Miller et al., Can Res 62:773-780 2002; Kievit et al., Can Res 59:1417-1421 1999).

RESUMEN

10 [0004] Se proporcionan aquí polipéptidos que se convierten de 5-FC a 5-FU. También se proporcionan moléculas de ácido nucleico que codifican tales polipéptidos, células que expresan tales polipéptidos, vectores que contienen tal polinucleótido y polipéptidos y métodos para sintetizar 5-FU o sus derivados de
15 uno adecuado utilizando tales polipéptidos. De acuerdo con lo anterior, en varias realizaciones, se proporcionan polipéptidos recombinantes o aislados que comprenden una desaminasa citosina que tiene una secuencia como se establece en la SEQ ID NO: 2. En otras realizaciones, el polipéptido comprende SEQ ID NO: 2 con
20 una mutación seleccionada del grupo que consiste de A23L, V108I, I140L y cualquier combinación de los mismos. En todavía otra realización, el polipéptido comprende la secuencia como se establece en la SEQ ID NO: 4 e incluye hasta 50, 25, 10, o 5 sustituciones de aminoácido conservadoras que excluyen los
25 residuos: (a) 23, 108 y 140. En general, los polipéptidos proporcionados aquí exhiben actividad desaminasa citosina útil para convertir 5-FC a 5-FU.

[0005] La descripción también proporciona un polipéptido que comprende una secuencia que es por lo menos 80%, 90%, 95%, 98% o 99% idénticas a la SEQ ID NO: 4, en donde el polipéptido tiene una leucina en la posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y en donde el polipéptido tiene actividad desaminasa citosina. En todavía otra realización, el polipéptido comprende la secuencia como se establece en la SEQ ID NO: 4.

10 [0006] La descripción proporciona adicionalmente construcciones de fusión que comprenden uno cualquiera de los polipéptidos anteriores ligados operablemente a una uracil fosforibosiltransferasa (UPRT) o un orotato fosforibosiltransferasa (OPRT). En una realización, la construcción de fusión comprende un primer polipéptido que comprende SEQ ID NO: 2 con una mutación seleccionada del grupo que consiste de A23L, V108I, I140L y cualquier combinación de los mismos ligado a un segundo polipéptido que tiene actividad UPRT u OPRT. En todavía otra realización, la construcción de fusión comprende un primer polipéptido que tiene la secuencia como se establece en la SEQ ID NO: 4 e incluye hasta 50, 25, 10, o 5 sustituciones de aminoácido conservadoras que excluyen los residuos: 23, 108 y 140 en donde el primer polipéptido se liga a un segundo polipéptido que tiene actividad UPRT o OPRT. La descripción también proporciona una construcción de fusión que comprende un primer polipéptido que tiene actividad desaminasa citosina y que comprende una secuencia que es por lo menos 80%, 90%, 95%, 98% o 99% idénticas a la SEQ ID NO:4, en donde el

polipéptido tiene una leucina en la posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y en donde el primer polipéptido se liga a un segundo polipéptido que comprende actividad UPRT o OPRT. En una realización, los
5 polipéptidos que tiene actividad UPRT u OPRT comprende una secuencia como se establece en la SEQ ID NO: 8 y 10, respectivamente, o sus variantes. En todavía otra realización, un primer polipéptido que comprende actividad desaminasa citosina se liga a un polipéptido que tiene actividad UPRT u OPRT, en donde
10 los polipéptidos se separan mediante un ligador de péptido. En otra realización, la construcción de fusión comprende una secuencia seleccionada del grupo que consiste de de la SEQ ID NO: 12, 14, 16 o 18.

15 [0007] La descripción proporciona adicionalmente polinucleótidos que codifican cualquiera de los polipéptidos anteriores. Por ejemplo, la descripción proporciona un polinucleótido que codifica un polipéptido que comprende SEQ ID NO: 2 con una mutación seleccionada del grupo que consiste de A23L, V108I,
20 I140L y cualquier combinación de los mismos. En todavía otra realización, el polinucleótido codifica un polipéptido que comprende una secuencia como se establece en la SEQ ID NO: 4 e incluye hasta 50, 25, 10, o 5 sustituciones de aminoácido conservadoras que excluyen los residuos: (a) 23, 108 y 140. En
25 una realización adicional, la descripción proporciona un polinucleótido que codifica un polipéptido que comprende una secuencia que es por lo menos 80%, 90%, 95%, 98% o 99% idénticas a la SEQ ID NO:4, en donde el polipéptido tiene una leucina en la

posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y en donde el polipéptido tiene actividad desaminasa citosina. En todavía otra realización, el polinucleótido codifica un polipéptido que comprende SEQ ID NO: 4. En una realización adicional, el polinucleótido codifica un polipéptido que comprende una secuencia como se establece en la SEQ ID NO: 4, 6, 8, 10, 11, 12, o 13.

[0008] La descripción proporciona un polinucleótido optimizado de codón humano que codifica una desaminasa citosina. En una realización, el polinucleótido de codón humano optimizado comprende una secuencia como se establece en la SEQ ID NO: 3. En todavía otra realización, el polinucleótido comprende una secuencia como se establece en la SEQ ID NO: 5, 7, o 9.

[0009] En otras realizaciones, una desaminasa citosina o construcción de fusión de la descripción se suministra utilizando un suministro del sistema de gen (GDS). En otro aspecto, el polinucleótido que codifica una desaminasa citosina se suministra con un GDS que es un vector vírico o vírico derivado. El vector vírico se puede replicar o no replicar, y puede ser un vector adenovírico, un vector sarampión, un vector de herpes, un vector retrovírico (que incluye un vector lentivírico), un vector rabdovírico tal como un vector vírico de Estomatitis Vesicular, un vector de reovirus, un vector del virus Seneca Valley, un vector poxvirus (que incluye vectores derivados de vacunas o viruela de animal), un vector parvovirus (que incluye un vector AAV), un vector alfavirus u otro vector vírico conocido por un

experto en la técnica. En una realización, el vector vírico puede ser un vector retrovírico de replicación competente capaz de infectar células de replicación de mamífero. El vector retrovírico de replicación competente puede comprender un
5 Ortoretrovirus o más típicamente un vector de retrovirus gama. En un aspecto, un vector retrovírico de replicación competente comprende un sitio de entrada ribosómica interna (IRES) 5' en el polinucleótido que codifica una desaminasa citosina de la descripción. En una realización, el polinucleótido que codifica
10 una desaminasa citosina es 3' en un polinucleótido ENV de un vector retrovírico.

[0010] En otras realizaciones, se proporcionan células anfitrionas transfectadas con una desaminasa citosina o
15 construcción de fusión de la descripción, o un vector que incluye un polinucleótido o construcción de fusión de la descripción. Las células anfitrionas incluyen células eucarióticas tal como células de levadura, células de insectos, o células de animal. Las células anfitrionas también incluyen células procarióticas
20 tal como células bacterianas.

[0011] La descripción también proporciona métodos para tratar un trastorno proliferativo de la célula que incluye un cáncer que comprende administrar un polinucleótido o polipéptido de la
25 descripción a un sujeto y poner en contacto el sujeto con un fármaco citotóxico que comprende 5- fluorocitosina (5-FC).

[0012] La descripción proporciona un retrovirus competente de replicación recombinante (RCR) que comprende: una proteína GAG retrovímica; una proteína POL retrovímica; una envoltura retrovímica; un polinucleótido retrovímico que comprende 5 Secuencias (LTR) de Repetición de Terminal Largo en el extremo 3' de la secuencia de polinucleótido retrovímica, una secuencia promotora en el extremo 5' del polinucleótido retrovímico, dicho promotor es adecuado para la expresión en una célula de mamífero, un dominio de ácido nucleico gag, un dominio de ácido nucleico 10 pol y un dominio de ácido nucleico env; un casete que comprende un sitio interno de entrada al ribosoma (IRES) ligado operablemente a un polinucleótido heterólogo, en donde el casete se posiciona 5' en el LTR 3' y 3' en el dominio de ácido nucleico env que codifica la envoltura retrovímico; y las secuencias que 15 actúan cis necesarias para transcripción inversa, empaque e integración en una célula objetivo, en donde el RCR mantiene competencia de replicación mayor después de 6 pasajes comparado con un vector pACE (SEQ ID NO:21) . En una realización, la secuencia de polinucleótido retrovímica se deriva de virus de 20 leucemia de murino (MLV), Virus de leucemia de murino de Moloney (MoMLV), Virus de leucemia felina o Virus de leucemia de mono Gibbon (GALV). En otra realización, el MLV es un MLV anfotrófico. En todavía otra realización, el retrovirus es un oncoretrovirus. En todavía otra realización, la célula objetivo es una célula que 25 tiene un trastorno proliferativo de la célula. El trastorno proliferativo de la célula se puede seleccionar del grupo que consiste de, pero no se limita a, cáncer de pulmón, cáncer de colon-recto, cáncer de mama, cáncer de próstata, cáncer del

tracto urinario, cáncer uterino, cáncer de cerebro, cáncer de cabeza y cuello, cáncer pancreático, melanoma, cáncer de estómago y cáncer de ovario, artritis reumatoide y otras enfermedades autoinmunes. En una realización, el promotor comprende un
5 promotor CMV que tiene una secuencia como se establece en la SEQ ID NO: 19, 20, o 22 a partir del nucleótido 1 a cerca del nucleótido 582 y puede incluir la modificación en una o más bases de ácido nucleico y que es capaz de dirigir y iniciar la transcripción. En todavía una realización adicional, el promotor
10 comprende una secuencia como se establece en la SEQ ID NO: 19, 20, o 22 a partir del nucleótido 1 a cerca del nucleótido 582. En una realización adicional, el promotor comprende un polinucleótido de dominio CMV-R-U5. En una realización, el dominio CMV-R-U5 comprende el promotor inmediatamente temprano
15 del citomegalovirus humano ligado a una región MLV R-U5. En todavía otra realización, el polinucleótido de dominio CMV-R-U5 comprende una secuencia como se establece en la SEQ ID NO: 19, 20, o 22 desde cerca del nucleótido 1 a cerca del nucleótido 1202 o las secuencias que son por lo menos 95% idénticas a una
20 secuencia como se establece en la SEQ ID NO: 19, 20, o 22, en donde el polinucleótido promueve la transcripción de una molécula de ácido nucleico ligada operablemente a esta. En otra realización, el gag y pol del polinucleótido se derivan de un oncoretrovirus. El dominio de ácido nucleico gag puede comprender
25 una secuencia desde cerca del nucleótido número 1203 a cerca del nucleótido 2819 de la SEQ ID NO: 19 o una secuencia que tiene por lo menos 95%, 98%, 99% o 99.8% de identidad de esta. El dominio pol puede comprender una secuencia desde cerca del nucleótido

número 2820 a cerca del nucleótido 6358 de la SEQ ID NO: 19 o una secuencia que tiene por lo menos 95%, 98%, 99% o 99.9% de identidad de esta. En una realización, el dominio env codifica una proteína env amfotérica. El dominio env puede comprender una

5 secuencia desde cerca del nucleótido número 6359 a cerca del nucleótido 8323 de la SEQ ID NO: 19 o una secuencia que tiene por lo menos 95%, 98%, 99% o 99.8% de identidad de esta. El dominio IRES del vector puede ser cualquier IRES, sin embargo, en una realización el IRES se deriva de un virus de encefalomiocarditis.

10 En una realización adicional, el IRES comprende una secuencia desde cerca del nucleótido número 8327 a cerca del nucleótido 8876 de la SEQ ID NO: 19 o una secuencia que tiene por lo menos 95%, 98%, o 99% de identidad de esta. En todavía una realización adicional, el polinucleótido heterólogo comprende un

15 polinucleótido que tiene una secuencia como se establece en la SEQ ID NO: 3, 5, 11, 13, 15 o 17. En una realización adicional, la secuencia heteróloga codifica un polipéptido que comprende una secuencia como se establece en la SEQ ID NO: 4. El ácido nucleico heterólogo es codón humano optimizado y codifica un polipéptido

20 como se establece en la SEQ ID NO: 4. En una realización adicional, el ácido nucleico heterólogo comprende una secuencia como se establece en la SEQ ID NO: 19 desde cerca del nucleótido número 8877 a cerca de 9353. En una realización, el LTR 3' se deriva de un oncoretrovirus. En una realización adicional, el LTR

25 3' comprende un dominio U3-R-U5. En todavía una realización adicional, el LTR 3' comprende una secuencia como se establece en la SEQ ID NO: 19 desde cerca del nucleótido 9405 a cerca de 9998

o una secuencia que es por lo menos 95%, 98% o 99.5% idénticas a esta.

[0013] La descripción proporciona un polinucleótido que comprende una secuencia como se establece en la SEQ ID NO: 19, 20, o 22.

[0014] La descripción proporciona un polinucleótido aislado que comprende de 5' a 3': una fusión CMV-R-U5 del promotor inmediato temprano del citomegalovirus humano a una región MLV R-U5; un PBS, el sitio de unión de cebador para transcriptasa inversa; un sitio de división 5'; señal de empaque ψ ; una secuencia codificante gag para el antígeno específico del grupo MLV; una secuencia codificante pol para la poliproteína de polimerasa MLV; un sitio de división 3'; una secuencia codificante env 4070A para la proteína de envoltura de la cepa MLV 4070A; un sitio interno de entrada al ribosoma (IRES) del virus de encefalomiocarditis; una secuencia codificante desaminasa citosina modificada; un tracto polipurina; y una repetición de terminal largo U3-R-U5 MLV.

20

[0015] La descripción proporciona un método para tratar un sujeto con un trastorno proliferativo de la célula que comprende poner en contacto el sujeto con un polinucleótido un polipéptido de la descripción que tiene actividad desaminasa citosina bajo condiciones de tal manera que el polinucleótido se expresa y se pone en contacto con el sujeto con 5- fluorocitosina.

21

[0016] La descripción también proporciona un método para tratar un trastorno proliferativo de la célula en un sujeto que comprende poner en contacto el sujeto con un retrovirus de la descripción, en donde la secuencia de ácido nucleico heterólogo
5 codifica una proteína terapéutica que inhibe la proliferación de una célula neoplásica. En una realización, el retrovirus comprende un polinucleótido que codifica un polipéptido que tiene una secuencia como se establece en la SEQ ID NO: 4, 12, 14, 16, o 18.

10

[0017] Los detalles de una o más realizaciones de la descripción se establecen en los dibujos que acompañan y la descripción adelante. Serán evidentes otras características, objetos y ventajas de la descripción y los dibujos, y de las
15 reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0018] La figura 1A-D muestra (a) un esquema de un vector
20 retroviral recombinante de la descripción; (b) un mapa de plásmido de un polinucleótido de la descripción; (c y d) una secuencia de un polinucleótido de la descripción (SEQ ID NO: 19).

[0019] La **figura 2A-D** muestra esquemas para la generación de
25 varias realizaciones de la descripción que comprende polipéptidos con actividad CD, OPRT y UPRT.

22

[0020] La **figura 3** muestra que se observan altos niveles de la proteína yCD2 comparado con la proteína yCD tipo natural en células infectadas U-87.

5 [0021] La **figura 4** muestra la estabilidad de un vector que comprende un polipéptido CD de la descripción y la comparación con otros vectores de la descripción.

[0022] La **figura 5A-B** muestra que una actividad desaminasa
10 citosina y el vector de la descripción proporciona mejor expresión o expresión comparable, y por lo tanto muerte, de la rata RG2 infectada (5A) o de las células U87 (5B) comparado con la actividad yCD tipo natural (T5.0007), cuando se exponen las células infectadas a niveles incrementados de 5-FC.

15

[0023] La **figura 6** muestra que la actividad específica en células U87 infectadas de T5.0002 (yCD2) es mayor que T5.0001 (yCD parcialmente optimizado) que es mayor que T5.0007 (yCD en peso).

20 [0024] Los símbolos de referencia en los varios dibujos indica elementos similares.

DESCRIPCIÓN DETALLADA

25 [0025] Como se utiliza aquí y en las reivindicaciones adjuntas, las formas singulares "un", "y", y "el" incluyen referentes de plural a menos que el contexto indique claramente otra cosa. Así, por ejemplo, la referencia a "una célula" incluye una pluralidad

23

de tales células y la referencia a "el agente" incluye la referencia a uno o más agentes conocidos por aquellos expertos en la técnica, y así establecidos.

5 [0026] También, el uso de "o" significa "y/o" a menos que de indique otra cosa. De forma similar, "comprende", "comprenden", "que comprende" "incluyen", "incluye", y "que incluye" son intercambiables y no están destinados a ser limitantes.

10 [0027] Se entiende adicionalmente que en donde las descripciones de varias realizaciones utilizan el término "que comprende," aquellos expertos en la técnica entenderán que en algunos casos específicos, una realización se puede describir alternativamente utilizando el idioma "que consiste esencialmente de" o "que
15 consiste de."

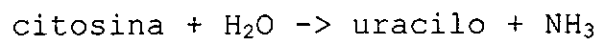
[0028] A menos que se defina otra cosa, todos los términos químicos o científicos utilizados aquí tienen el mismo significado como se entiende comúnmente por parte un experto en
20 la técnica a la cual pertenece esta descripción. Aunque se pueden utilizar los métodos y materiales similares o equivalentes a aquellos descritos aquí en la práctica de los métodos y composiciones descritas, se describen aquí los métodos de ejemplo, dispositivos y materiales.

25 [0029] Las publicaciones discutidas anteriormente y a través del texto se proporcionan únicamente para su descripción antes de la fecha de presentación de la solicitud actual. Nada aquí se

24

constituye como una admisión que los inventores no tienen derecho a anticipar tal descripción por virtud de la descripción anterior.

- 5 [0030] Una desaminasa citosina (EC 3.5.4.1) es una enzima que cataliza la reacción química



- 10 Así, los dos sustratos de esta enzima son citosina y H_2O , mientras que sus dos productos son uracilo y NH_3 . Esta enzima pertenece a la familia de las hidrolasas, aquellas que actúan en los enlaces de carbono-nitrógeno diferentes a los enlaces de péptido, específicamente en amidinas cíclicas. El nombre
15 sistemático de esta clase de enzima es citosina aminohidrolasa. Esta enzima también se denomina isodesaminasa citosina. Esta enzima participa en el metabolismo de la pirimidina.

- [0031] Más particularmente, la desaminasa citosina es una enzima
20 involucrada en la ruta metabólica para las pirimidinas, a través de las cuales se transforma la citosina exógena, por medio de la desaminación hidrolítica, dentro del uracilo. Las actividades de desaminasa citosina (CDasa o CD) se han demostrado en procariotes y eucariotes inferiores, pero ellos están ausentes en los
25 mamíferos (Koechlin et al., 1966, Biochem. Pharmacol. 15, 435-446; Polak et al., 1976, Chemotherapy 22, 137-153). El gen FCY1 de *Saccharomyces cerevisiae* (*S. cerevisiae*) y el gen codificante de *E. coli*, que codifica, respectivamente, la CDasa de estos dos

organismos, se conocen y sus secuencias se publican (EP 402 108; Erbs et al., 1997, Curr. Genet. 31, 1-6; WO 93/01281). La CDasa también desamina un análogo de citosina, 5-fluorocitosina (5-FC) a 5- fluorouracilo (5-FU), que es un compuesto altamente citotóxico, en particular cuando este se convierte a 5-fluoro-UMP (5-FUMP). Las células que carecen de actividad CDasa, debido a una mutación inactivante del gen que codifica la enzima o a su deficiencia natural para esta enzima (por ejemplo células de mamífero) son resistentes a 5-FC (Jund and Lacroute, 1970, J. Bacteriol. 102, 607-615; Kilstrup et al., 1989, J. Bacteriol., 171, 2124-2127). Por otra parte, se ha demostrado que es posible transmitir la sensibilidad de 5-FC a las células de mamífero dentro de las cuales se ha transferido la secuencia que codifica una actividad CDasa (Huber et al., 1993, Cancer Res. 53, 4619-4626; Mullen et al., 1992, Proc. Natl. Acad. Sci. USA 89, 33-37; WO 93/01281). De acuerdo con lo anterior, el uso de CD es ventajoso en el contexto de la terapia de gen, en particular terapia de gen antineoplásica.

[0032] Sin embargo, la sensibilidad del 5-FC varía a un gran grado dependiendo de las estirpes celulares. Se observa baja sensibilidad, por ejemplo, en PANC-I (carcinoma del páncreas) y estirpes celulares de tumor humano SK-BR-3 (adenocarcinoma de mama) se transducen con un retrovirus que expresa el gen codificante de E. coli (Harris et al., 1994, Gene Therapy 1, 170-175). Este fenómeno se explica mediante la ausencia o pobre conversión endógena del 5-FU formado por la acción enzimática de la CDasa, con el 5-FUMP citotóxico. Esta etapa, que se lleva a

cabo normalmente con células de mamífero mediante orotato fosforibosiltransferasa (OPRTasa), puede estar ausente en varios tumores y así hace terapia de gen con base en la CDasa inefectiva. En los procariotos y eucariotos inferiores, el uracilo se transforma dentro de UMP a través de la acción del uracil fosforibosiltransferasa (actividad UPRTasa). Esta enzima también se convierte de 5-FU a 5-FUMP. De manera importante, la uracil fosforibosiltransferasa bacteriana (UPRT) es funcionalmente equivalente a orotato fosforibosiltransferasa (OPRT) o sintasa uridina-5'-monofosfato de las células de mamífero. Estas enzimas median la conversión de 5-fluorouracilo (5-FU) (un análogo fluorinado de uracilo) a 5-fluorouridina 5' monofosfato (5-FUMP). La 5- fluorouridina 5' monofosfato se convierte posteriormente a 5-FdUDP y FdUTP por medio del mamífero de novo de la ruta pirimidina. Cada 5-FdUTP es un inhibidor irreversible de la sintasa timidilato (Thy-A) y resulta en la estarvación del dTTP. Se acepta ampliamente que esta conversión es una de las rutas requisito para alcanzar los efectos citotóxicos de 5-fluorouracilo y que el uracil fosforibosiltransferasa bacteriano de origen bacteriano es capaz de convertir el 5- fluorouracilo en el mismo metabolito activo como lo hace el orotato fosforibosiltransferasa de mamífero. En la ausencia de la actividad UPRTasa o la actividad OPRTasa, el 5-FU, originado de la desaminación del 5-FC mediante la CDasa, no se transforma en 5-FUMP citotóxico (Jund and Lacroute, 1970, J. Bacteriol. 102, 607-615).

27

[0033] Como se describe aquí, la descripción proporciona polinucleótidos que codifican los polipéptidos y los polipéptidos que comprenden una fusión de un primer polipéptido que tiene actividad desaminasa citosina y un segundo polipéptido que comprende actividad UPRT o OPRT. Tales construcciones de fusión son útiles para convertir uracilo, o un derivado del mismo, en un análogo de monofosfato, y en particular 5-FU en 5-FUMP.

[0034] Como se describirá en más detalle adelante, la descripción se basa, por lo menos en parte, en la generación y expresión de polipéptidos novedosos que catalizan la conversión de 5-FC a 5-FU. En una realización, se proporcionan polipéptidos novedosos que se han construido por ingeniería mejoran las propiedades catalíticas para la conversión de 5-FC a 5-FU. Tales polipéptidos incluyen variantes que se han alterado para incluir sustituciones de aminoácido en los residuos específicos. Aunque estas variantes se describirán en más detalle adelante, se entiende que los polipéptidos de la descripción pueden contener uno o más aminoácidos modificados. La presencia de los aminoácidos modificados puede ser ventajosa en, por ejemplo, (a) incrementar la vida útil del polipéptido in vivo, (b) reducir o incrementar la antigenicidad del polipéptido, y (c) incrementar la estabilidad de almacenamiento del polipéptido. Se modifican los aminoácidos, por ejemplo, co-traduccionalmente o post-traduccionalmente durante la producción recombinante (por ejemplo, glucosilación ligada a N en los motivos N--X- -S/T durante expresión en células de mamífero) o se modifica por medios sintéticos. De acuerdo con lo anterior, una proteína

"mutante", "variante" o "modificada", enzima, polinucleótido, gen, o célula, significa una proteína, enzima, polinucleótido, gen, o célula, que se ha alterado o derivado, o es en alguna forma diferente o cambiado, de una proteína, enzima, polinucleótido, gen, o célula progenitora. Una proteína o enzima mutante o modificada usualmente, aunque no necesariamente, se expresa a partir de un polinucleótido o gen mutante.

[0035] Una "mutación" significa cualquier proceso o mecanismo que resulta en una proteína, enzima, polinucleótido, gen, o célula mutante. Esto incluye cualquier mutación en el cual se altera una proteína, enzima, polinucleótido, o secuencia de gen, y cualquier cambio detectable en una célula que surge de tal una mutación. Típicamente, ocurre una mutación en un polinucleótido o secuencia de gen, mediante mutaciones de punto, eliminaciones, o inserciones de los únicos o múltiples residuos de nucleótido. Una mutación incluye alteraciones de polinucleótido que surgen dentro de una región que codifica la proteína de un gen así como también alteraciones en regiones por fuera de una secuencia que codifica la proteína, tal como, pero no limitado a, secuencias promotoras o reguladoras. Una mutación en un gen puede ser "inactivo", es decir, no se refleja en una alteración de aminoácido luego de la expresión, que conduce a una variante "conservadora de la secuencia" del en. Esto surge generalmente cuando un aminoácido corresponde a más de un codón.

[0036] Los ejemplos no limitantes de un aminoácido modificado incluyen un aminoácido glucosilado, un aminoácido sulfatado, un

aminoácido prenilado (por ejemplo, farnesilado, geranilgeranilado), un aminoácido acetilado, un aminoácido acilado, un aminoácido pegilado, un aminoácido biotinilado, un aminoácido carboxilatado, un aminoácido fosforilado, y similares.

5 Las referencias adecuadas para guiar a un experto en la modificación de aminoácidos se encuentran a lo largo de la literatura. Los protocolos de ejemplo se encuentran en Walker (1998) Protein Protocols on CD-ROM (Humana Press, Towata, N.J.).

10 [0037] Se describen aquí los métodos recombinantes para producir, aislar, y utilizar los polipéptidos CD modificados y los polinucleótidos de la descripción. En adición a la producción recombinante, los polipéptidos se pueden producir mediante síntesis de péptido directa utilizando técnicas de fase sólida
15 (por ejemplo, Stewart et al. (1969) Solid-Phase Peptide Synthesis (WH Freeman Co, San Francisco); y Merrifield (1963) J. Am. Chem. Soc. 85: 2149-2154). Se puede desarrollar síntesis de péptido utilizando las técnicas del manual o mediante automatización. Se puede lograr síntesis automática, por ejemplo, utilizando Applied
20 Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City, Calif.) de acuerdo con las instrucciones proporcionadas por el fabricante.

[0038] Por vía de ilustración, las secuencias de ácido nucleico
25 que codifican las UPRTasas de *E. coli* (Anderson et al., 1992, Eur. J. Biochem 204, 51-56), de *Lactococcus lactis* (Martinussen and Hammer, 1994, J. Bacteriol. 176, 6457-6463), de *Mycobacterium bovis* (Kim et al., 1997, Biochem MoI. Biol. Int 41, 1117-1124) y

30

de *Bacillus subtilis* (Martinussen et al., 1995, J. Bacteriol. 177, 271-274) se pueden utilizar en el contexto de la descripción. Sin embargo, el uso de una UPRTasa de levadura, y en particular que se codifica mediante el gen FUR1 de *S. cerevisiae*,
 5 cuya secuencia se describe en Kern et al. (1990, Gene 88, 149-157) se incorpora aquí como referencia. Por medio de la indicación, las secuencias de los genes y aquellas de las UPRTasas correspondientes se pueden encontrar en la literatura y los bancos de datos especializados (SWISSPROT, EMBL, Genbank
 10 Medline, etc.).

[0039] Una "proteína" o "polipéptido", cuyos términos se utilizan intercambiabilmente aquí, comprende una o más cadenas de bloques de construcción química denominados aminoácidos que se ligan
 15 mediante enlaces químicos denominados enlaces de péptido. Una "enzima" significa cualquier sustancia, preferiblemente compuesta completamente o mayormente de la proteína, que se cataliza o promueve, más o menos específicamente, una o más reacciones químicas o bioquímicas. El término "enzima" también se puede
 20 referir a un polinucleótido catalítico (por ejemplo, ARN o ADN). Una proteína "nativa" o "tipo natural", enzima, polinucleótido, gen, o célula, significa una proteína, enzima, polinucleótido, gen, o célula que ocurre en la naturaleza.

25 [0040] De acuerdo con lo anterior, en varias realizaciones, se proporcionan polipéptidos recombinantes o aislados que comprenden la SEQ ID NO: 2 con una mutación seleccionada del grupo que consiste de A23L, V108I, I140L y cualquier combinación de los

mismos. En todavía otra realización, el polipéptido comprende la secuencia como se establece en la SEQ ID NO: 4 e incluye hasta 50, 25, 10, o 5 sustituciones de aminoácido conservadoras que excluyen los residuos: (a) 23, 108 y 140. En general, los
 5 polipéptidos proporcionados aquí exhiben actividad desaminasa citosina útil para convertir 5-FC a 5-FU.

[0041] La descripción también proporciona un polipéptido que comprende una secuencia que es por lo menos 80%, 90%, 95%, 98% o
 10 99% idénticas a la SEQ ID NO: 4, en donde el polipéptido tiene una leucina en la posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y en donde el polipéptido tiene actividad desaminasa citosina. En todavía otra realización, el polipéptido comprende la secuencia como se establece en la SEQ ID
 15 NO: 4.

[0042] La descripción proporciona adicionalmente construcciones de fusión que comprenden uno cualquiera de los polipéptidos anteriores ligados operablemente a un uracil
 20 fosforibosiltransferasa (UPRT) o un orotato fosforibosiltransferasa (OPRT). En una realización, la construcción de fusión comprende un primer polipéptido que comprende SEQ ID NO: 2 con una mutación seleccionada del grupo que consiste de A23L, V108I, I140L y cualquier combinación de los
 25 mismos ligado a un segundo polipéptido que tiene actividad UPRT u OPRT. En todavía otra realización, la construcción de fusión comprende un primer polipéptido que tiene la secuencia como se establece en la SEQ ID NO: 4 e incluye hasta 50, 25, 10, o 5

sustituciones de aminoácido conservadoras que excluyen los residuos: 23, 108 y 140 en donde el primer polipéptido se liga a un segundo polipéptido que tiene actividad UPRT o OPRT. La descripción también proporciona una construcción de fusión que
5 comprende un primer polipéptido que tiene actividad desaminasa citosina y que comprende una secuencia que es por lo menos 80%, 90%, 95%, 98% o 99% idéntica a la SEQ ID NO: 4, en donde el polipéptido tiene una leucina en la posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y en donde
10 el primer polipéptido se liga a un segundo polipéptido que comprende actividad UPRT o OPRT. En una realización, los polipéptidos que tienen actividad UPRT o OPRT comprende una secuencia como se establece en la SEQ ID NO: 8 y 10, respectivamente, o sus variantes. En todavía otra realización, un
15 primer polipéptido que comprende actividad desaminasa citosina se liga a un polipéptido que tiene actividad UPRT o OPRT, en donde los polipéptidos se separan mediante un ligador de péptido. En otra realización, la construcción de fusión comprende una secuencia seleccionada del grupo que consiste de de la SEQ ID NO:
20 11, 12 o 13.

[0043] El término "ligado operablemente" o "asociado operablemente" se refiere a un ligado funcional o asociación entre una secuencia reguladora y el polinucleótido regulado por
25 la secuencia reguladora o entre dos polipéptidos distintos o los polinucleótidos que codifican tales polipéptidos.

[0044] Una construcción de fusión que comprende un polipéptido que tiene actividad CD y un polipéptido que comprende actividad UPRT o OPRT se puede construir por ingeniería por contener un sitio de división para ayudar en la recuperación de la proteína u otro grupo funcional ligador. Típicamente un ligador será un ligador del grupo funcional de péptido. La longitud del grupo funcional ligador se selecciona para optimizar la actividad biológica del polipéptido CD y el polipéptido UPRT o OPRT y se puede determinar empíricamente sin la debida experimentación. El grupo funcional ligador debe ser lo suficientemente larga y suficientemente flexible para permitir a un sustrato interactuar con el polipéptido CD y un sustrato con el polipéptido UPRT o OPRT. Un grupo funcional ligador es un péptido entre aproximadamente uno y 30 residuos de aminoácido en longitud, típicamente entre aproximadamente dos y 15 residuos de aminoácido. Ejemplos de grupos funcionales ligadores son --Gly--Gly--, GGGGS (SEQ ID NO:22), (GGGGS)_N=(SEQ ID NO:23), (SGGGG)_N=(SEQ ID NO: 24), GKSSGSGSESKS (SEQ ID NO:25), GSTSGSGKSSEGKG (SEQ ID NO:26), GSTSGSGKSSEGSGSTKG (SEQ ID NO: 27), GSTSGSGKPGSGEGSTKG (SEQ ID NO:28), o EGKSSGSGSESKEF (SEQ ID NO:29). Los grupos funcionales ligadores se describen, por ejemplo, en Huston et al., Proc. Nat ' l Acad. Sci 85:5879, 1988; Whitlow et al., Protein Engineering 6:989, 1993; y Newton et al., Biochemistry 35:545, 1996. Otros ligadores de péptido adecuados son aquellos descritos en las Patentes Estadounidenses Nos. 4,751,180 y 4,935,233, que se incorporan aquí como referencia. Una secuencia de ADN que codifica un ligador de péptido deseado se puede insertar entre, y en el mismo marco de lectura como, un

polinucleótido que codifica un polipéptido CD seguido por un polipéptido UPRT o OPRT, utilizando cualquier técnica convencional adecuada. Por ejemplo, un oligonucleótido sintetizado químicamente que codifica el ligador que se puede
5 ligar entre dos polinucleótidos codificantes. En realizaciones particulares, un polipéptido de fusión comprende de dos a cuatro dominios separados (por ejemplo, un dominio CD y un UPRT o OPRT) se separan mediante ligadores de péptido.

10 [0045] "sustitución de aminoácido conservadora" o, simplemente, "variaciones conservadores" de una secuencia particular se refiere al reemplazo de un aminoácido, o serie de aminoácidos, con secuencias de aminoácido esencialmente idénticas. Un experto reconocerá que sustituciones, eliminaciones o adiciones
15 individuales que alteran, agregan o eliminan un aminoácido único o un porcentaje de aminoácidos en una secuencia codificante resulta en "variaciones conservadoras" en donde las alteraciones resultan en la eliminación de un aminoácido, adición de un aminoácido, o sustitución de un aminoácido con un aminoácido
20 químicamente similar.

[0046] Las tablas de sustitución conservadora proporcionan aminoácidos funcionalmente similares que son bien conocidos en la técnica. Por ejemplo, un grupo de sustitución conservadora
25 incluye Alanina (A), Serina (S), y Treonina (T). Otro grupo de sustitución conservadora incluye ácido Aspártico (D) y ácido Glutámico (E). Otro grupo de sustitución conservadora incluye Asparagina (N) y Glutamina (Q). Todavía otro grupo de sustitución

35

conservadora incluye Arginina (R) y Lisina (K). Otro grupo de sustitución conservadora incluye Isoleucina, (I) Leucina (L), Metionina (M), y Valina (V). Otro grupo de sustitución conservadora incluye Fenilalanina (F), Tirosina (Y), y Triptofan
5 (W).

[0047] Así, "sustituciones de aminoácido conservadoras" de una secuencia de polipéptido listada (por ejemplo, SEQ ID NOs: 2, 4, 6, 8, 10, 11, 12, 13 etc.) incluyen las secuencias de un
10 porcentaje, típicamente menos de 10%, de los aminoácido de la secuencia de polipéptido, con un aminoácido conservativamente seleccionado del mismo grupo de sustitución conservadora. De acuerdo con lo anterior, una variación conservativamente sustituida de un polipéptido de la descripción puede contener
15 100, 75, 50, 25, o 10 sustituciones con una variación conservativamente sustituida del mismo grupo de sustitución conservadora.

[0048] La "actividad" de una enzima es una medición de su
20 capacidad para catalizar una reacción, es decir, para "funcionar", y se puede expresar como un índice en el cual se produce el producto de la reacción. Por ejemplo, la actividad de enzima se puede representar como la cantidad de producto producido por unidad de tiempo o por unidad de enzima (por
25 ejemplo, concentración o peso), o en términos de afinidad o las constantes de disociación. Como se utiliza intercambiamente aquí una "actividad desaminasa citosina", "actividad biológica de desaminasa citosina " o "actividad funcional de una desaminasa

36

citosa ", se refiere a una actividad ejercida mediante una proteína desaminasa citosina, o polipéptido de la descripción, en un sustrato desaminasa citosina, como se determina in vivo, o in vitro, de acuerdo con técnicas estándar. Los ensayos para medir la actividad desaminasa citosina se conocen en la técnica. Por ejemplo, se puede medir una actividad desaminasa citosina al determinar el índice de conversión de 5-FC a 5-FU o citosina a uracilo. La detección de 5-FC, 5-FU, citosina y uracilo se puede desarrollar mediante cromatografía y otros métodos conocidos en la técnica.

[0049] Un experto en la técnica apreciará que muchas variaciones conservadoras de las construcciones de ácido nucleico que se describen producen una construcción funcionalmente idéntica. Por ejemplo, como se discutió anteriormente, propio a la degeneración del código genético, "sustituciones inactivas" (es decir, sustituciones en una secuencia de ácido nucleico que no resultan en una alteración en un polipéptido codificado) son una característica implicada de cada secuencia de ácido nucleico que codifica un aminoácido. Se apreciará por aquellos expertos en la técnica que debido a la degeneración del código genético, una multitud de secuencias de nucleótido que codifican los polipéptidos desaminasa citosina modificados de la descripción se pueden producir, algunas de las cuales llevan identidad sustancial en las secuencias de ácido nucleico explícitamente descritas aquí. Por ejemplo, los codones AGA, AGG, CGA, CGC, CGG, y CGU codificarán la arginina de aminoácido. Así, en cada posición en los nucleótidos en la descripción en donde una

37

arginina se especifica mediante un codón, el codón se puede alterar en cualquiera de los codones correspondientes descritos anteriormente sin alterar el polipéptido codificado. Se entiende que en U en una secuencia de ARN corresponde a T en una secuencia de ADN.

[0050] De forma similar, "sustituciones de aminoácido conservadoras," en uno o unos pocos aminoácidos en una secuencia de aminoácido se sustituyen con diferentes aminoácidos con propiedades altamente similares, también se identifican fácilmente cuando es altamente similar a una construcción descrita. Tales variaciones conservadoras de cada secuencia descrita son una característica de los polipéptidos proporcionados aquí.

[0051] "Variantes conservadoras" son proteínas o enzimas en las que se ha cambiado un residuo de aminoácido dado sin alterar la conformación general y la función de la proteína o enzima que incluye, pero no se limita a, el reemplazo de un aminoácido con uno que tiene propiedades similares, que incluye carácter polar o no polar, tamaño, forma y carga. Los aminoácidos diferentes a aquellos indicados como conservados pueden diferir en una proteína o enzima ya que el porcentaje de la secuencia de proteína o aminoácido similarmente entre cualquiera de las dos proteínas de función similar puede variar y puede ser, por ejemplo, por lo menos 30%, por lo menos 50%, por lo menos 70%, por lo menos 80%, o por lo menos 90%, como se determina de acuerdo con un esquema de alineación. Como se denomina aquí,

"similitud de secuencia" significa el grado en el cual se relacionan las secuencias nucleótido o proteína. El grado de similitud entre las dos secuencias con base en el porcentaje de identidad de secuencia y/o conservación. "Identidad de secuencia" aquí significa el grado en el cual las dos secuencias de nucleótido o aminoácido son invariantes. "Alineación de secuencia" significa el proceso de alineación de dos o más secuencias para alcanzar niveles máximos de identidad (y, en el caso de las secuencias de aminoácido, la conservación) para el propósito de evaluar el grado de similitud. Numerosos métodos para las secuencias de alineación y para evaluar la similitud/identidad se conocen en la técnica tal como, por ejemplo, el método Cluster, en donde la similitud se basa en el algoritmo MEGALIGN, así como también BLASTN, BLASTP, y FASTA (Lipman and Pearson, 1985; Pearson and Lipman, 1988). Cuando se utilizan todos estos programas, las configuraciones preferidas son aquellas que resultan en la similitud de secuencia mayor. Por ejemplo, la "identidad" o "porcentaje de identidad" con respecto a un par particular de las secuencias de aminoácido alineadas puede referir al porcentaje de identidad de la secuencia de aminoácido que se obtiene mediante análisis ClustalW (versión W 1.8 disponible de European Bioinformatics Institute, Cambridge, UK), que cuenta el número de coincidencias idénticas en la alineación y divide tal número de coincidencias idénticas por la longitud mayor de (i) de las secuencias alineadas, y (ii) 96, y utilizando los siguientes parámetros de falta ClustalW para alcanzar alineaciones en forma de par lentas/exactas—Penalidad de Espacio Abierto: 10; Penalidad de extensión de Espacio: 0.10;

matriz de peso de la Proteína: Gonnet series; matriz de peso de ADN: IUB; alineaciones en forma de par Lentas/Rápidas Toggle = Alineación LENTA o COMPLETA.

5 [0052] Las dos secuencias se "alinean óptimamente" cuando ellas se alinean para la clasificación de similitud utilizando una matriz de sustitución de aminoácido definida (por ejemplo, BLOSUM62), penalidad de existencia de espacio y penalidad de extensión de espacio con el fin de arribar en la clasificación mayor posible para el par de las secuencias. Las matrices de sustitución de aminoácido y su uso en cuantificar la similitud entre las dos secuencias son bien conocidos en la técnica y se describen, por ejemplo, en Dayhoff et al. (1978) "A model of evolutionary change in proteins" en "Atlas of Protein Sequence and Structure", Vol. 5, Suppl. 3 (ed. M. O. Dayhoff), pp. 345-352. Natl. Biomed. Res. Found., Washington, D. C. y Henikoff et al. (1992) Proc. Nat ' l. Acad. Sci. USA 89: 10915-10919. La matriz BLOSUM62 (FIG. 10) se utiliza frecuentemente como una matriz de sustitución de clasificación de falta en los protocolos de alineación de secuencia tal como Gapped BLAST 2.0. La penalidad de existencia de espacio se impone a la introducción de un espacio de aminoácido único en una de las secuencias alineadas, y la penalidad de extensión de espacio se impone para cada posición de aminoácido vacía adicional insertada dentro del espacio ya abierto. La alineación se define por las posiciones de los aminoácidos de cada secuencia en la que la alineación empieza y finaliza, y opcionalmente mediante la inserción de un espacio o múltiples espacios en una o ambas secuencias con el fin de

arribar en la clasificación mayor posible. Mientras que se puede llevar a cabo manualmente la alineación y clasificación óptima, el proceso se facilita mediante el uso de un algoritmo de alineación implementado por computador, por ejemplo, separación
5 BLAST 2.0, descrito en Altschul et al. (1997) Nucl. Acids Res. 25: 3389-3402, y se encuentra disponible al público en el sitio web del Centro Nacional para la Información Biotecnológica (NCBI) (www.ncbi.nlm.nih.gov). Las alineaciones óptimas, que incluyen múltiples alineaciones, se pueden preparar utilizando, por
10 ejemplo, PSI-BLAST, disponible a través del sitio web NCBI y descrito por Altschul et al. (1997) Nucl. Acids Res. 25:3389-3402.

[0053] Con respecto a una secuencia de aminoácido que se alinea
15 óptimamente con una secuencia de referencia, un residuo de aminoácido "corresponde a" la posición en la secuencia de referencia con la cual el residuo se empareja en la alineación. La "posición" se denota por un número que identifica secuencialmente cada aminoácido en la secuencia de referencia con
20 base en su posición con relación al Terminal N. Propio a la eliminaciones, inserciones, truncaciones, fusiones, etc., que se puede tomar en cuenta cuando se determina una alineación óptima, en general el número de residuo de aminoácido en una secuencia de prueba como se determina mediante conteo simple del terminal N no
25 será necesariamente el mismo como el número de su posición correspondiente en la secuencia de referencia. Por ejemplo, en un caso en donde exista eliminación en una secuencia de prueba alineada, no existirá aminoácido que corresponda a una posición

u1

en la secuencia de referencia en el sitio de eliminación. En donde existe una inserción en la secuencia de referencia alineada, que se inserta no corresponderá a cualquier posición de aminoácido en la secuencia de referencia. En el caso de truncaciones o fusiones existe estiramiento de los aminoácidos en la secuencia de referencia o la secuencia alineada que no corresponde a cualquier aminoácido en la secuencia correspondiente.

