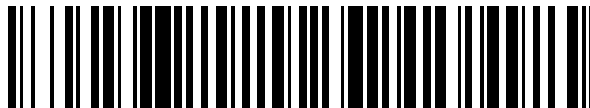


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 422 303**

51 Int. Cl.:

**C12N 15/63** (2006.01)  
**C12N 15/85** (2006.01)  
**C12N 15/12** (2006.01)  
**C12N 15/62** (2006.01)  
**C07K 14/705** (2006.01)  
**C07K 14/72** (2006.01)  
**A01H 5/00** (2006.01)  
**C12Q 1/68** (2006.01)  
**G01N 33/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **20.02.2002 E 02742489 (4)**

97 Fecha y número de publicación de la concesión europea: **24.04.2013 EP 1373470**

54 Título: **Receptores mutantes por sustitución novedosos y su uso en un sistema de expresión génica inducible basado en receptores nucleares**

30 Prioridad:

**20.02.2001 US 269799 P**  
**21.08.2001 US 313925 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**10.09.2013**

73 Titular/es:

**INTREXON CORPORATION (100.0%)**  
**1750 Kraft Drive, Suite 1400**  
**Blacksburg VA 24060, US**

72 Inventor/es:

**PALLI, SUBBA, REDDY;**  
**KUMAR, MOHAN, BASAVARAJU;**  
**CRESS, DEAN, ERVIN y**  
**FUJIMOTO, TED, TSUTOMU**

74 Agente/Representante:

**PÉREZ BARQUÍN, Eliana**

**ES 2 422 303 T3**

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

## DESCRIPCIÓN

Receptores mutantes por sustitución novedosos y su uso en un sistema de expresión génica inducible basado en receptores nucleares

5

**Campo de la invención**

Esta invención se refiere al campo de biotecnología o ingeniería genética. Específicamente, esta invención se refiere al campo de la expresión génica. Más específicamente, esta invención se refiere a receptores nucleares novedosos que comprenden una mutación por sustitución y a su uso en un sistema de expresión génica inducible basado en receptores nucleares y a métodos de modulación de la expresión de un gen dentro de una célula huésped usando este sistema de expresión génica inducible.

10

**Antecedentes de la invención**

15

La mención de cualquier referencia en el presente documento no debe interpretarse como una admisión de que tal referencia está disponible como "técnica anterior" para la presente solicitud.

20

En el campo de ingeniería genética, el control preciso de la expresión génica es una herramienta valiosa para estudiar, manipular y controlar el desarrollo y otros procesos fisiológicos. La expresión génica es un proceso biológico complejo que implica varias interacciones proteína-proteína específicas. Con el fin de desencadenar la expresión génica, de modo que se produzca el ARN necesario como primera etapa en la síntesis de proteínas, debe colocarse un activador de la transcripción en proximidad de un promotor que controla la transcripción génica. Normalmente, el propio activador de la transcripción está asociado con una proteína que tiene al menos un dominio de unión a ADN que se une a sitios de unión a ADN presentes en las regiones promotoras de genes. Por tanto, para que se produzca la expresión génica, una proteína que comprende un dominio de unión a ADN y un dominio de transactivación ubicados a una distancia apropiada del dominio de unión a ADN deben colocarse en la posición correcta en la región promotora del gen.

25

30

El enfoque transgénico tradicional utiliza un promotor específico de tipo celular para dirigir la expresión de un transgén diseñado. En primer lugar se incorpora un constructo de ADN que contiene el transgén en un genoma huésped. Cuando se desencadena por un activador de la transcripción, se produce la expresión del transgén en un tipo celular dado.

35

Otros medios para regular la expresión de genes foráneos en células es a través de promotores inducibles. Los ejemplos del uso de tales promotores inducibles incluyen el promotor PR1-a, sistemas de represor-operador procariotas, sistemas de inmunosupresor-inmunofilina y sistemas de activación de la transcripción de eucariotas superiores tales como sistemas de receptores de hormonas esteroideas y se describen a continuación.

40

El promotor PR1-a del tabaco se induce durante la respuesta de resistencia adquirida sistémica tras un ataque de patógenos. El uso de PR1-a puede limitarse debido a que a menudo responde a materiales endógenos y factores externos tales como patógenos, radiación UV-B y contaminantes. Se han descrito sistemas de regulación génica basados en promotores inducidos por choque térmico, interferón y metales pesados (Wum *et al.*, 1986, Proc. Natl. Acad. Sci. USA 83: 5414-5418; Amheiter *et al.*, 1990, Cell 62:51-61; Filmus *et al.*, 1992, Nucleic Acids Research 20: 27550-27560). Sin embargo, estos sistemas tienen limitaciones debido a su efecto sobre la expresión de genes no diana. Estos sistemas también tienen brechas.

45

50

Los sistemas de represor-operador procariotas utilizan proteínas represoras bacterianas y las secuencias de ADN de operador únicas a las que están unidas. Se han usado tanto los sistemas de represor-operador de tetraciclina ("Tet") como de lactosa ("Lac") de la bacteria *Escherichia coli* en plantas y animales para controlar la expresión génica. En el sistema Tet, la tetraciclina se une a la proteína represora TetR, dando como resultado un cambio conformacional que libera la proteína represora del operador que como resultado permite que se produzca la transcripción. En el sistema Lac, se activa un operón lac en respuesta a la presencia de lactosa, o análogos sintéticos tales como isopropil-b-D-tiogalactósido. Desafortunadamente, el uso de tales sistemas se ve restringido por la química inestable de los ligandos, es decir tetraciclina y lactosa, su toxicidad, su presencia natural o los niveles relativamente altos requeridos para la inducción o represión. Por motivos similares, la utilidad de tales sistemas en animales es limitada.

55

60

Moléculas inmunosupresoras tales como FK506, rapamicina y ciclosporina A pueden unirse a inmunofilinas FKBP12, ciclofilina, etc. Usando esta información, se ha ideado una estrategia general para juntar dos cualesquiera proteínas simplemente colocando FK506 en cada una de las dos proteínas o colocando FK506 en una y ciclosporina A en otra. Entonces puede usarse un homodímero sintético de FK506 (FK1012) o un compuesto resultante de la fusión de FK506-ciclosporina (FKCsA) para inducir la dimerización de estas moléculas (Spencer *et al.*, 1993, Science 262: 1019-24; Belshaw *et al.*, 1996 Proc Natl Acad Sci USA 93: 4604-7). Se usaron el dominio de unión a ADN de Gal4 fusionado al dominio activador FKBP12 y VP16 fusionado a ciclofilina, y el compuesto FKCsA, para mostrar la heterodimerización y activación de un gen indicador bajo el control de un promotor que contenía los sitios de unión de Gal4. Desafortunadamente, este sistema incluye inmunosupresores que pueden tener efectos secundarios no

65

deseados y, por tanto, se limita su uso para diversas aplicaciones de interruptor génico en mamíferos.

También se han empleado sistemas de activación de la transcripción de eucariotas superiores tales como sistemas de receptores de hormonas esteroideas. Los receptores de hormonas esteroideas son miembros de la superfamilia de receptores nucleares y se encuentran en células de vertebrados e invertebrados. Desafortunadamente, el uso de compuestos esteroideos que activan los receptores para la regulación de la expresión génica, particularmente en plantas y mamíferos, se ve limitada debido a su implicación en muchas otras rutas biológicas naturales en tales organismos. Con el fin de superar tales dificultades, se ha desarrollado un sistema alternativo usando receptores de ecdisona de insectos (EcR).

El crecimiento, la muda y el desarrollo en insectos se regulan por la hormona esteroidea ecdisona (hormona de la muda) y las hormonas juveniles (Dhadialla, *et al.*, 1998, Annu. Rev. Entomol. 43: 545-569). La diana molecular para la ecdisona en insectos consiste en al menos un receptor de ecdisona (EcR) y una proteína ultraespiráculo (USP). EcR es un miembro de la superfamilia de receptores esteroideos nucleares que se caracteriza por un ADN firma y dominios de unión a ligando, y un dominio de activación (Koelle *et al.* 1991, Cell, 67:59-77). Los receptores EcR son sensibles a varios compuestos esteroideos tales como ponasterona A y muristerona A. Recientemente, se han descrito compuestos no esteroideos con actividad agonista de ecdiesteroides, incluyendo los insecticidas disponibles comercialmente tebufenozida y metoxifenozida que se comercializan en todo el mundo por Rohm y Haas Company (véanse la solicitud internacional de patente n.º PCT/EP96/00686 y la patente de EE.UU. 5.530.028). Ambos análogos tienen perfiles de seguridad excepcionales frente a otros organismos.

El receptor de ecdisona de insecto (EcR) se heterodimeriza con ultraespiráculo (USP), el homólogo de insectos del RXR de mamíferos, y se une a ecdiesteroides y a elementos de respuesta a receptores de ecdisona y activan la transcripción de genes sensibles a ecdisona (Riddiford *et al.*, 2000). Los complejos EcR/USP/ligando desempeñan papeles importantes durante la reproducción y el desarrollo de insectos. El EcR es un miembro de la superfamilia de receptores de hormonas esteroideas y tiene cinco dominios modulares, dominios A/B (transactivación), C (unión a ADN, heterodimerización), D (bisagra, heterodimerización), E (unión a ligando, heterodimerización y transactivación) y F (transactivación). Algunos de estos dominios tales como A/B, C y E conservan su función cuando se fusionan con otras proteínas.

Sistemas de expresión génica inducible estrechamente regulados o "interruptores génicos" son útiles para diversas aplicaciones tales como terapia génica, producción a gran escala de proteínas en células, ensayos de selección de alto rendimiento basados en células, genómica funcional y regulación de los rasgos en plantas y animales transgénicos.

La primera versión del interruptor génico basado en EcR usaba EcR de *Drosophila melanogaster* (DmEcR) y RXR de *Mus musculus* (MmRXR) y mostró que estos receptores en presencia de esteroide, ponasterona A, transactivan genes indicadores en líneas celulares de mamíferos y ratones transgénicos (Christopherson *et al.*, 1992; No *et al.*, 1996). Posteriormente, Suhr *et al.*, 1998 mostraron que el agonista de ecdisona no esteroideo, tebufenozida, inducía un alto nivel de transactivación de genes indicadores en células de mamífero a través del EcR de *Bombyx mori* (BmEcR) en ausencia de pareja de heterodímero exógena.

Las solicitudes internacionales de patente n.º PCT/US97/05330 (documento WO 97/38117) y n.º PCT/US99/08381 (documento WO 99/58155) dan a conocer métodos para modular la expresión de un gen exógeno en el que un constructo de ADN que comprende el gen exógeno y un elemento de respuesta a ecdisona se activa por un segundo constructo de ADN que comprende un receptor de ecdisona que, en presencia de un ligando para el mismo, y opcionalmente en presencia de un receptor que puede actuar como pareja silenciosa, se une al elemento de respuesta a ecdisona induciendo expresión génica. Se aisló el receptor de ecdisona de elección de *Drosophila melanogaster*. Normalmente, tales sistemas requieren la presencia de la pareja silenciosa, preferiblemente receptor retinoide X (RXR), con el fin de proporcionar activación óptima. En células de mamíferos, el receptor de ecdisona de insectos (EcR) se heterodimeriza con el receptor retinoide X (RXR) y regula la expresión de genes diana de una manera dependiente del ligando. La solicitud internacional de patente n.º PCT/US98/14215 (documento WO 99/02683) da a conocer que el receptor de ecdisona aislado de la mariposa de la seda *Bombyx mori* es funcional en sistemas de mamíferos sin la necesidad de una pareja de dímero exógena.

La patente de EE.UU. n.º 6.265.173 B1 da a conocer que diversos miembros de la superfamilia esteroidea/tiroidea de receptores pueden combinarse con el receptor de ultraespiráculo (USP) de *Drosophila melanogaster* o fragmentos del mismo que comprenden al menos el dominio de dimerización de USP para su uso en un sistema de expresión génica. La patente de EE.UU. n.º 5.880.333 da a conocer un sistema de heterodímero de EcR y ultraespiráculo (USP) de *Drosophila melanogaster* usado en plantas en el que el dominio de transactivación y el dominio de unión a ADN están situados en dos proteínas híbridas diferentes. Desafortunadamente, estos sistemas basados en USP son constitutivos en células animales y, por tanto, no son eficaces para regular la expresión de genes indicadores.

En cada uno de estos casos, se incorporaron el dominio de transactivación y el dominio de unión a ADN (o bien como EcR nativo como en la solicitud internacional de patente n.º PCT/US98/14215 o bien como EcR modificado

como en la solicitud internacional de patente n.º PCT/US97/05330) en una molécula individual y se usaron las otras parejas heterodiméricas, o bien USP o bien RXR, en su estado nativo.

5 Los inconvenientes de los sistemas de regulación génica basados en EcR descritos anteriormente incluyen una considerable actividad de fondo en ausencia de ligandos y la no aplicabilidad de estos sistemas para su uso en tanto plantas como animales (véase la patente de EE.UU. n.º 5.880.333). Por tanto, existe una necesidad en la técnica de sistemas basados en EcR mejorados para modular de manera precisa la expresión de genes exógenos en tanto plantas como animales. Tales sistemas mejorados serían útiles para aplicaciones tales como terapia génica, producción de proteínas y anticuerpos a gran escala, ensayos de selección de alto rendimiento basado en células, genómica funcional y regulación de rasgos en animales transgénicos. Para determinadas aplicaciones tales como terapia génica, puede ser deseable tener un sistema de expresión génica inducible que responda bien a ligandos no esteroideos sintéticos y al mismo tiempo sea insensible a los esteroides naturales. Por tanto, sistemas mejorados que sean sencillos, compactos y dependientes de ligandos que sean relativamente económicos, fácilmente disponibles y de baja toxicidad para el huésped serían útiles para regular sistemas biológicos.

15 Recientemente, los solicitantes han mostrado que un sistema de expresión génica inducible basado en receptor de ecdisona en el que los dominios de unión a ADN y de transactivación están separados entre sí colocándolos en dos proteínas diferentes da como resultado una actividad de fondo enormemente reducida en ausencia de un ligando y una actividad significativamente aumentada con respecto al fondo en presencia de un ligando (solicitud pendiente PCT/US01/09050). Este sistema de dos híbridos es un sistema de modulación de la expresión génica inducible significativamente mejorado en comparación con los dos sistemas dados a conocer en las solicitudes PCT/US97/05330 y PCT/US98/14215. El sistema de dos híbridos se aprovecha de la capacidad de un par de proteínas de interacción para colocar el dominio de activación de la transcripción en una posición más favorable con respecto al dominio de unión a ADN de modo que cuando el dominio de unión a ADN se une al sitio de unión a ADN en el gen, el dominio de transactivación activa más eficazmente al promotor (véase, por ejemplo, la patente de EE.UU. n.º 5.283.173). En resumen, el sistema de expresión génica de dos híbridos comprende dos casetes de expresión génica; el primero que codifica para un dominio de unión a ADN fusionado a un polipéptido de receptor nuclear, y el segundo que codifica para un dominio de transactivación fusionado a un polipéptido de receptor nuclear diferente. En presencia de ligando, la interacción del primer polipéptido con el segundo polipéptido une eficazmente el dominio de unión a ADN al dominio de transactivación. Puesto que los dominios de transactivación y de unión a ADN residen en dos moléculas diferentes, la actividad de fondo en ausencia del ligando se reduce enormemente.

35 Un sistema de dos híbridos también proporciona sensibilidad mejorada a ligandos no esteroideos, por ejemplo, diacilhidrazinas, en comparación con ligandos esteroideos, por ejemplo, ponasterona A ("PonA") o muristerona A ("MurA"). Es decir, en comparación con esteroides, los ligandos no esteroideos proporcionan actividad superior a una concentración inferior. Además, puesto que la transactivación basada en interruptores génicos de EcR es a menudo dependiente de línea celular, es más fácil adaptar sistemas de interrupción para obtener una capacidad de transactivación máxima para cada aplicación. Además, el sistema de dos híbridos evita algunos efectos secundarios debido a la sobreexpresión de RXR que a menudo se producen cuando se usa RXR no modificado como pareja de interrupción. En un sistema de dos híbridos preferido, se eliminan los dominios de transactivación y de unión a ADN de EcR o RXR y, como resultado, estas moléculas híbridas tienen menos posibilidades de interactuar con otros receptores de hormonas esteroideas presentes en la célula, lo que da como resultado efectos secundarios reducidos.

45 El EcR es un miembro de la superfamilia de receptor nuclear y se clasifica en la subfamilia 1, grupo H (denominado en el presente documento "receptores nucleares de grupo H"). Los miembros de cada grupo comparten una identidad de aminoácidos del 40-60% en el dominio E (unión a ligando) (Laudet *et al.*, A Unified Nomenclature System for the Nuclear Receptor Subfamily, 1999; Cell 97: 161-163). Además del receptor de ecdisona, otros miembros de esta subfamilia 1 de receptores nucleares, grupo H, incluyen: receptor ubicuo (UR), receptor huérfano 1 (OR-1), receptor nuclear de hormonas esteroideas 1 (NER-1), proteína 15 de interacción de RXR (RIP-15), receptor hepático X  $\beta$  (LXR $\beta$ ), proteína similar al receptor de hormonas esteroideas (RLD-1), receptor hepático X (LXR), receptor hepático X  $\alpha$  (LXR $\alpha$ ), receptor farnesoide X (FXR), proteína 14 de interacción de receptor (RIP-14) y receptor de farnesol (HRR-1).

55 El documento WO 96/27673 da a conocer un método de control de la expresión génica en plantas. El método comprende obtener una planta transgénica que comprende al menos dos casetes de expresión de receptor y al menos un casete de expresión de diana.

60 Para desarrollar un sistema de expresión génica inducible basado en receptores nucleares de grupo H mejorado en el que se modifica la unión a ligando o especificidad de ligando, los solicitantes crearon varios EcR mutantes por sustitución que comprenden residuos de aminoácido sustituidos en el dominio de unión a ligando (LBD). Se usó un enfoque de acoplamiento y modelado de homología para predecir residuos críticos que median en la unión de ecdiesteroides y moléculas distintas de ecdiesteroides al LBD de EcR. Se evaluaron estos EcR mutantes por sustitución en ensayos de unión a ligando y transactivación. Tal como se presenta en el presente documento, los receptores nucleares mutantes por sustitución novedosos de los solicitantes y su uso en un sistema de expresión génica inducible basado en receptores nucleares proporcionan un sistema de expresión génica inducible mejorado

tanto en células huésped procariotas como eucariotas en el que la sensibilidad de ligando y la magnitud de la transactivación pueden seleccionarse según se desee, dependiendo de la aplicación.

**Breve descripción de los dibujos**

- 5
- Figura 1: Unión del ligando <sup>3</sup>H-RH2485 *in vitro* del mutante de CfEcR A110P de longitud completa mientras que la unión de esteroides se altera completamente. Los valores de unión de ligando se expresan como cuentas específicas (dpm específicas).
- 10
- Figura 2: Transactivación de genes indicadores a través de GAL4/CfEcR-ABCDEF (CfEcR de longitud completa) o sus constructos de versión mutante GAL4/A110P transfectados en células NIH3T3 junto con VP16LmUSP-EF y pFREcRE por PonA o GS<sup>TM</sup>-E. Los números en la parte superior de las barras indican el aumento en veces con respecto a los niveles de DMSO.
- 15
- Figura 3: Transactivación de genes indicadores a través de IE1VP16/CfEcRCDEF (ejemplo 1.5) o sus constructos de versión mutante VP16/A110P (ejemplo 1.6) transfectados en células L57 junto con indicador de pMK43.2 por 20E o GS<sup>TM</sup>-E. Los números en la parte superior de las barras indican el aumento en veces con respecto a los niveles de DMSO.
- 20
- Figura 4: Transactivación de genes indicadores a través de GAL4/CfEcR-A/BCDEF (CfEcR de longitud completa) o sus constructos de versión mutante GAL4/A110 (A110S, A110P, A110L, y A110M) transfectados en células NIH3T3 junto con VP16LmUSP-EF y pFREcRE por PonA o GS<sup>TM</sup>-E. Los números en la parte superior de las barras indican aumento en veces con respecto a los niveles de DMSO.
- 25
- Figura 5: Unión del ligando <sup>3</sup>H-PonA *in vitro* de CfEcR-A/BCDEF de tipo natural (CFEcR de longitud completa) o sus versiones mutantes A110 (A110S, A110P, A110L y A110M). Los valores de unión de ligando se expresan como cuentas específicas (dpm específicas).
- 30
- Figura 6: Unión del ligando <sup>3</sup>H-RHTA85 *in vitro* de CfEcR-A/BCDEF de tipo natural (CfEcR de longitud completa) o sus versiones mutantes A110 (A110S, A110P, A110L y A110M). Se expresaron estos valores como cuentas específicas (dpm específicas).

**Descripción detallada de la invención**

35

Los solicitantes describen en el presente documento la construcción de receptores nucleares de grupo H que comprenden mutaciones por sustitución (denominadas en el presente documento “mutantes por sustitución”) en residuos de aminoácido que están implicados en la unión de ligando a un dominio de unión a ligando de receptor nuclear de grupo H que afecta a la sensibilidad al ligando y a la magnitud de inducción del receptor nuclear de grupo H y la demostración de que estos receptores nucleares mutantes por sustitución son útiles en métodos de modulación de la expresión génica.

40

Específicamente, los solicitantes han desarrollado un sistema de expresión génica inducible basado en receptores nucleares novedoso que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución. Los solicitantes han mostrado que el efecto de una mutación por sustitución de este tipo puede aumentar o reducir la actividad de unión a ligando o sensibilidad al ligando y el ligando puede ser un compuesto esteroideo o no esteroideo específico. Por tanto, la invención de los solicitantes proporciona un sistema de expresión génica inducible basado en receptores nucleares de grupo H útil para modular la expresión de un gen de interés en una célula huésped. En una realización particularmente deseable, la descripción proporciona un sistema de expresión génica inducible basado en receptor de ecdisona que responde únicamente a o bien ligando esteroideo o bien ligando no esteroideo. Además, la presente invención también proporciona un sistema de expresión génica inducible basado en receptor de ecdisona sensible a ligando no esteroideo mejorado. Por tanto, el sistema de expresión génica inducible novedoso de los solicitantes y su uso en métodos de modulación de la expresión génica en una célula huésped superan las limitaciones de sistemas de expresión inducibles disponibles actualmente y proporcionan al experto medios eficaces para controlar la expresión génica.

45

50

55

La presente invención es útil para aplicaciones tales como terapia génica, producción de proteínas y anticuerpos a gran escala, ensayos de selección de alto rendimiento basados en células, ensayos de selección de ligandos ortogonales, genómica, proteómica, metabolómica funcional y regulación de rasgos en organismos transgénicos, en los que es deseable niveles de control de la expresión génica. Una ventaja de la invención de los solicitantes es que proporciona un medio para regular la expresión génica y adaptar los niveles de expresión para adecuarse a los requisitos del usuario.

60

La presente invención proporciona un sistema de modulación de la expresión génica que comprende un casete de expresión génica que puede expresarse en una célula huésped que comprende un polinucleótido que codifica para un polipéptido que comprende:

65

i) un dominio de transactivación;

ii) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse; y

5 iii) un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución; en el que el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor nuclear de grupo H seleccionado del grupo que consiste en un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un  
 10 NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína 15 de interacción de receptor retinoide X, receptores hepáticos X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor y un receptor de farnesol,

15 en el que el dominio de unión a ligando de receptor nuclear de grupo H está codificado por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido, estando el residuo de aminoácido en una posición equivalente o análoga a

a) el residuo de aminoácido 107 de SEQ ID NO: 1,

20 b) los residuos de aminoácido 107 y 175 de SEQ ID NO: 1,

c) los residuos de aminoácido 107 y 127 de SEQ ID NO: 1,

25 d) los residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o

e) los residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, y

30 en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad del receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

En otro aspecto, la presente invención proporciona un sistema de modulación de la expresión génica que comprende:

35 a) un primer casete de expresión génica que puede expresarse en una célula huésped que comprende un polinucleótido que codifica para un primer polipéptido que comprende:

40 i) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse; y

ii) un dominio de unión a ligando de receptor nuclear; y

45 b) un segundo casete de expresión génica que puede expresarse en la célula huésped que comprende un polinucleótido que codifica para un segundo polipéptido que comprende:

i) un dominio de transactivación; y

ii) un dominio de unión a ligando de receptor nuclear

50 en el que uno de los dominios de unión a ligando de receptor nuclear es un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución;

55 en el que el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor nuclear de grupo H seleccionado del grupo que consiste en un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína 15 de interacción de receptor retinoide X, un receptor hepático X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor y un receptor de farnesol, en el que el dominio de unión a ligando de receptor nuclear de grupo H está codificado por un polinucleótido que  
 60 comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido, estando el residuo de aminoácido en una posición equivalente o análoga a

a) el residuo de aminoácido 107 de SEQ ID NO: 1,

65 b) los residuos de aminoácido 107 y 175 de SEQ ID NO: 1,

c) los residuos de aminoácido 107 y 127 de SEQ ID NO: 1,

d) los residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o

e) los residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, y

5 en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

10 En un aspecto adicional, la presente invención proporciona un casete de expresión génica que comprende un polinucleótido que codifica para un polipéptido que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución; en el que el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor nuclear de grupo H seleccionado del grupo que consiste en un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína  
15 15 de interacción de receptor retinoide X, un receptor hepático X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, receptores hepáticos X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor y un receptor de farnesol, en el que el dominio de unión a ligando de receptor nuclear de grupo H está codificado por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido, estando el residuo de aminoácido en una posición equivalente o análoga a

20 a) el residuo de aminoácido 107 de SEQ ID NO: 1,

b) los residuos de aminoácido 107 y 175 de SEQ ID NO: 1,

25 c) los residuos de aminoácido 107 y 127 de SEQ ID NO: 1,

d) los residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o

e) los residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, y

30 en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

35 En otro aspecto, la presente invención proporciona un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, comprendiendo el polinucleótido aislado una mutación de codón que da como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga a

40 a) el residuo de aminoácido 107 de SEQ ID NO: 1,

b) los residuos de aminoácido 107 y 175 de SEQ ID NO: 1,

45 c) los residuos de aminoácido 107 y 127 de SEQ ID NO: 1,

d) los residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1,

e) los residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, o

50 f) los aminoácidos 107, 110 y 175 de SEQ ID NO: 1, y

en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

55 En un aspecto adicional, la presente invención proporciona un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución, en el que el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroideos y el polinucleótido que codifica para el dominio de unión a ligando de receptor de ecdisona se hibrida con un polinucleótido que comprende una mutación de codón que da como resultado la mutación por sustitución, mutación por sustitución V107I/A110P/R175E de SEQ ID NO: 1, en condiciones de hibridación que comprenden una etapa de hibridación en menos de 500 mM de sal y al menos 37 grados Celsius, y una etapa de lavado en 2XSSPE a al menos 63 grados Celsius.

65 La presente invención proporciona, en otro aspecto, un vector de expresión que comprende el polinucleótido aislado de la presente invención operativamente unido a un elemento regulador de la transcripción.

La presente invención proporciona, en un aspecto adicional, una célula huésped aislada que comprende el vector de expresión de la presente invención, en la que el elemento regulador de la transcripción es operativo en la célula huésped.

- 5 En otro aspecto, la presente invención proporciona un polipéptido aislado codificado por el polinucleótido aislado según la presente invención.

En un aspecto adicional, la presente invención proporciona un polipéptido aislado que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, estando la mutación por sustitución en una posición equivalente o análoga a

- 10 a) el residuo de aminoácido 107 de SEQ ID NO: 1,  
 b) los residuos de aminoácido 107 y 175 de SEQ ID NO: 1,  
 15 c) los residuos de aminoácido 107 y 127 de SEQ ID NO: 1,  
 d) los residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o  
 20 e) los residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, y

en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

- 25 En otro aspecto, la presente invención proporciona un método de modulación de la expresión de un gen en una célula huésped que comprende las etapas de:

- 30 a) introducir en la célula huésped un sistema de modulación de la expresión génica que comprende un casete de expresión génica que puede expresarse en una célula huésped que comprende un polinucleótido que codifica para un polipéptido que comprende:

- i) un dominio de transactivación;  
 35 ii) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse; y  
 iii) un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución;

- 40 en el que el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor nuclear de grupo H seleccionado del grupo que consiste en un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína 15 de interacción de receptor retinoide X, receptores hepáticos X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor y un receptor de farnesol,  
 45

en el que el dominio de unión a ligando de receptor nuclear de grupo H está codificado por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido, estando el residuo de aminoácido en una posición equivalente o análoga a

- 50 a) el residuo de aminoácido 107 de SEQ ID NO: 1,  
 b) los residuos de aminoácido 107 y 175 de SEQ ID NO: 1,  
 55 c) los residuos de aminoácido 107 y 127 de SEQ ID NO: 1,  
 d) los residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o  
 e) los residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, y  
 60

en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de receptor nuclear de grupo H no mutado, y

- 65 b) introducir en la célula huésped un ligando;



en el que el gen que va a modularse es un componente de un casete de expresión génica que comprende:

i) un elemento de respuesta reconocido por el dominio de unión a ADN;

5 ii) un promotor que se activa por el dominio de transactivación; y

iii) un gen cuya expresión ha de modularse;

mediante lo cual tras la introducción del ligando en la célula huésped, se modula la expresión del gen de b) iii).

10 La presente invención proporciona, en un aspecto adicional, una célula huésped aislada que comprende el sistema de modulación de la expresión génica según la presente invención.

15 La presente invención proporciona, en otro aspecto, un organismo no humano que comprende la célula huésped de la presente invención.

En otro aspecto, la presente invención proporciona un vector que comprende el sistema de modulación de la expresión génica de la presente invención, el polinucleótido aislado de la presente invención o el casete de expresión génica de la presente invención.

20 La presente invención proporciona, en otro aspecto, una célula huésped aislada que comprende el vector de la presente invención.

### Definiciones

25 En esta descripción, se usan varios términos y abreviaturas. Se proporcionan las siguientes definiciones y deben ser útiles en el entendimiento del alcance y la práctica de la presente invención.

30 En una realización específica, el término "aproximadamente" significa dentro del 20%, preferiblemente dentro del 10%, más preferiblemente dentro del 5% e incluso más preferiblemente dentro del 1% de un intervalo o valor dado.

35 El término "sustancialmente libre" significa que una composición que comprende "A" (en la que "A" es una proteína, molécula de ADN, vector, célula huésped recombinante individual, etc.) está sustancialmente libre de "B" (en la que "B" comprende una o más proteínas, moléculas de ADN, vectores contaminantes, etc.) cuando al menos aproximadamente el 75% en peso de las proteínas, ADN, vectores (dependiendo de la categoría de las especies a la que pertenecen A y B) en la composición es "A". Preferiblemente, "A" comprende al menos aproximadamente el 90% en peso de las especies A + B en la composición, lo más preferiblemente al menos aproximadamente el 99% en peso. También se prefiere que una composición, que está sustancialmente libre de contaminación, contenga sólo una única especie en peso molecular que tiene la actividad o característica de las especies de interés.

40 El término "aislado" para los fines de la presente invención designa un material biológico (ácido nucleico o proteína) que se ha retirado de su entorno original (el entorno en el que está presente de manera natural). Por ejemplo, un polinucleótido presente en el estado natural en una planta o un animal no está aislado, sin embargo, el mismo polinucleótido separado de los ácidos nucleicos adyacentes en los que está presente de manera natural, se considera "aislado". El término "purificado" no requiere que el material esté presente en una forma que presenta pureza absoluta, excluyendo la presencia de otros compuestos. Es más bien una definición relativa.

45 Un polinucleótido está en el estado "purificado" tras la purificación del material de partida o del material natural en al menos un orden de magnitud, preferiblemente 2 ó 3 y preferiblemente 4 ó 5 órdenes de magnitud.

50 Un "ácido nucleico" es un compuesto polimérico compuesto por subunidades unidas de manera covalente denominadas nucleótidos. El ácido nucleico incluye ácido polirribonucleico (ARN) y ácido polidesoxirribonucleico (ADN), ambos de los cuales pueden ser monocatenarios o bicatenarios. ADN incluye pero no se limita a ADNc, ADN genómico, ADN plasmídico, ADN sintético y ADN semisintético. El ADN puede ser lineal, circular o superenrollado.

55 Una "molécula de ácido nucleico" se refiere a la forma polimérica de éster de fosfato de ribonucleósidos (adenosina, guanosina, uridina o citidina; "moléculas de ARN") o desoxirribonucleósidos (desoxiadenosina, desoxiguanosina, desoxitimidina o desoxicitidina; "moléculas de ADN"), o cualquier análogo de fosfoéster de las mismas, tales como fosforotioatos y tioésteres, o bien en forma monocatenaria, o bien una hélice bicatenaria. Son posibles hélices ADN-ADN, ADN-ARN y ARN-ARN bicatenarias. El término molécula de ácido nucleico, y en particular molécula de ADN o ARN, se refiere sólo a la estructura primaria y secundaria de la molécula, y no se limita a ninguna forma terciaria particular. Por tanto, este término incluye ADN bicatenario hallado, entre otros, en moléculas de ADN lineal o circular (por ejemplo, fragmentos de restricción), plásmidos y cromosomas. Al discutir la estructura de moléculas de ADN bicatenarias particulares, las secuencias pueden describirse en el presente documento según la convención normal de proporcionar sólo la secuencia en la dirección de 5' a 3' a lo largo de la hebra no transcrita de ADN (es decir, la hebra que tiene una secuencia homóloga al ARNm). Una "molécula de ADN recombinante" es una molécula de ADN

que ha experimentado una manipulación biológica molecular.

Se entenderá que el término "fragmento" significa una secuencia de nucleótidos de longitud reducida con respecto al ácido nucleico de referencia y que comprende, a lo largo de la parte común, una secuencia de nucleótidos idéntica al ácido nucleico de referencia. Un fragmento de ácido nucleico de ese tipo según la descripción puede incluirse, cuando sea apropiado, en un polinucleótido más grande del cual es un constituyente. Tales fragmentos comprenden, o alternativamente consisten en, oligonucleótidos que oscilan en longitud entre al menos 6, 8, 9, 10, 12, 15, 18, 20, 21, 22, 23, 24, 25, 30, 39, 40, 42, 45, 48, 50, 51, 54, 57, 60, 63, 66, 70, 75, 78, 80, 90, 100, 105, 120, 135, 150, 200, 300, 500, 720, 900, 1000 ó 1500 nucleótidos consecutivos de un ácido nucleico según la descripción.

Tal como se usa en el presente documento, un "fragmento de ácido nucleico aislado" es un polímero de ARN o ADN que es mono o bicatenario, que contiene opcionalmente bases de nucleótidos sintéticas, no naturales o alteradas. Un fragmento de ácido nucleico aislado en forma de un polímero de ADN puede estar compuesto por uno o más segmentos de ADNc, ADN genómico o ADN sintético.

Un "gen" se refiere a un conjunto de nucleótidos que codifica para un polipéptido, e incluye ácidos nucleicos de ADNc y ADN genómico. "Gen" también se refiere a un fragmento de ácido nucleico que expresa una proteína o polipéptido específico, incluyendo secuencias reguladoras que preceden (secuencias no codificantes en 5') y que siguen (secuencias no codificantes en 3') a la secuencia codificante. "Gen nativo" se refiere a un gen tal como se encuentra en la naturaleza con sus propias secuencias reguladoras. "Gen quimérico" se refiere a cualquier gen que no es un gen nativo, que comprende secuencias codificantes y/o reguladoras que no se encuentran juntas en la naturaleza. Por consiguiente, un gen quimérico puede comprender secuencias reguladoras y secuencias codificantes que se derivan de diferentes fuentes, o secuencias reguladoras y secuencias codificantes derivadas de la misma fuente, pero dispuestas de manera diferente de la encontrada en la naturaleza. Un gen quimérico puede comprender secuencias codificantes derivadas de diferentes fuentes y/o secuencias reguladoras derivadas de diferentes fuentes. "Gen endógeno" se refiere a un gen nativo en su ubicación natural en el genoma de un organismo. Un gen "foráneo" o gen "heterólogo" se refiere a un gen no encontrado normalmente en el organismo huésped, pero que se introduce en el organismo huésped mediante transferencia génica. Los genes foráneos pueden comprender genes nativos insertados en un organismo no nativo, o genes quiméricos. Un "transgén" es un gen que se ha introducido en el genoma mediante un procedimiento de transformación.

ADN "heterólogo" se refiere a ADN no ubicado de manera natural en la célula, o en un sitio cromosómico de la célula. Preferiblemente, el ADN heterólogo incluye un gen foráneo a la célula.

El término "genoma" incluye ADN o ARN cromosómico así como mitocondrial, de cloroplastos y viral.

Una molécula de ácido nucleico "puede hibridarse" con otra molécula de ácido nucleico, tal como un ADNc, ADN genómico o ARN, cuando una forma monocatenaria de la molécula de ácido nucleico puede aparearse con la otra molécula de ácido nucleico en condiciones apropiadas de temperatura y fuerza iónica en disolución (véase Sambrook *et al.*, 1989 más adelante). Se conocen bien condiciones de hibridación y lavado y se muestran a modo de ejemplo en Sambrook, J., Fritsch, E. F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*, segunda edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989), particularmente el capítulo 11 y la tabla 11.1 en ese documento. Las condiciones de temperatura y fuerza iónica determinan la "rigurosidad" de la hibridación.

Las condiciones de rigurosidad pueden ajustarse para seleccionar de fragmentos moderadamente similares, tales como secuencias homólogas de organismos relacionados de forma distante, a fragmentos altamente similares, tales como genes que duplican enzimas funcionales de organismos estrechamente relacionados. Para la selección preliminar de ácidos nucleicos homólogos, pueden usarse condiciones de hibridación de baja rigurosidad, que corresponden a una  $T_m$  de 55°, por ejemplo, 5x SSC, SDS al 0,1%, leche al 0,25% y sin formamida; o formamida al 30%, 5x SSC, SDS al 0,5%. Las condiciones de hibridación de rigurosidad moderada corresponden a una  $T_m$  superior, por ejemplo, formamida al 40%, con 5x o 6x SCC. Las condiciones de hibridación de alta rigurosidad corresponden a la  $T_m$  más alta, por ejemplo, formamida al 50%, 5x o 6x SCC.

La hibridación requiere que los dos ácidos nucleicos contengan secuencias complementarias, aunque dependiendo de la rigurosidad de la hibridación, son posibles apareamientos erróneos entre bases. El término "complementario" se usa para describir la relación entre bases de nucleótidos que pueden hibridarse entre sí. Por ejemplo, con respecto al ADN, la adenosina es complementaria a la timina y la citosina es complementaria a la guanina. Por consiguiente, la descripción también incluye fragmentos de ácido nucleico aislados que son complementarios a las secuencias completas tal como se da a conocer o se usa en el presente documento así como las secuencias de ácido nucleico sustancialmente similares.

En una realización específica se detectan polinucleótidos empleando condiciones de hibridación que comprenden una etapa de hibridación a  $T_m$  de 55°C, y utilizando condiciones tal como se expuso anteriormente. En una realización preferida, la  $T_m$  es de 60°C; en una realización más preferida, la  $T_m$  es de 63°C; en una realización incluso más preferida, la  $T_m$  es de 65°C.

Los lavados tras la hibridación también determinan las condiciones de rigurosidad. Un conjunto de condiciones preferidas usa una serie de lavados que comienzan con 6X SSC, SDS al 0,5% a temperatura ambiente durante 15 minutos (min.), entonces se repiten con 2X SSC, SDS al 0,5% a 45°C durante 30 minutos y entonces se repiten dos veces con 0,2X SSC, SDS al 0,5% a 50°C durante 30 minutos. Un conjunto más preferido de condiciones rigurosas usa temperaturas superiores en las que los lavados son idénticos a los anteriores excepto porque la temperatura de los dos lavados finales de 30 min. en 0,2X SSC, SDS al 0,5% se aumentó hasta 60°C. Otro conjunto preferido de condiciones altamente rigurosas usa dos lavados finales en 0,1X SSC, SDS al 0,1% a 65°C. La hibridación requiere que los dos ácidos nucleicos comprendan secuencias complementares, aunque dependiendo de la rigurosidad de la hibridación, son posibles apareamientos erróneos entre bases.

La rigurosidad apropiada para hibridar ácidos nucleicos depende de la longitud de los ácidos nucleicos y el grado de complementación, variables bien conocidas en la técnica. Cuanto mayor sea el grado de similitud u homología entre dos secuencia de nucleótidos, mayor será el valor de  $T_m$  para híbridos de ácidos nucleicos que tienen esas secuencias. La estabilidad relativa (correspondiente a  $T_m$  superior) de las hibridaciones de ácidos nucleicos disminuye en el siguiente orden: ARN:ARN, ADN:ARN, ADN:ADN. Para híbridos de más de 100 nucleótidos de longitud, se han derivado ecuaciones para calcular la  $T_m$  (véase Sambrook *et al.*, citado anteriormente, 9.50-0.51). Para la hibridación con ácidos nucleicos más cortos, es decir, oligonucleótidos, la posición de los apareamientos erróneos se vuelve más importante, y la longitud de los oligonucleótido determina su especificidad (véase Sambrook *et al.*, citado anteriormente, 11.7-11.8).

En una realización específica se detectan polinucleótidos empleando condiciones de hibridación que comprenden una etapa de hibridación en menos de 500 mM de sal y al menos 37 grados Celsius, y una etapa de lavado en 2XSSPE a al menos 63 grados Celsius. En una realización preferida, las condiciones de hibridación comprenden menos de 200 mM de sal y al menos 37 grados Celsius para la etapa de hibridación. En una realización más preferida, las condiciones de hibridación comprenden 2XSSPE y 63 grados Celsius para las etapas de tanto hibridación como lavado.

En una realización, la longitud para un ácido nucleico hibridable es de al menos aproximadamente 10 nucleótidos. De manera preferible, una longitud mínima para un ácido nucleico hibridable es de al menos aproximadamente 15 nucleótidos; más preferiblemente al menos aproximadamente 20 nucleótidos; y lo más preferiblemente la longitud es de al menos 30 nucleótidos. Además, el experto reconocerá que la temperatura y concentración salina de la disolución de lavado pueden ajustarse según sea necesario según factores tales como la longitud de la sonda.

El término "sonda" se refiere a una molécula de ácido nucleico monocatenaria que puede experimentar apareamiento de bases con un ácido nucleico diana monocatenario complementario puede formar una molécula bicatenaria.

Tal como se usa en el presente documento, el término "oligonucleótido" se refiere a un ácido nucleico, en general de al menos 18 nucleótidos, que puede hibridarse con una molécula de ADN genómico, una molécula de ADNc, un ADN plasmídico o una molécula de ARNm. Los oligonucleótidos pueden marcarse, por ejemplo, con  $^{32}\text{P}$ -nucleótidos o nucleótidos en los que se ha conjugado de manera covalente un marcador, tal como biotina. Un oligonucleótido marcado puede usarse como sonda para detectar la presencia de un ácido nucleico. Pueden usarse oligonucleótidos (uno o ambos de los cuales puede estar marcado) como cebadores de PCR, o bien para clonar la longitud completa o un fragmento de un ácido nucleico, o bien para detectar la presencia de un ácido nucleico. Un oligonucleótido también puede usarse para formar una hélice triple con una molécula de ADN. En general, se preparan oligonucleótidos de manera sintética, preferiblemente en un sintetizador de ácido nucleico. Por consiguiente, pueden prepararse oligonucleótidos con enlaces análogos a fosfoéster que no se producen de manera natural, tales como enlaces tioéster, etc.

Un "cebador" es un oligonucleótido que se hibrida con una secuencia de ácido nucleico diana para crear una región de ácido nucleico bicatenario que puede servir como un punto de iniciación para la síntesis de ADN en condiciones adecuadas. Tales cebadores pueden usarse en una reacción en cadena de la polimerasa.

"Reacción en cadena de la polimerasa" se abrevia PCR y significa un método *in vitro* para amplificar de manera enzimática secuencias de ácido nucleico específicas. La PCR implica una serie repetitiva de ciclos de temperatura, comprendiendo cada ciclo tres fases: desnaturalización del ácido nucleico molde para separar las hebras de la molécula diana, apareamiento de un cebador de oligonucleótido de PCR monocatenario con el ácido nucleico molde, y extensión del/de los cebador(es) apareado(s) mediante ADN polimerasa. La PCR proporciona un medio para detectar la presencia de la molécula diana y, en condiciones cuantitativas o semicuantitativas, para determinar la cantidad relativa de esta molécula diana dentro del conjunto de partida de ácidos nucleicos.

"Reacción en cadena de la polimerasa con transcripción inversa" se abrevia RT-PCR y significa un método *in vitro* para producir de manera enzimática una molécula o moléculas de ADNc diana a partir de una molécula o moléculas de ARN, seguido por amplificación enzimática de una secuencia o secuencias de ácido nucleico específicas dentro de la molécula o moléculas de ADNc diana tal como se describió anteriormente. RT-PCR también proporciona un medio para detectar la presencia de la molécula diana y, en condiciones cuantitativas o semicuantitativas, para

determinar la cantidad relativa de esta molécula diana dentro del conjunto de partida de ácidos nucleicos.

Una "secuencia codificante" de ADN es una secuencia de ADN bicatenaria que se transcribe y se traduce para dar un polipéptido en una célula *in vitro* o *in vivo* cuando se coloca bajo el control de secuencias reguladoras apropiadas.

5 "Secuencias reguladoras adecuadas" se refiere a secuencias de nucleótidos ubicadas en el sentido de 5' (secuencias no codificantes en 5'), dentro o en el sentido de 3' (secuencias no codificantes en 3') de una secuencia codificante, y que influyen en la transcripción, estabilidad o procesamiento del ARN o traducción de la secuencia codificante asociada. Las secuencias reguladoras pueden incluir promotores, secuencias líder de traducción, intrones, secuencias de reconocimiento de poliadenilación, sitio de procesamiento de ARN, sitio de unión a efector y estructura de tallo-bucle. Los límites de la secuencia codificante están determinados por un codón de iniciación en el extremo terminal 5' (amino) y un codón de terminación de la traducción en el extremo terminal 3' (carboxilo). Una secuencia codificante puede incluir, pero no se limita a, secuencias procariotas, ADNc a partir de ARNm, secuencias de ADN genómico e incluso secuencias de ADN sintético. Si la secuencia codificante está destinada a la expresión en una célula eucariota, se ubicarán habitualmente una señal de poliadenilación y secuencia de terminación de la transcripción en 3' con respecto a la secuencia codificante.

20 "Marco de lectura abierto" se abrevia ORF y significa una longitud de secuencia de ácido nucleico, o bien ADN, ADNc o bien ARN, que comprende una señal de inicio de la traducción o codón de iniciación, tal como un ATG o AUG, y un codón de terminación y puede traducirse potencialmente para dar una secuencia polipeptídica.

El término "cabeza a cabeza" se usa en el presente documento para describir la orientación de dos secuencias de polinucleótido en relación entre sí. Se posicionan dos polinucleótidos en una orientación cabeza a cabeza cuando el extremo 5' de la hebra codificante de un polinucleótido es adyacente al extremo 5' de la hebra codificante del otro polinucleótido, mediante lo cual la dirección de transcripción de cada polinucleótido avanza desde el extremo 5' del otro polinucleótido. El término "cabeza a cabeza" puede abreviarse (5')-a-(5') y también puede indicarse por los símbolos ( $\leftarrow\rightarrow$ ) o ( $3'\leftarrow 5'5'\rightarrow 3'$ ).

30 El término "cola a cola" se usa en el presente documento para describir la orientación de dos secuencias de polinucleótido en relación entre sí. Se posicionan dos polinucleótidos en una orientación cola a cola cuando el extremo 3' de la hebra codificante de un polinucleótido es adyacente al extremo 3' de la hebra codificante del otro polinucleótido, mediante lo cual la dirección de transcripción de cada polinucleótido avanza hacia el otro polinucleótido. El término "cola a cola" puede abreviarse (3')-a-(3') y también puede indicarse por los símbolos ( $\rightarrow\leftarrow$ ) o ( $5'\rightarrow 3'3'\leftarrow 5'$ ).

35 El término "cabeza a cola" se usa en el presente documento para describir la orientación de dos secuencias de polinucleótido en relación entre sí. Se posicionan dos polinucleótidos en una orientación cabeza a cola cuando el extremo 5' de la hebra codificante de un polinucleótido es adyacente al extremo 3' de la hebra codificante del otro polinucleótido, mediante lo cual la dirección de transcripción de cada polinucleótido avanza en la misma dirección que la del otro polinucleótido. El término "cabeza a cola" puede abreviarse (5')-a-(3') y también puede indicarse por los símbolos ( $\rightarrow\rightarrow$ ) o ( $5'\rightarrow 3'5'\rightarrow 3'$ ).

45 El término "en el sentido de 3'" se refiere a una secuencia de nucleótidos que se ubica en 3' con respecto a la secuencia de nucleótidos de referencia. En particular, secuencias de nucleótidos en el sentido de 3' se refiere en general a secuencias que siguen el punto de inicio de la transcripción. Por ejemplo, el codón de iniciación de la traducción de un gen se ubica en el sentido de 3' del sitio de inicio de la transcripción.

50 El término "en el sentido de 5'" se refiere a una secuencia de nucleótidos que se ubica en 5' con respecto a la secuencia de nucleótidos de referencia. En particular, secuencias de nucleótidos en el sentido de 5' se refiere en general a secuencias que se ubican en el lado en 5' de una secuencia codificante o punto de inicio de la transcripción. Por ejemplo, la mayoría de los promotores se ubican en el sentido de 5' del sitio de inicio de la transcripción.

Los términos "endonucleasa de restricción" y "enzima de restricción" se refieren a una enzima que se une y corta dentro de una secuencia de nucleótidos específica dentro de un ADN bicatenario.

55 "Recombinación homóloga" se refiere a la inserción de una secuencia de ADN foráneo en otra molécula de ADN, por ejemplo, la inserción de un vector en un cromosoma. Preferiblemente, el vector selecciona como diana un sitio cromosómico específico para la recombinación homóloga. Para la recombinación homóloga específica, el vector contendrá regiones de homología suficientemente largas con respecto a secuencias del cromosoma para permitir la unión complementaria y la incorporación del vector en el cromosoma. Regiones de homología más largas, y mayores grados de similitud de secuencia, pueden aumentar la eficacia de la recombinación homóloga.

65 Pueden usarse varios métodos conocidos en la técnica para propagar un polinucleótido según la invención. Una vez que se establecen un sistema huésped y condiciones de crecimiento adecuados, los vectores de expresión recombinantes pueden propagarse y prepararse en cantidad. Tal como se describe en el presente documento, los vectores de expresión que pueden usarse incluyen, pero no se limitan a, los siguientes vectores o sus derivados:

virus humano o animal tal como virus vaccinia o adenovirus; virus de insecto tales como baculovirus; vectores de levadura; vectores de bacteriófagos (por ejemplo, lambda), y vectores de ADN de plásmido y cósmido, por citar algunos.

5 Un "vector" es cualquier medio para la clonación de y/o transferencia de un ácido nucleico a una célula huésped. Un vector puede ser un replicón al que puede unirse otro segmento de ADN de modo que se ocasiona la replicación del segmento unido. Un "replicón" es cualquier elemento genético (por ejemplo, plásmido, fago, cósmido, cromosoma, virus) que funciona como una unidad autónoma de replicación de ADN *in vivo*, es decir, que puede replicarse bajo su propio control. El término "vector" incluye tanto medios virales como no virales para introducir el ácido nucleico en una célula *in vitro*, *ex vivo* o *in vivo*. Puede usarse un gran número de vectores conocidos en la técnica para manipular ácidos nucleicos, incorporar elementos de respuesta y promotores en genes, etc. Los posibles vectores incluyen, por ejemplo, plásmidos o virus modificados incluyendo, por ejemplo, bacteriófagos tales como derivados de lambda, o plásmidos tales como derivados de plásmidos pBR322 o pUC, o el vector Bluescript. Por ejemplo, la inserción de los fragmentos de ADN correspondientes a elementos de respuesta y promotores en un vector adecuado puede realizarse ligando los fragmentos de ADN apropiados en un vector elegido que tiene extremos terminales cohesivos complementarios. Alternativamente, los extremos de las moléculas de ADN pueden modificarse de manera enzimática o puede producirse cualquier sitio ligando secuencias de nucleótidos (ligadores) en los extremos terminales de ADN. Tales vectores pueden modificarse por ingeniería para contener genes marcadores seleccionables que proporcionan la selección de células que han incorporado el marcador en el genoma celular. Tales marcadores permiten la identificación y/o selección de células huésped que incorporan y expresan las proteínas codificadas por el marcador.

Se han usado vectores virales, y particularmente vectores retrovirales, en una amplia variedad de aplicaciones de suministro de genes en células, así como sujetos animales vivos. Los vectores virales que pueden usarse incluyen pero no se limitan a vectores de retrovirus, virus adenoasociados, viruela, baculovirus, vaccinia, herpes simple, Epstein-Barr, adenovirus, geminivirus y caulimovirus. Los vectores no virales incluyen plásmidos, liposomas, lípidos cargados eléctricamente (citofectinas), complejos de ADN-proteína y biopolímeros. Además de un ácido nucleico, un vector también puede comprender una o más regiones reguladoras, y/o marcadores seleccionables útiles en la selección, medición y monitorización de los resultados de la transferencia de ácido nucleico (transferencia a qué tejidos, duración de la expresión, etc.).

El término "plásmido" se refiere a un elemento extracromosómico que porta a menudo un gen que no es parte del metabolismo central de la célula, y habitualmente en la forma de moléculas de ADN bicatenario circular. Tales elementos pueden ser secuencias de replicación autónoma, secuencias de integración en el genoma, secuencias de nucleótidos o fagos, secuencias lineales, circulares o superenrolladas de un ADN o ARN mono o bicatenario, derivado de cualquier fuente, en los que varias secuencias de nucleótidos se han unido o recombinado para dar un único constructo que puede introducir un fragmento promotor y una secuencia de ADN para un producto génico seleccionado junto con una secuencia no traducida en 3' apropiada en una célula.

Un "vector de clonación" es un "replicón", que es una longitud unitaria de un ácido nucleico, preferiblemente ADN, que se replica secuencialmente y que comprende un origen de replicación, tal como un plásmido, fago o cósmido, al que puede unirse otro segmento de ácido nucleico para ocasionar la replicación del segmento unido. Los vectores de clonación pueden replicarse en un tipo celular y expresarse en otro ("vector lanzadera").

Pueden introducirse vectores en las células huésped deseadas mediante métodos conocidos en la técnica, por ejemplo, transfección, electroporación, microinyección, transducción, fusión celular, DEAE dextrano, precipitación con fosfato de calcio, lipofección (fusión de liposomas), uso de un pistola génica o un transportador de vector de ADN (véanse, por ejemplo, Wu *et al.*, 1992, J. Biol. Chem. 267: 963-967; Wu y Wu, 1988, J. Biol. Chem. 263: 14621-14624; y Hartmut *et al.*, solicitud de patente canadiense n.º 2.012.311, presentada el 15 de marzo de 1990).

Un polinucleótido según la invención también puede introducirse *in vivo* mediante lipofección. Durante la última década, ha habido un uso creciente de liposomas para encapsulación y transfección de ácidos nucleicos *in vitro*. Pueden usarse lípidos catiónicos sintéticos diseñados para limitar las dificultades y peligros encontrados con la transfección mediada por liposomas para preparar liposomas para la transfección *in vivo* de un gen que codifica para un marcador (Felgner *et al.*, 1987, Proc. Natl. Acad. Sci. U.S.A. 84: 7413; Mackey, *et al.*, 1988, Proc. Natl. Acad. Sci. U.S.A. 85:8027-8031; y Ulmer *et al.*, 1993, Science 259: 1745-1748). El uso de lípidos catiónicos puede promover la encapsulación de ácidos nucleicos cargados negativamente, y también promover la fusión con membranas celulares cargadas negativamente (Felgner y Ringold, 1989, Science 337:387-388). Se describen composiciones y compuestos lipídicos particularmente útiles para la transferencia de ácidos nucleicos en las publicaciones internacionales de patente WO 95/18863 y WO 96/17823, y en la patente de EE.UU. n.º 5.459.127. El uso de lipofección para introducir genes exógenos en los órganos específicos *in vivo* tiene determinadas ventajas prácticas. El direccionamiento molecular de liposomas a células específicas representa un área de beneficio. Está claro que dirigir la transfección a tipos celulares particulares se preferiría particularmente en un tejido con heterogeneidad celular, tal como páncreas, hígado, riñón y el cerebro. Pueden acoplarse químicamente lípidos a otras moléculas para el fin de direccionamiento (Mackey, *et al.*, 1988, citado anteriormente). Péptidos dirigidos, por ejemplo, hormonas o neurotransmisores, y proteínas tales como anticuerpos, o moléculas no peptídicas podrían acoplarse a

liposomas químicamente.

Otras moléculas también son útiles para facilitar la transfección de un ácido nucleico *in vivo*, tal como un oligopéptido catiónico (por ejemplo, el documento WO 95/21931), péptidos derivados de proteínas de unión a ADN (por ejemplo, el documento WO 96/25508) o un polímero catiónico (por ejemplo, el documento WO 95/21931).

También es posible introducir un vector *in vivo* como un plásmido de ADN desnudo (véase las patentes de EE.UU. 5.693.622, 5.589.466 y 5.580.859). También pueden usarse enfoques de suministro de ADN mediado por receptor (Curiel *et al.*, 1992, Hum. Gen Ther. 3: 147-154; y Wu y Wu, 1987, J. Biol. Chem. 262: 4429-4432).

El término “transfección” significa la captación de ARN o ADN exógeno o heterólogo por una célula. Una célula se ha “transfectado” por ARN o ADN exógeno o heterólogo cuando se ha introducido tal ARN o ADN dentro de la célula. Una célula se ha “transformado” por ARN o ADN exógeno o heterólogo cuando el ARN o ADN transfectado efectúa un cambio fenotípico. El ARN o ADN transformante puede integrarse (unirse de manera covalente) en el ADN cromosómico que constituye el genoma de la célula.

“Transformación” se refiere a la transferencia de un fragmento de ácido nucleico al genoma de un organismo huésped, dando como resultado herencia genéticamente estable. Los organismos huésped que contienen los fragmentos de ácido nucleico transformados se denominan organismos “transformados” o “recombinantes” o “transformados”.

El término “región genética” se referirá a una región de una molécula de ácido nucleico o una secuencia de nucleótidos que comprende un gen que codifica para un polipéptido.

Además, el vector recombinante que comprende un polinucleótido según la invención puede incluir uno o más orígenes para la replicación en los huéspedes celulares en los que se busca su amplificación o su expresión, marcadores o marcadores seleccionables.

El término “marcador seleccionable” significa un factor de identificación, habitualmente un gen de resistencia a antibióticos o productos químicos, que puede seleccionarse basándose en el efectos del gen marcador, es decir, resistencia a un antibiótico, resistencia a un herbicida, marcadores colorimétricos, enzimas, marcadores fluorescentes, y similares, en los que se usa el efecto para rastrear la herencia de un ácido nucleico de interés y/o para identificar una célula u organismo que ha heredado el ácido nucleico de interés. Los ejemplos de genes marcadores seleccionables conocidos y usados en la técnica incluyen: genes que proporcionan resistencia a ampicilina, estreptomycin, gentamicina, kanamicina, higromicina, herbicida bialafos, sulfonamida, y similares; y genes que se usan como marcadores fenotípicos, es decir, genes reguladores de antocianina, gen de isopentanol transferasa, y similares.

El término “gen indicador” significa un ácido nucleico que codifica para un factor de identificación que puede identificarse basándose en el efecto del gen indicador, en el que se usa el efecto para rastrear la herencia de un ácido nucleico de interés, identificar una célula u organismo que ha heredado el ácido nucleico de interés y/o medir la inducción o transcripción de la expresión génica. Los ejemplos de genes indicadores conocidos y usados en la técnica incluyen: luciferasa (Luc), proteína fluorescente verde (GFP), cloranfenicol acetiltransferasa (CAT),  $\beta$ -galactosidasa (LacZ),  $\beta$ -glucuronidasa (Gus), y similares. Los genes marcadores seleccionables también pueden considerarse genes indicadores.

“Promotor” se refiere a una secuencia de ADN que puede controlar la expresión de una secuencia codificante o ARN funcional. En general, una secuencia codificante se ubica en 3' con respecto a una secuencia promotora. Los promotores pueden derivarse en su totalidad de un gen nativo, o pueden estar compuestos por diferentes elementos derivados de diferentes promotores encontrados en la naturaleza, o incluso pueden comprender segmentos de ADN sintético. Los expertos en la técnica entienden que diferentes promotores pueden dirigir la expresión de un gen en diferentes tejidos o tipos celulares, o en diferentes estadios de desarrollo, o en respuesta a condiciones ambientales o fisiológicas diferentes. Los promotores que hacen que un gen se exprese en la mayoría de los tipos celulares la mayoría de las veces se denominan comúnmente “promotores constitutivos”. Los promotores que hacen que un gen se exprese en un tipo celular específico se denominan comúnmente “promotores específicos de célula” o “promotores específicos de tejido”. Los promotores que hacen que un gen se exprese en un estadio específico de desarrollo o diferenciación celular se denominan comúnmente “promotores específicos del desarrollo” o “promotores específicos de diferenciación celular”. Los promotores que se inducen y hacen que un gen se exprese tras la exposición o el tratamiento de la célula con un agente, molécula biológica, producto químico, ligando, luz, o similares que induce el promotor se denominan comúnmente “promotores inducibles” o “promotores regulables”. Se reconoce además que ya que en la mayoría de los casos los límites exactos de secuencias reguladoras no se han definido completamente, fragmentos de ADN de diferentes longitudes pueden tener actividad promotora idéntica.

Una “secuencia promotora” es una región reguladora de ADN que puede unirse a ARN polimerasa en una célula e iniciar la transcripción de una secuencia codificante en el sentido de 3' (dirección 3'). Para fines de definición de la presente invención, la secuencia promotora está limitada en su extremo terminal 3' por el sitio de iniciación de la

transcripción y se extiende en el sentido de 5' (dirección 5') incluyendo el número mínimo de bases o elementos necesarios para iniciar la transcripción a niveles detectables por encima del fondo. Dentro de la secuencia promotora se encontrará un sitio de iniciación de la transcripción (definido convenientemente por ejemplo, mediante mapeo con nucleasa S1), así como dominios de unión a proteína (secuencias consenso) responsables de la unión de ARN polimerasa.

Una secuencia codificante está "bajo el control" de secuencias de control de la transcripción y traducción en una célula cuando la ARN polimerasa transcribe la secuencia codificante para dar ARNm, que entonces se corta y empalma en trans (si la secuencia codificante contiene intrones) y se traduce para dar la proteína codificada por la secuencia codificante.

"Secuencias de control de la transcripción y traducción" son secuencias reguladoras de ADN, tales como promotores, potenciadores, terminadores, y similares, que proporcionan la expresión de una secuencia codificante en una célula huésped. En células eucariotas, las señales de poliadenilación son secuencias de control.

El término "elemento de respuesta" significa uno o más elementos de ADN de actuación en cis que confieren receptividad en un promotor mediada a través de la interacción con los dominios de unión a ADN del primer gen quimérico. Este elemento de ADN puede ser o bien palindrómico (perfecto o imperfecto) en su secuencia o bien puede estar compuesto por motivos de secuencias o semisitios separados por un número variable de nucleótidos. Los semisitios pueden ser similares o idénticos y estar dispuestos como repeticiones o bien directas o bien invertidas o como un semisitio individual o multímeros de semisitios adyacentes en tándem. El elemento de respuesta puede comprender un promotor mínimo aislado de diferentes organismos dependiendo de la naturaleza de la célula u organismo en el que se incorporará el elemento de respuesta. El dominio de unión a ADN de la primera proteína híbrida se une, en presencia o ausencia de un ligando, a la secuencia de ADN de un elemento de respuesta para iniciar o suprimir la transcripción del/de los gen(es) en el sentido de 3' bajo la regulación de este elemento de respuesta. Los ejemplos de secuencias de ADN para elementos de respuesta del receptor de ecdisona natural incluyen: RGGG/TTCANTGAC/ACY (véase Cherbas L., *et. al.*, (1991), *Genes Dev.* 5,120-131); AGGTCAN<sub>(n)</sub>AGGTCA, en el que N<sub>(n)</sub> puede ser uno o más nucleótidos espaciadores (véase D'Avino PP., *et. al.*, (1995), *Mol. Cell. Endocrinol.* 113, 1-9); y GGGTTGAATGAATTT (véase Antoniewski C., *et. al.*, (1994). *Mol. Cell Biol.* 14,4465-4474).

El término "operativamente unido" se refiere a la asociación de secuencias de ácido nucleico en un único fragmento de ácido nucleico de modo que la función de uno se ve afectada por el otro. Por ejemplo, un promotor está operativamente unido a una secuencia codificante cuando puede afectar la expresión de la secuencia codificante (es decir, que la secuencia codificante está bajo el control de la transcripción del promotor). Las secuencias codificantes pueden estar operativamente unidas a secuencias reguladoras en orientación sentido o antisentido.

El término "expresión", tal como se usa en el presente documento, se refiere a la transcripción y acumulación estable de (ARNm) sentido o ARN antisentido derivado de un ácido nucleico o polinucleótido. Expresión también puede referirse a la traducción de ARNm para dar una proteína o polipéptido.

Los términos "casete", "casete de expresión" y "casete de expresión génica" se refieren a un segmento de ADN que puede insertarse en un ácido nucleico o polinucleótido en sitios de restricción específicos o mediante recombinación homóloga. El segmento de ADN comprende un polinucleótido que codifica para un polipéptido de interés, y el casete y los sitios de restricción se diseñan para garantizar la inserción del casete en el marco de lectura apropiado para la transcripción y traducción. "Casete de transformación" se refiere a un vector específico que comprende un polinucleótido que codifica para un polipéptido de interés y que tiene elementos además del polinucleótido que facilitan la transformación de una célula huésped particular. Casetes, casetes de expresión, casetes de expresión génica y casetes de transformación de la invención también pueden comprender elementos que permiten la expresión potenciada de un polinucleótido que codifica para un polipéptido de interés en una célula huésped. Estos elementos pueden incluir, pero no se limitan a: un promotor, un promotor mínimo, un potenciador, un elemento de respuesta, una secuencia terminadora, una secuencia de poliadenilación, y similares.

Para fines de esta invención, el término "interruptor génico" se refiere a la combinación de un elemento de respuesta asociado con un promotor, y un sistema basado en EcR que, en presencia de uno o más ligandos, modula la expresión de un gen en el que se incorporan el elemento de respuesta y el promotor.

Los términos "modular" y "modula" significan inducir, reducir o inhibir el ácido nucleico o la expresión génica, dando como resultado la respectiva inducción, reducción o inhibición de la producción de proteína o polipéptido.

Los plásmidos o vectores según la descripción pueden comprender además al menos un promotor adecuado para dirigir la expresión de un gen en una célula huésped. El término "vector de expresión" significa un vector, plásmido o vehículo diseñado para permitir la expresión de una secuencia de ácido nucleico insertada tras la transformación en el huésped. El gen clonado, es decir, la secuencia de ácido nucleico insertada, se coloca habitualmente bajo el control de elementos de control tales como un promotor, un promotor mínimo, un potenciador, o similares. Los promotores o regiones de control de la iniciación, que son útiles para dirigir la expresión de un ácido nucleico en la

célula huésped deseada son numerosos y familiares para los expertos en la técnica. Prácticamente cualquier promotor que pueda dirigir estos genes es adecuado para la presente invención incluyendo pero sin limitarse a: promotores virales, promotores bacterianos, promotores animales, promotores de mamíferos, promotores sintéticos, promotores constitutivos, promotor específico de tejido, promotores específicos del desarrollo, promotores inducibles, promotores regulados por la luz; CYC1, HIS3, GAL1, GAL4, GAL10, ADH1, PGK, PHO5, GAPDH, ADC1, TRP1, URA3, LEU2, ENO, TPI, promotores de fosfatasa alcalina (útiles para la expresión en *Saccharomyces*); promotor AOX1 (útil para la expresión en *Pichia*); promotores de  $\beta$ -lactamasa, lac, ara, tet, trp, IPL, LPR, T7, tac y trc (útiles para la expresión en *Escherichia coli*); promotores regulados por la luz, específicos de semilla, específicos de polen, específicos de ovarios, relacionados con patogénesis o enfermedad, 35S del virus del mosaico de la coliflor, mínimo 35S de CMV, del virus del mosaico de la mandioca (CsVMV), proteína de unión a clorofila a/b, ribulosa 1,5-bisfosfato carboxilasa, específicos de brotes, específicos de raíces, quitinasa, inducibles por estrés, del virus baciliforme del tungro del arroz, promotor superior de plantas, leucina aminopeptidasa de la patata, nitrato reductasa, manopina sintasa, nopalina sintasa, ubiquitina, proteína zeína y antocianina (útiles para la expresión en células vegetales); los promotores animales y de mamíferos conocidos en la técnica incluyen, pero no se limitan a, la región promotora temprana del SV40 (SV40e), el promotor contenido en la repetición terminal larga en 3' (LTR) del virus del sarcoma de Rous (RSV), los promotores de los genes tardíos principales E1A (MLP) de adenovirus (Ad), el promotor temprano del citomegalovirus (CMV), el promotor de timidina cinasa (TK) del virus del herpes simple (HSV), un promotor de baculovirus IE1, un promotor del factor de elongación 1 alfa (EF1), un promotor de la fosfoglicerato cinasa (PGK), un promotor de ubiquitina (Ubc), un promotor de albúmina, las secuencias reguladoras del promotor de la L-metalotioneína de ratón y regiones de control de la transcripción, los promotores ubicuos (HPRT, vimentina,  $\alpha$ -actina, tubulina y similares), los promotores de los filamentos intermedios (desmina, neurofilamentos, queratina, GFAP, y similares), los promotores de genes terapéuticos (del MDR, CFTR o tipo de factor VIII, y similares), promotores relacionados con enfermedad o patogénesis, y promotores que presentan especificidad de tejido y se han utilizado en animales transgénicos, tales como la región de control del gen I de la elastasa que es activa en células acinares pancreáticas; región de control del gen de la insulina activa en células beta pancreáticas, región de control del gen de la inmunoglobulina activa en células linfoides, región de control del virus de tumor mamario de ratón activa en células testiculares, de mama, linfoides y mastocitos; regiones de control del gen de la albúmina, Apo AI y Apo AII activas en el hígado, región de control del gen de la alfa-fetoproteína activa en el hígado, región de control del gen de la alfa 1-antitripsina activa en el hígado, región de control del gen de la beta-globina activa en células mieloides, región de control del gen de la proteína básica de la mielina activa en células de oligodendrocitos en el cerebro, región de control del gen de la cadena-2 ligera de la miosina activa en el músculo esquelético, y región de control del gen de la hormona liberadora de la gonadotropina activa en el hipotálamo, promotor de la piruvato cinasa, promotor de vilina, promotor de la proteína intestinal de unión a ácidos grasos, promotor de la  $\alpha$ -actina de células de músculo liso, y similares. Además, estas secuencias de expresión pueden modificarse mediante la adición de secuencias reguladoras o potenciadoras y similares.

Los potenciadores que pueden usarse en realizaciones de la invención incluyen pero no se limitan a: un potenciador de SV40, un potenciador de citomegalovirus (CMV), un potenciador de factor I de elongación (EF1), potenciadores de levadura, potenciadores de genes virales, y similares.

Las regiones de control de la terminación, es decir, secuencias terminadoras o de poliadenilación, también pueden derivarse de diversos genes nativos para los huéspedes preferidos. Opcionalmente, un sitio de terminación puede no ser necesario, sin embargo, se prefiere más si se incluye. En una realización preferida de la descripción, la región de control de la terminación puede estar compuesta por o derivarse de una secuencia sintética, señal de poliadenilación sintética, una señal de poliadenilación tardía de SV40, una señal de poliadenilación de SV40, una señal de poliadenilación de la hormona de crecimiento bovina (BGH), secuencias de terminadores virales, o similares.

Los términos "secuencias no codificantes en 3'" o "región no traducida en 3' (UTR)" se refieren a secuencias de ADN ubicadas en el sentido de 3' (3') de una secuencia codificante y pueden comprender secuencias de reconocimiento de la poliadenilación [poli(A)] y otras secuencias que codifican para señales reguladoras que pueden afectar al procesamiento de ARNm o a la expresión génica. La señal de poliadenilación se caracteriza habitualmente por afectar a la adición de tramos de ácido poliadenílico al extremo 3' del precursor de ARNm.

"Región reguladora" significa una secuencia de ácido nucleico que regula la expresión de una segunda secuencia de ácido nucleico. Una región reguladora puede incluir secuencias que son responsables de manera natural de la expresión de un ácido nucleico particular (una región homóloga) o pueden incluir secuencias de un origen diferente que son responsables de la expresión de diferentes proteínas o incluso proteínas sintéticas (una región heteróloga). En particular, las secuencias pueden ser secuencias de genes procariontes, eucariotas o virales o secuencias derivadas que estimulan o reprimen la transcripción de un gen de una manera específica o no específica y de una manera inducible o no inducible. Las regiones reguladoras incluyen orígenes de replicación, sitios de corte y empalme de ARN, promotores, potenciadores, secuencias de terminación de la transcripción y secuencias señal que dirigen el polipéptido a las rutas secretoras de la célula diana.

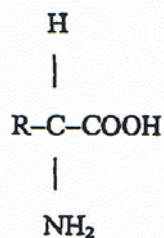
Una región reguladora de una "fuente heteróloga" es una región reguladora que no está asociada de manera natural con el ácido nucleico expresado. Se incluyen entre las regiones reguladoras heterólogas regiones reguladoras de una



especie diferente, regiones reguladoras de un gen diferente, secuencias reguladoras híbridas y secuencias reguladoras que no se producen en la naturaleza, pero que se diseñan por un experto en la técnica.

5 “Transcrito de ARN” se refiere al producto resultante de la transcripción catalizada por la ARN polimerasa de una secuencia de ADN. Cuando el transcrito de ARN es una copia complementaria perfecta de la secuencia de ADN, se denomina transcrito primario o puede ser una secuencia de ARN derivada del procesamiento postranscripcional del transcrito primario y se denomina ARN maduro. “ARN mensajero (ARN)” se refiere al ARN que no tiene intrones y que puede traducirse para dar una proteína por la célula. “ADNc” se refiere a un ADN bicatenario que es complementario a y se deriva de ARNm. ARN “sentido” se refiere al transcrito de ARN que incluye el ARNm y de ese modo puede traducirse para dar una proteína por la célula. “ARN antisentido” se refiere a un transcrito de ARN que es complementario a todo o parte de un transcrito primario diana o ARNm y que bloquea la expresión de un gen diana. La complementariedad de un ARN antisentido puede ser con cualquier parte del transcrito del gen específico, es decir, en la secuencia no codificante en 5', secuencia no codificante en 3' o la secuencia codificante. “ARN funcional” se refiere a un ARN antisentido, ARN de ribozima u otro ARN que no se traduce aunque tiene un efecto sobre procesos celulares.

Un “polipéptido” es un compuesto polimérico compuesto por residuos de aminoácido unidos de manera covalente. Los aminoácidos tienen la siguiente estructura general:



20 Los aminoácidos se clasifican en siete grupos basándose en la cadena lateral R: (1) cadenas laterales alifáticas, (2) cadenas laterales que contienen un grupo hidroxílico (OH), (3) cadenas laterales que contienen átomos de azufre, (4) cadenas laterales que contienen un grupo ácido o amida, (5) cadenas laterales que contienen un grupo básico, (6) cadenas laterales que contienen un anillo aromático, y (7) prolina, un iminoácido en el que la cadena lateral se condensa con el grupo amino. Un polipéptido de la descripción comprende preferiblemente al menos aproximadamente 14 aminoácidos.

30 Una “proteína” es un polipéptido que realiza un papel estructural o funcional en una célula viva.

Un “polipéptido aislado” o “proteína aislada” es un polipéptido o proteína que está sustancialmente libre de los compuestos que normalmente están asociados con los mismos en su estado natural (por ejemplo, otras proteínas o polipéptidos, ácidos nucleicos, hidratos de carbono, lípidos). “Aislado” no pretende excluir mezclas artificiales o sintéticas con otros compuestos, o la presencia de impurezas que no interfirieren con la actividad biológica, y que pueden estar presentes, por ejemplo, debido a purificación incompleta, adición de estabilizadores o composición en una preparación farmacéuticamente aceptable.

40 Un “polipéptido mutante por sustitución” o un “mutante por sustitución” se entenderá que significa un polipéptido mutante que comprende una sustitución de al menos un (1) aminoácido que se produce de manera natural o de tipo natural por un aminoácido diferente en relación con el polipéptido que se produce de manera natural o de tipo natural. Un polipéptido mutante por sustitución puede comprender sólo una (1) sustitución de aminoácido que se produce de manera natural o de tipo natural y puede denominarse un polipéptido “mutante puntual” o un “mutante en un único punto”. Alternativamente, un polipéptido mutante por sustitución puede comprender una sustitución de dos (2) o más aminoácidos que se producen de manera natural o de tipo natural con 2 o más aminoácidos en relación con el polipéptido que se produce de manera natural o de tipo natural. Según la invención, un polipéptido de dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución comprende una sustitución de al menos un (1) aminoácido que se produce de manera natural o de tipo natural con un aminoácido diferente en relación con el polipéptido de dominio de unión a ligando de receptor nuclear de grupo H que se produce de manera natural o de tipo natural.

50 Cuando el polipéptido mutante por sustitución comprende una sustitución de dos (2) o más aminoácidos que se producen de manera natural o de tipo natural, esta sustitución puede comprender o bien un número equivalente de aminoácidos que se producen de manera natural o de tipo natural delecionados para la sustitución, es decir, 2 aminoácidos que se producen de manera natural o de tipo natural reemplazados por 2 aminoácidos que no se producen de manera natural o de tipo natural, o bien un número no equivalente de aminoácidos de tipo natural delecionados para la sustitución, es decir, 2 aminoácidos de tipo natural reemplazados por 1 aminoácido de tipo no natural (una mutación por sustitución + delección), o 2 aminoácidos de tipo natural reemplazados por 3 aminoácidos de tipo no natural (una mutación por sustitución + inserción). Los mutantes por sustitución pueden describirse

- usando un sistema de nomenclatura abreviado para indicar el residuo de aminoácido y el número reemplazado dentro de la secuencia polipeptídica de referencia y el nuevo residuo de aminoácido sustituido. Por ejemplo, un mutante por sustitución en el que el vigésimo (20<sup>o</sup>) residuo de aminoácido de un polipéptido está sustituido puede abreviarse como “x20z”, en el que “x” es el aminoácido que va a reemplazarse, “20” es la posición o el número del residuo de aminoácido dentro del polipéptido y “z” es el nuevo aminoácido sustituido. Por tanto, un mutante por sustitución abreviado de manera intercambiable como “E20A” o “Glu20Ala” indica que el mutante comprende un residuo de alanina (comúnmente abreviado en la técnica como “A” o “Ala”) en lugar del ácido glutámico (comúnmente abreviado en la técnica como “E” o “Glu”) en la posición 20 del polipéptido.
- 10 Una mutación por sustitución puede prepararse mediante cualquier técnica para mutagénesis conocida en la técnica, incluyendo pero sin limitarse a, mutagénesis dirigida al sitio *in vitro* (Hutchinson, C., *et al.*, 1978, J. Biol. Chem. 253: 6551; Zoller y Smith, 1984, DNA 3: 479-488; Oliphant *et al.*, 1986, Gen 44: 177; Hutchinson *et al.*, 1986, Proc. Natl. Acad. Sci. U.S.A. 83: 710), uso de ligadores TAB® (Pharmacia), digestión por endonucleasas de restricción/delección de fragmentos y sustitución, mutagénesis dirigida por oligonucleótidos/mediada por PCR, y similares.
- 15 técnicas basadas en PCR para la mutagénesis dirigida al sitio (véase Higuchi, 1989, “Using PCR to Engineer DNA”, en PCR Technology: Principles and Applications for DNA Amplification, H. Erlich, ed., Stockton Press, capítulo 6, págs. 61-70).
- 20 “Fragmento” de un polipéptido según la invención se entenderá que significa un polipéptido cuya secuencia de aminoácidos es más corta que la del polipéptido de referencia y que comprende, a lo largo de toda la parte con estos polipéptidos de referencia, una secuencia de aminoácidos idéntica. Tales fragmentos pueden incluirse, cuando sea apropiado, en un polipéptido más largo del que forman una parte. Tales fragmentos de un polipéptido según la invención pueden tener una longitud de al menos 2, 3, 4, 5, 6, 8, 10, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 25, 26, 30, 35, 40, 45, 50, 100, 200, 240 ó 300 aminoácidos.
- 25 Una “variante” de un polipéptido o proteína es cualquier análogo, fragmento, derivado o mutante que se deriva de un polipéptido o proteína y que conserva al menos una propiedad biológica del polipéptido o proteína. Pueden existir en la naturaleza diferentes variantes del polipéptido o proteína. Estas variantes pueden ser variaciones alélicas caracterizadas por diferencias en las secuencias de nucleótidos del gen estructural que codifica para la proteína, o pueden implicar modificación postraduccional o por corte y empalme diferencial. El experto puede producir variantes que tienen sustituciones, deleciones, adiciones o reemplazos de un único o de múltiples aminoácidos. Estas variantes pueden incluir, entre otros: (a) variantes en las que uno o más residuos de aminoácido se sustituyen con aminoácidos conservativos o no conservativos, (b) variantes en las que se añaden uno o más aminoácidos al polipéptido o proteína, (c) variantes en las que uno o más de los aminoácidos incluye un grupo sustituyente y (d) variantes en las que el polipéptido o proteína se fusiona con otro polipéptido tal como albumina sérica. Los expertos habituales en la técnica conocen técnicas para obtener estas variantes, incluyendo técnicas genéticas (supresiones, deleciones, mutaciones, etc.), químicas y enzimáticas. Un polipéptido variante comprende preferiblemente al menos aproximadamente 14 aminoácidos.
- 30 Una “proteína heteróloga” se refiere a una proteína no producida de manera natural en la célula.
- Una “proteína madura” se refiere a un polipéptido procesado postraduccionalmente; es decir, uno del que se ha retirado cualquier pre o propéptido presente en el producto de traducción primario. Proteína “precursora” se refiere al producto de traducción primario de ARNm; es decir, con pre y propéptidos todavía presentes. Pre y propéptidos pueden ser pero no se limitan a señales de localización intracelular.
- 45 El término “péptido señal” se refiere a un polipéptido amino terminal que precede a la proteína madura secretada. El péptido señal se escinde de y por tanto no está presente en la proteína madura. Los péptidos señal tienen la función de dirigir y translocar proteínas secretadas a través de las membranas celulares. El péptido señal también se denomina proteína señal.
- 50 Se incluye una “secuencia señal” al comienzo de la secuencia codificante de una proteína que va a expresarse en la superficie de una célula. Esta secuencia codifica para un péptido señal, en posición N-terminal con respecto al polipéptido maduro, que dirige la célula huésped para translocar el polipéptido. El término “secuencia señal de translocación” se usa en el presente documento para referirse a esta clase de secuencia señal. Pueden encontrarse secuencias señal de translocación asociadas con una variedad de proteínas nativas para eucariotas y procariotas, y son a menudo funcionales en ambos tipos de organismos.
- 55 El término “homología,” se refiere al porcentaje de identidad entre dos restos de polinucleótido o dos de polipéptido. La correspondencia entre la secuencia de un resto con otro puede determinarse mediante técnicas conocidas en la técnica. Por ejemplo, puede determinarse la homología mediante una comparación directa de la información de secuencia entre dos moléculas de polipéptido alineando la información de secuencia y usando programas informáticos fácilmente disponibles. Alternativamente, puede determinarse la homología mediante hibridación de polinucleótidos en condiciones que forman dúplex estables entre regiones homólogas, seguido por digestión con nucleasa(s) específica (s) monocatenaria(s) y determinación del tamaño de los fragmentos digeridos.
- 60
- 65

Tal como se usa en el presente documento, el término “homólogo” en todas sus formas gramaticales y variaciones ortográficas se refiere a la relación entre proteínas que tienen un “origen evolutivo común”, incluyendo proteínas de superfamilias (por ejemplo, la superfamilia de las inmunoglobulinas) y proteínas homólogas de diferentes especies (por ejemplo, cadena ligera de la miosina, etc.) (Reeck *et al.*, 1987, Cell 50: 667). Tales proteínas (y sus genes codificantes) tienen homología de secuencia, tal como se refleja por su alto grado de similitud de secuencia. Sin embargo, en su uso común y en la presente solicitud, el término “homólogo,” cuando se modifica con un adverbio tal como “altamente”, puede referirse a similitud de secuencia y no un origen evolutivo común.

Por consiguiente, el término “similitud de secuencia” en todas las formas gramaticales se refiere al grado de identidad o correspondencia entre secuencias de ácido nucleico o aminoácidos de proteínas que pueden compartir o no un origen evolutivo común (véase Reeck *et al.*, 1987, Cell 50:667).

En una realización específica, dos secuencias de ADN son “sustancialmente homólogas” o “sustancialmente similares” cuando al menos aproximadamente el 50% (preferiblemente al menos aproximadamente el 75%, y lo más preferiblemente al menos aproximadamente el 90 o el 95%) de los nucleótidos coinciden a lo largo la longitud definida de las secuencias de ADN. Pueden identificarse secuencias que son sustancialmente homólogas comparando las secuencias usando software convencional disponible en bancos de datos de secuencias, o en un experimento de hibridación Southern, por ejemplo, en condiciones rigurosas tal como se define para ese sistema particular. La definición de condiciones de hibridación apropiadas está dentro de la experiencia de la técnica. Véase, por ejemplo, Sambrook *et al.*, 1989, citado anteriormente.

Tal como se usa en el presente documento, “sustancialmente similar” se refiere a fragmentos de ácido nucleico en los que los cambios en una o más bases de nucleótidos dan como resultado la sustitución de uno o más aminoácidos, pero no afectan a las propiedades funcionales de la proteína codificada por la secuencia de ADN. “Sustancialmente similar” también se refiere a fragmentos de ácido nucleico en los que los cambios en una o más bases de nucleótidos no afectan a la capacidad del fragmento de ácido nucleico para mediar en la alteración de la expresión génica mediante tecnología antisentido o co-supresión. “Sustancialmente similar” también se refiere a modificaciones de los fragmentos de ácido nucleico de la descripción tal como delección o inserción de una o más bases de nucleótidos que no afectan sustancialmente a las propiedades funcionales del transcrito resultante. Por tanto, se entiende que la descripción abarca más que las secuencias a modo de ejemplo específicas. Cada una de las modificaciones propuestas se encuentra bien dentro de la experiencia de rutina en la técnica, así como la determinación de la conservación de la actividad biológica de los productos codificados.