10 [0054] Las modificaciones no conservadoras de un polipéptido particular son aquellas que sustituyen cualquier aminoácido no caracterizado como una sustitución conservadora. Por ejemplo, cualquier sustitución que cruza los enlaces de los seis grupos establecidos anteriormente. Estos incluyen sustituciones de los aminoácidos acídicos o básicos para los aminoácidos neutros (por ejemplo, Asp, Glu, Asn, o Gln para Val, Ile, Leu o Met), aminoácido aromático para los aminoácidos acídicos o básicos (por ejemplo, Phe, Tyr o Trp for Asp, Asn, Glu o Gln) o cualquier otra sustitución que no reemplaza un aminoácido con un aminoácido similar. Las cadenas laterales básicas incluyen lisina (K), arginina (R), histidina (H) ; las cadenas laterales acídicas incluyen ácido aspártico (D), ácido glutámico (E); cadenas laterales polares no cargadas que incluyen glicina (G), asparagina (N), glutamina (Q), serina (S), treonina (T), tirosina (Y), cisteína (C); cadenas laterales no polares que incluyen alanina (A), valina (V), leucina (L), isoleucina (I), prolina (P), fenilalanina (F), metionina (M), triptofan (W); cadenas laterales ramificadas beta que incluyen treonina (T), valina (V),

u2

isoleucina (I); cadenas laterales aromáticas que incluyen tirosina (Y), fenilalanina (F), triptofan (W), histidina (H).

[0055] De acuerdo con lo anterior, algunos residuos de aminoácido
5 en las posiciones específicas en un polipéptido se "excluyen" de las sustituciones de aminoácido conservadoras. Por ejemplo, la descripción proporciona un polipéptido que comprende una secuencia como se establece en la SEQ ID NO: 4. En algunas realizaciones, el polipéptido se puede alterar con sustituciones
10 de aminoácido conservadoras como se describió anteriormente, sin embargo, ciertos residuos se dejan deseablemente no sustituidos tal como los residuos en las posiciones 23, 108, y 140.

[0056] Una proteína, enzima, polinucleótido, gen, o célula,
15 "progenitora" es cualquier proteína, enzima, polinucleótido, gen, o célula, de la cual se deriva o se hace cualquier otra proteína, enzima, polinucleótido, gen, o célula, utilizando cualesquier métodos, herramientas o técnicas, y si o no el progenitor es en sí mismo nativo o mutante. Un polinucleótido progenitor o gen que
20 codifica una proteína o enzima progenitora.

[0057] Un polinucleótido, polipéptido, u otro componente se "aisla" cuando este se separa parcialmente o completamente de los componentes con los cuales este se asocia normalmente (otras
25 proteínas, ácidos nucleicos, células, reactivos sintéticos, etc.). Un ácido nucleico o polipéptido es "recombinante" cuando este es artificial o se construye por ingeniería, o se deriva de una proteína o ácido nucleico artificial o construido por

ingeniería. Por ejemplo, un polinucleótido que se inserta dentro de un vector o cualquier otra ubicación heteróloga, por ejemplo, en un genoma de un organismo recombinante, de tal manera que este no se asocia con las secuencias de nucleótido que flanquean normalmente el polinucleótido cuando este se encuentra en la naturaleza es un polinucleótido recombinante. Una proteína expresada in vitro o in vivo de un polinucleótido recombinante es un ejemplo de un polipéptido recombinante. De esta forma, una secuencia de polinucleótido que no aparece en la naturaleza, por ejemplo una variante de un gen de ocurrencia natural, es recombinante.

[0058] En otras realizaciones, los polinucleótidos aislados se proporcionan que codifican un polipéptido desaminasa citosina o construcción de fusión de la descripción. En un aspecto, la descripción proporciona un polinucleótido aislado o recombinante denominado aquí como "polinucleótidos CD optimizados", "polinucleótidos CD" o "polinucleótidos de fusión CD." Los términos "polinucleótido", "secuencia de nucleótido", y "molécula de ácido nucleico" se utilizan para referirse a un polímero de los nucleótidos (A, C, T, U, G, etc. o los análogos de nucleótido de ocurrencia natural o artificiales), por ejemplo, ADN o ARN, o una representación de los mismos, por ejemplo, una cadena de caracteres, etc., dependiendo del contexto relevante. Un polinucleótido dado o polinucleótido complementario se puede determinar de cualquier secuencia de nucleótido específica.

[0059] Una realización de la descripción pertenece a polinucleótidos aislados que codifican una desamina citosina un polipéptido desaminasa citosina mutante o sus porciones biológicamente activas. Como se utiliza aquí, el término

5 "molécula de ácido nucleico" o "polinucleótido" está destinado a incluir moléculas de ADN (por ejemplo, cADN o ADN genómico) y moléculas de ARN (por ejemplo, mARN) y análogos del ADN o ARN generados utilizando análogos de nucleótido. La molécula de ácido nucleico puede ser monocatenaria o bicatenaria.

10

[0060] En una realización, un polinucleótido CD optimizado o polinucleótido CD comprende formas aisladas, recombinantes o construidas por ingeniería de ácidos nucleicos de ocurrencia natural aislados del organismo, por ejemplo, una bacteria o cepa

15 de levadura. Los polinucleótidos CD de ejemplo incluyen aquellos que codifican los polipéptidos tipo natural establecidos en la SEQ ID NO: 2. En otra realización de la descripción, los polinucleótidos CD se producen por diversificación, por ejemplo, recombinación y/o mutación de uno o más polinucleótidos CD

20 recombinantes, aislados o de ocurrencia natural. Como se describe en más detalle en cualquier parte aquí, es posible generar polinucleótidos CD diversificados que codifican los polipéptidos CD con atributos funcionales superiores, por ejemplo, función catalítica incrementada, estabilidad incrementada, o nivel de

25 expresión mayor, que un polinucleótido CD utilizado como un sustrato o progenitor en el proceso de diversificación. Los polinucleótidos de ejemplo incluyen la SEQ ID NO: 3 y 5 y

US

aquellos que codifican los polipéptidos CD variantes de la descripción.

[0061] Los polinucleótidos de la descripción tienen una variedad de usos en, por ejemplo, producción recombinante (es decir, expresión) de los polipéptidos CD de la descripción y como sustratos para generación de diversidad adicional, por ejemplo, reacciones de recombinación o reacciones de mutación para producir nuevos y/o mejorados homólogos CD, y similares.

10

[0062] Es importante notar que ciertas utilidades específicas, sustanciales y creíbles de los polinucleótidos CD no requieren que el polinucleótido codifique un polipéptido con actividad CD sustancial o aún actividad CD variante. Por ejemplo, los polinucleótidos CD que no codifican las enzimas activas pueden ser fuentes valiosas de los polinucleótidos progenitores para uso en los procedimientos de diversificación para arribar a las variantes de polinucleótido de CD, o polinucleótidos no CD, con propiedades deseablemente funcionales (por ejemplo, kcat alto o kcat/Km, Kn bajo, alta estabilidad hacia el calor u otros factores ambientales, los altos índices de transcripción o traducción, resistencia a la división proteolítica, etc.).

20

[0063] Los polinucleótidos CD, que incluyen las secuencias de nucleótido que codifican los polipéptidos CD y sus variantes, fragmentos de polipéptidos CD, proteínas de fusión relacionadas, o sus equivalentes funcionales, se utilizan en moléculas de ADN recombinante que dirigen la expresión de los polipéptidos CD en

25

46

las células anfitrionas apropiadas, tal como las células bacterianas. Debido a la degeneración inherente del código genético, otras secuencias de ácido nucleico que codifican sustancialmente la misma o una secuencia de aminoácido funcionalmente equivalente también se puede utilizar para clonar y expresar los polinucleótidos CD. El término "célula anfitriona", como se utiliza aquí, incluye cualquier tipo de célula que es susceptible a la transformación con una construcción de ácido nucleico. El término "transformación" significa la introducción de un gen extraño (es decir, extrínseco o extracelular), una secuencia de ADN o ARN con una célula anfitriona, ya que la célula anfitriona expresará el gen introducido o la secuencia para producir una sustancia deseada, típicamente una proteína o enzima codificada por el gen o la secuencia introducida. El gen introducido o la secuencia pueden incluir las secuencias reguladoras o de control, tal como las secuencias de inicio, detención, promotoras, de señal, de secreción, u otras secuencias utilizadas por la maquinaria genética de la célula. Una célula anfitriona que recibe y se expresa en el ADN o ARN introducido se ha "transformado" y es un "transformante" o un "clon." El ADN o ARN introducido a una célula anfitriona puede venir de cualquier fuente, que incluye células de los mismos géneros o especies como la célula anfitriona, o células de una especie o género diferente.

[0064] Como se entenderá por parte de aquellos expertos en la técnica, puede ser ventajoso modificar una secuencia codificante para mejorar su expresión en un anfitrión particular. El código

UA

genético es redundante con 64 codones posibles, pero más organismos utilizan preferiblemente un subconjunto de estos codones. Los codones que se utilizan más frecuentemente en una especie se denominan codones óptimos, y aquellos que no se utilizan muy frecuentemente se clasifican como codones de uso raro o de bajo uso (ver, por ejemplo, Zhang et al. (1991) Gene 105:61-72). Se pueden sustituir codones para reflejar el uso del codón preferido del anfitrión, un proceso algunas veces denominado "optimización del codón" o "controlar el sesgo de codón de las especies."

[0065] Las secuencias codificantes optimizadas que contienen los codones preferidos por un anfitrión eucariótico o procariótico particular (ver también, Murray et al. (1989) Nucl. Acids Res. 17:477-508) se puede preparar, por ejemplo, para incrementar el índice de traducción o para producir los transcriptos de ARN recombinante que tiene propiedades deseables, tal como una vida útil mayor, cuando se compara con los transcriptos producidos de una secuencia no optimizada. Los codones de detención de traducción también se pueden modificar para reflejar la preferencia del anfitrión. Por ejemplo, los codones de detención preferidos para *S. cerevisiae* y los mamíferos son UAA y UGA, respectivamente. El codón de detención preferido para las plantas monocotiledóneas es UGA, mientras que los insectos y *E. coli* prefieren utilizar UAA como el codón de detención (Dalphin et al. (1996) Nucl. Acids Res. 24: 216-218).

[0066] El sesgo de uso de Codones se refiere a diferencias entre los organismos en la frecuencia de ocurrencia de codones en las secuencias de ADN que codifican la proteína (genes). Un codón es una serie de tres nucleótidos (tripletes) que codifican un residuo de aminoácido específico en una cadena de polipéptido. Debido a que existen cuatro nucleótidos en ADN, adenina (A), guanina (G), citosina (C) y timina (T), existen 64 tripletes posibles que codifican 20 aminoácidos, y tres codones de terminación de traducción (no codificantes). Debido a esta redundancia, todos pero dos aminoácidos se codifican por más de un triplete. Diferentes organismos frecuentemente muestran preferencias particulares para uno de los varios codones que codifican el mismo aminoácido. Como estas preferencias elevadas es un área mucho más debatida de evolución molecular.

[0067] Es de conocimiento general que las preferencias de codón reflejan un balance entre sesgos mutacionales y selección natural para la optimización traduccional. Los codones óptimos en los microorganismos de crecimiento rápido, como *Escherichia coli* o *Saccharomyces cerevisiae* (levadura de panadería), refleja la composición de su grupo de tARN genómico respectivo. Esto es que los codones óptimos ayudan a alcanzar índices de traducción más rápidos y alta exactitud. Como un resultado de estos factores, se espera que la selección traduccional sea más fuerte en genes altamente expresados, como es de hecho el caso para los organismos mencionados anteriormente. En otros organismos que no muestran índices de alto crecimiento o que presentan genomas pequeños, la optimización del uso de codón está normalmente

ua

ausente, y las preferencias de codón se determinan mediante sesgos mutacionales características vistas en este genoma particular. Ejemplos de estos son *Homo sapiens* (humano) y *Helicobacter pylori*. Los organismos que muestran un nivel
 5 intermedio de la optimización de uso de codón incluyen *Drosophila melanogaster* (mosca de los frutos), *Caenorhabditis elegans* (gusano nemátodo) o *Arabidopsis thaliana* (oruga).

[0068] El término "secuencias optimizadas de codón" generalmente
 10 se refiere a las secuencias de nucleótido que se han optimizado para una especie anfitriona particular al reemplazar cualesquier codones que tiene una frecuencia de uso de menos de aproximadamente 20%. Las secuencias de nucleótido que se han optimizado para la expresión en una especie anfitriona dada
 15 mediante la eliminación de las secuencias de poliadenilación espuria, eliminación de la señales de división de exón/intrón, eliminación de las repeticiones similares a transposón y/o la optimización del contenido de GC en adición a la optimización de codón se denomina aquí como "secuencias que mejoran la
 20 expresión."

[0069] Tabla 1: Uso de codón humano y la preferencia de codón. Para cada codón, la tabla exhibe la frecuencia del uso de cada codón (por miles) en las regiones codificantes humanas (primera
 25 columna) y la frecuencia relativa de cada codón entre los codones sinónimos (segunda columna).

La Tabla de Uso del Codón Humano															
Gly	GGG	17.08	0.23	Arg	AGG	12.09	0.22	Trp	TGG	14.74	1.00	Arg	CGG	10.40	0.19
Gly	GGA	19.31	0.26	Arg	AGA	11.73	0.21	End	TGA	2.64	0.61	Arg	CGA	5.63	0.10
Gly	GGT	13.66	0.18	Ser	AGT	10.18	0.14	Cys	TGT	9.99	0.42	Arg	CGT	5.16	0.09
Gly	GGC	24.94	0.33	Ser	AGC	18.54	0.25	Cys	TGC	13.86	0.58	Arg	CGC	10.82	0.19
Glu	GAG	38.82	0.59	Lys	AAG	33.79	0.60	End	TAG	0.73	0.17	Gln	CAG	32.95	0.73
Glu	GAA	27.51	0.41	Lys	AAA	22.32	0.40	End	TAA	0.95	0.22	Gln	CAA	11.94	0.27
Asp	GAT	21.45	0.44	Asn	AAT	16.43	0.44	Tyr	TAT	11.80	0.42	His	CAT	9.56	0.41
Asp	GAC	27.06	0.56	Asn	AAC	21.30	0.56	Tyr	TAC	16.48	0.58	His	CAC	14.00	0.59
Val	GTG	28.60	0.48	Met	ATG	21.86	1.00	Leu	TTG	11.43	0.12	Leu	CTG	39.93	0.43
Val	GTA	6.09	0.10	Ile	ATA	6.05	0.14	Leu	TTA	5.55	0.06	Leu	CTA	6.42	0.07
Val	GTT	10.30	0.17	Ile	ATT	15.03	0.35	Phe	TIT	15.36	0.43	Leu	CTT	11.24	0.12
Val	GTC	15.01	0.25	Ile	ATC	22.47	0.52	Phe	TTC	20.72	0.57	Leu	CTC	19.14	0.20
Ala	GCG	7.27	0.10	Thr	ACG	6.80	0.12	Ser	TCG	4.38	0.06	Pro	CCG	7.02	0.11
Ala	GCA	15.50	0.22	Thr	ACA	15.04	0.27	Ser	TCA	10.96	0.15	Pro	CCA	17.11	0.27
Ala	GCT	20.23	0.28	Thr	ACT	13.24	0.23	Ser	TCT	13.51	0.18	Pro	CCT	18.03	0.29
Ala	GCC	28.43	0.40	Thr	ACC	21.52	0.38	Ser	TCC	17.37	0.23	Pro	CCC	20.51	0.33

[0070] De acuerdo con lo anterior, en algunas realizaciones de la descripción un polinucleótido comprende un codón de molécula para la traducción en una célula humana. Por ejemplo, SEQ ID NO: 3 y 5 comprende las secuencias que se han optimizado para producir desaminasa citosina en una célula anfitriona humana.

[0071] La descripción así proporciona un polinucleótido optimizado de codón humano que codifica un polipéptido que tiene actividad desaminasa citosina. Un polinucleótido optimizado de codón humano de la descripción (por ejemplo, SEQ ID NO:5) puede incluir adicionalmente mutaciones adicionales que resultan en sustitución de aminoácido conservadora y o actividad mejorada o estabilidad del polipéptido codificado. Por ejemplo, mutaciones en las posiciones 23, 108 y 140 del polipéptido que comprende SEQ

ID NO: 6 proporciona termoestabilidad incrementada. De acuerdo con lo anterior, la descripción proporciona un polinucleótido optimizado de codón humano que codifica un polipéptido que tiene actividad desaminasa citosina y termoestabilidad incrementada. En una realización, el polinucleótido de codón humano optimizado comprende una secuencia que codifica un polipéptido de la SEQ ID NO: 4, en donde los codones que codifican los aminoácidos del polipéptido se han optimizado para la expresión en una célula humana. En otra realización, un polinucleótido de la descripción comprende SEQ ID NO: 3 o 5.

[0072] En una realización, la descripción proporciona un polinucleótido que comprende una desaminasa citosina polinucleótido o un polinucleótido de codón optimizado o mutante ligado a un polinucleótido heterólogo que codifica un polipéptido que tiene actividad UPRT o OPRT.

[0073] Un polinucleótido de la descripción que comprende, por ejemplo, una secuencia que codifica el polipéptido de la SEQ ID NO: 4, 8, 10, 12, 14, 16, o 18; o que tiene la secuencia de nucleótido de la establecida en cualquier de la SEQ ID NOS:3, 5, 7, 9, 11, 13, 15, o 17, o una porción de las mismas, se puede aislar y generar utilizando técnicas de biología molecular estándar y la información de secuencia proporcionada aquí.

[0074] Un polinucleótido de la descripción se puede amplificar utilizando cADN, mRNA o alternativamente, ADN genómico, como una plantilla y los cebadores de oligonucleótido apropiados de

52

acuerdo con las técnicas de amplificación de PCR estándar. El ácido nucleico así amplificado se puede clonar en un vector apropiado y se caracteriza por un análisis de secuencia de ADN. Adicionalmente, los oligonucleótidos que corresponden a las

5 secuencias de nucleótido se pueden preparar mediante técnicas sintéticas estándar, por ejemplo, utilizando un sintetizador de ADN automático. En algunas realizaciones, una molécula de ácido nucleico aislada de la descripción comprende una molécula de ácido nucleico que es un complemento de una secuencia de

10 nucleótido que codifica un polipéptido establecida en cualquiera de la SEQ NOs: 2, 4, 6, 8, 10, 12, 14, 16, o 18, o que tiene la secuencia de nucleótido de la establecida en cualquiera de la SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, o 17.

15 [0075] En otra realización, un polinucleótido aislado de la descripción comprende una molécula de ácido nucleico que hibrida bajo condiciones exigentes en una molécula de ácido nucleico que consiste de la secuencia de nucleótido que codifica un polipéptido establecida en cualquiera de la SEQ NOs: 4, 12, 14,

20 16 o 18, o que consiste de la secuencia de nucleótido establecida en cualquiera de la SEQ ID NOs: 3, 5, 11, 13, 15, o 17. En otras realizaciones, el ácido nucleico es por lo menos 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, o 600 nucleótidos en longitud. Las moléculas de ácido nucleico son "hibridizables" uno

25 con el otro cuando por lo menos una cepa de un polinucleótido puede hibridar en otro polinucleótido bajo condiciones exigentes definidas. La exigencia de la hibridación se determina, por ejemplo, mediante (a) la temperatura en la cual se desarrolla la

hibridación y/o el lavado, y (b) la resistencia iónica y polaridad (por ejemplo, formamida) de la hibridación y las soluciones de lavado, así como también como otros parámetros. La hibridación requiere que los dos polinucleótidos contengan

5 secuencias sustancialmente complementarias; dependiendo de la exigencia de hibridación, sin embargo, se pueden tolerar desfases. Típicamente, la hibridación de las dos secuencias en alta exigencia (tal como, por ejemplo, en una solución acuosa de 0.5 X SSC a 65°C) requiere que las secuencias exhiban algún alto

10 grado de complementariedad sobre su secuencia completa. Las condiciones de exigencia intermedia (tal como, por ejemplo, una solución acuosa de 2 X SSC a 65°C) y baja exigencia (tal como, por ejemplo, una solución acuosa de 2 X SSC a 55°C), requiere correspondientemente menos complementariedad general entre las

15 secuencias hibridizantes (1 X SSC es 0.15 M de NaCl, 0.015 M Na de citrato). Las moléculas de ácido nucleico que hibridan incluyen aquellas que hibridan bajo condiciones exigentes adecuadas y que codifican los polipéptidos o enzimas que tienen la misma función, tal como la capacidad de catalizar la

20 conversión de 5-Fluorocitosina (5-FC) a 5-flurouracilo (5-FU), de la descripción. Adicionalmente, el término "hibrida bajo condiciones exigentes" está destinado a describir condiciones para la hibridación y lavado cuyas secuencias de nucleótido por lo menos 30%, 40%, 50%, o 60% homólogas a cada uno típicamente

25 permanecen hibridadas una con la otra. Preferiblemente, las condiciones son de tal manera que las secuencias por lo menos aproximadamente 70%, más preferiblemente por lo menos aproximadamente 80%, aún más preferiblemente por lo menos

aproximadamente 85% o 90% homólogas una con la otra típicamente permanecen hibridadas una con la otra. En algunos casos, una molécula de ácido nucleico aislado de la descripción híbrida bajo condiciones exigentes a una secuencia de ácido nucleico que
5 codifica un polipéptido establecido en cualquiera de la SEQ ID NOs: 4, 12, 14, 16, o 18, o que tiene la secuencia de nucleótido establecida en cualquiera de la SEQ ID NOs: 3, 5, 11, 13, 15, o 17. Como se utiliza aquí, una molécula "de ocurrencia natural" de ácido nucleico se refiere a una molécula de ADN o de ARN que
10 tiene una secuencia de nucleótido que ocurre en la naturaleza (por ejemplo, que codifica una proteína natural).

[0076] El artesano experto apreciará que cambios se pueden introducir mediante mutación dentro de las secuencias de
15 nucleótido de cualquiera de los polinucleótidos que codifican un polipéptido establecido en cualquiera de la SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, o 18, o que tiene la secuencia de nucleótido establecida en cualquiera de la SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, o 17, que conduce por lo tanto a cambios en la secuencia
20 de aminoácido de las proteínas codificadas. En algunos casos la alteración conducirá a la función alterada del polipéptido. En otros casos el cambio no alterará la capacidad funcional del polipéptido codificado. En general, las sustituciones que no alteran la función de un polipéptido incluyen sustituciones de
25 nucleótido que conducen a sustituciones de aminoácido en los residuos de aminoácido "no esenciales". Un residuo de aminoácido "no esencial" es un residuo que se puede alterar a partir de la secuencia progenitora sin alterar la actividad biológica del

SS

polipéptido resultante, por ejemplo, catalizando la conversión de 5-FC a 5-FU.

[0077] También se contemplan aquellas situaciones que son deseables para alterar la actividad de un polipéptido progenitor de tal manera que el polipéptido tiene nueva o actividad incrementada en un sustrato particular o estabilidad incrementada o degradación reducida. Se entiende que estas sustituciones de aminoácido no constituirán generalmente sustituciones "conservadoras". En su lugar, estas sustituciones constituyen sustituciones no conservadoras introducidas en una secuencia con el fin de obtener una o mejorada actividad. Por ejemplo, la SEQ ID NO:1 proporciona la secuencia de ácido nucleico progenitora para desaminasa citosina *S. cerevisiae* (SEQ ID NO: 2 proporciona el polipéptido correspondiente). La SEQ ID NO: 3 proporciona la secuencia de ácido nucleico de una secuencia mutante que incluye las sustituciones de aminoácido que imparten estabilidad incrementada al polipéptido. De acuerdo con lo anterior, la molécula de ácido nucleico que codifica la secuencia de aminoácido de la SEQ ID NO: 2 proporciona una molécula de ácido nucleico "progenitora" de la cual se pueden hacer mutaciones para obtener una molécula de ácido nucleico que codifica un polipéptido modificado que incluye sustituciones de aminoácido.

[0078] También se entiende que un polipéptido modificado puede constituir un polipéptido "progenitor" del cual se pueden hacer sustituciones adicionales. De acuerdo con lo anterior, un polipéptido progenitor, y una molécula de ácido nucleico que

codifica un polipéptido progenitor, incluye polipéptidos modificados y no justo secuencias "tipo natural". Por ejemplo, el polinucleótido de la SEQ ID NO: 5 es un polinucleótido modificado con respecto a la SEQ ID NO: 1 (es decir, el polinucleótido "progenitor"). De forma similar, el polinucleótido de la SEQ ID NO: 3 es un polinucleótido modificado con respecto a la SEQ ID NO: 5. De acuerdo con lo anterior, la SEQ ID NO: 5 es la secuencia progenitora de la SEQ ID NO: 3.

10 [0079] Los métodos mutacionales para generar la diversidad incluyen, por ejemplo, mutagenia dirigida a sitio (Ling et al. (1997) "Approaches to DNA mutagenesis: an overview" Anal Biochem. 254(2): 157-178; Dale et al. (1996) "Oligonucleótido-directed random mutagenesis using the fosforothioate method" Methods Mol. Biol. 57:369-374; Smith (1985) "In vitro mutagenesis" Ann. Rev. Genet. 19:423-462; Botstein & Shortle (1985) "Strategies and applications of in vitro mutagenesis" Science 229:1193-1201; Carter (1986) "Site-directed mutagenesis" Biochem. J. 237:1-7; y Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" in Nucleic Acids & Molecular Biology (Eckstein, F. y Lilley, D. M. J. eds., Springer Verlag, Berlin)); mutagenia utilizando plantillas que contienen uracilo (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without phenotypic selection" Proc. Natl. Acad. Sci. USA 82:488-492; Kunkel et al. 20 (1987) "Rapid and efficient site-specific mutagenesis without phenotypic selection" Methods in Enzymol. 154, 367-382; y Bass et al. (1988) "Mutant Trp repressors with new DNA-binding specificities" Science 242:240-245); mutagenia dirigida a

oligonucleótido (Methods in Enzymol. 100: 468-500 (1983); Methods
in Enzymol. 154: 329-350 (1987); Zoller & Smith (1982)
"Oligonucleotide-directed mutagenesis using M13-derived vectors:
an efficient and general procedure for the production of point
5 mutations in any DNA fragment" Nucleic Acids Res. 10:6487-6500;
Zoller & Smith (1983) "Oligonucleotide-directed mutagenesis of
DNA fragments cloned into M13 vectors" Methods in Enzymol.
100:468-500; y Zoller & Smith (1987) "Oligonucleotide-directed
mutagenesis: a simple method using two oligonucleotide primers y
10 a single- stranded DNA template" Methods in Enzymol. 154:329-
350); mutagenia de ADN modificada por fosforotioate (Taylor et
al. (1985) "The use of fosforothioate-modified DNA in restriction
enzyme reactions to prepare nicked DNA" Nucl. Acids Res. 13:
8749-8764; Taylor et al. (1985) "The rapid generation of
15 oligonucleótido-directed mutations at high frequency using
fosforothioate-modified DNA" Nucl. Acids Res. 13: 8765-8787;
Nakamaye & Eckstein (1986) "Inhibition of restriction
endonuclease Nci I cleavage by fosforothioate groups y its
application to oligonucleotide-directed mutagenesis" Nucl. Acids
20 Res. 14: 9679-9698; Sayers et al. (1988) "Y-T Exonucleases in
fosforothioate-based oligonucleotide-directed mutagenesis" Nucl.
Acids Res. 16:791-802; y Sayers et al. (1988) "Strand specific
cleavage of fosforothioate-containing DNA by reaction with
restriction endonucleases in the presence of ethidium bromide"
25 Nucl. Acids Res. 16: 803-814); mutagenia utilizando ADN dúplex
espaciado (Kramer et al. (1984) "The gapped duplex DNA approach
to oligonucleotide- directed mutation construction" Nucl. Acids
Res. 12: 9441-9456; Kramer & Fritz (1987) Methods in Enzymol.

"Oligonucleotide-directed construction of mutations via gapped duplex DNA" 154:350-367; Kramer et al. (1988) "Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations" Nucl. Acids Res. 16: 7207; y Fritz et al. (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro" Nucl. Acids Res. 16: 6987-6999).

[0080] Los métodos adicionales adecuados incluyen reparación de desfases de punto (Kramer et al. (1984) "Point Mismatch Repair" Cell 38:879-887), mutagenia utilizando cepas anfitrionas deficientes de reparación (Carter et al. (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" Nucl. Acids Res. 13: 4431-4443; y Carter (1987) "Improved oligonucleotide-directed mutagenesis using M13 vectors" Methods in Enzymol. 154: 382-403), mutagenia de eliminación (Eghtedarzadeh & Henikoff (1986) "Use of oligonucleotides to generate large deletions" Nucl. Acids Res. 14: 5115), purificación de restricción y selección de restricción (Wells et al. (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" Phil. Trans. R. Soc. Lond. A 317: 415-423), mutagenia mediante síntesis de gen total (Nambiar et al. (1984) "Total synthesis y cloning of a gene coding for the ribonuclease S protein" Science 223: 1299-1301; Sakamar and Khorana (1988) "Total synthesis and expression of a gene for the α -subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin) " Nucl. Acids Res. 14: 6361-6372; Wells et al. (1985) "Cassette mutagenesis: an efficient method for

generation of multiple mutations at defined sites" Gene 34:315-323; y Grundstrom et al. (1985) "Oligonucleotide-directed mutagenesis by microscale Nshot-gunN gene synthesis" Nucl. Acids Res. 13: 3305-3316); reparación de rompimiento bicatenario (Mandecki (1986); Arnold (1993) "Protein engineering for unusual environments" Current Opinion in Biotechnology 4:450-455; y "Oligonucleotide-directed double-strand break repair in plasmides of Escherichia coli: a method for site-specific mutagenesis" Proc. Natl. Acad. Sci. USA, 83:7177-7181). Los detalles adicionales en muchos de los métodos anteriores se pueden encontrar en Methods in Enzymology Volumen 154, que también describe controles útiles para la solución de problemas con varios métodos de mutagenia.

[0081] Los detalles adicionales con respecto a los varios métodos de generación diversificados se pueden encontrar siguiendo las patentes Estadounidenses, publicaciones PCT, y publicaciones EP0: la patente Estadounidense No. 5,605,793 de Stemmer (Feb. 25, 1997), "Methods for In vitro Recombination;" la patente Estadounidense No. 5,811,238 de Stemmer et al. (Sep. 22, 1998) methods for Generating Polinucleotides having Desired Characteristics by Iterative Selection and Recombination;" la patente Estadounidense No. 5,830,721 de Stemmer et al. (Nov. 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly;" la patente Estadounidense No. 5,834,252 de Stemmer, et al. (Nov. 10, 1998) "End-Complementary Polymerase Reaction;" la patente Estadounidense No. 5,837,458 de Minshull, et al. (Nov. 17, 1998), methods and Compositions for Cellular y Metabolic Engineering;"

la WO 95/22625, Stemmer and Cramer, "Mutagenesis by Random Fragmentation and Reassembly;" la WO 96/33207 por Stemmer and Lipschutz "End Complementary Polymerase Chain Reaction;" la WO 97/20078 por Stemmer and Cramer methods for Generating
5 Polinucleotides having Desired Characteristics by Iterative Selection and Recombination;" la WO 97/35966 por Minshull and Stemmer, methods and Compositions for Cellular and Metabolic Engineering;" la WO 99/41402 por Punnonen et al. "Targeting of Genetic Vaccine Vectors;" la WO 99/41383 por Punnonen et al.
10 "Antigen Library Immunization;" la WO 99/41369 por Punnonen et al. "Genetic Vaccine Vector Engineering;" la WO 99/41368 por Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines;" la EP 752008 por Stemmer and Cramer, "DNA Mutagenesis by Random Fragmentation and Reassembly;" la EP
15 0932670 por Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination;" la WO 99/23107 por Stemmer et al., "Modification of Virus Tropism and Host Range by Viral Genome Shuffling;" la WO 99/21979 por Apt et al., "Human Papillomavirus Vectors;" la WO 98/31837 por del Cardayre et al. "Evolution of
20 Whole Cells and Organisms by Recursive Sequence Recombination;" la WO 98/27230 por Patten and Stemmer, methods and Compositions for Polypeptide Engineering;" la WO 98/13487 por Stemmer et al., methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection;" la WO 00/00632, methods for Generating
25 Highly Diverse Libraries;" la WO 00/09679, methods for Obtaining in vitro Recombined Polinucleotide Sequence Banks and Resulting Sequences;" la WO 98/42832 por Arnold et al., "Recombination of Polinucleotide Sequences Using Random or Defined Primers;" la WO

99/29902 por Arnold et al., "Method for Creating Polinucleotide and Polipéptido Sequences;" la WO 98/41653 por Vind, "An in vitro Method for Construction of a DNA Library;" la WO 98/41622 por Borchert et al., "Method for Constructing a Library Using DNA Shuffling;" la WO 98/42727 por Pati and Zarling, "Sequence Alterations using Homologous Recombination;" la WO 00/18906 por Patten et al., "Shuffling of Codon- Altered Genes;" la WO 00/04190 por del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Recombination;" la WO 00/42561 por Cramer et al., "Oligonucleotide Mediated Nucleic Acid Recombination;" la WO 00/42559 por Selifonov and Stemmer methods of Populating Data Structures for Use in Evolutionary Simulations;" la WO 00/42560 por Selifonov et al., methods for Making Character Strings, Polinucleotides & Polypeptides Having Desired Characteristics;" la WO 01/23401 por Welch et al., "Use of Codon-Varied Oligonucleotide Synthesis for Synthetic Shuffling;" y la WO 01/64864 "Single-Stranded Nucleic Acid Template-Mediated Recombination and Nucleic Acid Fragment Isolation" por Affholter.

[0082] También se proporcionan construcciones recombinantes que comprenden una o más de las secuencias de ácido nucleico como se describió anteriormente ampliamente. Las construcciones comprenden un vector, tal como, un plásmido, un cósmido, un fago, un virus, un cromosoma artificial bacteriano (BAC), un cromosoma artificial de levadura (YAC), un vector vírico o similares, en los que un polinucleótido de la descripción se ha insertado, en una orientación inversa o hacia adelante. Grandes números de vectores y promotores adecuados se conocen por aquellos expertos

en la técnica, y están comercialmente disponibles. En una realización, el vector vírico es un vector retrovírico.

[0083] El términos "vector", "construcción de vector" y "vector de expresión" significa el vehículo por el cual una secuencia de ADN o de ARN (por ejemplo un gen extraño) se puede introducir dentro de una célula anfitriona, con el fin de transformar la expresión de promotor y anfitrión (por ejemplo la transcripción y traducción) de la secuencia introducida. Los vectores típicamente comprenden el ADN de un agente transmisible, en el cual el ADN extraño que codifica una proteína se inserta mediante tecnología de enzima de restricción. Un tipo común de vector es un "plásmido", que generalmente es una molécula auto contenida de un ADN bicatenario que puede aceptar fácilmente ADN adicional (extraño) que se puede introducir fácilmente dentro de una célula anfitriona adecuada. Un gran número de vectores, que incluyen plásmido y vectores fúngicos, se han descrito para la replicación y/o expresión en una variedad de anfitriones eucarióticos y procarióticos. Los ejemplos no limitantes incluyen los plásmidos pKK (Clonetech), los plásmidos pUC, los plásmidos pET (Novagen, Inc., Madison, Wis.), los plásmidos pRSET o pREP (Invitrogen, San Diego, Calif.), o plásmidos pMAL (New England Biolabs, Beverly, Mass.), y muchas células anfitrionas apropiadas, utilizando los métodos descritos o citados aquí o de otra forma conocidos por aquellos expertos en la técnica relevante. Los vectores de clonación recombinante frecuentemente incluirán uno o más sistemas de replicación para clonación o expresión, uno o más

marcadores para selección en el anfitrión, por ejemplo, resistencia antibiótica, y uno o más casetes de expresión.

[0084] Los términos "expresa" y "expresión" significan que
5 permite o que origina la información en un gen o una secuencia de
ADN para llegar a manifestar, por ejemplo producir una proteína
al activar las funciones celulares involucradas en la
transcripción y traducción de un gen correspondiente o una
secuencia de ADN. Una secuencia de ADN se expresa en o mediante
10 una célula para formar un "producto de expresión" tal como una
proteína. El producto de expresión en sí mismo, por ejemplo la
proteína resultante, también puede ser dicha que se "expresa"
mediante la célula. Un polinucleótido o polipéptido se expresa
recombinantemente, por ejemplo, cuando este se expresa o se
15 produce en una célula anfitriona extraña bajo el control de un
promotor nativo o extraño, o en una célula anfitriona nativa bajo
el control de un promotor extraño.

[0085] Los polinucleótidos proporcionados aquí se pueden
20 incorporar dentro de una cualquier variedad de vectores de
expresión adecuados para la expresión de un polipéptido. Los
vectores adecuados incluyen secuencias de ADN sintéticas,
cromosómicas y no cromosómicas, por ejemplo, derivados de SV40;
plásmidos bacterianos; ADN de fago; baculovirus; plásmidos de
25 levadura; vectores derivados de combinaciones de plásmidos y ADN
de fago, ADN vírico tal como vacunas, adenovirus, virus de la
viruela aviar, seudorrabias, adenovirus, adenovirus asociados ,
retrovirus y muchos otros. Cualquier vector que transluce el

material genético dentro de una célula, y, si se desea se puede utilizar la replicación, que es replicable y viable en el anfitrión relevante. Estos vectores de ácido nucleico se pueden suministrar utilizando un sistema de suministro no vírico tal como se describe en P. Midoux et al. Brit J Pharm 157: 166-178 2009 o Kamimura & Liu J AAPS 10:589-595 (2008) o un sistema vírico como se describe adelante.

[0086] En otro aspecto, el polinucleótido que codifica una desaminasa citosina o un mutante del mismo se suministra con un suministro de sistema de gen que es un vector derivado viral o vírico. El vector vírico se puede replicar o no replicar, y puede ser un vector adenovírico, un vector de sarampión, un vector de herpes, un vector retrovírico (que incluye un vector lentivírico), un vector rabdovírico tal como un vector vírico de Estomatitis Vesicular, un vector de reovirus, un Vector del virus Seneca Valley, un vector poxvirus (que incluye vectores derivados de vacunas o viruela de animal), un vector parvovirus (que incluye un vector AAV), un vector alfavirus u otro vector vírico conocido por un experto en la técnica (ver también, por ejemplo, Concepts in Genetic Medicine, ed. Boro Dropulic and Barrie Carter, Wiley, 2008, Hoboken, NJ.; The Development of Human Gene Therapy, ed. Theodore Friedmann, Cold Springs Harbor Laboratory Press, Cold springs Harbor, New York, 1999; Gene and Cell Therapy, ed. Nancy Smyth Templeton, Marcel Dekker Inc., New York, New York, 2000 and Gene Therapy: Therapeutic Mechanism and Strategies, ed. Nancy Smyth Templeton and Danilo D Lasic, Marcel

Dekker, Inc., New York, New York, 2000; cuyas descripciones se incorporan aquí como referencia).

[0087] En una realización, el vector vírico puede ser un vector
5 retrovírico de replicación competente capaz de infectar solo
células de replicación de mamífero. Los retrovirus se han
clasificado en varias formas pero la nomenclatura se ha
estandarizado en la última década (ver ICTVdB - La Base de Datos
de Virus Universal, v 4 en la Red Amplia Mundial (www) en
10 ncbi.nlm.nih.gov/ICTVdb/ICTVdB/ y el texto "Retroviruses" Eds
Coffin, Hughs and Varmus, Cold Spring Harbor Press 1997; cuya
descripción se incorpora aquí como referencia). El vector
retrovírico de replicación competente puede comprender un
Ortoretrovirus o más típicamente un vector retrovirus gama. En un
15 aspecto, un vector retrovírico de replicación competente
comprende un sitio de entrada ribosómica interna (IRES) 5' en el
polinucleótido que codifica una desaminasa citosina. En una
realización, el polinucleótido que codifica una desaminasa
citosina es 3' en un polinucleótido ENV de un vector retrovírico.

20

[0088] De acuerdo con lo anterior, en otras realizaciones, se
proporcionan vectores que incluyen un polinucleótido de la
descripción. En otras realizaciones, se proporcionan células
anfitrionas transfectadas con una molécula de ácido nucleico de
25 la descripción, o un vector que incluye un polinucleótido de la
descripción. Las células anfitrionas incluyen células
eucarióticas tal como células de levadura, células de insectos, o

células de animal. Las células anfitrionas también incluyen células procarióticas tal como células bacterianas.

[0089] Como se discutió previamente, los textos generales que describen técnicas de biología molecular útiles aquí, que incluyen el uso de vectores, promotores y muchos otros tópicos relevantes, incluyen Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology Volumen 152, (Academic Press, Inc., San Diego, Calif.) ("Berger"); Sambrook et al., Molecular Cloning--A Laboratory Manual, 2d ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y., 1989 ("Sambrook") y Current Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (complementado a lo largo de 1999) ("Ausubel"). Ejemplos de protocolos suficientes para dirigir las personas expertas a través de métodos de amplificación in vitro, que incluyen la reacción de cadena polimerasa (PCR), la reacción de cadena ligasa (LCR), amplificación de Q β -replicasa y otras técnicas mediadas por polimerasa de ARN (por ejemplo, NASBA), por ejemplo, para la producción de los ácidos nucleicos homólogos de la descripción se encuentran en Berger, Sambrook, and Ausubel, así como también en Mullis et al. (1987) la patente Estadounidense No. 4,683,202; Innis et al., eds. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press Inc. San Diego, Calif.) ("Innis"); Arnheim & Levinson (Oct. 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3: 81-94; Kwoh et al. (1989) Proc. Natl. Acad. Sci. USA 86: 1173; Guatelli et al. (1990) Proc. Nat'l. Acad. Sci.

67

USA 87: 1874; Lomell et al. (1989) J. Clin. Chem 35: 1826; Landegren et al. (1988) Science 241: 1077-1080; Van Brunt (1990) Biotechnology 8: 291-294; Wu and Wallace (1989) Gene 4:560; Barringer et al. (1990) Gene 89:117; y Sooknanan and Malek (1995) 5 Biotechnology 13: 563-564. Los métodos mejorados para la clonación in Vitro que amplifican los ácidos nucleicos se describen en Wallace et al., la patente Estadounidense No. 5,426,039. Los métodos mejorados para amplificar los ácidos nucleicos grandes mediante PCR se resumen en Cheng et al. (1994) 10 Nature 369: 684-685 y las referencias citadas aquí, en las que se generan amplicones PCR de hasta 40 kb. Un experto apreciará que esencialmente cualquier ARN se puede convertir en un ADN bicatenario adecuado la digestión de restricción, expansión y secuenciamiento PCR utilizando transcriptasa inversa y una 15 polimerasa. Ver, por ejemplo, Ausubel, Sambrook and Berger, todos supra.

[0090] También se proporcionan las células anfitrionas 20 construidas por ingeniería que se transducen (transforman o transfectan) con un vector proporcionado aquí (por ejemplo, un vector de clonación o un vector de expresión), así como también la producción de polipéptidos de la descripción mediante técnicas recombinantes. El vector puede ser, por ejemplo, un plásmido, una partícula vírica, un fago, etc. Las células anfitrionas 25 construidas por ingeniería se pueden cultivar en medios de nutriente convencionales modificados según sea apropiado para promotores activantes, transformantes de selección, o amplificar un polinucleótido codificante. Las condiciones de cultivo, tal

como temperatura, pH y similares, son aquellas previamente utilizadas con la célula anfitriona seleccionada para expresión, y serán evidentes para aquellos expertos en la técnica y en las referencias citadas allí, que incluyen, por ejemplo, Sambrook, Ausubel and Berger, así como también por ejemplo, Freshney (1994) Culture of Animal Cells: A Manual of Basic Technique, 3rd ed. (Wiley-Liss, New York) y las referencias citadas aquí.

[0091] Se pueden emplear vectores para transformar un anfitrión apropiado para permitir al anfitrión expresar una proteína o polipéptido. Ejemplos de anfitriones de expresión apropiados incluyen: células bacterianas, tal como *E. coli*, *B. subtilis*, *Streptomyces*, y *Salmonella typhimurium*; células fúngicas, tal como *Saccharomyces cerevisiae*, *Pichia pastons*, y *Neurospora crassa*; células de insecto tal como *Drosophila* y *Spodoptera frugiperda*; células de mamífero tal como CHO, COS, BHK, HEK 293 o melanoma de Bowes; o células de planta o explantes, etc.