Además, el experto reconoce que también se definen secuencias sustancialmente similares abarcadas por esta descripción por su capacidad para hibridarse, en condiciones rigurosas (0,1X SSC, SDS al 0,1%, 65°C y lavado con 2X SSC, SDS al 0,1% seguido por 0,1X SSC, SDS al 0,1%), con las secuencias mostradas a modo de ejemplo en el presente documento. Fragmentos de ácido nucleico sustancialmente similares de la descripción son los fragmentos de ácido nucleico cuyas secuencias de ADN son al menos el 70% idénticas a la secuencia de ADN de los fragmentos de ácido nucleico notificados en el presente documento. Fragmentos de ácido nucleico sustancialmente preferidos de la descripción son los fragmentos de ácido nucleico cuyas secuencias de ADN son al menos el 80% idénticas a la secuencia de ADN de los fragmentos de ácido nucleico notificados en el presente documento. Fragmentos de ácido nucleico más preferidos son al menos el 90% idénticos a la secuencia de ADN de los fragmentos de ácido nucleico notificados en el presente documento. Se prefieren incluso más fragmentos de ácido nucleico que son al menos el 95% idénticos a la secuencia de ADN de los fragmentos de ácido nucleico notificados en el presente documento.

Dos secuencias de aminoácido son “sustancialmente homólogas” o “sustancialmente similares” cuando más de aproximadamente el 40% de los aminoácidos son idénticos, o más del 60% son similares (funcionalmente idénticos). Preferiblemente, se identifican secuencias similares u homólogas mediante alineación usando, por ejemplo, el programa Pileup GCG (Genetics Computer Group, Program Manual for the GCG Package, versión 7, Madison, Wisconsin).

El término “correspondiente a” se usa en el presente documento para referirse a secuencias similares u homólogas, ya sea la posición exacta idéntica o diferente de la molécula con la que se mide la similitud u homología. Una alineación de secuencia de ácido nucleico o aminoácidos puede incluir espacios. Por tanto, el término “correspondiente a” se refiere a la similitud de secuencia, y no a la numeración de los residuos de aminoácido o bases de nucleótidos.

Una “parte sustancial” de un secuencia de nucleótidos o aminoácidos comprende lo suficiente de la secuencia de aminoácidos de un polipéptido o la secuencia de nucleótidos de un gen para identificar supuestamente ese polipéptido o gen, o bien mediante evaluación manual de la secuencia por un experto en la técnica, o bien mediante comparación de secuencias automatizada por ordenador e identificación usando algoritmos tales como BLAST (Basic Local Alignment Search Tool; Altschul, S. F., *et al.*, (1993) J. Mol. Biol. 215: 403-410; véase también [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)). En general, se necesita una secuencia de diez o más aminoácidos contiguos o treinta o más nucleótidos con el fin de identificar supuestamente una secuencia polipeptídica o de ácido nucleico como homóloga con respecto a una proteína o gen conocido. Además, con respecto a secuencias de nucleótidos,

pueden usarse sondas oligonucleotídicas específicas de gen que comprenden 20-30 nucleótidos contiguos en métodos dependientes de secuencia de identificación (por ejemplo, hibridación Southern) y aislamiento (por ejemplo, hibridación *in situ* de colonias bacterianas o placas de bacteriófagos) de genes. Además, pueden usarse oligonucleótidos cortos de 12-15 bases como cebadores de amplificación en PCR con el fin de obtener un fragmento de ácido nucleico particular que comprende los cebadores. Por consiguiente, una “parte sustancial” de una secuencia de nucleótidos comprende lo suficiente de la secuencia para identificar y/o aislar específicamente un fragmento de ácido nucleico que comprende la secuencia.

El término “porcentaje de identidad”, tal como se conoce en la técnica, es una relación entre dos o más secuencias polipeptídicas o dos o más secuencias de polinucleótido, tal como se determina comparando las secuencias. En la técnica, “identidad” también significa el grado de relación de secuencia entre secuencias polipeptídicas o de polinucleótido, según sea el caso, tal como se determina por la coincidencia entre cadenas de tales secuencias. La “identidad” y “similitud” pueden calcularse fácilmente mediante métodos conocidos, incluyendo pero sin limitarse a los descritos en: Computational Molecular Biology (Lesk, A. M., ed.) Oxford University Press, Nueva York (1988); Biocomputing: Informatics and Genome Projects (Smith, D. W., ed.) Academic Press, Nueva York (1993); Computer Analysis of Sequence Data, parte I (Griffin, A. M., y Griffin, H. G., eds.) Humana Press, Nueva Jersey (1994); Sequence Analysis in Molecular Biology (von Heinje, G., ed.) Academic Press (1987); y Sequence Analysis Primer (Gibbskov, M. y Devereux, J., eds.) Stockton Press, Nueva York (1991). Se diseñan métodos preferidos para determinar la identidad para proporcionar la mejor coincidencia entre las secuencias sometidas a prueba. Los métodos para determinar la identidad y similitud están codificados en programas informáticos disponibles para el público. Pueden realizarse alineaciones de secuencias y cálculos del porcentaje de identidad usando el programa Megalign del conjunto de análisis bioinformático LASERGENE (DNASTAR Inc., Madison, WI). Pueden realizarse alineaciones múltiples de las secuencias usando el método de alineación Clustal (Higgins y Sharp (1989) CABIOS. 5:151-153) con los parámetros por defecto (PENALIZACIÓN POR HUECO=10, PENALIZACIÓN POR LONGITUD DE HUECO=10). Pueden seleccionarse parámetros por defecto para alineaciones por parejas usando el método Clustal: KTUPL0 1, PENALIZACIÓN POR HUECO=3, VENTANA= 5 y DIAGONALES GUARDADAS=5.

El término “software de análisis de secuencias” se refiere a cualquier algoritmo informático o programa de software que es útil para el análisis de secuencias de nucleótidos o aminoácidos. El “software de análisis de secuencias” puede desarrollarse independientemente o está disponible comercialmente. El software de análisis de secuencias típico incluirá pero no se limita al conjunto de programas GCG (Wisconsin Package versión 9.0, Genetics Computer Group (GCG), Madison, WI), BLASTP, BLASTN, BLASTX (Altschul *et al.*, J. Mol. Biol. 215: 403-410 (1990) y DNASTAR (DNASTAR, Inc. 1228 S. Park St. Madison, WI 53715 EE.UU.). Dentro del contexto de esta solicitud se entenderá que cuando el software de análisis de secuencias se usa para el análisis, los resultados del análisis se basarán en los “valores por defecto” del programa al que se hace referencia, a menos que se especifique lo contrario. Tal como se usa en el presente documento, “valores por defecto” significará cualquier conjunto de valores o parámetros que se cargan de manera original con el software cuando se inicia por primera vez.

Pueden ensamblarse “genes sintéticos” a partir de bloques de construcción de oligonucleótidos que se sintetizan químicamente usando procedimientos conocidos por los expertos en la técnica. Estos bloques de construcción se ligan y se aparean para formar segmentos génicos que entonces se ensamblan de manera enzimática para construir el gen entero. “Sintetizado químicamente”, en relación con una secuencia de ADN, significa que los nucleótidos componentes se ensamblaron *in vitro*. Puede lograrse la síntesis química manual de ADN usando procedimientos bien establecidos, o puede realizarse síntesis química automatizada usando una de varias máquinas disponibles comercialmente. Por consiguiente, los genes pueden adaptarse para expresión génica óptima basándose en la optimización de la secuencia de nucleótidos para reflejar el sesgo de codones de la célula huésped. El experto aprecia la probabilidad de expresión génica satisfactoria si el uso de codones experimenta un sesgo hacia los codones favorecidos por el huésped. La determinación de codones preferidos puede basarse en un estudio de genes derivados de la célula huésped en la que está disponible información de secuencias.

Tal como se usa en el presente documento, se dice que dos o más sistemas de regulación génica que pueden funcionar individualmente son “ortogonales” cuando; a) la modulación de cada uno de los sistemas dados por su respectivo ligando, a una concentración elegida, da como resultado un cambio medible en la magnitud de la expresión del gen de este sistema, y b) el cambio es diferente de manera estadísticamente significativa que el cambio en la expresión de todos los demás sistemas que pueden funcionar simultáneamente en la célula, tejido u organismo, independientemente de la simultaneidad o secuencialidad de la modulación real. Preferiblemente, la modulación de cada sistema de regulación génica que puede funcionar individualmente efectúa un cambio en la expresión génica al menos 2 veces mayor que todos los demás sistemas que pueden funcionar en la célula, tejido u organismo. Más preferiblemente, el cambio es al menos 5 veces mayor. Incluso más preferiblemente, el cambio es al menos 10 veces mayor. Todavía más preferiblemente, el cambio es al menos 100 veces mayor. Incluso todavía más preferiblemente, el cambio es al menos 500 veces mayor. Idealmente, la modulación de cada uno de los sistemas dados por su respectivo ligando a una concentración elegida da como resultado un cambio medible en la magnitud de la expresión del gen de ese sistema y ningún cambio medible en la expresión de todos los demás sistemas que pueden funcionar en la célula, tejido u organismo. En tales casos, se dice que el sistema de regulación génica inducible múltiple es “completamente ortogonal”. La presente descripción es útil para buscar ligandos ortogonales y sistemas de expresión génica basados en receptores ortogonales tales como los descritos en la

solicitud estadounidense en tramitación junto con la presente 60/237446.

Sistema de modulación de la expresión génica de la invención

5 Los solicitantes han identificado en el presente documento residuos de aminoácido que están implicados en la unión de ligando a un dominio de unión a ligando de receptor nuclear de grupo H que afecta a la sensibilidad a ligando y la magnitud de la inducción en un sistema de expresión génica inducible basado en receptor de ecdisona. Los solicitantes describen en el presente documento la construcción de receptores nucleares de grupo H que comprenden mutaciones por sustitución (denominado en el presente documento “mutantes por sustituciones”) a  
 10 estos residuos críticos y la demostración de que estos receptores nucleares mutantes por sustitución son útiles en métodos para modular la expresión génica. Tal como se presenta en el presente documento, los receptores nucleares mutantes por sustitución novedosos de los solicitantes y su uso en un sistema de expresión génica inducible basado en receptores nucleares proporciona un sistema de expresión génica inducible mejorado tanto en las células huésped procariontas como eucariotas en las que la sensibilidad a ligando y la magnitud de la transactivación puede seleccionarse según se desea, dependiendo de la aplicación.  
 15

Por tanto, la presente invención se refiere a polinucleótidos y polipéptidos receptores nucleares de grupo H mutantes por sustitución novedosos, un sistema de expresión génica inducible basado en receptores nucleares que comprende tales polinucleótidos y polipéptidos del receptor nuclear de grupo H mutado, y métodos para modular la expresión de un gen dentro de una célula huésped usando un sistema de expresión génica inducible basado en receptores nucleares de ese tipo.  
 20

En particular, la presente invención se refiere a un sistema de modulación de la expresión génica que comprende al menos un casete de expresión génica que puede expresarse en una célula huésped que comprende un polinucleótido que codifica para un polipéptido que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución. Preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución procede de un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína 15 de interacción de receptor retinoide X, un receptor hepático X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor y un receptor de farnesol. Más preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución procede de un receptor de ecdisona.  
 25  
 30

En una realización específica, el sistema de modulación de la expresión génica comprende un casete de expresión génica que comprende un polinucleótido que codifica para un polipéptido que comprende un dominio de transactivación, un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión ha de modularse; y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución. El sistema de modulación de la expresión génica puede comprender además un segundo casete de expresión génica que comprende: i) un elemento de respuesta reconocido por el dominio de unión a ADN del polipéptido codificado del primer casete de expresión génica; ii) un promotor que se activa por el dominio de transactivación del polipéptido codificado del primer casete de expresión génica; y iii) un gen cuya expresión ha de modularse.  
 35  
 40

En otra realización específica, el sistema de modulación de la expresión génica comprende un casete de expresión génica que comprende a) un polinucleótido que codifica para un polipéptido que comprende un dominio de transactivación, un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión ha de modularse; y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución y b) un segundo dominio de unión a ligando de receptor nuclear seleccionado del grupo que consiste en un dominio de unión a ligando de receptor retinoide X de vertebrado, un dominio de unión a ligando de receptor retinoide X de invertebrado, un dominio de unión a ligando de proteína ultraespiráculo, y un dominio de unión a ligando quimérico que comprende dos fragmentos de polipéptido, en el que el primer fragmento de polipéptido procede de un dominio de unión a ligando de receptor retinoide X de vertebrado, un dominio de unión a ligando de receptor retinoide X de invertebrado, o un dominio de unión a ligando de proteína ultraespiráculo, y el segundo fragmento de polipéptido procede de un dominio de unión a ligando de receptor retinoide X de vertebrado, dominio de unión a ligando de receptor retinoide X de invertebrado, o dominio de unión a ligando de proteína ultraespiráculo diferente. El sistema de modulación de la expresión génica puede comprender además un segundo casete de expresión génica que comprende: i) un elemento de respuesta reconocido por el dominio de unión a ADN del polipéptido codificado del primer casete de expresión génica; ii) un promotor que se activa por el dominio de transactivación del polipéptido codificado del primer casete de expresión génica; y iii) un gen cuya expresión ha de modularse.  
 45  
 50  
 55  
 60

En otra realización específica, el sistema de modulación de la expresión génica comprende un primer casete de expresión génica que comprende un polinucleótido que codifica para un primer polipéptido que comprende un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión ha de modularse y un dominio de unión a ligando de receptor nuclear, y un segundo casete de expresión génica que comprende un polinucleótido que codifica para un segundo polipéptido que comprende un dominio de  
 65

transactivación y un dominio de unión a ligando de receptor nuclear, en el que uno de los dominios de unión a ligando de receptor nuclear es un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución. En una realización preferida, el primer polipéptido está sustancialmente libre de un dominio de transactivación y el segundo polipéptido está sustancialmente libre de un dominio de unión a ADN. Para fines de la invención, “sustancialmente libre” significa que la proteína en cuestión no contiene una secuencia suficiente del dominio en cuestión para proporcionar la activación o actividad de unión. El sistema de modulación de la expresión génica puede comprender además un tercer casete de expresión génica que comprende: i) un elemento de respuesta reconocido por el dominio de unión a ADN del primer polipéptido del primer casete de expresión génica; ii) un promotor que se activa por el dominio de transactivación del segundo polipéptido del segundo casete de expresión génica; y iii) un gen cuya expresión ha de modularse.

Cuando sólo un dominio de unión a ligando de receptor nuclear es un dominio de unión a ligando de grupo H que comprende una mutación por sustitución, el otro dominio de unión a ligando de receptor nuclear puede ser de cualquier otro receptor nuclear que forma un dímero con el dominio de unión a ligando de grupo H que comprende la mutación por sustitución. Por ejemplo, cuando el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución, el otro dominio de unión a ligando de receptor nuclear (“pareja”) puede ser de un receptor de ecdisona, un receptor retinoide X (RXR) de vertebrado, un RXR de invertebrado, una proteína ultraespiráculo (USP), o un receptor nuclear quimérico que comprende al menos dos fragmentos de polipéptido de dominio de unión a ligando de receptor nuclear diferentes seleccionados del grupo que consiste en un RXR de vertebrado, un RXR de invertebrado, y una USP (véanse las solicitudes de patente en tramitación junto con la presente PCT/US01/09050, US 60/294814 y US 60/294819). El dominio de unión a ligando de receptor nuclear de “pareja” puede comprender además una mutación por truncamiento, una mutación por delección, una mutación por sustitución u otra modificación.

Preferiblemente, el dominio de unión a ligando de RXR de vertebrado procede de un RXR de ser humano *Homo sapiens*, ratón *Mus musculus*, rata *Rattus norvegicus*, pollo *Gallus gallus*, cerdo *Sus scrofa domestica*, rana *Xenopus laevis*, pez cebra *Danio rerio*, tunicado *Polyandrocarpa misakiensis* o medusa *Tripedalia cysophara*.

Preferiblemente, el dominio de unión a ligando de RXR de invertebrado procede de un polipéptido de ultraespiráculo de langosta *Locusta migratoria* (“LmUSP”), un homólogo de RXR de garrapata solitaria *Amblyomma americanum* 1 (“AmaRXR1”), un homólogo de RXR de garrapata solitaria *Amblyomma americanum* 2 (“AmaRXR2”), un homólogo de RXR del cangrejo violinista *Celuca pugilator* (“CpRXR”), un homólogo de RXR del escarabajo *Tenebrio molitor* (“TmRXR”), un homólogo de RXR de la abeja *Apis mellifera* (“AmRXR”), un homólogo de RXR del áfido *Myzus persicae* (“MpRXR”), o un homólogo de RXR de no dípteros/no lepidópteros.

Preferiblemente, el dominio de unión a ligando de RXR quimérico comprende al menos dos fragmentos de polipéptido seleccionados del grupo que consiste en un fragmento de polipéptido de RXR de especie de vertebrado, un fragmento de polipéptido de RXR de especie de invertebrado, y un fragmento de polipéptido de homología de RXR de especie de invertebrado de no dípteros/no lepidópteros. Un dominio de unión a ligando de RXR quimérico para su uso en la presente invención puede comprender al menos dos fragmentos de polipéptido de RXR de especies diferentes, o cuando la especie es la misma, los dos o más fragmentos de polipéptido pueden ser de dos o más isoformas diferentes del fragmento de polipéptido de RXR de la especie.

En una realización preferida, el dominio de unión a ligando de RXR quimérico comprende al menos un fragmento de polipéptido de RXR de especie de vertebrado y un fragmento de polipéptido de RXR de especie de invertebrado.

En una realización más preferida, el dominio de unión a ligando de RXR quimérico comprende al menos un fragmento de polipéptido de RXR de especie de vertebrado y un fragmento de polipéptido homólogo de RXR de especie de invertebrado no dípteros/no lepidópteros.

En una realización específica, el gen cuya expresión ha de modularse es un gen homólogo con respecto a la célula huésped. En otra realización específica, el gen cuya expresión ha de modularse es un gen heterólogo con respecto a la célula huésped.

Los ligandos para su uso en la presente invención tal como se da a conocer a continuación, cuando se combinan con el dominio de unión a ligando del/de los receptor(es) nuclear(es), que a su vez se unen al elemento de respuesta unido a un gen, proporcionan el medio para la regulación de expresión temporal externa del gen. El mecanismo de unión o el orden en el que los diversos componentes de esta invención se unen entre sí, es decir, por ejemplo, el ligando al dominio de unión a ligando, dominio de unión a ADN al elemento de respuesta, dominio de transactivación al promotor, etc., no es crítico.

En un ejemplo específico, la unión del ligando al dominio de unión a ligando de un receptor nuclear de grupo H y su pareja de dominio de unión a ligando de receptor nuclear permite la expresión o supresión del gen. Este mecanismo no excluye la posibilidad de que el ligando se una al receptor nuclear de grupo H (GHNR) o su pareja, y la formación resultante de complejos de homodímero activos (por ejemplo GHNR + GHNR o pareja+pareja). Preferiblemente, uno

- o más de los dominios de receptor se varía produciendo un interruptor génico híbrido. Normalmente, uno o más de los tres dominios, DBD, LBD y el dominio de transactivación, puede elegirse de una fuente diferente de la fuente de los demás dominios de modo que los genes híbridos y las proteínas híbridas resultantes se optimizan en la célula huésped u organismo elegido para la actividad de transactivación, unión complementaria del ligando y reconocimiento de un elemento de respuesta específico. Además, el propio elemento de respuesta puede modificarse o sustituirse con elementos de respuesta para otros dominios de proteína de unión a ADN tales como la proteína GAL-4 de levadura (véase Sadowski, *et al.* (1988), *Nature*, 335: 563-564) o proteína LexA de *Escherichia coli* (véase Brent y Ptashne (1985), *Cell*, 43: 729-736), o elementos de respuesta sintéticos específicos para las interacciones dirigidas con proteínas diseñadas, modificadas y seleccionadas para tales interacciones específicas (véase, por ejemplo, Kim, *et al.* (1997), *Proc. Natl. Acad. Sci., USA*, 94:3 616-3620) para acomodar los receptores híbridos. Otra ventaja de sistemas híbridos dobles es que permiten la elección de un promotor usado para dirigir la expresión génica según un resultado final deseado. Tal doble control puede ser particularmente importante en áreas de terapia génica, especialmente cuando se producen proteínas citotóxicas, debido a que puede controlarse tanto la temporización de la expresión así como las células en las que se produce la expresión. Cuando se introducen genes, operativamente unidos a un promotor adecuado, en las células del sujeto, la expresión de los genes exógenos se controla por la presencia del sistema de esta invención. Los promotores pueden regularse de manera constitutiva o inducible o pueden ser específicos de tejidos (es decir, expresados sólo en un tipo particular de células) o específicos para determinadas fases de desarrollo del organismo.
- El receptor de ecdisona es un miembro de la superfamilia de receptor nuclear y se clasifica en la subfamilia 1, grupo H (denominado en el presente documento "receptores nucleares de grupo H"). Los miembros de cada grupo comparten el 40-60% de identidad de aminoácidos en el dominio E (unión a ligando) (Laudet *et al.*, A Unified Nomenclature System for the Nuclear Receptor Subfamily, 1999; *Cell* 97: 161-163). Además del receptor de ecdisona, otros miembros de esta subfamilia de receptor nuclear 1, grupo H incluyen: receptor ubicuo (UR), receptor huérfano 1 (OR-1), receptor nuclear de hormonas esteroideas 1 (NER-1), proteína 15 de interacción de receptor retinoide X (RIP-15), receptor hepático X  $\beta$  (LXR $\beta$ ), proteína similar al receptor de hormonas esteroideas (RLD-1), receptor hepático X (LXR), receptor hepático X  $\alpha$  (LXR $\alpha$ ), receptor farnesoide X (FXR), proteína 14 de interacción de receptor (RIP-14) y receptor de farnesol (HRR-1).
- Los solicitantes han desarrollado un modelo de homología CfEcR y han usado este modelo de homología junto con un modelo de homología de receptor de ecdisona de *Chironomus tentans* ("CtEcR") publicado (Wurtz *et al.*, 2000) para identificar residuos críticos implicados en la unión a esteroides y compuestos no esteroideos. Se ha mostrado que los compuestos no esteroideos sintéticos, diacilhidrazinas, se unen a EcR de lepidópteros con alta afinidad e inducen muda incompleta precoz en estos insectos (Wing *et al.*, 1988) y varios de estos compuestos se comercializan actualmente como insecticidas. La cavidad de unión a ligando o "bolsa" de EcRs ha evolucionado para adaptarse a las estructuras de esqueleto largas de ecdiesteroides tales como 20-hidroxiecdisona (20E). Las diacilhidrazinas tienen una estructura compacta en comparación con esteroides y sólo ocupan la parte inferior de la bolsa de unión a EcR. Esto deja pocos residuos críticos en la parte superior de la bolsa de unión que entra en contacto con esteroides pero no con compuestos no esteroideos tales como bisacilhidrazinas. Los solicitantes describen en el presente documento la construcción de receptores de ecdisona mutantes que comprenden una mutación por sustitución en estos residuos de la bolsa de unión y han identificado diversas clases de receptores de ecdisona mutantes por sustitución con características de transactivación y unión a ligando modificado.
- Dada la estrecha relación del receptor de ecdisona con otros receptores nucleares de grupo H, también se espera que las mutaciones por sustitución del dominio de unión a ligando de receptor de ecdisona identificadas por los solicitantes funcionen cuando se introduzcan en la posición análoga de los dominios de unión a ligando de otros receptores nucleares de grupo H para modificar su unión a ligando o sensibilidad a ligando. Los polinucleótidos y polipéptidos del receptor nuclear de grupo H mutado por sustitución novedoso de los solicitantes son útiles en un sistema de modulación génica inducible basado en receptores nucleares para diversas aplicaciones incluyendo terapia génica, expresión de proteínas de interés en células huésped, producción de organismos transgénicos y ensayos basados en células.
- En particular, los solicitantes describen en el presente documento un sistema de modulación de la expresión génica novedoso que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución. Este sistema de expresión génica puede ser un sistema de expresión génica basado en "interruptor único" en el que el dominio de transactivación, dominio de unión a ADN y dominio de unión a ligando están en un polipéptido codificado. Alternativamente, el sistema de modulación de la expresión génica puede ser un sistema de modulación de la expresión génica basado en "interruptor doble" o "doble híbrido" en el que el dominio de transactivación y dominio de unión a ADN se ubican en dos polipéptidos codificados diferentes. Los solicitantes han demostrado por primera vez que puede usarse un receptor nuclear mutado por sustitución como un componente de un sistema de expresión génica inducible basado en receptores nucleares para modificar la actividad de unión a ligando y/o especificidad de ligando tanto en células procariontas como eucariotas. Tal como se discute en el presente documento, los hallazgos de los solicitantes son tanto inesperados como sorprendentes.
- Un sistema de modulación de la expresión génica basado en receptor de ecdisona de la presente invención puede ser o bien heterodimérico o bien homodimérico. Un complejo de EcR funcional en general se refiere a un complejo

de proteína heterodimérica que consiste en dos miembros de la familia de receptor de esteroides, una proteína de receptor de ecdisona obtenida a partir de diversos insectos y una proteína ultraespiráculo (USP) o el homólogo de USP de vertebrado, proteína de receptor retinoide X (véase Yao, *et al.* (1993) *Nature* 366: 476-479; Yao, *et al.*, (1992) *Cell* 71: 63-72). Sin embargo, el complejo también puede ser un homodímero tal como se detalla a continuación. El complejo de receptor de ecdiesteroides funcional también puede incluir proteína(s) adicional(es) tal como inmunofilinas. Los miembros adicionales de la familia de proteínas de receptor de esteroides, conocidos como factores de la transcripción (tal como DHR38 o betaFTZ-1), también pueden ser parejas dependientes o independientes de ligando para EcR, USP, y/o RXR. Adicionalmente, pueden requerirse otros cofactores tal como proteínas conocidas generalmente como coactivadores (también denominados adaptadores o mediadores). Estas proteínas no se unen de manera específica de secuencia a ADN y no están implicadas en la transcripción basal. Pueden ejercer su efecto sobre la activación por transcripción a través de diversos mecanismos, incluyendo estimulación de unión a ADN de activadores, afectando a la estructura de cromatina, o mediando las interacciones de complejo de iniciación-activador. Los ejemplos de tales coactivadores incluyen RIP140, TIF1, RAP46/Bag-1, ARA70, SRC-1/NCoA-1, TIF2/GRIP/NCoA-2, ACTR/AIB1/RAC3/pCIP así como la proteína B de unión al elemento de respuesta al coactivador C promiscuo, CBP/p300 (para revisión véase Glass *et al.*, *Curr. Opin. Cell Biol.* 9: 222-232, 1997). Además, los cofactores de proteínas conocidos generalmente como correpresores (también conocidos como represores, silenciadores o mediadores de silenciamiento) pueden requerirse para inhibir de manera eficaz la activación de la transcripción en ausencia de ligando. Estos correpresores pueden interactuar con el receptor de ecdisona sin ligando para silenciar la actividad al elemento de respuesta. Las pruebas actuales sugieren que la unión de ligando cambia la conformación del receptor, lo que da como resultado la liberación del correpresor y reclutamiento de los coactivadores descritos anteriormente, eliminando de ese modo su actividad de silenciamiento. Los ejemplos de correpresores incluyen N-CoR y SMRT (para revisión, véase Horwitz *et al.* *Mol Endocrinol.* 10: 1167-1177, 1996). Estos cofactores pueden ser o bien endógenos dentro de la célula u organismo, o pueden añadirse de manera exógena como transgenes que van a expresarse de modo o bien regulado o bien no regulado. Los complejos de homodímero de la proteína de receptor de ecdisona, USP o RXR también pueden ser funcionales en algunas circunstancias.

El complejo de receptor de ecdisona normalmente incluye proteínas que son miembros de la superfamilia de receptor nuclear en la que todos los miembros se caracterizan en general por la presencia de un dominio de transactivación amino-terminal, un dominio de unión a ADN ("DBD") y un dominio de unión a ligando ("LBD") separado del DBD mediante una región bisagra. Tal como se usa en el presente documento, el término "dominio de unión a ADN" comprende una secuencia de polipéptidos mínima de una proteína de unión a ADN, hasta la longitud entera de una proteína de unión a ADN, siempre que el dominio de unión a ADN funcione para asociarse con un elemento de respuesta particular. Los miembros de la superfamilia de receptores nucleares también se caracterizan por la presencia de cuatro o cinco dominios: A/B, C, D, E, y en algunos miembros F (véase la patente de EE.UU. 4.981.784 y Evans, *Science* 240:889-895 (1988)). El dominio "A/B" corresponde al dominio de transactivación, "C" corresponde al dominio de unión a ADN, "D" corresponde a la región bisagra y "E" corresponde al dominio de unión a ligando. Algunos miembros de la familia también pueden tener otro dominio de transactivación en el lado carboxi-terminal del LBD correspondiente a "F".

El DBD se caracteriza por la presencia de dos dedos de zinc cisteína entre los que están dos motivos de aminoácido, la caja P y la caja D, que confieren especificidad para el elemento de respuesta a ecdisonas. Estos dominios pueden ser o bien nativos o bien modificados o bien quimeras de diferentes dominios de proteínas de receptores heterólogos. El receptor EcR, como un subconjunto de la familia de receptor de esteroides, también tiene regiones menos definidas responsables de las propiedades de heterodimerización. Debido a que los dominios de receptores nucleares son modulares en la naturaleza, los dominios de LBD, DBD y de transactivación pueden intercambiarse.

Se sabe que los sistemas de interruptor génico incorporan los componentes del complejo de receptor de ecdisona. Sin embargo, en estos sistemas conocidos, siempre que se usa EcR se asocia con dominios de unión a ADN nativos o modificados y dominios de transactivación en la misma molécula. USP o RXR se usan normalmente como parejas silenciosas. Los solicitantes han mostrado previamente que cuando los dominios de unión a ADN y los dominios de transactivación están en la misma molécula la actividad de fondo en ausencia de ligando es alta y que tal actividad se reduce drásticamente cuando los dominios de unión a ADN y los dominios de transactivación están en diferentes moléculas, es decir, en cada una de dos parejas de un complejo heterodimérico u homodimérico (véase el documento PCT/US01/09050).

#### Casetes de expresión génica de la invención

El sistema de expresión génica inducible basado en receptores nucleares novedoso de la invención comprende al menos un casete de expresión génica que puede expresarse en una célula huésped, en el que el casete de expresión génica comprende un polinucleótido que codifica para un polipéptido que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución. Por tanto, la invención de los solicitantes también proporciona casetes de expresión génica novedosos para su uso en el sistema de expresión génica de la invención.



En una realización específica, el casete de expresión génica que puede expresarse en una célula huésped comprende un polinucleótido que codifica para un polipéptido seleccionado del grupo que consiste en a) un polipéptido que comprende un dominio de transactivación, un dominio de unión a ADN y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución; b) un polipéptido que comprende un dominio de unión a ADN y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución; y c) un polipéptido que comprende un dominio de transactivación y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución.

En otra realización específica, la presente descripción proporciona un casete de expresión génica que puede expresarse en una célula huésped, comprendiendo el casete de expresión génica un polinucleótido que codifica para un polipéptido híbrido seleccionado del grupo que consiste en a) un polipéptido híbrido que comprende un dominio de transactivación, un dominio de unión a ADN y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución; b) un polipéptido híbrido que comprende un dominio de unión a ADN y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución; y c) un polipéptido híbrido que comprende un dominio de transactivación y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución. Un polipéptido híbrido según la descripción comprende al menos dos fragmentos de polipéptido, en los que cada fragmento de polipéptido procede de una fuente diferente, es decir, un polipéptido diferente, un receptor nuclear diferente, una especie diferente, etc. El polipéptido híbrido según la descripción puede comprender al menos dos dominios de polipéptido, en los que cada dominio de polipéptido procede de una fuente diferente.

En una realización específica, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución procede de un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína 15 de interacción de receptor retinoide X, un receptor hepático X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor y un receptor de farnesol. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H procede de un receptor de ecdisona.

Por tanto, la presente descripción también proporciona un casete de expresión génica que comprende un polinucleótido que codifica para un polipéptido seleccionado del grupo que consiste en a) un polipéptido que comprende un dominio de transactivación, un dominio de unión a ADN y un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución; b) un polipéptido que comprende un dominio de unión a ADN y un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución; y c) un polipéptido que comprende un dominio de transactivación y un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución. Preferiblemente, el casete de expresión génica comprende un polinucleótido que codifica para un polipéptido híbrido seleccionado del grupo que consiste en a) un polipéptido híbrido que comprende un dominio de transactivación, un dominio de unión a ADN y un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución; b) un polipéptido híbrido que comprende un dominio de unión a ADN y un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución; y c) un polipéptido híbrido que comprende un dominio de transactivación y un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución; en el que el polipéptido híbrido codificado comprende al menos dos fragmentos de polipéptido, en los que cada fragmento de polipéptido procede de una fuente diferente.

El dominio de unión a ligando (LBD) de receptor de ecdisona (EcR) puede ser de un EcR de invertebrado, preferiblemente seleccionado de la clase de EcR de artrópodo. Preferiblemente el EcR se selecciona del grupo que consiste en un EcR de lepidópteros, un EcR de dípteros, un EcR de ortópteros, un EcR de homópteros y un EcR de hemípteros. Más preferiblemente, el dominio de unión a ligando de EcR para su uso en la presente invención es de un EcR de tortrix de las yemas de la picea *Choristoneura fumiferana* ("CfEcR"), un EcR de escarabajo *Tenebrio molitor* ("TmEcR"), un EcR de *Manduca sexta* ("MsEcR"), un EcR de *Heliethies virescens* ("HvEcR"), un EcR de mosquilla *Chironomus tentans* ("CtEcR"), un EcR de mariposa de la seda *Bombyx mori* ("BmEcR"), un EcR de mariposa *Bicyclus anynana* ("BanEcR"), un EcR de mariposa de color castaño *Junonia coenia* ("JcEcR"), un EcR de mosca de la fruta *Drosophila melanogaster* ("DmEcR"), un EcR de mosquito *Aedes aegypti* ("AaEcR"), un EcR de mosca azul *Lucilia capitata* ("palanca"), un EcR de mosca azul *Lucilia cuprina* ("LucEcR"), un EcR de mosca azul *Calliphora vicina* ("CvEcR"), un EcR de mosca de la fruta mediterránea *Ceratitis capitata* ("CcEcR"), un EcR de langosta *Locusta migratoria* ("LmEcR"), un EcR de áfido *Myzus persicae* ("MpEcR"), un EcR de cangrejo violinista *Celaca pugillator* ("CpEcR"), un EcR de garrapata solitaria *Amblyomma americanum* ("AmaEcR"), un EcR de mosca blanca *Bamecia argentifoli* ("BaBcR", SEQ ID NO: 112) o un EcR de saltahojas *Nephotetix cincticeps* ("NcEcR", SEQ ID NO: 113). Más preferiblemente, el LBD es de un CfEcR, un DmEcR o un AmaEcR.

En una realización específica, el LBD es de un polipéptido de EcR truncado. La truncación del polipéptido de EcR da como resultado una delección de al menos 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260 ó 265 aminoácidos. Preferiblemente, la truncación del polipéptido de EcR da como resultado una delección de al menos un dominio de polipéptido parcial. Más

preferiblemente, la truncación del polipéptido de EcR da como resultado una delección de al menos un dominio de polipéptido entero. En una realización específica, la truncación del polipéptido de EcR da como resultado una delección de al menos un dominio A/B, un dominio C, un dominio D, un dominio F, dominios A/B/C, dominios A/B/1/2-C, dominio A/B/C/D, dominios A/B/C/D/F, dominios A/B/F, dominios A/B/C/F, un dominio E parcial o un dominio F parcial. También puede realizarse una combinación de varias delecciones de dominio completo y/o parcial.

En una realización específica, el dominio de unión a ligando de receptor nuclear de grupo H se codifica por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de a) residuo de aminoácido 20, 21, 48, 51, 52, 55, 58, 59, 61, 62, 92, 93, 95, 96, 107, 109, 110, 120, 123, 125, 175, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) residuos de aminoácido 95 y 110 de SEQ ID NO: 1, c) residuos de aminoácido 218 y 219 de SEQ ID NO: 1, d) residuos de aminoácido 107 y 175 de SEQ ID NO: 1, e) residuos de aminoácido 127 y 175 de SEQ ID NO: 1, f) residuos de aminoácido 107 y 127 de SEQ ID NO: 1, g) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, h) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, i) residuos de aminoácido 96, 107 y 175 de SEQ ID NO: 1, j) residuos de aminoácido 107, 110 y 175 de SEQ ID NO: 1, k) residuo de aminoácido 107, 121, 213 ó 217 de SEQ ID NO: 2 o l) residuo de aminoácido 91 ó 105 de SEQ ID NO: 3. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H se codifica por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 20, 21, 48, 51, 55, 58, 59, 61, 62, 92, 93, 95, 109, 120, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) un residuo de alanina, valina, isoleucina o leucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, c) un residuo de alanina, treonina, ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, d) un residuo de prolina, serina, metionina o leucina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, e) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, f) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO: 1 y un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, g) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, h) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, i) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 175 de SEQ ID NO: 1, j) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, k) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO: 1, l) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 127 de SEQ ID NO: 1, m) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO: 1, n) un residuo de valina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, o) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, p) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, q) un residuo de treonina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, r) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, un residuo de prolina en una posición equivalente o análoga al aminoácido 110 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al aminoácido 175 de SEQ ID NO: 1, s) una prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, t) una arginina o una leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, u) una alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, v) una alanina o una serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, w) una alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3, o x) una prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución codificada por un polinucleótido que comprende una mutación de codón que da como resultado una mutación por sustitución seleccionada del grupo que consiste en a) mutación por sustitución E20A, Q21A, F48A, 151A, T52A, T52V, T52I, T52L, T55A, T58A, V59A, L61A, I62A, M92A, M93A, R95A, V96A, V96T, V96D, V96M,

V107I, F109A, A110P, A110S, A110M, A110L, Y120A, A123F, M125A, R175E, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P, M218A/C219A, V107I/R175E, Y127E/R175E, V107I/Y127E, V107I/Y127E/R175E, T52V/V107I/R175E, V96A/V107I/R175E, T52A/V107I/R175E, V96T/V107I/R175E o V107I/A110P/R175E de SEQ ID NO: 1, b) mutación por sustitución A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2 y c) mutación por sustitución G91A o A105P de SEQ ID NO: 3.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un polipéptido de dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución codificada por un polinucleótido que se hibrida con un polinucleótido que comprende una mutación de codón que da como resultado una mutación por sustitución seleccionada del grupo que consiste en a) T58A, A110P, A110L, A110S o A110M de SEQ ID NO: 1, b) A107P de SEQ ID NO: 2 y c) A105P de SEQ ID NO: 3 en condiciones de hibridación que comprenden una etapa de hibridación en menos de 500 mM de sal y al menos 37 grados Celsius y una etapa de lavado en 2XSSPE a al menos 63 grados Celsius. En una realización preferida, las condiciones de hibridación comprenden menos de 200 mM de sal y al menos 37 grados Celsius para la etapa de hibridación. En otra realización preferida, las condiciones de hibridación comprenden 2XSSPE y 63 grados Celsius para las etapas tanto de hibridación como de lavado. En otra realización preferida, el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión esteroide, tal como actividad de unión a 20-hidroxiecdisona, actividad de unión a ponasterona A o actividad de unión a muristerona A.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H comprende una mutación por sustitución en una posición equivalente o análoga a a) residuo de aminoácido 20, 21, 48, 51, 52, 55, 58, 59, 61, 62, 92, 93, 95, 96, 107, 109, 110, 120, 123, 125, 175, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) residuos de aminoácido 95 y 110 de SEQ ID NO: 1, c) residuos de aminoácido 218 y 219 de SEQ ID NO: 1, d) residuos de aminoácido 107 y 175 de SEQ ID NO: 1, e) residuos de aminoácido 127 y 175 de SEQ ID NO: 1, f) residuos de aminoácido 107 y 127 de SEQ ID NO: 1, g) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, h) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, i) residuos de aminoácido 96, 107 y 175 de SEQ ID NO: 1, j) residuos de aminoácido 107, 110, y 175 de SEQ ID NO: 1, k) residuo de aminoácido 107, 121, 213 ó 217 de SEQ ID NO: 2, o l) residuo de aminoácido 91 ó 105 de SEQ ID NO: 3. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

Preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo H comprende una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 20, 21, 48, 51, 55, 58, 59, 61, 62, 92, 93, 95, 109, 120, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) un residuo de alanina, valina, isoleucina o leucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, c) un residuo de alanina, treonina, ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, d) un residuo de prolina, serina, metionina o leucina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, e) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, f) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO: 1 y un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, g) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, h) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, i) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 175, j) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, k) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO: 1, l) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 127 de SEQ ID NO: 1, m) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO: 1, n) un residuo de valina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, o) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, p) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, q) un residuo de treonina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, r) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, s) una prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, t) una arginina o una leucina en una posición equivalente o análoga al

residuo de aminoácido 121 de SEQ ID NO: 2, u) una alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, v) una alanina o una serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, w) una alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3, o x) una prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un polipéptido de dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución, en la que la mutación por sustitución se selecciona del grupo que consiste en a) mutación por sustitución E20A, Q21A, F48A, 151A, T52A, T52V, T52I, T52L, T55A, T58A, V59A, L61A, I62A, M92A, M93A, R95A, V96A, V96T, V96D, V96M, V107I, F109A, A110P, A110S, A110M, A110L, Y120A, A123F, M125A, R175E, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P, M218A/C219A, V107I/R175E, Y127E/R175E, V107I/Y127B, V107I/Y127E/R175E, T52V/V107I/R175E, V96A/V107I/R175E, T52A/V107I/R175E, V96T/V107I/R175E o V107I/A110P/R175E de SEQ ID NO: 1, b) mutación por sustitución A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, y c) mutación por sustitución G91A o A105P de SEQ ID NO: 3.

El dominio de unión a ADN puede ser cualquier dominio de unión a ADN con un elemento de respuesta conocido, incluyendo dominios de unión a ADN sintéticos y quiméricos, o análogos, combinaciones o modificaciones de los mismos. Preferiblemente, el DBD es un DBD de GAL4, un DBD de LexA, un DBD de factor de transcripción, un DBD de miembro de receptor nuclear de grupo H, un DBD de miembro de la superfamilia de los receptores nucleares de hormonas esteroideas/tiroideas, o un DBD de LacZ bacteriano. Más preferiblemente, el DBD es un DBD de EcR [SEQ ID NO: 4 (polinucleótido) o SEQ ID NO: 5 (polipéptido)], un DBD de GAL4 [SEQ ID NO: 6 (polinucleótido) o SEQ ID NO: 7 (polipéptido)], o un DBD de LexA [(SEQ ID NO: 8 (polinucleótido) o SEQ ID NO: 9 (polipéptido)].

El dominio de transactivación (abreviado "AD" o "TA") puede ser cualquier AD de miembro de receptor nuclear de grupo H, AD de receptor nuclear de hormonas esteroideas/tiroideas, AD sintético o quimérico, AD de poliglutamina, AD de aminoácido básico o ácido, un AD de VP16, un AD de GAL4, un AD de NP-κB, un AD de BP64, un dominio de activación ácido B42 (B42AD), un dominio de transactivación p65 (p65AD), o un análogo, combinación o modificación de los mismos. En una realización específica, el AD es un AD sintético o quimérico, o se obtiene a partir de un EcR, un receptor de glucocorticoides, VP16, GAL4, NF-κB, o AD de dominio de activación ácido B42. Preferiblemente, el AD es un AD de EcR [SEQ ID NO: 10 (polinucleótido) o SEQ ID NO: 11 (polipéptido)], un AD de VP16 [SEQ ID NO: 12 (polinucleótido) o SEQ ID NO: 13 (polipéptido)], un AD de B42 [SEQ ID NO: 14 (polinucleótido) o SEQ ID NO: 15 (polipéptido)], o un AD de p65 [SEQ ID NO: 16 (polinucleótido) o SEQ ID NO: 17 (polipéptido)].

En una realización específica, el casete de expresión génica codifica para un polipéptido híbrido que comprende o bien a) un dominio de unión a ADN codificado por un polinucleótido que comprende una secuencia de ácido nucleico de SEQ ID NO: 4, SEQ ID NO: 6, o SEQ ID NO: 8, o bien b) un dominio de transactivación codificado por un polinucleótido que comprende una secuencia de ácido nucleico de SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, o SEQ ID NO: 16; y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución codificada por un polinucleótido según la invención. Preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución codificada por un polinucleótido según la invención.

En otra realización específica, el casete de expresión génica codifica para un polipéptido híbrido que comprende o bien a) un dominio de unión a ADN que comprende una secuencia de aminoácidos de SEQ ID NO: 5, SEQ ID NO: 7, o SEQ ID NO: 9, o bien b) un dominio de transactivación que comprende una secuencia de aminoácidos de SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, o SEQ ID NO: 17; y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción. Preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución según la descripción.

La presente descripción también proporciona un casete de expresión génica que comprende: i) un elemento de respuesta que comprende un dominio reconocido por un polipéptido que comprende un dominio de unión a ADN; ii) un promotor que se activa por un polipéptido que comprende un dominio de transactivación; y iii) un gen cuya expresión ha de modularse.

El elemento de respuesta ("RE") puede ser cualquier elemento de respuesta con un dominio de unión a ADN conocido, o un análogo, combinación o modificación del mismo. En la presente invención puede emplearse un RE individual o pueden usarse múltiples RE, o bien múltiples copias del mismo RE o bien dos o más RE diferentes. En una realización específica, el RE es un RE de GAL4 ("GAL4RE"), LexA, un RE de receptor nuclear de grupo H, un RE de receptor nuclear de hormonas esteroideas/tiroideas, o un RE sintético que reconoce un dominio de unión a ADN sintético. Preferiblemente, el RE es un elemento de respuesta a ecdisona (EcRE) que comprende una secuencia de polinucleótidos de SEQ ID NO: 18, un GAL4RE que comprende una secuencia de polinucleótidos de

SEQ ID NO: 19, o un RE de LexA (operón, "op") que comprende una secuencia de polinucleótidos de SEQ ID NO: 20 ("2XLexAopRE").

- Según la descripción puede obtenerse un dominio de unión a ADN, dominio de activación o elemento de respuesta de receptor nuclear de hormonas esteroideas/tiroideas, a partir de un receptor nuclear de hormonas esteroideas/tiroideas seleccionado del grupo que consiste en receptor de la hormona tiroidea  $\alpha$  ( $TR\alpha$ ), receptor tiroideo 1 (c-erbA-1), receptor de la hormona tiroidea  $\beta$  ( $TR\beta$ ), receptor de ácido retinoico  $\alpha$  ( $RAR\alpha$ ), receptor de ácido retinoico  $\beta$  ( $RAR\beta$ , HAP), receptor de ácido retinoico  $\gamma$  ( $RAR\gamma$ ), receptor de ácido retinoico tipo gamma (RARD), receptor activado por el proliferador de peroxisomas  $\alpha$  ( $PPAR\alpha$ ), receptor activado por el proliferador de peroxisomas  $\beta$  ( $PPAR\beta$ ), receptor activado por el proliferador de peroxisomas  $\delta$  ( $PPAR\delta$ , NUC-1), receptor relacionado con el activador-proliferador de peroxisomas (FFAR), receptor activado por el proliferador de peroxisomas  $\gamma$  ( $PPAR\gamma$ ), receptor huérfano codificado por la hebra no codificante de receptor de la hormona tiroidea  $\alpha$  ( $REVERB\alpha$ ), receptor relacionado con v-erb A (EAR-1), receptor relacionado con v-erb ( $EAR-1A$ ,  $\gamma$ ), receptor huérfano codificado por la hebra no codificante de receptor de la hormona tiroidea  $\beta$  ( $REVERB\beta$ ), receptor relacionado con v-erb ( $EAR-1\beta$ ), receptor nuclear huérfano BD73 (BD73), receptor relacionado con rev-erbA (RVR), proteína en dedos de zinc 126 (HZF2), proteína inducible por ecdisona E75 (E75), proteína inducible por ecdisona E78 (E78), receptor de *Drosophila* 78 (DR-78), receptor huérfano relacionado con retinoide  $\alpha$  ( $ROR\alpha$ ), receptor de retinoide Z  $\alpha$  ( $RZR\alpha$ ), receptor huérfano relacionado con retinoide  $\beta$  ( $ROR\beta$ ), receptor de retinoide Z  $\beta$  ( $RZR\beta$ ), receptor huérfano relacionado con retinoide  $\gamma$  ( $ROR\gamma$ ), receptor de retinoide Z  $\gamma$  ( $RZR\gamma$ ), receptor huérfano relacionado con retinoide (TOR), receptor hormonal 3 (HR-3), receptor hormonal de *Drosophila* 3 (DHR-3), receptor hormonal de *Manduca* (MHR-3), receptor hormonal de *Galleria* 3 (GHR-3), receptor nuclear de *C. elegans* 3 (CNR-3), receptor hormonal de *Choristoneura* 3 (CHR-3), receptor nuclear de *C. elegans* 14 (CNR-14), receptor de ecdisona (ECR), receptor ubicuo (UR), receptor nuclear huérfano (OR-1), NER-1, proteína 15 de interacción de receptor (RIP-15), receptor hepático X  $\beta$  ( $LXR\beta$ ), proteína similar al receptor de hormonas esteroideas (RLD-1), receptor hepático X ( $LXR$ ), receptor hepático X  $\alpha$  ( $LXR\alpha$ ), receptor farnesoide X (FXR), proteína 14 de interacción de receptor (RIP-14), HRR-1, receptor de vitamina D (VDR), receptor nuclear huérfano (ONR-1), receptor de pregnano X (PXR), receptor esteroide y xenobiótico (SXR), receptor de benzoato X (BXR), receptor nuclear (MB-67), receptor constitutivo de androstano 1 (CAR-1), receptor constitutivo de androstano  $\alpha$  ( $CAR\alpha$ ), receptor constitutivo de androstano 2 (CAR-2), receptor constitutivo de androstano  $\beta$  ( $CAR\beta$ ), receptor hormonal de *Drosophila* 96 (DHR-96), receptor hormonal nuclear 1 (NHR-1), factor nuclear de hepatocitos 4 (HNF-4), factor nuclear de hepatocitos 4G (HNF-4G), factor nuclear de hepatocitos 4B (HNF-4B), factor nuclear de hepatocitos 4D (HNF-4D, DHNF-4), receptor retinoide X  $\alpha$  ( $RXR\alpha$ ), receptor retinoide X  $\beta$  ( $RXR\beta$ ), proteína de unión a la región II de H-2 (H-2RIIBP), co-regulador-1 de receptor nuclear (RCoR-1), receptor retinoide X  $\gamma$  ( $RXR\gamma$ ), ultraespiráculo (USP), receptor nuclear 2Cl, factor coriónico 1 (CF-1), receptor testicular 2 (TR-2), receptor testicular 2-11 (TR2-11), receptor testicular 4 (TR4), TAK-1, receptor hormonal de *Drosophila* (DHR78), Tailless (TLL), homólogo de Tailless (TLX), XTLL, factor de transcripción del promotor en el sentido de 5' de la ovalbúmina de pollo I (COUP-TFI), factor de transcripción del promotor en el sentido de 5' de la ovalbúmina de pollo A (COUP-TFA), BAR-3, SVP-44, factor de transcripción del promotor en el sentido de 5' de la ovalbúmina de pollo II (COUP-TFII), factor de transcripción del promotor en el sentido de 5' de la ovalbúmina de pollo B (COUP-TFB), ARP-1, SVP-40, SVP, factor de transcripción del promotor en el sentido de 5' de la ovalbúmina de pollo III (COUP-TFIII), factor de transcripción del promotor en el sentido de 5' de la ovalbúmina de pollo G (COUP-TFG), SVP-46, EAR-2, receptor de estrógeno  $\alpha$  ( $ER\alpha$ ), receptor de estrógeno  $\beta$  ( $ER\beta$ ), receptor relacionado con estrógeno 1 (ERR1), receptor relacionado con estrógeno  $\alpha$  ( $ERR\alpha$ ), receptor relacionado con estrógeno 2 (ERR2), receptor relacionado con estrógeno  $\beta$  ( $ERR\beta$ ), receptor de glucocorticoides (GR), receptor de mineralocorticoides (MR), receptor de progesterona (PR), receptor de andrógenos (AR), gen inducido por el factor de crecimiento nervioso B (NGFI-B), receptor nuclear similar a Nur-77 (TRS), N10, receptor huérfano (NUR-77), gen humano de respuesta temprana (NAK-1), factor relacionado con Nurr 1 (NURR-1), un gen humano de respuesta temprana inmediata (NOT), receptor nuclear hepático regenerante 1 (RNR-1), dedo de zinc hematopoyético 3 (HZF-3), proteína-1 relacionada con Nur (TINOR), receptor huérfano nuclear 1 (NOR-1), receptor relacionado con NOR1 (MINOR), receptor hormonal de *Drosophila* 38 (DHR-38), receptor nuclear de *C. elegans* 8 (CNR-8), C48D5, factor esteroideogénico 1 (SF1), péptido similar a endozepina (ELP), factor fushi tarazu 1 (FTZ-F1), proteína de unión adrenal 4 (AD4BP), homólogo de receptor hepático (LRH-1), receptor huérfano relacionado con Ftz-F1 A (xFFrA), receptor huérfano relacionado con Ftz-F1 B (xFFrB), receptor nuclear relacionado con LRH-1 (FFLR), receptor nuclear relacionado con LRH-1 (PHR), factor de transcripción de fetoproteína (FTF), factor nuclear de células germinales (GCNFM), receptor asociado a testículos relacionado con el receptor retinoide (RTR), Knirps (KNI), relacionados con Knirps (KNRL), gónada embrionaria (EGON), gen de *Drosophila* para el receptor nuclear dependiente de ligando (EAGLE), receptor nuclear similar a Trithorax (ODR7), Trithorax, gen del cromosoma X en la región crítica de la hipoplasia adrenal congénita de reversión de sexo sensible a dosificación (DAX-1), hipoplasia adrenal congénita e hipogonadismo hipogonadotrópico (AHCH) y pareja de heterodímero corto (SHP).
- Para fines de esta descripción los receptores nucleares y receptores nucleares de grupo H también incluyen receptores nucleares sintéticos y quiméricos y receptores nucleares de grupo H y sus homólogos.

Los genes de interés para su uso en casetes de expresión génica de los solicitantes pueden ser genes endógenos o heterólogos. La información de secuencia de ácido nucleico o aminoácidos para una proteína o gen deseado puede

ubicarse en una de muchas bases de datos de acceso público, por ejemplo, GENBANK, BMBL, Swiss-Prot y PIR, o en muchas publicaciones de revista relacionadas con biología. Por tanto, los expertos en la técnica tienen acceso a la información de secuencia de ácido nucleico para prácticamente todos los genes conocidos. Entonces puede usarse tal información para construir los constructos deseados para la inserción del gen de interés dentro de los casetes de expresión génica usados en métodos de los solicitantes descritos en el presente documento.

Los ejemplos de genes de interés para su uso en casetes de expresión génica de los solicitantes incluyen, pero no se limitan a: genes que codifican para productos o polipéptidos terapéuticamente deseados que pueden usarse para tratar un estado, una enfermedad, un trastorno, una disfunción, un defecto genético, tal como anticuerpos monoclonales, enzimas, proteasas, citocinas, interferones, insulina, eritropoyetina, factores de coagulación, otros componentes o factores sanguíneos, vectores virales para la terapia génica, virus para vacunas, dianas para descubrimiento de fármacos, análisis y aplicaciones de genómica funcional y proteómica, y similares

#### Polinucleótidos de la invención

El sistema de expresión génica inducible basado en receptores nucleares novedoso de la invención comprende al menos un casete de expresión génica que comprende un polinucleótido que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución. Estos casetes de expresión génica, los polinucleótidos que comprenden, y los polipéptidos que codifican son útiles como componentes de un sistema de expresión génica basado en receptores nucleares para modular la expresión de un gen dentro de una célula huésped.

Por tanto, la presente descripción proporciona un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución.

En una realización específica, el dominio de unión a ligando de receptor nuclear de grupo H se codifica por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 20, 21, 48, 51, 52, 55, 58, 59, 61, 62, 92, 93, 95, 96, 107, 109, 110, 120, 123, 125, 175, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) residuos de aminoácido 95 y 110 de SEQ ID NO:1, c) residuos de aminoácido 218 y 219 de SEQ ID NO: 1, d) residuos de aminoácido 107 y 175 de SEQ ID NO: 1. e) residuos de aminoácido 127 y 175 de SEQ ID NO: 1, f) residuos de aminoácido 107 y 127 de SEQ ID NO:1, g) residuos de aminoácido 107,127 y 175 de SEQ ID NO:1, h) residuos de aminoácido 52,107 y 175 de SEQ ID NO: 1, i) residuos de aminoácido 96, 107 y 175 de SEQ ID NO: 1, j) residuos de aminoácido 107, 110 y 175 de SEQ ID NO: 1. k) residuo de aminoácido 107, 121, 213 ó 217 de SEQ ID NO: 2, o l) residuo de aminoácido 91 ó 105 de SEQ ID NO: 3. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H se codifica por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 20, 21, 48, 51, 55, 58, 59, 61, 62, 92, 93, 95, 109,120, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) un residuo de alanina, valina, isoleucina o leucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, c) un residuo de alanina, treonina, ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, d) un residuo de prolina, serina, metionina o leucina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, e) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, f) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO: 1 y un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, g) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, h) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, i) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 175, j) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, k) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO: 1, l) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 127 de SEQ ID NO: 1, m) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO: 1, n) un residuo de valina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, o) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, p) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al

residuo de aminoácido 175 de SEQ ID NO: 1, q) un residuo de treonina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, r) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, un residuo de prolina en una posición equivalente o análoga al aminoácido 110 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al aminoácido 175 de SEQ ID NO: 1, s) una prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, t) una arginina o una leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, u) una alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, v) una alanina o una serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, w) una alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3, o x) una prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución codificada por un polinucleótido que comprende una mutación de codón que da como resultado una mutación por sustitución seleccionada del grupo que consiste en a) mutación por sustitución E20A, Q21A, F48A, I51A, T52A, T52V, T52I, T52L, T55A, T58A, V59A, L61A, I62A, M92A, M93A, R95A, V96A, V96T, V96D, V96M, V107I, F109A, A110P, A110S, A110M, A101L, Y120A, A123F, M125A, R175E, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P, M218A/C219A, V107I/R175E, Y127E/R175E, V107I/Y127E, V107I/Y127E/R175E, T52V/V107I/R175E, V96A/V107I/R175E, T52A/V107I/R175E, V96T/V107I/R175E o V107I/A110P/R175E de SEQ ID NO: 1, b) mutación por sustitución A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2 y c) mutación por sustitución G91A o A105P de SEQ ID NO: 3.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución codificada por un polinucleótido que se hibrida con un polinucleótido que comprende una mutación de codón que da como resultado una mutación por sustitución seleccionada del grupo que consiste en a) T58A, A110P, A110L, A110S o A110M de SEQ ID NO: 1, b) A107P de SEQ ID NO: 2, y c) A105P de SEQ ID NO: 3 en condiciones de hibridación que comprenden una etapa de hibridación en menos de 500 mM de sal y al menos 37 grados Celsius y una etapa de lavado en 2XSSPE a al menos 63 grados Celsius. En una realización preferida, las condiciones de hibridación comprenden menos de 200 mM de sal y al menos 37 grados Celsius para la etapa de hibridación. En otra realización preferida, las condiciones de hibridación comprenden 2XSSPE y 63 grados Celsius para las etapas tanto de hibridación como de lavado. En otra realización preferida, el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides tal como 20-hidroxicdisona, ponasterona A o muristerona A.

La presente descripción también proporciona un polinucleótido aislado que codifica para un polipéptido seleccionado del grupo que consiste en a) un polipéptido que comprende un dominio de transactivación, un dominio de unión a ADN, y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción b) un polipéptido que comprende un dominio de unión a ADN y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción; y c) un polipéptido que comprende un dominio de transactivación y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción

En una realización específica, el polinucleótido aislado codifica para un polipéptido híbrido seleccionado del grupo que consiste en a) un polipéptido híbrido que comprende un dominio de transactivación, un dominio de unión a ADN, y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción: b) un polipéptido híbrido que comprende un dominio de unión a ADN y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción; y c) un polipéptido híbrido que comprende un dominio de transactivación y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción.

La presente descripción también se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución afecta a la actividad de unión a ligando o la sensibilidad a ligando del dominio de unión a ligando de receptor nuclear de grupo H.

En particular, la presente descripción se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución reduce la actividad de unión a ligando o la sensibilidad a ligando del dominio de unión a ligando de receptor nuclear de grupo H.

En una realización específica, la presente descripción se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución reduce la actividad de unión a esteroides o la sensibilidad a esteroides del dominio de

- unión a ligando de receptor nuclear de grupo H. Preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 20, 21, 48, 51, 52, 55, 58, 59, 62, 92, 93, 95, 109, 110, 120, 123, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) residuos de aminoácido 95 y 110 de SEQ ID NO: 1, c) residuos de aminoácido 218 y 219 de SEQ ID NO: 1, d) residuo de aminoácido 107, 121, 213, o 217 de SEQ ID NO: 2, o e) residuo de aminoácido 105 de SEQ ID NO: 3. Más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 20, 21, 48, 51, 52, 55, 58, 59, 62, 92, 93, 95, 109, 120, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO:1, b) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO:1, c) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, d) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO: 1 y un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, e) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, f) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, g) un residuo de arginina o leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, h) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, i) un residuo de alanina o serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o j) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una mutación por sustitución de a) E20A, Q21A, F48A, I51A, T52A, T55A, T58A, V59A, I62A, M92A, M93A, R95A, F109A, A110P, Y120A, A123F, M125A, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P o M218A/C219A de SEQ ID NO: 1, b) A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, o c) A105P de SEQ ID NO: 3.
- En otra realización específica, la presente descripción se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución elimina la actividad de unión a esteroides o la sensibilidad a esteroides del dominio de unión a ligando de grupo H. Preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 58 ó 110 de SEQ ID NO:1, b) residuos de aminoácido 107, 110 y 175 de SEQ ID NO:1, c) residuo de aminoácido 107,121, 213 ó 217 de SEQ ID NO: 2, o d) residuo de aminoácido 105 de SEQ ID NO: 3. Más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de a) una alanina en una posición equivalente o análoga al residuo de aminoácido 58 de SEQ ID NO:1, b) un residuo de prolina, leucina, serina o metionina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO:1, c) una isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO:1, una prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, y una glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, d) una prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, e) una arginina o una leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, f) una alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, g) una alanina o una serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o h) una prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una mutación por sustitución seleccionada del grupo que consiste en a) mutación por sustitución T58A, A110P, A110L, A110S, A110M, o V107I/A110P/R175E de SEQ ID NO:1, b) mutación por sustitución A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, y c) mutación por sustitución A105P de SEQ ID NO: 3.

La presente descripción también se refiere a un polinucleótido aislado que codifica para un polipéptido que comprende un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución, en la que el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides. Preferiblemente, el dominio de unión a ligando de receptor de ecdisona comprende una mutación de codón que da como resultado una mutación por sustitución en una posición equivalente o análoga al residuo de aminoácido a a) residuo de aminoácido 58 ó 110 de SEQ ID NO: 1, b) residuos de aminoácido 107, 110 y 175 de SEQ ID NO:1, b) residuo de aminoácido 107, 121, 213 ó 217 de SEQ ID NO: 2, o d) residuo de aminoácido 105 de SEQ ID NO: 3. Más preferiblemente, el dominio de unión a ligando de receptor de ecdisona comprende una mutación de codón que da como resultado una sustitución de a) una alanina en una posición equivalente o análoga al residuo de aminoácido 58 de SEQ ID NO:1, b) un residuo de prolina, leucina, serina o metionina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, c) una isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO:1, una prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, una glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, d) una prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, e) una arginina o una leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, f) una alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, g) una alanina o una serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o h) una prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el dominio de unión a ligando de receptor de ecdisona aislado comprende una mutación de codón que da como



resultado una mutación por sustitución seleccionada del grupo que consiste en a) mutación por sustitución T58A, A110P, A110L, A110S, A110M o V107I/A110P/R175E de SEQ ID NO: 1, b) mutación por sustitución A107P, 0121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, y c) mutación por sustitución A105P de SEQ ID NO: 3. En una realización específica, el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides seleccionada del grupo que consiste en actividad de unión a ecdisona, actividad de unión a 20-hidroxiecdisona, actividad de unión a ponasterona A y actividad de unión a muristerona A.

En otra realización específica, el polinucleótido aislado que codifica para un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución, en la que el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides, se hibrida con polinucleótido que comprende una mutación de codón que da como resultado una mutación por sustitución seleccionada del grupo que consiste en a) mutación por sustitución T58A, A110P, A110L, A110S, A110M o V107I/A110P/R175E de SEQ ID NO: 1, b) mutación por sustitución A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, y c) mutación por sustitución A105P de SEQ ID NO: 3 en condiciones de hibridación que comprenden una etapa de hibridación en menos de 500 mM de sal y al menos 37 grados Celsius y una etapa de lavado en 2XSSPE a al menos 63 grados Celsius. En una realización preferida, las condiciones de hibridación comprenden menos de 200 mM de sal y al menos 37 grados Celsius para la etapa de hibridación. En otra realización preferida, las condiciones de hibridación comprenden 2XSSPE y 63 grados Celsius para las etapas tanto de hibridación como de lavado. En otra realización preferida, el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides seleccionada del grupo que consiste en actividad de unión a ecdisona, actividad de unión a 20-hidroxiecdisona, actividad de unión a ponasterona A y actividad de unión a muristerona A.

En otra realización específica, la presente descripción se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución reduce la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de receptor nuclear de grupo H. Preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga al residuo de aminoácido a) 21, 48, 51, 52, 59, 62, 93, 95, 96, 109, 120, 123, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO:1, b) 121, 213 ó 217 de SEQ ID NO:2 o c) 105 de SEQ ID NO: 3. Más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 21, 48, 51, 59, 62, 93, 95, 96, 109,120,125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) un residuo de leucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, c) un residuo de treonina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, d) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, e) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO:1 y un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, f) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, g) una arginina o residuo de leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, h) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, i) un residuo de alanina o serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o j) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una mutación por sustitución de a) Q21A, F48A, I51A, T52L, V59A, I62A, M93A, R95A, V96A, V96T, F109A, Y120A, A123F, M125A, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P o M218/C219A de SEQ ID NO: 1, b) G121R, 01211, N213A, C217A o C217S de SEQ ID NO: 2, o c) A105P de SEQ ID NO: 3.

En otra realización específica, la presente descripción se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de polipéptido de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución elimina la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de grupo H.

En otra realización específica, la presente descripción se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de polipéptido de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución reduce tanto la actividad de unión a esteroides como la sensibilidad a esteroides y la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de grupo H. Preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 21, 48, 51, 59, 62, 93, 95, 109, 120,123, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) residuos de aminoácido 95 y 110 de SEQ ID NO:1, c) residuos de aminoácido 218 y 219 de SEQ ID NO: 1, d) residuo de aminoácido 121, 213 ó 217 de SEQ ID NO: 2, o e) residuo de aminoácido 105 de SEQ ID NO: 3. Más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 21, 48, 51, 59, 62, 93, 95, 109, 120, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, c) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO:1 y un residuo de prolina en una posición

equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, d) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, e) un residuo de arginina o leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, f) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, g) un residuo de alanina o serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o h) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una mutación por sustitución de Q21A, F48A, I51A, V59A, I62A, M93A, R95A, F109A, Y120A, A123F, M125A, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P o M218A/C219A de SEQ ID NO: 1, b) G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, o c) A105P de SEQ ID NO: 3.

Además, la presente descripción también se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución potencia la actividad de unión a ligando o la sensibilidad a ligando del dominio de unión a ligando de receptor nuclear de grupo H.

En una realización específica, la presente descripción se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución potencia la actividad de unión a esteroides o la sensibilidad a esteroides del dominio de unión a ligando de receptor nuclear de grupo H. Preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 52 ó 96 de SEQ ID NO: 1 o b) residuo de aminoácido 91 de SEQ ID NO: 3. Más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de a) un residuo de leucina, valina o isoleucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, b) un residuo de alanina, treonina, ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, c) un residuo de treonina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de (SEQ ID NO: 1, o d) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3. Incluso más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado a mutación por sustitución de a) T52L, T52V, T52I, V96A, V96T, V96D, o V96M de SEQ ID NO: 1 o b) G91A de SEQ ID NO: 3.

En otra realización específica, la presente descripción se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución potencia la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de receptor nuclear de grupo H. Preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga al residuo de aminoácido 52, 5 ó 96 de SEQ ID NO: 1 o b) residuo de aminoácido 91 de SEQ ID NO: 3. Más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de a) un residuo de alanina, valina o isoleucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, b) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 55 de SEQ ID NO: 1, c) un residuo de ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, o d) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3. Incluso más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una mutación por sustitución de a) T52A, T52V, TS2I, T55A, V96D o V96M de SEQ ID NO: 1 o b) G91A de SEQ ID NO: 3.

En otra realización específica, la presente descripción se refiere a un polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en la que la mutación por sustitución potencia tanto la actividad de unión a esteroides o la sensibilidad a esteroides como la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de grupo H. Preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 52, 96, 107 ó 175 de SEQ ID NO: 1, b) residuos de aminoácido 107 y 175 de SEQ ID NO: 1, c) residuos de aminoácido 127 y 175 de SEQ ID NO: 1, d) residuos de aminoácido 107 y 127 de SEQ ID NO: 1, e) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, f) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, g) residuos de aminoácido 96, 107 y 175 de SEQ ID NO: 1, o h) residuo de aminoácido 91 de SEQ ID NO: 3. Más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de a) un residuo de valina o isoleucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, b) un residuo de ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, c) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, d) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 175, e) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, f) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y

175 de SEQ ID NO: 1, g) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 127 de SEQ ID NO:1, h) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO:1 y un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO: 1, i) un residuo de valina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, j) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO:1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, k) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, o l) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3. Incluso más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una mutación por sustitución de a) T52V, T52I, V96D, V96M, V107I, R175E, V107I/R175E, Y127E/R175E, V107I/Y127E, V107I/Y127E/R175E, T52V/V10I/VR175E, V96A/V107I/R175E o T52A/V107I/R175E de SEQ ID NO: 1 o b) G91A de SEQ ID NO: 3.

Además, la presente descripción se refiere a un vector de expresión que comprende un polinucleótido según la invención, operativamente unido a un elemento regulador de la transcripción. Preferiblemente, el polinucleótido que codifica para un dominio de unión a ligando de receptor nuclear que comprende una mutación por sustitución está operativamente unido con una secuencia de control de la expresión que permite la expresión del dominio de unión a ligando de receptor nuclear en una célula huésped competente para la expresión. La secuencia de control de la expresión puede comprender un promotor que es funcional en la célula huésped en la que se desea la expresión. El vector puede ser una molécula de ADN de plásmido o un vector viral. Los vectores virales preferidos incluyen retrovirus, adenovirus, virus adenoasociado, virus del herpes y virus vaccinia. La descripción se refiere además a un virus recombinante de replicación defectuosa que comprende en su genoma, el polinucleótido que codifica para un dominio de unión a ligando de receptor nuclear que comprende una mutación por sustitución tal como se describió anteriormente. Por tanto, la presente descripción también se refiere a una célula huésped aislada que comprende un vector de expresión de ese tipo, en el que el elemento regulador de la transcripción es operativo en la célula huésped.

La presente invención también se refiere a un polipéptido aislado codificado por un polinucleótido según la invención.

#### Polipéptidos de la invención

El sistema de expresión génica inducible basado en receptores nucleares novedoso de la invención comprende al menos un casete de expresión génica que comprende un polinucleótido que codifica para un polipéptido que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución. Por tanto, la presente invención también proporciona un polipéptido aislado que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H comprende una mutación por sustitución en una posición equivalente o análoga a a) residuo de aminoácido 20, 21, 48, 51, 52, 55, 58, 59, 61, 62, 92, 93, 95, 96, 107, 109, 110, 120, 123, 125, 175, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO:1, b) residuos de aminoácido 95 y 110 de SEQ ID NO:1, c) residuos de aminoácido 218 y 219 de SEQ ID NO:1, d) residuos de aminoácido 107 y 175 de SEQ ID NO: 1, e) residuos de aminoácido 127 y 175 de SEQ ID NO: 1, f) residuos de aminoácido 107 y 127 de SEQ ID NO: 1, g) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, h) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, i) residuos de aminoácido 96,107 y 175 de SEQ ID NO: 1, j) residuos de aminoácido 107, 110 y 175 de SEQ ID NO: 1, k) residuo de aminoácido 107, 121, 213, o 217 de SEQ ID NO: 2, o l) residuo de aminoácido 91 ó 105 de SEQ ID NO: 3. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

Preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo H comprende una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 20, 21, 48, 51, 55, 58, 59, 61, 62, 92, 93, 95, 109, 120, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) un residuo de alanina, valina, isoleucina o leucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO:1, c) un residuo de alanina, treonina, ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO:1, d) un residuo de prolina, serina, metionina o leucina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, e) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, f) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO: 1 y un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, g) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, h) un residuo de isoleucina en una posición

equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, i) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 175, j) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, k) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO: 1, l) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 127 de SEQ ID NO: 1, m) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO: 1, n) un residuo de valina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, o) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, p) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, q) un residuo de treonina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, r) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, una prolina en una posición equivalente o análoga al aminoácido 110 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, s) una prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, t) una arginina o una leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, u) una alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, v) una alanina o una serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, w) una alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3, o x) una prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un polipéptido de dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución, en la que la mutación por sustitución se selecciona del grupo que consiste en a) mutación por sustitución E20A, Q21A, F48A, I51A, T52A, T52V, T52I, T52L, T55A, T58A, V59A, L61A, I62A, M92A, M93A, R95A, V96A, V96T, V96D, V96M, V107I, F109A, A110P, A110S, A110M, A110L, Y120A, A123F, M125A, R175E, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P, M21BA/C219A, V107I/R175E, Y127E/R175E, V107I/Y127E, V107I/Y127E/R175E, T52V/V107R175E, V96A/V107I/R175E, T52A/V107I/R175E V96T/V107I/R175E o V107I/A110P/R175E de SEQ ID NO: 1, b) mutación por sustitución A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, y c) mutación por sustitución G91A o A105P de SEQ ID NO: 3.

La presente descripción también proporciona un polipéptido aislado seleccionado del grupo que consiste en a) un polipéptido aislado que comprende un dominio de transactivación, un dominio de unión a ADN, y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción; b) un polipéptido aislado que comprende un dominio de unión a ADN y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción; y c) un polipéptido aislado que comprende un dominio de transactivación y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

La presente descripción también proporciona un polipéptido híbrido aislado seleccionado del grupo que consiste en a) un polipéptido híbrido aislado que comprende un dominio de transactivación, un dominio de unión a ADN, y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción b) un polipéptido híbrido aislado que comprende un dominio de unión a ADN y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción; y c) un polipéptido híbrido aislado que comprende un dominio de transactivación y un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución según la descripción. En una realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H es de un receptor de ecdisona.

La presente descripción también proporciona un polipéptido aislado que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución que afecta a la actividad de unión a ligando o a la sensibilidad a ligando del dominio de unión a ligando de receptor nuclear de grupo H.

En particular, la presente descripción se refiere a un polipéptido de receptor nuclear de grupo H aislado que comprende un dominio de unión a ligando que comprende una mutación por sustitución que reduce la actividad de

unión a ligando o la sensibilidad a ligando del dominio de unión a ligando de receptor nuclear de grupo H.

En una realización específica, la presente descripción se refiere a un polipéptido aislado que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución que reduce la actividad de unión a esteroides o la sensibilidad a esteroides del dominio de unión a ligando de receptor nuclear de grupo H. Preferiblemente, el polipéptido aislado comprende una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 20, 21, 48, 51, 52, 55, 58, 59, 62, 92, 93, 95, 109, 110, 120, 123, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) residuo de aminoácido 107, 121, 213 ó 217 de SEQ ID NO: 2, o c) residuo de aminoácido 105 de SEQ ID NO: 3. Más preferiblemente, el polipéptido aislado comprende una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 20, 21, 48, 51, 52, 55, 58, 59, 62, 92, 93, 95, 109, 120, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, c) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, d) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO: 1 y un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, e) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, f) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, g) un residuo de arginina o leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, h) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, i) un residuo de alanina o serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o j) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el polipéptido aislado comprende una mutación por sustitución de a) E20A, Q21A, F48A, 151A, T52A, T55A, T58A, V59A, I62A, M92A, M93A, R95A, F109A, A110P, Y120A, A123F, M125A, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P o M218A/C219A de SEQ ID NO: 1, b) A107P, G121R G121L, N213A, C217A o C217S de SEQ ID NO: 2, o c) A105P de SEQ ID NO: 3.

En otra realización específica, la presente descripción se refiere a un polipéptido aislado que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución que elimina la actividad de unión a esteroides o la sensibilidad a esteroides del dominio de unión a ligando de grupo H. Preferiblemente, el polipéptido aislado comprende una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 58 ó 110 de SEQ ID NO: 1, b) residuos de aminoácido 107, 110 y 175 de SEQ ID NO: 1, c) residuo de aminoácido 107, 121, 213 ó 217 de SEQ ID NO: 2, o d) residuo de aminoácido 105 de SEQ ID NO: 3. Más preferiblemente, el polinucleótido aislado comprende una mutación de codón que da como resultado una sustitución de a) una alanina en una posición equivalente o análoga al residuo de aminoácido 58 de SEQ ID NO: 1, b) un residuo de prolina, leucina, serina o metionina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, c) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, un residuo de prolina en una posición equivalente o análoga al aminoácido 110 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, d) una prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, e) una arginina o una leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, f) una alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, g) una alanina o una serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o h) una prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el polipéptido aislado comprende una mutación por sustitución seleccionado del grupo que consiste en a) mutación por sustitución T58A, A110P, A110L, A110S, A110M o V107I/A110P/R175E de SEQ ID NO: 1, b) mutación por sustitución A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, y c) mutación por sustitución A105P de SEQ ID NO: 3.

La presente descripción también se refiere a un polipéptido aislado que comprende un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución, en la que el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides. Preferiblemente, el dominio de unión a ligando de receptor de ecdisona comprende una mutación por sustitución a una posición equivalente o análoga al residuo de aminoácido a a) residuo de aminoácido 58 ó 110 de SEQ ID NO: 1, b) residuos de aminoácido 107, 110 y 175 de SEQ ID NO: 1, c) residuo de aminoácido 107, 121, 213 ó 217 de SEQ ID NO: 2, o d) residuo de aminoácido 105 de SEQ ID NO: 3. Más preferiblemente, el dominio de unión a ligando de receptor de ecdisona comprende una sustitución de a) una alanina en una posición equivalente o análoga al residuo de aminoácido 58 de SEQ ID NO: 1, b) un residuo de prolina, leucina, serina o metionina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, c) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, d) una prolina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 2, e) una arginina o una leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, f) una alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, g) una alanina o una serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o h) una prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el dominio de unión a ligando de receptor de ecdisona comprende una mutación por sustitución seleccionada del

grupo que consiste en a) mutación por sustitución T58A, A110f, A110L, A110S, A110M o V107I/A110P/R175E de SEQ ID NO: 1, b) mutación por sustitución A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, y c) mutación por sustitución A105P de SEQ ID NO: 3. En una realización específica, el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides seleccionada del grupo que consiste en actividad de unión a ecdisona, actividad de unión a 20-hidroxiectdisona, actividad de unión a ponasterona A y actividad de unión a muristerona A.

En otra realización específica, el polipéptido aislado que comprende un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución, en la que el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides y se codifica por un polinucleótido que se hibrida con un polinucleótido que comprende una mutación de codón que da como resultado una mutación por sustitución seleccionada del grupo que consiste en a) mutación por sustitución T58A, A110P, A110L, A110S, A110M o V107I/A110P/R175E de SEQ ID NO:1, b) mutación por sustitución A107P, G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, y c) mutación por sustitución A105P de SEQ ID NO: 3 en condiciones de hibridación que comprenden una etapa de hibridación en menos de 500 mM de sal y al menos 37 grados Celsius y una etapa de lavado en 2XSSPE a al menos 63 grados Celsius. En una realización preferida, las condiciones de hibridación comprenden menos de 200 mM de sal y al menos 37 grados Celsius para la etapa de hibridación. En otra realización preferida, las condiciones de hibridación comprenden 2XSSPE y 63 grados Celsius para las etapas tanto de hibridación como de lavado. En otra realización preferida, el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides seleccionada del grupo que consiste en actividad de unión a ecdisona, actividad de unión a 20-hidroxiectdisona, actividad de unión a ponasterona A y actividad de unión a muristerona A.

En otra realización específica, la presente descripción se refiere a un polipéptido aislado que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución que reduce la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de receptor nuclear de grupo H. Preferiblemente, el polipéptido aislado comprende una sustitución de un residuo de aminoácido en una posición equivalente o análoga al residuo de aminoácido a) 21, 48, 51, 52, 59, 62, 93, 95, 96, 109, 120, 123, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO:1, b) 121, 213 ó 217 de SEQ ID NO: 2, o c) 105 de SEQ ID NO: 3. Más preferiblemente, el polipéptido aislado comprende una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 21, 48, 51, 59, 62, 93, 95, 96, 109, 120, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) un residuo de leucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, c) un residuo de treonina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, d) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, e) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO: 1 y un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, f) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, g) una arginina o un residuo de leucina en una posición equivalente o análoga al residuo de aminoácido 121 de SEQ ID NO: 2, h) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, i) un residuo de alanina o serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o j) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el polipéptido aislado comprende una mutación por sustitución de a) Q21A, F48A, I51A, T52L, V59A, I62A, M93A, R95A, V96A, V96T, F109A, Y120A, A123F, M125A, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P o M218/C219A de SEQ ID NO: 1, b) G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, o c) A105P de SEQ ID NO: 3.

En otra realización específica, la presente descripción se refiere a un polipéptido aislado que comprende un dominio de unión a ligando de polipéptido de receptor nuclear de grupo H que comprende una mutación por sustitución que elimina la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de grupo H.

En otra realización específica, la presente descripción se refiere a un polipéptido aislado que comprende un dominio de unión a ligando de polipéptido de receptor nuclear de grupo H que comprende una mutación por sustitución que reduce tanto la actividad de unión a esteroides o la sensibilidad a esteroides como la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de grupo H. Preferiblemente, el polipéptido aislado comprende una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 21, 48, 51, 59, 62, 93, 95, 109, 120, 123, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) residuos de aminoácido 95 y 110 de SEQ ID NO:1, c) residuos de aminoácido 218 y 219 de SEQ ID NO: 1, d) residuo de aminoácido 121, 213 ó 217 de SEQ ID NO: 2, o e) residuo de aminoácido 105 de SEQ ID NO: 3. Más preferiblemente, el polipéptido aislado comprende una sustitución de a) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 21, 48, 51, 59, 62, 93, 95, 109, 120, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1, b) un residuo de fenilalanina en una posición equivalente o análoga al residuo de aminoácido 123 de SEQ ID NO: 1, c) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 95 de SEQ ID NO: 1 y un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 110 de SEQ ID NO: 1, d) un residuo de alanina en una posición equivalente o análoga a los residuos de aminoácido 218 y 219 de SEQ ID NO: 1, e) un residuo de arginina o leucina en una posición equivalente o análoga

al residuo de aminoácido 121 de SEQ ID NO: 2, f) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 213 de SEQ ID NO: 2, g) un residuo de alanina o serina en una posición equivalente o análoga al residuo de aminoácido 217 de SEQ ID NO: 2, o h) un residuo de prolina en una posición equivalente o análoga al residuo de aminoácido 105 de SEQ ID NO: 3. Incluso más preferiblemente, el polipéptido aislado  
 5 comprende una mutación por sustitución de Q21A, F48A, I51A, V59A, I62A, M93A, R95A, F109A, Y120A, A123F, M125A, M218A, C219A, L223A, L230A, L234A, W238A, R95A/A110P o M218A/C219A de SEQ ID NO: 1, b) G121R, G121L, N213A, C217A o C217S de SEQ ID NO: 2, o c) A105P de SEQ ID NO: 3.

Además, la presente descripción también se refiere a un polipéptido aislado que comprende un dominio de unión a  
 10 ligando de receptor nuclear de grupo H que comprende una mutación por sustitución que potencia la actividad de unión a ligando o sensibilidad a ligando del dominio de unión a ligando de receptor nuclear de grupo H.

En una realización específica, la presente descripción se refiere a un polipéptido aislado que comprende un dominio  
 15 de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución que potencia la actividad de unión a esteroides o la sensibilidad a esteroides del dominio de unión a ligando de receptor nuclear de grupo H. Preferiblemente, el polipéptido aislado comprende una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 52 ó 96 de SEQ ID NO: 1, b) residuos de aminoácido 96, 107 y 175 de SEQ ID NO: 1, o c) residuo de aminoácido 91 de SEQ ID NO: 3. Más preferiblemente, el polipéptido  
 20 aislado comprende una sustitución de a) un residuo de leucina, valina o isoleucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, b) un residuo de alanina, treonina, ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, c) un residuo de treonina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, o d) un residuo de  
 25 alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3. Incluso más preferiblemente, el polipéptido aislado comprende una mutación por sustitución de a) T52L, T52V, T52I, V96A, V96T, V96D, V96M o V96T/V107I/R175E de SEQ ID NO:1 o b) G91A de SEQ ID NO: 3.

En otra realización específica, la presente descripción se refiere a un polipéptido aislado que comprende un dominio  
 30 de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución que potencia la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de receptor nuclear de grupo H. Preferiblemente, el polipéptido aislado comprende una sustitución de un residuo de aminoácido en una posición equivalente o análoga al residuo de aminoácido 52, 55 ó 96 de SEQ ID NO: 1 o b) residuo de aminoácido 91 de SEQ ID NO: 3. Más preferiblemente, el polipéptido aislado comprende una  
 35 sustitución de a) un residuo de alanina, valina o isoleucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, b) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 55 de SEQ ID NO:1, c) un residuo de ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO:1, o d) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3. Incluso más preferiblemente, el polipéptido aislado comprende una  
 40 mutación por sustitución de a) T52A, T52V, T52I, T55A, V96D o V96M de SEQ ID NO: 1 o b) G91A de SEQ ID NO: 3.