[0092] En los sistemas bacterianos, se puede seleccionar un número de vectores de expresión dependiendo del uso destinado para el polipéptido desaminasa citosina. Por ejemplo, cuando grandes cantidades de polipéptido CD o sus fragmentos se necesitan para la producción comercial o para la inducción de anticuerpos, pueden ser deseables vectores que dirigen alto nivel de expresión de las proteínas de fusión que son fácilmente purificadas. Tales vectores incluyen, pero no se limitan a, clonación *E. coli* multifuncional y vectores de expresión tal como BLUESCRIPT (Stratagene), en el cual la secuencia codificante de

polipéptido CD se puede ligar dentro del vector en estructura con secuencias para el terminal amino Met y los 7 residuos posteriores de beta-galactosidasa ya que se produce una proteína híbrida; los vectores pIN (Van Heeke & Schuster (1989) J. Biol. Chem. 264: 5503-5509); vectores pET (Novagen, Madison Wis.); y similares.

[0093] De forma similar, en el *Saccharomyces cerevisiae* de levadura un número de vectores que contienen promotores inducibles constitutivos o inducibles tal como factor alfa, alcohol oxidasa y PGH se pueden utilizar para la producción de los polipéptidos CD de la descripción. Para revisiones, ver Ausubel (supra) y Grant et al. (1987) Methods in Enzymology 153:516-544.

[0094] La descripción también proporciona vectores retrovíricos de replicación competente que tiene estabilidad incrementada con relación a los vectores retrovíricos anteriores. Tal estabilidad incrementada durante la infección y replicación es importante para el tratamiento de trastornos proliferativos de célula. La combinación de la eficiencia de transducción, la estabilidad del transgen y la selectividad objetivo se proporciona por el retrovirus de replicación competente. Las composiciones y métodos proporcionan estabilidad de inserto y mantienen la actividad de transcripción del transgen y la viabilidad translacional del polipéptido codificado.

20

[0095] Los retrovirus se pueden transmitir horizontalmente y verticalmente. La transmisión infecciosa eficiente de retrovirus requiera la expresión en la célula objetivo de receptores que reconocen específicamente las proteínas de envoltura víricas, mientras que los virus pueden utilizar rutas no específicas, independientes del receptor de entrada de baja eficiencia. Adicionalmente, el tipo de célula objetivo puede ser capaz de soportar todas las etapas del ciclo de replicación después que virus se ha unido y penetrado. La transmisión vertical ocurre cuando el genoma vírico se integra en la línea germinal del anfitrión. El provirus luego se pasará de generación a generación cuando este es un gen celular. Por lo tanto los provirus endógenos se llegan a establecer que frecuentemente permanecen latentes, pero se pueden llegar a activar cuando el anfitrión se expone a agentes apropiados.

[0096] Como se mencionó anteriormente, El intermedio de ADN integrado se refiere a provirus. Antes de la terapia de gen o los sistemas de suministro de gen utilizan métodos y retrovirus que requieren la transcripción del provirus y el ensamble dentro del virus infeccioso aunque en la presencia de un virus auxiliar apropiado o en una línea celular que contiene las secuencias apropiadas permite la encapsidación sin producción coincidente de un virus auxiliar contaminante. Como se describe adelante, no se requiere un virus auxiliar para la producción de los retrovirus recombinantes de la descripción, ya que las secuencias para encapsidación se proporcionan en el genoma proporcionando así un

21

vector retrovítico de replicación competente para el suministro de gen o la terapia.

[0097] UN genoma retrovítico útil en los métodos y composiciones de la descripción comprenden un ADN provítico que tiene por lo menos tres genes: el gag, el pol, y el env, estos genes se pueden flanquear por uno o dos repeticiones de Terminal largo (LTR), o en el provirus se flanquean por dos repeticiones de terminal largo (LTR) y las secuencias que contienen las secuencias que actúan cis tal como psi. El gen gag codifica las proteínas de estructura interna (matriz, cápsido, y nucleocápsido); el gen pol codifica la polimerasa de ADN dirigida por ARN (transcriptasa inversa), proteasa e integrasa; y el gen env codifica las glucoproteínas de envoltura víricas. Los LTR 5' y/o 3' sirven para promover la transcripción y poliadenilación de los ARN de virión. El LTR contiene todas las otras secuencias que actúan cis necesarias para la replicación vírica. Los lentivirus tienen genes adicionales que incluyen vif, vpr, tat, rev, vpu, nef, y vpx (en HIV-I, HIV-2 y/o SIV).

20

[0098] Adyacentes al LTR 5' están las secuencias necesarias para la transcripción inversa del genoma (el sitio de unión de cebador tRNA) y para la encapsidación eficiente del ARN vírico dentro de las partículas (el sitio Psi). Si se pierden las secuencias necesarias para encapsidación (o empaque del ARN retrovítico dentro del virión infeccioso) del genoma vírico, el resultado es un defecto cis que evita la encapsidación del ARN vírico genómico. Este tipo de vector modificado es que se ha utilizado

72

típicamente en sistemas de suministro de gen anteriores (es decir, sistemas que carecen de elementos que se requieren para la encapsidación del virión).

5 [0099] En una realización, la descripción proporciona un retrovirus recombinante capaz de infectar una célula no dividida, una célula dividida, o una célula que tiene un trastorno proliferativo de la célula. El retrovirus recombinante de replicación competente de la descripción comprende una secuencia
10 de polinucleótido que codifica un GAG vírico, un POL vírico, un ENV vírico, un polinucleótido heterólogo precedido por un sitio interno de entrada al ribosoma (IRES) encapsulado dentro de un virión.

15 [00100] La secuencia de ácido nucleico heterólogo se liga operablemente a un IRES. Como se utiliza aquí, el término secuencias o transgen de ácido nucleico "heterólogo" se refiere a (i) una secuencia que no existe normalmente en un retrovirus tipo natural, (ii) una secuencia que se origina de una especie
20 extraña, o (iii) si de la misma especie, esta se puede modificar sustancialmente de su forma original. Alternativamente, una secuencia de ácido nucleico no cambiada que no se expresa normalmente en una célula es una secuencia de ácido nucleico heterólogo. En una realización específica, el polinucleótido
25 heterólogo es un polipéptido de la descripción que tiene actividad desaminasa citosina.

[00101] La frase célula "no dividida" se refiere a una célula que no va a través de mitosis. Las células no divididas se pueden bloquear en cualquier punto en el ciclo celular, (por ejemplo, G0ZG1, G1/s, G2 /M), tanto como la célula no se divide
5 activamente. Para la infección ex vivo, se puede tratar una célula dividida para bloquear la división celular mediante técnicas estándar utilizadas por aquellos expertos en la técnica, que incluyen, irradiación, tratamiento con afidocolina, suero de inanición, e inhibición por contacto. Sin embargo, se debe
10 entender que la infección ex vivo se desarrolla frecuentemente sin bloquear las células debido a que muchas células ya se han disminuido (por ejemplo, células madre). Por ejemplo, un vector lentivirus recombinante es capaz de infectar cualquier célula no dividida, independiente del mecanismo utilizado para bloquear la
15 división celular o el punto en el ciclo celular en el cual se bloquea la célula. Ejemplos de células no divididas preexistentes en el cuerpo incluyen células neuronales, de músculo, de hígado, de piel, de corazón, de pulmón, y células de médula ósea, y sus derivados. Para dividir las células se pueden utilizar vectores
20 retrovíricos de onco-vectores.

[00102] Célula "dividida" significa una célula que experimenta mitosis activa, o meiosis. Tales células divididas incluyen células madre, células de piel (por ejemplo, fibroblastos y
25 queratinocitos), gametos, y otras células divididas conocidas en la técnica. Son de interés particular y se abarcan por el término célula dividida las células que tienen trastornos proliferativos de célula, tal como células neoplásicas. El término "trastorno

m

proliferativo de la célula" se refiere a una afección caracterizada por un número anormal de células. La afección puede incluir el crecimiento celular hipertrófico (la multiplicación continúa de células que resultan en un crecimiento ascendente de una población celular dentro de un tejido) e hipotrófico (una carencia o deficiencia de células dentro de un tejido) o un influjo o migración excesiva de las células dentro de un área del cuerpo. Las poblaciones celulares no se transforman necesariamente, células tumorigénicas o malignas, pero pueden incluir células normales muy bien. Los trastornos proliferativos de célula incluyen trastornos asociados con un crecimiento ascendente de los tejidos conectivos, tal como varias afecciones fibróticas, que incluyen escleroderma, artritis y cirrosis del hígado. Los trastornos proliferativos de célula incluyen trastornos neoplásicos tal como carcinomas de cabeza y cuello. Los carcinomas de cabeza y cuello incluyen, por ejemplo, carcinoma de la boca, esófago, garganta, laringe, glándula tiroides, lengua, labios, glándulas salivares, nariz, senos paranasales, nasofaringe, bóveda nasal superior y tumores del seno, esteseoneuroblastoma, cáncer de célula escamosa, melanoma maligno, carcinoma no diferenciado sinonasal (SNUC), neoplasia de cerebro (que incluye glioblastomas) o neoplasia de sangre. También se incluyen carcinomas de los nodos linfáticos regionales que incluyen nodos linfáticos cervicales, nodos linfáticos prelaríngeos, nodos linfáticos yuxtaesofágicos pulmonares y nodos linfáticos submandibulares (Harrison's Principles of Internal Medicine (eds., Isselbacher, et al., McGraw-Hill, Inc., 13ava Edición, ppl850-1853, 1994). Otros tipos de cáncer, incluyen,

pero no se limitan a, cáncer de pulmón, cáncer de colon-recto, cáncer de mama, cáncer de próstata, cáncer del tracto urinario, linfoma de cáncer uterino, cáncer oral, cáncer pancreático, leucemia, melanoma, cáncer de estómago, cáncer de piel y cáncer
5 de ovario.

[00103] En una realización, el polinucleótido heterólogo dentro del vector comprende una desaminasa citosina que se ha optimizado para la expresión en una célula humana. En una realización
10 adicional, la desaminasa citosina comprende una secuencia que ha sido codón humano optimizado y comprende mutaciones que incrementan la estabilidad de la desaminasa citosina (por ejemplo, degradación reducida o termoestabilidad incrementada) comparado con una desaminasa citosina tipo natural. En todavía
15 otra realización, el polinucleótido heterólogo codifica una construcción de fusión que comprende una desaminasa citosina (codón humano optimizado o no optimizado, mutado o no mutado) ligado operablemente a un polinucleótido que codifica un polipéptido que tiene actividad UPRT o OPRT. En otra realización,
20 el polinucleótido heterólogo comprende un polinucleótido CD o construcción de fusión de la descripción (por ejemplo, SEQ ID NO:3, 5, 11, 13, 15, o 17).

[00104] En otra realización, el vector retrovílico de replicación
25 competente puede comprender un polinucleótido heterólogo que codifica un polipéptido que comprende una desaminasa citosina (como se describe aquí) y puede comprender adicionalmente un

polinucleótido que comprende una molécula de miARN o siARN ligado a un promotor específico de tejido o de tipo celular.

[00105] El término "secuencia de ácido nucleico reguladora" se refiere colectivamente a las secuencias promotoras, las señales de poliadenilación, las secuencias de terminación de transcripción, los dominios reguladores en la dirección 5', orígenes de replicación, mejoradores y similares, que proporcionan colectivamente la replicación, la transcripción y traducción de una secuencia codificante en una célula receptora. No todas estas secuencias de control necesitan siempre estar presentes tanto como la secuencia codificante seleccionada que es capaz de ser replicada, transcrita y traducida en una célula anfitriona apropiada. Un experto en la técnica puede identificar fácilmente la secuencia de ácido nucleico reguladora de la base de datos pública y los materiales. Adicionalmente, un experto en la técnica puede identificar una secuencia reguladora que es aplicable al uso destinado, por ejemplo, in vivo, ex vivo, o in vitro.

20

[00106] El término "región promotora" se utiliza aquí en su sentido ordinario para referirse a una región de nucleótido que comprende una secuencia reguladora de ADN, en donde la secuencia reguladora se deriva de un gen que es capaz de unirse a la polimerasa de ARN e iniciar la transcripción de una secuencia codificante en la dirección 3' (dirección 3'). La secuencia reguladora puede ser homóloga o heteróloga para la secuencia de gen deseada. Por ejemplo, se puede utilizar un amplio rango de

77

promotores, que incluye el promotor vírico o de mamífero como se describió anteriormente.

[00107] Un sitio interno de entrada al ribosomas ("IRES") se
5 refiere a un segmento del ácido nucleico que promueve la entrada
o la retención de una ribosoma durante la traducción de una
secuencia codificante usualmente 3' en el IRES. En algunas
realizaciones el IRES puede comprender un sitio de división
receptor/donante, sin embargo, los IRES preferidos carecen de un
10 sitio de división receptor/donante. Normalmente, la entrada de
las ribosomas dentro del ARN mensajero tiene lugar por medio del
casquete localizado en el extremo 5' de todos los mRNA
eucarióticos. Sin embargo, existen excepciones en su regla
universal. La ausencia de un casquete en algunos mRNA vírico
15 sugiere que la existencia de las estructuras alternativas permite
la entrada de ribosomas en un sitio interno de estos ARN. A la
fecha, un número de estas estructuras, designadas IRES en cuenta
de su función, se han identificado en la región no codificante 5'
de los mRNA víricos sin casquete, tal como aquellos, en
20 particular, de los picornavirus tal como el virus de
poliomielitis (Pelletier et al., 1988, Mol. Cell. Biol., 8, 1103-
1112) y el virus EMCV (virus de encéfalo-miocarditis (Jang et
al., J. Virol., 1988, 62, 2636-2643). La descripción proporciona
el uso de un IRES en el contexto de un vector retrovírico de
25 replicación competente.

[00108] La secuencia de ácido nucleico heterólogo está
típicamente bajo el control de las señales de promotor-mejorador

LTR víricas o un promotor interno, y las señales retenidas dentro del LTR retrovírico puede aún traer aproximadamente eficiente integración del vector dentro del genoma de la célula anfitriona. De acuerdo con lo anterior, los vectores retrovíricos recombina-
5 tes de la descripción, las secuencias deseadas, genes y/o fragmentos de gen se pueden insertar en varios sitios y bajo diferentes secuencias reguladoras. Por ejemplo, un sitio para inserción puede ser el sitio próximo mejorador/promotor vírico (es decir, el sitio de gen que conduce LTR 5'). Alternativamente, las
10 secuencias deseadas se pueden insertar dentro de un sitio distal (por ejemplo, la secuencia IRES 3' en el gen env) o en donde dos o más secuencias heterólogas están presentes una secuencia heteróloga puede estar bajo el control de una primera región reguladora y una segunda secuencia heteróloga bajo el control de
15 una segunda región reguladora. Otros sitios distales incluyen secuencias de promotor vírico, en donde la expresión de la secuencia o secuencias deseadas es a través de la división del cistrón de promotor próximo, un promotor heterólogo interno como SV40 o CMV, o se puede utilizar un sitio interno de entrada al
20 ribosoma (IRES).

[00109] En una realización, el genoma retrovírico de la descripción contiene un IRES que comprende un sitio de clonación para la inserción de una secuencia de polinucleótido deseada. En
25 una realización, el IRES se ubica en 3' del gen env en el vector retrovírico, pero en 5' en el ácido nucleico heterólogo deseado. De acuerdo con lo anterior, una secuencia de polinucleótido

heteróloga que codifica un polipéptido deseado se puede ligar operablemente al IRES.

[00110] En otra realización una secuencia de polinucleótido
5 objetivo se incluye como parte del vector retrovítico
recombinante de la descripción. La secuencia de polinucleótido
objetivo es un ligando objetivo (por ejemplo, hormonas de péptido
tal como heregulina, anticuerpos de cadena sencilla, un receptor
o un ligando para un receptor), un elemento regulador específico
10 de tipo celular o específico de tejido (por ejemplo, un promotor
o mejoradores específico de tipo celular o específico de tejido),
o una combinación de un ligando objetivo y un elemento regulador
específico de tipo celular/específico de tejido. Preferiblemente,
el ligando objetivo se liga operablemente a la proteína env del
15 retrovirus, creando una proteína env retrovítica quimérica. Las
proteínas de GAG vírico, POL vírico y ENV vírico se pueden
derivar de cualquier retrovirus adecuado (por ejemplo, MLV o
derivado de lentivirus). En otra realización, la proteína ENV
retrovítica no se deriva de retrovirus (por ejemplo, CMV o VSV).

20

[00111] Los retrovirus recombinantes de la descripción se
modifica por lo tanto genéticamente en tal una forma que el virus
se objetiva en un tipo celular particular (por ejemplo, células
de músculo liso, células hepáticas, células renales,
25 fibroblastos, queratinocitos, células de tallo mesenquimal,
células de médula ósea, condrocito, células epiteliales, células
intestinales, células neoplásicas, células de glioma, células
neuronales y otras conocidas en la técnica) de tal manera que el

80

genoma del ácido nucleico se suministra a una célula dividida
objetivo, o no dividida objetivo, o una célula objetivo que tiene
un trastorno proliferativo de la célula. El objetivo se puede
alcanzar en dos formas. La primera forma dirige el retrovirus a
5 una célula objetivo mediante la unión a células que tiene una
molécula en la superficie externa de la célula. Este método del
retrovirus objetivo utiliza la expresión de un ligando objetivo
en el recubrimiento del retrovirus para asistir en objetivar el
virus a las células o tejidos que tienen un receptor o molécula
10 de unión que interactúa con el ligando objetivo en la superficie
del retrovirus. Después de la infección de una célula por el
virus, el virus inyecta su ácido nucleico dentro de la célula y
el material genético de retrovirus se puede integrar dentro del
genoma de la célula anfitriona. El segundo método para objetivar
15 utiliza elementos reguladores específicos de tejido o de célula
para promover la expresión y la transcripción del genoma vírico en
una célula objetivo que utiliza activamente los elementos
reguladores, como se describe más completamente adelante. El
material genético de retrovirus transferido luego se transcribe y
20 se traduce dentro de las proteínas dentro de la célula
anfitriona. El elemento regulador objetivo típicamente se liga al
LTR 5' y/o 3', creando un LTR quimérico.

[00112] Al insertar una secuencia de ácido nucleico heterólogo de
25 interés dentro del vector vírico de la descripción, junto con
otro gen que codifica, por ejemplo, el ligando para un receptor
en una célula objetivo específica, el vector es ahora objetivo
específico. Los vectores víricos pueden hacer el objetivo

específico al adherir, por ejemplo, un azúcar, un glucolípido, o una proteína. Se puede llevar a cabo el objetivo al utilizando un anticuerpo para el vector vírico objetivo. Aquellos expertos en la técnica conocerán de, o pueden comprobar fácilmente, las
5 secuencias de polinucleótido objetivo que se pueden insertar dentro del genoma vírico o las proteínas que se pueden adherir a una envoltura vírica para permitir el suministro específico objetivo del vector vírico que contiene la secuencia de ácido nucleico de interés.

10

[00113] Así, la descripción incluye en una realización, una proteína env quimérica que comprende una proteína env retrovítica ligada operablemente a un polipéptido objetivo. El polipéptido objetivo puede ser una molécula de receptor específica de célula,
15 un ligando para un receptor específico de célula, un anticuerpo o fragmento de anticuerpo para un epítipo antigénico específico de célula o cualquier otro ligando identificado fácilmente en la técnica que es capaz de unir o interactuar con una célula objetivo. Ejemplos de los polipéptidos o moléculas objetivo
20 incluyen anticuerpos bivalentes utilizando biotina-estreptavidina como ligadores (Etienne-Julan et al., J. Of General Virol., 73, 3251-3255 (1992); Roux et al., Proc. Natl. Acad. Sci USA 86, 9079-9083 (1989)), virus recombinante que contiene en su envoltura una secuencia que codifica un anticuerpo de región
25 variable de cadena sencilla contra un hapteno (Russell et al., Nucleic Acids Research, 21, 1081-1085 (1993)), clonación de los ligandos de hormona de péptido dentro de la envoltura del retrovirus (Kasahara et al., Science, 266, 1373-1376 (1994)),

82

construcciones EPO/env quiméricas (Kasahara et al., 1994), anticuerpo de cadena sencilla contra el receptor de lipoproteína (LDL) de baja densidad en la envoltura MLV ecotrópico, que resulta en una infección específica de células HeLa que expresan el receptor LDL (Somia et al., Proc. Natl. Acad. Sci USA, 92, 7570-7574 (1995)), de forma similar el rango anfitrión de ALV se puede alterar mediante la incorporación de un ligando integrina, que permite al virus ahora cruzar especies para infectar específicamente las células de glioblastoma de rata (Valsesia-Wittmann et al., J. Virol. 68, 4609- 4619 (1994)), y Dornberg y colegas (Chu and Dornburg, J. Virol 69, 2659-2663 (1995)) han reportado objetivo específico de tejido del virus de necrosis de bazo (SNV), un retrovirus aviar, que utiliza envolturas que contienen anticuerpos de cadena sencilla dirigidos contra los marcadores de tumor.

[00114] La descripción proporciona un método para producir un retrovirus recombinante capaz de infectar una célula objetivo que comprende transfectar una célula anfitriona adecuada con lo siguiente: un vector que comprende una secuencia de polinucleótido que codifica un gag vírico, un pol vírico y un env vírico, en donde el vector contiene un sitio de clonación para la introducción de un gen heterólogo, ligado operablemente a una secuencia de ácido nucleico reguladora, y recuperar el virus recombinante.

[00115] El retrovirus y los métodos de la descripción proporcionan un retrovirus de replicación competente que no

requiere virus auxiliar o la secuencia de ácido nucleico adicional o proteínas con el fin de propagar y producir virión. Por ejemplo, las secuencias de ácido nucleico del retrovirus de la descripción codifican, por ejemplo, un antígeno específico de grupo y transcriptasa inversa, (y enzimas de integrasa y proteasa necesarias para la maduración y transcripción inversa), respectivamente, como se discutió anteriormente. El gag y pol vírico se puede derivar de un lentivirus, tal como HIV o un oncovirus tal como MoMLV. Adicionalmente, el genoma del ácido nucleico del retrovirus de la descripción incluye una secuencia que codifica una proteína de envoltura vírica (ENV). El gen env se puede derivar de cualesquier retrovirus. El env puede ser una proteína de envoltura anfotrófica que permite la traducción de las células de especie humana y otras especies, o puede ser una proteína de envoltura ecotrófica, que es capaz de translucir solo las células de ratón y rata. Adicionalmente, puede ser deseable objetivar el virus recombinante mediante el ligado de la proteína de envoltura con un anticuerpo o un ligando particular para objetivar un receptor de un tipo celular particular. Como se mencionó anteriormente, los vectores retrovíricos pueden ser objetivos específicos al insertar, por ejemplo, un glucolípido, o una proteína. El objetivo se lleva a cabo frecuentemente al utilizar un anticuerpo para objetivar el vector retrovírico en un antígeno en un tipo de célula particular (por ejemplo, un tipo de célula encontrado en cierto tejido, o un tipo de célula de cáncer). Aquellos expertos en la técnica conocerán, o pueden comprobar fácilmente sin la debida experimentación, los métodos específicos para alcanzar el suministro de un vector retrovírico

gm

en un objetivo específico. En una realización, el gen env se deriva de un no retrovirus (por ejemplo, CMV o VSV). Ejemplos de genes env derivados de retrovirus incluyen, pero no se limitan a: Virus de leucemia de murino de Moloney (MoMuLV), virus de sarcoma murino de Harvey (HaMuSV), virus de tumor mamario de murino (MuMTV), virus de leucemia de mono Gibbon (GaLV), virus de inmunodeficiencia humana (VIH) y virus de sarcoma de Rous (RSV). También se pueden utilizar otros genes env tal como virus de Estomatitis Vesicular (VSV) (Proteína G), envoltura de citomegalovirus (CMV), o virus de influenza hemaglutinina (HA).

[00116] En una realización, el genoma retroviral se deriva de un onco-retrovirus o retrovirus gama, y más particularmente un onco-retrovirus de mamífero o retrovirus gama. "Derivado" significa que la secuencia de polinucleótido progenitor es un oncovirus tipo natural que se ha modificado mediante la inserción o remoción de las secuencias de ocurrencia natural (por ejemplo, inserción de un IRES, inserción de un polinucleótido heterólogo que codifica, por ejemplo, un polipéptido que tiene actividad desaminasa citosina de la descripción y similares).

[00117] En otra realización, la descripción proporciona vectores retrovirales que se objetivan utilizando secuencias reguladoras. Las secuencias reguladoras específicas de tejido o específicas de célula (por ejemplo, promotores) se pueden utilizar para la expresión objetivo de las secuencias de gen en las poblaciones celulares específicas. Los promotores víricos y de mamífero adecuados para la descripción se describen en cualquier parte

aquí. De acuerdo con lo anterior, en una realización, la descripción proporciona un retrovirus que tiene elementos promotores específicos de tejido en el extremo 5' del genoma retroviral. Preferiblemente, los elementos/secuencias reguladoras específicas de tejido están en la región U3 del LTR del genoma retroviral, que incluye por ejemplo promotores específicos de tejido o específicos de célula y mejoradores para las células neoplásicas (por ejemplo, mejoradores y promotores específicos de célula de tumor), y promotores inducibles (por ejemplo, tetraciclina).

[00118] En algunas circunstancias, puede ser deseable regular la expresión. Por ejemplo, diferentes promotores víricos con estiramientos variados de la actividad se pueden utilizar dependiendo del nivel de expresión deseado. En las células de mamífero, el promotor temprano inmediato CMV se utiliza frecuentemente para proporcionar la fuerte activación transcripcional. Las versiones modificadas del promotor CMV que son menos potentes también se han utilizado cuando se desean niveles reducidos de la expresión del transgen. Cuando se desea la expresión de un transgen en las células hematopoyéticas, se pueden utilizar promotores retrovirales tal como los LTR de MLV o MMTV. Otros promotores víricos que se pueden utilizar incluyen SV40, RSV LTR, HIV-1 y HIV-2 LTR, promotores de adenovirus tal como de la región E1A, E2A, o MLP, LTR AAV, virus mosaico de coliflor, HSV-TK, y virus de sarcoma aviar.

[00119] De forma similar se pueden utilizar promotores selectivos o específicos de tejido para efectuar la transcripción en tejidos o células específicas con el fin de reducir la toxicidad potencial o los efectos indeseables para los tejidos no objetivo.

5 Por ejemplo, los promotores tal como el PSA, probasina, fosfatasa de ácido prostático o calicreína glandular específica de próstata (hK2) se puede utilizar para objetivar la expresión de gen en la próstata. Otros dominios de promotores/reguladores que se pueden utilizar se establecen en la Tabla 1.

10

[00120] En ciertas indicaciones, puede ser deseable activar la transcripción en momentos específicos después de la administración del vector de terapia de gen. Esto se puede hacer con tales promotores como aquellos que son regulables de hormona o

15 citoquina. Por ejemplo en las aplicaciones terapéuticas en donde la indicación es un tejido gonadal en donde los esteroides

específicos se producen o se enrután para, puede ser ventajoso el uso de promotores regulados por andrógeno o estrógeno. Tales

promotores que son regulables por hormona incluyen MMTV, MT-I, ecdysone y RuBisco. Se pueden utilizar otros promotores regulados

20 por hormona tal como aquellos que responden a las hormonas

tiroides, pituitaria y adrenal. Los promotores que responden a la proteína inflamatoria y citoquina que se pueden utilizar incluyen

K y T Kininogen (Kageyama et al., 1987), c-fos, TNF-alfa,

25 proteína C reactiva (Arcone et al., 1988), haptoglobina (Oliviero

et al., 1987), suero amiloide A2, alfa C/EBP, IL-1, IL-6 (PoIi

and Cortese, 1989), Complemento C3 (Wilson et al., 1990), IL-8,

ácido alfa-1 glucoproteína (Prowse and Baumann, 1988), alfa-1

antitripsina, lipoproteína lipasa (Zechner et al., 1988),
 angiotensinogena (Ron et al., 1990), fibrinógeno, c-jun
 (inducible por ésteres forbol, TNF-alfa, radiación UV, ácido
 retinoico, y peróxido de hidrógeno), colagenasa (inducido por
 5 ésteres forbol y ácido retinoico), metalotioneina (metal pesado y
 glucocorticoide inducible), Estromelisina (inducible por éster
 forbol, interleuquina-1 y EGF), alfa-2 macroglobulina y alfa-1
 antiquimotripsina. Los promotores específicos de tumor tal como
 osteocalcina, elementos que responde a hipoxia (HRE), MAGE-4,
 10 CEA, alfa-fetoproteína, GRP78/BiP y tirosinasa también se pueden
 utilizar para regular la expresión de gen en las células de
 tumor.

[00121] Adicionalmente, esta lista de promotores no se debe
 15 construir como que es exhaustiva o limitante, aquellos expertos
 en la técnica conocerán de otros promotores que se pueden
 utilizar en conjunto con los promotores y métodos descritos aquí.

TABLA 1 PROMOTORES ESPECÍFICOS DE TEJIDO

20	Tejido	Promotor
	Páncreas	Insulina Elastina Amilasa pdr-1 pdx-1 glucoquinasa
	Hígado	Albúmina PEPCK mejorador de HBV
25		α fetoproteína apolipoproteína C α -1 antitripsina vitelogenina, NF-AB Transtiretina

músculo esquelético quinasas creatininas de Músculo de cadena de
Miosina H

Distrofina Calpaina p94 troponina 1 rápida de
alfa-actina esquelética

5 Piel queratina K6 queratina K1

Pulmón citoqueratina 18 humana CFTR (K18) proteínas
A, B y C de tensoactivo pulmonar CC-10 P1

Músculo liso sm22 α SM-alfa-actina

Endotelio Endotelina-1 E-selectina factor von

10 Willebrand TIE, KDR/flk-1 Melanocitos
Tirosinasa

Tejido adiposo Lipoproteína lipasa (Zechner et al., 1988)
Adipsina (Spiegelman et al., 1989) acetil-CoA
carboxilasa (Pape and Kim, 1989)

15 glicerofosfato deshidrogenasa (Dani et al.,
1989) adipocito P2 (Hunt et al., 1986)

Mama Proteína Ácida Whey (WAP) (Andres et al.
PNAS 84:1299-1303 1987)

Sangre β -globina

20

[00122] "Elementos reguladores específicos de tejido " son
elementos reguladores (por ejemplo, promotores) que son capaces
de conducir la transcripción de un gen en un tejido mientras
permanece mayormente "inactivo" en otros tipos de tejido. Se
25 entenderá, sin embargo, que los promotores específicos de tejido
pueden tener una cantidad detectable de la actividad
"antecedente" o "base" en aquellos tejidos en donde ellos se
inactivan. El grado en el cual un promotor se activa

selectivamente en un tejido objetivo se puede expresar como una relación de selectividad (actividad en un tejido/actividad objetivo en un tejido de control). A este respecto, un promotor específico de tejido útil en la práctica de la descripción típicamente tiene una relación de selectividad de más de aproximadamente 5. Preferiblemente, la relación de selectividad es mayor que aproximadamente 15.

[00123] Se entenderá adicionalmente que ciertos promotores, mientras no se restringe en la actividad para un tipo de tejido único, puede no obstante mostrar selectividad en que ellos pueden ser activos en un grupo de tejidos, y menos activos o inactivos en otro grupo. Tales promotores también se denominan "específico de tejido", y se contemplan para uso con la descripción. De acuerdo con lo anterior, los elementos reguladores específicos de tejido utilizados en la descripción, tienen aplicabilidad para la regulación de las proteínas heterólogas tal como los polipéptidos que tienen actividad desaminasa citosina de la descripción así como también una aplicabilidad como una secuencia de polinucleótido objetivo en los vectores retrovíricos.

[00124] Los vectores retrovíricos y los polinucleótidos CD y polipéptidos de la descripción se pueden utilizar para tratar un amplio rango de enfermedades y trastornos que incluyen un número de enfermedades proliferativas de célula y trastornos (ver, por ejemplo, la patente Estadounidense Nos. 4,405,712 y 4,650,764; Friedmann, 1989, Science, 244:1275-1281; Mulligan, 1993, Science, 260:926-932, R. Crystal, 1995, Science 270:404-410, cada uno de

los cuales se incorpora aquí como referencia en su totalidad, ver también, The Development of Human Gene Therapy, Theodore Friedmann, Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y., 1999. ISBN 0-87969-528-5, que se incorpora aquí
5 como referencia en su totalidad).

[00125] La descripción también proporciona terapia de gen para el tratamiento de trastornos proliferativos de célula. Tal terapia alcanzará su efecto terapéutico mediante la introducción de una
10 secuencia de polinucleótido terapéutica apropiada (por ejemplo, un polipéptido de la descripción que tiene actividad desaminasa citosina), dentro de las células del sujeto que tiene el trastorno proliferativo. El suministro de las construcciones de polinucleótido se puede lograr utilizando el vector retroviral
15 recombinante de la descripción.

[00126] Adicionalmente, los métodos terapéuticos (por ejemplo, la terapia de gen o métodos de suministro de gen) como se describe aquí se pueden desarrollar in vivo o ex vivo. Por ejemplo, en los
20 métodos para el tratamiento de enfermedades proliferativas de célula o trastornos puede ser útil remover la mayor parte de un tumor antes de la terapia de gen, por ejemplo quirúrgicamente o mediante radiación. En algunas realizaciones, la terapia retroviral se puede preceder o seguir por quimioterapia.

25 [00127] Así, la descripción proporciona un retrovirus recombinante capaz de infectar una célula no dividida, una célula dividida o una célula neoplásica, en donde el retrovirus

recombinante comprende un GAG vírico; un POL vírico; un ENV vírico; un ácido nucleico heterólogo (por ejemplo, que comprende un polipéptido de la descripción que tiene actividad desaminasa citosina) ligado operablemente a un IRES; y las secuencias de
5 ácido nucleico que actúan cis necesarias para empaque, transcripción inversa e integración. El retrovirus recombinante puede ser un lentivirus, tal como HIV, o puede ser un oncovirus. Como se describió anteriormente para el método de producir un retrovirus recombinante, el retrovirus recombinante de la
10 descripción pueden incluir adicionalmente por lo menos una de las proteínas VPR, VIF, NEF, VPX, TAT, REV, y VPU. Mientras que no se espera estar vinculado por una teoría particular, se considera que uno o más de estos productos de gene/proteína son importantes para incrementar el título vírico del retrovirus recombinante
15 producido (por ejemplo, NEF) o puede ser necesario para infección y empaque de virión.

[00128] La descripción también proporciona un método de
20 transferencia de ácido nucleico a una célula objetivo para proporcionar la expresión de un ácido nucleico particular (por ejemplo, una secuencia heteróloga tal como un polinucleótido que codifica un polipéptido que tiene actividad desaminasa citosina). Por lo tanto, en otra realización, la descripción proporciona un método para la introducción y la expresión de un ácido nucleico
25 heterólogo en una célula objetivo que comprende infectar la célula objetivo con el virus recombinante de la descripción y que expresa el ácido nucleico heterólogo en la célula objetivo. Como se mencionó anteriormente, la célula objetivo puede ser cualquier

tipo de célula que incluye célula dividida, no dividida, neoplásica, inmortalizada, modificada y otros tipos de células reconocidos por aquellos expertos en la técnica, tan grande cuando ellos son capaces de infección mediante un retrovirus.

5

[00129] La descripción también proporciona terapia de gen para el tratamiento de trastornos inmunológicos o proliferativos de célula. En una realización, un trastorno proliferativo de la célula se trata al introducir un polinucleótido CD de la descripción, que expresa el polinucleótido para producir un polipéptido que comprende actividad desaminasa citosina y poner en contacto la célula con 5-fluorocitosina en una cantidad y durante un periodo para producir una cantidad citotóxica de 5-FU.

15 [00130] Adicionalmente, la descripción proporciona una secuencia de polinucleótido que codifica un vector retrovítico recombinante de la descripción. La secuencia de polinucleótido se puede incorporar dentro de varias partículas víricas. Por ejemplo, varios vectores víricos que se pueden utilizar para terapia de gen incluyen adenovirus, virus herpes, vaccinia, o, preferiblemente, un virus de ARN tal como un retrovirus y más particularmente un oncovirus de mamífero. El vector retrovítico puede ser un derivado de un retrovirus de murino, simio o humano. Ejemplos de vectores retrovíricos en los cuales un gen extraño (por ejemplo, una secuencia de polinucleótido heterólogo) se puede insertar que incluye, pero no se limita a: derivados de Virus de leucemia de murino de Moloney (MoMuLV), virus de sarcoma murino de Harvey (HaMuSV), virus de tumor mamario de murino

(MuMTV), y Virus de Sarcoma Rous (RSV). Todos estos vectores pueden transferir o incorporar un gen para un marcador seleccionable ya que las células traducidas se pueden identificar y generar. En todavía otra realización, la descripción
 5 proporciona plásmidos que comprenden una construcción derivada retrovímica recombinante. El plásmido se puede introducir directamente dentro de una célula objetivo o un cultivo celular tal como NIH 3T3, HT1080 (humano), CF2 (perro) u otras células de cultivo de tejido. Las células resultantes se liberan del vector
 10 retrovímico dentro del medio de cultivo.

[00131] La descripción proporciona una construcción de polinucleótido que comprende de 5' a 3': un promotor o región reguladora útil para iniciar la transcripción ; una señal de
 15 empaque psi; un gag que codifica la secuencia de ácido nucleico, un pol que codifica la secuencia de ácido nucleico; un env que codifica la secuencia de ácido nucleico; un sitio interno de entrada al ribosoma la secuencia de ácido nucleico; un polinucleótido heterólogo que codifica un marcador, polipéptido
 20 diagnóstico o terapéutico (por ejemplo, un polipéptido que tiene actividad desaminasa citosina); y una secuencia de ácido nucleico LTR. En realizaciones específicas, un polinucleótido heterólogo que codifica un polipéptido que tiene actividad desaminasa citosina puede comprender adicionalmente un dominio que codifica
 25 un polipéptido que comprende actividad UPRT u OPRT.

[00132] Como se describe aquí en cualquier parte y como sigue los varios segmentos de la construcción de polinucleótido de la

descripción (por ejemplo, un polinucleótido retrovítico competente de replicación recombinante) se construyen por ingeniería dependiendo en parte de la célula anfitriona deseada, tiempo o cantidad de expresión, y el polinucleótido heterólogo.

5 Una construcción retrovítica de replicación competente de la descripción se puede dividir en un número de dominios que se pueden modificar individualmente por aquellos expertos en la técnica.

10 [00133] Por ejemplo, el promotor puede comprender un promotor CMV que tiene una secuencia como se establece en la SEQ ID NO: 19, 20, o 22 a partir del nucleótido 1 a cerca del nucleótido 582 y puede incluir la modificación en una o más (por ejemplo, 2-5, 5-10, 10-20, 20-30, 30-50, 50-100 o más bases de ácido nucleico) tanto como el promotor modificado es capaz de dirigir y iniciar la transcripción. En una realización, el promotor o región reguladora comprende un polinucleótido de dominio CMV-R-U5. El dominio CMV-R-U5 comprende el promotor inmediatamente temprano del citomegalovirus humano en la región MLV R-U5. En una
20 realización, el polinucleótido de dominio CMV-R-U5 comprende una secuencia como se establece en la SEQ ID NO: 19, 20, o 22 desde cerca del nucleótido 1 a cerca del nucleótido 1202 o las secuencias que son por lo menos 95% idénticas a una secuencia como se establece en la SEQ ID NO: 19, 20, o 22, en donde el
25 polinucleótido promueve la transcripción de una molécula de ácido nucleico ligada operablemente a esta. El dominio gag del polinucleótido se puede derivar de cualquier número de retrovirus, pero típicamente se derivará de un oncoretrovirus y

más particularmente de un oncoretrovirus de mamífero. En una realización el dominio gag comprende una secuencia desde cerca del nucleótido número 1203 a cerca del nucleótido 2819 o una secuencia que tiene por lo menos 95%, 98%, 99% o 99.8% (rounded to the nearest 10th) de identidad de la misma. El dominio pol del polinucleótido se puede derivar del número de retrovirus, pero típicamente se derivará de un oncoretrovirus y más particularmente de un oncoretrovirus de mamífero. En una realización el dominio pol comprende una secuencia desde cerca del nucleótido número 2820 a cerca del nucleótido 6358 o una secuencia que tiene por lo menos 95%, 98%, 99% o 99.9% (roundest to the nearest 10th) de identidad de esta. El dominio env del polinucleótido se puede derivar de cualquier número de retrovirus, pero típicamente se derivará de un oncoretrovirus y más particularmente de un oncoretrovirus de mamífero. En algunas realizaciones el dominio de codificación env comprende un dominio env amfotrófico. En una realización el dominio env comprende una secuencia desde cerca del nucleótido número 6359 a cerca del nucleótido 8323 o una secuencia que tiene por lo menos 95%, 98%, 99% o 99.8% (roundest to the nearest 10th) de identidad de esta. El dominio IRES del polinucleótido se puede obtener de cualquier número de sitios de entrada de ribosoma internos. En una realización, el IRES se deriva de un virus de encefalomiocarditis. En una realización el dominio IRES comprende una secuencia desde cerca del nucleótido número 8327 a cerca del nucleótido 8876 o una secuencia que tiene por lo menos 95%, 98%, o 99% (roundest to the nearest 10th) de identidad de la misma tanto como el dominio permite la entrada de una ribosoma. El

dominio heterólogo puede comprender una desaminasa citosina de la descripción. En una realización, el polinucleótido CD comprende una secuencia de codón humano optimizado. En todavía otra realización, el polinucleótido CD codifica un polipéptido mutante

5 que tiene desaminasa citosina, en donde las mutaciones confieren estabilización térmica incrementada que incrementa la temperatura de ebullición (T_m) por 10°C que permite la actividad cinética sostenida sobre un amplio rango de temperatura y niveles acumulados incrementados de la proteína. En una realización, la

10 desaminasa citosina comprende una secuencia como se establece en la SEQ ID NO: 19 desde cerca del nucleótido número 8877 a cerca de 9353. El dominio heterólogo se puede seguir por un dominio rico en polipurina. El LTR 3' se puede derivar de cualquier número de retrovirus, típicamente un oncoretrovirus y

15 preferiblemente un oncoretrovirus de mamífero. En una realización, el LTR 3' comprende un dominio U3-R-U5. En todavía otra realización el LTR comprende una secuencia como se establece en la SEQ ID NO: 19 desde cerca del nucleótido 9405 a cerca de 9998 o una secuencia que es por lo menos 95%, 98% o 99.5%

20 (rounded to the nearest 10th) idéntica de la misma.

[00134] La descripción también proporciona un vector retrovítico recombinante que comprende de 5' a 3' a CMV-R-U5, fusión del promotor temprano inmediato del citomegalovirus humano en la

25 región MLV R-U5; un PBS, el sitio de unión de cebador para transcriptasa inversa; un sitio de división 5'; una señal de empaque ψ ; un gag, ORF para el antígeno específico del grupo MLV; un pol, ORF para la poliproteína de polimerasa MLV; un sitio de

división 3'; un 4070A env, ORF para la proteína de envoltura de la cepa MLV 4070A; un IRES, sitio interno de entrada al ribosoma del virus de encefalomiocarditis; una desaminasa citosina modificada (termoestabilizada y optimizada de codón) ; un PPT, 5 tracto polipurina; y un U3-R-U5, repetición de Terminal largo MLV. Esta estructura se describe adicionalmente en la figura 1.

[00135] La descripción también proporciona un vector retrovítico que comprende una secuencia como se establece en la SEQ ID NO: 10 19, 20, o 22.

[00136] En otra realización, la descripción proporciona un método para tratar un sujeto que tiene un trastorno proliferativo de la célula. El sujeto puede ser cualquier mamífero, y es 15 preferiblemente un humano. El sujeto se pone en contacto con un vector retrovítico recombinante de replicación competente de la descripción. El poner en contacto puede ser in vivo o ex vivo. Los métodos para administrar el vector retrovítico de la descripción se conocen en la técnica e incluyen, por ejemplo, 20 administración sistémica, administración tópica, administración intraperitoneal, administración intramuscular, intracraneal, cerebrospinal, así como también administración directamente en el sitio de un trastorno proliferativo de célula o tumor. Se conocen en la técnica otras rutas de administración.

25 [00137] Así, la descripción incluye varias composiciones farmacéuticas útiles para tratar un trastorno proliferativo de la célula. Las composiciones farmacéuticas de acuerdo con la

descripción se preparan al traer un vector retrovítico que
 contiene una secuencia de polinucleótido heterólogo útil en
 tratar o modular un trastorno proliferativo de la célula de
 acuerdo con la descripción dentro de una forma adecuada de
 5 administración a un sujeto utilizando portadores, excipientes y
 aditivos o auxiliares. Los portadores o auxiliares frecuentemente
 utilizados incluyen carbonato de magnesio, dióxido de titanio,
 lactosa, manitol y otros azúcares, talco, proteína de leche,
 gelatina, almidón, vitaminas, celulosa y sus derivados, aceites
 10 vegetales y de animal, polietilenglicoles y disolventes, tal como
 agua estéril, alcoholes, glicerol y alcoholes polihídricos. Los
 vehículos intravenosos incluyen replenishers de fluido y
 nutrientes. Los conservantes incluyen antimicrobianos, anti-
 oxidantes, agentes quelantes y gases inertes. Otros portadores
 15 farmacéuticamente aceptables incluyen soluciones acuosas,
 excipientes no tóxicos, que incluye sales, conservantes,
 amortiguantes y similares, como se describe, por ejemplo, en
 Remington's Pharmaceutical Sciences, 15th ed. Easton: Mack
 Publishing Co., 1405-1412, 1461-1487 (1975) y The National
 20 Formulary XIV., 14th ed. Washington: American Pharmaceutical
 Association (1975), cuyos contenidos se incorporan aquí como
 referencia. El pH y la concentración exacta de los varios
 componentes de la composición farmacéutica se ajustan de acuerdo
 con las técnicas de rutina en el arte. Ver Goodman and Gilman's
 25 The Pharmacological Basis for Therapeutics (7th ed.).