En otra realización específica, la presente descripción se refiere a un polipéptido aislado que comprende un dominio  
 45 de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución que potencia tanto la actividad de unión a esteroides o la sensibilidad a esteroides como la actividad de unión a compuestos no esteroideos o la sensibilidad a compuestos no esteroideos del dominio de unión a ligando de grupo H. Preferiblemente, el polipéptido aislado comprende una sustitución de un residuo de aminoácido en una posición equivalente o análoga a a) residuo de aminoácido 52, 96, 107 ó 175 de SEQ ID NO. 1, b) residuos de aminoácido 107 y 175 de SEQ ID NO: 1, c) residuos de aminoácido 127 y 175 de SEQ ID NO: 1, d) residuos de aminoácido 107  
 50 y 127 de SEQ ID NO: 1, e) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, f) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, g) residuos de aminoácido 96, 107 y 175 de SEQ ID NO: 1, o h) residuo de aminoácido 91 de SEQ ID NO: 3. Más preferiblemente, el polipéptido aislado comprende una sustitución de a) un residuo de valina o isoleucina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO:1, b) un residuo de ácido aspártico o metionina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ  
 55 ID NO: 1, c) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1, d) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 175, e) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, f) un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO:  
 60 1, g) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 127 de SEQ ID NO: 1, h) un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga a los residuos de aminoácido 127 y 175 de SEQ ID NO:  
 65 1, i) un residuo de valina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, j) un

residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 96 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, k) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 52 de SEQ ID NO: 1, un residuo de isoleucina en una posición equivalente o análoga al residuo de aminoácido 107 de SEQ ID NO: 1 y un residuo de glutamina en una posición equivalente o análoga al residuo de aminoácido 175 de SEQ ID NO: 1, o l) un residuo de alanina en una posición equivalente o análoga al residuo de aminoácido 91 de SEQ ID NO: 3. Incluso más preferiblemente, el polipéptido aislado comprende una mutación por sustitución de a) T52V, T52I, V96D, V96M, V107I, R175E, V107I/R175E, Y127E/R175E, V107I/Y127E, V107I/Y127E/R175E, T52V/V107I/R175E, V96A/V107I/R175B o T52A/V107I/R175E de SEQ ID NO: 1 o b) G91A de SEQ ID NO: 3.

La presente descripción también se refiere a composiciones que comprenden un polipéptido aislado según la invención.

#### 15 Método para modular la expresión génica de la invención

La invención de los solicitantes también se refiere a métodos para modular la expresión génica en una célula huésped usando un sistema de modulación de la expresión génica según la invención. Específicamente, los solicitantes proporcionan un método para modular la expresión de un gen en una célula huésped que comprende las etapas de: a) introducir en la célula huésped un sistema de modulación de la expresión génica según la descripción y b) introducir en la célula huésped un ligando; en el que el gen que va a modularse es un componente de un casete de expresión génica que comprende: i) un elemento de respuesta que comprende un dominio reconocido mediante el dominio de unión a ADN del sistema de expresión génica; ii) un promotor que se activa por el dominio de transactivación del sistema de expresión génica; y iii) un gen cuya expresión ha de modularse, mediante lo cual tras la introducción del ligando en la célula huésped, se modula la expresión del gen.

La descripción también proporciona un método para modular la expresión de un gen en una célula huésped que comprende las etapas de: a) introducir en la célula huésped un sistema de modulación de la expresión génica según la descripción; b) introducir en la célula huésped un casete de expresión génica según la descripción en el que el casete de expresión génica comprende i) un elemento de respuesta que comprende un dominio reconocido por el dominio de unión a ADN del sistema de expresión génica; ii) un promotor que se activa por el dominio de transactivación del sistema de expresión génica; y iii) un gen cuya expresión ha de modularse; y c) introducir en la célula huésped un ligando; mediante lo cual tras la introducción del ligando en la célula huésped, se modula la expresión del gen.

Los solicitantes también proporcionan un método para modular la expresión de un gen en una célula huésped que comprende un casete de expresión génica que comprende un elemento de respuesta que comprende un dominio al cual se une el dominio de unión a ADN del primer polipéptido híbrido del sistema de modulación de la expresión génica; un promotor que se activa por el dominio de transactivación del segundo polipéptido híbrido del sistema de modulación de la expresión génica; y un gen cuya expresión ha de modularse; en el que el método comprende las etapas de: a) introducir en la célula huésped un sistema de modulación de la expresión génica según la descripción y b) introducir en la célula huésped un ligando; mediante lo cual tras la introducción del ligando en el huésped, se modula la expresión del gen.

Los genes de interés para la expresión en una célula huésped usando métodos de los solicitantes pueden ser genes endógenos o genes heterólogos. La información de secuencia de ácido nucleico o de aminoácidos para una proteína o gen deseado puede ubicarse en una de muchas bases de datos de acceso público, por ejemplo, GENBANK, EMBL, Swiss-Prot y PIR, o en muchas publicaciones de revista relacionadas con biología. Por tanto, los expertos en la técnica tienen acceso a información de secuencia de ácido nucleico para prácticamente todos los genes conocidos. Entonces puede usarse tal información para construir los constructos deseados para la inserción del gen de interés dentro de los casetes de expresión génica usados en los métodos de los solicitantes descritos en el presente documento.

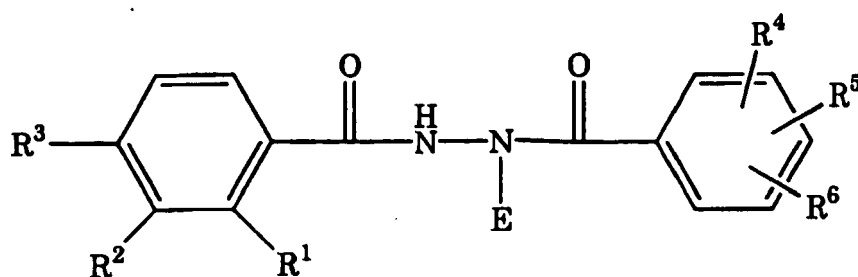
Los ejemplos de genes de interés para la expresión en una célula huésped usando métodos de los solicitantes incluyen, pero no se limitan a: antígenos producidos en plantas como vacunas, enzimas como alfa-amilasa, fitasa, glucanos, xilasa y xilanasa, genes para resistencia contra insectos, nematodos, hongos, bacterias, virus y estreses abióticos, nutracéuticos, productos farmacéuticos, vitaminas, genes para modificar el contenido de aminoácidos, resistencia a herbicidas, tolerancia al frío, la sequía y el calor, productos industriales, aceites, proteína, hidratos de carbono, antioxidantes, plantas estériles masculinas, flores, combustibles, otros rasgos de producción, genes que codifican polipéptidos o productos terapéuticamente deseados que pueden usarse para tratar un estado, una enfermedad, un trastorno, una disfunción, un defecto genético, tal como anticuerpos monoclonales, enzimas, proteasas, citocinas, interferones, insulina, eritropoyetina, factores de coagulación, otros componentes o factores sanguíneos, vectores virales para la terapia génica, virus para vacunas, dianas para descubrimiento de fármacos, análisis y aplicaciones de genómica funcional y proteómica, y similares.

Ligandos aceptables son uno cualquiera que module la expresión del gen cuando la unión del dominio de unión a



ADN del sistema de expresión génica según la descripción al elemento de respuesta en presencia del ligando da como resultado la activación o supresión de la expresión de los genes. Los ligandos preferidos incluyen un ecdisteroide, tal como ecdisona, 20-hidroxiectdisona, ponasterona A, muristerona A, y similares, ácido 9-cis-retinoico, análogos sintéticos de ácido retinoico, N,N'-diacilhidrazinas tales como las dadas a conocer en las patentes de EE.UU. n.º 6.013.836, n.º 5.117.057, n.º 5.530.028 y n.º 5.378.726; dibenzoilalquil-cianohidrazinas tales como las dadas a conocer en la solicitud europea n.º 461809; N-alquil-N,N'-diaroihidrazinas tales como las dadas a conocer en la patente de EE.UU. n.º 5.225.443; N-acil-N-alquilcarbonilhidrazinas tales como las dadas a conocer en la solicitud europea n.º 234994; N-aroi-N-alquil-N'-aroihidrazinas tales como las descritas en la patente de EE.UU. n.º 4.985.461; y otros materiales similares incluyendo 3,5-di-terc-butil-4-hidroxi-N-isobutil-benzamida, 8-O-acetilharpagida, oxiesteroles, 22(R) hidroxicolesterol, 24(S) hidroxicolesterol, 25-epoxicolesterol, T0901317, 5-alfa-6-alfa-epoxicolesterol-3-sulfato (ECHS), 7-cetocolesterol-3-sulfato, farnesol, ácidos biliares, ésteres de 1,1-bifosfonato, hormona juvenil III, y similares.

En una realización preferida, el ligando para su uso en el método de los solicitantes para modular la expresión de gen es un compuesto de fórmula:



en la que:

E es un alquilo (C<sub>4</sub>-C<sub>6</sub>) que contiene un carbono terciario o un cianoalquilo (C<sub>3</sub>-O<sub>5</sub>) que contiene un carbono terciario;

R<sup>1</sup> es H, Me, Et, i-Pr, F, formilo, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, OH, OMe, OEt, ciclopropilo, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, alilo, azido, SCN o SCHF<sub>2</sub>;

R<sup>2</sup> es H, Me, Et, n-Pr, i-Pr, formilo, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc, NMe<sub>2</sub>, NEt<sub>2</sub>, SMe, SEt, SOCF<sub>3</sub>, OCF<sub>2</sub>CF<sub>2</sub>H, COET, ciclopropilo, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, alilo, azido, OCF<sub>3</sub>, OCHF<sub>2</sub>, O-i-Pr, SCN, SCHF<sub>2</sub>, SOMe, NH-CN, o está unido a R<sup>3</sup> y los carbonos de fenilo a los que están unidos R<sup>2</sup> y R<sup>3</sup> para formar un etilendioxilo, un anillo dihidropirilo con el oxígeno adyacente a un carbono de fenilo, o un anillo dihidropirilo con el oxígeno adyacente a un carbono de fenilo;

R<sup>3</sup> es H, Et, o está unido a R<sup>2</sup> y los carbonos de fenilo a los que están unidos R<sup>2</sup> y R<sup>3</sup> para formar un etilendioxilo, un anillo dihidrofurilo con el oxígeno adyacente a un carbono de fenilo, o un anillo dihidropirilo con el oxígeno adyacente a un carbono de fenilo;

R<sup>4</sup>, R<sup>5</sup> y R<sup>6</sup> son independientemente H, Me, Et, F, Cl, Br, formilo, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, OMe, OEt, SMe o SEt.

En otra realización preferida, el ligando para su uso en el método de los solicitantes para modular la expresión de gen es una ecdisona, 20-hidroxiectdisona, ponasterona A, muristerona A, oxiesterol, un 22(R)-hidroxicolesterol, 24(S)-hidroxicolesterol, 25-epoxicolesterol, T0901317, 5-alfa-6-alfa-epoxicolesterol-3-sulfato (ECHS), 7-cetocolesterol-3-sulfato, farnesol, ácidos biliares, ésteres de 1,1-bifosfonato u hormona juvenil III.

En otra realización preferida, puede usarse un segundo ligando además del primer ligando comentado anteriormente en el método de los solicitantes para modular la expresión de un gen. Preferiblemente, este segundo ligando es ácido 9-cis-retinoico o un análogo sintético de ácido retinoico.

#### Células huésped y organismos no humanos de la invención

Tal como se describió anteriormente, el sistema de modulación de la expresión génica de la presente invención puede usarse para modular la expresión génica en una célula huésped. La expresión en células huésped transgénicas puede ser útil para la expresión de diversos genes de interés. La invención de los solicitantes proporciona modulación de expresión la génica en células huésped eucariotas y procariotas. La expresión en células huésped transgénicas es útil para la expresión de diversos polipéptidos de interés incluyendo pero sin limitarse a antígenos producidos en plantas como vacunas, enzimas como alfa-amilasa, fitasa, glucanos, xilasa y xilanas, genes para resistencia contra insectos, nematodos, hongos, bacterias, virus, y estreses abióticos, antígenos,

5 nutracéuticos, productos farmacéuticos, vitaminas, genes para modificar el contenido de aminoácidos, resistencia a herbicidas, tolerancia al frío, la sequía y el calor, productos industriales, aceites, proteína, hidratos de carbono, antioxidantes, plantas estériles masculinas, flores, combustibles, otros rasgos de producción, polipéptidos terapéuticos, productos intermedios de ruta; para la modulación de rutas ya existentes en el huésped para la síntesis de nuevos productos no posibles hasta ahora usando el huésped; ensayos basados en células; ensayos de genómica funcional, producción de proteínas bioterapéuticas, ensayos proteómicos, y similares. Adicionalmente los productos génicos pueden ser útiles para conferir mayores rendimientos de crecimiento del huésped o para permitir que se utilice un modo de crecimiento alternativo.

10 Por tanto, la invención de los solicitantes proporciona una célula huésped aislada que comprende un sistema de expresión génica según la invención. La presente descripción también proporciona una célula huésped aislada que comprende un casete de expresión génica según la invención. La invención de los solicitantes también proporciona una célula huésped aislada que comprende un polinucleótido o un polipéptido según la descripción. La presente descripción también se refiere a una célula huésped transfectada con un vector de expresión según la descripción.

15 La célula huésped puede ser una célula bacteriana, una célula fúngica, una célula de nematodo, una célula de insecto, una célula de pez, una célula vegetal, una célula aviar, una célula animal o una célula de mamífero. Aún en otra realización, la descripción se refiere a un método para producir un dominio de unión a ligando de receptor nuclear que comprende una mutación por sustitución, en la que el método comprende cultivar la célula huésped tal como se describió anteriormente en medio de cultivo en condiciones que permiten la expresión de un polinucleótido que codifica para el dominio de unión a ligando de receptor nuclear que comprende una mutación por sustitución, y aislar el dominio de unión a ligando de receptor nuclear que comprende una mutación por sustitución a partir del cultivo.

20

En una realización específica, la célula huésped aislada es una célula huésped procariota o una célula huésped eucariota. En otra realización específica, la célula huésped aislada es una célula huésped de invertebrado o una célula huésped de vertebrado. Preferiblemente, la célula huésped se selecciona del grupo que consiste en una célula bacteriana, una célula fúngica, una célula de levadura, una célula de nematodo, una célula de insecto, una célula de pez, una célula vegetal, una célula aviar, una célula animal y una célula de mamífero. Más preferiblemente, la célula huésped es una célula de levadura, una célula de nematodo, una célula de insecto, una célula vegetal, una célula de pez cebra, una célula de pollo, una célula de hámster, una célula de ratón, una célula de rata, una célula de conejo, una célula de gato, una célula de perro, una célula bovina, una célula de cabra, una célula de vaca, una célula de cerdo, una célula de caballo, una célula de oveja, una célula de simio, una célula de mono, una célula de chimpancé o una célula humana. Los ejemplos de células huésped preferidas incluyen, pero no se limitan a, especies fúngicas o de levaduras tales como *Aspergillus*, *Trichoderma*, *Saccharomyces*, *Pichia*, *Candida*, *Hansenula*, o especies bacterianas tales como las de los géneros *Synechocystis*, *Synechococcus*, *Salmonella*, *Bacillus*, *Acinetobacter*, *Rhodococcus*, *Streptomyces*, *Escherichia*, *Pseudomonas*, *Methylobacter*, *Alcaligenes*, *Synechocystis*, *Anabaena*. Thiobacillus, *Methanobacterium* y *Klebsiella*; especies vegetales seleccionadas del grupo que consiste en una manzana, *Arabidopsis*, mijo perla, banana, cebada, judías, remolacha, lenteja negra, garbanzo, chile, pepino, berenjena, habas, maíz, melón, mijo, judía mung, avena, quingombó, *Panicum*, papaya, cacahuete, guisante, pimiento, judía de la India, piña, *Phaseolus*, patata, calabaza, arroz, sorgo, soja, calabacín, caña de azúcar, remolacha azucarera, girasol, boniato, té, tomate, tabaco, sandía y trigo; células huésped animales y de mamíferos.

25

30

35

40

En una realización específica, la célula huésped es una célula de levadura seleccionada del grupo que consiste en una célula huésped de *Saccharomyces*, de *Pichia* y de *Candida*.

45

En otra realización específica, la célula huésped es una célula de nematodo *Caenorhabditis elegans*.

En otra realización específica, la célula huésped es una célula de insecto.

50

En otra realización específica, la célula huésped es una célula vegetal seleccionada del grupo que consiste en una célula de manzana, *Arabidopsis*, mijo perla, banana, cebada, judías, remolacha, lenteja negra, garbanzo, chile, pepino, berenjena, habas, maíz, melón, mijo, judía mung, avena, quingombó, *Panicum*, papaya, cacahuete, guisante, pimiento, judía de la India, piña, *Phaseolus*, patata, calabaza, arroz, sorgo, soja, calabacín, caña de azúcar, remolacha azucarera, girasol, boniato, té, tomate, tabaco, sandía y trigo.

55

En otra realización específica, la célula huésped es una célula de pez cebra.

En otra realización específica, la célula huésped es una célula de pollo.

60

En otra realización específica, la célula huésped es una célula de mamíferos seleccionada del grupo que consiste en una célula de hámster, una célula de ratón, una célula de rata, una célula de conejo, una célula de gato, una célula de perro, una célula bovina, una célula de cabra, una célula de vaca, una célula de cerdo, una célula de caballo, una célula de oveja, una célula de mono, una célula de chimpancé y una célula humana.

65

La transformación de células huésped se conoce bien en la técnica y puede lograrse mediante una variedad de

métodos incluyendo pero sin limitarse a electroporación, infección viral, transfección de plásmido/vector, transfección mediada por vector no viral, transformación mediada por *Agrobacterium*, bombardeo de partículas, y similares. La expresión de productos génicos deseados implica cultivar las células huésped transformadas en condiciones adecuadas e inducir la expresión del gen transformado. Se conocen bien en la técnica las condiciones de cultivo y los protocolos de expresión génica en células procariotas y eucariotas (véase la sección de métodos generales de los ejemplos). Pueden recogerse las células y los productos génicos aislados según los protocolos específicos para el producto génico.

Además, puede elegirse una célula huésped que modula la expresión del polinucleótido insertado, o modifica y procesa el producto de polipéptido de modo específico deseado. Las células huésped diferentes tienen mecanismos característicos y específicos para el procesamiento traduccional y post-traduccional y la modificación [por ejemplo, glicosilación, escisión (por ejemplo, de secuencia señal)] de proteínas. Las líneas celulares o sistemas de huésped apropiados pueden elegirse para garantizar la modificación deseada y el procesamiento de la proteína foránea expresada. Por ejemplo, la expresión en un sistema bacteriano puede usarse para producir un producto de proteína de núcleo no glicosilado. Sin embargo, un polipéptido expresado en bacterias puede no estar plegado de manera apropiada. La expresión en levadura puede producir un producto glicosilado. La expresión en células eucariotas puede aumentar la probabilidad de glicosilación "nativa" y plegado de una proteína heteróloga. Además, la expresión en células de mamíferos puede proporcionar una herramienta para reconstituir, o constituir, la actividad del polipéptido. Además, diferentes sistemas de expresión de huésped/vector pueden afectar a las reacciones de procesamiento, tales como escisiones proteolíticas, a un grado diferente.

La invención de los solicitantes también se refiere a un organismo no humano que comprende una célula huésped aislada según la invención. En una realización específica, el organismo no humano es un organismo procariota o un organismo eucariota. En otra realización específica, el organismo no humano es un organismo de invertebrado o un organismo de vertebrado.

Preferiblemente, el organismo no humano se selecciona del grupo que consiste en una bacteria, un hongo, una levadura, un nematodo, un insecto, un pez, una planta, un pájaro, un animal y un mamífero. Más preferiblemente, el organismo no humano es una levadura, un nematodo, un insecto, una planta, un pez cebra, un pollo, un hámster, un ratón, una rata, un conejo, un gato, un perro, un bovino, una cabra, una vaca, un cerdo, un caballo, una oveja, un simio, un mono o un chimpancé.

En una realización específica, el organismo no humano es una levadura seleccionada del grupo que consiste en *Saccharomyces*, *Pichia* y *Candida*.

En otra realización específica, el organismo no humano es un nematodo *Caenorhabdus elegans*.

En otra realización específica, el organismo no humano es una planta seleccionada del grupo que consiste en una manzana, *Arabidopsis*, mijo perla, banana, cebada, judías, remolacha, lenteja negra, garbanzo, chile, pepino, berenjena, habas, maíz, melón, mijo, judía mung, avena, quingombó, *Panicum*, papaya, cacahuete, guisante, pimiento, judía de la India, piña, *Phaseolus*, patata, calabaza, arroz, sorgo, soja, calabacín, caña de azúcar, remolacha azucarera, girasol, boniato, té, tomate, tabaco, sandía, y trigo.

En otra realización específica, el organismo no humano es un ratón *Mus musculus*.

#### Medición de la expresión/transcripción génica

Una medición útil de los métodos de los solicitantes de la invención es la del estado transcripcional de la célula incluyendo las identidades y abundancias de ARN, preferentemente especies de ARNm. Tales mediciones se realizan convenientemente midiendo las abundancias de ADNc mediante cualquiera de varias tecnologías de expresión génica existentes.

La tecnología de matriz de ácido nucleico es una técnica útil para determinar la expresión de ARNm diferencial. Tal tecnología incluye, por ejemplo, chips de oligonucleótidos y micromatrices de ADN. Estas técnicas se basan en fragmentos de ADN u oligonucleótidos que corresponden a diferentes genes o ADNc que se inmovilizan en un soporte sólido y se hibridan con sondas preparadas a partir de combinaciones de ARNm total extraídas de células, tejidos u organismos completos y se convierten en ADNc. Los chips de oligonucleótidos son matrices de oligonucleótidos sintetizados en un sustrato usando técnicas fotolitográficas. Se han producido chips que pueden analizar hasta 1700 genes. Las micromatrices de ADN son matrices de muestras de ADN, normalmente productos de PCR, que se imprimen de forma robótica en un portaobjetos de microscopio. Cada gen se analiza por una secuencia de ADN diana de longitud completa o parcial. Ahora se preparan comercialmente de manera rutinaria micromatrices con hasta 10.000 genes. La diferencia principal entre estas dos técnicas es que los chips de oligonucleótidos normalmente utilizan oligonucleótidos de 25 meros lo que permite el fraccionamiento de moléculas de ADN cortas mientras que las dianas de ADN mayores de micromatrices, de aproximadamente 1000 pares de bases, pueden proporcionar más sensibilidad en el fraccionamiento de mezclas de ADN complejas.

Otra medición útil de los métodos de los solicitantes es la de determinar el estado de traducción de la célula midiendo las abundancias de las especies de proteínas constituyentes presentes en la célula usando procedimientos bien conocidos en la técnica.

5 Cuando se desea identificación de genes asociados con diversas funciones fisiológicas, puede emplearse un ensayo en el que se miden cambios en funciones tales como crecimiento celular, apoptosis, senescencia, diferenciación, adhesión, unión a moléculas específicas, unión a otra célula, organización celular, organogénesis, transporte intracelular, facilitación del transporte, conversión de energía, metabolismo, miogénesis, neurogénesis y/o hematopoyesis.

10 Además, puede usarse la expresión de gen marcador o indicador seleccionable para medir la modulación de la expresión génica usando la invención de los solicitantes.

15 Otros métodos para detectar los productos de expresión génica se conocen bien en la técnica e incluyen transferencias de tipo Southern (detección de ADN), transferencias puntuales o por ranuras (ADN, ARN), transferencias de tipo Northern (ARN), RT-PCR (ARN), inmunotransferencias de tipo Western (detección de polipéptidos) y análisis de ELISA (polipéptido). Aunque se prefieren menos, pueden usarse proteínas marcadas para detectar una secuencia de ácido nucleico particular con la cual se hibridan.

20 En algunos casos es necesario amplificar la cantidad de una secuencia de ácido nucleico. Esto puede llevarse a cabo usando uno o más de varios métodos adecuados incluyendo, por ejemplo, reacción en cadena de la polimerasa ("PCR"), reacción en cadena de la ligasa ("LCR"), amplificación por desplazamiento de hebra ("SDA"), amplificación basada en transcripción y similares. La PCR se lleva a cabo según técnicas conocidas en las que, por ejemplo, se trata una muestra de ácido nucleico en presencia de una ADN polimerasa estable al calor, en condiciones de hibridación, con un par de cebadores de oligonucleótidos, hibridándose un cebador con una hebra (molde) de la secuencia específica que va a detectarse. Los cebadores son suficientemente complementarios a cada hebra de molde de la secuencia específica como para hibridarse con la misma. Se sintetiza un producto de extensión de cada cebador y es complementario a la hebra de molde de ácido nucleico con la que se hibrida. El producto de extensión sintetizado a partir de cada cebador también puede servir como un molde para la síntesis adicional de productos de extensión usando los mismos cebadores. Tras un número suficiente de ciclos de síntesis de productos de extensión, la muestra puede analizarse tal como se describió anteriormente para evaluar si la secuencia o secuencias que van a detectarse están presentes.

#### 35 Ensayos de examen de ligandos

La presente descripción también se refiere a métodos para seleccionar un compuesto que induce o reprime la transactivación de un dominio de unión a ligando de receptor nuclear que comprende una mutación por sustitución en una célula poniendo en contacto un dominio de unión a ligando de receptor nuclear con una molécula candidata y detectando la actividad de gen indicador en presencia del ligando. Los compuestos candidatos pueden ser o bien agonistas o bien antagonistas del dominio de unión a ligando de receptor nuclear. En una realización preferida, el dominio de unión a ligando de receptor nuclear se expresa a partir de un polinucleótido en la célula y se mide la actividad de transactivación (es decir, expresión o represión de un gen indicador) o actividad de unión al compuesto.

45 Por consiguiente, además del diseño racional de agonistas y antagonistas basándose en la estructura de un dominio de unión a ligando de receptor nuclear, la presente descripción contempla un método alternativo para identificar ligandos específicos de un dominio de unión a ligando del receptor nuclear usando diversos ensayos de selección conocidos en la técnica.

50 Puede usarse cualquier técnica de selección conocida en la técnica para examinar agonistas o antagonistas del dominio de unión a ligando de receptor nuclear de grupo H. Por ejemplo, una línea celular adecuada que comprende un sistema de expresión génica basado en receptores nucleares según la descripción puede transfectarse con un casete de expresión génica que codifica para un gen marcador operativamente unido a un promotor inducible o reprimible. Entonces se exponen las células transfectadas a una disolución de prueba que comprende un compuesto agonista o antagonista candidato, y entonces se somete a ensayo para detectar la expresión o represión del gen marcador. La presencia de más expresión de gen marcador con respecto a células control no expuestas a la disolución de prueba es una indicación de la presencia de un compuesto agonista en la disolución de prueba. Por el contrario, la presencia de menos expresión de gen marcador con respecto a células de control no expuestas a la disolución de prueba es una indicación de la presencia de un compuesto antagonista en la disolución de prueba.

60 La presente descripción contempla exámenes de ligandos de molécula pequeña o análogos de ligando y miméticos, así como exámenes de ligandos naturales que se unen a y agonizan o antagonizan un dominio de unión a ligando de receptor nuclear de grupo H según la descripción *in vivo*. Por ejemplo, pueden examinarse bibliotecas de productos naturales usando ensayos de la descripción para seleccionar moléculas que agonizan o antagonizan la actividad del sistema de expresión génica basado en receptores nucleares.

65 La identificación y el examen de antagonistas se facilitan adicionalmente determinando las características

estructurales de la proteína, por ejemplo, usando cristalografía de rayos X, difracción de neutrones, espectrometría de resonancia magnética nuclear y otras técnicas para la determinación de estructura. Estas técnicas proporcionan el diseño racional o la identificación de agonistas y antagonistas.

5 Otro enfoque usa bacteriófagos recombinantes para producir bibliotecas grandes. Usando el "método de fagos" [Scott y Smith, 1990, *Science* 249: 386-390 (1990); Cwirla, *et al.*, *Proc. Natl. Acad. Sci.*, 87: 6378-6382 (1990); Devlin *et al.*, *Science*. 249: 404-406 (1990)], pueden construirse bibliotecas muy grandes ( $10^6$ - $10^8$  entidades químicas). Un segundo enfoque usa principalmente métodos químicos, de los cuales son ejemplos el método de Geysen [Geysen *et al.*, *Molecular Immunology* 23: 709-715 (1986); Geysen *et al.* *J. Immunologic Method* 102: 259-274 (1987)] y el método de Fodor *et al.* [*Science* 251: 767-773 (1991)]. Furka *et al.* [14th International Congress of Biochemistry, volumen 5, resumen FR:013 (1988); Furka, *Int. J. Peptide Protein Res.* 37:487-493 (1991)], Houghton [patente de EE.UU. n.º 4.631.211, presentada en diciembre de 1986] y Rutter *et al.* [patente de EE.UU. n.º 5.010.175, presentada el 23 de abril de 1991] describen métodos para producir una mezcla de péptidos que pueden someterse a prueba como agonistas o antagonistas.

15 En otro aspecto, pueden usarse bibliotecas sintéticas [Needels *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 10700-4 (1993); Ohlmeyer *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 10922-10926 (1993); Lam *et al.*, publicación internacional de patente n.º WO 92/00252; Kocis *et al.*, publicación internacional de patente n.º WO 9428028, y similares para seleccionar ligandos candidatos según la presente descripción.

20 Puede realizarse la selección con células recombinantes que expresan un dominio de unión a ligando de receptor nuclear según la descripción, o alternativamente, usando proteína purificada, por ejemplo, producida de manera recombinante, tal como se describió anteriormente. Por ejemplo, pueden usarse dominios de unión a ligando de receptor nuclear solubles marcados para examinar bibliotecas, tal como se describe en las referencias anteriores.

25 En una realización, un dominio de unión a ligando de receptor nuclear de grupo H según la descripción puede marcarse directamente. En otra realización, puede usarse un reactivo secundario marcado para detectar la unión de un dominio de unión a ligando de receptor nuclear de la descripción a una molécula de interés, por ejemplo, una molécula unida a un soporte de fase sólida. Puede detectarse la unión mediante la formación *in situ* de un cromóforo mediante un marcador enzimático. Las enzimas adecuadas incluyen, pero no se limitan a, fosfatasa alcalina y peroxidasas del rábano. En una realización adicional, puede usarse un ensayo de dos colores, usando dos sustratos cromogénicos con dos marcadores enzimáticos en diferentes moléculas aceptoras de interés. Pueden identificarse ligandos de reacción cruzada y reacción sencilla con un ensayo de dos colores.

30 Otros marcadores para su uso en la descripción incluyen perlas de látex coloreadas, perlas magnéticas, marcadores fluorescentes (por ejemplo, isotiocianato de fluoresceína (FITC), ficoeritrina (PE), Texas red (TR), rodamina, sales de la serie de lantánidos libres o queladas, especialmente  $\text{Eu}^{3+}$ , por citar algunos fluoróforos), moléculas quimioluminiscentes, radioisótopos o marcadores de formación de imágenes de resonancia magnética. Pueden realizarse ensayos de dos colores con dos o más perlas de látex coloreadas, o fluoróforos que emiten a diferentes longitudes de onda. Moléculas o células marcadas pueden detectarse visualmente o mediante medios mecánicos/ópticos. Los medios mecánicos/ópticos incluyen clasificación activada por fluorescencia, es decir, análogos a FACS, y medios de retirada de micromanipulador.

35 La presente invención puede entenderse mejor con referencia a los siguientes ejemplos no limitativos, que se proporcionan a modo de ejemplo de la invención.

### Ejemplos

40 Los solicitantes han desarrollado un modelo de homología de CfEcR y han usado este modelo de homología junto con un modelo de homología de receptor de ecdisona de *Chironomous tetans* ("CtEcR") publicado (Wurtz *et al.*, 2000) para identificar residuos críticos implicados en la unión a esteroides y compuestos no esteroideos. Se ha mostrado que los compuestos no esteroideos sintéticos, diacilhidrazinas, se unen a EcR de lepidópteros con alta afinidad e inducen muda incompleta precoz en estos insectos (Wing *et al.*, 1988) y varios de estos compuestos se comercializan actualmente como insecticidas. La cavidad de unión a ligando de EcR ha evolucionado para adaptarse a las estructuras de esqueleto largas de ecdiesteroides tales como 20E. Las diacilhidrazinas tienen una estructura compacta en comparación con esteroides y sólo ocupan la parte inferior de la bolsa de unión a EcR. Esto deja pocos residuos críticos en la parte superior de la bolsa de unión que entra en contacto con esteroides pero no con compuestos no esteroideos tales como diacilhidrazinas. Los solicitantes realizaron mutaciones por sustitución de los residuos que entran en contacto con esteroides y/o compuestos no esteroideos y determinaron el efecto de la mutación sobre la unión a ligando. Los solicitantes describen en el presente documento mutaciones por sustitución en varios de estos residuos y han identificado varias clases de receptores mutantes por sustitución basándose en sus características de unión y transactivación. Los polinucleótidos y polipéptidos de receptor nuclear mutado por sustitución novedosos de los solicitantes son útiles en un sistema de modulación génica inducible basado en receptores nucleares para diversas aplicaciones incluyendo terapia génica, expresión de proteínas de interés en células huésped, producción de organismos transgénicos y ensayos basados en células.

MÉTODOS GENERALES

- Se conocen bien en la técnica técnicas de clonación molecular y ADN recombinante convencionales usadas en el presente documento y se describen por Sambrook, J., Fritsch, E. F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y. (1989) (Maniatis) y por T. J. Silhavy, M. L. Bannan, y L. W. Enquist, *Experiments with Gen Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1984) y por Ausubel, F. M. *et al.*, *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley-Interscience (1987).
- Se conocen bien en la técnica materiales y métodos adecuados para el mantenimiento y crecimiento de cultivos bacterianos. Pueden encontrarse técnicas adecuadas para su uso en los siguientes ejemplos tal como se expone en *Manual of Methods for General Bacteriology* (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds), American Society for Microbiology, Washington, DC. (1994)) o en Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, segunda edición, Sinauer Associates, Inc., Sunderland, MA (1989). Todos los reactivos, enzimas de restricción y materiales usados para el crecimiento y mantenimiento de células huésped se obtuvieron de Aldrich Chemicals (Milwaukee, WI), DIFCO Laboratories (Detroit, MI), GIBCO/BRL (Gaithersburg, MD), o Sigma Chemical Company (St. Louis, MO) a menos que se especifique lo contrario.
- Pueden realizarse manipulaciones de secuencias genéticas usando la serie de programas disponible de Genetics Computer Group Inc. (Wisconsin Package versión 9.0, Genetics Computer Group (GCG), Madison, WI). Cuando se usa el programa de GCG "Pileup" pueden usarse el valor por defecto de creación de hueco de 12, y el valor por defecto de extensión de hueco de 4. Cuando se usa el programa de GCG "Gap" o "Bestfit" pueden usarse la penalización de creación de hueco por defecto de 50 y la penalización de extensión de hueco por defecto de 3. En cualquier caso, cuando no se piden parámetros del programa de GCG, en estos o cualquier otro programa de GCG, pueden usarse valores por defecto.

El significado de abreviaturas es tal como sigue: "h" significa hora(s), "min." significa minuto(s), "s" significa segundo(s), "d" significa día(s), "µl" significa microlitro(s), "ml" significa mililitro(s), "l" significa litro(s), "µM" significa micromolar, "mM" significa milimolar, "µg" significa microgramo(s), "mg" significa miligramo(s), "A" significa adenina o adenosina, "T" significa timina o timidina, "G" significa guanina o guanosa, "C" significa citidina o citosina, "xg" significa veces de gravedad, "nt" significa nucleótido(s), "aa" significa aminoácido(s), "pb" significa par(es) de bases, "kb" significa kilobase(s), "k" significa kilo, "µ" significa micro, y "°C" significa grados Celsius.

## EJEMPLO 1

Este ejemplo describe la construcción de varios casetes de expresión génica que comprenden polinucleótidos y polipéptidos de receptor nuclear de grupo H mutado por sustitución novedosos de la invención para su uso en un sistema de expresión génica inducible basado en receptores nucleares. Los solicitantes construyeron varios casetes de expresión génica basados en el EcR de tortrix de las yemas de la picea *Choristoneura fumiferana* ("CfEcR"), EcR de mosca de la fruta *Drosophila melanogaster* ("DmEcR"), EcR de garrapata solitaria *Amblyomma americanum* ("AmaEcR"), proteína ultraespiráculo de langosta *Locusta migratoria* ("LmUSP"), un homólogo de RXR de invertebrado de RXR de vertebrado, y USP de *C. fumiferana* ("CfUSP"). Los constructos receptores preparados comprenden un dominio de unión a un ligando de un EcR, una USP de invertebrado o un RXR de invertebrado; y un dominio de unión a ADN (DBD) de GAL4 o un dominio de transactivación (AD) de VP16. Los constructos indicadores incluyen un gen indicador, luciferasa o LacZ ( $\beta$ -galactosidasa), operativamente unido a un constructo promotor sintético que comprende un elemento de respuesta a GAL4 al cual se une el DBD de Gal4. Se cotransfectaron diversas combinaciones de estos constructos receptor e indicador en células de mamífero tal como se describe en los ejemplos 2-10 a continuación.

Casetes de expresión génica: Se construyeron los casetes de expresión génica (interruptores) basados en receptor de ecdisona tal como sigue, usando métodos de clonación convencionales disponibles en la técnica. Lo siguiente es una breve descripción de la preparación y la composición de cada interruptor usado en los ejemplos descritos en el presente documento.

1.1 - GAL4CfEcR-DEF/VP16LmUSP-EF: Se fusionaron los dominios de D, E y F de tipo natural de EcR de tortrix de las yemas de la picea *Choristoneura fumiferana* ("CfEcR-DEF"; SEQ ID NO: 21) con un dominio de unión a ADN de GAL4 ("Gal4DNABD" o "Gal4DBD"; SEQ ID NO: 6) y se colocaron bajo el control de un promotor de SV40e (SEQ ID NO: 22). Se fusionaron los dominios E y F de la proteína ultraespiráculo de langosta *Locusta migratoria* ("LmUSP-EF"; SEQ ID NO: 23) con el dominio de transactivación de VP16 ("VP16AD"; SEQ ID NO: 12) y se colocaron bajo el control de un promotor de SV40e (SEQ ID NO: 22). Se fusionaron cinco sitios de unión a elemento de respuesta a GAL4 consenso ("5XGAL4RE"; que comprende 5 copias de un GAL4RE que comprende SEQ ID NO: 19) con un promotor mínimo de E1b sintético (SEQ ID NO: 24) y se colocaron en el sentido de 5' del gen de la luciferasa (SEQ ID NO: 25).

1.2-GAL4/CfEcR-DEF mutante/VP16LmUSP-EF: Se preparó este constructo de la misma manera que en el interruptor 1.1 anterior excepto porque se reemplazó CfEcR-DEF de tipo natural por CfEcR-DEF mutante que comprende un dominio de unión a ligando que comprende una mutación por sustitución seleccionada de la tabla 1 a continuación.

5

Tabla 1. Mutantes por sustitución de dominio de unión a ligando (LBD) de receptor de ecdisona de *Choristoneura fumiferana* ("CfEcR").

<b>Mutación de LBD de CfEcR-DEF</b>	<b>Sustitución de aminoácido "de tipo natural a mutante" resultante</b>	<b>Aminoácido correspondiente en CfEcR de longitud completa (SEQ ID NO: 26)</b>
E20A	De ácido glutámico (E) a alanina (A)	303
Q21 A	De glutamina (Q) a alanina (A)	304
F48A	De fenilalanina (F) a alanina (A)	331
I51A	De isoleucina (I) a alanina (A)	334
T52A	De treonina (T) a alanina (A)	335
T52L	De treonina (T) a leucina (L)	335
T52V	De treonina (T) a valina (V)	335
T52I	De treonina (T) a isoleucina (I)	335
T55A	De treonina (T) a alanina (A)	338
T58A	De treonina (T) a alanina (A)	341
V59A	De valina (V) a alanina (A)	342
L61A	De leucina (L) a alanina (A)	344
I62A	De isoleucina (I) a alanina (A)	345
M92A	De metionina (M) a alanina (A)	375
M93A	De metionina (M) a alanina (A)	376
R95A	De arginina (R) a alanina (A)	378
V96A	De valina (V) a alanina (A)	379
V96T	De valina (V) a treonina (T)	379
V96D	De valina (V) a ácido aspártico (D)	379
V96M	De valina (V) a metionina (M)	379
V107I	De valina (V) a isoleucina	390
F109A	De fenilalanina (F) a alanina (A)	392
A110P	De alanina (A) a prolina (P)	393
A110S	De alanina (A) a serina (S)	393
A110L	De alanina (A) a leucina (L)	393
A110M	De alanina (A) a metionina (M)	393
Y120A	De tirosina (Y) a alanina (A)	403
A123F	De alanina (A) a fenilalanina (F)	406
M125A	De metionina (M) a alanina (A)	408
R175E	De arginina (R) a glutamina (E)	458
M218A	De metionina (M) a alanina (A)	501
C219A	De cisteína (C) a alanina (A)	502
L223A	De leucina (L) a alanina (A)	506
L230A	De leucina (L) a alanina (A)	513

L234A	De leucina (L) a alanina (A)	517
W238A	De triptófano (W) a alanina (A)	521
Mutante doble R95A y A110P	De arginina (R) a alanina (A) y de alanina (A) a prolina (P), respectivamente	378 y 393, respectivamente
Mutante doble V107I y R175E	De valina (V) a isoleucina (I) y de arginina (R) a glutamina (E), respectivamente	390 y 458, respectivamente
Mutante doble V107E e Y127E	De valina (V) a isoleucina (I) y de tirosina (Y) a glutamina (E), respectivamente	390 y 410, respectivamente
Mutante doble Y127E y R175E	De tirosina (Y) a glutamina (E) y de arginina (R) a glutamina (E), respectivamente	410 y 458, respectivamente
Mutante doble M218A y C219A	De metionina (M) a alanina (A) y de cisteína (C) a alanina (A), respectivamente	501 y 502, respectivamente
Mutante triple T52V, V107I y R175E	De treonina (T) a valina (V), de valina (V) a isoleucina (I) y de arginina (R) a glutamina (E), respectivamente	335, 390 y 458, respectivamente
Mutante triple T52A, V107I y R175E	De treonina (T) a alanina (A), de valina (V) a isoleucina (I) y de arginina (R) a glutamina (E), respectivamente	335, 390 y 458, respectivamente
Mutante triple V96A, V107I y R175E	De valina (V) a alanina (A), de valina (V) a isoleucina (I) y de arginina (R) a glutamina (E), respectivamente	379, 390 y 458, respectivamente
Mutante triple V96T, V107I y R175E	De valina (V) a treonina (T), de valina (V) a isoleucina (I) y de arginina (R) a glutamina (E), respectivamente	379, 390 y 458, respectivamente
Mutante triple V107I, Y127E y R175E	De valina (V) a isoleucina (I), de tirosina (Y) a glutamina (E) y de arginina (R) a glutamina (E), respectivamente	390, 410 y 458, respectivamente
Mutante triple V107I, A110P y R175E	De valina (V) a isoleucina (I), de alanina (A) a prolina (P) y de arginina (R) a glutamina (E), respectivamente	390, 393, y 458, respectivamente

1.3 - GAL4CfEcR-ABCDEF/VP16LmUSP-EF: Se fusionó el EcR de tortrix de las yemas de la picea *Choristoneura fumiferana* de longitud completa ("CfEcRABCDEF"; SEQ ID NO: 27) con un dominio de unión a ADN de GAL4 ("Gal4DNABD" o "Gal4DBD"; SEQ ID NO: 6) y se colocó bajo el control de un promotor de SV40e (SEQ ID NO: 22). Se fusionaron los dominios E y F de ultraespiráculo de *Locusta migratoria* ("LmUSP-EF"; SEQ ID NO: 23) con el dominio de transactivación de VP16 ("VP16AD"; SEQ ID NO: 12) y se colocaron bajo el control de un promotor de SV40e (SEQ ID NO: 22). Se fusionaron cinco sitios de unión a elemento de respuesta a GAL4 consenso ("5XGAL4RE"; que comprenden 5 copias de un RE a GAL4 que comprende SEQ ID NO: 19) con un promotor mínimo de E1b sintético (SEQ ID NO: 24) y se colocaron en el sentido de 5' del gen de la luciferasa (SEQ ID NO: 25).

1.4 - GAL4/CfEcR-ABCDEF mutante A110P/VP16LmUSP-EF: Se preparó este constructo de la misma manera que en el interruptor 1.3 anterior excepto porque se reemplazó CfEcR-ABCDEF de tipo natural por CfEcR-ABCDEF mutante que comprende un dominio de unión a ligando que comprende una mutación por sustitución de A110P tal como se describió en la tabla 1 anteriormente.

1.5 - VP16/CfEcR-CDEF: Se preparó este constructo de la misma manera que el interruptor 1.1 excepto porque se reemplazó el dominio de unión a ADN de GAL4 por el dominio de transactivación de VP16 ("VP16AD"; SEQ ID NO: 12) y se colocó bajo el control de un promotor IE1 de baculovirus (SEQ ID NO: 28). Se fusionaron seis sitios de unión a elemento de respuesta a ecdisona consenso ("6XEcRE"; que comprende 6 copias de un RE a ecdisona que comprende SEQ ID NO: 18) con un promotor mínimo de E1b sintético (SEQ ID NO: 24) y se colocaron en el sentido de 5' del gen de la  $\beta$ -galactosidasa (SEQ ID NO: 29). Este constructo usa proteína ultraespiráculo endógena como componente de heterodimerización.

1.6 - VP16/CfEcR-CDEF mutante A110P: Se preparó este constructo de la misma manera que en el interruptor 1.5 anterior excepto porque se reemplazó CfEcR-CDEF de tipo natural por CfEcR-CDEF mutante que comprende un dominio de unión a ligando que comprende una mutación por sustitución de A110P tal como se describió en la tabla 1 anteriormente.



1.7 - CfUSP-A/BCDEF expresado por bacterias: Se preparó este constructo con los dominios A/BCDEF de USP de tortrix de las yemas de la picea de *C. fumiferana* ("CfUSP-A/BCDEF"; SEQ ID NO: 30).

5 1.8 - GAL4/DmEcR-CDEF/VP16LmUSP-EF: Se fusionaron los dominios C, D, E y F de tipo natural de EcR de la mosca de la fruta *Drosophila melanogaster* ("DmEcR-CDEF"; SEQ ID NO: 31) con un dominio de unión a ADN de GAL4 ("Gal4DNABD" o "Gal4DBD"; SEQ ID NO: 6) y se colocaron bajo el control de un promotor de SV40e (SEQ ID NO: 22). Se fusionaron los dominios de E y F de la proteína ultraespiráculo de langosta *Locusta migratoria* ("LmUSP-EF"; SEQ ID NO: 23) con el dominio de transactivación de VP16 ("VP16AD"; SEQ ID NO: 12) y se colocaron bajo el control de un promotor de SV40e (SEQ ID NO: 22). Se fusionaron cinco sitios de unión a elemento de respuesta a GAL4 consenso ("5XGAL4RE"; que comprende 5 copias de un GAL4RE que comprende SEQ ID NO: 19) con un promotor mínimo de E1b sintético (SEQ ID NO: 24) y se colocaron en el sentido de 5' del gen de la luciferasa (SEQ ID NO: 25).

15 1.9 - GAL4/DmEcR-CDEF mutante/VP16LmUSP-EF: Se preparó este constructo de la misma manera que en el interruptor 1.8 anterior excepto porque se reemplazó DmEcR-CDEF de tipo natural por DmEcR-CDEF mutante que comprende un dominio de unión a ligando que comprende una mutación por sustitución seleccionada de la tabla 2 a continuación.

20 Tabla 2. Mutantes por sustitución de dominio de unión a ligando (LBD) de receptor de ecdisona de *Drosophila melanogaster* ("DmEcR").

<b>Mutación de LBD de DmEcR-CDEF</b>	<b>Sustitución de aminoácido "de tipo natural a mutante" resultante</b>	<b>Aminoácido correspondiente en DmEcR de longitud completa (SEQ ID NO: 32)</b>
A107P	De alanina (A) a prolina (P)	522
G121R	De glicina (G) a arginina (R)	536
G121L	De glicina (G) a leucina (L)	536
N213A	De asparagina (N) a alanina (A)	628
C217A	De cisteína (C) a alanina (A)	632
C217S	De cisteína (C) a serina (S)	632

25 1.10 - VP16/DmEcR-CDEF: Se preparó este constructo de la misma manera que el interruptor 1.8 excepto porque se reemplazó el dominio de unión a ADN de GAL4 por el dominio de transactivación de VP16 ("VP16AD"; SEQ ID NO: 12) y se colocó bajo el control de un promotor IE1 de baculovirus (SEQ ID NO: 28). Se fusionaron seis sitios de unión a elemento de respuesta a ecdisona consenso ("6XEcRE"; que comprende 6 copias de un RE a ecdisona que comprende SEQ ID NO: 18) en un promotor mínimo de E1b sintético (SEQ ID NO: 24) y se colocaron en el sentido de 5' del gen de la  $\beta$ -galactosidasa (SEQ ID NO: 29). Este constructo usa proteína ultraespiráculo endógena como componente de heterodimerización.

30 1.11 - VP16/DmEcR-CDEF mutante: Se preparó este constructo de la misma manera que en el interruptor 1.10 anterior excepto porque se reemplazó DmEcR-CDEF de tipo natural por DmEcR-CDEF mutante que comprende un dominio de unión a ligando que comprende una mutación por sustitución seleccionada de la tabla 2 anteriormente.

35 1.12 - GAL4/AmaEcR-DEFNP16LmUSP-EF: Se fusionaron los dominios D, E y F de tipo natural de EcR de garrapata solitaria *Amblyomma americanum* ("AmaEcR-DEF"; SEQ ID NO: 33) con un dominio de unión a ADN de GAL4 ("Gal4DNABD" o "Gal4DBD"; SEQ ID NO: 6) y se colocaron bajo el control de un promotor de SV40e (SEQ ID NO: 22). Se fusionaron los dominios de E y F de la proteína ultraespiráculo de langosta *Locusta migratoria* ("LmUSP-EF"; SEQ ID NO: 23) con el dominio de transactivación de VP16 ("VP16AD"; SEQ ID NO: 12) y se colocaron bajo el control de un promotor de SV40e (SEQ ID NO: 22). Se fusionaron cinco sitios de unión a elemento de respuesta a GAL4 consenso ("5XGAL4RE"; que comprende 5 copias de un GAL4RE que comprende SEQ ID NO: 19) con un promotor mínimo de E1b sintético (SEQ ID NO: 24) y se colocaron en el sentido de 5' del gen de la luciferasa (SEQ ID NO: 25).

45 1.13 - GAL4/AmaEcR-DEF mutante/VP16LmUSP-EF: Se preparó este constructo de la misma manera que en el interruptor 1.12 anterior excepto porque se reemplazó AmaEcR-DEF de tipo natural por AmaEcR-DEF mutante que comprende un dominio de unión a ligando que comprende una mutación por sustitución seleccionada de la tabla 3 a continuación.

Tabla 3. Mutantes por sustitución de dominio de unión a ligando (LBD) de receptor de ecdisona de *Amblyomma americanum* ("AmaEcR").

Mutación de LBD de AmaEcR-DEF	Sustitución de aminoácido "de tipo natural a mutante" resultante	Aminoácido correspondiente en AmaEcR de longitud completa (SEQ ID NO: 34)
G91A	De glicina (G) a alanina (A)	417
A105P	De alanina (A) a prolina (P)	431

5 Construcción de dominios de unión a ligando de receptor de ecdisona que comprenden una mutación por sustitución:

10 En un intento de modificar la unión a ligando de EcR, se mutaron residuos dentro de los dominios de unión a ligando de EcR que se predijo que eran importantes para la unión a ligando basándose en un análisis de modelización molecular en EcR de tres clases diferentes de organismos. Las tablas 1-3 indican los residuos de aminoácido dentro del dominio de unión a ligando de CfEcR (EcR de lepidóptero), DmEcR (EcR de dípteros) y AmaEcR (EcR de artrópodos), respectivamente que se mutaron y se examinaron para determinar la modificación de la unión a esteroides y compuestos no esteroideos.

15 Se construyó cada una de las mutaciones por sustitución de aminoácido indicadas en las tablas 1-3 en un ADNc de EcR mediante mutagénesis dirigida a sitio mediada por PCR a alanina (o a prolina o fenilalanina en el caso de un residuo de alanina de tipo natural, por ejemplo residuos A110 y A123 de CfEcR, respectivamente). Se mutaron los aminoácidos T52, V96 y A110 de CfEcR a cuatro aminoácidos diferentes. También se prepararon cinco CfEcR mutantes puntuales dobles diferentes: uno que comprendía las sustituciones tanto R95A como A110P (R95A + A110P, "DM"), un segundo que comprendía las sustituciones tanto M218A como C219A (M218A + C219A), un tercero que comprendía las sustituciones tanto V107I como R175E (V107I + R175E), un cuarto que comprendía las sustituciones Y127E y R175E (Y127E + R175E), y un quinto que comprendía las sustituciones V107I e Y127E (V107I + Y127E). También se prepararon seis CfEcR mutantes puntuales triples diferentes: uno que comprendía las sustituciones tanto V107I como R175E y una sustitución Y127E (V107I + Y127E + R175E), un segundo que comprendía una sustitución T52V y las sustituciones V107I y R175E (T52V + V107I + R175E), un tercero que comprendía las sustituciones V96A, V107I y R175E (V96A + V107I + R175E), un cuarto que comprendía las sustituciones T52A, V107I y R175E (T52A + V107I + R175E), un quinto que comprendía una sustitución V96T y las sustituciones V107I y R175E (V96T + V107I + R175E), y un sexto que comprendía las sustituciones A110P, V107I y R175E.

30 Se realizó mutagénesis dirigida a sitio por PCR usando el kit de mutagénesis dirigida a sitio de Quikchange (Stratagene, La Jolla, CA) usando las siguientes condiciones de reacción y parámetros de ciclación. Se realizó mutagénesis dirigida a sitio por PCR usando tampón de reacción 1x (suministrado por el fabricante), 50 ng de molde de ADNbc, 125 ng de cebador directo (FP), 125 ng de cebador complementario inverso (RCP), y 1 µl de mezcla de dNTP (suministrada por el fabricante) en un volumen de reacción final de 50 µl. En las tablas 4-6 se presentan el cebador directo y el cebador complementario inverso usados para producir cada mutante de EcR. Los parámetros de ciclación usados consistieron en un ciclo de desnaturalización a 95°C durante 30 segundos, seguido por 16 ciclos de desnaturalización a 95°C durante 30 segundos, apareamiento a 55°C durante 1 minuto, y extensión a 68°C durante 22 minutos.

40 Tabla 4: Cebadores de PCR para la construcción de dominio de unión a ligando de CfEcR mutante por sustitución

MUTANTE	CEBADOR (SEQ ID NO:)	SECUENCIA DE NUCLEÓTIDOS DE CEBADOR (DE 5' A 3')
E20A	FP (SEQ ID NO: 35)	gtaccaggacgggtacgagcagccttctgatgaagattg
E20A	RCP (SEQ ID NO: 36)	caaatcttcatcagaaggctgagcgtaccctctctggtac
Q21A	FP (SEQ ID NO: 37)	ccaggacgggtacgaggcgccttctgatgaagattg
Q21A	RCP (SEQ ID NO: 38)	caaatcttcatcagaaggcgcctctgaccctctctg
F48A	FP (SEQ ID NO: 39)	gagtctgacactcccgccccagatcacag
F48A	RCP (SEQ ID NO: 40)	ctgtgatctggcggcgaggagtgctcagactc
I51A	FP (SEQ ID NO: 41)	cactcccttccgccaggccacagagatgac
I51A	RCP (SEQ ID NO: 42)	gtcatctctgtggcctggcgggaaggagtg
T52A	FP (SEQ ID NO: 43)	cactcccttccgccagatcgagagatgac
T52A	RCP (SEQ ID NO: 44)	gtcatctctgagatctggcgggaaggagtg
T55A	FP (SEQ ID NO: 45)	cgccagatcacagagatggctatctcaggtcc

ES 2 422 303 T3

T55A	RCP (SEQ ID NO: 46)	ggaccgtgaggatagccatctctgtgatctggcg
T58A	FP (SEQ ID NO: 47)	gagatgactatcctcgcggccaacttatctgtg
T58A	RCP (SEQ ID NO: 48)	cacgataagttggaccgagaggatagcatctc
V59A	FP (SEQ ID NO: 49)	gatgactatcctcacggccaacttatctgtgg
V59A	RCP (SEQ ID NO: 50)	ccacgataagttggccgtgaggatagtcac
L61A	FP (SEQ ID NO: 51)	ctatcctcacggtccaagctatctgtgagttcgcg
L61A	RCP (SEQ ID NO: 52)	cgcgaaactccacgatagcttgaccgtgaggatag
I62A	FP (SEQ ID NO: 53)	ctatcctcacggtccaacttgccgtgagttcgcg
I62A	RCP (SEQ ID NO: 54)	cgcgaaactccacggaagttgaccgtgaggatag
M92A	FP (SEQ ID NO: 55)	gctcaagtgaggtagcgtgctccgagtcgc
M92A	RCP (SEQ ID NO: 56)	gcgactcggagcatcgctactcacttgagc
M93A	FP (SEQ ID NO: 57)	gctcaagtgaggtaatggcgctccgagtcgc
M93A	RCP (SEQ ID NO: 58)	gcgactcggagcgcattactcacttgagc
V96A	FP (SEQ ID NO: 59)	gtaatgatgctccgagccgcgacgatac
V96A	RCP (SEQ ID NO: 60)	gtatcgtcgcggtcggagcatcattac
R95A	FP (SEQ ID NO: 61)	gtgaggtaatgatgctcgcagtcgcgacgatac
R95A	RCP (SEQ ID NO: 62)	cgatcgtcgcgactcgcgagcatcattactcac
F108A	FP (SEQ ID NO: 63)	cagacagtgttctggccggaacaaccaagcg
F108A	RCP (SEQ ID NO: 64)	cgcttggtgttcgcgccagaacactgtctg
F109A	FP (SEQ ID NO: 65)	tcagacagtgttctggccggaacaaccaagcg
F109A	RCP (SEQ ID NO: 66)	cgcttggtgttcgcgccagaacactgtctga
A110P	FP (SEQ ID NO: 67)	cagacagtgttctgttccgaacaaccaagcg
A110P	RCP (SEQ ID NO: 68)	cgcttggtgttcgggaacagaacactgtctg
Y120A	FP (SEQ ID NO: 69)	cactcgcgacaacgcccgaaggctggcatg
Y120A	RCP (SEQ ID NO: 70)	catgccagcctgcccggcgtgtcgcgagtg
A123F	FP (SEQ ID NO: 71)	cgacaactaccgcaagttggcatggcctacgtc
A123F	RCP (SEQ ID NO: 72)	gacgtaggcatgccaactgcggtagtgtctg
M125A	FP (SEQ ID NO: 73)	ctaccgcaaggctggcggcctacgtcatc
M125A	RCP (SEQ ID NO: 74)	gatgacgtaggccgcccagccttgcggtag
L230A	FP (SEQ ID NO: 75)	gctcaagaacagaaaggcgcgcttctcctcg
L230A	RCP (SEQ ID NO: 76)	cgaggaaggcggcgccttctgttcttgagc
L223A	FP (SEQ ID NO: 77)	ctccaacatgtgatctccgccaagctcaagaacag
L223A	RCP (SEQ ID NO: 78)	ctgttcttgagctggcggagatgcacatgtggag
L234A	FP (SEQ ID NO: 79)	gaaagctccgccttccgagggagatctg
L234A	RCP (SEQ ID NO: 80)	cagatctcctcggcgaaaggcggcagcttc
W238A	FP (SEQ ID NO: 81)	cttctcggaggatcgcggatgtggcagg
W238A	RCP (SEQ ID NO: 82)	cctgccacatccgcatctcctcggaggaaag
A110n	FP (SEQ ID NO: 83) ALEATORIO	cagacagtgttctgttgncaacaaccaagcg
A110n	RCP (SEQ ID NO: 84) ALEATORIO	cgcttggtgttcgncaacagaacactgtctg
A110n	2 FP (SEQ ID NO: 85) ALEATORIO	cagacagtgttctgttgnngaacaaccaagcg
A110n	2 RCP (SEQ ID NO: 86) ALEATORIO	cgcttggtgttcnnaacagaacactgtctg

T52n	FP (SEQ ID NO: 87) ALEATORIO	cactccctccgccagatcnnngagatgactatcctcacg
T52n	RCP (SEQ ID NO: 88) ALEATORIO	cgtgaggatagtcattcnnngatctggcggaagggagt
V96n	FP (SEQ ID NO: 89) ALEATORIO	gtaatgatgctccgannngcgcgacgatacgtgcggc
V96n	RCP (SEQ ID NO: 90) ALEATORIO	gccgcatcgatcgtcgcgcnntcggagcatcattac
V107I	FP (SEQ ID NO: 107)	gcggcctcagacagtattctgttcgcaac
R175E	FP (SEQ ID NO: 108)	ggtggaagaaatccaggagtactacctgaatacgtcc
Y127E	FP (SEQ ID NO: 109)	caaggctggcatggccgaggtcatcgagg
T52V	FP (SEQ ID NO: 110)	ccctccgccagatcgtagagatgactatcctcac
V96T	FP (SEQ ID NO: 111)	ggtaatgatgctcegaaccgcgcgacgatacg

Tabla 5: Cebadores de PCR para la construcción de dominio de unión a ligando de DmEcR mutante por sustitución

MUTANTE	CEBADOR (SEQ ID NO:)	SECUENCIA DE NUCLEÓTIDOS DE CEBADOR (DE 5' A 3')
A107P	FP (SEQ ID NO: 91)	tccgactcaatattctcccgaataatagatcatatac
A107P	RCP (SEQ ID NO: 92)	gtatatgatctattattcgggaagaatattgagtcgga
G121R	FP (SEQ ID NO: 93)	tcttacaanaatggcccgaatggctgataacattg
G121R	RCP (SEQ ID NO: 94)	caatgttatcagccattcggccattttgtaaga
G121L	FP (SEQ ID NO: 95)	tcttacaanaatggccctaattggctgataacattg
G121L	RCP (SEQ ID NO: 96)	caatgttatcagccattaggccattttgtaaga
N213A	FP (SEQ ID NO: 97)	acgctgggcaaccaggccgccgagatgtgtttc
N213A	RCP (SEQ ID NO: 98)	gaaacacatctcggcgccctggttcccagcgt
C217A	FP (SEQ ID NO: 99)	cagaacgccgagatggctttctactaaagctc
C217A	RCP (SEQ ID NO: 100)	gagctttagtgagaaagccatctcggcgttctg
C217S	FP (SEQ ID NO: 101)	cagaacgccgagatgtctttctactaaagctc
C217S	RCP (SEQ ID NO: 102)	gagctttagtgagaaagacatctcggcgttctg

Tabla 6: Cebadores de PCR para la construcción de dominio de unión a ligando de AmaEcR mutante por sustitución

MUTANTE	CEBADOR (SEQ ID NO:)	SECUENCIA DE NUCLEÓTIDOS DE CEBADOR (DE 5' A 3')
G91A	FP (SEQ ID NO: 103)	gtgatgatgctgagagctgcccggaaatgatg
G91A	RCP (SEQ ID NO: 104)	catcatattccgggcagctctcagcatcatcac
A105P	FP (SEQ ID NO: 105)	acagattctatagtttccaataaccagccgtacac
A105P	RCP (SEQ ID NO: 106)	gtgtacggctggttattgggaacactatagaatctgt

5 Entonces se fusionaron cada uno de los productos de ácido nucleico de PCR resultantes que codificaban para los dominios de unión a ligando de EcR mutante con un dominio de unión a ADN de GAL4 tal como se describió en los ejemplos 1.2, 1.4, 1.9 y 1.13 anteriores. Se sometieron a prueba los constructos receptores de GAL 4/EcR mutante para determinar la actividad transfectándolos en células NIH3T3 junto con CP16/LmUSP-EF y pFRLuc en presencia de ligando esteroideo o no esteroideo.

10 También se fusionaron cada uno de los ácidos nucleicos resultantes que codificaban para los dominios de unión a ligando de EcR mutante con un dominio de transactivación de VP16 tal como se describió en los ejemplos 1.6 y 1.11 anteriores. Se sometieron a prueba los constructos receptores de VP16/CfEcR-DEF mutante y VP 16/DmEcR-CDEF mutante para determinar la actividad transfectándolos en células de insecto L57 junto con un 6XEcRE/gen indicador de la  $\beta$ -galactosidasa en presencia de 20-hidroxiecdisona (20E).

15 Ligandos: Se adquirieron los ligandos esteroideos muristerona A, ponasterona A,  $\alpha$ -ecdisona y 20-hidroxiecdisona de Sigma Chemical Company e Invitrogen. Los ligandos no esteroideos: N-(2-etil-3-metoxibenzoil)-N'-(3,5-dimetilbenzoil)-N'-terc-butilhidrazina (ligando no esteroideo GS<sup>TM</sup>-E); N'-terc-butil-N'-(3,5-dimetilbenzoil)-3-metoxi-2-metilbenzohidrazida (RH-2485); N-terc-butil-N'-(4-etilbenzoil)-3,5-dimetilbenzohidrazida (RH-5992) y N'-terc-butil-N'-(3,5-dimetilbenzoil)-3,4-(1,2-etilendioxi)-2-metilbenzohidrazida (RH-125020) son ligandos ecdiesteroides estables

sintéticos sintetizados en Rohm and Haas Company. Se disolvieron todos los ligandos en DMSO y se mantuvo la concentración final de DMSO al 0,1% tanto en los controles como en los tratamientos. Se adquirieron <sup>3</sup>H-PonA y <sup>3</sup>H- $\alpha$ -ecdisona de New England Nuclear. Se sintetizó <sup>3</sup>H-RH2485 en Rohm and Haas Company.

5 Transfecciones: Se transfectaron ADN correspondientes a los diversos constructos de interruptor expuestos en el ejemplo 1, específicamente los interruptores 1.1-1.13, en células NIH3T3 (ATCC) de ratón o células L57 (Dr. Peter Cherbas; Indiana University) tal como sigue. Se siguieron métodos convencionales para el cultivo y mantenimiento de las células. Se recogieron las células cuando alcanzaron el 50% de confluencia y se sembraron en placas de 6, 10 12 ó 24 pocillos a 125.000, 50.000 ó 25.000 células, respectivamente, en 2,5, 1,0 ó 0,5 ml de medio de crecimiento que contenía suero bovino fetal al 10% (FBS), respectivamente. El día siguiente, se aclararon las células con medio de crecimiento y se transfectaron durante cuatro horas. Se usó Superfect™ (Qiagen Inc.) para células 3T3 y se usó Lipofectamine™ (LifeTechnologies) para células L57 como reactivos de transfección. Para las placas de 12 pocillos, se mezclaron 4  $\mu$ l de Superfect™ o Lipofectamine™ con 100  $\mu$ l de medio de crecimiento. Se añadieron un  $\mu$ g de constructo indicador y 0,25  $\mu$ g de cada constructo receptor del par receptor que va a analizarse a la mezcla de transfección. Se añadió un segundo constructo indicador [pTKRL (Promega), 0,1  $\mu$ g/mezcla de transfección] que comprendía un gen de la luciferasa de *Renilla* operativamente unido y colocado bajo el control de un promotor constitutivo de timidina cinasa (TK) y se usó para la normalización. Se mezcló el contenido de la mezcla de transfección en un mezclador con vórtex y se dejó a temperatura ambiente durante 30 minutos. Al final de la incubación, se añadió la mezcla de transfección a las células mantenidas en 400  $\mu$ l de medio de crecimiento. Se mantuvieron las células a 37°C y CO<sub>2</sub> al 5% durante cuatro horas. Al final de la incubación, se añadieron 500  $\mu$ l de medio de crecimiento que contenía FBS al 20% y o bien dimetilsulfóxido (DMSO; control) o bien una disolución en DMSO de ligando esteroideo o no esteroideo y se mantuvieron las células a 37°C y CO<sub>2</sub> al 5% durante 48 horas. Se recogieron las células y se sometió a ensayo la actividad indicadora. Se siguió el mismo procedimiento para placas de 6 y 24 pocillos excepto porque se duplicaron todos los reactivos para placas de 6 pocillos y se redujeron a la mitad para placas de 24 pocillos.

Ensayos de indicador: Se recogieron células 40 horas tras añadir ligandos. Se añadieron 125  $\mu$ l de tampón de lisis pasivo (parte del sistema de ensayo de indicador Dual-luciferase™ de Promega Corporation) a cada pocillo de la placa de 24 pocillos. Se colocaron las placas en un agitador rotatorio durante 15 minutos. Se sometieron a ensayo 30 veinte  $\mu$ l de lisado. Se midió la actividad luciferasa usando el sistema de ensayo de indicador Dual-luciferase™ de Promega Corporation siguiendo las instrucciones del fabricante. Se midió la  $\beta$ -galactosidasa usando un kit de ensayo Galacto-Star™ de TROPIX siguiendo las instrucciones del fabricante. Se normalizaron todas las actividades luciferasa y  $\beta$ -galactosidasa usando luciferasa de *Renilla* como patrón. Se calcularon las actividades en veces dividiendo las unidades de luz relativas ("RLU") normalizadas en células tratadas con ligando entre RLU normalizada en células tratadas con DMSO (control sin tratamiento).

## EJEMPLO 2

Este ejemplo describe la identificación de mutantes por sustitución de dominio de unión a ligando de CfEcR 40 sensibles a compuestos no esteroideos mejorados que presentan aumento de actividad en respuesta a ligando no esteroideo y disminución de actividad en respuesta a ligando esteroideo. Brevemente, los solicitantes mutaron residuos de aminoácido que se prevé que son críticos para la unión a ecdiesteroides a alanina y crearon casetes de expresión génica de ADNc de GAL4/CfEcR-DEF mutante tal como se describió en el ejemplo 1 anterior usando el kit de mutagénesis dirigido a sitio mediado por PCR Quikchange (Stratagene, La Jolla, CA). Se sometieron a prueba los 45 ADNc mutados y de tipo natural en ensayos de indicador de luciferasa dirigido por GAL4.

Transfecciones: Se transfectaron los ADN correspondientes a los diversos constructos de interruptor expuestos en el ejemplo 1, específicamente los interruptores 1.1-1.2, en células NIH3T3 (ATCC) de ratón tal como sigue. Se recogieron células cuando alcanzaron el 50% de confluencia y se sembraron en placas de 24 pocillos a 12.500 50 células/pocillo en 0,5 ml de medio de crecimiento que contenía suero bovino fetal al 10% (FBS). El día siguiente, se aclararon las células con medio de crecimiento y se transfectaron durante cuatro horas. Se encontró que Superfect™ (Qiagen Inc.) era el mejor reactivo de transfección para células 3T3. Se mezclaron dos  $\mu$ l de Superfect™ con 100  $\mu$ l de medio de crecimiento y se añadieron 50 ng de casete o bien GAL4/EcR de tipo natural o bien Gal4/EcR mutante, 50 ng de VP16/LmUSP-EF y 200 ng de pFRLuc a la mezcla de transfección. Se añadió un 55 segundo constructo indicador [pTKRL (Promega), 0,05  $\mu$ g/ mezcla de transfección] que comprendía un gen de luciferasa de *Renilla* operativamente unido y colocado bajo el control de un promotor constitutivo de timidina cinasa (TK) y se usó para la normalización. Se mezcló el contenido de la mezcla de transfección en un mezclador con vórtex y se dejó reposar a temperatura ambiente durante 30 min. Al final de la incubación, se añadió la mezcla de transfección a las células mantenidas en 200  $\mu$ l de medio de crecimiento. Se mantuvieron las células a 37°C y CO<sub>2</sub> al 5% durante cuatro horas. Al final de la incubación, se añadieron 250  $\mu$ l de medio de crecimiento que contenía FBS al 20% y o bien dimetilsulfóxido (DMSO; control) o bien una disolución en DMSO de ligando esteroideo PonA 10 nM o 2,5  $\mu$ M o ligando no esteroideo GS™-E [N-(2-etil-3-metoxibenzoil)-N'-(3,5-dimetilbenzoil)-N'-terc-butilhidrazina] y se mantuvieron las células a 37 °C y CO<sub>2</sub> al 5% durante 40 horas. Se recogieron las células y se sometió a ensayo la actividad indicadora tal como se describió anteriormente. Se calcularon las actividades en veces dividiendo las 65 unidades de luz relativas ("RLU") normalizadas en células tratadas con ligando entre RLU normalizadas en células

tratadas con DMSO (control sin tratar).

Se identificaron dos residuos de aminoácido que, cuando se sustituyen, producen un receptor de ecdisona mutante que presenta aumento de actividad en respuesta a un ligando no esteroideo y disminución de actividad en respuesta a un ligando esteroideo. El efecto de sustitución de alanina en el residuo de aminoácido 52 ó 55 de SEQ ID NO: 1 sobre la actividad del receptor CfEcR-DEF mutado se presenta en la tabla 7 como aumento en veces con respecto a la actividad de interruptor Gal4/CtEcR-DEF de tipo natural (WT).

Tabla 7. Mutantes de CtEcR-DEF que muestran aumento de actividad no esteroidea y disminución de actividad esteroidea. Aumento en veces con respecto al tipo natural

MUTANTES	GS <sup>TM</sup> -E 2,5 µM	PonA 2,5 µM
T52A	1,5	0,5
T55A	1,7	0,13

### EJEMPLO 3

Este ejemplo describe la identificación de mutantes por sustitución de dominio de unión a ligando de CfEcR sensibles a esteroides que presentan aumento de actividad en respuesta a ligando esteroideo y disminución significativa de actividad en respuesta a ligando no esteroideo. En un intento de identificar mutaciones por sustitución en el CfEcR que aumentan la actividad de ligando esteroideo, pero disminuyen la actividad de ligando no esteroideo, los solicitantes mutaron residuos de aminoácido que se prevé que son críticos para la unión a ecdiesteroides y crearon casetes de expresión génica de ADNc de GAL4/CtEcR-DEF mutante tal como se describió en el ejemplo 1 anterior usando el kit de mutagénesis dirigida a sitio mediada por PCR. Se prepararon los ADNc mutados y de tipo natural correspondientes a los diversos constructos de interruptor expuestos anteriormente en los ejemplos 1.1 y 1.2 y se sometieron a prueba en ensayos de indicador de luciferasa dirigido por GAL4 tal como se describió en el ejemplo 2 anterior. Se calculó la actividad en veces dividiendo RLU en presencia de ligando entre RLU en ausencia del ligando.

Se identificaron residuos de aminoácido específicos que, cuando se sustituyen, producen un receptor de ecdisona mutante que presenta aumento de actividad en respuesta a un ligando esteroideo y disminución de actividad en respuesta a un ligando no esteroideo. El efecto de una sustitución de aminoácido en el residuo de aminoácido 52 ó 96 de SEQ ID NO: 1 y sustitución de aminoácido en los residuos de aminoácido 96, 107 y 175 de SEQ ID NO: 1 sobre la actividad del receptor CfEcR-DEF mutado se presenta en la tabla 8 como aumento en veces con respecto a la actividad de interruptor de Gal4/CfEcR-DEF de tipo natural (WT).

Tabla 8. Mutantes que muestran aumento de actividad esteroidea y disminución de actividad no esteroidea

Aumento en veces con respecto al tipo natural				
MUTANTES	GS <sup>TM</sup> E 2,5 µM	PonA 2,5 µM	GS <sup>TM</sup> E 10 nM	PonA 10 nM
T52L	0,26	3,4		
V96A	0,35	408		
V96T	0,018	45		
V96T/V107I/R175E			0,4	485,7

### EJEMPLO 4

Este ejemplo describe la identificación de mutantes por sustitución de dominio de unión a ligando de CfEcR sensibles a esteroides y a compuestos no esteroides mejorada que presentan aumento de actividad en respuesta tanto a un ligando esteroideo como a un ligando no esteroideo. En un intento de identificar las mutaciones por sustitución en el CfEcR que aumentan la actividad de ligando tanto esteroideo como no esteroideo, los solicitantes mutaron residuos de aminoácido que se prevé que son críticos para la unión a esteroides y crearon casetes de expresión génica de ADNc de GAL4/CfEcR-DEF mutante tal como se describió en el ejemplo 1 anterior usando mutagénesis dirigida a sitio mediada por PCR. Se prepararon ADNc mutados y de tipo natural correspondientes a los diversos constructos de interruptor expuestos anteriormente en los ejemplos 1.1 y 1.2 y se sometieron a prueba en ensayos de indicador de luciferasa dirigido por GAL4 tal como se describió en el ejemplo 2 anterior. Se calculó la actividad en veces dividiendo RLU en presencia de ligando entre RLU en ausencia del ligando.

Se identificaron residuos de aminoácido específicos que, cuando se sustituyen, producen un receptor de ecdisona mutante que presenta aumento de actividad en respuesta a ligandos tanto no esteroides como esteroides. El efecto de una sustitución de aminoácido en el residuo de aminoácido 52, 96, 107 ó 175 de SEQ ID NO: 1, sustitución de aminoácido en los residuos de aminoácido 107 y 175 de SEQ ID NO: 1, sustitución de aminoácido en los

- residuos de aminoácido 127 y 175 de SEQ ID NO: 1, sustitución de aminoácido en los residuos de aminoácido 107 y 127 de SEQ ID NO: 1, sustitución de aminoácido en los residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, sustitución de aminoácido en los residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1 o sustitución de aminoácido en los residuos de aminoácido 96, 107 y 175 de SEQ ID NO: 1 sobre la actividad del receptor CfEcR-DEF mutado se presenta en la tabla 9 como aumento en veces con respecto a la actividad de interruptor de Gal4/CfEcR-DEF de tipo natural (WT).

Tabla 9. Mutantes que muestran aumento de actividad esteroidea y no esteroidea

Aumento en veces con respecto al tipo natural				
MUTANTES	GS <sup>TME</sup> 2,5 µM	PonA 2,5 µM	GS <sup>TME</sup> 10 nM	PonA 10 nM
T52V	17,3	35,7		
T52I	8	8		
V96D	3,07	3,1		
V96M	122	3,37		
V107I			12,4	26,6
R175E			22,0	11,3
V107I/R175E			386,4	1194,4
Y127E/R175E			622,8	42,2
V107I/Y127E			314,6	35,8
V107I/Y127E/R175E			124,3	122,3
T52V/V107I/R175E			62,8	136,6
V96A/V107I/R175E			21,1	1005,1
T52A/V107I/R175E			2,3	20,3

## 10 EJEMPLO 5

Este ejemplo describe la identificación de mutantes por sustitución de dominio de unión a ligando de CfEcR sensibles a compuestos no esteroideos que presentan disminución significativa de actividad en respuesta a ligando esteroideo pero no afectan a la actividad en respuesta a ligando no esteroideo. En un intento de identificar mutaciones por sustitución en el CfEcR que disminuyen la actividad de ligando esteroideo, pero no afectan a la actividad de ligando no esteroideo, los solicitantes mutaron residuos de aminoácido que se prevé que son críticos para la unión a ecdisteroides y crearon casetes de expresión génica de ADNc de GAL4/CfEcR-DEF mutante tal como se describió en el ejemplo 1 anterior usando mutagénesis dirigida a sitio mediada por PCR. Se prepararon ADNc mutados y de tipo natural correspondientes a los diversos constructos de interruptor expuestos anteriormente en los ejemplos 1.1 y 1.2 y se sometieron a prueba en ensayos de indicador de luciferasa dirigido por GAL4 tal como se describió en el ejemplo 2 anterior. Se calculó la actividad en veces dividiendo RLU en presencia de ligando entre RLU en ausencia del ligando.

Se identificaron cuatro residuos de aminoácido que, cuando se sustituyen, producen receptor de ecdisona mutante que presentan disminución de actividad en respuesta a un ligando esteroideo y efecto mínimo sobre la actividad en respuesta a un ligando no esteroideo. El efecto de una sustitución de aminoácido en el residuo de aminoácido 20, 58, 92 ó 110 de SEQ ID NO: 1 sobre la actividad del receptor CfEcR-DEF mutado se presenta en la tabla 10 como aumento en veces con respecto a la actividad de interruptor de Gal4/ CfEcR-DEF de tipo natural (WT).

Tabla 10. Mutantes que muestran disminución de actividad esteroidea, pero la actividad no esteroidea no se ve afectada. Aumento en veces con respecto al tipo natural

MUTANTES	GS <sup>TME</sup> 2,5 µM	PonA 2,5 µM
E20A	0,9	0,35
T58A	0,8	0,008
M92A	0,7	0,39
A110P	0,8	0,005

Tal como se describe en la tabla 10, los solicitantes han identificado mutaciones puntuales en el dominio de unión a ligando de CfEcR que reducen significativamente la actividad de unión a esteroides. Los mutantes puntuales de

CfEcR T58A y A110P eliminaron esencialmente la actividad de unión a esteroides. De manera interesante, la actividad no esteroidea de estos mutantes puntuales no se vio significativamente afectada.

#### EJEMPLO 6

5 Los solicitantes han caracterizado adicionalmente el receptor CfEcR A110P no esteroideo identificado en el ejemplo 5 anterior. Este ejemplo demuestra que la mutación de un residuo de alanina crítico (A110) conduce a la ruptura de la unión a esteroide y por tanto la transactivación mediante el EcR en presencia de esteroides. Sin embargo, la unión así como la transactivación mediante compuestos no esteroideos no se ve alterada.

#### 10 Ensayo de unión a ligando

Los solicitantes sometieron a prueba el receptor CfEcR mutante A110P en un ensayo de unión a ligando esteroideo y no esteroideo para confirmar que se eliminó la unión a esteroides. Brevemente, se determinó la actividad de unión a PonA usando un ensayo de unión a ligando (LBA) *in vitro*. Se realizó el ensayo de unión a ligando esteroideo (LBA) usando <sup>3</sup>H-PonA (200 Ci/mmol). En el ensayo se usaron Gal4/CfEcR-DEF de tipo natural o mutante A110P traducido *in vitro* y GST-CfUSP-A/BCDEF expresado por bacterias. Se realizó el ensayo con 8 µl de Gal4/CfEcR-DEF de tipo natural o mutante A110P, 2,5 µl de GST-CfUSP-A/BCDEF, 1 µl de <sup>3</sup>HPonA y 2 µl de PonA ("frío") no marcado como competidor en presencia de tampón de T [Tris 90 mM pH 8,0, DTT 10 µM que comprendía un cóctel de inhibidor de proteasa Complete™ usado según las instrucciones del fabricante (Boehringer Mannheim)]. Se llevó a cabo la reacción a temperatura ambiente durante 1 hora seguido por la adición de carbón recubierto por dextrano (Sigma). Se centrifugó la mezcla a 7000 xg durante 10 minutos, se midió la cantidad de <sup>3</sup>H-PonA en el sobrenadante. Se realizaron las reacciones por triplicado. También se tradujeron *in vitro* el EcR de tipo natural de longitud completa o su mutante A110P y se transcribieron usando el sistema TNT (Promega) según las instrucciones del fabricante y se sometieron a prueba en ensayos de unión a ligando *in vitro* usando <sup>3</sup>H-RH2485, con 20E frío o compuestos no esteroideos (RH2485 y GS™-E) como competidores. Además, se sometieron a ensayo 5 µl de las traducciones *in vitro* para determinar la eficiencia de traducción usando SDS-PAGE siguiendo métodos convencionales (Maniatis, 1989). Se calcularon los resultados de unión a ligando tanto para los receptores CfEcR-DEF de tipo natural como mutante A110P y se muestran en la figura 1.

30 Los dos ligandos no esteroideos, RH2485 y GS™-E, y el ligando esteroideo 20E sometidos a prueba pudieron competir de manera eficaz con <sup>3</sup>H-RH2485 unido, lo que sugiere que pueden unirse a EcR de longitud completa de tipo natural de manera eficaz (véase la figura 1). Sin embargo, cuando se examinó la unión de los mismo ligandos por el mutante A110P, se rompió completamente la unión del esteroide 20E pero la unión de los compuestos no esteroideos no se vio afectada (véase la figura 1). Estos resultados indican que la ausencia de unión a esteroides en el caso de la proteína de fusión GAL4-CfEcR no es un artefacto del truncamiento o de la fusión y demuestra que el CfEcR mutante A110P es un receptor no esteroideo selectivo.

40 Afinidades de ligando de CfEcR mutante A110P a diversos ligandos esteroideos y no esteroideos: Se midieron las afinidades de unión a ligando del mutante de GAL4/CfEcR A110P uniendo <sup>3</sup>H-RH2485 y haciéndolo competir con diferentes concentraciones de esteroides o compuestos no esteroideos fríos. Brevemente, se unió <sup>3</sup>H-RH2485 a CfEcR de longitud completa traducido *in vitro* y CfUSP expresada por bacterias y se hizo competir con concentraciones crecientes de ligandos esteroideos o no esteroideos fríos. Se llevó a cabo la reacción a temperatura ambiente durante 1 hora seguido por adición de carbón recubierto por dextrano activado y la centrifugación a 7000 xg durante 10 minutos a 4°C. Se midió el <sup>3</sup>H-RH2485 residual en el sobrenadante tras la centrifugación usando un contador de centelleo. Se determinaron los valores de fracción unida (f unida) y se trazaron frente a la concentración de ligando (en mM). Se determinaron los valores de CI<sub>50</sub> para cada uno de los ligandos esteroideos PonA y MurA y no esteroideos N-(2-etil-3-metoxibenzoil)-N'-(3,5-dimetilbenzoil)-N'-terc-butilhidrazina (ligando no esteroideo GS™-E); N'-terc-butil-N'-(3,5-dimetilbenzoil)-3-metoxi-2-metilbenzohidrazida (RH-2485); N-terc-butil-N'-(4-etilbenzoil)-3,5-dimetilbenzohidrazida (RH-5992) y N'-terc-butil-N'-(3,5-dimetilbenzoil)-3,4-(1-,2-etilendioxi)-2-metilbenzohidrazida (RH-125020) para el tipo natural y el mutante trazando la fracción unida frente a la concentración de cada ligando.

55 Tal como se muestra en la tabla 11, los valores de CI<sub>50</sub> para esteroides PonA y MurA aumentaron (más de 1 mM) en el caso del mutante A110P en comparación con los valores nanomolares observados para el receptor de tipo natural, lo que sugiere que se alteró la unión de esteroides en el CfEcR mutante A110P. Por otro lado, los valores de CI<sub>50</sub> de compuestos no esteroideos fueron similares tanto para los receptores mutante A110P como de tipo natural (véase la tabla 11). Estos resultados confirman los hallazgos de los solicitantes presentados en el ejemplo 5 anterior de que la mutación por sustitución A110P da como resultado un dominio de unión a ligando de receptor de ecdisona no esteroideo que ha perdido la capacidad de unirse a ligando esteroideo.



Tabla 11. Valores de  $CI_{50}$  determinados para CfEcR-A/BCDEF de tipo natural y mutante A110P usando varios ligandos esteroideos y no esteroideos.