[00138] Por ejemplo, y no por vía de limitación, un vector retrovítico útil en tratar un trastorno proliferativo de la

célula incluirán una proteína ENV amfotrópica, proteínas GAG, y proteínas POL, una secuencia promotora en el genoma retrovílico de la región U3, y todas las secuencias que actúan cis necesarias para replicación, empaque e integración del genoma retrovílico dentro de la célula objetivo.

[00139] Los siguientes Ejemplos están destinados a ilustrar, pero no a limitar la descripción. Mientras que tales Ejemplos son típicos para aquellos que los pueden utilizar, se pueden utilizar alternativamente otros procedimientos conocidos por aquellos expertos en la técnica.

EJEMPLOS

15 Ejemplo 1 Construcción de genes CD modificados e inserción dentro de los vectores de plásmido.

[00140] Los mejoramientos genéticos para el gen de desaminasa citosina de levadura tipo natural se ha hecho e incluye: (1) tres mutaciones posicionales que cambian tres aminoácidos (A23L, I140L y V108I) para incrementar la estabilidad térmica de la proteína de desaminasa citosina de levadura y (2) modificaciones de secuencia de gen adicionales para mejorar las secuencias de uso de codón humano para mejorar la eficiencia de traducción de proteína en células humanas sin cambios adicionales para la secuencia de aminoácido.

[00141] El diseño de secuencia para el CD incluye CD optimizado, CD-UPRT (ligador +/-) y CD-OPRTasa (ligados +/-). La secuencia codificante de desaminasa citosina final que puede comprender en el extremo 5' un sitio PSII (longitud completa) y el extremo 3' del sitio NotI más la cola poli A para la estrategia basada en el casete PSII/NotI.

[00142] La siguiente secuencia que comprende una desaminasa citosina de levadura se utiliza para clonación, optimización y mutación (los ácidos nucleicos boxed que comprende los sitios de restricción utilizados en los métodos posteriores para clonación:

```

AACACGAATTATAAATGGTGACAGGGGGAATGGCAAGCAAGTGGGATCAGAAGGGTATGGACA
TTGCCTATGAGGAGGCGGCCTTAGGTTACAAAGAGGGTGGTGTTCCTATTGGCGGATGTCTT
ATCAATAACAAAGACGGAAGTGTTCTCGGTCGTGGTCACAACATGAGATTTCAAAGGGATC
CGCCACACTACATGGTGAGATCTCCACTTTGGAAAAGTGTGGGAGATTAGAGGGCAAAGTGT
ACAAAGATACCACTTTGTATACGACGCTGTCTCCATGCGACATGTGTACAGGTGCCATCATC
ATGTATGGTATTCCACGCTGTGTTGTCGGTGAGAACGTTAATTTCAAAGTAAGGGCGAGAA
ATATTTACAACTAGAGGTCACGAGGTTGTTGTTGTTGACGATGAGAGGTGTAAAAGATCA
TGAAACAATTTATCGATGAAAGACCTCAGGATTGGTTTGAAGATATTGGTGAGTAGGCGGCC
GCGCCATAGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG (SEQ ID
NO:31)

```

15 La siguiente Tabla resume los genes y los vectores de plásmido resultantes que se hacen y sus nombres.

Tabla: Construcciones de vector y nombres

Código Tocage n	Nombre de referenc ia	Nombre original	Pro m 5' LTR	envoltu ra	Vecto r	IRE S	transgen	3'LT R
T5.000 0	pACE-γCD	pACE-CD (Tai et al. 2005)	CMV	Amfo (4070A)	pACE	EMC V	Levadura CD en peso	MLV U3
T5.000 1	pAC3- γCD1	Secuenci a CDopt	CMV	Amfo (4070A)	pAC3	EMC V	CD modifica do	MLV U3
T5.000 2	pAC3- γCD2	CDopt+3p t	CMV	Amfo (4070A)	pAC3	EMC V	CD modifica do	MLV U3
T5.000 3	pAC3- γCD2-U	CDopt+3p t-UPRT	CMV	Amfo (4070A)	pAC3	EMC V	CD2-UPRT	MLV U3
T5.000 4	pAC3- γCD2-O	CDopt+3p t-OPRT	CMV	Amfo (4070A)	pAC3	EMC V	CD2-OPRT	MLV U3
T5.000 5	pAC3- γCD2-LO	CDopt+3p t-LINK- OPRT	CMV	Amfo (4070A)	pAC3	EMC V	CD2-L- OPRT	MLV U3
T5.000 6	pAC3- eGFP	pAC3- emd, pAC3GFP	CMV	Amfo (4070A)	pAC3	EMC V	Emera id GFP	MLV U3
T5.000 7	pAC3- γCD	pAC3- γCD	CMV	Amfo (4070A)	pAC3	EMC V	Levadura CD en peso	MLV U3

[00143] El vector retrovífico de replicación competente descrito
5 por Kasahara et al. pACE-CD (Patente Estadounidense No.
6,899,871, cuya descripción se incorpora aquí) se utiliza como
una base para modificaciones adicionales. Un vector (pAC3-γCD) se
modifica para expresar un gen de desaminasa citosina de levadura

102

modificado como se describe aquí y se utiliza en las construcciones. Ver IA adelante para un diagrama de la construcción de vector para el retrovirus competente de replicación transfectada inicial. El CMV es el promotor temprano inmediato CMV humano, U3, R y U5 son las regiones correspondientes de la repetición de Terminal largo vírico (LTR). El Gag, pol y env son las regiones codificantes de proteína vírica. El IB y ID muestra la estructura de plásmido y una secuencia de la descripción. Después que se sintetizan los genes en un contractor (Bio Basic Inc., Markham, Ontario, Canadá) ellos se insertan dentro del sitio Psil -NotI de la estructura del vector pAC3 (figura 1). La estructura de plásmido se genera normalmente al cortar el plásmido pAC3-eGFP con Psil y NotI y purificar el fragmento grande (aproximadamente 11kb) que codifica el plásmido y la estructura retroviral)

[00144] A. Gen CD optimizado de codón humanizado (CDopt, aka CD1, T5.0001) Una comparación de una desaminasa citosina de codón humano optimizado de Conrad et al. y PCT WO 99/60008 indica 91 codones totales optimizados en ambos, 36 codones idénticos, 47 codones que tiene cambios de tercer par base (todos codifican el mismo aminoácido) y 9 codones son diferentes (sin embargo ellos codifican el mismo aminoácido). De los 9 codones que difieren:

AGC (Ser) a TCC (Ser)
CGT (Arg) a AGG (Arg)
CCA (Pro) a CCT (Pro)

103

104

20 descripción: codificante: 5'-
tcgaggatatcggcgagtgaaaccggtattctttttggc-3' (SEQ ID NO: 25)
anticodificante: 5'-gccaaaaagaataacgggtttcactcgccgatatcctcga-3'
(SEQ ID NO:26)

codificante:

5'tcggcgagtgatccggcgggcggcctccggcgggcggcctccggcgggcggcgc
tccggcgggcgggcgccaacccggttatt-3' (SEQ ID NO:27)

anticodificante:

5'-

5 aataacggggttggcgcccgcgcggaggcgccgcccggaggcgccgcgc
cggaggcgccgcccgcggatcactcgccga-3' (SEQ ID NO: 28)

[00147] Para incrementar la capacidad del CD de levadura nativa,
se construyen por ingeniería tres sustituciones de aminoácido
dentro de la proteína. Estas sustituciones se hacen solas o en
combinación con la optimización del codón humano.

[00148] Las tres sustituciones de aminoácido son: A23L, V108I,
I140L. Una secuencia que codifica estas sustituciones se muestra
adelante.

ATGGTGACAGGGGGAATGGCAAGCAAGTGGGATCAGAAGGGTATGGACATTGCCTATGAGGA
GGCG**TTA**TTAGGTTACAAAGAGGGTGGTGTTCCTATTGGCGGATGTCTTATCAATAACAAAG
ACGGAAGTGTTCTCGGTGCTGGTCACAACATGAGATTTCAAAGGGATCCGCCACACTACAT
GGTGAGATCTCCACTTTGGAAAAGTGTGGGAGATTAGAGGGCAAAGTGTACAAAGATACCAC
TTTGTATACGACGCTGTCTCCATGCGACATGTGTACAGGTGCCATCATCATGTATGGTATTC
CACGCTGTGT**CATC**GGTGAGAACGTTAATTTCAAAGTAAGGGCGAGAAATATTTACAAACT
AGAGGTCACGAGGTTGTTGTTGTTGACGATGAGAGGTGTAAGAG**TTA**ATGAAACAATTTAT
CGATGAAAGACCTCAGGATTGGTTTGAAGATATTGGTGAGTAG**GCGGCCGC**GCCATAGATAA
AATAAGATTTTATTATTAGTCTCCAGAAAAGGGGGG (SEQ ID NO:33)

[00149] El polipéptido codificado comprende la siguiente
secuencia (aminoácidos sustituidos subrayados en negro):

1 MVTGGMASKWDQKGM¹DIAYEEALLGYKEGGVPIGGCLINNKG²SVLGRGHNMRFQKGSAT
61 LHGEISTLENCGRLEGKVYKDTTLYTTLSPCDMCTGAIIMYGI³PRCVIGENVNFKSKGEK
121 YLQTRGHEVVVDDDERCKKLMKQFIDERPDQWFEDIGE-

105

[00150] El diseño de construcción final que integra 3 sustituciones de aminoácido A23L/V108I/I140L utilizando codones preferidos y los usos preferidos de codón humano utilizados para la secuencia completa (este gen se denomina CDopt+3pt [aka CD2] y T5.0002 cuando se inserta dentro del vector de plásmido pAC3 que codifica el RCR con IRES) (SEQ ID NO:34):

```

1 ATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTTACGAG
61 GAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAAC
121 AAGGACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACC
181 CTGCACGGCGAGATCTCACCCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAG
241 GACACCACCCTGTACACCACCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATG
301 TACGGCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCCAAGGGCGAGAAG
361 TACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGCTG
421 ATGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTGATAA

```

10 Los codones subrayados denotan codones preferidos para las sustituciones de aminoácido.

[00151] El diseño de secuencia optimizado CD (preferencia de codón humano + 3 sustituciones de aminoácido)

15

```

AACACGATTTATAAATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATA
TCGCTTACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTG
ATCAACAACAAGGACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTC
CGCCACCCTGCACGGCGAGATCTCACCCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGT
ACAAGGACACCACCCTGTACACCACCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATC
ATGTACGGCATCCCTAGGTGTGTGATCGGCGAGAACGTGAACTTCAAGTCCAAGGGCGAGAA
GTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGCTGA
TGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTAAAGCGGCC
GGGCCATAGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG (SEQ ID
NO:35)

```

[00152] C. Construcción de gen de fusión CD-UPRT (CDopt+3pt-UPRT, [aka CDopt-UPRT y CD2-UPRT], T5.0003 en el vector RCR de plásmido

20 pAC3) Una construcción de fusión también se desarrolla que

comprende un polipéptido CD como se describió anteriormente
ligado a un polipéptido UPRT para generar una secuencia UPRT
optimizada CD utilizando el Esquema I como se estableció en la
figura 2A. Se utilizan los siguientes cebadores para eliminar el
5 inicio-detención entre el CD y UPRT.

Secuencias cebadoras:

Nombre de Cebador	Secuencia Cebadora (5' a 3') (SEQ ID No:)			
Del 118-123	5'-tcgaggatatcggcgagtgaaacccgttattctttttggc-3' (36)			
Del 118-123-anticodificante	5'-gccaaaaagaataacgggtttcactcgccgatatcctcga-3' (37)			
Nombre de Cebador	Longitud (nt.)	Tm	Energía Dúplex a 68° C	Costo de energía de mismatches
Del 118-123	40	79.06° C	-44.37 kcal/mol	21.1%
Del 118-123-anticodificante	40	79.06° C	-47.95 kcal/mol	20.3%
Nombre de Cebador	Cebador-Plantilla Dúplex			
Del 118-123	5'-tcgaggatatcggcgagtg-----aacccgttattctttttggc-3' ccaagctcctatagccgctcactatctacttgggcaataagaaaaaccgaag			
Del 118-123-anticodificante	ggttcgaggatatcggcgagtgatagatgaacccgttattctttttggcttc 3'-agctcctatagccgctcact-----ttgggcaataagaaaaaccg			

10 [00153] El polinucleótido de fusión resultante comprende 1296 bp
y la secuencia se establece inmediatamente adelante:

107

AACACGATTATAAATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTT
ACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAAG
GACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGCACGGCGA
GATCTCCACCCTGGAGAAGTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACCA
CCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACGGCATCCCTAGGTGTGTGATCGGC
GAGAACGTGAACTTCAAGTCCAAGGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT
TGACGATGAGAGGTGTAAGAAGCTGATGAAGCAGTTCATCGACGAGAGGCCCTCAGGACTGGTTCGAGG
ATATCGGCGAGAACCCGTTATTCTTTTGGCTTCTCCATTCTTGTACCTTACATATCTTATATATTAT
CCAAACAAAGGGTCTTTTCGTTAGCAAACCTAGAAATCTGCAAAAATGTCTTCGGAACCATTTAAGAA
CGTCTACTTGCTACCTCAAACAAACCAATTGCTGGGTTTGTACACCATCATCAGAAATAAGAATACAA
CTAGACCTGATTTTCAATTTTCTACTCCGATAGAATCATCAGATTGTTGGTTGAAGAAGGTTTGAACCAT
CTACCTGTGCAAAAGCAAATTGTGGAACTGACACCAACGAAACTTCGAAGGTGTCTCATTTCATGGG
TAAATCTGTGGTGTTCATTGTGAGAGCTGGTGAATCGATGGAGCAAGGATTAAGAGACTGTTGTA
GGTCTGTGCGTATCGGTAAAATTTAATTCAAAGGGACGAGGAGACTGCTTTACCAAAGTTATTCTAC
GAAAAATTACCAGAGGATATATCTGAAAGGTATGTCTTCTTATTAGACCCAATGCTGGCCACCGGTGG
TAGTGCTATCATGGCTACAGAAGTCTTGATTAAGAGAGGTGTTAAGCCAGAGAGAATTTACTTCTTAA
ACCTAATCTGTAGTAAGGAAGGGATTGAAAAATACCATGCCGCCTTCCCAGAGGTCAGAATTGTTACT
GGTGCCCTCGACAGAGGTCTAGATGAAAACAAGTATCTAGTTCAGGGTTGGGTGACTTTGGTGACAG
ATACTACTGTGTTTAA[CGGGCCGC]GCCATAGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGG
GGGG (SEQ ID NO:38)

[00154] D. Construcción del gen de fusión UPRT de ligador CD
(CDopt+3pt- LINK-UPRT [aka CDopt-LINKER-UPRT y CD2-L-UPRT]). Una

5 construcción de fusión también se desarrolla mediante clonación
de un dominio 4 ligador (Ser-Gly-Gly-Gly- GIy) entre y en
estructura con el polipéptido CD y el polipéptido UPRT para
generar una secuencia UPRT de ligador optimizado CD utilizando el
Esquema II como se describe en 2B. Se utilizan los siguientes
10 cebadores para insertar el ligador.

Nombre de Cebador	Secuencia Cebadora (5' a 3') (SEQ ID No:)			
Ins_60nt_después_477	5'- tcggcgagtgatccggcgggcggcctccggcgggcggcctccggcg gcggcgccctccggcgggcgccaaccggtatt-3' (39)			
Ins_60nt_después_477- anticodificante	5'- aataacgggttggcgccgcgcggaggcgccgcgcgcggaggcgcc			
	gccgcggaggcgccgcgcgcggatcactcgccga-3' (40)			
Nombre de Cebador	Longitud (nt.)	Tm	Energía Dúplex a 68° C	Costo de energía de mismatches
Ins_60nt_después_477	82	79.77°	-30.19	83.3%

		C	kcal/mol	
Ins_60nt_después_477- anticodificante	82	79.77° C	-32.31 kcal/mol	82.2%

[00155] La construcción resultante tiene el tamaño: 1356 bp y la secuencia inmediatamente adelante:

AACACGATTATATAATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTT
ACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAAG
GACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGCACGGCGA
GATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACCA
CCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACGGCATCCCTAGGTGTGTGATCGGC
GAGAACGTGAACTTCAAGTCCAAGGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT
TGACGATGAGAGGTGTAAGAAGCTGATGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGG
ATATCGGCGAGTCCGGCGGCGGCGCCTCCGGCGGCGGCGCCTCCGGCGGCGGCGCCTCCGGCGGCGGCGC
GCCAACCCGTTATTCTTTTTGGCTTCTCCATTCTTGACCTTACATATCTTATATATTATCCAAACAA
AGGGTCTTTTCGTTAGCAAACCTAGAAATCTGCAAAAAATGTCTTCGGAACCATTTAAGAACGTCTACT
TGCTACCTCAAACAAACCAATTGCTGGGTTTGTACACCATCATCAGAAATAAGAATACAACCTAGACCT
GATTTTCATTTTCTACTCCGATAGAATCATCAGATTGTTGGTTGAAGAAGGTTTGAACCATCTACCTGT
GCAAAAGCAAATTTGTGAAACTGACACCAACGAAAACCTCGAAGGTGTCTCATTCATGGGTAAATCT
GTGGTGTTCATTGTGAGAGCTGGTGAATCGATGGAGCAAGGATTAAGAGACTGTTGTAGGTCTGTG
CGTATCGGTAAATTTTAATTCAAAGGGACGAGGAGACTGCTTTACCAAAGTTATTCTACGAAAAATT
ACCAGAGGATATATCTGAAAGGTATGTCTTCTATTAGACCCAATGCTGGCCACCGGTGGTAGTGCTA
TCATGGCTACAGAAGTCTTGATTAAGAGAGGTGTTAAGCCAGAGAGAATTTACTTCTTAAACCTAATC
TGTAGTAAGGAAGGGATTGAAAAATACCATGCCGCCTTCCCAGAGGTGAGAATTGTTACTGGTGCCCT
CGACAGAGGTCTAGATGAAAACAAGTATCTAGTTCCAGGGTTGGGTGACTTTGGTGACAGATACTACT
GTGTTTAAAGCGCGCGCGCCATAGATAAAATAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG
(SEQ ID NO:41)

[00156] E. Construcción de gen de fusión CD-OPRT (CDopt+3pt-OPRT [aka CDopt-OPRT y CD2-OPRT], T5.0004 cuando se inserta dentro del vector RCR de plásmido pAC3) Una construcción de fusión también se desarrolla que comprende un polipéptido CD como se describió anteriormente ligado a un polipéptido OPRT para generar un CD optimizado-OPRTasa (dominio funcional humano CD humanizado+3ptmutación + OPRTasa) utilizando el Esquema III como se muestra en la figura 2C.

[00157] La construcción resultante comprende un tamaño de 1269 bp y la secuencia inmediatamente adelante:

109

AACACGA**TTATAA**ATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTT
 ACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGGCGGCTGTCTGATCAACAACAAG
 GACGGCAGTGTGCTGGGCGAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGCACGGCGA
 GATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACCA
 CCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACGGCATCCCTAGGTGTGTGATCGGC
 GAGAACGTGAACTTCAAGTCCAAGGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT
 TGACGATGAGAGGTGTAAGAAGCTGATGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGG
 ATATCGGGCGAGGG**CGGTCGCTCGTG**cagctttggggccattgggtgacgggtctgtacgacgtgcaggct
 ttcaagtttggggacttcgtgctgaagagcgggctttcctcccccatctacatcgatctgcggggcat

cgtgtctcgaccgcgtcttctgagtcaggttgcagatatatttattccaaactgccccaaatgcaggca
 tcagttttgacaccgtgtgtggagtgcttatacagctttgccattggctacagttatctgttcaacc
 aatcaaattccaatgcttattagaaggaaagaaacaaaggattatggaactaagcgtcttgtagaagg
 aactattaatccaggagaaacctgtttaatcattgaagatgttgtcaccagtggtatctagtgttttg
 aaactgttgaggttcttcagaaggagggttgaagggtcactgatgccatagtgtgttggacagagag
 cagggaggcaaggacaagttgcaggcgacgggatccgcctccaactcagtggtgtacattgtccaaaat
 gctggagattctcgagcagcagaaaaaagttgatgctgagacagttgggagagtgaagaggtttattc
 aggagaatgtctttgtggcagcgaatcataatgggttctccctttctataaaggaagcacccaaagaa
 ctcaGCTTCGGTGCACGTGCAGAGCTGCCAGGATCCACCCAGTTGCATCGAAGTAA**GCGGCCG**GCC
 ATAGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGG (SEQ ID NO:42)

- 5 [00158] F. Construcción del gen de fusión CD-ligador-OPRT
(CDopt+3pt- LINK-OPRT, [aka CDopt-LINKER-OPRT y CD2-L-QPRT],
T5.0005 en el vector RCR de plásmido pAC3) Una construcción de
 fusión también se desarrolla mediante clonación de un dominio
 ligador (Ser-Gly-Gly-Gly-Gly) 4) entre y en estructura con el
 10 polipéptido CD y el polipéptido OPRT para generar una secuencia
 CD optimizada-ligador-OPRT utilizando el Esquema IV como se
 muestra en la figura 2D.

- [00159] La construcción resultante comprende un tamaño de 1329 bp
 15 y la secuencia inmediatamente adelante:

AACACGATTATAAATGGTGACCGGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTT
 ACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAAG
 GACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGCACGGCGA
 GATCTCCACCCTGGAGAAGTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACCA
 CCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACGGCATCCCTAGGTGTGTGATCGGC
 GAGAACGTGAACTTCAAGTCCAAGGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT
 TGACGATGAGAGGTGTAAGAAGCTGATGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGG
 ATATCGGCGAGTCCGGCGGCGGCGCTCCGGCGGCGGCGCTCCGGCGGCGGCGCTCCGGCGGCGGCG
 GCCGCGGTGCTCGTGcagctttggggccattggtgacgggtctgtacgacgtgcaggctttcaagtt
 tggggacttcgtgctgaagagcgggcttttccctcccccatctacatcgatctgcggggcatcgtgtctc
 gaccgcgtcttctgagtcagggttcagatatattttattccaaactgccccaaatgcaggcatcagtttt
 gacaccgtgtgtggagtgccttatacagctttgccattggctacagttatctgttcaaccaatcaaat
 tccaatgcttattagaaggaaagaaacaaaggattatggaactaagcgtctttagagaaggaaactatta
 atccaggagaaacctgtttaatcattgaagatgttgtcaccagtggtctagtgttttgaaactgtt
 gaggttcttcagaaggagggttgaaggtcactgatgccatagtgtgttggacagagagcagggagg
 caaggacaagttgcaggcgcacgggacccgctccactcagtggtgtacattgtccaaaatgctggaga
 ttctcgagcagcagaaaaaagttgatgctgagacagttgggagagtgaagaggtttattcaggagaat
 gtctttgtggcagcgaatcataatggttctccctttctataaaggaagcaccacaaagaactcaGCTT
 CGGTGCACGTGCAGAGCTGCCAGGATCCACCCAGTTGCATCGAAGTAAGCGGCCGGCCATAGATAA
 AATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG. (SEQ ID NO:43)

Ejemplo 2 Productor de Vector Infeccioso

- 5 Se puede producir el vector en un número de formas, pero la etapa de cebador es para introducir el vector de ADN de plásmido dentro de las células para permitir la producción de las partículas infecciosas, que se pueden cosechar del sobrenadante celular. Una vez se han generado partículas infecciosas se pueden implementar
- 10 otros métodos de producción por aquellos expertos en la técnica. Se generan partículas de vector mediante transfección transitoria de las células 293T (Pear et al. Proc Natl Acad Sci U S A. 90:8392-8396 1993). Las células 293T se descongelan y se ponen dentro del cultivo, luego se pasa dos veces en matraces T-75 que
- 15 contienen 15 mL del medio DMEM que se prepara al mezclar medio Alto en Glucosa DMEM (Hyclone# 30081, 500 mL) con FBS (Hyclone* SH30070, 50 mL), L-Glutamina (Cellgro* 25-005-CI, 5 mL), NEAA (Hyclone #SH30238, 5 mL), y Penicilina-strep (Cellgro# 30-002-CI, 5 mL). Los matraces se incuban a 37 C y 5% de CO₂. Después se
- 20 siembra 3 células de pasaje en 6 T-25's, cada una contiene 5 mL

del medio, en una densidad celular 1.8×10^6 células/T-25 (o 7.2×10 células/cm). Un día después de sembrar el T-25's, las células se transfectan con el plásmido T5.0002 que expresa el vector vírico utilizando el Equipo de Transfección de Fosfato de Calcio de Promega (Cat# E1200). Dieciocho horas después de la transfección, el medio es uno establecido de los matraces (3 matraces cada conjunto) se reemplazan con medio fresco que contiene 10 mM nd de NaB. El medio en el 2 conjunto de los matraces no se reemplaza, que sirve como un control (cero NaB).

10 Ocho horas post-tratamiento NaB, el medio en todos los matraces se reemplazan con medio fresco que no contiene NaB. La expresión se le permite continuar para ambos conjuntos de matraces hasta el siguiente día (22 horas de duración). Los sobrenadantes de ambos conjuntos de matraces se cosechan y se evalúan para sus títulos

15 mediante qPCR expresado en las Unidades de Transducción (TU) /ml (ver ejemplo 3).

Los resultados del título se muestran en la siguiente tabla.

Condición	Primer título	Segundo título (después de almacenamiento a - 80° C durante 68 días)
Sin NaB	1.5 (± 0.05) $\times 10^6$ TU/mL	1.2 (± 0.2) $\times 10^6$ TU/mL
10 mM de NaB	1.4 (± 0.3) $\times 10^6$ TU/mL	7.0 (± 0.14) $\times 10^5$ TU/mL

Se producen preparaciones de vector posteriores en esta forma, sin butirato de sodio. Los otros plásmidos de vector se han utilizado en la misma forma para generar las preparaciones de vector con títulos entre 10^5 TU/ml y 10^7 TU/ml. Tal material se puede purificar adicionalmente y se concentra, si se desea, como se describe adelante y se ve también: US 5792643; T. Rodriguez et al. J Gene Med 9:233 2007; Aplicación Estadounidense 61218063. En ciertas realizaciones de la descripción se calcula la dosificación mediante gramos de peso del cerebro. En tales realizaciones, la dosificación de un vector retrovítico de replicación competente de la descripción útil en los métodos para el tratamiento puede variar de 10 a 10 TU por gramo de peso de cerebro.

Ejemplo 3 Ensayo de titulación PCR cuantitativo

[00247] La concentración de vector funcional, o título, se determina utilizando un método con base en PCR cuantitativo (qPCR). En este método, el vector se titula al infectar una estirpe celular anfitriona transducible (por ejemplo células de carcinoma prostático humano PC-3, ATCC Cat# CRL-1435) con un volumen estándar del vector y se mide la cantidad resultante del provirus presente dentro de las células anfitrionas después de transducción. Las células y el vector se incuban bajo condición de cultivo estándar (37°C , 5% de CO_2) durante 24 hr para permitir completar la infección antes de la adición del AZT anti-retrovítico para detener la replicación del vector. Luego, las células se cosechan del plato de cultivo y el ADN genómico (gADN)

se purifica utilizando un equipo de Purificación gADN Invitrogen Purelink y se eluye de la columna de purificación con agua RNasa-/DNasa-agua libre estéril. La relación de absorbancia A260/A280 se mide en un espectrofotómetro para determinar la concentración y pureza relativa de la muestra. Las concentraciones de gADN se normalizan con RNasa-/DNasa-agua libre adicional en la concentración inferior de cualquier conjunto de preparaciones de gADN de tal manera que el ADN de entrada para el qPCR es constante para todas las muestras analizadas. La pureza de ADN genómica se evalúa adicionalmente mediante electrofóresis de una alícuota de cada muestra en un bromuro de etidio teñido con 0.8% de gel de agarosa. Si la muestra pasa un rango de absorbancia A260/A280 de 1.8-2.0 y muestra una banda única de gADN, entonces la muestra está lista para análisis qPCR del número de copia de provirus del vector. Utilizando cebadores que interrogan la región LTR del provirus (el ADN de vector transcrito inverso y el ADN vector que se integra dentro del gADN anfitrión), se desarrolla qPCR para estimar el número total de eventos de transducción que ocurren cuando el volumen conocido del vector se utiliza para transducir el número conocido de células. El número de eventos de transducción por reacción se calcula de una curva estándar que utiliza un plásmido que lleva objetivo del número de copia conocido que es serial diluido de $1E7$ a $1E1$ copias y medido bajo condiciones qPCR idénticas como las muestras. Se utilizan muchos equivalentes genómicos como se conocen para cada reacción qPCR (a partir de la concentración previamente determinada) y como muchos eventos de transducción que ocurren por reacción, determinamos el número total de eventos de transducción que

114

ocurren con base en el número total de células que están presentes en el momento de transducción. Este valor es el título del vector después de dilución dentro del medio que contiene las células durante la transducción inicial. Para calcular el valor de título correcto, la dilución se corrige al multiplicar por el volumen de cultivo y el volumen de título dividido por el volumen de título. Estos experimentos se desarrollan en cultivos de replicado y se analizan mediante qPCR utilizando mediciones de triplicados para cada condición para determinar un título promedio y con su desviación estándar asociada y coeficiente de varianza.

Ejemplo 4 Niveles de expresión medidos por Western Blot

[00160] La figura 3 demuestra que se observan altos niveles del codón humano optimizado con las tres mutaciones para mayor estabilidad comparado con proteína yCD tipo natural en un análisis Western blot de células U-87 infectadas con virus que codifica los genes desaminasa citosina tipo natural (ACE-yCD) o completamente optimizados (AC3-yCD2).

Ejemplo 5 Estabilidad genética de vectores víricos.

[00161] Se reconoce que después de la transcripción inversa y el primer evento de integración dentro de las células tratadas, el provirus de ADN y cualquier retrovirus progenie posterior tiene una estructura LTR convencional de MLV al final. Esta

configuración se ha mostrado que es estable después de ciclos múltiples de infección (Ver figura 4 adelante).

[00162] Aproximadamente 10^6 células, naive U-87 se infectan inicialmente con el vector vírico en un MOI de 0.01, y se hacen crecer hasta que se infectan completamente para completar un único ciclo de infección. El sobrenadante luego se repasa dentro de las células no infectadas y se repite el ciclo. La estabilidad genómica de la secuencia yCD2 se evalúa mediante amplificación PCR del provirus integrado de las células infectadas utilizando cebadores específicos MLV que flanquean el sitio de inserción de transgen. Para cada conjunto de infecciones, la amplificación del plásmido vector (pAC3-yCD2 y el vector pACE-CD de Kasahara et al.) también se desarrolla to track el tamaño de amplicón de longitud completa en el gel. La aparición de cualesquier bandas más pequeñas que el amplicón de longitud completa será un indicador de inestabilidad del vector. Tales experimentos demuestran que un vector de la descripción (T5.0002 - que comprende el vector modificado y el polinucleótido heterólogo CDopt+3pt (CD2) mantiene estabilidad para más pasajes que pACE-CD o T5.007 ambos de los cuales llevan a cabo la levadura tipo natural.

Ejemplo 6 Experimentos de Muerte Celular

25

[00163] En los experimentos de cultivo celular in-vitro, los experimentos demuestran que la desaminasa citosina en las células que expresan la proteína yCD2 es por lo menos tan activa como

aquella de las células que expresan la proteína yCD tipo natural, al desarrollar titulaciones 5-FC en las células de rata RG2 (figura 5A) o células U-87 (figura 5B) infectadas con el virus hecho de pAC3-yCD2/T5.0002) [AC3-yCD2 (V)] de pAC3-yCD/T5.0007 [AC3-yCD(V)] o los otros vectores. En resumen, para las células U-87, 5 días post-infección en una multiplicidad de infección de 0.1 (es decir 100% infectado) con el vector AC3-yCD (CD tipo natural) o el vector AC3-yCD2 (termoestabilizado & optimizado por codón) se someten a cantidades incrementadas de 5-FC o 0.1 mM de 5-FU como un control positivo durante 8 días. En el día 8 de tratamiento 5-FC, se evalúan los cultivos celulares para viabilidad utilizando un ensayo MTS (Promega CellTiter 96 AQUEOUS Un Ensayo de Proliferación de Solución). Los datos muestran muerte comparable entre los dos vectores retrovíricos en dosis incrementadas de tratamiento 5-FC. Los cultivos RG2 se tratan de forma similar y también muestran si cualquier cosa cambia ligeramente en la curva de muerte hacia baja concentración de 5-FC para el virus del ensayo de expresión T5.0002CD. Se transluen células U87 en una multiplicidad de infección (MOI) de 0.1, cultivado durante 5 días para permitir rociado vírico inicial y se cosechan células del día 5 de post-transducción. Las células luego se recolectan mediante centrifugación a 800 x g durante 5 min. El sobrenadante se aspira lejos del glóbulo celular y se lava con 5 mL de solución salina amortiguada con fosfato (PBS) y de nuevo se centrifuga a 800 x g durante 5 min. El glóbulo celular resultante se toma en 1.5 mL de PBS, se resuspende mediante el pesaje a través de a pipette tip y se coloca en un congelador a -20C. Las células se lisan mediante un método de

congelado/descongelado. A las células previamente resuspendidas se les permite descongelar a temperatura ambiente, se pasa a través de a pipette tip, se mezcla con cóctel de inhibidor de proteasa y de nuevo se vuelve a congelar a -20C. Previamente al
5 ensayo de enzima, la muestra se descongela de nuevo a temperatura ambiente y se pasa a través de a pipette tip. La suspensión luego se centrifuga a 14,000 rpm en una centrifuga de mesa durante 5 min. El sobrenadante se decanta lejos del glóbulo y se coloca en un tubo eppendorf fresco y se coloca en hielo. La actividad de la
10 enzima yCD se evalúa al utilizar un ensayo HPLC.

[00164] Se desarrolla ensayo HPLC en una unidad Shimadzu LC20AT conectada en serie con un detector de fotoarreglo y autoinyector. La fase sólida es una columna Hypersil BDS C18 HPLC con un tamaño
15 de esfera de 5 μ m y dimensiones de columna de 4.0 x 250 mm. La fase móvil es 50 mM de fosfato de amonio, pH 2.1, que contiene 0.01% de perclorato terc-butilamonio y 5% de metanol; el sistema se equilibra a 22 C. Todos los reactivos son grado ACS y los
disolventes son grado HPLC. Se hace una mezcla de reacción que
20 consiste de 800 μ L con una concentración final de 0.125 mg/mL 5FC (1 mM) en PBS y se coloca en un frasco de automuestreo de 1.5 mL. La reacción luego se inicia al agregar 200uL de cada lisato celular. Se colocan frascos de reaacción/automuestreo en el automuestreador y se inyecta 5uL de la mezcla de reacción. Los
25 puntos de tiempo se toman periódicamente by retrieving una alícuota de 5 uL de cada frasco de reacción y se analiza en la columna HPLC. Los índices de conversión de 5FC a 5FU se calculan al comparar las áreas pico con cantidades conocidas de una curva

estándar generada previamente de 5FU. El índice de conversión 5FC a 5FU se deriva al graficar la cantidad de 5FU (en nmol) generada contra su intervalo de tiempo correspondiente. La concentración de la proteína para la muestra celular se deriva y la Actividad Específica de las muestras de lisato celular se calcula al dividir el índice de conversión (nmol/min) mediante la cantidad de proteína utilizada en el ensayo en mg. La figura 6 muestra la actividad específica de varios vectores después de 5 días en transducción en un MOI de 0.1. Los datos demuestran que pACE-yCD (T5.0000) < pAC3-yCD1 (T5.0001) < pAC3-CD2 (T5.0002) en términos de la actividad específica de desaminasa citosina in en células de cultivo de tejido.

Ejemplo 7. Los tumores tratados con el gen CD completamente modificado (yCD2) se eliminan más eficientemente que el tumor tratado con el gen CD de levadura no modificada, y no recurre (PR-01-08-001).

[00165] Para determinar que construcción de vector surge para el vector más efectivo en un modelo de xenoinjerto de ratón/humano subcutáneo. Se evalúan tres construcciones diferentes: T5.0001 (CD parcialmente modificado); T5.0007 (gen CD de levadura no modificado); T5.0002 (gen yCD2 completamente modificado). El crecimiento del tumor, la supervivencia y la regresión del tumor se evalúa por post-administración de profármaco 5-FC en un modelo subcutáneo de glioma humano (U87) en ratones inmunodeficientes. Se evalúan dos concentraciones diferentes 5-FC para determinar

una respuesta de dosis entre 5-FC y las construcciones de tumor-vector.

[00166] se estudia un total de 12 grupos que consisten de 9-11
5 ratones hembra por grupo. Todos los ratones experimentan
implantación de flank dorsal derecho en el Día 0 con: una mezcla
de 98% de una estirpe celular de tumor U-87 no infectada y una
mezcla de 2% mixture de una estirpe celular U-87 que se ha
infectado con una de las tres construcciones de vector TOCA 511,
10 T5.0002, T5.0001, T5.0007; o una estirpe celular de control no
infectado U-87 (100%). Los ratones se inoculan con 2×10^6
células/ratón. Los ratones en los grupos 1-3 tienen células U87
no infectadas, los grupos 4-6 tienen mezclas de células U87 que
contienen los grupos transducidos T5.0002 U87, los grupos 7-9
15 tienen mezclas de célula U87 que contienen transducido con
T5.0001, y los grupos 10-12 tienen mezcla de célula U87 que
contiene transducido con T5.0007. A los tumores se les permite
crecer durante 6 días, hasta que el tamaño del tumor es
aproximadamente 100mm^3 . Cada grupo de dosis de vector de los
20 ratones se aleatoriza para recibir una de las dos dosis de 5-FC
(200 o 500 mg/kg/día), se administra como una inyección IP única,
empezando en el Día 6, o no 5-FC. La administración 5-FC continúa
diariamente durante 28 días consecutivos. Los ratones que
sobreviven después del día 29 se evalúan para el tamaño de tumor
25 y luego se sacrifican en varios días con tumores, si están
presentes, se extraen para análisis. En el día 29, los ratones
del grupo 4-12 se vuelven a aleatorizar y los ratones se
subdividen en 2 subgrupos (continúa en 5-FC para monitorear la

120

regresión del tumor o discontinúa el tratamiento 5-FC para monitorear el recrecimiento del tumor).

Resultados: Los tumores tratados con T5.0002, T5.0000 y T5.0007 demuestran todos regresión de tumor para no detectabilidad en 200 y 500 mg/kg dosis en el día 29. Todos los animales tratados T5.0002 no reforman los tumores cuando se discontinúa el tratamiento 5FC, fuera de 39 días. En contraste algunos tumores tratados con T5.0001 y T5.0007 recurren mediante el punto de tiempo del día 39. El experimento se termina en el día 39.

Conclusión: El T5.0002 más 5-FC es una terapia antitumor más efectiva que T5.0001 o T5.0007 más 5-FC.

[00167] Se ha descrito un número de realizaciones de la descripción. No obstante, se entenderá que se pueden hacer varias modificaciones sin apartarse del espíritu y alcance de la descripción. De acuerdo con lo anterior, otras realizaciones están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

LO QUE SE REIVINDICA ES:

5 1. Un polinucleótido aislado que comprende:

un polinucleótido optimizado de codón humano que
codifica un polipéptido de la SEQ ID NO:4, en donde el
polipéptido comprende actividad desaminasa citosina
10 (CD).

2. El polinucleótido aislado de la reivindicación 1, en donde
el polinucleótido comprende una secuencia como se establece
en la SEQ ID NO: 3 y que comprende 1-50 mutaciones
15 inactivas.

3. El polinucleótido aislado de la reivindicación 1 o 2, en
donde el polinucleótido comprende la SEQ ID NO: 3.

20 4. El polinucleótido aislado de la reivindicación 1, que
comprende adicionalmente un polinucleótido UPRT OPRT que
codifica un polipéptido que tiene actividad UPRT o OPRT
ligada operablemente al polipéptido que tiene actividad CD.

25 5. El polinucleótido aislado de la reivindicación 4, en donde
el polinucleótido UPRT OPRT es el codón optimizado.

122

6. El polinucleótido aislado de la reivindicación 4, en donde el polinucleótido UPRT OPRT se separa del polinucleótido que codifica el polipéptido que tiene actividad CD mediante un ligador.

5

7. Un polinucleótido aislado que comprende:

una primera secuencia codificante que codifica un polipéptido que tiene actividad desaminasa citosina; y

10

una segunda secuencia codificante que codifica un polipéptido que tiene actividad URPT o OPRT, en donde la primera secuencia codificante se liga operablemente a la segunda secuencia codificante.

15

8. El polinucleótido aislado de la reivindicación 7, en donde la primera y la segunda secuencias codificantes se separan mediante una secuencia que codifica un ligador de péptido.

- 20 9. El polinucleótido aislado de la reivindicación 7, en donde el polinucleótido codifica un polipéptido que comprende una secuencia como se establece en la SEQ ID N0:12, 14, o 16.

10. Un polipéptido sustancialmente purificado producido por la
25 expresión del polinucleótido de la reivindicación 1, 2, 4 o 7.

123

11. Un vector que comprende un polinucleótido de la reivindicación 1, 2, 4 o 7.
12. El vector de la reivindicación 11, en donde el vector es un plásmido.
13. El vector de la reivindicación 11, en donde el vector es un vector de expresión.
14. El vector de la reivindicación 11, en donde el vector comprende un vector vírico.
15. El vector de la reivindicación 14, en donde el vector vírico comprende un retrovirus de replicación competente.
16. Una célula anfitriona que contiene el vector de la reivindicación 12, 13, 14, o 15.
17. Una célula anfitriona que comprende un polinucleótido de la reivindicación 1, 4 o 7.
18. La célula anfitriona de la reivindicación 17, en donde la célula anfitriona comprende una célula humana.
19. La célula anfitriona de la reivindicación 18, en donde la célula anfitriona comprende una célula que comprende un trastorno proliferativo de la célula.

124

20. La célula anfitriona de la reivindicación 17, en donde la célula anfitriona comprende una célula humana establemente transformada.

5 21. La célula anfitriona de la reivindicación 20, en donde la célula anfitriona comprende un trastorno proliferativo de la célula.

10 22. La célula anfitriona de la reivindicación 19 o 21, en donde la célula anfitriona comprende glioblastoma multiforme.

23. La célula anfitriona de la reivindicación 17, en donde la célula anfitriona se transforma in vivo.

15 24. Un retrovirus competente de replicación recombinante (RCR) que comprende:

una proteína GAG retrovímica;

20 una proteína POL retrovímica;

una envoltura retrovímica;

25 un polinucleótido retrovímico que comprende Secuencias (LTR) de Repetición de Terminal Largo en el extremo 3' de la secuencia de polinucleótido retrovímico, una secuencia promotora en el extremo 5' de el polinucleótido retrovímico, dicho promotor es adecuado para la expresión en una célula

de mamífero, un dominio de ácido nucleico gag, un dominio de ácido nucleico pol y un dominio de ácido nucleico env;

un casete que comprende un sitio interno de entrada al
5 ribosoma (IRES) ligado operablemente a un ácido nucleico heterólogo que comprende un polinucleótido de la reivindicación 1, en donde el casete se posiciona 5' en el 3' LTR y 3' en el dominio de ácido nucleico env que codifica la envoltura retrovímica; y

10 las secuencias que actúan cis necesarias para transcripción inversa, empaque e integración en una célula objetivo.

25. El retrovirus de la reivindicación 24, en donde la secuencia
15 de polinucleótido retrovímico se deriva de virus de leucemia de murino (MLV), Virus de leucemia de murino de Moloney (MoMLV), Virus de leucemia felina o Virus de leucemia de mono Gibbon (GALV).

20 26. El retrovirus de la reivindicación 24 o 25, en donde el MLV es un MLV anfotrófico.