	$CI_{50}$ de tipo natural (nM)	$CI_{50}$ de mutante A110P (nM)
Esteroides:		
Ponasterona A	345,95	> 1 mM
Muristerona A	423,99	> 1 mM
Compuestos no esteroideos:		
GS <sup>TM</sup> -E	85,26	12,88
CH-5992	132,81	322,34
RH-2485	1,80 x 10 <sup>3</sup>	350,42
RH-125020	10,11	25,71

#### A110P en CfEcR truncado o fondo de CfEcR de longitud completa

5 Para eliminar la posibilidad de que la pérdida de actividad esteroidea pueda deberse al receptor CfEcR-DEF truncado o a un artefacto de la proteína de fusión de GAL4, los solicitantes introdujeron la mutación A110P en el CfEcR (CfEcRA/BCDEF) de longitud completa (FL), lo fusionaron con el dominio de unión a ADN de GAL4 tal como se describió en el ejemplo 1.4, y lo sometieron a ensayo de manera comparativa en células NIH3T3 en placas de 24 pocillos tal como se describió en el ejemplo 2 anterior con CfEcR de tipo natural de longitud completa (ejemplo 1.3) en combinación con VP16/LmUSP-EF y pFREcRE Luc que comprenden el gen indicador de luciferasa operativamente unido a seis copias del elemento de respuesta a ecdisona (6X EcRE) y un TATAA sintético. Se hicieron crecer las células transfectadas en presencia de PonA o GS<sup>TM</sup>-E 0,25 ó 10  $\mu$ M y se midió la actividad indicadora a 40 horas tras añadir ligando. Se recogieron las células y se sometieron los extractos a ensayo para determinar la actividad de luciferasa. Los resultados se presentan en la figura 2. Los números en la parte superior de las barras indican el aumento de veces con respecto a niveles de DMSO.

20 Tal como se muestra en figura 2, la mutación A110P tuvo un efecto similar cuando se introdujo en el receptor de ecdisona de longitud completa en el contexto de un gen indicador dirigido por EcRE. Específicamente, se perdió completamente la actividad de PonA para el mutante de CfEcR A110P de longitud completa. Sin embargo, no hubo diferencia significativa en la actividad no esteroidea entre CfEcR de tipo natural de longitud completa y el CfEcR mutante A110P de longitud completa en presencia de GS<sup>TM</sup>-E. Estos resultados indican que la ausencia de sensibilidad del mutante A110P frente a esteroides observada con el CfEcR de fusión con GAL4 en el ejemplo 5 no era un artefacto de la fusión de GAL4 o del truncamiento de EcR. Por tanto, los solicitantes han determinado que el residuo de aminoácido A110 es crítico para la actividad esteroidea en el receptor de ecdisona de longitud completa.

#### El residuo A110 es crítico para la actividad esteroidea en células de insecto.

30 En células de mamífero, el ligando natural de EcR, 20-hidroxiecdisona (20E), no induce la transactivación a través de un sistema de expresión génica basado en CfEcR. Para determinar si el mutante A110P puede responder a 20E, los solicitantes sometieron a prueba este sistema de expresión génica basado en mutante A110P en un ensayo de indicador dirigido por EcRE en células L57 de insecto (una línea celular de *Drosophila melanogaster* que carece de isoforma B de EcR, el homólogo de isoforma de EcR de *Drosophila melanogaster* de la isoforma B de CfEcR de la cual se deriva el mutante A110P, sin embargo las células L57 aún contienen la isoforma A de EcR). Se introdujo la mutación en la proteína de fusión VP16/CfEcR-CDEF y se unió operativamente a un promotor IE1 de baculovirus y se transfectaron las células L57 con IE1VP16CfEcRCDEF (ejemplo 1.5) o su ADN de versión de mutante A110P (ejemplo 1.6) junto con un gen indicador de  $\beta$ -galactosidasa pMK43.2 bajo el control de 6X elementos de respuesta a ecdisona ("6XEcRE"; se obtuvo el constructo pMK43.2 de Michael Koelle en la Stanford University) en placas de 24 pocillos. Se midió la actividad indicadora para la transactivación del sistema de expresión génica basado en mutante A110P tras 40 horas de tratamiento de las células transfectadas con 20E o GS<sup>TM</sup>-E 0, 1, 10, 100 ó 1000 nM. Se recogieron las células y se sometieron a ensayo los extractos para determinar la actividad de  $\beta$ -galactosidasa ( $\beta$ -gal) y luciferasa. Se midió la  $\beta$ -galactosidasa usando el kit de ensayo Galacto-Star<sup>TM</sup> de TROPIX siguiendo las instrucciones del fabricante. Los números en la parte superior de las barras indican el aumento de veces con respecto a niveles de DMSO. Los resultados se presentan en la figura 3.

45 En células de insecto, el sistema de expresión génica basado en CfEcR de tipo natural indujo actividad de  $\beta$ -gal de una manera dependiente de la dosis en respuesta tanto a 20E como GS<sup>TM</sup>-E, mientras que el sistema de expresión génica basado en el mutante A110P transactivó la expresión de gen indicador en presencia de GS<sup>TM</sup>-E, sin embargo, las células L57 transfectadas con el mutante A110P mostraron un ligero aumento de actividad de gen indicador en presencia de 20E (véase la figura 3). Esta actividad de nivel bajo en presencia de 20E se debe lo más

probablemente a una actividad endógena de la isoforma A de EcR dentro de las células L57 puesto que los solicitantes han demostrado que el mutante A110P derivado de la isoforma B de CfEcR no 20E (datos no mostrados).

- 5 La mutación A110P tuvo un efecto similar cuando se introdujo en el receptor de longitud completa en el contexto de un gen indicador dirigido por EcRE en las células L57 indicando que la mutación tiene un efecto análogo en células de insecto, supuestamente en presencia de co-factores de la transcripción de insecto (datos no mostrados). Estos datos confirman los resultados de los solicitantes de células de mamífero y establecen que la mutación A110P da como resultado un efecto drástico sobre la sensibilidad esteroidea de CfEcR pero no afecta a la sensibilidad no esteroidea de CfEcR.

#### EJEMPLO 7

- 15 Los resultados de los solicitantes presentados anteriormente en los ejemplos 5 y 6 describen la identificación de un residuo de alanina en la posición 110 que es un residuo crítico para la actividad esteroidea pero no para la no esteroidea del dominio de unión a ligando de CfEcR. Para caracterizar adicionalmente el papel del residuo A110 en la transactivación esteroidea y no esteroidea de CfEcR de genes indicadores, se preparó una mini-biblioteca de receptores CfEcR-DEF mutando A110 usando cebadores degenerados. Se diseñaron estos cebadores de PCR degenerados (pares de cebadores aleatorios de A110P que comprenden o bien SEQ ID NO: 83 y SEQ ID NO: 84 o bien SEQ ID NO: 85 y SEQ ID NO: 86; véase la tabla 4) para reemplazar A110 por diversos residuos de aminoácido. Las condiciones de mutagénesis por PCR usadas fueron tal como se describieron anteriormente en el ejemplo 1. Se secuenciaron los clones resultantes para identificar los mutantes.

#### Transactivación del gen indicador de mutantes de A110

- 25 Se obtuvieron cuatro mutaciones: A110S, A110P, A110L y A110M. Se sometieron a ensayo estos cuatro receptores mutantes y de tipo natural en células NIH3T3. Se transfirieron fusiones de GAL4 de cada uno de los cuatro receptores CfEcR-DEF mutantes o de tipo natural, VP16LmUSPEF y pFRLUC en células NIH3T3, se hicieron crecer las células en presencia de 0, 0,04, 0,2, 1, 5 ó 25  $\mu$ M de PonA, MurA, N-(2-etil-3-metoxibenzoil)-N'-(3,5-dimetilbenzoil)-N'-terc-butildiazina (ligando no esteroideo GS<sup>TM</sup>-E), o N'-terc-butyl-N'-(3,5-dimetilbenzoil)-3,4-(1,2-etilendioxi)-2-metilbenzohidrazida (RH-125020) durante 48 horas y se midió la actividad indicadora. Tal como se muestra en la figura 4, el receptor de ecdisona de tipo natural mostró actividad indicadora en presencia tanto de esteroides como de compuestos no esteroideos. Sin embargo, todos los receptores mutantes mostraron actividad indicadora sólo en presencia de ligandos no esteroideos pero no en presencia de ligandos esteroideos. El mutante A110P presentó actividad no esteroidea similar en comparación con receptor de tipo natural, sin embargo los mutantes A110S, A110L y A110M demostraron sensibilidad inferior y transactivación no detectable a las dos concentraciones más bajas. Estos resultados confirman que un dominio de unión a ligando de EcR mutante por sustitución de A110 se caracteriza por una respuesta significativamente reducida a esteroides pero permanece sensible a compuestos no esteroideos.

- 40 Unión a ligandos mediante los mutantes de A110: Los solicitantes han realizado ensayos de unión a ligando usando ligandos marcados de manera radioactiva de <sup>3</sup>H-PonA o 3HRH2485 (N'-terc-butyl-N'-(3,5-dimetilbenzoil)-3-metoxi-2-metilbenzohidrazida) para determinar si las diferencias en la respuesta a ligando de receptores de ecdisona mutantes por sustitución y de tipo natural se deben a diferencias al nivel de unión a ligando. Los receptores CfEcR de tipo natural y mutante se transcribieron *in vitro* y se tradujeron y se sometieron a ensayo en presencia de GST-CfUSP-A/BCDEF expresado por bacterias (longitud completa). La unión de PonA a mutante A110P no fue detectable mientras que otros mutantes por sustitución de A110 mostraron el 5-10% de unión a receptor de tipo natural (véase la figura 5). EcR de tipo natural y todos los mutantes sometidos a prueba mostraron unión similar a RH-2485, un análogo próximo de GS<sup>TM</sup>-E (véase la figura 6). Puede ser que el residuo de prolina inflexible del mutante A110P dificulte la unión de ligandos esteroideos pero no de los no esteroideos.

- 55 Se analizaron las proteínas traducidas *in vitro* mediante SDS-PAGE y se encontró que se traducían de manera similar al receptor de tipo natural, indicando que las diferencias en la unión observada entre los mutantes y receptor de tipo natural no se deben a la variación en la cantidad de proteínas presentes en el ensayo (datos no mostrados). La actividad de unión a ligando se correlaciona con la actividad del gen indicador en la mayoría de los casos, proporcionando prueba adicional del descubrimiento de los solicitantes de que el residuo A110 desempeña un papel crítico en la unión de ecdiesteroides, pero no de compuestos no esteroideos.

- 60 Todos de los mutantes por sustitución de A110 estaban alterados en cuanto a la unión a esteroides así como en su capacidad de transactivar genes indicadores en presencia de esteroides en células de mamífero y de insecto. Estos mutantes mantuvieron los niveles de tipo natural de la unión no esteroidea y la transactivación del gen indicador en presencia de compuestos no esteroideos.

- 65 El residuo A110, encontrado junto a la lámina  $\beta$  predicha entre la hélice 5 y 6, está sumamente conservado en EcR de diversas especies de insectos, y en RXR, receptor de progesterona (PR) y receptor de estrógenos (ER) destacando adicionalmente que este residuo es crítico para la unión a ligando y por tanto la transactivación.

Además, el residuo A110 está flanqueado por otros residuos sumamente conservados, algunos de los cuales también pueden ser críticos para la unión a ligando y/o transactivación. La comparación estrecha de las estructuras tridimensionales de receptores nucleares muestra que hay grandes cambios estructurales incluso en el LBD conservado entre diferentes receptores nucleares. Se observan importantes cambios no sólo entre diferentes estructuras de receptores nucleares, sino también en complejos del mismo receptor cuando se une a ligandos naturales y sintéticos como en el caso de ER. La unión de esteroides y compuestos no esteroideos puede reflejar una situación similar. Los modelos de homología generados sugieren que la unión de los dos ligandos es diferente, en cuanto a las hélices implicadas y a los residuos de contacto. El examen estrecho de los resultados del ensayo de transactivación sugiere que las respuestas a la dosis son ligeramente diferentes en las dos situaciones. Los esteroides son menos activos a concentraciones inferiores mientras que la actividad inducida por compuestos no esteroideos es varias veces mayor. Sin embargo, a dosis superiores las actividades esteroidea y no esteroidea son similares. La actividad superior de compuestos no esteroideos a concentraciones inferiores puede reflejar la afinidad superior de los compuestos no esteroideos con respecto al EcR. Se ha sugerido que la presencia del grupo terciario permite que algunos compuestos no esteroideos formen contactos de van der Waals extensos con el LBD de EcR y por tanto se ajusten en un ranura que no se ocupa por ecdisteroides. Esto puede explicar en algún grado las diferencias observadas en las actividades de los esteroides y los compuestos no esteroideos en ensayos de gen indicador a bajas concentraciones ligando. Aún no se conoce el mecanismo mediante el cual la unión afecta a los cambios conformacionales y por tanto al potencial de transactivación. Sin embargo, en el caso del ER, se ha mostrado que la unión de agonistas y antagonistas induce diferentes cambios conformacionales dando como resultado el desplazamiento de la hélice 12. La hélice 12 desempeña un papel activo en la inclusión e interacción de coactivadores con respecto al receptor. En el caso de agonistas como dietilstilbestrol (DES), el coactivador GRIP1 se une a una ranura hidrófoba sobre la superficie del LBD formado por las hélices 3, 4, 5 y 12 y el giro entre las hélices 3 y 4. Por otro lado, en presencia de un antagonista parcial, 4-hidroxitamoxifeno (OHT), la hélice 12 bloquea la ranura de reconocimiento de coactivador imitando las interacciones de la caja GRIP NR con el LBD. La unión de esteroides y de compuestos no esteroideos también podría inducir cambios conformacionales sutiles que afectan a la inclusión de coactivador y por tanto a la transactivación.

El residuo A110 parece interactuar con la cadena lateral del ligando esteroideo. La introducción de residuos más voluminosos o inflexibles en esta posición podría romper potencialmente estas interacciones y por tanto el acoplamiento del ligando al LBD. Esto a su vez da como resultado la ausencia de activación del EcR. La identificación de un mutante que da como resultado la ruptura de la unión a ecdisteroides sin afectar a la unión a compuestos no ecdisteroides y la activación proporciona medios para la evolución sistemática del EcR para desarrollar un sistema inducible por ecdisona que puede regularse de manera precisa para su uso en sistemas de mamíferos.

#### 35 EJEMPLO 8

Este ejemplo describe la identificación de mutantes por sustitución de dominio de unión a ligando de CfEcR que presentan disminución de actividad en respuesta tanto a ligando esteroideo como a ligando no esteroideo. Estos mutantes por sustitución son útiles en ensayos de examen de ligandos para la identificación de ligandos ortogonales. En un intento de identificar mutaciones por sustitución en el CfEcR que disminuyen la actividad de ligando tanto esteroideo como no esteroideo, los solicitantes mutaron residuos de aminoácido que se prevé que son críticos para la unión a ecdisteroides y crearon casetes de expresión génica de ADNc de GAL4/CfEcR-DEF mutante tal como se describió en el ejemplo 1 anteriormente usando mutagénesis dirigida a sitio mediada por PCR. Se prepararon ADNc mutados y de tipo natural correspondientes a los diversos constructos de interruptor expuestos anteriormente en los ejemplos 1.1 y 1.2 y se sometieron a prueba en ensayos de indicador de luciferasa dirigido por GAL4 tal como se describió en el ejemplo 2 anteriormente. Se calculó la actividad en veces dividiendo RLU en presencia de ligando entre RLU en ausencia del ligando.

50 Se identificaron diecisiete residuos de aminoácido que, cuando se sustituyen, producen un receptor de ecdisona mutante que presenta disminución de actividad en respuesta a ligandos tanto no esteroideos como esteroideos. El efecto de una sustitución de aminoácido en el residuo de aminoácido 21, 48, 51, 59, 62, 93, 95, 109, 120, 123, 125, 218, 219, 223, 230, 234 ó 238 de SEQ ID NO: 1 sobre la actividad del receptor CfEcR-DEF mutado se presenta en la tabla 12 como aumento en veces con respecto a la actividad de interruptor de Gal4/ CfEcRDEF de tipo natural (WT). Además, se prepararon dos mutantes dobles (R95A/A110P y M218A/C219A) y un mutante triple (V107I/A110P/R175E) y también se identificaron como receptores CfEcR-DEF mutados que presentan disminución de actividad en respuesta a ligandos tanto no esteroideos como ligandos esteroideos (véase la tabla 12).

Tabla 12. Mutantes que muestran disminución de actividad esteroidea y no esteroidea. Aumento en veces con respecto al tipo natural

MUTANTES	GS™-E 2,5 µM	PonA 2,5 µM	GS™-E 10 nM	PonA 10 nM
Q21A	0,32	0,37		
F48A	0,007	0,007		
I51A	0,003	0,004		
V59A	0,47	0,002		
I62A	0,12	0,004		
M93A	0,46	0,07		
R95A	0,4	0,006		
F109A	0,22	0,005		
Y120A	0,001	0,006		
A123F	0,09	0,005		
M125A	0,005	0,007		
M218A	0,001	0,001		
C219A	0,001	0,001		
L223A	0,118	0,007		
L230A	0,001	0,006		
L234A	0,001	0,006		
W238A	0,002	0,013		
R95A/A110P	0,4	0,007		
M218A/C219A V107I/A110P/R175E	0,001	0,001	0,345	nd*

\* No detectable

## 5 EJEMPLO 9

Este ejemplo describe la introducción de mutaciones por sustitución dentro del EcR de *Drosophila melanogaster* (DmEcR) en residuos de aminoácido dentro del dominio de unión a ligando de DmEcR que son análogos a los mutantes por sustitución del dominio de unión a ligando de CfEcR identificados anteriormente. Específicamente, se introdujeron mutaciones por sustitución en residuos de aminoácido de DmEcR 107, 121, 213 y 217 de SEQ ID NO: 2, correspondiente a residuos de aminoácido de CfEcR 110, 124, 211 y 219 de SEQ ID NO: 1, respectivamente.

Los solicitantes mutaron residuos de aminoácido que se prevé que son críticos para la unión a ecdisteroides y crearon casetes de expresión génica de ADNc de GAL4/DmEcR-CDEF mutante tal como se describió en el ejemplo 1 anteriormente usando mutagénesis dirigida a sitio mediada por PCR. Se prepararon ADNc mutados y de tipo natural correspondientes a los diversos constructos de interruptor expuestos anteriormente en los ejemplos 1.8 y 1.9 y se sometieron a prueba en ensayos de indicador en células NIH3T3 tal como se describió en el ejemplo 2. Se transfectoron cada constructo de GAL4/DmEcRCDEF, VP16/LmUSP-EF y pFRLUC en células NIH3T3 y se trataron las células transfectadas con GS™-E o ponasterona A 2,5 µM. Se recogieron las células y se midió la actividad indicadora a 48 horas tras la adición de ligando. Se calculó la inducción en veces dividiendo la actividad indicadora en presencia de ligando entre la actividad indicadora en ausencia de ligando. A partir de la inducción en veces, se calculó el porcentaje de actividad de tipo natural para cada mutante. Los resultados se presentan en la tabla 13.

Tabla 13. Tipo natural de GAL4/DmEcR-CDEF y mutantes por sustitución G121R, G121L, G217A y C217S sometidos a prueba para la transactivación en células NIH3T3. Aumento en veces con respecto al tipo natural:

Mutante de DmEcR-CDEF	Ponasterona A 2,5 µM	GS™-E 2,5 µM
G121R	0,05	0,0075
G121L	0,001	0,008
C217A	0,022	0,008
C217S	0,0064	0,014

Tal como se observa en la tabla 13, se disminuyeron de manera significativa las actividades tanto no esteroidea como esteroidea cuando se mutó el dominio de unión a ligando de DmEcR en los residuos de aminoácido 121 ó 217, indicando que estos residuos son residuos importantes en la bolsa de unión a ligando de DmEcR.

5 También se usaron los receptores DmEcR-CDEF mutante y de tipo natural para preparar constructos de VP16/DmEcR-CDEF mutante o de tipo natural tal como se describió en los ejemplos 1.10 y 1.11. Se transfectoron VP16DmEcR-CDEF y un indicador de 6XE $\beta$ -gal en células L57 y se trataron las células transfectadas con 20-hidroxicdisona (20E) o GS<sup>TM</sup>-E 1  $\mu$ M. Se recogieron las células, se lisaron y se midió la actividad indicadora tal como se describió anteriormente en el ejemplo 6. Se calculó la inducción en veces dividiendo la actividad indicadora en presencia de ligando entre la actividad indicadora en ausencia de ligando. A partir de la inducción en veces, se calculó el porcentaje de actividad de tipo natural para cada mutante. Los resultados se presentan en la tabla 14.

15 Tabla 14. Tipo natural de VP16/DmEcR-CDEF y mutantes por sustitución A107P, G121R, G121L, N213A, G217A y C217S sometidos a prueba para determinar la transactivación en células L57 de insecto. Aumento en veces con respecto a tipo natural:

Mutante DmEcR-CDEF	20-hidroxicdisona 1 $\mu$ M	GS <sup>TM</sup> -E 1 $\mu$ M
A107P	0,09	0,9
G121R	0,5	0,92
G121L	0,09	0,15
N213A	0,01	0,08
C217A	0,48	0,70
C217S	0,39	0,92

20 La mutación A107P de DmEcR provocó la pérdida de la mayor parte de la actividad esteroidea pero tuvo muy poco efecto sobre la actividad no esteroidea. Las mutaciones G121R y C217S de DmEcR dieron como resultado reducciones del 50% y del 61% respectivamente en la actividad esteroidea pero un efecto mínimo sobre la actividad no esteroidea. La mutación C217A de DmEcR dio como resultado la reducción de actividades no esteroidea y esteroidea, y los mutantes G121L y N213A de DmEcR perdieron la sensibilidad tanto para esteroides como para compuestos no esteroideos, indicando que estos residuos están implicados en la unión tanto a esteroides como a compuestos no esteroideos.

#### 25 EJEMPLO 10

30 Este ejemplo describe la introducción de mutaciones por sustitución dentro del EcR de *Amblyomma americanum* (AmaEcR) en residuos de aminoácido dentro del dominio de unión a ligando de AmaEcR que son análogos a los mutantes por sustitución de dominio de unión a ligando CfEcR identificados anteriormente. Específicamente, se introdujeron mutaciones por sustitución en los residuos de aminoácido de AmaEcR 91 y 105 de SEQ ID NO: 3, correspondiente a residuos de aminoácido de CfEcR 96 y 110 de SEQ ID NO: 1, respectivamente.

35 Los solicitantes mutaron residuos de aminoácido que se prevé que son críticos para la unión a ecdisteroides y crearon casetes de expresión génica de ADNc de GAL4/AmaEcR-DEF mutante tal como se describió en el ejemplo 1 anterior usando mutagénesis dirigida a sitio mediada por PCR. Se sometieron a prueba ADNc mutados y de tipo natural correspondientes a los diversos constructos de interruptor expuestos anteriormente en los ejemplos 1.12 y 1.13 en ensayos de indicador de luciferasa dirigido por GAL4 en células NIH3T3 tal como se describió en el ejemplo 2. Se transfectoron GAL4/AmaEcR-DEF, VP16LmUSP-EF y pFRLUC en células NIH3T3 y se trataron las células transfectadas o bien con ligando esteroideo ponasterona A 0,2  $\mu$ M o bien con ligando no esteroideo GS<sup>TM</sup>-E 1  $\mu$ M. Se recogieron las células y se midió la actividad indicadora a 48 horas tras la adición de ligando. Se calculó la inducción en veces dividiendo la actividad indicadora en presencia de ligando entre la actividad indicadora en ausencia de ligando. A partir de la inducción en veces, se calculó el porcentaje de actividad de tipo natural para cada mutante. Los resultados se presentan en la tabla 15.

45 Tabla 15. Mutantes por sustitución de AmaEcR-DEF en G91 y A105 en células NH3T3. Aumento en veces con respecto al tipo natural:

Mutante de AmaEcR-DEF	Ponasterona A 0,2 $\mu$ M	GS <sup>TM</sup> -E 1 $\mu$ M
G91A	1,29	1,22
A105P	0,11	0,01

La mutación G91A de AmaEcR en la posición de residuo de aminoácido homóloga a V96 en CfEcR dio como resultado el aumento en las actividades esteroidea y no esteroidea. La mutación A105P de AmaEcR en la posición de residuo de aminoácido homóloga a A110 de CfEcR provocó la pérdida de la mayor parte de la actividad esteroidea y eliminó esencialmente la actividad no esteroidea.

5 Debe entenderse además que los tamaños de bases o tamaños de aminoácidos, y todos los valores de peso molecular o masa molecular, dados para ácidos nucleicos o polipéptidos son aproximados, y se proporcionan para la descripción.

10 **Lista de secuencias**

<110> Rohm and Haas Company

Palli, Subba R.

Cress, Dean E.

15 Fujimoto, Ted T.

Kumar, Mohan B.

<120> Receptores mutantes por sustitución novedosos y su uso en un sistema de expresión génica inducible basado en receptores nucleares

<130> A01247

20 <150> Documento US 60/313925

<151> 21-08-2001

<160> 113

<170> Patent In versión 3.1

<210> 1

25 <211> 240

<212> PRT

<213> *Choristoneura fumiferana*

<400> 1

```

Leu Thr Ala Asn Gln Gln Phe Leu Ile Ala Arg Leu Ile Trp Tyr Gln
 1          5          10          15

Asp Gly Tyr Glu Gln Pro Ser Asp Glu Asp Leu Lys Arg Ile Thr Gln
          20          25          30

Thr Trp Gln Gln Ala Asp Asp Glu Asn Glu Glu Ser Asp Thr Pro Phe
          35          40          45

Arg Gln Ile Thr Glu Met Thr Ile Leu Thr Val Gln Leu Ile Val Glu
          50          55          60

Phe Ala Lys Gly Leu Pro Gly Phe Ala Lys Ile Ser Gln Pro Asp Gln
 65          70          75          80

Ile Thr Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met Leu Arg Val
          85          90          95

Ala Arg Arg Tyr Asp Ala Ala Ser Asp Ser Val Leu Phe Ala Asn Asn
          100          105          110

Gln Ala Tyr Thr Arg Asp Asn Tyr Arg Lys Ala Gly Met Ala Tyr Val
          115          120          125

Ile Glu Asp Leu Leu His Phe Cys Arg Cys Met Tyr Ser Met Ala Leu
          130          135          140
    
```

ES 2 422 303 T3

Asp Asn Ile His Tyr Ala Leu Leu Thr Ala Val Val Ile Phe Ser Asp  
 145 150 155 160

Arg Pro Gly Leu Glu Gln Pro Gln Leu Val Glu Glu Ile Gln Arg Tyr  
 165 170 175

Tyr Leu Asn Thr Leu Arg Ile Tyr Ile Leu Asn Gln Leu Ser Gly Ser  
 180 185 190

Ala Arg Ser Ser Val Ile Tyr Gly Lys Ile Leu Ser Ile Leu Ser Glu  
 195 200 205

Leu Arg Thr Leu Gly Met Gln Asn Ser Asn Met Cys Ile Ser Leu Lys  
 210 215 220

Leu Lys Asn Arg Lys Leu Pro Pro Phe Leu Glu Glu Ile Trp Asp Val  
 225 230 235 240

<210> 2

<211> 237

<212> PRT

5 <213> *Drosophila melanogaster*

<400> 2

Leu Thr Tyr Asn Gln Leu Ala Val Ile Tyr Lys Leu Ile Trp Tyr Gln  
 1 5 10 15

Asp Gly Tyr Glu Gln Pro Ser Glu Glu Asp Leu Arg Arg Ile Met Ser  
 20 25 30

Gln Pro Asp Glu Asn Glu Ser Gln Thr Asp Val Ser Phe Arg His Ile  
 35 40 45

Thr Glu Ile Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala Lys  
 50 55 60

Gly Leu Pro Ala Phe Thr Lys Ile Pro Gln Glu Asp Gln Ile Thr Leu  
 65 70 75 80

Leu Lys Ala Cys Ser Ser Glu Val Met Met Leu Arg Met Ala Arg Arg  
 85 90 95

Tyr Asp His Ser Ser Asp Ser Ile Phe Phe Ala Asn Asn Arg Ser Tyr  
 100 105 110

ES 2 422 303 T3

Thr Arg Asp Ser Tyr Lys Met Ala Gly Met Ala Asp Asn Ile Glu Asp  
115 120 125

Leu Leu His Phe Cys Arg Gln Met Phe Ser Met Lys Val Asp Asn Val  
130 135 140

Glu Tyr Ala Leu Leu Thr Ala Ile Val Ile Phe Ser Asp Arg Pro Gly  
145 150 155 160

Leu Glu Lys Ala Gln Leu Val Glu Ala Ile Gln Ser Tyr Tyr Ile Asp  
165 170 175

Thr Leu Arg Ile Tyr Ile Leu Asn Arg His Cys Gly Asp Ser Met Ser  
180 185 190

Leu Val Phe Tyr Ala Lys Leu Leu Ser Ile Leu Thr Glu Leu Arg Thr  
195 200 205

Leu Gly Asn Gln Asn Ala Glu Met Cys Phe Ser Leu Lys Leu Lys Asn  
210 215 220

Arg Lys Leu Pro Lys Phe Leu Glu Glu Ile Trp Asp Val  
225 230 235

<210> 3

<211> 240

5 <212> PRT

<213> *Amblyomma americanum*

<400> 3

Pro Gly Val Lys Pro Leu Ser Ser Ser Gln Glu Asp Leu Ile Asn Lys  
1 5 10 15

Leu Val Tyr Tyr Gln Gln Glu Phe Glu Ser Pro Ser Glu Glu Asp Met  
20 25 30

Lys Lys Thr Thr Pro Phe Pro Leu Gly Asp Ser Glu Glu Asp Asn Gln  
35 40 45

Arg Arg Phe Gln His Ile Thr Glu Ile Thr Ile Leu Thr Val Gln Leu  
50 55 60

Ile Val Glu Phe Ser Lys Arg Val Pro Gly Phe Asp Thr Leu Ala Arg  
65 70 75 80

Glu Asp Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met  
85 90 95



ES 2 422 303 T3

Leu Arg Gly Ala Arg Lys Tyr Asp Val Lys Thr Asp Ser Ile Val Phe  
 100 105 110

Ala Asn Asn Gln Pro Tyr Thr Arg Asp Asn Tyr Arg Ser Ala Ser Val  
 115 120 125

Gly Asp Ser Ala Asp Ala Leu Phe Arg Phe Cys Arg Lys Met Cys Gln  
 130 135 140

Leu Arg Val Asp Asn Ala Glu Tyr Ala Leu Leu Thr Ala Ile Val Ile  
 145 150 155 160

Phe Ser Glu Arg Pro Ser Leu Val Asp Pro His Lys Val Glu Arg Ile  
 165 170 175

Gln Glu Tyr Tyr Ile Glu Thr Leu Arg Met Tyr Ser Glu Asn His Arg  
 180 185 190

Pro Pro Gly Lys Asn Tyr Phe Ala Arg Leu Leu Ser Ile Leu Thr Glu  
 195 200 205

Leu Arg Thr Leu Gly Asn Met Asn Ala Glu Met Cys Phe Ser Leu Lys  
 210 215 220

Val Gln Asn Lys Lys Leu Pro Pro Phe Leu Ala Glu Ile Trp Asp Ile  
 225 230 235 240

<210> 4

<211> 198

<212> ADN

5 <213> *Choristoneura fumiferana*

<400> 4

tgctctggtat gcggggacag agcctccgga taccactaca atgcgctcac gtgtgaaggg 60

tgtaaagggt tcttcagacg gagtgttacc aaaaatgctgg tttatatttg taaattcggg 120

cacgcttgcg aaatggacat gtacatgcca cggaaatgcc aggagtgccg cctgaagaag 180

tgettagctg taggcatg 198

<210> 5

<211> 66

10 <212> PRT

<213> *Choristoneura fumiferana*

<400> 5

Cys Leu Val Cys Gly Asp Arg Ala Ser Gly Tyr His Tyr Asn Ala Leu

1 5 10 15

Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Val Thr Lys Asn  
20 25 30

Ala Val Tyr Ile Cys Lys Phe Gly His Ala Cys Glu Met Asp Met Tyr  
35 40 45

Met Arg Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Ala Val  
50 55 60

Gly Met  
65

<210> 6

<211> 441

<212> ADN

5 <213> *Saccharomyces cerevisiae*

<400> 6

atgaagctac tgtcttctat cgaacaagca tgcgatattt gccgacttaa aaagctcaag 60  
tgctccaaag aaaaaccgaa gtgcgccaag tgtctgaaga acaactggga gtgtcgctac 120  
tctcccaaaa ccaaagggtc tccgctgact agggcacatc tgacagaagt ggaatcaagg 180  
ctagaaagac tggaacagct atttctactg atttttcttc gagaagacct tgacatgatt 240  
ttgaaaatgg attctttaca ggatataaaa gcattgttaa caggattatt tgtacaagat 300  
aatgtgaata aagatgccgt cacagataga ttggcttcag tggagactga tatgcctcta 360  
acattgagac agcatagaat aagtgcgaca tcatcatcgg aagagagtag taacaaaggt 420  
caaagacagt tgactgtatc g 441

<210> 7

<211> 147

<212> PRT

10 <213> *Saccharomyces cerevisiae*

<400> 7

Met Lys Leu Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu  
 1 5 10 15

Lys Lys Leu Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu  
 20 25 30

Lys Asn Asn Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro  
 35 40 45

Leu Thr Arg Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu  
 50 55 60

Glu Gln Leu Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile  
 65 70 75 80

Leu Lys Met Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu  
 85 90 95

Phe Val Gln Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala  
 100 105 110

Ser Val Glu Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser  
 115 120 125

Ala Thr Ser Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu  
 130 135 140

Thr Val Ser  
 145

- <210> 8
- <211> 606
- <212> ADN
- <213> *Escherichia coli*
- <400> 8

5

ES 2 422 303 T3

```

atgaaagcgt taacggccag gcaacaagag gtgtttgatc tcatccgtga tcacatcagc    60
cagacaggta tgccgccgac gcgtgcggaa atcgcgcagc gtttgggggt ccgttcccca    120
aacgcggctg aagaacatct gaaggcgctg gcacgcaaag gcgttattga aattgtttcc    180
ggcgcatcac gcgggattcg tctgttgagc gaagaggaag aagggttgcc gctggtaggt    240
cgtgtggctg ccggtgaacc acttctggcg caacagcata ttgaaggcca ttatcaggtc    300
gatccttctt tattcaagcc gaatgctgat ttctgctgc gcgtcagcgg gatgtcgatg    360
aaagatatcg gcattatgga tgggtgacttg ctggcagtgc ataaaactca ggatgtacgt    420
aacggtcagg tcgttgctgc acgtattgat gacgaagtta ccgtaagcg cctgaaaaaa    480
cagggcaata aagtcgaact gttgccagaa aatagcgagt ttaaaccaat tgtcgtagat    540
cttcgtcagc agagcttcac cattgaaggg ctggcggttg gggttattcg caacggcgac    600
_tggctg                                     . . . . .                    606

```

```

<210> 9
<211> 202
<212> PRT
5 <213> Escherichia coli
<400> 9

```

ES 2 422 303 T3

Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp Leu Ile Arg  
 1 5 10 15

Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala  
 20 25 30

Gln Arg Leu Gly Phe Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys  
 35 40 45

Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly Ala Ser Arg  
 50 55 60

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

Arg Val Ala Ala Gly Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly  
 85 90 95

His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu  
 100 105 110

Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly  
 115 120 125

Asp Leu Leu Ala Val His Lys Thr Gln Asp Val Arg Asn Gly Gln Val  
 130 135 140

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

Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro  
 165 170 175

Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala  
 180 185 190

Val Gly Val Ile Arg Asn Gly Asp Trp Leu  
 195 200

<210> 10

<211> 420

<212> ADN

5 <213> *Choristoneura fumiferana*

<400> 10

ES 2 422 303 T3

atgagacgcc gctgggtccaa caacgggggc ttccagacgc tgcgaatgct cgaggagagc 60  
 tcgtccgaag tgacgtcgtc ctcagctctg ggtctgccgg ccgcatggt tatgtctccg 120  
 gagtcgctcg cctcggcaga gtacggcggg ctcgagctct ggggatacga cgatgggttg 180  
 tcatacaaca cggcgcagtc cttgctgggc aatacttgca cgatgcagca gcagcaacag 240  
 acgcagccgc tgccgtcgat gccgttgct atgccgccga ccacgccgaa gtctgaaaac 300  
 gagtctattt cctcaggccg tgaggaactg tcgccagctt caagtataaa tgggtgcagt 360  
 acagatggcg aggcacgacg tcagaagaag ggcctgccc cccgtcagca agaggaactg 420

<210> 11

<211> 140

<212> PRT

5 <213> *Choristoneura fumiferana*

<400> 11

Met Arg Arg Arg Trp Ser Asn Asn Gly Gly Phe Gln Thr Leu Arg Met  
 1 5 10 15  
 Leu Glu Glu Ser Ser Ser Glu Val Thr Ser Ser Ser Ala Leu Gly Leu  
 20 25 30  
 Pro Ala Ala Met Val Met Ser Pro Glu Ser Leu Ala Ser Pro Glu Tyr  
 35 40 45  
 Gly Gly Leu Glu Leu Trp Gly Tyr Asp Asp Gly Leu Ser Tyr Asn Thr  
 50 55 60  
 Ala Gln Ser Leu Leu Gly Asn Thr Cys Thr Met Gln Gln Gln Gln Gln  
 65 70 75 80  
 Thr Gln Pro Leu Pro Ser Met Pro Leu Pro Met Pro Pro Thr Thr Pro  
 85 90 95  
 Lys Ser Glu Asn Glu Ser Ile Ser Ser Gly Arg Glu Glu Leu Ser Pro  
 100 105 110  
 Ala Ser Ser Ile Asn Gly Cys Ser Thr Asp Gly Glu Ala Arg Arg Gln  
 115 120 125  
 Lys Lys Gly Pro Ala Pro Arg Gln Gln Glu Glu Leu  
 130 135 140

<210> 12

<211> 271

<212> ADN

10 <213> virus del herpes simple 7

<400> 12

atggggccta aaaagaagcg taaagtcgcc cccccgaccg atgtcagcct gggggacgag 60  
 ctccacttag acggcgagga cgtggcgatg gcgcatgccg acgcgctaga cgatttcgat 120  
 ctggacatgt tgggggacgg ggattccccg gggccgggat ttacccccca cgactccgcc 180  
 ccctacggcg ctctggatat ggccgacttc gagtttgagc agatgtttac cgatgcctt 240  
 ggaattgacg agtacggtgg ggaattccccg g 271

<210> 13

<211> 90

<212> PRT

5 <213> virus del herpes simple 7

<400> 13

Met Gly Pro Lys Lys Lys Arg Lys Val Ala Pro Pro Thr Asp Val Ser  
 1 5 10 15

Leu Gly Asp Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His  
 20 25 30

Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp  
 35 40 45

Ser Pro Gly Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala  
 50 55 60

Leu Asp Met Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu  
 65 70 75 80

Gly Ile Asp Glu Tyr Gly Gly Glu Phe Pro  
 85 90

<210> 14

<211> 307

10 <212> ADN

<213> *Saccharomyces cerevisiae*

<400> 14

atgggtgctc ctccaaaaaa gaagagaaag gtagctggta tcaataaaga tatcgaggag 60  
 tgcaatgcca tcattgagca gtttatcgac tacctgcgca cggacagga gatgccgatg 120  
 gaaatggcgg atcaggcgat taacgtggtg ccgggcatga cgccgaaaac cattcttcac 180  
 gccgggcccgc cgatccagcc tgactggctg aaatcgaatg gttttcatga aattgaagcg 240  
 gatgttaacg ataccagcct cttgctgagt ggagatgcct cctaccctta tgatgtgcca 300  
 gattatg 307

<210> 15

15 <211> 102

<212> PRT

ES 2 422 303 T3

<213> *Saccharomyces cerevisiae*

<400> 15

Met Gly Ala Pro Pro Lys Lys Lys Arg Lys Val Ala Gly Ile Asn Lys  
 1 5 10 15

Asp Ile Glu Glu Cys Asn Ala Ile Ile Glu Gln Phe Ile Asp Tyr Leu  
 20 25 30

Arg Thr Gly Gln Glu Met Pro Met Glu Met Ala Asp Gln Ala Ile Asn  
 35 40 45

Val Val Pro Gly Met Thr Pro Lys Thr Ile Leu His Ala Gly Pro Pro  
 50 55 60

Ile Gln Pro Asp Trp Leu Lys Ser Asn Gly Phe His Glu Ile Glu Ala  
 65 70 75 80

Asp Val Asn Asp Thr Ser Leu Leu Leu Ser Gly Asp Ala Ser Tyr Pro  
 85 90 95

Tyr Asp Val Pro Asp Tyr  
 100

<210> 16

5 <211> 807

<212> ADN

<213> *Homo sapiens*

<400> 16



ES 2 422 303 T3

```

cccatggaat tccagtacct gccagataca gacgatcgtc accggattga ggagaaacgt      60
aaaaggacat atgagacctt caagagcadc atgaagaaga gtcctttcag cggacccacc      120
gacccccggc ctcacactcg acgcattgct gtgccttccc gcagctcagc ttctgtcccc      180
aagccagcac cccagcccta tccctttacg tcatccctga gcacatcaa ctatgatgag      240
tttcccacca tgggtgtttcc ttctgggcag atcagccagg cctcggcctt ggccccggcc      300
cctcccccaag tcttgccccca ggtccagcc cctgcccctg ctccagccat ggtatcagct      360
ctggcccagg ccccagcccc tgtcccagtc ctagccccag gccctcctca ggctgtggcc      420
ccacctgccc ccaagcccac ccaggctggg gaaggaacgc tgtcagaggc cctgctgcag      480
ctgcagtttg atgatgaaga cctggggggcc ttgcttgcca acagcacaga cccagctgtg      540

ttcacagacc tggcatccgt cgacaactcc gagtttcagc agctgctgaa ccagggcata      600
cctgtggccc cccacacaac tgagcccatg ctgatggagt accctgaggc tataactcgc      660
ctagtgcacag gggcccagag gccccccgac ccagctcctg ctccactggg ggccccgggg      720
ctccccaatg gcctccttc aggagatgaa gacttctcct ccattgcgga catggacttc      780
tcagccctgc tgagtcagat cagctcc                                          807

```

<210> 17

<211> 269

<212> PRT

5 <213> *Homo sapiens*

<400> 17

ES 2 422 303 T3

Pro Met Glu Phe Gln Tyr Leu Pro Asp Thr Asp Asp Arg His Arg Ile  
 1 5 10 15

Glu Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys  
 20 25 30

Lys Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg Pro Pro Arg Arg  
 35 40 45

Ile Ala Val Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala Pro  
 50 55 60

Gln Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu  
 65 70 75 80

Phe Pro Thr Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala  
 85 90 95

Leu Ala Pro Ala Pro Pro Gln Val Leu Pro Gln Ala Pro Ala Pro Ala  
 100 105 110

Pro Ala Pro Ala Met Val Ser Ala Leu Ala Gln Ala Pro Ala Pro Val  
 115 120 125

Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val Ala Pro Pro Ala Pro  
 130 135 140

Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln  
 145 150 155 160

Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr  
 165 170 175

Asp Pro Ala Val Phe Thr Asp Leu Ala Ser Val Asp Asn Ser Glu Phe  
 180 185 190

Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr Glu  
 195 200 205

Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr Arg Leu Val Thr Gly  
 210 215 220

Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly  
 225 230 235 240

Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala  
 245 250 255

Asp Met Asp Phe Ser Ala Leu Leu Ser Gln Ile Ser Ser  
 260 265

<210> 18

<211> 34

5 <212> ADN

<213> Secuencia artificial

<220>

<223> 1X elemento de respuesta a ecdisona (EcRE)