27. El retrovirus de la reivindicación 24, en donde el retrovirus es un retrovirus gama.

25 28. El retrovirus de la reivindicación 24, en donde la célula objetivo es una célula que tiene un trastorno proliferativo de la célula.

126

29. El retrovirus de la reivindicación 24, en donde la célula objetivo es una célula neoplásica.

5 30. El retrovirus de la reivindicación 28, en donde el trastorno proliferativo de la célula se selecciona del grupo que consiste de cáncer de pulmón, cáncer de colon-recto, cáncer de mama, cáncer de próstata, cáncer del tracto urinario, cáncer uterino, cáncer de cerebro, cáncer de cabeza y
10 cuello, cáncer pancreático, melanoma, cáncer de estómago y cáncer de ovario, artritis reumatoide u otras enfermedades autoinmunes.

15 31. El retrovirus de la reivindicación 24, en donde la secuencia promotora se asocia con un gen regulador del crecimiento.

32. El retrovirus de acuerdo con la reivindicación 24, en donde la secuencia promotora comprende una secuencia promotora específica de tejido.

20 33. El retrovirus de acuerdo con la reivindicación 32, en donde la secuencia promotora específica de tejido comprende por lo menos un elemento de respuesta a andrógeno (ARE).

25 34. El retrovirus de acuerdo con la reivindicación 33, en donde el elemento de respuesta a andrógeno se deriva del promotor probasina.

127

35. El retrovirus de acuerdo con la reivindicación 32, en donde la secuencia promotora específica de tejido comprende un promotor probasina.

5 36. El retrovirus de acuerdo con la reivindicación 24, en donde el promotor comprende un promotor CMV que tiene una secuencia como se establece en la SEQ ID NO: 19, 20, o 22 a partir del nucleótido 1 a cerca del nucleótido 582 y puede incluir la modificación en una o más bases de ácido nucleico y que es capaz de dirigir e iniciar la transcripción .

10

37. El retrovirus de la reivindicación 24, en donde el promotor comprende una secuencia como se establece en la SEQ ID NO: 19 o 20 a partir del nucleótido 1 a cerca del nucleótido

15 582.

38. El retrovirus de la reivindicación 24, en donde el promotor comprende un polinucleótido de dominio CMV-R-U5.

20 39. El retrovirus de la reivindicación 38, en donde el dominio CMV-R-U5 comprende el promotor inmediatamente temprano del citomegalovirus humano ligado a una región MLV R-U5.

25 40. El retrovirus de la reivindicación 39, en donde el polinucleótido de dominio CMV-R-U5 comprende una secuencia como se establece en la SEQ ID NO: 19, 20, o 22 desde cerca del nucleótido 1 a cerca del nucleótido 1202 o las secuencias que son por lo menos 95% idénticas a una

secuencia como se establece en la SEQ ID NO: 19, 20, o 22, en donde el polinucleótido promueve la transcripción de una molécula de ácido nucleico ligada operablemente a esta.

- 5 41. El retrovirus de la reivindicación 24, en donde the gag del polinucleótido se deriva de un retrovirus gama.
- 10 42. El retrovirus de la reivindicación 41, en donde el dominio de ácido nucleico gag comprende una secuencia desde cerca del nucleótido número 1203 a cerca del nucleótido 2819 de la SEQ ID NO: 19 o una secuencia que tiene por lo menos 95%, 98%, 99% o 99.8% de identidad de esta.
- 15 43. El retrovirus de la reivindicación 24, en donde el dominio pol del polinucleótido se deriva de un retrovirus gama.
- 20 44. El retrovirus de la reivindicación 43, en donde el dominio pol comprende una secuencia desde cerca del nucleótido número 2820 a cerca del nucleótido 6358 de la SEQ ID NO: 19 o una secuencia que tiene por lo menos 95%, 98%, 99% o 99.9% de identidad de esta.
- 25 45. El retrovirus de la reivindicación 24, en donde el dominio env comprende una secuencia desde cerca del nucleótido número 6359 a cerca del nucleótido 8323 de la SEQ ID NO:19 o una secuencia que tiene por lo menos 95%, 98%, 99% o 99.8% de identidad de esta.

46. El retrovirus de la reivindicación 24, en donde el IRES se deriva de un virus de encefalomiocarditis.
47. El retrovirus de la reivindicación 46, en donde el IRES
5 comprende una secuencia desde cerca del nucleótido número 8327 a cerca del nucleótido 8876 de la SEQ ID NO:19 o una secuencia que tiene por lo menos 95%, 98%, o 99% de identidad de esta.
- 10 48. El retrovirus de la reivindicación 24, en donde el ácido nucleico heterólogo comprende un polinucleótido que tiene una secuencia como se establece en la SEQ ID NO:3, 5, 11, 13, 15 o 17.
- 15 49. El retrovirus de la reivindicación 24, en donde el ácido nucleico heterólogo codifica un polipéptido que comprende una secuencia como se establece en la SEQ ID NO:4.
- 20 50. El retrovirus de la reivindicación 24, en donde el ácido nucleico heterólogo es codón humano optimizado y codifica un polipéptido como se establece en la SEQ ID NO: 4.
- 25 51. El retrovirus de la reivindicación 24, en donde el ácido nucleico heterólogo comprende una secuencia como se establece en la SEQ ID NO: 19 desde cerca del nucleótido número 8877 a cerca de 9353.

52. El retrovirus de la reivindicación 51, en donde el LTR 3' se deriva de un retrovirus gama.

53. El retrovirus de la reivindicación 52, en donde el LTR 3' comprende un dominio U3-R-U5.

54. El retrovirus de la reivindicación 52, en donde el LTR 3' comprende una secuencia como se establece en la SEQ ID NO: 19 desde cerca del nucleótido 9405 a cerca de 9998 o una secuencia que es por lo menos 95%, 98% o 99.5% idénticas a esta.

55. El retrovirus de la reivindicación 24, en donde el polinucleótido retrovítico comprende una secuencia como se establece en la SEQ ID NO: 19, 20, o 22.

56. Un polinucleótido aislado que comprende de 5' a 3':

una fusión CMV-R-U5 del promotor inmediato temprano del citomegalovirus humano con una región MLV R-U5;

un PBS, el sitio de unión de cebador para transcriptasa inversa;

un sitio de división 5';

señal de empaque ψ ;

131

una secuencia codificante gag para el antígeno específico del grupo MLV;

5 una secuencia codificante pol para la poliproteína de polimerasa MLV;

un sitio de división 3';

10 una secuencia codificante env 4070A para la proteína de envoltura de la cepa MLV 407 OA;

un sitio interno de entrada al ribosoma (IRES) del virus de encefalomiocarditis;

15 una secuencia codificante desaminasa citosina modificada;

un tracto polipurina; y

20 una repetición de terminal largo U3-R-U5 MLV.

57. El polinucleótido de la reivindicación 56, en donde el polinucleótido retrovítico comprende una secuencia como se establece en la SEQ ID NO: 19, 20, o 22.

25 58. El polinucleótido de la reivindicación 56, en donde la secuencia de polinucleótido retrovítico se deriva de virus de leucemia de murino (MLV), Virus de leucemia de murino de

132

Moloney (MoMLV), Virus de leucemia felina o Virus de leucemia de mono Gibbon (GALV).

59. El polinucleótido de la reivindicación 58, en donde el MoMLV
5 es un MoMLV amfotrófico.

60. Un método para tratar un sujeto con un trastorno
proliferativo de la célula que comprende poner en contacto
el sujeto con un polinucleótido de la reivindicación 1, 7 o
10 56 bajo condiciones de tal manera que el polinucleótido se
expresa y se pone en contacto el sujeto con 5-
fluorocitosina.

61. El método de la reivindicación 60, en donde el
15 polinucleótido se integra dentro de la célula del sujeto.

62. El método de la reivindicación 60, en donde el
polinucleótido se suministra mediante un vector retrovílico.

20 63. El método de la reivindicación 62, en donde el vector
retrovílico comprende una secuencia como se establece en la
SEQ ID NO: 19, 20, o 22.

64. El método de la reivindicación 60, en donde el trastorno
25 proliferativo de la célula es glioblasoma multiforme.

65. El método de la reivindicación 60, en donde el trastorno
proliferativo de la célula se selecciona del grupo que

consiste de cáncer de pulmón, cáncer de colon-recto, cáncer de mama, cáncer de próstata, cáncer del tracto urinario, cáncer uterino, cáncer de cerebro, cáncer de cabeza y cuello, cáncer pancreático, melanoma, cáncer de estómago y
5 cáncer de ovario, artritis reumatoide u otras enfermedades autoinmunes.

66. Un método para tratar un trastorno proliferativo de la célula en un sujeto que comprende poner en contacto el
10 sujeto con un retrovirus de la reivindicación 24, y poner en contacto el sujeto con 5-fluorocitosina.

67. El método de la reivindicación 66, en donde el ácido nucleico heterólogo codifica un polipéptido que comprende
15 una secuencia como se establece en la SEQ ID NO: 4, 12, 14, 16, o 18.

68. El método de la reivindicación 66, en donde el polinucleótido retrovítico comprende una secuencia como se
20 establece en la SEQ ID NO: 19, 20, o 22.

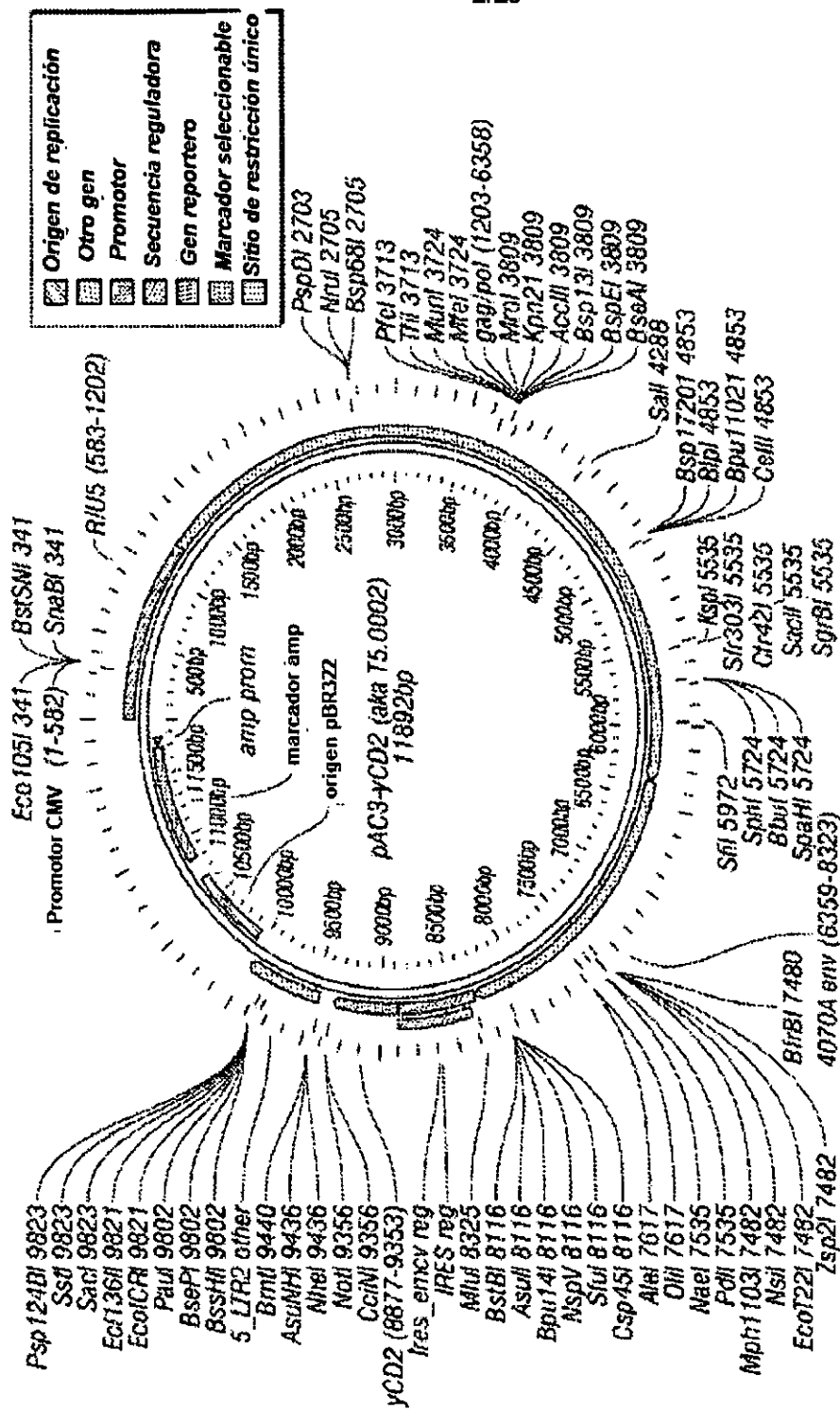


FIGURA 1C

136

7/29

Promotor CMV (1-582)>>>

1 TAGTATTAAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG 60
ATCAATAAATATCATTAGTTAATGCCCCAGTAATCAASTATCGGGTATATACCTCAAGGC

61 CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCCTT 120
GCAATGTATTGAATGCCAATTACCGGGCGGACCGACTGGCGGGTTGCTGGGGGGGGTAA

121 GACGTCATAATGACGTATGTTCCCATAGTAAAGCCAAATAGGGACTTTCCATTGACGTCA 180
CTGCAGTTATTACTGCATACAAGGGTATCATTCCGGTTATCCCTGAAAGGTAACTGCAGT

181 ATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC 240
TACCCACCTCATAAATGCCAATTGACCGGTGAACCGTCATGTAGTTTACATAGTATACGG

241 AAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTA 300
TTCATGCGGGGGATAACTGCAGTTACTGGCCATTTACCGGGCGGACCGTAATACGGGTGAT

Eco105I
|
SnaBI
|
BstSI
|

301 CATGACCTTATGGGACTTTCTTACTTGGCAGTACATCTACGTAATTAGTCATCGCTATTAC 360
GTACTGGAATACCTGAAAGGATGAACCGTCAITGTAGATGCATAATCAGTAGCGATAATG

361 CATGGTGATGGGTTTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGG 420
GTACCACTACGCCAAAACCGTCAATAGTGTACCGGCACCTATCGCCAAACTGAGTGCCCC

421 ATTTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG 480
TAAAGGTTTCAAGGTTGGGGTAACTGCAGTTACCTCAAAACAAAACCGTGGTTTTAGTTGC

481 GGACTTTCCAAAATGTCTAACAACCTCCGCCCCATTGACGCAATGGGCGGTAGGCGTGT 540
CCTGAAAGGTTTTACAGCATTGTTGAGGCGGGTAACTGCGTTTACCCGCCATCCGCACA

R region (583-550)>>>

541 ACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTACTGAACCGGCGCCAGTCTCCCATTTG 600
TGCCACCCCTCCAGATATATTGCTCTGACCAAATCACTGGCCCGGGTCAGGAGGGTAAC

FIGURA 1D-4

HOJA SUSTITUTA (REGLA 26)

141

8/29

región U5 (651-1202)>>>

601 ACTGAGTCGGCCGGGTACCCGGTATCCAAATAACCCCTCTGCAGTTGCATCCGACTTGT 660
TCACTCAGCGGGCCCATGGGACATAGGTATTTGGGAGAACGTCACGTAGGCTGAACA
661 GGTCTCCCTGTTCTTGGGAGGGTCTCTCTGAGTGAATGACTACCCGTCAGCGGGGTC 720
CCAGAGCGACAAAGGAACCCCTCCAGAGGAGACTCACTAAGTATGGGCAGTCGCCCCCAG
721 TTTCATTTGGGGGCTCTCTCGGGATCGGGAGACCCCTGCCAGGGACCAACGACCCACCA 780
AAAGTAACCCCGAGCAGGCCCTAGCCCTCTGGGGAAGGTCCTCTGGTGGCTGGGTGGT
5'SS (738)
781 CCGGGAGGTAAAGCTGGCCAGCAACTTATCTGTGTCTGTCGATTGTCTAGTGTCTATGAC 840
GGCCCTCCATTCGACCGGTCTTGAATAGACACAGACAGGCTAACAGATCAGATACCTG
841 TGATTTTATGGCCCTGGGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCG 900

FIGURA 1D-5

142

9/29

ACTAAATACGGGACGCGAGCCATGATCAATCGATTGATCGAGACATAGACCGCCTGGGC

901 TGGTGGAACTGACGAGTTGCGACACCCGCGCGCAACCCCTGGGAGACGTCCCAGGGACTT 960
ACCACCTTGACTGCTCAAGCCTTGTGGGCGGGCTTGGGACCCTCTGCAGGCTCCCTGAA

961 CCGGGGCGCTTTTGTGCCCCGACCTGAGTCCAAAAATCCCGATCGTTTGGACTCTTTG 1020
GCCCCCGGAAAAACACCGGGCTGGACTCAGGTTTATAGGGCTAGCAAAACCTGAGAAC

1021 CTGCACCCCCCTTAGAGSAGGGATATGTGTTCTGCTAGGAGACGAGAACCTAAACAGT 1080
CACGTGGGGGAATCTCCTCCCTATACACCAAGACCATCTCTGCTCTTGGATTTTGTCA

1031 TCCCGCCTCGTCTGAATTTTGGCTTTCGGTTTGGGACCGAAGCCGCGCGCGCGCTCTTG 1140
AGGGCGGAGGCAGACTTAAAAACGAAGCCAAACCTGGCTTCGGCGCGCGCGCAGAAC

1141 TCTGCTGCAGCATCGTCTCTGTGTGTCTGTCTGACTGTCTTCTGTATTTGTCTGAGA 1200
AGACGAGCTGTAGCAAGACACACACAGACAGACTGACACAAAGACATAACAGACTCT

gag(1203, 2819) >>>

1201 ATATGGGCCAGACTGTACCACTCCCTTAAGTTTGACCTTAGGTCACTGGAAAGATGTG 1260
TATACCCGCTCTGACAAATGGT GAGGGAATTCAAAC TGGATCCAGTGACCTTCTACAGC

1261 AGCGGATCGCTCACAACCACTCGCTAGATGTCAAGAAGAGACGTTGGGTACCTTCTGCT 1320
TCGCCCTAGCGAGTGTGGTLAGCCATCTACAGTTCTTCTCTGCAACCAATGGAAGACGA

1321 CTGCAGAATGGCCAACCTTTAACGTCGGATGGCGCGAGACGGCACCTTTAACCGAGACC 1380
GACGTCTTACCGGTTGGAAATTCAGCCTACCGGCGCTCTGCCGTGGAAATTTGGCTCTGG

1381 TCATCACCCAGGTTAAGATCAAGGTCTTTTCACCTGGCCCGCATGGACACCCAGACCAGG 1440
AGTAGTGGGTCCAAATTTAGTTTCAGAAAAGTGGACCGGGCGTACCTGTGGGTCTGGTCC

1441 TCCCTACATCGTGACTGGGAAGCCTTGGCTTTTGACCCCTCTCCCTGGGTCAAGCCCT 1500
AGGGGATGTAGCACTGGACCTTGGGAACCGAAACTGGGGGAGGGACCCAGTTCGGGA

1501 TTGTACACCTTAAGCCTCCGCTCCTCTTCCATCCGCCCCGTCTCTCCCCCTTGAAC 1560
AACATGTGGGATTCGGAGCGGAGGAGAAGGAGGTAGGCGGGCAGAGAGGGGAACTTG

1561 CTCTCTGTTGACCCCCGCTCGATCCTCCCTTTATCCAGCCCTCACTCCTTCTCTAGGGG 1620
GAGGAGCAAGCTGGGGCGGAGCTAGGAGGGAAATAGGTGGGAGTGAGGAAGAGATCCGC

1621 CCAAACTTAACCTCAAGTTCTTTCTGACAGTGGGGGGCGCTCATOGACCTACTTACAG 1680
GGTTTGGATTTCAGGATTCAGAAAGACTGTCAACCCCGGCGAGTAGCTGGATGAATGTC

1681 AAGACCCCGGCTTATAGGAGCCCAAGACCACCCCTTCCGACAGGGACGGAATGGTG 1740
TTCTGGGGGGCGGAATATCCCTGGGTCTGTGTGGGGAAGGCTGTCCCTGCCCTTTACCAC

FIGURA 1D-6

10/29

1741 GAGAAGCGACCCCTGCGGGAGAGGCACCGGACCCCTCCCCAATGGCATCTCGCCTACGTG 1800
 CTCTTCGCTGGGGACGCCCTCTCCGTGGCCGCGGGAGGGGTTACCGTAGAGCGGATGCAC
 1801 GGAGACGGGAGCCCCCTGTGGCGGACTCCACTACCTCGCAGGCATTCGCCCTCCGGGCAG 1860
 CCTCTGCCCTCGGGGGACACCGGCTGAGGTGATGGAGCGTCCGTAAGGGGGAGGCGCGTC
 1861 GAGGAAACGGACAGCTTCAATACTGGCCGTTCTCTCTTCTGACCTTTACAACCTGAAAA 1920
 CTCCTTTGCCTGTGGAAGTTATGACCGGCAAGAGGAGAAGACTGGAATGTTGACCTTTT
 1921 ATAATAACCCCTTCTTTTTTCTGAAGATCCAGGTAAACTGACAGCTCTGATCGAGTCTGTTC 1980
 TATTATTGGGAAGAAAAAGACTTCTAGGTCCATTTGACTGTGAGACTAGCTCAGACAAG
 1981 TCATCACCCATCAGCCACCTGGGACGACTGTCAGCAGCTGTTGGGGACTCTGCTGACCG 2040
 AGTAGTGGGTAGTGGGTGGACCCCTGCTGACAGTGGTGGACAACCCCTGAGACGACTGGC
 3121 ACTTTGAGGGATCAGGAGCCACGTTATGGGACCAATGGGGCAGCCCCCTGCAAGTGTGA 3180
 TGAAACTCCCTAGTCCCTCGGGTCCAATACCCCTGGTTACCCCGTCGGGACGTTACACACT
 3181 CCCTAAATATAGAAGATGAGCATCGGCTACATGACACCTCAAAAGAGCCAGATGTTTCTC 3240
 GGGATTTATATCTTCTACTCGTAGCCGATGTACTCTGGAGTTTCTCGGTCTACAAAGAG
 3241 TAGGGTCCACATGGCTGTCTGATTTTCTCTAGGCTGGGCGGAACCGGGGGCATGGGAC 3300
 ATCCCAAGGTGTACCGACAGACTAAAGGAGTCCGACCCGCCCTTTGGCCCCCGTACCCCTG
 3' SS (3314)
 3301 TGGCAGTTGCCAAGCTCTCTGATCATACCTCTGAAAGCAACCTCTACCCCCGTGTCCA 3360
 ACCGTCAAGCGGTTCAGGAGACTAGTATGGAGACTTTCTGTTGGAGATGGGGGCACAGGT
 3361 TAAACAATACCCCATGTCAAGAAGCCAGACTGGGGATCAAGCCCCACATACAGAGAC 3420
 ATTTTGTATGGGTACAGTGTCTCTCGGTCTGACCCCTAGTTCCGGGTGTATGTCTCTG
 3421 TGTTGGACAGGGAATACTGGTACCCCTGCCAGTCCCCCTGGAACAGCCCCCTGCTACCCG 3480
 ACAACCTGGTCCCTTATGACCATGGGACGGTCAGGGGGACCTTGTGCGGGGACGATGGGC
 3481 TTAAGAAACCAGGGACTAATGATTATAGGCCTGTCCAGGATCTGAGAGAGTCAACAAGC 3540
 AATTCTTTGGTCCCTGATTACTAATATCCGGACAGGTCTAGACTCTCTTCAGTTGTTCG
 3541 GGGTGGAGACATCCACCCACCGTGCCCAACCCCTTACAACTCTTGAGCGGGCTCCAC 3600
 CCCACCTTCTGTAGGTGGGGTGGCAGGGGTGGGAATGTTGGAGAACTCGCCCGAGGGTG
 3601 CGTCCACACAGTGGTACACTGTCTTGAATTAAGGATGCCCTTTTCTGCCTGAGACTCC 3660
 GCAGGGTGGTCAACATGTGACACGAACTAAATTTCTACGGAAAAAGACGGACTCTGAGG
 Pfl
 |
 Tfl
 |
 3661 ACCCAACAGTCAGCCTCTCTTCGCCCTTGAGTGGAGAGATCCAGAGATGGGAATCTCAG 3720
 TGGGGTGGTCACTCGGAGAGAGCGGAAACTCACCTCTCTAGGTCTCTACCCCTAGAGTC

FIGURA 1D-7

HOJA SUSTITUTA (REGLA 26)

144

11/29

MfeI
 |
 MunI
 |

3721 GACAAATTGACCTGGACCCAGACTCCACAGGGTTTCAAAAACAGTCCACCCCTGTTTGATG 3780
 CTGTTAACTGGACCTGGTCTGAGGGTGTCCAAAGTTTGTGTCAGGGTGGGACAACTAC

MroI
 |
 BseAI
 |
 Bsp13I
 |
 BspEI
 |
 Kpn2I
 |
 AccIII
 |

3781 AGGCACTGCACAGAGACCTAGCGACTTCCGGATCCAGCACCCAGACTTGATCCTGCTAC 3840
 TCCGTGACGTGTCTCTGGATCGTCTGAAGGCCTAGGTGCTGGGTCTGAAGTACGACGATG

3841 AGTACGTGGATGACTTACTGCTGGCCGCCACTTCTGAGCTAGACTGCCAACAGGGTACTC 3900
 TCATGCACCTACTGATGACGACCGCGGTGAAGACTCGATCTGACGGTTGTTCCATGAG

3121 ACTTTGAGGGATCAGGAGCCAGGTTATGGGACCAATGGGGCAGCCCTGCAAGTCTTGA 3180
 TGAAGCTCCCTAGTCTCTCGGGTCCAATACCCCTGGTTACCCCGTCCGGGACGTTACAACT

3181 CCTTAAATATAGAAGATGAGCATCGGCTACATGAGACCTCAAAAGAGCCAGATGTTTCTC 3240
 GGGATTTATATCTTCTACTCGTAGCCGATGTACTCTGAGTTTTCTCGGTCTACAAAGAG

3241 TAGGGTCCACATGCTGTCTGATTTTCTCAGGCTGGGGCGAAACCGGGGGCATGGGAC 3300
 ATUCCAGGTGTACGACAGACTAAAGGAGTCCGACCCGCTTTGGCCCCCGTACCCCTG

3' SS (3314)
 |

3301 TGGCAGTTCGCCAAGCTCCTCTGATCATACTCTGAAAGCAACCTCTACCCCGTGTCCA 3360
 ACCGTCAAGCGGTTCGAGGAGACTAGTATGGAGACTTTCGTTGGAGATGGGGGCACAGGT

3361 TAAAACAATACCCCATGTCAAGAGAGCCAGACTGGGGATCAAGCCCCACATACAGAGAC 3420
 ATTTTGTATGGGGTACAGTGTCTTCGGTCTGACCCCTAGTTCCGGGTGTATGTCTCTG

3421 TGTGGACCCAGGGAATACTGGTACCCCTGCCAGTCCCCCTGGACACGCCCCCTGCTACCCG 3480
 ACAACCTGGTCCCTTATGACCATGGGACGGTCAAGGGGACCTTGTGCGGGGACGATGGGC

3481 TTAAGAAACCCAGGACTAATGATTATAGGCTGTCCAGGATCTGAGAGAGTCAACAAGC 3540
 AATTCTTTGGTCCCTGATTACTAATATCCGGACAGGTCCCTAGACTCTCTTCAGTTGTTG

FIGURA 1D-8

HOJA SUSTITUTA (REGLA 26)

145

12/29

3541 GGGTGGGAAGACATCCACCCCAACCGTGCCCAACCCCTTACAACTCTTGAGCGGGCTCCAC 3600
CCCACTTCTGTAGGTGGGTGGCAGGGGTGGGAATGTTGGGAAGTCCGCCGAGGGTG

3601 CGTCCCACCACTGGTACACTGTGCTTGATTTAAAGGATGCCTTTTCTGCCTGAGACTCC 3660
GCAGGGTGGTCACCATGTGACACGAACTAAATTTCTACGGAAGAACGAGACTCTGAGG

PfeI
|
TliI
|

3661 ACCCCACCACTCAGCCTCTCTTCGCCCTTGAGTGGAGAGATCCAGAGATGGGAATCTCAG 3720
TGGGGTGGTCAGTCGGAGAGAGCGGAACTCACCTCTCTAGGTCTCTACCCCTAGAGTC

MfeI
|
MunI
|

3721 GACAATTGACCTGGACCCAGACTCCACAGGGTTTCAAAAACAGTCCACCCCTGTTTGATG 3780
CTGTTAACTGACCTGGTCTGAGGGTGTCCTCAAGCTTTTGTGAGGGTGGGACAACTAC

MroI
|
BseAI
|
Bsp13I
|
BspEI
|
Kpn2I
|
AccIII
|

3781 AGGCACTGCACAGAGACCTAGCAGACTTCGGATCCAGCAACCAGACTTGATCCTGCTAC 3840
TCCGTGACGTGTCTCTGGATCGTCTGAAGGCCTAGGTCTGGGTCTGAAGTAGGACGATG

3841 AGTACGTGGATGACTTACTGCTGGCCGCCACTTCTGAGCTAGACTGCCAACAGGTACTC 3900
TCATGCACCTACTGAATGACGACGGCGGTGAAGACTCGATCTGACGGTTGTTCCATGAG

FIGURA 1D-9

HOJA SUSTITUTA (REGLA 26)

146

13/29

3901 GGGCCCTGTTACAAACCCCTAGGGAACTCCTGGGTATCGGGCCTCGGCCAAGAAAGCCCAA 3960
 CCGGGGACAAATGTTTGGGATCCCTTGGAGCCCATAGCCCGGAGCCGGTTCTTTCGGGTTT
 3961 TTTGCCAGAAACAGGTCAAGTATCTGGGGTATCTTCTAAAAGAGGGTCAGAGATGGCTGA 4020
 AACCGGTCTTTGTCCAGTTCATAGACCCCATAGAAGATTTCTCCTCAGTCTCTACCGACT
 4021 CTGAGGCCAGAAAGAGACTGTGATGGGGCAGCCTACTCCGAAGACCCCTCGACAACTAA 4080
 GACTCCGGTCTTTTCTCTGACACTACCCCGTCGGATGAGGCTTCTGGGGAGCTGTTGATT
 4081 GGGAGTTCTTAGGGACGGCAGGCTTCTGTCTGCTCTGGATCCCTGGGTTTSCAGAAATGG 4140
 CCTCAAGGATCCCTGCCGTCCGAAGACAGCGGAGACCTAGGGACCCAAACGTCTTTACC
 4141 CAGCCCCCTTGTACCTCTCACCAAAACGGGGACTCTGTTTAATTGGGGCCAGACCAAC 4200
 GTCGGGGGAACATGGGAGAGTGGTTTGGCCCTGAGACAAATTAACCCCGGGTCTGTTG
 4201 AAAAGGCCATCAAGAAATCAAGCAAGCTCTTCTAACTGCCCCAGCCCTGGGGTTGCCAG 4260
 TTTCCGGATAGTTCTTTAGTTTCGTTCGAGAAGATTGACGGGTCTGGGACCCCAACGGTC
 Sal I
 |
 4261 ATTTGACTAAGCCCTTTGAACCTCTTGTGACGAGAGAAGCAAGGCTACGCCAAGGTGTCC 4320
 TAAACTGATTCGGGAAACTTGAGAAACAGCTGCTCTTCGTCCCGATGCGGTTTCCACAGG
 4321 TAACGCAAAAACCTGGGACCTTGGCGTCGGCCGGTGCCCTACCTGTCCAAAAGCTAGACC 4380
 ATTGCGTTTTTTGACCCCTGGAAACCGCAGCCGGCCACCGGATGGACAGGTTTTTCGATCTGG
 4381 CAGTAGCAGCTGGGTGGCCCCCTTGCCTACGGATGGTAGCAGCCATTGCCGTACTGACAA 4440
 GTCATCTGTCGACCCACCGGGGGAACGGATGCCCTACCATCGTCGGTAACGGCATGACTGTT
 4441 AGGATGCAGGCAAGCTAACCATGGGACAGCCACTAGTCATTCTGGCCCCCATGCAGTAG 4500
 TCCTACGTCCGTTTCGATTGGTACCCGTGCGGTGATCAGTAAGACCAGGGGGTACGTCATC
 4501 AGGCACTAGTCAAACAACCCCGGACCGCTGGCTTTCCAACGCCCGGATGACTCACTATC 4560
 TCCGTGATCAGTTTGTGGGGGGCTGGCGACCGAAAGGTTGCGGGCTACTGAGTGATAG
 4561 AGGCCTTGCTTTTGGACACGGACCGGGTCCAGTTCCGACCGGTGGTAGCCCTGAACCGG 4620
 TCCGGAACGAAACCTGTCCCTGGCCAGGTCAAGCCTGSCCACCATCGGGACTTGGGCC
 4621 CTACGCTGCTCCCACTGCTGAGGAAGGGCTGCAACACAACCTGCTTGATATCCTGGCCG 4680
 GATGCGACGAGGCTGACGGACTCCTTCCCGACGTTGTGTTGACGGAACTATAGGACCGGC
 4681 AAGCCACGGAAACCGACCCGACCTAACGGACGAGCCGCTCCAGACCGGACCAACCT 4740
 TTGGGGTGCCCTTGGGCTGGGCTGGATTGCTGGTCGGGAGGGTCTGGGCTGGTGTGGA
 4741 GGTACACGGATGGAGCGTCTCTTACAAGAGGGACAGCGTAAGGCGGGAGCTGCGGTGA 4800
 CCATGTGCTACCTTCGTCAGAGATGTTCTCCCTGTCGCATTCCGCCCTCGACGCCACT

FIGURA 1D-10

HOJA SUSTITUTA (REGLA 26)

147

14/29

BlpI
|
CelIII
|
Bpu1102I
|
Bsp1720I
|

4801 CCACCGAGACCGAGGTAACTCTGGGCTAAAGCCCTGCCAGCCGGGACATCCGCTCAGCGGG 4860
GGTGGCTCTGGCTCCATTAGACCCGATTTCGGGACGGTCGGCCCTGTAGGCGAGTCGCC

4861 CTGAAGTGTATAGCACTCAACCCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCTAAATGTTT 4920
GACTTGACTATCGTGAGTGGGTCCGGGATTTCTACCGTCTTCCATTCCTCGATTTACAAA

4921 ATACTGATAGCCGTTATGCTTTTGGCTACTGCCCATATCCATGGAGAAATATACAGAAGGC 4980
TATGACTATCGGCAATACGAAAACGATGACGGGTATAGGTACCTCTTTATATGCTCTCCG

4981 GTGGGTTGCTCACATCAGAAGGCCAAAGAGATCAAAATAAAGACGAGATCTTGGCCCTAC 5040
CACCAACGAGTGTAGTCTTCCGTTCTCTAGTTTATTTCTGCTCTAGAACCGGGATG

5041 TAAAAGCCCTCTTTCGCCCAAAAGACTTAGCATAATCCATTGTCCAGGACATCAAAAGG 5100
ATTTTCGGGAGAAAGACGGGTTTTCTGAATCGTATTAGGTAAAGGTCCCTGTAGTTTTCC

5101 GACACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGGCCGAAAGGCAGCCA 5160
CTGTGTCGGCGCTCCGATCTCCGTTGGCCTACCGACTGGTTGCGCGGGCTTTCCGTCGGT

5161 TCACAGAGACTCCAGACACCTCTACCTCTCTATAGAAAATTCTACCCCTACACCTCAG 5220
AGTGTCTCTGAGGTCTGTGGAGATGGGAGGAGTATCTTTAAGTAGTGGGATGTGGAGTC

5221 AACATTTTCATTACACAGTGAATATAAAGGACCTAACCAAGTTGGGGGCCATTTATG 5280
TTGTAAAAGTAATGTGTCACTGACTATATTTCCCTGGATTGGTTCAACCCCGGTAATAC

5281 ATAAACAAAGAAGTATTGGGTCTACCAAGGAAACCTGTGATGCCCTGACAGTTTACTT 5340
TATTTTGTTCCTCATAAACCAGATGGTTCCCTTTGGACACTACGGACTGGTCAAATGAA

5341 TTGAATTATTAGACTTTCTTCATCAGCTGACTCACCTCAGCTTCTCAAAAATGAAGGCTC 5400
AACTTAATAATCTGAAAGAAGTAGTCGACTGAGTGGAGTCGAAGAGTTTTTACTTCCGAG

5401 TCCTAGAGAGAAGCCACAGTCCCTACTACATGCTGAACCGGGATCGAAGACTCAAAAATA 5460
AGGATCTCTCTCGGTGTCAAGGATGATGTACGACTTGGCCCTAGCTTGTGAGTTTTTAT

FIGURA1D-11

15/29

5461 TCACTGAGACCTGCAAAGCTTGTGCAAGTCAACGCCAGCAAGTCTGCCGTTAAACAGG 5520
 AGTGACTCTGGACGTTTCGAACACGTGTTCAAGTTGCCGTCGTTACAGACGCAATTTGTCC

SacII
 |
 SgrBI
 |
 Cfr42I
 |
 Sfr303I
 |
 KspI
 |

5521 GAACTAGGGTCCGCGGGCATCGGCCGGCACTCATTTGGGAGATCGATTTCACCGAGATAA 5580
 CTTGATCCCAGGCCCGGTAGCCGGGCCGTGAGTAACCTCTAGCTAAAGTGGCTCTATT

5581 AGCCCGGATTGTATGGCTATAAATATCTTCTAGTTTTATAGATACCTTTCTGGCTGGA 5640
 TCGGGCTTAACATACCGATATTTATAGAAGATCAAAAATATCTATGGAAAAGACCGACCT

5641 TAGAAGCCTTCCCAACCAAGAAGAAACCGCCAGGTCGTAACCAAGAAGCTACTAGAGG 5700
 ATCTTCGGAGGGTTGGTTCTTTCTTGGCGGTTCCAGCATTGGTTCTTCGATGATCTCC

PaeI
 |
 EbuI
 |
 SpaHI
 |
 SphI
 |

5701 AGATCTTCCCCAGGTTCCGGCATGCCTCAGGTATTGGGAACGACAATGGGCCCTGCCCTTCG 5760
 TCTAGAAGGGGTCCAAGCCGTACGGAGTCCATAACCTTGACTGTTACCCGGACGGAAGC

6901 CTTCCAGGGGCTACTCGAGGGGGCAGATGCAACCCCTCTAGTCTAGAAATTCAGTGTATGC 6960
 GAAGGTTCCCGATGAGCTCCCCCGTCTACGTTGGGAGATCAGGATCTTAGTGACTACG

6961 AGGAAAAAGGCTAACTGGGACGGGCCAAATCGTGGGGACTGAGACTGTACCGGACAGG 7020
 TCCTTTTTTCCGATTGACCTGCCCGGTTTTCAGCACCCCTGACTCTGACATGGCCTGTCC

7021 AACAGATCCTATTACCATGTTCTCCCTGACCCGGCAGGTCTTAATGTGGGACCCCGAGT 7080
 TTGTC TAGGATAATGGTACAAGAGGGACTGGGCCGTCCAGGAATTACACCTGGGGCTCA

7081 CCCCATAGGGCCCAACCCAGTATTACCCGACCAAGACTCCCTTCTCAACCAATAGAGAT 7140
 GGGGTATCCCGGTTGGGTCAATCGGCTGGTTCTGAGGGAAGGAGTGGTTATCTCTA

7141 TGTACCGGCTCCACAGCCACCTAGCCCCCTCAATACAGTTACCCCCCTTCCACTACCAG 7200
 ACNTGGCCGAGGTGTGGGTGGATCGGGGAGTTATGGTCAATGGGGGAAGGTGATGGTC

7201 TACACCTCAACCTCCCTACAGTCCAGTGTCCACAGCCACCCCGAGGACTGGAGA 7260
 ATGTGGGAGTTGAGGGGATGTTCAAGTTACAGGGGTGTGGGTGGGGGTCTTGACCTCT

7261 TAGACTACTAGCTCTAGTCAAAGGAGCCTATCAGGCGCTTAACCTCACCATCCCGACAA 7320
 ATCTGATGATCGAGATCAGTTTCCTCGGATAGTCCCGAATTGGAGTGGTTAGGGCTGTT

FIGURA1D-12

HOJA SUSTITUTA (REGLA 26)

149

16/29

7321 GACCCAAGAATGTTGGCTGTGCTTAGTGTCGGGACCTCCTTATTACGAAGGAGTAGCGGT 7380
CTGGGTTCCTTACAACCGACACGAATCACAGCCCTGGAGGAATAATGCTTCCTCATCGCCA

7381 CGTGGGCACCTTATACCAATCATTCCACCGCTCCGGCCAACTGTACGGCCACTTCCCAACA 7440
GCACCCGTGAATATGGTTAGTAAGGTGGCGAGGCCGGTTGACATGCCGGTGAAGGGTTGT

EcoT22I
|
NsiI
|
Mph1103I
|
Zsp2I
|
BfrBI
| |

7441 TAAGCTTACCCTATCTGAAGTGACAGGACAGGGCCTATGCATGGGGGCAGTACCTAAAAC 7500
ATTGGAATGGGATAGACTTCACTGTCTCTGTCGCGGATACGTACCCCGTCATGGATTTG

NaeI
|
PdiI
|
MroNI
| |
NgoMIV
| |

7501 TCACCAAGGCCTTATGTAACACCAACCAAAGGGCGGCTCAGGATCCTACTACCTTGCAGC 7560
AGTGGTCCGGAATACATTGTGGTGGGTTTCGCGGCGAGTCTAGGATGATGGAACGTCC

AleI
|
OliI
|

7561 ACCCGCCGGAACAATGTGGGCTTGAGCACTGGATTGACTCCCTGCTTGTCCACCAAGGT 7620
TGGCGCGCCTTGTACACCCGACGTCGTGACCTAAGTACGAGGACGAACAGGTGGTGCCA

7621 GCTCAATCTAACCAAGATTATTGTGTATTAGTTGAACCTCTGSCCAGAGTAATTACCA 7680
CGAGTTAGATTGGTGTCTAATAACACATAATCAACTTGAGACCGGTCTCATTAAATGGT

FIGURA 1D-13

150

17/29

6901 CTTCGAAGGGGCTACTGAGGGGGCAGATGCAACCCCTCTAGTCTAGAAATTCACCTGATGC 6960
GAAGGTTCCCCGATGAGCTCCCCCGTCTACGTTGGGAGATCAGGATCTTAAGTGACTACG

6961 AGCAAAAAGGCTAACTGGGACGGGCCCCAAATCGTGGGACTGAGACTGTACCGGACAGG 7020
TCCTTTTTC CGATTGACCTGCCCCGGTTTAGCACCCCTGACTCTGACATGGCCTGTCC

7021 AACAGATCCTATTACCATGTTCTCCCTGACCCGGCAGGTCCTTAATGTGGGACCCCGAGT 7080
TTGCTTAGCATATGCTACAAGAGGGACTGGGCGTCCAGGAATTACACCCTGGGGCTCA

7081 CCCCATAGGGCCCCAACCCAGTATTACCCGACCAAGAGACTCCCTTCCTCACCATAGAGAT 7140
GGGGTATCCCCGGTGGGTCAATAGGGTGGTTTCTGAGGGAGGAGTGGTTATCTCTA

7141 TGTACCGGCTCCACAGCCACCTAGCCCCCTCAATACCAGTTACCCCTTCCACTACCAG 7200
ACATGGCCGAGGTGTCGGTGGATCGGGGGAGTTATGGTCAATGGGGGAAGGTGATGGTC

7201 TACACCCTCAACCTCCCTACAAGTCCAAGTGTCCACAGCCACCCCGAGGAAGTGGAGA 7260
ATGTGGGAGTTGGAGGGGATGTTCAAGTTCACAGGGTGTCTGGTGGGGTCCCTGACCTCT

7261 TAGACTACTAGCTCTAGTCAAAGGAGCCTATCAGGCGCTTAACCTCACCATCCCACAA 7320
ATCTGATGATCGAGATCAGTTTCCTCGGATAGTCCGCGAATTGGAGTGGTTAGGGCTGTT

7321 GACCCAAGAATGTTGGCTGTGCTTAGTGTGGGACCTCCTTATTAAGAGGAGTAGCGGT 7380
CTGGGTCTTACAACCGACACGAATCACAGCCCTGGAGGAATATGCTTCCTCATCGCA

7381 CGTGGGCACTTATAACATCAATCCACCGCTCCGGCCAACTGTAGGGGCACTTCCCAACA 7440
GCACCCGTGAATATGGTTAGTAAGGTGGCGAGGCCGGTTGACATCCCGGTGAAGGGTTGT

EcoT22I

|

NsiI

|

MphI103I

|

Zsp2I

|

BfrBI

| |

7441 TAAGCTTACCTATCTGAGTGACAGGACAGGGCCTATGCAATGGGGCAGTACCTAABAC 7500
ATTGGAATGGGATAGACTTCACTGTCTGTCCCGGATACGTACCCCGTCATGGAATTTG

NaeI

|

PstI

|

MroNI

| |

NgoMIV

| |

7501 TCACCAGGCTTATGTRACACCAACCAAGCGCCGGCTCAGGATCCTACTACCTTGACAGC 7560
AGTGGTCCGGATACATTGTGGTGGTTTCGGGCCGAGTCTAGGATGATGGAACTCG

FIGURA 1D-14

HOJA SUSTITUTA (REGLA 26)

151

18/29

A1eI
 |
 011I
 |

7561 ACCCGCCGGAACAAATGTGGGCTTGCAGCACTGGATTGACTCCCTGCTTGTCCACCACGGT 7620
 TGGGCGGCTTGTACACCCGAACGTCGTGACCTAACTGAGGGACGAACAGGTGGTGCCA
 7621 GCTCAATCTAACCACAGATTATTGTGTATTAGTTGAACTCTGGCCCAGAGTAATTTACCA 7680
 CGAGTTAGATTGGTGTCTAATAACACATAATCAACTTGAGACCGGGTCTCATTAAATGGT
 7681 CTCCCCCGATTATATGTATGGTCAGCTTGAACAGCGTACCAAATATAAAAGAGAGCCAGT 7740
 GAGGGGGCTAATATACATACCACTCCAACTTGTGCGATGGTTTATATTTCTCTCGGTCA
 7741 ATCATTGACCCTGGCCCTTCTACTAGGAGGATTAAACATGGGAGGGATTCCAGCTGGAAT 7800
 TAGTAACCTGGGACCGGGAAGATGATCCTCCTAATGCTACCTCCCTAAGCTCGACCTTA
 7801 AGGGACGGGGACCACTGCCTTAATTAAAACCCAGCAGTTTGAGCAGCTTCATGCCGCTAT 7860
 TCCCTGCCCCCTGGTGACGGGAATTAAATTTGGGTGCTCAAACTCGTCAAGTACGGCGATA
 7861 CCAGACAGACCTCAACGAAGTCGAAAAGTCAATTACCAACCTAGAAAAGTCACTGACCTC 7920
 GGTCTGTCTGGAGTTGCTTCAGCTTTTCAGTTAATGGTTGGATCTTTTCAGTGACTGGAG
 7921 GTTCTCTGAAGTAGTCTACAGAACCGCAGAGGCTTAGATTTECTATTCCCTAAAGGAGGG 7980
 CAACAGACTTCATCAGGATGTCTTCCCGTCTCCGATCTAAACGATAAGGATTTCTCCC
 7981 AGGTCTCTGCGCAGCCCTAAAAGAAGAAATGTTGTTTTATCCAGACCACACGGGGCTAGT 8040
 TCCAGAGACGGCGTGGGATTTTCTTCTTACAACAAAATACGCTGTGCTGCCCGGATCA
 8041 GAGAGACAGCATGGCCAAATTAAGAGAAAGGCTTAATCAGAGACAAAACTATTTGAGAC 8100
 CTCTCTGTGCTACCGTTTAATTCTCTTTCCGAATTAGTCTCTGTTTTTGATAAACTCTG

NspV

|

BstBI

|

Esp119I

|

AseII

|

Csp45I

|

SfuI

|

Bpu14I

|

BspT104I

|

8101 AGGCCAAGGATGGTTTGAAGGGCTGTTTAAATGATCCCCCTGGTTTACCACCTTAATCTC 8160
 TCCGGTTCTTACCAAGCTTCCCGACAAATTATCTAGGGGGACCAATGGTGAATTAGAG

FIGURA 1D-15

HOJA SUSTITUTA (REGLA 26)

152

19/29

```

8161 CACCATCATGGGACCTCTAATAGTACTCTTACTGATCTTACTCTTTGGACCTTGCATTCT 8220
      GTGCTAGTACCTGGAGATTATCATGAGAACTGACTAGAAATGAGAAACCTGGAACTGTAAGA

8221 CAATCGATTGGTCCAAATTTGTTAAAGACAGGATCTCAGTGGTCCAGGCTCTGCTTTTGAC 8280
      GTTAGCTAACAGGTTAAACAATTTCTGTCTAGAGTCACCAGSTCCGAGACCAAACTG
                                     IRES reg(8327,8876)>>>
                                     |
                                     M1uI (8325)
                                     | |
8281 TCAGCAATATCACCAGCTAAAACCCATAGAGTACGAGCCATGAACGGCTTACTGGCCGAA 8340
      AGTCGTTATAGTGGTCCGATTTTGGGTATCTCATGCTCGGTACTTGGCAATGACCGGCTT

                                     ires_emov reg(8378,8876)>>>
                                     |
8341 GCGGCTTGGAAATAGGCGCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCGGT 8400
      GCGCGAACCTTATTCGGGCCACACGCAACAGATATACAAATAAAGGTGGTATACGGCA

8401 CTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGACATTCCTAGGG 8460
      GAAACCGTTACACTCCCGGTCCTTTGGACCGGGACAGAGACTTCTCGTAAGGATCCC

8461 GTCTTTCCCTCTCCGCAAAAGGAATGCAAGTCTGTGAATGTCGTGAAGGAGGCACTTC 8520
      CAGAAAGGGGAGAGCGGTTTCTTACGTTCCAGACAACTTACAGCACTTCTTTCGTCAAG

8521 CTCTGGAGCTTCTTGAAGACAAACAAGTCTGTAGCGACCCCTTTCAGGCGACCGGAACC 8580
      GAGACCTTCGAGAACTTCTGTTTGTTCAGACATCGCTGGGAAACGTCCTGCGCTTGG

8581 CDDACCTGGGACAGGTGCTCTGCGGCCAAAAGCCAGCTGTATTAAGATACACCTGCAA 8640
      GGGGTGGACCGCTGTCCAGGGAGACGCGGTTTTTGGTGCACATATCTATGTGGACGTT

8641 AGCGGGCACAACCCCACTGCCACGTTGTGAGTTGGATAGTTCTCGAAAGAGTCAAATGGC 8700
      TCCGCCGTGTGGGGTCAGGCTGCAACACTCAACCTATCAACACCTTTCTCAGTTTACCG

8701 TCTCCTCAAGCGTATTCACAAGGGGCTGAAGGATGCCAGAGGTACCCCATTTGTATGG 8760
      AGAGGAGTTCCATAAGTGTGTTCCCGACTTCCTACGGGTCTTCCATGGGGTAACATACC

8761 GATCTGATCTGGGGCTCGGTCCACATGCTTTACATGTGTTTGTGAGGTTAAAAAAC 8820
      CTAGACTAGACCCCGGAGCCACGTGTACGAATGTACACAAATCAGCTCCAATTTTTTG

```

FIGURA 1D-16

HOJA SUSTITUTA (REGLA 26)

153

20/29

yCD2 (8877, 9353)>>>

PstI (8874)

8821 GTCTAGGCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAACACGATATAAATGG 8880
 CAGATCOGGGGGCTTGGTSCCCCTGCACCAAAAGGAACTTTTGTGCTAATATTTACC

8881 TGACCGGGGGCATGGCCTCCAAGTGGGATCAAAGGGCATGGATATCGCTTACGAGGAGG 8940
 ACTGGCCGGCGTACCGGAGGTTCACCCTAGTTTTCCTTACCTATAGCGAATGCTCCTCC

8941 CCCTGCTCGCTACAAAGAGGGGGGGCTCCCTATCGCGCGCTGCTCTGATCAACAACAAGG 9000
 GCGACGACCGGATCTTCTCCCGCGGACGGATAGCGCGGACAGACTAGTTGTTGTTCC

9001 ACGGCAGTGTGCTGGGACGGGGCCACACATGAGGTTCCAGAGGGCTCCGCCACCCCTGG 9060
 TGCCGTCACACGACCCGTCGCCGCTGTGTACTCCAGGCTCTTCGAGGGCGGTGGGAOG

9061 ACGGCCAGATCTCCACCTCGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACN 9120
 TCGCGCTCTAGAGCTGGGAGCTCTTGACACCGTCCGACCTCCCGTTCCACATGTTCTGT

9121 CCACCTGTACACCACTGTCCCTTGTGACATGTGTACCGCGCTATCATCATGTACG 9180
 GGTGGGACATGTGGTGGGACAGGGGAACACTGTACACATGGCGCGATAGTAGTACATGC

9181 ECATCCCTAGGTGTGTGATCGGCGAAGCTGAACCTCAAGTCCAGGGCGAAGTACG 9240
 CGTAGGGATCCACACACTAGCGGCTCTTCACTTGAAGTTCAAGTTCCCGCTTTTCATGG

9241 TGCATAACAGGGGGCCACGAGCTGGTSCCTTGTGACGATGACAGGTGTAAAGAACTGATG 9300
 ACGTTTGGTCCCGGCTGCTCCACCAACAACACTGCTACTCTCCACATTCTTCCACTACT

NotI (9356)

CciNI

9301 AGCAGTTTCATCGACGAGAGCCCTCAGCACTCGTTCCAGGATATCGCCGAGTACCGGCGG 9360
 TCGTCAAGTAGCTGCTCTCGGAGTCTCAACCAAGCTCCTATAGCCGCTCATTGCGCGG

U3 Region(9405, 9854)>>>

9361 CAGATAAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGAATGAAAGACCCACC 9420
 TCTATTTTATTTTCTAAATAAATCAGAGGTCTTTTCCCCCTTACTTTCTGGGGTGG

FIGURA 1D-17

HOJA SUSTITUTA (REGLA 26)

154

21/29

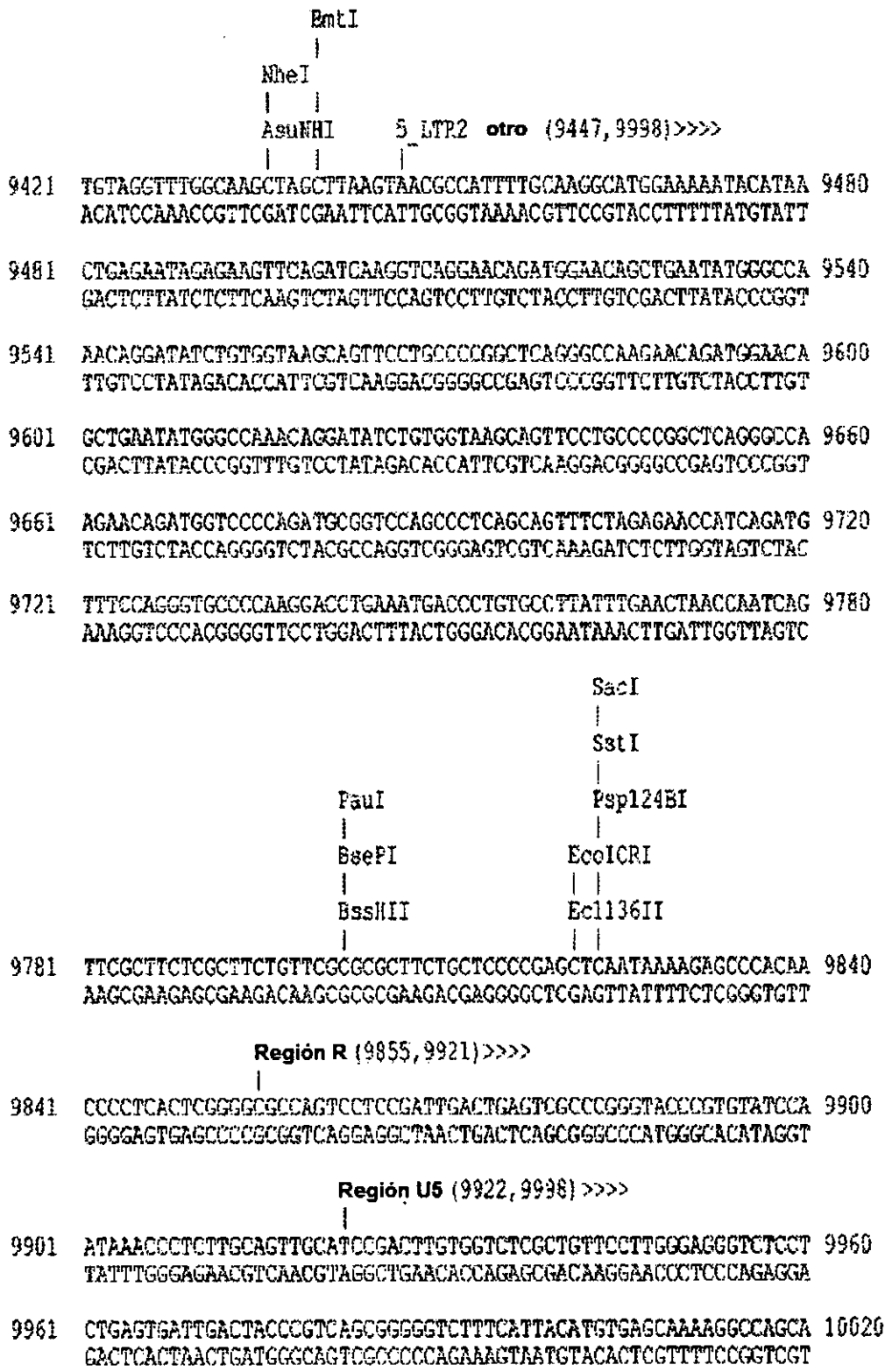


FIGURA 1D-18

HOJA SUSTITUTA (REGLA 26)

155

22/29

origen pBR322 (10045, 10654) <<<

1

10021 AAAGGCCAGGAACCGTAAAAAGGCCGCTTGGCTGGCGTTTTTCCATAGGCTCCGCCCCCC 10080
TTTCCGGCTCCTTGGCATTTTTCCGGCGCAACGACCGCAAAAGGTATCCGAGGCGGGGG

10081 TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCGACAGGACTATA 10140
ACTGCTCGTAGTGTTTTGTAGCTGGGAGTTTCTCTCCACCGCTTTGGGCTGTCTGATAT

10141 AAGATACCAGGCGTTTCCCCCTGGAAGCTCCTCGTGCGCTCTCCTGTTCCGACCCCTGCC 10200
TTCTATGGTCCGCAAGGGGGACCTTCGAGGGAGCAGCGGAGAGGACAAAGGCTGGGACGG

10201 GCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 10250
CGAATGGCCTATGGACAGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTTACGAG

10261 ACGCTGTAGGTATCTCAGTTCCGCTGTAGGTCTGCTCCAGCTGGGCTGTGTGCACGA 10320
TGGGACATCCATAGAGTCAAGCCACATCCAGCAGCGAGGTTCCGACCCGACACAGTCT

10321 ACCCCCCGTTCCAGCCGACCGCTGGCGCTTATCCGCTAACTATCGTCTTGAGTCCAAACC 10380
TGGGGGCAAGTCCGGCTGGCGACCGGAATAGGCCATTGATAGCAGAACTCAGGTTGGG

10381 GGTAAAGACAGGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 10440
CCATTCTGTGCTGAATAGCGGTGACCGTCTGTGGTGAACATTGTCTTAATCGTCTCGCTC

10441 GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAG 10500
CATACATCCGCCACGATGTCTCAGAACTTCACCACCGGATTGATGCCGATGTGATCTC

10501 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 10560
CTGTCAAAACCATAGACGCGAGAGGACTTCGGTCAATGGAGCCTTTTTCTCAACCATC

10561 CTCTTGATCCGGCAAAACCAACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCA 10620
GAGAACTAGGCCGTTTGTGTTGGTGGCGACCATCGCCACCAAAAAACAAACGTTCTGCTGT

10621 GATTACGGCGAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGA 10680
CTAATGCGCGTCTTTTTTCTTAGAGTTCTTCTAGGAACTAGAAAAGATGCCCGAGACT

10681 CGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGAT 10740
GCGAGTCACCTTGCTTTTGGAGTCAATTCCCTAAAAACAGTACTCTAATAGTTTTTCTTA

10741 CTTCACTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGA 10800
GAAGTGGATCTAGGAAAATTTAATTTTTTACTTCAAAATTTAGTTAGATTTCAATATATCT

FIGURA 1D-19

HOJA SUSTITUTA (REGLA 26)

156

23/29

marcador amp (10819, 11679) <<<

10801 GTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTATCTCAGCGATCTG 10860
CATTGGAACCAGACTGTCAATGGTTACGAATTAGTCACTCCGTGATAGAGTGGCTAGAC

10861 TCTATTTCGTTTCATCCATAGTTGCCGTGACTCCCCGTCGTGTAGATAACTAOCATAGGGA 10920
AGATAAAGCAAGTAGGTATCAACGGACTGAGGGGCAGCACA TCTATTGATGCTATGGCT

10921 GGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCCTCTCC 10980
CCCGAATGCTAGACCGGGGTCCAGACGTTACTATGGCGCTCTGGGTGCGAGTGGCCGAGG

10981 AGATTTATCAGCAATAAACCAGCCAGCCGGAGGGCCGAGCCAGAAAGTGGTCTGCAAC 11040
TCTAAATAGTCGTTAFTTGGTCCGTGGCCCTTCCGGGCTCGGTCTTCAACAGGACGTTG

11041 TTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGC 11100
AAATAGGGGGAGGTAGGTGAGATAATTAAACAACGGCCCTTCGATCTCATTCATCAAGCGG

11101 AGTTAATAGTTTGGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTACGGCTGTC 11160
TCAATTATCAACCGCTTGCAACAACGTTACGACGTCGGTAGCACCAGAGTGGGAGCAG

11161 GTTGGTATCGCTTCATTTCAGCTCCGCTTCCCAACGATCAAGCCGAGTTACATGATCCCC 11220
CAAACCATACCGAAGTAAGTGGAGGCCAAGGCTTGCTAGTTCGGCTCAATGTACTAGGGG

11221 CATGTTGTGCAAAAAGCGGTTAGCTCCTTCCGCTCCGATCGTTGTGAGAAGTAAGTT 11280
GTACAACACGTTTTTTCGCCAATCGAGGAGGCCAGGAGGCTAGCAACAGTCTTCATTCAA

11281 GGCCGCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCTCTTACTGTATGCC 11340
CCGGCGTCACAATAGTGAGTACCAATACCGTCTGACGTATTAAAGAGATGACAGTACGG

11341 ATCCGTAAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 11400
TAGGCATTCTACGAAAAGACACTGACCACCTCATGAGTTGGTTCAAGTAACTCTTATCAC

11401 TATGCGGGGACCGASTTGTCTTGGCCGGGCTCAACACGGGATAATACCGCGCCACATAG 11460
ATACGCGGCTGGCTCAACGAGAACGGGCGCAGTTGTGCCCTATTATGGCGCGGTGTATC

11461 CAGAACTTTAAAAGTGCTCATCATTTGAAAAAGTTCTTCGGGGCGAAAACCTCTCAAGGAT 11520
GTCTTGAAATTTTACGAGTAGTAACCTTTTGCAAGAGCCCCGCTTTTGAGAGTTCCTA

11521 CTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTC GTGCACCCACTGATCTTCAGC 11580
GAATGGCGACAACTCTAGGTCAAGCTACATTGGGTGAGCAOGTGGGTTGACTAGAGTUG

11581 ATCTTTTACTTTACACAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGAAA 11640
TAGAAAATGAAAGTGGTCCGAAAGACCCACTCGTTTTTGTCTTCCGTTTTACGGCGTTT

11641 AAAGGGCAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTCAATATTA 11700
TTTCCCTTATTCGGCTGTGCTTTACAACTTATGAGTATGAGAAGGAAAAGTTATAAT

FIGURA 1D-20

HOJA SUSTITUTA (REGLA 26)

157

24/29

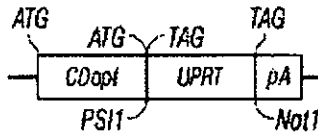
amp prom(11721,11749)<<<
|
11701 TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTAGAA 11760
AACTTCGTAAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAACTTACATAAATCTT
11761 AAATAACAAATAGGGGTTCCGGGACATTTCCCGAAAAGTCCACCTGACGTCTAAGA 11820
TTTATTTGTTTATCCCAAGGCGGTGTAAAGGGGCTTTTCACGGTGGACTGCAGATTCT
11821 AACCATTTATTATCATGACATTACCTATAPPAATAGGGGTATCAGAGGCCCTTTCTCT 11880
TTGGTAATAATAGTACTGTAAATTCATATTTTATCCGCATAGTCTCCGGAAAGCAGA
11881 TCAAGAATTCAT 11892
AGTTCTTAAGTA

FIGURA 1D-21

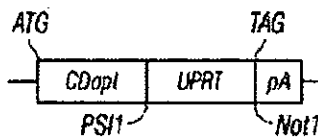
1. Síntesis separada de genes CDopt y UPRT con sitios PSII/NotI



2. Digerido con PSII y desarrollo de ligación de extremo romo



3. Eliminación del codón de INICIO y DETENCIÓN UPRT del gen CDopt mediante mutagenia dirigida a sitio



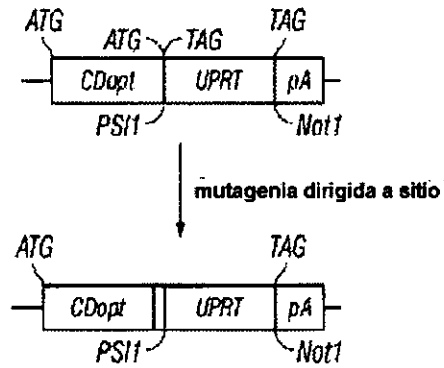
4. Tamaño del gen CDopt-UPRT: 1.3 kb

FIGURA 2A

158

25/29

1. Síntesis de CDopt-LIGADOR-UPRT: Inserción de mutagenia dirigida a sitio del ligador



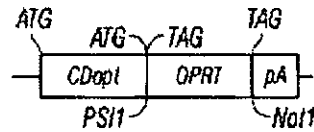
2. Tamaño del gen CDpot-Ligador-UPRT: 1.36 kb

FIGURA 2B

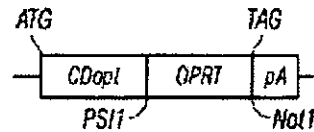
1. Síntesis separada del gen CDopt y el dominio OPRT con sitios PSI1/NotI



2. Digerido con PSI1 y desarrolla ligación de extremo romo



3. Eliminación del codón de INICIO y DETENCIÓN OPRT del gen CDopt mediante mutagenia dirigida a sitio



4. Tamaño del gen CDopt-OPRT: 1.3 kb

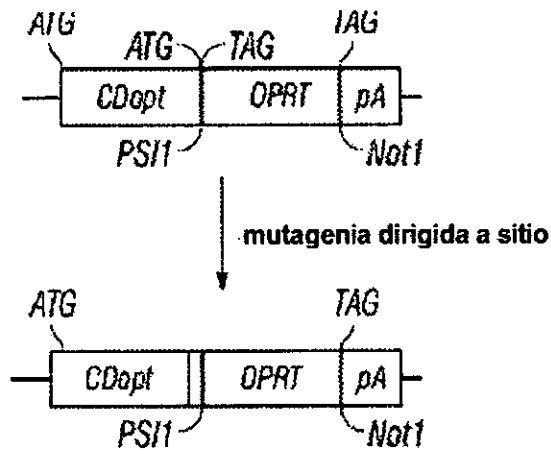
FIGURA 2C

HOJA SUSTITUTA (REGLA 26)

159

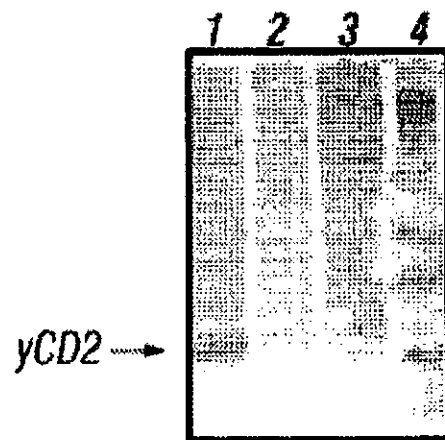
26/29

1. Síntesis de CDopt-LIGADOR-OPRT: Inserción de mutagenia dirigida a sitio del ligador



2. Tamaño del gen CD-OPRT: 1.33 kb

FIGURA 2D



- 1 - U87 + AC3-yCD2 (V)
- 2 - U87 + ACE-yCD (V)
- 3 - U87 (no infectado)
- 4 - Estándares de Peso Molecular

FIGURA 3

27/29

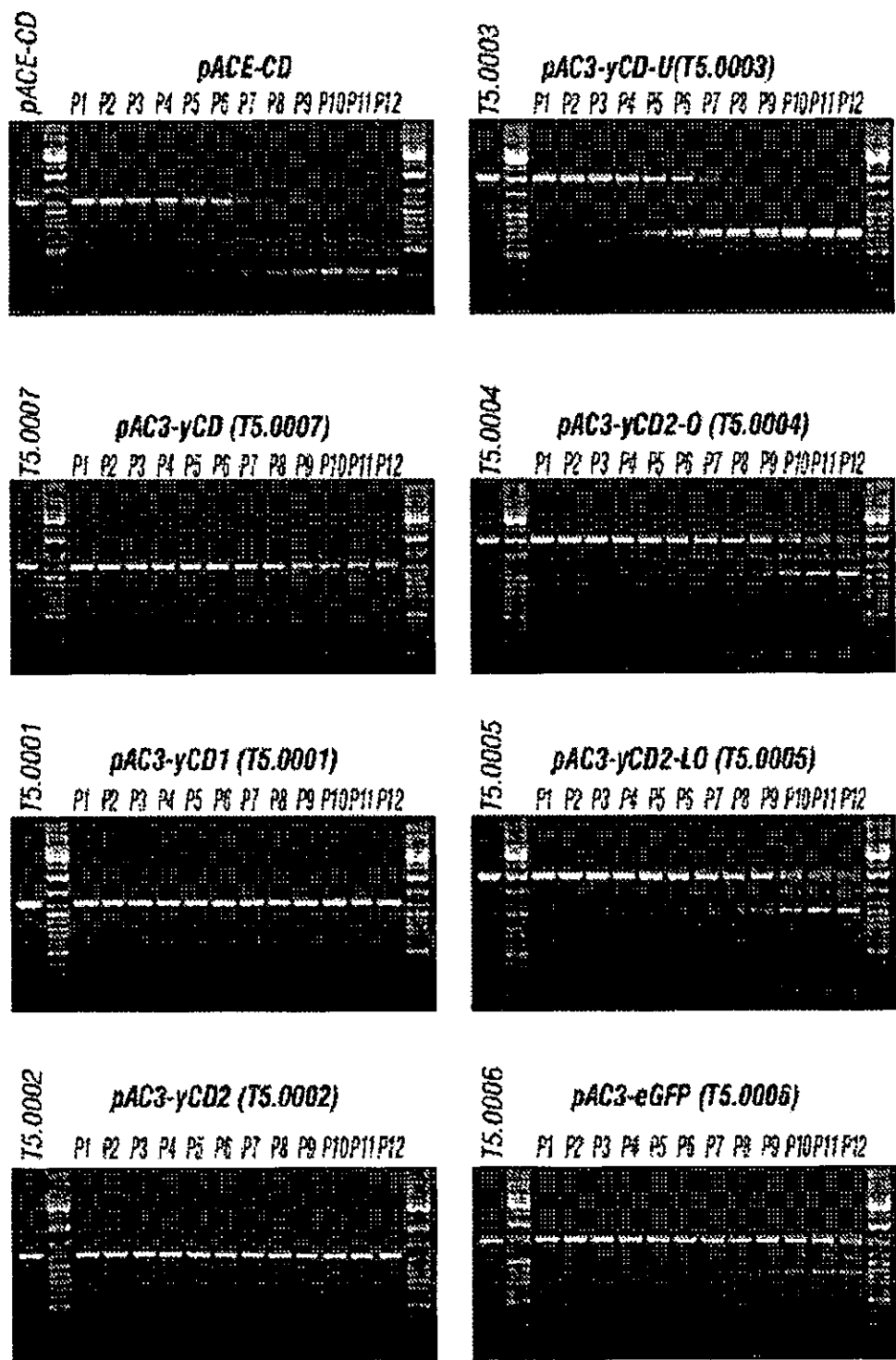


FIGURA 4

161

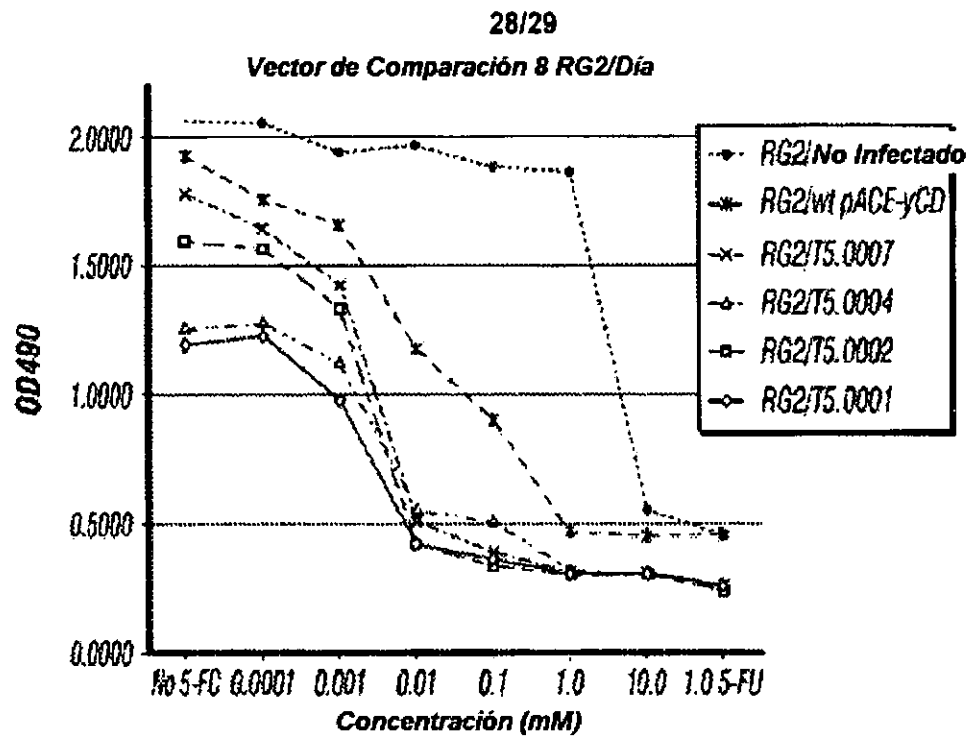


FIGURA 5A

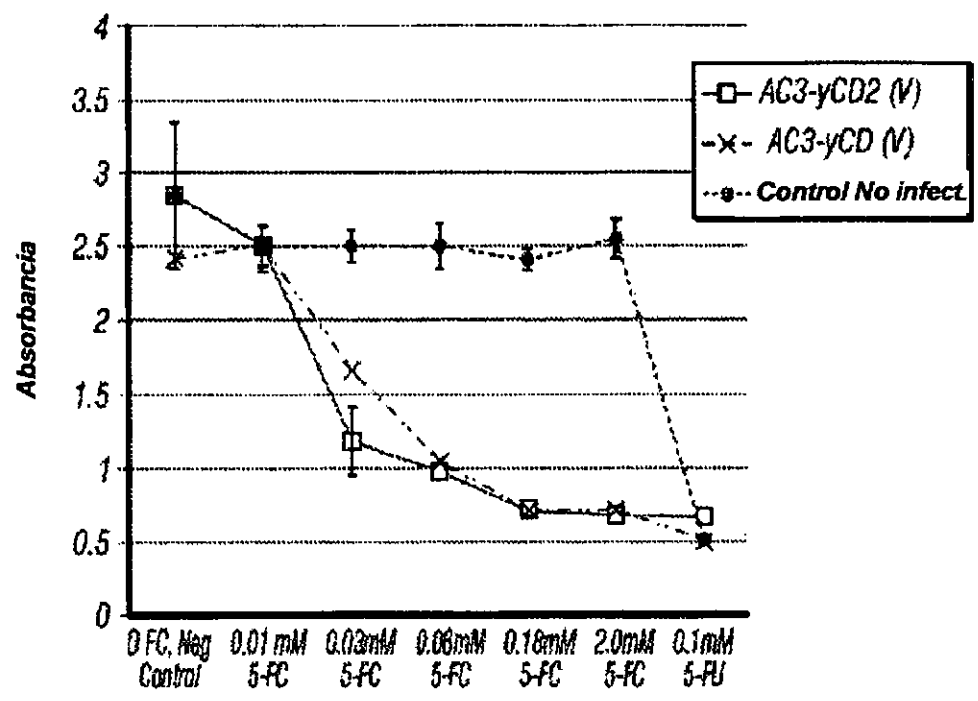


FIGURA 5B

162

LISTADO DE SECUENCIAS

<110> Tocagen Inc.
 Gruber, Harry E.
 Jolly, Douglas
 Perez, Omar

<120> VECTORES DE TERAPIA DE GEN Y DESAMINASAS CITOSINA

<130> 00014-007WO1

<140> No se conoce aún
 <141> 2009-09-26

<150> US 61/100,666
 <151> 2008-09-26

<150> US 61/120,618
 <151> 2008-12-08

<150> US 61/186,823
 <151> 2009-06-13

<160> 22

<170> PatentIn version 3.5

<210> 1
 <211> 477
 <212> ADN
 <213> Saccharomyces cerevisiae

<220>
 <221> CDS
 <222> (1)..(477)

<400> 1
 atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac 48
 Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15
 att gcc tat gag gag gcg gcc tta ggt tac aaa gag ggt ggt gtt cct 96
 Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30
 att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt 144
 Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45
 ggt cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag 192
 Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60
 atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa 240
 Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80
 gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt 288
 Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

gcc atc atc atg tat ggt att cca cgc tgt gtt gtc ggt gag aac gtt 336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
100 105 110

aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag 384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

gtt gtt gtt gtt gac gat gag agg tgt aaa aag atc atg aaa caa ttt 432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
130 135 140

atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag 477
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 2
<211> 158
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 2

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
 145 150 155

<210> 3
 <211> 477
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Desaminasa citosina construida por ingenieria

<220>
 <221> CDS
 <222> (1)..(477)

<400> 3
 atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac 48
 Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15
 att gcc tat gag gag gcg tta tta ggt tac aaa gag ggt ggt gtt cct 96
 Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30
 att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt 144
 Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45
 ggt cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag 192
 Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60
 atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa 240
 Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80
 gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt 288
 Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95
 gcc atc atc atg tat ggt att cca cgc tgt gtc atc ggt gag aac gtt 336
 Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110
 aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag 384
 Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125
 gtt gtt gtt gtt gac gat gag agg tgt aaa aag tta atg aaa caa ttt 432
 Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
 130 135 140
 atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag 477
 Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
 145 150 155

<210> 4
 <211> 158

<212> PRT

<213> **Secuencia Artificial**

<220>

<223> **Construcción Sintética**

<400> 4

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 5

<211> 480

<212> **ADN**

<213> **Secuencia Artificial**

<220>

<223> **Desaminasa citosina optimizada de codón humano**

<220>

<221> CDS

<222> (1) .. (480)

167

<400> 5
 atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat 48
 Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

atc gct tac gag gag gcc gca ctg ggc tac aag gag ggc ggc gtg cct 96
 Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg 144
 Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag 192
 Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag 240
 Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc 288
 Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

gct atc atc atg tac ggc atc cct agg tgt gtg gtc ggc gag aac gtg 336
 Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
 100 105 110

aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag 384
 Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

gtg gtg gtt gtt gac gat gag agg tgt aag aag atc atg aag cag ttc 432
 Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
 130 135 140

atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tga taa 480
 Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
 145 150 155

<210> 6
 <211> 158
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Construcción Sintética

<400> 6
 Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg

35 40 45
 Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60
 Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80
 Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95
 Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
 100 105 110
 Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125
 Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
 130 135 140
 Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
 145 150 155

<210> 7
 <211> 756
 <212> ADN
 <213> *Saccharomyces cerevisiae*

<220>
 <221> CDS
 <222> (1)..(756)

<400> 7
 atg aac ccg tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca tat 48
 Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr
 1 5 10 15
 ctt ata tat tat cca aac aaa ggg tct ttc gtt agc aaa cct aga aat 96
 Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn
 20 25 30
 ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct 144
 Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro
 35 40 45
 caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat 192
 Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn
 50 55 60
 aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga ttg 240
 Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu
 65 70 75 80

169

ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val 85 90 95	288
gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys 100 105 110	336
atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly 115 120 125	384
tta aga gac tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln 130 135 140	432
agg gac gag gag act gct tta cca aag tta ttc tac gaa aaa tta cca Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro 145 150 155 160	480
gag gat ata tct gaa agg tat gtc ttc cta tta gac cca atg ctg gcc Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala 165 170 175	528
acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga ggt Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly 180 185 190	576
gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag gaa Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu 195 200 205	624
ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr 210 215 220	672
ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly 225 230 235 240	720
ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val 245 250	756

<210> 8
 <211> 251
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 8

Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr 1 5 10 15
--

Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn 20 25 30

170

Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro
 35 40 45

Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn
 50 55 60

Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu
 65 70 75 80

Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val
 85 90 95

Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys
 100 105 110

Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly
 115 120 125

Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln
 130 135 140

Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro
 145 150 155 160

Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala
 165 170 175

Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly
 180 185 190

Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu
 195 200 205

Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr
 210 215 220

Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly
 225 230 235 240

Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
 245 250

<210> 9
 <211> 1443
 <212> ADN
 <213> HOMO sapiens

171

<220>

<221> CDS

<222> (1)..(1443)

<400> 9

atg gct gtt gct cgt gct gct ctt ggt cct ctt gtt act ggt ctt tat	48
Met Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr	
1 5 10 15	
gat gtt caa gct ttt aaa ttt ggt gat ttt gtt ctt aaa tct ggt ctt	96
Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu	
20 25 30	
tct tct cct att tat att gat ctt cgt ggt att gtt tct cgt cct cgt	144
Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg	
35 40 45	
ctt ctt tct caa gtt gct gat att ctt ttt caa act gct caa aat gct	192
Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala	
50 55 60	
ggt att tct ttt gat act gtt tgt ggt gtt cct tat act gct ctt cct	240
Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro	
65 70 75 80	
ctt gct act gtt att tgt tct act aat caa att cct atg ctt att cgt	288
Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg	
85 90 95	
cgt aaa gaa act aaa gat tat ggt act aaa cgt ctt gtt gaa ggt act	336
Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr	
100 105 110	
att aat cct ggt gaa act tgt ctt att att gaa gat gtt gtt act tct	384
Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser	
115 120 125	
ggt tct tct gtt ctt gaa act gtt gaa gtt ctt caa aaa gaa ggt ctt	432
Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu	
130 135 140	
aaa gtt act gat gct att gtt ctt ctt gat cgt gaa caa ggt ggt aaa	480
Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys	
145 150 155 160	
gat aaa ctt caa gct cat ggt att cgt ctt cat tct gtt tgt act ctt	528
Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu	
165 170 175	
tct aaa atg ctt gaa att ctt gaa caa caa aaa aaa gtt gat gct gaa	576
Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu	
180 185 190	
act gtt ggt cgt gtt aaa cgt ttt att caa gaa aat gtt ttt gtt gct	624
Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala	
195 200 205	
gct aat cat aat ggt tct cct ctt tct att aaa gaa gct cct aaa gaa	672
Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu	
210 215 220	

172

ctt tct ttt ggt gct cgt gct gaa ctt cct cgt att cat cct gtt gct Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala 225 230 235 240	720
tct aaa ctt ctt cgt ctt atg caa aaa aaa gaa act aat ctt tgt ctt Ser Lys Leu Leu Arg Leu Met Gln Lys Lys Glu Thr Asn Leu Cys Leu 245 250 255	768
tct gct gat gtt tct ctt gct cgt gaa ctt ctt caa ctt gct gat gct Ser Ala Asp Val Ser Leu Ala Arg Glu Leu Leu Gln Leu Ala Asp Ala 260 265 270	816
ctt ggt cct tct att tgt atg ctt aaa act cat gtt gat att ctt aat Leu Gly Pro Ser Ile Cys Met Leu Lys Thr His Val Asp Ile Leu Asn 275 280 285	864
gat ttt act ctt gat gtt atg aaa gaa ctt att act ctt gct aaa tgt Asp Phe Thr Leu Asp Val Met Lys Glu Leu Ile Thr Leu Ala Lys Cys 290 295 300	912
cat gaa ttt ctt att ttt gaa gat cgt aaa ttt gct gat att ggt aat His Glu Phe Leu Ile Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly Asn 305 310 315 320	960
act gtt aaa aaa caa tat gaa ggt ggt att ttt aaa att gct tct tgg Thr Val Lys Lys Gln Tyr Glu Gly Gly Ile Phe Lys Ile Ala Ser Trp 325 330 335	1008
gct gat ctt gtt aat gct cat gtt gtt cct ggt tct ggt gtt gtt aaa Ala Asp Leu Val Asn Ala His Val Val Pro Gly Ser Gly Val Val Lys 340 345 350	1056
ggt ctt caa gaa gtt ggt ctt cct ctt cat cgt ggt tgt ctt ctt att Gly Leu Gln Glu Val Gly Leu Pro Leu His Arg Gly Cys Leu Leu Ile 355 360 365	1104
gct gaa atg tct tct act ggt tct ctt gct act ggt gat tat act cgt Ala Glu Met Ser Ser Thr Gly Ser Leu Ala Thr Gly Asp Tyr Thr Arg 370 375 380	1152
gct gct gtt cgt atg gct gaa gaa cat tct gaa ttt gtt gtt ggt ttt Ala Ala Val Arg Met Ala Glu Glu His Ser Glu Phe Val Val Gly Phe 385 390 395 400	1200
att tct ggt tct cgt gtt tct atg aaa cct gaa ttt ctt cat ctt act Ile Ser Gly Ser Arg Val Ser Met Lys Pro Glu Phe Leu His Leu Thr 405 410 415	1248
cct ggt gtt caa ctt gaa gct ggt ggt gat aat ctt ggt caa caa tat Pro Gly Val Gln Leu Glu Ala Gly Gly Asp Asn Leu Gly Gln Gln Tyr 420 425 430	1296
aat tct cct caa gaa gtt att ggt aaa cgt ggt tct gat att att att Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile 435 440 445	1344
gtt ggt cgt ggt att att tct gct gct gat cgt ctt gaa gct gct gaa Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu 450 455 460	1392

atg tat cgt aaa gct gct tgg gaa gct tat ctt tct cgt ctt ggt gtt 1440
 Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val
 465 470 475 480

taa 1443

<210> 10
 <211> 480
 <212> PRT
 <213> homo sapiens

<400> 10

Met Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr
 1 5 10 15

Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu
 20 25 30

Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg
 35 40 45

Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala
 50 55 60

Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro
 65 70 75 80

Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg
 85 90 95

Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr
 100 105 110

Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser
 115 120 125

Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu
 130 135 140

Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys
 145 150 155 160

Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu
 165 170 175

Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu
 180 185 190

Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala
 195 200 205
 Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu
 210 215 220
 Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala
 225 230 235 240
 Ser Lys Leu Leu Arg Leu Met Gln Lys Lys Glu Thr Asn Leu Cys Leu
 245 250 255
 Ser Ala Asp Val Ser Leu Ala Arg Glu Leu Leu Gln Leu Ala Asp Ala
 260 265 270
 Leu Gly Pro Ser Ile Cys Met Leu Lys Thr His Val Asp Ile Leu Asn
 275 280 285
 Asp Phe Thr Leu Asp Val Met Lys Glu Leu Ile Thr Leu Ala Lys Cys
 290 295 300
 His Glu Phe Leu Ile Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly Asn
 305 310 315 320
 Thr Val Lys Lys Gln Tyr Glu Gly Gly Ile Phe Lys Ile Ala Ser Trp
 325 330 335
 Ala Asp Leu Val Asn Ala His Val Val Pro Gly Ser Gly Val Val Lys
 340 345 350
 Gly Leu Gln Glu Val Gly Leu Pro Leu His Arg Gly Cys Leu Leu Ile
 355 360 365
 Ala Glu Met Ser Ser Thr Gly Ser Leu Ala Thr Gly Asp Tyr Thr Arg
 370 375 380
 Ala Ala Val Arg Met Ala Glu Glu His Ser Glu Phe Val Val Gly Phe
 385 390 395 400
 Ile Ser Gly Ser Arg Val Ser Met Lys Pro Glu Phe Leu His Leu Thr
 405 410 415
 Pro Gly Val Gln Leu Glu Ala Gly Gly Asp Asn Leu Gly Gln Gln Tyr
 420 425 430

175

Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile
435 440 445

Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu
450 455 460

Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val
465 470 475 480

<210> 11

<211> 1227

<212> ADN

<213> Secuencia Artificial

<220>

<223> Construcción de fusión CDopt-UPRT

<220>

<221> CDS

<222> (1)..(1227)

<400> 11

atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat 48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct 96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg 144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag 192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag 240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc 288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg 336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag 384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc 432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

1726

atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag aac ccg Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro 145 150 155 160	480
tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca tat ctt ata tat Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr 165 170 175	528
tat cca aac aaa ggg tct ttc gtt agc aaa cct aga aat ctg caa aaa Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn Leu Gln Lys 180 185 190	576
atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct caa aca aac Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn 195 200 205	624
caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat aca act aga Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg 210 215 220	672
cct gat ttc att ttc tac tcc gat aga atc atc aga ttg ttg gtt gaa Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu 225 230 235 240	720
gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg gaa act gac Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp 245 250 255	768
acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa atc tgt ggt Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly 260 265 270	816
gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga tta aga gac Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp 275 280 285	864
tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa agg gac gag Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu 290 295 300	912
gag act gct tta cca aag tta ttc tac gaa aaa tta cca gag gat ata Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile 305 310 315 320	960
tct gaa agg tat gtc ttc cta tta gac cca atg ctg gcc acc ggt ggt Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly 325 330 335	1008
agt gct atc atg gct aca gaa gtc ttg att aag aga ggt gtt aag cca Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro 340 345 350	1056
gag aga att tac ttc tta aac cta atc tgt agt aag gaa ggg att gaa Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu 355 360 365	1104
aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act ggt gcc ctc Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu 370 375 380	1152

127

gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg ttg ggt gac 1200
 Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
 385 390 395 400

ttt ggt gac aga tac tac tgt gtt taa 1227
 Phe Gly Asp Arg Tyr Tyr Cys Val
 405

<210> 12
 <211> 408
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Construcción Sintética

<400> 12

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
 130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro
 145 150 155 160

Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr
 165 170 175

178

Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn Leu Gln Lys
180 185 190

Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn
195 200 205

Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg
210 215 220

Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu
225 230 235 240

Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp
245 250 255

Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly
260 265 270

Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp
275 280 285

Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu
290 295 300

Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile
305 310 315 320

Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly
325 330 335

Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro
340 345 350

Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu
355 360 365

Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu
370 375 380

Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
385 390 395 400

Phe Gly Asp Arg Tyr Tyr Cys Val
405

179

<210> 13
 <211> 1287
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Construcción de fusión-CDopt-ligador-UPRT

<220>
 <221> CDS
 <222> (1)..(1287)

<400> 13
 atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat 48
 Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15
 atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct 96
 Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30
 atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg 144
 Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45
 ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag 192
 Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60
 atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag 240
 Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80
 gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc 288
 Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95
 gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg 336
 Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110
 aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag 384
 Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125
 gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc 432
 Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
 130 135 140
 atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tcc ggc 480
 Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
 145 150 155 160
 ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc 528
 Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
 165 170 175
 ggc gcc aac ccg tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca 576
 Gly Ala Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr

180	185	190	
tat ctt ata tat tat cca aac aaa ggg tct ttc gtt agc aaa cct aga Tyr Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg 195 200 205			624
aat ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta Asn Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu 210 215 220			672
cct caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag Pro Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys 225 230 235 240			720
aat aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga Asn Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg 245 250 255			768
ttg ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att Leu Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile 260 265 270			816
gtg gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt Val Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly 275 280 285			864
aaa atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa Lys Ile Cys Gly Val Ser Ile Val Arg Ala Gly Ser Met Glu Gln 290 295 300			912
gga tta aga gac tgt tgt agg tct gtg cgt atc ggt aaa att tta att Gly Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile 305 310 315 320			960
caa agg gac gag gag act gct tta cca aag tta ttc tac gaa aaa tta Gln Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu 325 330 335			1008
cca gag gat ata tct gaa agg tat gtc ttc cta tta gac cca atg ctg Pro Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu 340 345 350			1056
gcc acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga Ala Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg 355 360 365			1104
ggt gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag Gly Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys 370 375 380			1152
gaa ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt Glu Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val 385 390 395 400			1200
act ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca Thr Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro 405 410 415			1248
ggg ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa Gly Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val 420 425			1287

<210> 14
 <211> 428
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Construcción Sintética

<400> 14

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
 130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
 145 150 155 160

Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
 165 170 175

Gly Ala Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr
 180 185 190

Tyr Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg

182

195					200					205					
Asn	Leu	Gln	Lys	Met	Ser	Ser	Glu	Pro	Phe	Lys	Asn	Val	Tyr	Leu	Leu
210						215					220				
Pro	Gln	Thr	Asn	Gln	Leu	Leu	Gly	Leu	Tyr	Thr	Ile	Ile	Arg	Asn	Lys
225					230					235					240
Asn	Thr	Thr	Arg	Pro	Asp	Phe	Ile	Phe	Tyr	Ser	Asp	Arg	Ile	Ile	Arg
				245					250					255	
Leu	Leu	Val	Glu	Glu	Gly	Leu	Asn	His	Leu	Pro	Val	Gln	Lys	Gln	Ile
			260					265					270		
Val	Glu	Thr	Asp	Thr	Asn	Glu	Asn	Phe	Glu	Gly	Val	Ser	Phe	Met	Gly
			275					280					285		
Lys	Ile	Cys	Gly	Val	Ser	Ile	Val	Arg	Ala	Gly	Glu	Ser	Met	Glu	Gln
	290					295					300				
Gly	Leu	Arg	Asp	Cys	Cys	Arg	Ser	Val	Arg	Ile	Gly	Lys	Ile	Leu	Ile
305						310					315				320
Gln	Arg	Asp	Glu	Glu	Thr	Ala	Leu	Pro	Lys	Leu	Phe	Tyr	Glu	Lys	Leu
				325					330					335	
Pro	Glu	Asp	Ile	Ser	Glu	Arg	Tyr	Val	Phe	Leu	Leu	Asp	Pro	Met	Leu
				340					345					350	
Ala	Thr	Gly	Gly	Ser	Ala	Ile	Met	Ala	Thr	Glu	Val	Leu	Ile	Lys	Arg
			355				360					365			
Gly	Val	Lys	Pro	Glu	Arg	Ile	Tyr	Phe	Leu	Asn	Leu	Ile	Cys	Ser	Lys
	370					375					380				
Glu	Gly	Ile	Glu	Lys	Tyr	His	Ala	Ala	Phe	Pro	Glu	Val	Arg	Ile	Val
385						390					395				400
Thr	Gly	Ala	Leu	Asp	Arg	Gly	Leu	Asp	Glu	Asn	Lys	Tyr	Leu	Val	Pro
				405					410					415	
Gly	Leu	Gly	Asp	Phe	Gly	Asp	Arg	Tyr	Tyr	Cys	Val				
			420					425							

<210> 15
 <211> 1200

183

<212> ADN

<213> Secuencia Artificial

<220>

<223> Construcción de fusión - CDopt3 - OPRT

<220>

<221> CDS

<222> (1)..(1200)

<400> 15

atg	gtg	acc	ggc	ggc	atg	gcc	tcc	aag	tgg	gat	caa	aag	ggc	atg	gat	48
Met	Val	Thr	Gly	Gly	Met	Ala	Ser	Lys	Trp	Asp	Gln	Lys	Gly	Met	Asp	
1			5						10					15		

atc	gct	tac	gag	gag	gcc	ctg	ctg	ggc	tac	aag	gag	ggc	ggc	gtg	cct	96
Ile	Ala	Tyr	Glu	Glu	Ala	Leu	Leu	Gly	Tyr	Lys	Glu	Gly	Gly	Val	Pro	
		20						25					30			

atc	ggc	ggc	tgt	ctg	atc	aac	aac	aag	gac	ggc	agt	gtg	ctg	ggc	agg	144
Ile	Gly	Gly	Cys	Leu	Ile	Asn	Asn	Lys	Asp	Gly	Ser	Val	Leu	Gly	Arg	
	35					40						45				

ggc	cac	aac	atg	agg	ttc	cag	aag	ggc	tcc	gcc	acc	ctg	cac	ggc	gag	192
Gly	His	Asn	Met	Arg	Phe	Gln	Lys	Gly	Ser	Ala	Thr	Leu	His	Gly	Glu	
	50					55					60					

atc	tcc	acc	ctg	gag	aac	tgt	ggc	agg	ctg	gag	ggc	aag	gtg	tac	aag	240
Ile	Ser	Thr	Leu	Glu	Asn	Cys	Gly	Arg	Leu	Glu	Gly	Lys	Val	Tyr	Lys	
65					70				75					80		

gac	acc	acc	ctg	tac	acc	acc	ctg	tcc	cct	tgt	gac	atg	tgt	acc	ggc	288
Asp	Thr	Thr	Leu	Tyr	Thr	Thr	Leu	Ser	Pro	Cys	Asp	Met	Cys	Thr	Gly	
			85					90						95		

gct	atc	atc	atg	tac	ggc	atc	cct	agg	tgt	gtg	atc	ggc	gag	aac	gtg	336
Ala	Ile	Ile	Met	Tyr	Gly	Ile	Pro	Arg	Cys	Val	Ile	Gly	Glu	Asn	Val	
			100					105					110			

aac	ttc	aag	tcc	aag	ggc	gag	aag	tac	ctg	caa	acc	agg	ggc	cac	gag	384
Asn	Phe	Lys	Ser	Lys	Gly	Glu	Lys	Tyr	Leu	Gln	Thr	Arg	Gly	His	Glu	
	115						120					125				

gtg	gtg	gtt	gtt	gac	gat	gag	agg	tgt	aag	aag	ctg	atg	aag	cag	ttc	432
Val	Val	Val	Val	Asp	Asp	Glu	Arg	Cys	Lys	Lys	Leu	Met	Lys	Gln	Phe	
	130					135					140					

atc	gac	gag	agg	cct	cag	gac	tgg	ttc	gag	gat	atc	ggc	gag	gag	gtc	480
Ile	Asp	Glu	Arg	Pro	Gln	Asp	Trp	Phe	Glu	Asp	Ile	Gly	Glu	Ala	Val	
145					150				155					160		

gct	cgt	gca	gct	ttg	ggg	cca	ttg	gtg	acg	ggt	ctg	tac	gac	gtg	cag	528
Ala	Arg	Ala	Ala	Leu	Gly	Pro	Leu	Val	Thr	Gly	Leu	Tyr	Asp	Val	Gln	
			165					170						175		

gct	ttc	aag	ttt	ggg	gac	ttc	gtg	ctg	aag	agc	ggg	ctt	tcc	tcc	ccc	576
Ala	Phe	Lys	Phe	Gly	Asp	Phe	Val	Leu	Lys	Ser	Gly	Leu	Ser	Ser	Pro	
		180						185					190			

atc	tac	atc	gat	ctg	cgg	ggc	atc	gtg	tct	cga	ccg	cgt	ctt	ctg	agt	624
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

184

Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg Leu Leu Ser	
195 200 205	
cag gtt gca gat att tta ttc caa act gcc caa aat gca ggc atc agt	672
Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala Gly Ile Ser	
210 215 220	
ttt gac acc gtg tgt gga gtg cct tat aca gct ttg cca ttg gct aca	720
Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro Leu Ala Thr	
225 230 235 240	
gtt atc tgt tca acc aat caa att cca atg ctt att aga agg aaa gaa	768
Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg Arg Lys Glu	
245 250 255	
aca aag gat tat gga act aag cgt ctt gta gaa gga act att aat cca	816
Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr Ile Asn Pro	
260 265 270	
gga gaa acc tgt tta atc att gaa gat gtt gtc acc agt gga tct agt	864
Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser Gly Ser Ser	
275 280 285	
gtt ttg gaa act gtt gag gtt ctt cag aag gag ggc ttg aag gtc act	912
Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu Lys Val Thr	
290 295 300	
gat gcc ata gtg ctg ttg gac aga gag cag gga ggc aag gac aag ttg	960
Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys Asp Lys Leu	
305 310 315 320	
cag gcg cac ggg atc cgc ctc cac tca gtg tgt aca ttg tcc aaa atg	1008
Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu Ser Lys Met	
325 330 335	
ctg gag att ctc gag cag cag aaa aaa gtt gat gct gag aca gtt ggg	1056
Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu Thr Val Gly	
340 345 350	
aga gtg aag agg ttt att cag gag aat gtc ttt gtg gca gcg aat cat	1104
Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala Ala Asn His	
355 360 365	
aat ggt tct ccc ctt tct ata aag gaa gca ccc aaa gaa ctc agc ttc	1152
Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu Leu Ser Phe	
370 375 380	
ggt gca cgt gca gag ctg ccc agg atc cac cca gtt gca tcg aag taa	1200
Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala Ser Lys	
385 390 395	

<210> 16
 <211> 399
 <212> PRT
 <213> **Secuencia Artificial**

 <220>
 <223> **Construcción Sintética**

 <400> 16

185

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ala Val
145 150 155 160

Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr Asp Val Gln
165 170 175

Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu Ser Ser Pro
180 185 190

Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg Leu Leu Ser
195 200 205

Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala Gly Ile Ser
210 215 220

Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro Leu Ala Thr
225 230 235 240

Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg Arg Lys Glu
245 250 255

Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr Ile Asn Pro
260 265 270

Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser Gly Ser Ser
275 280 285

Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu Lys Val Thr
290 295 300

Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys Asp Lys Leu
305 310 315 320

Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu Ser Lys Met
325 330 335

Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu Thr Val Gly
340 345 350

Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala Ala Asn His
355 360 365

Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu Leu Ser Phe
370 375 380

Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala Ser Lys
385 390 395

<210> 17
<211> 1260
<212> ADN
<213> Secuencia Artificial

<220>
<223> Construcción de fusión - CDopt3 - Ligador - OPRT

<220>
<221> CDS
<222> (1)..(1260)

<400> 17
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat 48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct 96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg 35 40 45	144
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu 50 55 60	192
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys 65 70 75 80	240
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly 85 90 95	288
gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val 100 105 110	336
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu 115 120 125	384
gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe 130 135 140	432
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tcc ggc Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly 145 150 155 160	480
ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly 165 170 175	528
ggc gcc gcg gtc gct cgt gca gct ttg ggg cca ttg gtg acg ggt ctg Gly Ala Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu 180 185 190	576
tac gac gtg cag gct ttc aag ttt ggg gac ttc gtg ctg aag agc ggg Tyr Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly 195 200 205	624
ctt tcc tcc ccc atc tac atc gat ctg cgg ggc atc gtg tct cga ccg Leu Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro 210 215 220	672
cgt ctt ctg agt cag gtt gca gat att tta ttc caa act gcc caa aat Arg Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn 225 230 235 240	720
gca ggc atc agt ttt gac acc gtg tgt gga gtg cct tat aca gct ttg Ala Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu 245 250 255	768
cca ttg gct aca gtt atc tgt tca acc aat caa att cca atg ctt att Pro Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile 260 265 270	816

aga agg aaa gaa aca aag gat tat gga act aag cgt ctt gta gaa gga Arg Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly 275 280 285	864
act att aat cca gga gaa acc tgt tta atc att gaa gat gtt gtc acc Thr Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr 290 295 300	912
agt gga tct agt gtt ttg gaa act gtt gag gtt ctt cag aag gag ggc Ser Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly 305 310 315 320	960
ttg aag gtc act gat gcc ata gtg ctg ttg gac aga gag cag gga ggc Leu Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly 325 330 335	1008
aag gac aag ttg cag gcg cac ggg atc cgc ctc cac tca gtg tgt aca Lys Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr 340 345 350	1056
ttg tcc aaa atg ctg gag att ctc gag cag cag aaa aaa gtt gat gct Leu Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala 355 360 365	1104
gag aca gtt ggg aga gtg aag agg ttt att cag gag aat gtc ttt gtg Glu Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val 370 375 380	1152
gca gcg aat cat aat ggt tct ccc ctt tct ata aag gaa gca ccc aaa Ala Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys 385 390 395 400	1200
gaa ctc agc ttc ggt gca cgt gca gag ctg ccc agg atc cac cca gtt Glu Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val 405 410 415	1248
gca tcg aag taa Ala Ser Lys	1260

<210> 18
 <211> 419
 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Construcción Sintética

 <400> 18

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp 1 5 10 15
--

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg 35 40 45

Gly	His	Asn	Met	Arg	Phe	Gln	Lys	Gly	Ser	Ala	Thr	Leu	His	Gly	Glu	50	55	60	
Ile	Ser	Thr	Leu	Glu	Asn	Cys	Gly	Arg	Leu	Glu	Gly	Lys	Val	Tyr	Lys	65	70	75	80
Asp	Thr	Thr	Leu	Tyr	Thr	Thr	Leu	Ser	Pro	Cys	Asp	Met	Cys	Thr	Gly	85	90	95	
Ala	Ile	Ile	Met	Tyr	Gly	Ile	Pro	Arg	Cys	Val	Ile	Gly	Glu	Asn	Val	100	105	110	
Asn	Phe	Lys	Ser	Lys	Gly	Glu	Lys	Tyr	Leu	Gln	Thr	Arg	Gly	His	Glu	115	120	125	
Val	Val	Val	Val	Asp	Asp	Glu	Arg	Cys	Lys	Lys	Leu	Met	Lys	Gln	Phe	130	135	140	
Ile	Asp	Glu	Arg	Pro	Gln	Asp	Trp	Phe	Glu	Asp	Ile	Gly	Glu	Ser	Gly	145	150	155	160
Gly	Gly	Ala	Ser	Gly	Gly	Gly	Ala	Ser	Gly	Gly	Gly	Ala	Ser	Gly	Gly	165	170	175	
Gly	Ala	Ala	Val	Ala	Arg	Ala	Ala	Leu	Gly	Pro	Leu	Val	Thr	Gly	Leu	180	185	190	
Tyr	Asp	Val	Gln	Ala	Phe	Lys	Phe	Gly	Asp	Phe	Val	Leu	Lys	Ser	Gly	195	200	205	
Leu	Ser	Ser	Pro	Ile	Tyr	Ile	Asp	Leu	Arg	Gly	Ile	Val	Ser	Arg	Pro	210	215	220	
Arg	Leu	Leu	Ser	Gln	Val	Ala	Asp	Ile	Leu	Phe	Gln	Thr	Ala	Gln	Asn	225	230	235	240
Ala	Gly	Ile	Ser	Phe	Asp	Thr	Val	Cys	Gly	Val	Pro	Tyr	Thr	Ala	Leu	245	250	255	
Pro	Leu	Ala	Thr	Val	Ile	Cys	Ser	Thr	Asn	Gln	Ile	Pro	Met	Leu	Ile	260	265	270	
Arg	Arg	Lys	Glu	Thr	Lys	Asp	Tyr	Gly	Thr	Lys	Arg	Leu	Val	Glu	Gly	275	280	285	

Thr Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr
290 295 300

Ser Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly
305 310 315 320

Leu Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly
325 330 335

Lys Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr
340 345 350

Leu Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala
355 360 365

Glu Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val
370 375 380

Ala Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys
385 390 395 400

Glu Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val
405 410 415

Ala Ser Lys

<210> 19
<211> 11892
<212> ADN
<213> Sequencia Artificial

<220>
<223> Vector RCR - pAC3 - yCD2

<400> 19
tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc ccgcccatt 120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180
atgggtggag tatttacggt aaactgcccc cttggcagta catcaagtgt atcatatgcc 240
aagtaagccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgtattac 360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480

ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaatgggcg	gtaggcgtgt	540
acgggtgggag	gtctatataa	gcagagctgg	tttagtgaac	cggcgccagt	cctccgattg	600
actgagtcgc	ccgggtaccc	gtgtatccaa	taaaccctct	tgcagttgca	tccgacttgt	660
ggctctcgctg	ttccttggga	gggtctctct	tgagtgattg	actaccctgc	agcgggggtc	720
tttcatttgg	gggctcgtec	gggatcggga	gacccctgcc	cagggaccac	cgacccacca	780
cggggaggta	agctggccag	caacttatct	gtgtctgtcc	gattgtctag	tgtctatgac	840
tgattttatg	cgcttgcgtc	ggtactagtt	agctaactag	ctctgtatct	ggcggaccgc	900
tggtggaact	gacgagttcg	gaacaccccg	ccgcaaccct	gggagacgtc	ccagggactt	960
cggggggccgt	ttttgtggcc	cgacctgagt	ccaaaaatcc	cgatcgtttt	ggactctttg	1020
gtgcaccccc	cttagaggag	ggatatgtgg	ttctggtagg	agacgagAAC	ctaaaacagt	1080
tcccgctcc	gtctgaattt	ttgctttcgg	tttgggaccg	aagccgcgcc	gcgcgtcttg	1140
tctgctgcag	catcgttctg	tgttgtctct	gtctgactgt	gtttctgtat	ttgtctgaga	1200
atatgggcca	gactgttacc	actcccttaa	gtttgacctt	aggtcactgg	aaagatgtcg	1260
agcggatcgc	tcacaaccag	tcggtagatg	tcaagaagag	acgttgggtt	accttctgct	1320
ctgcagaatg	gccaaccttt	aacgtcggat	ggccgcgaga	cggcaccttt	aaccgagacc	1380
tcatcaccca	ggttaagatc	aaggtctttt	cacctggccc	gcattggacac	ccagaccagg	1440
tccctacat	cgtgacctgg	gaagccttgg	cttttgaccc	ccctccctgg	gtcaagccct	1500
ttgtacaccc	taagcctccg	cctcctcttc	ctccatccgc	cccgctctct	cccttgaac	1560
ctcctcgttc	gaccccgccct	cgatccctcc	tttatccagc	cctcactcct	tctctaggcg	1620
ccaaacctaa	acctcaagtt	ctttctgaca	gtggggggcc	gtcctcgcac	ctacttacag	1680
aagaccccc	gccttatagg	gacccaagac	caccccttcc	cgacagggac	ggaaatggtg	1740
gagaagcgac	ccctgcggga	gaggcaaccg	acccctcccc	aatggcatct	cgcctacgtg	1800
ggagacggga	gccccctgtg	gccgactcca	ctacctcgca	ggcattcccc	ctccgcgcag	1860
gaggaaacgg	acagcttcaa	tactggccgt	tctcctcttc	tgacctttac	aactggaaaa	1920
ataataaccc	ttctttttct	gaagatccag	gtaaactgac	agctctgac	gagttctgtc	1980
tcatcaccca	tcagcccacc	tgggacgact	gtcagcagct	gttggggact	ctgctgaccg	2040
gagaagaaaa	acaacgggtg	ctcttagagg	ctagaaaggc	ggtgcggggc	gatgatgggc	2100
gccccactca	actgcccact	gaagtccatg	ccgtttttcc	cctcgagcgc	ccagactggg	2160
attacaccac	ccaggcaggt	aggaaccacc	tagtccacta	tcgccagttg	ctcctagcgg	2220
gtctccaaaa	cgcgggcaga	agccccacca	atttggccaa	ggtaaaagga	ataacacaag	2280

ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca	2340
ctccttatga cctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc	2400
agtctgcccc agacattggg agaaagttag agagggttaga agatttataaa aacaagacgc	2460
ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa	2520
gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg	2580
agcagaaaga gaaagaaaga gatcgttaga gacatagaga gatgagcaag ctattggcca	2640
ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg	2700
atcgcgacca gtgtgcctac tgcaaagaaa aggggcaactg ggctaaagat tgtcccaaga	2760
aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgactagg	2820
gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac	2880
ccgtcacctt cctggtagat actggggccc aacactccgt gctgaccca aatcctggac	2940
ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaagcgg tatcgttga	3000
ccacggatcg caaagtacat ctagctacgg gtaaggctac ccactcttct ctccatgtac	3060
cagactgtcc ctatcctctg ttaggaagag atttgcctgac taaactaaaa gcccaaatcc	3120
actttgaggg atcaggagcc cagggttatg gaccaatggg gcagcccctg caagtgttga	3180
ccctaaatat agaagatgag catcggtac atgagacctc aaaagagcca gatgtttctc	3240
tagggccac atggctgtct gattttctct aggcctgggc ggaaaccggg ggcattggac	3300
tggcagttcg ccaagctcct ctgatcatc ctctgaaagc aacctctacc cccgtgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420
tgttggaaca gggaatactg gtacctgccc agtccccctg gaacacgccc ctgctaccgg	3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc	3540
gggtggaaga catccacccc accgtgccc acccttaca cctcttgagc gggctccac	3600
cgccccacca gtggtacact gtgcttgatt taaaggatgc cttttctgct ctgagactcc	3660
accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag	3720
gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg	3780
aggcactgca cagagacctc gcagacttcc ggatccagca cccagacttg atcctgctac	3840
agtaagtga tgaactactg ctggccgcca cttctgagct agactgcca caaggtactc	3900
gggccctgtt acaaacccta gggaacctcg ggtatcgggc ctcggccaaag aaagcccaaa	3960
tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga	4020
ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccctc cgacaactaa	4080
gggagtccct agggacggca ggcttctgct gcctctggat cctggggtt gcagaaatgg	4140

cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac	4200
aaaaggccta tcaagaaatc aagcaagctc ttctaactgc ccagccctg gggttgccag	4260
atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaagggtgcc	4320
taacgcaaaa actgggacct tggcgctggc cggtaggcta cctgtccaaa aagctagacc	4380
cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgacaa	4440
aggatgcagg caagctaacc atgggacagc cactagtcac tctggccccc catgcagtag	4500
aggcactagt caaacaaccc cccgaccgct ggctttccaa cggccggatg actcactatc	4560
aggccttgct tttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaacccgg	4620
ctacgtgctt cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg	4680
aagcccacgg aacccgacct gacctaacgg accagccgct cccagacgcc gaccacacct	4740
ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga	4800
ccaccgagac cgaggtaatc tgggctaaag cctgtccagc cgggacatcc gctcagcggg	4860
ctgaactgat agcactcacc caggccctaa agatggcaga aggtaagaag ctaaattgtt	4920
atactgatag ccgttatgct tttgtactg cccatatcca tggagaaata tacagaaggc	4980
gtgggttgct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac	5040
taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg	5100
gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccgga aaggcagcca	5160
tcacagagac tccagacacc tctaccctcc tcatagaaaa ttcacacccc tacacctcag	5220
aacattttca ttacacagtg actgatataa aggacctaac caagttgggg gccatttatg	5280
ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt	5340
ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc	5400
tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata	5460
tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg	5520
gaactagggc ccgcgggcat cggcccgga ctcattggga gatcgatttc accgagataa	5580
agcccggtt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga	5640
tagaagcctt cccaaccaag aaagaaaccg ccaaggctcg aaccaagaag ctactagagg	5700
agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg	5760
tctccaaggc gagtcagaca gtggccgata tgttggggat tgattggaaa ttacattgtg	5820
catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt	5880
taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc	5940

tgtaccgagc cgcgaacacg cggggccccc atggcctcac cccatatgag atcttatatg	6000
gggcaccccc gcccccttgta aacttccctg accctgacat gacaagagtt actaacagcc	6060
cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc	6120
tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagtcg	6180
gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac	6240
cttacacagt cctgctgacc acccccaccg cctcaaaagt agacggcctc gcagcttggg	6300
tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat	6360
ggcgcgttca acgctctcaa aacccctca agataagatt aaccctgga agcccttaat	6420
agtcattgga gtccctgttag gagtagggat ggacagagagc ccccatcagg tctttaatgt	6480
aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctcctggg	6540
aactgtacaa gatgccttcc caaaattata tttgatcta tgtgatctgg tcggagagga	6600
gtgggacctt tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag	6660
acagcggacc cggacttttg acttttacgt gtgcctggg cataccgtaa agtcgggggtg	6720
tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta	6780
ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca cccctggga	6840
cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc	6900
cttccaaggg gctactcgag ggggcagatg caacctcta gtccatagaat tcaactgatgc	6960
aggaaaaaag gctaactggg acgggcccac atcgtgggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt	7080
cccataggg cccaaccag tattaccga ccaaagactc ccttctcac caatagagat	7140
tgtaccggct ccacagccac ctgccccct caataccagt tccccctt ccactaccag	7200
tacacctca acctcccta caagtccaag tgtccacag ccaccccag gaactggaga	7260
tagactacta gctctagtca aaggagcta tcaggcgctt aacctacca atcccacaa	7320
gacccaagaa tgttggctgt gcttagtgtc gggacctct tattacgaag gagtagcgt	7380
cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca ctcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggag tacctaaac	7500
tcaccaggcc ttatgtaaca ccaccaaag cgcggctca ggatcctact acctgcagc	7560
acccgcggga acaatgtggg cttgcagcac tggattgact cctgcttgt ccaccagggt	7620
gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca	7680
ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attraaccatg ggagggattg cagctggaat	7800

agggacgggg accactgcct taattaaaac ccagcagttt gagcagette atgccgctat	7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc	7920
gttgctctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg	7980
aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt	8040
gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
aggccaagga tggttogaag ggctgtttaa tagatcccc tggtttacca ccttaatctc	8160
caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct	8220
caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac	8280
tcagcaatat caccagctaa aaccataga gtacgagcca tgaacgcgtt actggccgaa	8340
gccgcttga ataaggccgg tgtgcgtttg tctatatgtt atttccacc atattgccgt	8400
cttttgcaa tgtgagggcc cggaaacctg gccctgtctt cttgacgagc attcctaggg	8460
gtctttcccc tctcgccaaa ggaatgcaag gtctgttgaa tgtcgtgaag gaagcagttc	8520
ctctggaagc ttcttgaaga caaacaacgt ctgtagcgac cctttgcagg cagcggaaac	8580
ccccacctgg cgacagggtg ctctgcggcc aaaagccacg tgtataagat acacctgcaa	8640
aggcggcaca accccagtgc cacgttgtga gttggatagt tgtggaaaga gtcaaatggc	8700
tctcctcaag cgtattcaac aaggggctga aggatgcccc gaagggtacc cattgtatgg	8760
gatctgatct ggggcctcgg tgcacatgct ttacatgtgt ttagtcgagg ttaaaaaaac	8820
gtctaggccc ccggaaccac ggggacgtgg ttttcctttg aaaaacacga ttataaatgg	8880
tgaccggcgg catggcctcc aagtgggac aaaagggcat ggatatecgt tacgaggagg	8940
cctgtctggg ctacaaggag ggcggcgtgc ctatcggcgg ctgtctgac aacaacaagg	9000
acggcagtgt gctgggcagg ggccacaaca tgagggtcca gaagggtcc gccacctgc	9060
acggcgagat ctccacctg gagaactgtg gcaggctgga gggcaagggtg tacaaggaca	9120
ccacctgta caccacctg tccccctgtg acatgtgtac cggcgctatc atcatgtacg	9180
gcatccctag gtgtgtgac ggcgagaacg tgaacttcaa gtccaagggc gagaagtacc	9240
tgcaaacag gggccacgag gtggtggttg ttgacgatga gaggtgtaag aagctgatga	9300
agcagttcat cgacgagagg cctcaggact ggttcgagga tatcggcgag taagcggccg	9360
cagataaaat aaaagatttt atttagtctc cagaaaaagg ggggaatgaa agacccacc	9420
tgtaggtttg gcaagctagc ttaagtaacg ccattttgca aggcattgaa aaatacataa	9480
ctgagaatag agaagttcag atcaaggta ggaacagatg gaacagctga atatgggcca	9540
aacaggatat ctgtggtaag cagttcctgc cccggctcag ggccaagaac agatggaaca	9600

gctgaatatg ggccaaacag gatatctgtg gtaagcagtt cctgccccgg ctgagggcca	9660
agaacagatg gtccccagat gcgggtccagc cctcagcagt ttctagagaa ccatcagatg	9720
tttccagggg gccccaagga cctgaaatga ccctgtgcct tatttgaact aaccaatcag	9780
ttcgctttct gcttctgttc gcgcgcttct gctccccgag ctcaataaaa gagcccacaa	9840
ccctcactc ggggcgccag tctctcgatt gactgagtcg cccgggtacc cgtgtatcca	9900
ataaacctc ttgcagttgc atccgacttg tggctctcgt gttccttggg aggggtctct	9960
ctgagtgatt gactaccgt cagcgggggt ctttcattac atgtgagcaa aaggccagca	10020
aaaggccagg aaccgtaaaa aggcgcggtt gctggcggtt ttccataggc tccgcccccc	10080
tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaaccgga caggactata	10140
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc	10200
gcttaccgga tacctgtccg cctttctccc ttccgggaagc gtggcgcttt ctcaatgtc	10260
acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga	10320
acccccggt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc	10380
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag	10440
gtatgtaggc ggtgctacag agttcttgaa gtgggtggct aactacggct acactagaag	10500
gacagtatth ggtatctgcg ctctgctgaa gccagttacc ttccgaaaaa gagttggtag	10560
ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgt ttttttgttt gcaagcagca	10620
gattacgcgc agaaaaaag gatctcaaga agatcctttg atcttttcta cggggctctga	10680
cgctcagtg aacgaaaact cacgttaagg gattttggct atgagattat caaaaaggat	10740
cttcacctag atccttttaa attaaaaatg aagtttttaa tcaatctaaa gtatatatga	10800
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg	10860
tctatttcgt tcatccatag ttgcctgaat ccccgctgtg tagataacta cgatacggga	10920
gggcttacca tctggccccg gtgctgcaat gataccgcga gaccacgct caccggctcc	10980
agatttatca gcaataaacc agccagccgg aaggggccgag cgcagaaagt gtccctgcaac	11040
tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagtccgcc	11100
agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtggtgt cacgctcgtc	11160
gtttggtagt gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc	11220
catgttgtgc aaaaaagcgg ttagctcctt cggctctccg atcgttgtca gaagtaagtt	11280
ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc	11340
atccgtaaga tgcttttctg tgactgggtg gtactcaacc aagtcattct gagaatagtg	11400
tatggcgga ccgagttgct cttgcccggc gtcaacacgg gataataccg cgccacatag	11460

cagaacttta aaagtgetca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat 11520
 cttaccgctg ttgagatcca gttegatgta acccactcgt gcacccaact gatcttcagc 11580
 atcttttact ttcaccagcg tttctgggtg agcaaaaaaca ggaaggcaaa atgccgcaaa 11640
 aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta 11700
 ttgaagcatt tatcaggggtt attgtctcat gagcggatac atatttgaat gtatttagaa 11760
 aaataaacia ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga 11820
 aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggg cctttcgtct 11880
 tcaagaattc at 11892

<210> 20
 <211> 11892
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Vector RCR - pAC3 - yCD

<400> 20
 tagttattaa tagtaatcaa ttacgggggc attagttcat agcccatata tggagttccg 60
 cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 120
 gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180
 atgggtggag tatttacggg aaactgccc cttggcagta catcaagtgt atcatatgcc 240
 aagtaagccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta 300
 catgacctta tgggactttc ctacttgcca gtacatctac gtattagtca tcgctattac 360
 catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420
 atttccaagt ctcaccccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480
 ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggag gtaggcgtgt 540
 acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600
 actgagtcgc ccgggtaccc gtgtatccaa taaacctct tgcagttgca tccgacttgt 660
 ggtctcgtg ttccttggga gggctctctc tgagtgttg actaccgctc agcgggggtc 720
 tttcatttgg gggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca 780
 ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
 tgattttatg cgctgcgtc ggtactagtt agctaactag ctctgtatct ggcggaccgg 900
 tggtggaact gacgagttcg gaacacccgg ccgcaacct gggagacgtc ccagggactt 960
 cgggggcccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020

gtgcaccccc	cttagaggag	ggatatgtgg	ttctggtagg	agacgagaac	ctaaaacagt	1080
tccgcctcc	gtctgaattt	ttgctttcgg	tttgggaccg	aagccgcgcc	gcgcgtcttg	1140
tctgctgcag	catcgttctg	tgttgctctt	gtctgactgt	gtttctgtat	ttgtctgaga	1200
atatgggcca	gactgttacc	actcccttaa	gtttgacctt	aggtcactgg	aaagatgtcg	1260
agcggatcgc	tcacaaccag	tcggtagatg	tcaagaagag	acgttggggt	accttctgct	1320
ctgcagaatg	gccaaccttt	aacgtcggat	ggccgcgaga	cggcaccttt	aaccgagacc	1380
tcatacccca	ggttaagatc	aaggtctttt	cacctggccc	gcattggacac	ccagaccagg	1440
tcccctacat	cgtgacctgg	gaagccttgg	cttttgacct	ccctccctgg	gtcaagccct	1500
ttgtacaccc	taagcctccg	cctcctcttc	ctccatccgc	ccgtctcttc	cccttgaac	1560
ctcctcgttc	gaccccgct	cgatcctccc	tttatccagc	cctcactcct	tctctaggcg	1620
ccaaacctaa	acctcaagtt	ctttctgaca	gtggggggcc	gctcatcgac	ctacttacag	1680
aagaccccc	gccttatagg	gacccaagac	caccccttc	cgacagggac	ggaaatggtg	1740
gagaagcgac	ccctgcggga	gaggcaccgg	acccctcccc	aatggcatct	cgcctacgtg	1800
ggagacggga	gccccctgtg	gccgactcca	ctacctcgca	ggcattcccc	ctccgcgcag	1860
gaggaaacgg	acagcttcaa	tactggcctg	tctcctcttc	tgacctttac	aactggaaaa	1920
ataataaccc	ttctttttct	gaagatccag	gtaaactgac	agctctgac	gagtctgttc	1980
tcatacccca	tcagcccacc	tgggacgaat	gtcagcagct	gttggggact	ctgctgaccg	2040
gagaagaaaa	acaacgggtg	ctcttagagg	ctagaaaggc	ggtgcggggc	gatgatgggc	2100
gccccactca	actgcccatt	gaagtccgat	ccgcttttcc	cctcgagcgc	ccagactggg	2160
attacaccac	ccaggcaggt	aggaaccacc	tagtccacta	tcgccagttg	ctcctagcgg	2220
gtctccaaaa	cgcgggcaga	agccccacca	atttggccaa	ggtaaaagga	ataacacaag	2280
ggcccaatga	gtctccctcg	gccttcctag	agagacttaa	ggaagcctat	cgacaggtaca	2340
ctccttatga	ccctgaggac	ccagggcaag	aaactaatgt	gtctatgtct	ttcatttggc	2400
agtctgcccc	agacattggg	agaaagttag	agagggttaga	agatttataa	aacaagacgc	2460
ttggagattt	ggttagagag	gcagaaaaa	tctttaataa	acgagaaacc	ccggaagaaa	2520
gagaggaacg	tatcaggaga	gaaacagagg	aaaaagaaga	acgccgtagg	acagaggatg	2580
agcagaaaga	gaaagaaaga	gatcgttaga	gacatagaga	gatgagcaag	ctattggcca	2640
ctgtcgttag	tggacagaaa	caggatagac	agggaggaga	acgaaggagg	tcccaactcg	2700
atcgcgacca	gtgtgcctac	tgcaagaaa	aggggcactg	ggctaaagat	tgtcccaaga	2760
aaccacgagg	acctcgggga	ccaagacccc	agacctccct	cctgacccta	gatgactagg	2820
gaggtcaggg	tcaggagccc	ccccctgaac	ccaggataac	cctcaaagtc	ggggggcaac	2880

ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac	2940
ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgctgga	3000
ccacggatcg caaagtacat ctgctaccg gtaaggtcac ccactcttcc ctccatgtac	3060
cagactgtcc ctatcctctg ttaggaagag atttgtgtac taaactaaaa gcccaaattcc	3120
actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga	3180
ccctaaatat agaagatgag catcggttac atgagacctc aaaagagcca gatgtttctc	3240
taggggccac atggctgtct gattttctc aggccctggg ggaaaccggg ggcctgggac	3300
tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420
tgttggacca gggaatactg gtacctgtcc agtccccctg gaacacgccc ctgctacctg	3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc	3540
gggtggaaga catccacccc accgtgccc acccttaca cctcttgagc gggctccac	3600
cgteccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc	3660
accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag	3720
gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtccacc ctgtttgatg	3780
aggcactgca cagagacctc gcagacttcc ggatccagca cccagacttg atcctgtac	3840
agtacgtgga tgacttactg ctggccgcca cttctgagct agactgcca caaggtactc	3900
gggcccctgtt acaaacctc gggaacctc ggtatcggg ctcggccaag aaagcccaaa	3960
tttgccagaa acaggtcaag tatctgggt atcttctaaa agagggtcag agatggctga	4020
ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccct cgacaactaa	4080
gggagttcct agggacggca ggcttctgt gcctctggat ccctgggttt gcagaaatgg	4140
cagccccctt gtacctctc accaaaacgg ggactctgtt taattggggc ccagaccaac	4200
aaaaggccta tcaagaaatc aagcaagtc ttctaactgc cccagccctg gggttgccag	4260
atttgactaa gccctttgaa ctctttgtc acgagaagca gggctacgcc aaagggtgtcc	4320
taacgcaaaa actgggacct tggcgtcggc cgggtggccta cctgtccaaa aagctagacc	4380
cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgacaa	4440
aggatgcagg caagctaacc atgggacagc cactagtcat tctggcccc catgcagtag	4500
aggcactagt caaacaaccc cccgaccgt ggctttccaa cggccggatg actcactatc	4560
aggccttgct tttggacag gaccgggtcc agttcggacc ggtggtagcc ctgaaccgg	4620
ctacgtgct cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg	4680

aagcccacgg	aacccgaccc	gacctaacgg	accagccgct	cccagacgcc	gaccacacct	4740
ggtacacgga	tggaagcagt	ctcttacaag	agggacagcg	taaggcgggg	gctgcggtga	4800
ccaccgagac	cgaggtaatc	tgggctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
ctgaactgat	agcactcacc	caggccctaa	agatggcaga	aggtaagaag	ctaaatgttt	4920
atactgatag	cogttatgct	tttgctactg	cccatatcca	tggagaaata	tacagaaggc	4980
gtgggttgct	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100
gacacagcgc	cgaggctaga	ggcaaccgga	tggctgacca	agcggcccca	aaggcagcca	5160
tcacagagac	tccagacacc	tctaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
aacattttca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccatttatg	5280
ataaaacaaa	gaagtattgg	gtctaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
ttgaattatt	agactttctt	catcagctga	ctcacctcag	cttctcaaaa	atgaaggctc	5400
tcctagagag	aagccacagt	ccctactaca	tgtgaaccg	ggatcgaaca	ctcaaaaata	5460
tcactgagac	ctgcaaagct	tgtgcacaag	tcaacgccag	caagtctgcc	gttaaacagg	5520
gaactagggg	ccgcgggcat	cggcccgga	ctcatlggga	gatcgatttc	accgagataa	5580
agcccggatt	gtatggctat	aaatatcttc	tagtttttat	agataccttt	tctggctgga	5640
tagaagcett	cccaaccaag	aaagaaaccg	ccaaggctgt	aaccaagaag	ctactagagg	5700
agatcttccc	caggttcggc	atgcctcagg	tattgggaac	tgacaatggg	cctgccttcg	5760
tctccaaggt	gagtcagaca	gtggccgac	tgttggggat	tgattggaaa	ttacattgtg	5820
catacagacc	ccaaagctca	ggccaggtag	aaagaatgaa	tagaaccatc	aaggagactt	5880
taactaaatt	aacgcttgca	actggctcta	gagactgggt	gctcctactc	cccttagccc	5940
tgtaccgagc	ccgcaacacg	ccgggcccc	atggcctcac	cccatatgag	atcttatatg	6000
gggcaccccc	gccccttgta	aacttccttg	accctgacat	gacaagagtt	actaacagcc	6060
cctctctcca	agctcactta	caggctctct	acttagtcca	gcacgaagtc	tggagacctc	6120
tggcggcagc	ctaccaagaa	caactggacc	gaccggtggt	acctcaccct	taccgagtcg	6180
gcgacacagt	gtgggtccgc	cgacaccaga	ctaagaacct	agaacctcgc	tggaaaggac	6240
cttacacagt	cctgctgacc	acccccaccg	ccctcaaagt	agacggcatc	gcagcttgga	6300
tacacgcgcg	ccacgtgaag	gctgccgacc	ccgggggtgg	accatcctct	agactgacat	6360
ggcgcgttca	acgctctcaa	aacccctca	agataagatt	aacccgtgga	agcccttaat	6420
agtcatggga	gtcctgttag	gagtagggat	ggcagagagc	cccatcagg	tctttaatgt	6480
aacctggaga	gtcaccaacc	tgatgactgg	gcgtaccgcc	aatgccacct	ccctcctggg	6540

201

aactgtacaa gatgccttcc caaaattata ttttgcata tgtgatctgg tcggagagga	6600
gtgggaccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag	6660
acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcgggggtg	6720
tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta	6780
ctggaagccc acatcatcgt gggacctaast ctcccttaag cgcggtaaca cccctggga	6840
cacgggatgc tctaaagtgt cctgtggccc ctgctacgac ctctccaaag tatccaattc	6900
cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tcaactgatgc	6960
aggaaaaaag gctaactggg acgggcccac atcgtgggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcagggtc cttaatgtgg gaccccgagt	7080
ccccataggg cccaaccagc tattaccgca ccaagactc ccttctcac caatagagat	7140
tgtaccggct ccacagccac ctagccccct caataccagt taccctcctt ccactaccag	7200
tacacctca acctcccta caagtccaag tgtccacag ccaccccgag gaactggaga	7260
tagactacta gctctagtca aaggagccta tcaggcgctt aacctacca atcccgacaa	7320
gaccaagaa tgttggtgt gcttagtgtc gggacctcct tattacgaag gagtagcgtt	7380
cgtgggcaact tataccaatc attccaccgc tccggccaac tgtacggcca cttcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac	7500
tcaccaggcc ttatgtaaca ccacccaag cgcgggtca ggatcctact accttgacgc	7560
acccgcggga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacggt	7620
gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taattttacca	7680
ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attaacctatg ggagggattg cagctggaat	7800
agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat	7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc	7920
gttgtctgaa gtagtcctac agaaccgcag aggctagat ttgctattcc taaaggaggg	7980
aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt	8040
gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
aggccaagga tggttcgaag ggctgtttta tagatcccc tggttttacca ccttaatctc	8160
caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct	8220
caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac	8280
tcagcaatat caccagctaa aacctataga gtacgagcca tgaacgcgtt actggccgaa	8340

202

gcccgttggg	ataaggccgg	tgtgcgtttg	tctatatgtt	atittccacc	atattgccgt	8400
cttttggcaa	tgtgagggcc	cggaaacctg	gccctgtctt	cttgacgagc	attcctaggg	8460
gtctttcccc	tctcgccaaa	ggaatgcaag	gtctgttgaa	tgtcgtgaag	gaagcagttc	8520
ctctggaagc	ttottgaaga	caaacaacgt	ctgtagcgac	cctttgcagg	cagcgggaacc	8580
ccccacctgg	cgacaggtgc	ctctgcggcc	aaaagccacg	tgtataagat	acacctgcaa	8640
aggcggcaca	accccagtgc	cacgtttgtg	gttgatagt	tgtggaaaga	gtcaaattgg	8700
tctcctcaag	cgtattcaac	aaggggctga	aggatgcccc	gaaggtaacc	cattgtatgg	8760
gatctgatct	ggggcctcgg	tgcacatgct	ttacatgtgt	ttagtcgagg	ttaaaaaac	8820
gtctaggccc	cccgaaccac	ggggacgtgg	ttttcctttg	aaaaacacga	ttataaatgg	8880
tgacaggggg	aattggcaagc	aagtgggatc	agaagggtat	ggacattgcc	tatgaggagg	8940
cggccttagg	ttacaaagag	ggtggtgttc	ctattggcgg	atgtcttctc	aataacaaag	9000
acggaagtgt	tctcggtcgt	ggtcacaaca	tgagatttca	aaagggatcc	gccacactac	9060
atggtgagat	ctccactttg	gaaaactgtg	ggagattaga	gggcaaagtg	tacaaagata	9120
ccactttgta	tacgacgctg	tctccatgog	acatgtgtac	agggtgccatc	atcatgtatg	9180
gtattccacg	ctgtgttgtc	ggtgagaacg	ttaatttcaa	aagtaagggc	gagaaatatt	9240
tacaaactag	aggtcacgag	gttgttgttg	ttgacgatga	gaggtgtaaa	aagatcatga	9300
aacaatttat	cgatgaaaga	cctcaggatt	ggtttgaaga	tattggtgag	taggcggccg	9360
cagataaaat	aaaagatttt	atttagtctc	cagaaaaagg	ggggaatgaa	agaccccacc	9420
tgtaggtttg	gcaagctagc	ttaagtaacg	ccattttgca	aggcatggaa	aaatacataa	9480
ctgagaatag	agaagttcag	atcaaggcca	ggaacagatg	gaacagctga	atatgggcca	9540
aacaggatat	ctgtggtaag	cagttcctgc	cccggctcag	ggccaagaac	agatggaaca	9600
gctgaatatg	ggccaaacag	gatctctgtg	gtaagcagtt	cctgccccgg	ctcaggggcca	9660
agaacagatg	gtccccagat	gcggtccagc	cctcagcagt	ttctagagaa	ccatcagatg	9720
tttccagggg	gccccaaagg	cctgaaatga	ccctgtgcct	tatttgaact	aaccaatcag	9780
ttcgcttctc	gcttctgttc	gcgcgcttct	gtccccgag	ctcaataaaa	gagcccacaa	9840
cccctcactc	ggggcgccag	tcctccgatt	gactgagtcg	cccgggtacc	cgtgtatcca	9900
ataaaccttc	ttgcagttgc	atccgacttg	tggtctcctc	gttccttggg	agggtctcct	9960
ctgagtgatt	gactaccctg	cagcgggggt	ctttcattac	atgtgagcaa	aaggccagca	10020
aaaggccagg	aaccgtaaaa	aggccgcgtt	gctggcgttt	ttccataggc	tcgccecccc	10080
tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaacccga	caggactata	10140
aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgacctgcc	10200

gcttacccgga tacctgtccg cttttctccc ttccgggaagc gtggcgcttt ctcaatgctc	10260
acgctgtagg tatctcagtt cgggtgtaggt cggtcgctcc aagctgggct gtgtgcacga	10320
acccccggt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc	10380
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag	10440
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag	10500
gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttccgaaaaa gagttggtag	10560
ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgt tttttgttt gcaagcagca	10620
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga	10680
cgctcagtg aacgaaaact cacgttaagg gattttgtc atgagattat caaaaaggat	10740
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga	10800
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg	10860
tctatttcgt tcatccatag ttgcctgact ccccgctcgt tagataacta cgatacggga	10920
gggcttacca tctggcccca gtgctgcaat gataccgcga gaccacgct caccgctcc	10980
agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtctgcaac	11040
tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc	11100
agttaatagt ttgcgcaacg ttgttgcct tctgagcagg atcgtggtgt cacgctcgtc	11160
gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc	11220
catgttgctc aaaaaagcgg ttagctcctt cggctcctcg atcgttgctc gaagtaagtt	11280
ggcgcagtg ttatcactca tgggtatggc agcactgcat aattctctta ctgtcatgcc	11340
atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagt	11400
tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg cgccacatag	11460
cagaacttta aaagtgtcga tcattggaaa acgttcttcg gggcgaaaac tctcaaggat	11520
cttacgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc	11580
atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa	11640
aaagggata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta	11700
ttgaagcatt tatcaggggt attgtctcat gagcggatac atatttgaat gtatttagaa	11760
aaataaacia ataggggttc cgcgcacatt tccccgaaa gtgccacctg acgtctaaga	11820
aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct	11880
tcaagaattc at	11892

204

<211> 12007
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> RCR Vector - pACE-CD

<400> 21
 tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60
 cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc ccgcccatt 120
 gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180
 atgggtggag tatttacggg aaactgcccc ctgggcagta catcaagtgt atcatatgcc 240
 aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta 300
 catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360
 catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420
 atttccaagt ctccacccca ttgacgtcaa tgggagtgtg ttttggccc aaaatcaacg 480
 ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt 540
 acgggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600
 actgagtcgc ccgggtaccc gtgtatccaa taaacctct tgcagttgca tccgacttgt 660
 ggtctcgctg ttccctggga gggctctctc tgagtgttg actaccgtc agcgggggtc 720
 ttccatttg gggctcgtcc gggatcggga gacctcgtcc cagggaccac cgaaccacca 780
 ccgggaggtg agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
 tgattttatg cgctcgctc ggtactagt agctaactag ctctgtatct ggcggaccgc 900
 ttgtggaact gacgagttcg gaacaccgcg ccgcaacct gggagacgtc ccagggaact 960
 cggggggcgt ttttgtggc cgacctgagt ccaaaatcc cgatcgtttt ggactctttg 1020
 gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080
 tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140
 tctgtgcag catcgttctg tgttgtctct gtctgactgt gtttctgtat ttgtctgaga 1200
 atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260
 agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320
 ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380
 tcacaccca ggtaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg 1440
 tcccctacat cgtgacctgg gaagccttgg cttttgaccc cctccctgg gtcaagccct 1500
 ttgtacaccc taagcctccg cctctcttc ctccatccgc ccgtctctc ccccttgaa 1560
 ctccctgttc gaccocgect cgatccctcc tttatccagc cctcactcct tctctaggcg 1620

205

ccaaacctaa acctcaagtt ctttctgaca gtgggggggc gctcatcgac ctacttacag	1680
aagaccccc gccttatagg gacccaagac caccaccttc cgacagggac ggaaatgggtg	1740
gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg	1800
ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag	1860
gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa	1920
ataatacccc ttctttttct gaagatccag gtaaactgac agctctgatc gagtctgttc	1980
tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg	2040
gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc	2100
gccccactca actgcccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg	2160
attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagttg ctcttagcgg	2220
gtctccaaaa cgcgggcaga agccccacca atttgccaa ggtaaaagga ataacacaag	2280
ggcccaatga gtctccctcg gccttccctag agagacttaa ggaagcctat cgcaggtaca	2340
ctccttatga ccctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc	2400
agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc	2460
ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa	2520
gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg	2580
agcagaaaga gaaagaaga gatcgtagga gacatagaga gatgagcaag ctattggcca	2640
ctgtcggttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg	2700
atcgcgacca gtgtgcctac tgcaaagaaa aggggcaactg ggctaaagat tgtcccaaga	2760
aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgactagg	2820
gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac	2880
ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac	2940
ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgttggg	3000
ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactcttct ctccatgtac	3060
cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc	3120
actttgaggg atcaggagcc caggttatgg gaccaatggg gcagccctg caagtgttga	3180
ccctaaatat agaagatgag catcggtac atgagacctc aaaagagcca gatgtttctc	3240
tagggtccac atggctgtct gattttctct aggcctgggc ggaaaccggg ggcattggac	3300
tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc ccggtgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420

tgttggacca ggggaatactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg	3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc	3540
gggtggaaga catccacccc accgtgccc aaccttaciaa cctcttgagc gggctccccc	3600
cgtcccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc	3660
accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag	3720
gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg	3780
aggcactgca cagagacctc gcagacttcc ggatccagca cccagacttg atcctgctac	3840
agtacgtgga tgacttactg ctggccgcca cttctgagct agactgccaa caaggtactc	3900
gggcccgtgtt acaaacccta gggaacctcg ggtatcgggc ctgggccaag aaagcccaa	3960
tttgccagaa acaggtcaag tatctggggg atctttctaaa agaggggtcag agatggctga	4020
ctgaggccag aaaagagact gtgatggggc agcctactcc gaagacccct cgacaactaa	4080
gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg	4140
cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac	4200
aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag	4260
atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc	4320
taacgcaaaa actgggacct tggcgtcggc cgggtggcta cctgtccaaa aagctagacc	4380
cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgacaa	4440
aggatgcagg caagctaacc atgggacagc cactagtcat tctggccccc catgcagtag	4500
aggcactagt caaacaaccc ccgaccgct ggctttccaa cggccggatg actcactatc	4560
aggccttget tttggacagc gaccgggtcc agttcggacc ggtggtagcc ctgaaccgg	4620
ctacgtgct cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg	4680
aagcccacgg aacccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct	4740
ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga	4800
ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc gctcagcggg	4860
ctgaactgat agcactcacc caggccctaa agatggcaga aggtaagaag ctaaagtgtt	4920
atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc	4980
gtgggttget cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac	5040
taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg	5100
gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccgga aaggcagcca	5160
tcacagagac tccagacacc tctaccctcc tcatagaaaa ttcacacccc tacacctcag	5220
aacattttca ttacacagtg actgatataa aggacctaac caagttgggg gccatttatg	5280

207

ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt	5340
ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc	5400
tcctagagag aagccacagt cctactaca tgctgaaccg ggatcgaaca ctcaaaaata	5460
tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg	5520
gaactagggg ccgcgggcat cggcccgga ctcattggga gatcgatttc accgagataa	5580
agcccggtt gtatggctat aaatatctt tagtttttat agataccttt tctggctgga	5640
tagaagcctt cccaaccaag aaagaaaccg ccaaggctgt aaccaagaag ctactagagg	5700
agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg	5760
tctccaaggt gagtcagaca gtggccgatc tgttggggat tgattggaaa ttacattgtg	5820
catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt	5880
taactaaatt aacgttgca actggctcta gagactgggt gctcctactc cccttagccc	5940
tgtaccgagc ccgcaacacg ccggggcccc atggcctcac cccatatgag atcttatatg	6000
gggcaccccc gcccttgta aacttcctg accctgacat gacaagagtt actaacagcc	6060
cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc	6120
tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagtcg	6180
gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac	6240
cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttgga	6300
tacacgccgc ccacgtgaag gctgcgacc ccgggggtgg accatcctct agactgacat	6360
ggcgcggtca acgtctcaa aacccccca agataagatt aaccctgga agcccttaat	6420
agtcatggga gtccgtttag gagtagggat ggcagagagc ccccatcagg tctttaatgt	6480
aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctcctggg	6540
aactgtacaa gatgccttcc caaaattata tttgatcta tgtgatctgg tcggagagga	6600
gtgggacctc tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag	6660
acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg	6720
tgggggacca ggagagggt actgtggtaa atgggggtgt gaaaccaccg gacaggctta	6780
ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca cccctggga	6840
cacgggatgc tetaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc	6900
cttccaaggg gctactcgag ggggcagatg caaccctcta gtccatagaat tcactgatgc	6960
aggaaaaaag gctaactggg acgggcccac atcgtgggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggct cttaatgtgg gaccccgagt	7080

ccccataggg	cccaaccag	tattacccga	ccaaagactc	ccttcctcac	caatagagat	7140
tgtaccgggt	ccacagccac	ctagccccc	caataccagt	tacccccctt	ccactaccag	7200
tacaccctca	acctccccta	caagtccaag	tgtcccacag	ccacccccag	gaactggaga	7260
tagactacta	gctctagtca	aaggagccta	tcaggcgctt	aacctcacca	atcccgcaca	7320
gaccaagaa	tgttggtgt	gcttagtgtc	gggacctcct	tattacgaag	gagtagcggt	7380
cgtgggcact	tataccaatc	attccaccgc	tccggccaac	tgtacggcca	cttcccaaca	7440
taagcttacc	ctatctgaag	tgacaggaca	gggcctatgc	atgggggcag	tacctaaaac	7500
tcaccaggcc	ttatgtaaca	ccacccaaag	cgccggctca	ggatcctact	accttgacgc	7560
acccgcggga	acaatgtggg	cttgacgac	tggattgact	ccctgcttgt	ccaccacggt	7620
gctcaatcta	accacagatt	attgtgtatt	agttgaactc	tggcccagag	taattttacca	7680
ctccccgat	tatatgtatg	gtcagcttga	acagcgtacc	aaatataaaa	gagagccagt	7740
atcattgacc	ctggcccttc	tactaggagg	attaaccatg	ggagggattg	cagctggaat	7800
agggacgggg	accactgcct	taattaaaac	ccagcagttt	gagcagcttc	atgccgctat	7860
ccagacagac	ctcaacgaag	tcgaaaagtc	aattaccaac	ctagaaaagt	cactgacctc	7920
gttgtctgaa	gtagtcctac	agaaccgcag	aggcctagat	ttgctattcc	taaaggaggg	7980
aggtctctgc	gcagccctaa	aagaagaatg	ttgtttttat	gcagaccaca	cggggctagt	8040
gagagacagc	atggccaaat	taagagaaag	gcttaatcag	agacaaaaac	tatttgagac	8100
aggccaagga	tggttcgaag	ggctgtttta	tagatcccc	tggtttacca	ccttaatctc	8160
caccatcatg	ggacctctaa	tagtactctt	actgatctta	ctctttggac	cttgcatctc	8220
caatcgatta	gtccaatttg	ttaaagacag	gatatcagtg	gtccaggctc	tagttttgac	8280
tcaacaatat	caccagctga	agcctataga	gtacgagcca	tgacgtacgt	tactggccga	8340
agccgcttgg	aataaggccg	gtgtgcgttt	gtctatatgt	tattttccac	catattgccg	8400
tcttttgcca	atgtgagggc	ccggaaacct	ggccctgtct	tcttgacgag	cattcctagg	8460
ggtctttccc	ctctcgccaa	aggaatgcaa	ggtctgttga	atgtcgtgaa	ggaagcagtt	8520
cctctggaag	cttcttgaag	acaaacaacg	tctgtagcga	ccctttgcag	gcagcggaac	8580
ccccacctg	gcgacagggt	cctctgcggc	caaaagccac	gtgtataaga	tacacctgca	8640
aaggcgccac	aacccagtg	ccacgttgtg	agttggatag	ttgtggaaag	agtcaaatgg	8700
ctctcctcaa	gcgtattcaa	caaggggctg	aaggatgccc	agaagggtacc	ccattgtatg	8760
ggatctgac	tggggcctcg	gtgcacatgc	ttlacatgtg	tttagtcgag	gttaaaaaaa	8820
cgtctaggcc	ccccgaacca	cggggacgtg	gttttccttt	gaaaaacacg	ataataccat	8880
ggtgacaggg	ggaatggcaa	gcaagtggga	tcagaagggt	atggacattg	cctatgagga	8940

ggcggcctta ggttacaaag aggggtgggtgt tcttattggc ggatgtctta tcaataacaa	9000
agacggaagt gttctcggtc gtggtcacaa catgagattt caaaagggat ccgccacact	9060
acatggtgag atctccactt tgaaaaactg tgggagatta gagggcaaag tgtacaaaga	9120
taccactttg tatacgacgc tgtctccatg cgacatgtgt acaggtgcc a tcatcatgta	9180
tggtattcca cgctgtgttg tcggtgagaa cgttaatttc aaaagtaagg gcgagaaata	9240
tttacaaact agaggtcacg aggttggtgt tgttgacgat gagaggtgta aaaagatcat	9300
gaaacaattt atcgatgaaa gacctcagga ttggtttgaa gatattgggtg agtaggcggc	9360
cgcgccatag ataaaaataa agattttatt tagtctccag aaaaaggggg gaatgaaaga	9420
ccccacctgt aggtttggca agctagctta agtaacgcc ttttgcaagg catggaaaaa	9480
tacataactg agaatagaga agttcagatc aaggtcagga acagatggaa cagctgaata	9540
tgggccaaac aggatatctg tggtaaacag ttctgcccc ggctcagggc caagaacaga	9600
tggaacagct gaatatgggc caaacaggat atctgtggta agcagttcct gccccggctc	9660
agggccaaga acagatggtc ccagatgag gtccagccct cagcagtttc tagagaacca	9720
tcagatgttt ccagggtgcc ccaaggacct gaaatgacct tgtgccttgt ttaactaac	9780
caatcagttc gcttctcgtc tctgttcgag cgcttctgct ccccgagctc aataaaagag	9840
cccacaacct ctcaactcggg gcgccagtc tccgattgac tgagtcgccc gggtagccgt	9900
gtatccaata aacctctttg cagttgcac cgacttgtag tctcgtgtt ccttgggagg	9960
gtctcctctg agtgattgac taccgctcag cgggggtctt tcatttgggg gctcgtccgg	10020
gacggggaga cccctgcccc gggaccaccg acccaccacc gggaggtaag ctggtgcct	10080
cgcgcgtttc ggtgatgac gtgaaaacct ctgacatgtg agcaaaaggc cagcaaaagg	10140
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggtccgc cccctgacg	10200
agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat	10260
accaggcgtt tccccctgga agtccctcg tgcgtctcc tgttcgacc ctgccgtta	10320
cggatacct gtccgccttt ctccctcgg gaagcgtggc gctttctcaa tgcacagct	10380
gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc	10440
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtc aacccggtaa	10500
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg	10560
taggcgggtgc tacagagttc ttgaagtggg ggcctaacta cggtacact agaaggacag	10620
tatttggtat ctgcgtctg ctgaagccag ttacctcgg aaaaagagtt ggtagctctt	10680
gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta	10740

210

```

cgcgagagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 10800
agtgggaacga aaactcacgt taagggattt tgggtcatgag attatcaaaa aggatcttca 10860
cctagatcct tttaaattaa aaatgaagtt ttaaataaat ctaaagtata tatgagtata 10920
cttgggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 10980
ttcggttcac catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct 11040
taccatctgg cccagtgct gcaatgatac cgcgagacc acgctcacc gctccagatt 11100
tatcagcaat aaaccagcca gccggaagg cggagcgag aagtggctct gcaactttat 11160
ccgcctccat ccagtcatt aattgttgcc gggaaagctag agtaagtagt tcgccagtta 11220
atagtttgcg caocgttggt gccattgctg caggcatcgt ggtgtcacgc tcgtcgtttg 11280
gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt 11340
tgtgcaaaaa agcgggttagc tccttcgggc ctccgatcgt tgcagaagt aagttggcgg 11400
cagtgttate actcatggtt atggcagcac tgcataatc tcttactgtc atgccatccg 11460
taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc 11520
ggcgaccgag ttgctcttgc ccggcgtaa cacgggataa taccgcgcca catagcagaa 11580
ctttaaaagt gtcctcatt ggaaaacgtt ctccggggcg aaaactctca aggatcttac 11640
cgctgttgag atccagttcg atgtaacca ctctgcacc caactgatct tcagcatctt 11700
ttactttcac cagcgtttct ggggtgagca aaacaggaag gcaaaatgcc gcaaaaaagg 11760
gaataaggcg gacacggaaa tgttgaatac tcatactctt cttttttcaa tattattgaa 11820
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 11880
aacaatatag ggttcgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca 11940
ttattatcat gacattaacc tataaaaaata ggcgtatcac gaggcccttt cgtcttcaag 12000
aattcat 12007

```

```

<210> 22
<211> 11893
<212> ADN
<213> Secuencia Artificial

<220>
<223> Vector RCR - pAC3 - yCD2

```

```

<400> 22
tagttattaa tagtaatcaa ttacgggggc attagttcat agcccatata tggagtccg 60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc ccgcccatt 120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180
atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc 240

```

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	360
catggtgatg cgggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
ggactttcca aaatgtcgta acaactccgc ccatttgacg caaatgggcg gtaggcgtgt	540
acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg	600
actgagtcgc ccgggtaccc gtgtatccaa taaacctctc tgcagttgca tccgacttgt	660
ggctctcgctg ttccctggga gggctctctc tgagtgattg actaccgctc agcgggggctc	720
tttcatttgg gggctcgctc gggatcggga gacctctgcc cagggaccac cgaccaccca	780
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
tgattttatg cgctcgctc ggtactagtt agctaactag ctctgtatct ggcggaccgc	900
tgggtggaact gacgagttcg gaacacccgg ccgcaacctt gggagacgct ccagggaactt	960
cggggggcgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg	1020
gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt	1080
tccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg	1140
tctgctgcag catcgttctg tgttctctct gtctgactgt gtttctgtat ttgtctgaaa	1200
atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg	1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acggtgggtt acctctgct	1320
ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc	1380
tcatcacca ggtaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg	1440
tccctacat cgtgacctgg gaagccttgg cttttgaccc cctccctgg gtcaagccct	1500
ttgtacaccc taagcctcgg cctcctcttc ctccatccgc cccgtctctc ccccttgaa	1560
ctcctcgctt gaccccgctt cgatccctcc tttatccagc cctcactcct tctctaggcg	1620
ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag	1680
aagaccccc gccttatagg gacccaagac caccctctc cgacagggaac ggaaatggtg	1740
gagaagcgac ccctgcggga gaggaaccgg accctccccc aatggcatct cgcctacgtg	1800
ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag	1860
gaggaaacgg acagcttcaa tactggcgtt tctcctcttc tgacctttac aactggaaaa	1920
ataataaccc ttctttttct gaagatccag gtaaaactgac agctctgate gagtctgtcc	1980
tcatcacca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg	2040

gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc	2100
gccccactca actgccccat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg	2160
attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagttg ctccatagcg	2220
gtctccaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag	2280
ggcccaatga gtctccctcg gccttcttag agagacttaa ggaagcctat cgcaggtaca	2340
ctccttatga ccctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc	2400
agtctgcccc agacattggg agaaagttag agagggttag agatttaaaa aacaagacgc	2460
ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa	2520
gagagggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg	2580
agcagaaaaga gaaagaaaaga gatcgtagga gacatagaga gatgagcaag ctatttgcca	2640
ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg	2700
atcgcgacca gtgtgcctac tgcaaagaaa aggggcaactg ggctaaagat tgtcccaaga	2760
aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgactagg	2820
gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac	2880
ccgtcacctt cctggttagat actggggccc aacctccgt gctgacccaa aatcctggac	2940
ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgttga	3000
ccacggatcg caaagtacat ctagtaccg gtaaggtcac ccactcttcc ctccatgtac	3060
cagactgtcc ctatcctctg ttaggaagag atttgtgac taaactaaaa gcccaaatcc	3120
actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga	3180
ccctaaatat agaagatgag tatcggtac atgagacctc aaaagagcca gatgtttctc	3240
taggttccac atggctgtct gattttctct aggcctgggc ggaaaccggg ggcattggac	3300
tggcagttcg ccaagctcct ctgatacac ctctgaaagc aacctctacc ccgtgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420
tgttggaaca ggaataactg gtacctgcc agtccccctg gaacacgccc ctgctaccgg	3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc	3540
gggtggaaga catccacccc accgtgccc acccttaca cctcttgagc gggctcccac	3600
cgtcccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc	3660
accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag	3720
gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg	3780
aggcactgca cagagacctc gcagacttcc ggatccagca ccagacttg atcctgctac	3840
agtacgtgga tgacttactg ctggccgcca cttctgagct agactgcca caaggctactc	3900

gggccctgtt	acaaacccta	gggaacctcg	ggtatcgggc	ctcggccaag	aaagcccaaa	3960
tttgccagaa	acaggtcaag	tatctggggt	atcttctaaa	agagggtcag	agatggctga	4020
ctgaggccag	aaaagagact	gtgatggggc	agcctactcc	gaagaccctt	cgacaactaa	4080
gggagttcct	agggacggca	ggcttctgtc	gcctctggat	ccctgggttt	gcagaaatgg	4140
cagccccctt	gtaccctctc	accaaaacgg	ggactctgtt	taattggggc	ccagaccaac	4200
aaaaggccta	tcaagaaatc	aagcaagctc	ttctaactgc	cccagccctg	gggttgccag	4260
atttgactaa	gccctttgaa	ctctttgtcg	acgagaagca	gggctacgcc	aaaggtgtcc	4320
taacgcaaaa	actgggacct	tggcgtcggc	cgggtggccta	cctgtccaaa	aagctagacc	4380
cagtagcagc	tgggtggccc	ccttgccctac	ggatggtagc	agccattgcc	gtactgacaa	4440
aggatgcagg	caagetaaac	atgggacagc	cactagtcac	tctggccccc	catgcagtag	4500
aggcactagt	caaacaaccc	cccgaccgct	ggctttccaa	cgcccggatg	actcactatc	4560
aggccttgct	tttggaacag	gaccgggtcc	agttcggacc	ggtggtagcc	ctgaaccggg	4620
ctacgctgct	cccactgcct	gaggaagggc	tgcaacacaa	ctgccttgat	atcctggccg	4680
aagcccacgg	aacccgaccc	gacctaacgg	accagccgct	cccagacgcc	gaccacacct	4740
ggtaacacga	tggaagcagt	ctcttacaag	agggacagcg	taaggcggga	gctgcgggtg	4800
ccaccgagac	cgaggtaatc	tgggctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
ctgaactgat	agcactcacc	caggccctaa	agatggcaga	aggtaagaag	ctaaatgttt	4920
atactgatag	ccgttatgct	tttgcactcg	cccatatcca	tggagaaata	tacagaaggc	4980
gtgggttgct	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100
gacacagcgc	cgaggctaga	ggcaaccgga	tggctgacca	agcggcccca	aaggcagcca	5160
tcacagagac	tccagacacc	tctaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
aacattttca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccattttatg	5280
ataaaacaaa	gaagtattgg	gtctaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
ttgaattatt	agactttctt	catcagctga	ctcacctcag	cttctcaaaa	atgaaggctc	5400
tccatagagag	aagccacagt	ccctactaca	tgctgaaccg	ggatcgaaca	ctcaaaaata	5460
tcactgagac	ctgcaaagct	tgtgcacaag	tcaacgccag	caagtctgcc	gttaaacagg	5520
gaactagggg	ccgcgggcat	cggcccgcca	ctcattggga	gacgatttc	accgagataa	5580
agcccggtat	gtatggctat	aaatatcttc	tagtttttat	agataccttt	tctggctgga	5640
tagaagcctt	cccaaccaag	aaagaaaccg	ccaaggtcgt	aaccaagaag	ctactagagg	5700

agatcttccc	cagggttcggc	atgcctcagg	tattgggaac	tgacaatggg	cctgccttcg	5760
tctccaaggt	gagtcagaca	gtggccgac	tggtggggat	tgattggaaa	ttacattgtg	5820
catacagacc	ccaaagctca	ggccaggtag	aaagaatgaa	tagaaccatc	aaggagactt	5880
taactaaatt	aacgcttgca	actggctcta	gagactgggt	gctcctactc	cccttagccc	5940
tgtaccgagc	ccgcaacacg	ccgggcccc	atggcctcac	cccatatgag	atcttatatg	6000
gggcaccccc	gccccttgta	aacttccctg	accctgacat	gacaagagtt	actaacagcc	6060
cctctctcca	agctcactta	caggctctct	acttagtcca	gcacgaagtc	tggagacctc	6120
tggcggcagc	ctaccaagaa	caactggacc	gaccgggtgt	acctcaccct	taccgagtcg	6180
gcgacacagt	gtgggtccgc	cgacaccaga	ctaagaacct	agaacctcgc	tggaaaggac	6240
cttacacagt	cctgctgacc	acccccacg	ccctcaaagt	agacggcatc	gcagcttgga	6300
tacacgccgc	ccacgtgaag	gctgccgacc	ccgggggtgg	accatcctct	agactgacat	6360
ggcgcggtca	acgctctcaa	aacccccctc	agataagatt	aacccgtgga	agcccttaat	6420
agtcattgga	gtcctgttag	gagtagggat	ggcagagagc	ccccatcagg	tctttaatgt	6480
aacctggaga	gtcaccaacc	tgatgactgg	gcgtaccgcc	aatgccacct	ccctcctggg	6540
aactgtacaa	gatgccttcc	caaaattata	ttttgatcta	tgtgatctgg	tgggagagga	6600
gtgggacctt	tcagaccagg	aaccgtatgt	cggttatggc	tgcaagtacc	ccgcagggag	6660
acagcggacc	cggacttttg	acttttacgt	gtgccctggg	cataccgtaa	agtcgggggtg	6720
tgggggacca	ggagagggct	actgtggtaa	atgggggtgt	gaaaccaccg	gacaggctta	6780
ctggaagccc	acatcatcgt	gggacctaat	ctcccttaag	cgcggttaaca	ccccctggga	6840
cacgggatgc	tctaaagttg	cctgtggccc	ctgctacgac	ctctccaaag	tatccaattc	6900
cttccaaggg	gctactcgag	ggggcagatg	caaccctcta	gtcctagaat	tcactgatgc	6960
aggaaaaaag	gctaactggg	acgggcccac	atcgtgggga	ctgagactgt	accggacagg	7020
aacagatcct	attaccatgt	tctccctgac	ccggcaggtc	cttaatgtgg	gaccccgagt	7080
ccccataggg	cccaaccacg	tattaccoga	ccaaagactc	ccttccctcac	caatagagat	7140
tgtaccggct	ccacagccac	ctagccccct	caataccagt	tacccccctt	ccactaccag	7200
tacaccctca	acctcccta	caagtccaag	tgtccacacg	ccacccccag	gaactggaga	7260
tagactacta	gctctagtca	aaggagccta	tcaggcgctt	aacctcacca	atcccgacaa	7320
gaccaagaa	tggtggctgt	gcttagtgtc	gggacctcct	tattacgaag	gagtagcggt	7380
cgtgggcact	tataccaatc	attccaccgc	tccggccaac	tgtacggcca	cttccaaca	7440
taagcttacc	ctatctgaag	tgacaggaca	gggcctatgc	atgggggcag	tacctaaaac	7500
tcaccaggcc	ttatgtaaca	ccacccaaag	cgcgggtca	ggatcctact	accttgacgc	7560

acccgcggga acaatgtggg cttgcagcac tggattgact ccttgettgt ccaccacggt	7620
gctcaatcta accacagatt atttgtgtatt agttgaactc tggcccagag taatttacca	7680
ctccccgat tatatgtatg gtcagcttga acagcgtacc aatataaaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattg cagctggaat	7800
agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat	7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc	7920
gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg	7980
aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt	8040
gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
aggccaagga tggttcgaag ggctgtttta tagatcccc tggtttacca ccttaatctc	8160
caccatcatg ggacctctaa tagtactctt actgatetta ctctttggac cttgcattct	8220
caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac	8280
tcagcaatat caccagctaa aaccataga gtacgagcca tgaacgcgtt actggccgaa	8340
gccgcttgga ataaggccgg tgtgcgtttg tctatatgtt atttccacc atattgccgt	8400
cttttgcaa tgtgagggcc cggaaacctg gcctgtctt cttgacgagc attcctaggg	8460
gtctttcccc tctgcctaaa ggaatgcaag gtctgttgaa tgcgtgaag gaagcagttc	8520
ctctggaagc ttcttgaaga caaacaacgt ctgtagcgac cctttgcagg cagcggaacc	8580
ccccacctgg cgacaggtgc ctctgcggcc aaaagccacg tgtataagat acacctgaa	8640
agggcgacaca acccagtgcc caggttgtga gttggatagt tgtggaaaga gtcaaatggc	8700
tctcctcaag cgtattcaac aaggggctga aggatgcca gaaggtaacc cattgtatgg	8760
gatctgatct ggggcctcgg tgcacatgct ttacatgtgt ttagtcgagg ttaaaaaaac	8820
gtctagggcc ccgaaccac ggggacgtgg ttttcctttg aaaaacacga ttataaatgg	8880
tgaccggcgg catggcctcc aagtgggac aaaagggcat ggatacgcct tacgaggagg	8940
ccttgetggg ctacaaggag gggggcgtgc ctatcggcgg ctgtctgatc aacaacaagg	9000
acggcagtgt gctgggcagg ggcacaca ttaggttcca gaagggtcc gccacctgc	9060
acggcgagat ctccacctg gagaactgtg gcaggctgga gggcaagggtg tacaaggaca	9120
ccacctgta caccacctg tccccttgtg acatgtgtac cggcgctatc atcatgtacg	9180
gcatccctag gtgtgtgatc ggcgagaacg tgaacttcaa gtccaagggc gagaagtacc	9240
tgcaaaccag gggccacgag gtggtggttg ttgacgatga gaggtgtaag aagctgatga	9300
agcagttcat cgacgagagg cctcaggact ggttcgagga tatcggcgag taagcggccg	9360

cagataaaat aaaagatttt atttagtctc cagaaaaagg ggggaatgaa agaccccacc	9420
tgtagggttg gcaagctagc ttaagtaacg ccattttgca aggcattgaa aaatacataa	9480
ctgagaatag agaagttcag atcaaggcca ggaacagatg gaacagctga atatgggcca	9540
aacaggatat ctgtggttaag cagttcctgc cccggctcag ggccaagaac agatggaaca	9600
gctgaatatg ggccaaacag gatatctgtg gtaagcagtt cctgccccgg ctgaggcca	9660
agaacagatg gtccccagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg	9720
ttccagggtt gcccgaaggc cctgaaatga cctgtgcct tatttgaact aaccaatcag	9780
ttcgcttctc gcttctgttc gcgcgcttct gctccccgag ctcaataaaa gagcccacaa	9840
cccctcactc ggggagccag tctccgatt gactgagtcg cccgggtacc cgtgtatcca	9900
ataaaccttc ttgcagttgc atccgacttg tgggtctcgt gtctcttggg agggctctct	9960
ctgagtgatt gactaccctg cagcgggggt ctttcattac atgtgagcaa aaggccagca	10020
aaaggccagg aaccgtaaaa aggcgcgctt gctggcgctt ttccataggc tccgcccccc	10080
tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaaccgga caggactata	10140
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc	10200
gcttacggga taectgtccg cttttctccc ttccgggaagc gtggcgcttt ctcatagctc	10260
acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga	10320
acccccgctt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc	10380
ggtaagacac gaattatcgc cactggcagc agccactggt aacaggatta gcagagcgag	10440
gtatgtaggc ggtgctacag agttcttgaa gtgggtggct aactacggct acactagaag	10500
gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttccgaaaaa gagttgtag	10560
ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgt tttttgttt gcaagcagca	10620
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga	10680
cgctcagtg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat	10740
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga	10800
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg	10860
tctatttcgt tcatccatag ttgcctgact ccccgctgtg tagataacta cgatacggga	10920
gggcttacca tctggcccca gtgctgcaat gataccgcga gaccacgct caccggctcc	10980
agatttatca gcaataaacc agccagccgg aaggggcgag cgcagaagtg gtctgcaac	11040
tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtatctgcc	11100
agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtggtgt cagctcgtc	11160
gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc	11220

catgttggtgc aaaaaagcgg ttagctcctt cggctcctccg atcgttggtca gaagtaagtt	11280
ggccgcagtg ttatcactca tgggtatggc agcactgcat aattctctta ctgtcatgcc	11340
atccgtaaga tgcttttctg tgactgggtga gtactcaecc aagtcattct gagaatagt	11400
tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg cgccacatag	11460
cagaacttta aaagtgtca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat	11520
cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc	11580
atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaa	11640
aaagggaata agggcgacac ggaaatggtg aatactcata ctcttccttt ttcaatatta	11700
ttgaagcatt tatcagggtt attgtotcat gagcggatac atatttgaat gtatttagaa	11760
aaataaacia ataggggttc cgcgcacatt tccccgaaa gtgccacctg acgtctaaga	11820
aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct	11880
tcaagaattc cat	11893



(43) International Publication Date
22 April 2010 (22.04.2010)

(51) International Patent Classification:
C12N 15/52 (2006.01) C12N 15/867 (2006.01)
C12N 5/10 (2006.01)

Encinitas, CA 92024 (US). PEREZ, Omar [US/US]; 219
Brannan Street #2C, San Francisco, CA 94107 (US).

(21) International Application Number: PCT/US2009/058510

(74) Agent: **BAKER, Joseph, R.**; Gavrilovich, Dodd & Lindsey, LLP, 4660 La Jolla Village Drive, Suite 750, San Diego, CA 92122 (US).

(22) International Filing Date: 26 September 2009 (26.09.2009)

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

61/100,666	26 September 2008 (26.09.2008)	US
61/120,618	8 December 2008 (08.12.2008)	US
61/186,823	13 June 2009 (13.06.2009)	US

(71) **Applicant** (for all designated States except US): **TOCAGEN INC.** [US/US]; 3030 Bunker Hill Street, Suite 230, San Diego, CA 92109 (US).

(72) Inventors; and

(75) **Inventors/Applicants (for US only): GRUBER, Harry, E.** [US/US]; P.O. Box 675272, Rancho Santa Fe, 92067 (US). **JOLLY, Douglas** [US/US]; 382 Sunset Drive.

(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LI, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM,

[Continued on next page]

(54) Title: GENE THERAPY VECTORS AND CYTOSINE DEAMINASES

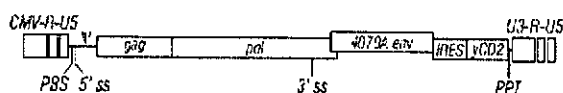


FIGURE 1A

(57) Abstract: This disclosure provides modified cytosine deaminases (CDs). The disclosure further relates to cells and vector expressing or comprising such modified CDs and methods of using such modified CDs in the treatment of disease and disorders.

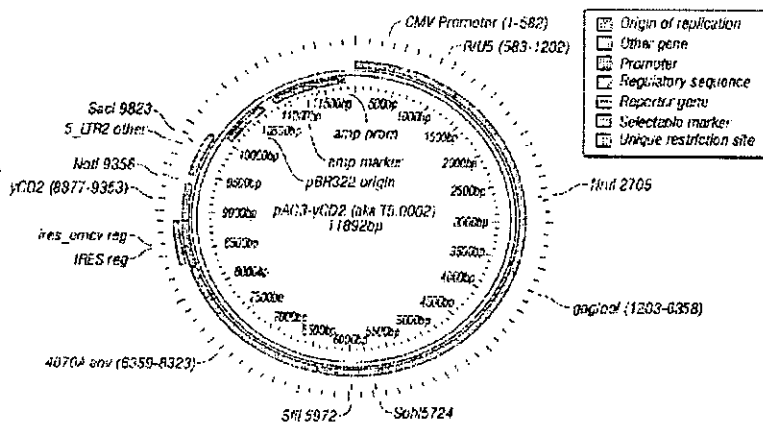


FIGURE 1B



No. 11-050300- -00000-0000

Fecha: 2011-04-25 15:43:45 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 220

PATENTE DE INVENCION ☐

MODELO I

- ☐ Indicación que se solicita una patente.
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Descripción de la invención
- ☐ Dibujos de ser estos pertinentes
- ☐ Comprobante de pago de las tasas establecidas (De ser el caso formato de descuento)
- Completa ☐ Incompleta ☐

PATENTE DE INVENCION PCT ☒

MODELO DE UTILIDAD PCT ☐

Art.33 Decisión 486/00, Circular Única

- ☒ Indicación que se solicita una PCT
- ☒ Copia de la solicitud en español, tal como fue presentada inicialmente (capítulo descriptivo, reivindicatorio, resumen)
- ☐ Dibujos de ser estos pertinentes
- ☒ Comprobante de pago de las tasas establecidas (de ser el caso formato de descuento)
- Completa ☐ Incompleta ☐

DISEÑO INDUSTRIAL ☐

(Art. 119 Decisión 486/00)

- ☐ Indicación que se solicita Diseño industrial
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Representación gráfica y fotográfica del Diseño industrial o muestra del material que incorpora el diseño
- ☐ Comprobante de pago de las tasas establecidas
- Completa ☐ Incompleta ☐

ESQUEMA DE TRAZADO ☐

(Art. 92 Decisión 486/00)

- ☐ Indicación que se solicita un esquema de trazado
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Representación gráfica de un esquema de trazado
- ☐ Comprobante de pago de las tasas establecidas
- Completa ☐ Incompleta ☐



Bogotá D.C.,
2020

Doctor(a)
RODRIGUEZ D' ALEMAN DILIA MARIA
Apoderado(a) y/o Representante de
TOCAGEN INC.

REFERENCIA	Radicación No.	11 050300
	Solicitud internacional	PCT/US2009/58510
	Fecha inicio fase nacional	26/04/2011
	Trámite	11
	Evento	378
	Actuación	430
	Oficio No.	08265

Como resultado del examen de forma, realizado con fundamento en el artículo 38 de la Decisión 486 de la Comisión de la Comunidad Andina, artículo 27 del Tratado de Cooperación en Materia de Patentes (PCT), regla 51 bis del Reglamento y Circular Única de la Superintendencia de Industria y Comercio, se le comunica que:

1. La solicitud no contiene el poder debidamente otorgado, con el lleno de los requisitos exigidos por los artículos 63 y subsiguientes del Código de Procedimiento Civil. Adicionalmente, deberá presentar el documento que acredita la existencia y representación legal de la sociedad solicitante, cuando ésta fuere persona jurídica.

2. La solicitud no contiene la copia del documento en que consta la cesión del derecho a la patente del inventor al solicitante o a su causante, conforme a lo establecido por el artículo 26 literal k) de la Decisión 486 de la Comisión de la Comunidad Andina. Cuando se trate de documentos otorgados en el extranjero, estos deberán legalizarse de conformidad con lo dispuesto por los artículos 259 y 260 del Código de Procedimiento Civil. Cuando se trate de las declaraciones a que hace referencia la Regla 4.17 ii) del Tratado de Cooperación en Materia de Patentes PCT., el alcance de las tales declaraciones se determinará según indiquen claramente una cesión de derechos del inventor al solicitante o su causante, de conformidad con lo dispuesto en el numeral 5.2.15 del Capítulo Quinto, Título X de la Circular Única de Superintendencia de Industria y Comercio.

En cumplimiento de lo dispuesto por el artículo 39 de la Decisión 486, se le requiere para que dentro del término de dos meses siguientes a la fecha de notificación del presente oficio, complete los requisitos anteriormente enunciados. En caso de no dar cumplimiento al presente requerimiento, la solicitud se considerará abandonada y perderá su prelación.

Continúa en la página siguiente...



Notifíquese el presente oficio conforme a lo dispuesto en el numeral 5.2, literal e), del capítulo quinto del título primero de la Circular Unica No.10 de 2001.

JOSE LUIS SALAZAR LOPEZ
Director de Nuevas Creaciones (c)

12 0 MAYO 2011

El anterior requerimiento fue notificado por fijación en lista de fecha _____
adpa11