<400> 18

tcgagagaca agggttcaat gcactgtcc aatg 34  
 <210> 19  
 <211> 19  
 <212> ADN  
 5 <213> Secuencia artificial  
 <220>  
 <223> elemento de respuesta a GAL4  
 <400> 19  
 ggagtactgt cctccgagc 19  
 10 <210> 20  
 <211> 36  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 15 <223> 2x elemento de respuesta a LexAop  
 <400> 20  
 ctgctgtata taaaaccagt ggttatatgt acagta 36  
 <210> 21  
 <211> 1054  
 20 <212> ADN  
 <213> *Choristoneura fumiferana*  
 <400> 21

```

cctgagtgcg tagtaccgga gactcagtgc gccatgaagc ggaaagagaa gaaagcacag      60
aaggagaagg aaaaactgcc tgtcagcacg acgacgggtgg acgaccacat gccgcccatt      120
atgcagtgtg aacctccacc tcctgaagca gcaaggattc acgaagtggc cccaagggtt      180
ctctccgaca agctggttga gacaaaccgg cagaaaaaca tccccagtt gacagccaac      240
cagcagttcc ttatcgccag gctcatctgg taccaggacg ggtacgagca gccttctgat      300
gaagatttga agaggattac gcagacgtgg cagcaagcgg acgatgaaaa cgaagagtct      360
gacactccct tccgccagat cacagagatg actatcctca cggccaact tatcgtggag      420
ttcgcaagg gattgccagg gttcgccaag atctcgcagc ctgatcaaat tacgctgctt      480
aaggcttgct caagtgaggc aatgatgctc cgagtcgcgc gacgatacga tgcggcctca      540
gacagtgttc tgttcgcgaa caaccaagcg tacactcgcg acaactaccg caaggctggc      600
atggcctacg tcategagga tctactgcac ttctgccggt gcatgtactc tatggcgttg      660
gacaacatcc attaegcgtc gctcacggct gtcgtcatct tttctgaccg gccagggttg      720
gagcagccgc aactggttga agaaatccag cggtaactacc tgaatacgtc ccgcatctat      780
atcctgaacc agctgagcgg gtcggcgcgt tcgtccgtca tatacggcaa gatcctctca      840
atcctctctg agctacgcac gctcggcatg caaaactcca acatgtgcat ctcctcaag      900
ctcaagaaca gaaagctgcc gcctttcctc gaggagatct gggatgtggc ggacatgtcg      960
cacaccaaac cgccgcctat cctcgagtec cccacgaatc tctagcccct gcgcgcacgc     1020
atcgccgatg ccgcgtccgg ccgcgtgct ctga                                     1054

```

25 <210> 22  
 <211> 309  
 <212> ADN  
 <213> Virus del simio 40  
 <400> 22

ES 2 422 303 T3

gggtgtgaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 60  
 agtcagcaac caggtgtgga aagtccccag gctccccage aggcagaagt atgcaaagca 120  
 tgcattctcaa ttagtcagca accatagtcg cgccccctaac tccgccccatc ccgccccctaa 180  
 ctccgcccag tccgccccat tctccgcccc atggctgact aatttttttt atttatgcag 240  
 aggccgaggc cgcctcggcc tctgagctat tccagaagta gtgaggaggc ttttttggag 300  
 gcctaggt 309

<210> 23

<211> 635

<212> ADN

5 <213> *Locusta migratoria*

<400> 23

tgcatacaga catgcctggt gaacgcatac ttgaagctga aaaacgagtg gagtgcaaag 60  
 cagaaaacca agtggaaatag gagctgggtg agtgggctaa acacatcccg cacttcacat 120  
 ccctacctct ggaggaccag gttctcctcc tcagagcagg ttggaatgaa ctgctaattg 180  
 cagcattttc acatcgatct gtagatgtta aagatggcat agtacttgcc actggtctca 240  
 cagtgcacag aaattctgcc catcaagctg gagtcggcac aatatttgac agagttttga 300  
 cagaactggt agcaaagatg agagaaatga aatggataa aactgaactt ggctgcttgc 360  
 gatctgttat tcttttcaat ccagaggtga ggggtttgaa atccgcccag gaagttgaac 420  
 ttctacgtga aaaagtatat gccgctttgg aagaatatac tagaacaaca catcccgatg 480  
 aaccaggaag atttgcaaaa cttttgcttc gtctgccttc tttacgttcc ataggcctta 540  
 agtgtttggg gcatttgttt ttctttcgcc ttattggaga tgttccaatt gatacgttcc 600  
 tgatggagat gcttgaatca ccttctgatt cataa 635

<210> 24

<211> 24

10 <212> ADN

<213> Secuencia artificial

<220>

<223> Promotor mínimo de E1b sintético

<400> 24

15 tatataatgg atccccggg accg 24

<210> 25

<211> 1653

<212> ADN

<213> *Photinus pyralis*

20 <400> 25

atggaagacg ccaaaaacat aaagaaaggc cggcgccat tctatcctct agaggatgga 60  
 accgctggag agcaactgca taaggctatg aagagatacg ccttgggtcc tggaacaatt 120  
 gcttttacag atgcacatat cgaggtgaac atcacgtacg cggaatactt cgaaatgtcc 180  
 gttcggttgg cagaagctat gaaacgatat gggctgaata caaatcacag aatcgctgta 240  
 tgcagtgaaa actctcttca attctttatg ccggtgttgg gcgcttatt tatcggagtt 300  
 gcagttgcgc ccgogaacga catttataat gaacgtgaat tgctcaacag tatgaacatt 360  
  
 tgcgagccta cegtagtgtt tgtttccaaa aaggggttgc aaaaaathtt gaacgtgcaa 420  
 aaaaaattac caataatcca gaaaattatt atcatggatt ctaaacgga ttaccagggga 480  
 tttcagtcga tgtacacgtt cgtcacatct catctacctc ccggttttaa tgaatacgat 540  
 tttgtaccag agtcttttga tcgtgacaaa acaattgcac tgataatgaa ttctcttggga 600  
 tctactgggt tacctaaggg tgtggccctt ccgcatagaa ctgcctgcgt cagattctcg 660  
 catgccagag atcctathtt tggcaatcaa atcattccgg atactgcat ttaagtgtt 720  
 gttccattcc atcacggttt tggaatgttt actacactcg gatatttgat atgtggattt 780  
 cgagtcgtct taatgtatag atttgaagaa gagctgtttt tacgatccct tcaggattac 840  
 aaaattcaaa gtgcgttgcg agtaccaccc ctathttcat tcttcgcaa aagcactctg 900  
 attgacaaat acgatttatc taatttacac gaaattgctt ctgggggcgc acctcttctg 960  
 aaagaagtcg ggggaagcgt tgcaaacgc ttccatcttc caggatagc acaaggatat 1020  
 gggctcactg agactacatc agctattctg attacaccg aggggatga taaaccgggc 1080  
 gcggtcggta aagttgttcc attttttgaa gcgaaggtt tggatctgga taccgggaaa 1140  
 acgctgggcg ttaatcagag aggcgaatta tgtgtcagag gacctatgat tatgtccggt 1200  
 tatgtaaaca atccggaagc gaccaacgcc ttgattgaca aggatggatg gctacattct 1260  
 ggagacatag cttactggga cgaagacgaa cacttctca tagtgaccg cttgaagtct 1320  
 ttaattaaat acaaaggata tcaggtggcc ccgctgaat tggaatcgat attgttacia 1380  
 caccccaaca tcttcgacgc gggcgtggca ggtcttccc acgatgacgc cggatgaactt 1440  
 cccgccgccc ttgtgtttt ggagcacgga aagacgatga cggaaaaaga gatcggtgat 1500  
 tacgtcgcca gtcaagtaac aaccgcgaaa aagttgcgcg gaggagtgt gtttgtggac 1560  
 gaagtaccga aaggtcttac cggaaaactc gacgcaagaa aatcagaga gatcctcata 1620  
 aaggccaaga agggcggaaa gtccaaattg taa 1653

<210> 26

<211> 541

5 <212> PRT

<213> *Choristoneura fumiferana*

<400> 26

ES 2 422 303 T3

Met Arg Arg Arg Trp Ser Asn Asn Gly Gly Phe Gln Thr Leu Arg Met  
1 5 10 15

Leu Glu Glu Ser Ser Ser Glu Val Thr Ser Ser Ser Ala Leu Gly Leu  
20 25 30

Pro Ala Ala Met Val Met Ser Pro Glu Ser Leu Ala Ser Pro Glu Tyr

ES 2 422 303 T3

	35		40		45															
Gly	Gly	Leu	Glu	Leu	Trp	Gly	Tyr	Asp	Asp	Gly	Leu	Ser	Tyr	Asn	Thr					
	50					55					60									
Ala	Gln	Ser	Leu	Leu	Gly	Asn	Thr	Cys	Thr	Met	Gln	Gln	Gln	Gln	Gln					
65					70					75										80
Thr	Gln	Pro	Leu	Pro	Ser	Met	Pro	Leu	Pro	Met	Pro	Pro	Thr	Thr	Pro					
				85					90						95					
Lys	Ser	Glu	Asn	Glu	Ser	Ile	Ser	Ser	Gly	Arg	Glu	Glu	Leu	Ser	Pro					
			100					105						110						
Ala	Ser	Ser	Ile	Asn	Gly	Cys	Ser	Thr	Asp	Gly	Glu	Ala	Arg	Arg	Gln					
			115					120					125							
Lys	Lys	Gly	Pro	Ala	Pro	Arg	Gln	Gln	Glu	Glu	Leu	Cys	Leu	Val	Cys					
			130			135					140									
Gly	Asp	Arg	Ala	Ser	Gly	Tyr	His	Tyr	Asn	Ala	Leu	Thr	Cys	Glu	Gly					
145					150					155					160					
Cys	Lys	Gly	Phe	Phe	Arg	Arg	Ser	Val	Thr	Lys	Asn	Ala	Val	Tyr	Ile					
				165					170						175					
Cys	Lys	Phe	Gly	His	Ala	Cys	Glu	Met	Asp	Met	Tyr	Met	Arg	Arg	Lys					
			180					185						190						
Cys	Gln	Glu	Cys	Arg	Leu	Lys	Lys	Cys	Leu	Ala	Val	Gly	Met	Arg	Pro					
		195					200					205								
Glu	Cys	Val	Val	Pro	Glu	Thr	Gln	Cys	Ala	Met	Lys	Arg	Lys	Glu	Lys					
		210					215							220						
Lys	Ala	Gln	Lys	Glu	Lys	Asp	Lys	Leu	Pro	Val	Ser	Thr	Thr	Thr	Val					
225					230					235					240					
Asp	Asp	His	Met	Pro	Pro	Ile	Met	Gln	Cys	Glu	Pro	Pro	Pro	Pro	Glu					
				245					250						255					
Ala	Ala	Arg	Ile	His	Glu	Val	Val	Pro	Arg	Phe	Leu	Ser	Asp	Lys	Leu					
			260					265						270						
Leu	Glu	Thr	Asn	Arg	Gln	Lys	Asn	Ile	Pro	Gln	Leu	Thr	Ala	Asn	Gln					
		275					280						285							

ES 2 422 303 T3

Gln Phe Leu Ile Ala Arg Leu Ile Trp Tyr Gln Asp Gly Tyr Glu Gln  
 290 295 300

Pro Ser Asp Glu Asp Leu Lys Arg Ile Thr Gln Thr Trp Gln Gln Ala  
 305 310 315 320

Asp Asp Glu Asn Glu Glu Ser Asp Thr Pro Phe Arg Gln Ile Thr Glu  
 325 330 335

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

Pro Gly Phe Ala Lys Ile Ser Gln Pro Asp Gln Ile Thr Leu Leu Lys  
 355 360 365

Ala Cys Ser Ser Glu Val Met Met Leu Arg Val Ala Arg Arg Tyr Asp  
 370 375 380

Ala Ala Ser Asp Ser Val Leu Phe Ala Asn Asn Gln Ala Tyr Thr Arg  
 385 390 395 400

Asp Asn Tyr Arg Lys Ala Gly Met Ala Tyr Val Ile Glu Asp Leu Leu  
 405 410 415

His Phe Cys Arg Cys Met Tyr Ser Met Ala Leu Asp Asn Ile His Tyr  
 420 425 430

Ala Leu Leu Thr Ala Val Val Ile Phe Ser Asp Arg Pro Gly Leu Glu  
 435 440 445

Gln Pro Gln Leu Val Glu Glu Ile Gln Arg Tyr Tyr Leu Asn Thr Leu  
 450 455 460

Arg Ile Tyr Ile Leu Asn Gln Leu Ser Gly Ser Ala Arg Ser Ser Val  
 465 470 475 480

Ile Tyr Gly Lys Ile Leu Ser Ile Leu Ser Glu Leu Arg Thr Leu Gly  
 485 490 495

Met Gln Asn Ser Asn Met Cys Ile Ser Leu Lys Leu Lys Asn Arg Lys  
 500 505 510

Leu Pro Pro Phe Leu Glu Glu Ile Trp Asp Val Ala Asp Met Ser His  
 515 520 525

Thr Gln Pro Pro Pro Ile Leu Glu Ser Pro Thr Asn Leu  
 530 535 540

<210> 27  
 <211> 1623  
 <212> ADN  
 <213> *Choristoneura fumiferana*  
 <400> 27

5



ES 2 422 303 T3

```

atgagacgcc gctgggccaa caacgggggc ttccagacgc tgcgaatgct cgaggagagc 60
tcgtccgaag tgacgtcgtc ctcagctctg ggtctgccgg ccgcatggt tatgtctccg 120
gagtcgctcg cctcggcaga gtacggcggg ctcgagctct ggggatacga cgatgggttg 180
tcatacaaca cggcgcagtc cttgctgggc aatacttgca cgatgcagca gcagcaacag 240
acgcagccgc tgccgtcgat gccgttgcc atgccgccga ccacgccga gtctgaaac 300
gagtctattt cctcaggccg tgaggaactg tcgccagctt caagtataaa tgggtgcagt 360
acagatggcg aggcacgacg tcagaagaag ggccctgccc cccgtcagca agaggaactg 420
tgtctgggat gcggggacag agcctccgga taccactaca atgcgctcac gtgtgaaggg 480
tgtaaagggg tcttcagacg gagtgttacc aaaaatgccg tttatatttg taaattcggg 540
cacgcttgcc aatggacat gtacatgcca cggaaatgcc aggagtgcc cctgaagaag 600
tgcttagctg taggcagtag gcctgagtc gtagtaccgc agactcagtg cgccatgaag 660
cggaaagaga agaaagcaca gaaggagaag gacaaactgc ctgtcagcac gacgacggg 720
gacgaccaca tgccgcccat tatgcagtg gaacctccac ctctgaagc agcaaggatt 780
cacgaagtgg tcccaagggt tctctccgac aagctgttgg agacaaaccg gcagaaaaac 840
atccccagc tgacagccaa ccagcagtc cttatcgcca ggctcatctg gtaccaggac 900
gggtacgagc agccttctga tgaagattg aagaggatta cgcagacgtg gcagcaagcg 960
gacgatgaaa acgaagagtc tgacactccc tccgccaga tcacagagat gactatectc 1020
acgggtccaa ttatcgtgga gttcgcgaag ggattgccag ggttcgccaa gatctcgag 1080
cctgatcaaa ttacgctgct taaggcttgc tcaagtgagg taatgatgct ccgagtcgcg 1140
cgacgatacg atgcggcctc agacagtggt ctgttcgcca acaaccaagc gtacactcgc 1200
gacaactacc gcaaggctgg catggcctac gtcacgagg atctactgca cttctgccgg 1260
tgcatgtact ctatggcgtt ggacaacatc cattacgccc tgctcacggc tgtcgtcatc 1320
ttttctgacc ggccagggtt ggagcagccg caactgggtg aagaaatcca gcggtactac 1380
ctgaatacgc tccgcatcta tctctgaac cagctgagcg ggtcggcgcg ttcgtccgtc 1440
atatacggca agatcctctc aatcctctct gagctacgca cgctcggcat gcaaaactcc 1500

aacatgtgca tctcctcaa gctcaagaac agaaagctgc cgccttctc cgaggagatc 1560
tgggatgtgg cggacatgtc gcacacccaa ccgccgcta tctcagatc cccacgaat 1620
ctc 1623

```

- <210> 28
- <211> 702
- 5 <212> ADN
- <213> Secuencia artificial
- <220>
- <223> Promotor de IE1 de baculovirus
- <400> 28

ES 2 422 303 T3

```

cccgggccag ttgcacaaca ctattatcga tttgcagttc gggacataaa tgtttaaata    60
tategatgtc tttgtgatgc gcgcgacatt tttgtagggtt attgataaaa tgaacggata    120
cgttgcccga cattatcatt aaatccttgg cgtagaatth gtcgggtcca ttgtccgtgt    180
gcgctagcat gcccgtaacg gacctcgtac ttttggcttc aaaggttttg cgcacagaca    240
aatgtgcca cacttgacg tctgcatgtg tgcgcgttac cacaaatccc aacggcgcag    300
tgtacttggt gtatgcaaat aaatctcgat aaaggcgcgg cgcgcgaatg cagctgatca    360
cgtacgctcc tcgtgttccg ttcaaggacg gtgttatcga cctcagatta atgtttatcg    420
gccgactggt ttcgtatccg ctcaccaaac gcgtttttgc attaacattg tatgtcggcg    480
gatgttctat atctaatttg aataaataaa cgataaccgc gttggtttta gagggcataa    540
taaaagaaat attgttatcg tgttcgcat tagggcagta taaattgacg ttcatgttgg    600
atattgtttc agttgcaagt tgacactggc ggcgacaaga tcgtgaacaa ccaagtgact    660
atagaattca ctcgaggcta gcataagatc taagctagcg cc                                702

```

<210> 29

<211> 3157

<212> ADN

5 <213> *Escherichia coli*

<400> 29

```

atgggggggt ctcacatca tcatcatcat ggtatggcta gcatgactgg tggacagcaa    60
atgggctcggg atctgtacga cgatgacgat aaggtaacct aggatcagct tggagttgat    120
cccgtcgttt tacaacgtcg tgactgggaa aacctggcgg ttaccaact taatcgccct    180
gcagcacatc cccctttcgc cagctggcgt aatagcgaag aggcccgcac cgatcgccct    240
tcccaacagt tgcgcagcct gaatggcgaa tggcgctttg cctggtttcc ggcaccagaa    300
gcggtgccgg aaagctggct ggagtgcgat ctctctgagg ccgatactgt cgtcgtcccc    360
tcaaactggc agatgcacgg ttacgatgcg cccatctaca ccaacgtaac ctatccatt    420

```

ES 2 422 303 T3

acggtcaatc cgccgtttgt tcccacggag aatccgacgg gttgttactc gctcacattt 480  
 aatgttgatg aaagctggct acaggaaggc cagacgcgaa ttatTTTTga tggcgtaac 540  
 tcggcgtttc atctgtggtg caacgggcgc tgggtcggtt acggccagga cagtcgtttg 600  
 ccgtctgaat ttgacctgag cgcattttta cgcgccggag aaaaccgcct cgcggtgatg 660  
 gtgctgcggt ggagtgacgg cagttatctg gaagatcagg atatgtggcg gatgagcggc 720  
 atttccgctg acgtctcgtt gctgcataaa ccgactacac aatcagcga tttccatggt 780  
 gccactcgcct ttaatgatga tttcagccgc gctgtactgg aggctgaagt tcagatgtgc 840  
 ggcgagttgc gtgactacct acgggtaaca gtttctttat ggcaggggtga aacgcaggtc 900  
 gccagcggca ccgcgccctt cggcggtgaa attatcgatg agcgtggtgg ttatgccgat 960  
 cgcgtcacac tacgtctgaa cgtcgaaaac ccgaaactgt ggagcgcga aatcccgat 1020  
 ctctatcgtg cgggtggtga actgcacacc gccacggca cgctgattga agcagaagcc 1080  
 tgcgatgtcg gtttccgcga ggtgcggatt gaaaatggtc tgctgctgct gaacggcaag 1140  
 ccgttgctga ttcgaggcgt taaccgtcac gagcatcacc ctctgcatgg tcaggtcatg 1200  
 gatgagcaga cgatggtgca ggatatactg ctgatgaagc agaacaactt taacgcctg 1260  
 cgctgttcgc attatccgaa ccatacgcctg tggtagacgc tgtgcgaccg ctacggcctg 1320  
 tatgtggtgg atgaagccaa tattgaaacc cacggcatgg tgccaatgaa tcgtctgacc 1380  
 gatgatccgc gctggetacc ggcgatgagc gaacgcgtaa cgcgaatggg gcagcgcgat 1440  
 cgtaatcacc cgagtgtgat catctggctg ctggggaatg aatcaggcca cggcgtaat 1500  
 cacgacgcgc tgtatcgcctg gatcaaatct gtcgatcctt ccgccccggt gcagtatgaa 1560  
 ggcggcggag ccgacaccac ggccaccgat attatttgc c gatgtacgc gcgcgtggat 1620  
 gaagaccagc ccttccccgc tgtgccgaaa tggtagcatca aaaaatggct ttcgctacct 1680  
 ggagagacgc gcccgcctgat ccttgcgaa tacgcccacg cgatgggtaa cagtcttggc 1740  
 ggtttcgcta aatactggca ggcgtttcgt cagtatcccc gtttacaggg cggcttcgtc 1800  
 tgggactggg tggatcagtc gctgattaaa tatgatgaaa acggcaaccc gtggtcggct 1860  
 tacggcggtg attttggcga tacgccgaac gatcggcagt tctgtatgaa cggctcggtc 1920  
 tttgccgacc gcacgccgca tccagcgcctg acggaagcaa aacaccagca gcagtttttc 1980  
 cagttccggt tatccgggca aaccatcgaa gtgaccagcg aatacctggt ccgtcatagc 2040  
 gataacgagc tcctgcactg gatggtggcg ctggatggta agccgctggc aagcggtgaa 2100  
 gtgcctctgg atgtcgtcc acaaggtaaa cagttgattg aactgcctga actaccgcag 2160  
 ccggagagcg ccgggcaact ctggctcaca gtacgcgtag tgcaaccgaa cgcgaccgca 2220  
 tggtcagaag ccgggcacat cagcgcctgg cagcagtggc gtctggcgga aaacctcagt 2280

gtgacgctcc ccgcccggtc ccacgccatc ccgcatctga ccaccagcga aatggathtt 2340  
 tgcacgagc tgggtaataa gcgttgcaa tttaaccgcc agtcaggctt tctttcacag 2400  
 atgtggattg gcgataaaaa acaactgctg acgccgctgc gcgatcagtt caccctgca 2460  
 ccgctggata acgacattgg cgtaagtga gcgaccgcga ttgaccctaa cgcctgggtc 2520  
 gaacgctgga aggcggcggg ccattaccag gccgaagcag cgttggtgca gtgcacggca 2580  
 gatacacttg ctgatgcggt gctgattacg accgctcacg cgtggcagca tcaggggaaa 2640  
 accttattta tcagccggaa aacctaccgg attgatggta gtggtaaat ggcgattacc 2700  
 gttgatgttg aagtggcggag cgatacaccg catccggcgc ggattggcct gaactgccag 2760  
 ctggcgcagg tagcagagcg ggtaaactgg ctccgattag ggccgcaaga aaactatccc 2820  
 gaccgcctta ctgccgcctg ttttgaccgc tgggatctgc cattgtcaga catgtatacc 2880  
 ccgtacgtct tcccagcga aaacggctc cgtgcggga cgcgcgaatt gaattatggc 2940  
 ccacaccagt ggcgcggcga cttccagttc aacatcagcc gctacagtca acagcaactg 3000  
 atggaaacca gccatcgcca tctgctgcac gcggaagaag gcacatggct gaatatcgac 3060  
 ggtttccata tggggattgg tggcgacgac tcttgagcc cgtcagtatc ggcggaatta 3120  
 cagctgagcg ccggtcgcta ccattaccag ttggtct 3157

<210> 30

<211> 4375

<212> ADN

5 <213> *Choristoneura fumiferana*

<400> 30

tgtaattttg atgggcccgg tgatgcaccg tgtgccatat tgccatccag tcgaatagaa 60  
 aaaaaaaaaa aaaaaaaaaat atcagttggt ttgtccctcg ctccgctttcg agtgtattcg 120  
 gaatattaga cgtcataatt cagcagtgct ttttaaattt atatagcgat tagcggggcc 180  
 gtttgttggg cgtgcgcttg cgttttagtg agtgcaggga tagtgaggcg agtatggtag 240  
 ttcgtgggtca tgtcaagtgt ggcgaagaaa gacaagccga cgatgtcggg gacggcgtg 300  
 atcaactggg cgcggccggc gccgccaggc ccgccgcagc cgcagtcagc gtcgcctgcg 360  
 ccggcagcca tgetgcagca gctcccagc cagtcaatgc agtcgttaa ccacatccca 420  
 actgtcgatt gctcgtcga tatgcagtg ctttaattag aacctggatt catgtcgcct 480  
 atgtcacctc ctgagatgaa accagacacc gccatgcttg atgggctacg agacgacgcc 540  
 acttcgccgc ctaacttcaa gaactaccgc cctaatacacc ccctgagtggt ctccaaacac 600  
 ctatgctcta tatgcggcga cagggcgtct gggaagcact atgggggtgta cagttgcgaa 660  
 ggatgcaagg gtttcttcaa gcggaccgct cggaaggacc tgtcgtacgc ttgccgggag 720

ES 2 422 303 T3

gagcggaaact gcacatagaa caagcgacaa aggaaccgat gccagtactg ccgctatcaa 780  
aagtgtttgg cttgcggtat gaagcgagag gcggtgcaag aggagcgcca gaggaatgct 840  
cgcgcgcggg aggatgcgca cccgagtagc tcggtgcagg taagcgatga gctgtcaatc 900  
gagcgcctaa cggagatgga gtctttggtg gcagatccca gcgaggagt ccagttcctc 960  
cgctgtgggc ctgacagcaa cgtgcctcca cgttaccgag cgcccgtctc ctccctctgc 1020  
caaataggca acaagcaaat agcggcggtg gtggtatggg cgcgcgacat ccctcatttc 1080  
gggcagctgg agctggacga tcaagtggta ctcatcaagg cctcctggaa tgagctgcta 1140  
ctcttcgcca tcgcctggcg ctctatggag tatttggaag atgagagga gaacggggac 1200  
ggaacgcgga gcaccactca gccacaactg atgtgtctca tgctggcat gacgttgac 1260  
cgcaactcgg cgcagcagge gggcgtgggc gccatcttcg accgcgtgct gtccgagctc 1320  
agtctgaaga tgcgcacctt gcgcatggac caggccgagt acgtcgcgct caaagccatc 1380  
gtgctgctca accctgatgt gaaaggactg aagaatcggc aagaagttga cgttttgcca 1440  
gaaaaaatgt tctcttgccg ggacgactac tgccggcggt cgcggaagca cgaggaaggc 1500  
cggtttgctt ccttgctgct gggctgcca gctctccgct ccatctcgct caagagcttc 1560  
gaacacctct acttcttcca cctcgtggcc gaaggtcca tcagcggata catacgagag 1620  
gcgctccgaa accacgcgcc tccgatcgac gtcaatgcca tgatgtaaag tgcgatacac 1680  
gccctgccga tgtgagaaga actatggcta atagaagcga aactgaatac atctagggtg 1740  
ggacttaact tgggactatc attaaagtat cacgcaaatt atgcgtagtc agaaagtccg 1800  
gtcgatcaaa cttttttata aacgaattga gtttctaacg actgcaacac agcggagttt 1860  
tgctctgat agtttttatt ctaatggtta agatgcttta cacgggcatt attgacattc 1920  
aagtgtaagt ggaagttgac aaccttgaca tttatatcac gtttgtaatt ggttaaataa 1980  
attaattaat cacaagtaag actaacatca acgtcacgat actaacgcca tttagtgata 2040  
ttttcatgt caagaaactc attgttttga taaaatattt ttctaattac tccagtgaac 2100  
tcatccaaat gtgaccaggt tcccgcaga gttgcccggt taaaatcacc tttagggaca 2160  
tatccccgc tatctcatga aattccaagg atcagtaggg gccaatccc ccgatgtgtt 2220  
gggaggcaga attttcgata atctacgact attgttagcc tacgaattag ttgaattttt 2280  
tgaaattatt tttattaagt cgccactttc caaacacacc agcaggggat atgtgcaatt 2340  
ttgtaacgat aactctattc atttctgata tttatcgaaa ttttatctta cataacatgc 2400  
tggctggctc aggtgtttgg tagttacata tgtatctacg gtttgtttta aattatagct 2460  
ttttattgt aatctgtata aaattgagtt atcttacttc aactacgat cgagtaaacc 2520

ES 2 422 303 T3

catcgtcagc tacgaaaaac taatcgtata aggcgtaaga gtaaataact aattgacaac 2580  
 cagcaacgag gaccacctca gtcctcgtgc ttacattgtg ccgtagctta atatgatgga 2640  
 agctgtcgtc gttacgacat tagataaagt gcatgaatac caaaaatgta ccatcccgtta 2700  
 ctgatctctc atgctctcgc tgcgtgggac ccggtgctgag tgcgtgaagg actgactaat 2760  
 attttagact aggcgtctat gcttcagtaa ttccttatac atattataag tcatccaaat 2820  
 aacgagtaag gcggcatggt gagatcagca ttcgagagt caaagagccc ctaacgtgac 2880  
 tgagaagtag agacaataca ctgattttct gagatgaacg caaccgagat tgacactaaa 2940  
 aatctattta tggatttcaa aatggcgatg cttgattgtc tgcggcgtgg atagactgaa 3000  
 atgggtttgc ttaacactgg atattgtttt tattagttaa tagtcttaca ttgcaagttg 3060  
 gtaattcggg gctaataatcg accggtttgt taactatcta acggttccca gtgtcaggca 3120  
 cacatcttcc ccaagcagac aacgcaagag tgtacaaaat gtacatgtta caaaaataagg 3180  
 aacattcgtc ggataagtgt aacagttgat aggtaaagaa aatggggccg cctctttatt 3240  
 attacgtagc cgtaaaatta ttaacgtatt tagtttagat gttcagctaa ttaggataat 3300  
 tctatttgtc gagtacctag atgtccatag tgaattaata taataattag actgttacgc 3360  
 gtaggtaatt ataaagttaa ccaaatctct cttcaaagca aaaactttgt acacttccgt 3420  
 actgagacgt cgtagcttat tctgattcac gaaatatttg gatcacattg ttacaaggcg 3480  
 accgtcacgt agtatatgat tatttacaaa tgacacgtat gtatcaatgc tataagtgtt 3540  
 ttcggtacat atgtcgggtc tttaacgtgc atttcgtatg gcagattaaa aatagcaaga 3600  
 aatcttgaaa ttgttttaga aaatatttga tttccttatt gaaagttatt tttaaatgta 3660  
 aatatttcgt aatcataata attatgtatt gtgtagttat ttcacctta cggttgggat 3720  
 attatttaat ggtggcctac gaaagtgatt ataaccatcc gcgtcctcaa aaaggccagt 3780  
 ttatttttgt acctacata tactaattac gtaagtaata tcaggcgaat ggttgactaa 3840  
 caactaacca gtattaaaa ttaaaagact tcgtcctaataaaaatgtaat atctatgtat 3900  
 aaaaatgaaa aatctggcgt ataataggta aaattaaact agattgttaa tgaatgtgat 3960  
 gtctcataaa cgtttagttt ttaatgagaa acatgttttag tcgcctacta taagacgaga 4020  
 cggcaagctc accgagttaa ctcgtaaaca ggaatgttga aaaagatgac acaatttata 4080  
 tttggtattg aaattatgac taaccatgcg ctctatcgtt tgttatggat gcatagtatt 4140  
 gctggtgaaa ataatggaat taggtaatta ctgcattaat gttgaaaact tgatattatt 4200  
 ctatgggtgg gtatgaatc tatgttgga gttgtgcagc ggttgtaaag atgatttata 4260  
 atgatgttca ctaaatatct gactaaatgt aagttatttt tttttgtata gacatagctt 4320  
 taagatgaag gtgattaaac tttatcctta tcacaataaa aaaaaaaaaa aaaaa 4375

<210> 31

<211> 1878

<212> ADN

5 <213> *Drosophila melanogaster*

<400> 31

ES 2 422 303 T3

ggacctgctc cacgggtgca agaggagctg tgcctggttt gcggcgacag ggcctccggc 60  
 taccactaca acgccctcac ctgtgagggc tgcaaggggt tctttcgacg cagcgttacg 120  
 aagagcgccg tctactgctg caagttcggg cgcgcctgcg aaatggacat gtacatgagg 180  
 cgaaagtgtc aggagtgcgg cctgaaaaag tgcctggccg tgggtatgcg gccggaatgc 240  
 gtcgtcccgg agaaccaatg tgcgatgaag cggcgcgaaa agaaggcca gaaggagaag 300  
 gacaaaatga ccaactcgcc gagctctcag catggcggca atggcagctt ggcctctggt 360  
 ggcggccaag actttgtaa gaaggagatt cttgacctta tgacatgcga gccgccccag 420  
 catgccacta ttccgctact acctgatgaa atattggcca agtgtcaagc gcgcaatata 480  
 ccttccttaa cgtacaatca gttggcggtt atatacaagt taatttggtta ccaggatggc 540  
 tatgagcagc catctgaaga ggatctcagg cgtataatga gtcaaccgga tgagaacgag 600  
 agccaaacgg acgtcagctt tcggcatata accgagataa ccatactcac ggtccagttg 660  
 attgttgagt ttgctaaagg tctaccagcg tttacaaga taccacagga ggaccagatc 720  
 acgttactaa aggctgctc gtcggaggtg atgatgctgc gtatggcacg acgctatgac 780  
 cacagctcgg actcaatatt ctctcgcaat aatagatcat atacgcggga ttcttataaaa 840  
 atggccggaa tggctgataa cattgaagac ctgtctgatt tctgccgcca aatggtctcg 900  
 atgaagggtg acaacgtcga atacgcgctt ctactgcca ttgtgatctt ctcgaccgg 960  
 ccgggcctgg agaaggccca actagtcgaa gcgatccaga gctactacat cgacacgcta 1020  
 cgcatttata tactcaaccg ccaactcgcc gactcaatga gcctcgtctt ctacgcaaag 1080  
 ctgctctcga tctcaccga gctgctacg ctgggcaacc agaacgcca gatgtgtttc 1140  
 tcaactaaagc tcaaaaaccg caaactgccc aagtctctcg aggagatctg ggacgttcat 1200  
 gccatcccgc catcgggtcca gtcgcacctt cagattaccg aggaggagaa cgagcgtctc 1260  
 gagcgggctg agcgtatgcg ggcacggtt gggggcgcca ttaccgcccg cattgattgc 1320  
 gactctgect ccaactcggc gggggcagcc gcggcccagc atcagcctca gcctcagccc 1380  
 cagccccaac cctcctcctt gaccagaac gattcccagc accagacaca gccgcagcta 1440  
 caacctcagc taccacctca gctgcaaggc caactgcaac cccagctcca accacagctt 1500  
 cagacgcaac tccagccaca gattcaacca cagccacagc tccttcccgt ctccgctccc 1560  
 gtgcccgcct ccgtaaccgc acctggttcc ttgtccgagg tcagtacgag cagcgaatac 1620  
  
 atgggcggaa gtgcggccat aggaccatc acgccggcaa ccaccagcag tatcacggct 1680  
 gccgttaccg ctactccac cacatcagcg gtaccgatgg gcaacggagt tggagtccgt 1740  
 gttggggtgg gcggcaacgt cagcatgtat gcgaacgccc agacggcgat ggccttgatg 1800  
 ggtgtagccc tgcattcgca ccaagagcag cttatcgggg gagtggcggg taagtcggag 1860  
 cactcgacga ctgcatag 1878

<210> 32  
 <211> 878  
 5 <212> PRT  
 <213> *Drosophila melanogaster*  
 <400> 32

ES 2 422 303 T3

Met Lys Arg Arg Trp Ser Asn Asn Gly Gly Phe Met Arg Leu Pro Glu  
 1 5 10 15

Glu Ser Ser Ser Glu Val Thr Ser Ser Ser Asn Gly Leu Val Leu Pro  
 20 25 30

Ser Gly Val Asn Met Ser Pro Ser Ser Leu Asp Ser His Asp Tyr Cys  
 35 40 45

Asp Gln Asp Leu Trp Leu Cys Gly Asn Glu Ser Gly Ser Phe Gly Gly  
 50 55 60

Ser Asn Gly His Gly Leu Ser Gln Gln Gln Gln Ser Val Ile Thr Leu  
 65 70 75 80

Ala Met His Gly Cys Ser Ser Thr Leu Pro Ala Gln Thr Thr Ile Ile  
 85 90 95

Pro Ile Asn Gly Asn Ala Asn Gly Asn Gly Gly Ser Thr Asn Gly Gln  
 100 105 110

Tyr Val Pro Gly Ala Thr Asn Leu Gly Ala Leu Ala Asn Gly Met Leu  
 115 120 125

Asn Gly Gly Phe Asn Gly Met Gln Gln Gln Ile Gln Asn Gly His Gly  
 130 135 140

Leu Ile Asn Ser Thr Thr Pro Ser Thr Pro Thr Thr Pro Leu His Leu  
 145 150 155 160

Gln Gln Asn Leu Gly Gly Ala Gly Gly Gly Gly Ile Gly Gly Met Gly  
 165 170 175



ES 2 422 303 T3

Ile Leu His His Ala Asn Gly Thr Pro Asn Gly Leu Ile Gly Val Val  
 180 185 190

Gly Gly Gly Gly Gly Val Gly Leu Gly Val Gly Gly Gly Gly Val Gly  
 195 200 205

Gly Leu Gly Met Gln His Thr Pro Arg Ser Asp Ser Val Asn Ser Ile  
 210 215 220

Ser Ser Gly Arg Asp Asp Leu Ser Pro Ser Ser Ser Leu Asn Gly Tyr  
 225 230 235 240

Ser Ala Asn Glu Ser Cys Asp Ala Lys Lys Ser Lys Lys Gly Pro Ala  
 245 250 255

Pro Arg Val Gln Glu Glu Leu Cys Leu Val Cys Gly Asp Arg Ala Ser  
 260 265 270

Gly Tyr His Tyr Asn Ala Leu Thr Cys Glu Gly Cys Lys Gly Phe Phe  
 275 280 285

Arg Arg Ser Val Thr Lys Ser Ala Val Tyr Cys Cys Lys Phe Gly Arg  
 290 295 300

Ala Cys Glu Met Asp Met Tyr Met Arg Arg Lys Cys Gln Glu Cys Arg  
 305 310 315 320

Leu Lys Lys Cys Leu Ala Val Gly Met Arg Pro Glu Cys Val Val Pro  
 325 330 335

Glu Asn Gln Cys Ala Met Lys Arg Arg Glu Lys Lys Ala Gln Lys Glu  
 340 345 350

Lys Asp Lys Met Thr Thr Ser Pro Ser Ser Gln His Gly Gly Asn Gly  
 355 360 365

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

Asp Leu Met Thr Cys Glu Pro Pro Gln His Ala Thr Ile Pro Leu Leu  
 385 390 395 400

Pro Asp Glu Ile Leu Ala Lys Cys Gln Ala Arg Asn Ile Pro Ser Leu  
 405 410 415

Thr Tyr Asn Gln Leu Ala Val Ile Tyr Lys Leu Ile Trp Tyr Gln Asp  
 420 425 430

Gly Tyr Glu Gln Pro Ser Glu Glu Asp Leu Arg Arg Ile Met Ser Gln  
 435 440 445

Pro Asp Glu Asn Glu Ser Gln Thr Asp Val Ser Phe Arg His Ile Thr  
 450 455 460

Glu Ile Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala Lys Gly  
 465 470 475 480

Leu Pro Ala Phe Thr Lys Ile Pro Gln Glu Asp Gln Ile Thr Leu Leu  
 485 490 495

Lys Ala Cys Ser Ser Glu Val Met Met Leu Arg Met Ala Arg Arg Tyr  
 500 505 510

Asp His Ser Ser Asp Ser Ile Phe Phe Ala Asn Asn Arg Ser Tyr Thr  
 515 520 525

Arg Asp Ser Tyr Lys Met Ala Gly Met Ala Asp Asn Ile Glu Asp Leu  
 530 535 540

Leu His Phe Cys Arg Gln Met Phe Ser Met Lys Val Asp Asn Val Glu  
 545 550 555 560

Tyr Ala Leu Leu Thr Ala Ile Val Ile Phe Ser Asp Arg Pro Gly Leu  
 565 570 575

Glu Lys Ala Gln Leu Val Glu Ala Ile Gln Ser Tyr Tyr Ile Asp Thr  
 580 585 590

Leu Arg Ile Tyr Ile Leu Asn Arg His Cys Gly Asp Ser Met Ser Leu  
 595 600 605

Val Phe Tyr Ala Lys Leu Leu Ser Ile Leu Thr Glu Leu Arg Thr Leu  
 610 615 620

Gly Asn Gln Asn Ala Glu Met Cys Phe Ser Leu Lys Leu Lys Asn Arg  
 625 630 635 640

Lys Leu Pro Lys Phe Leu Glu Glu Ile Trp Asp Val His Ala Ile Pro  
 645 650 655

ES 2 422 303 T3

Pro Ser Val Gln Ser His Leu Gln Ile Thr Gln Glu Glu Asn Glu Arg  
660 665 670

Leu Glu Arg Ala Glu Arg Met Arg Ala Ser Val Gly Gly Ala Ile Thr  
675 680 685

Ala Gly Ile Asp Cys Asp Ser Ala Ser Thr Ser Ala Ala Ala Ala Ala  
690 695 700

Ala Gln His Gln Pro Gln Pro Gln Pro Gln Pro Gln Pro Ser Ser Leu  
705 710 715 720

Thr Gln Asn Asp Ser Gln His Gln Thr Gln Pro Gln Leu Gln Pro Gln  
725 730 735

Leu Pro Pro Gln Leu Gln Gly Gln Leu Gln Pro Gln Leu Gln Pro Gln  
740 745 750

Leu Gln Thr Gln Leu Gln Pro Gln Ile Gln Pro Gln Pro Gln Leu Leu  
755 760 765

Pro Val Ser Ala Pro Val Pro Ala Ser Val Thr Ala Pro Gly Ser Leu  
770 775 780

Ser Ala Val Ser Thr Ser Ser Glu Tyr Met Gly Gly Ser Ala Ala Ile  
785 790 795 800

Gly Pro Ile Thr Pro Ala Thr Thr Ser Ser Ile Thr Ala Ala Val Thr  
805 810 815

Ala Ser Ser Thr Thr Ser Ala Val Pro Met Gly Asn Gly Val Gly Val  
820 825 830

Gly Val Gly Val Gly Gly Asn Val Ser Met Tyr Ala Asn Ala Gln Thr  
835 840 845

Ala Met Ala Leu Met Gly Val Ala Leu His Ser His Gln Glu Gln Leu  
850 855 860

Ile Gly Gly Val Ala Val Lys Ser Glu His Ser Thr Thr Ala  
865 870 875

<210> 33

<211> 1282

<212> ADN

5 <213> *Amblyomma americanum*

<400> 33

ES 2 422 303 T3

cggcggaaagt gccaggagtg cggctcaag aagtgcctgg ccgtcgggat gcggccggag 60  
 tgcgtcgtgc cggagaacca gtgcgccatc aagcggaaagg agaagaaagc ccagaaggag 120  
 aaggacaagg tgcaaacgaa cgccaccgtc agtacaacga acagcaccta ccggtcggag 180  
 atactgccga tcctgatgaa atgtgatcca ccgccgcacc aagcgatacc tctactaccg 240  
 gaaaagctcc tgcaggagaa taggctaaga aacatacctc tactgacggc gaaccaaatg 300  
 gccgtcattt aaaaactcat ctggtaccag gacgggtacg agcaaccctc ggaggaagat 360  
 ctcaaacgga taatgatcgg ttcacccaac gaggaggaag atcaacatga cgtgcacttc 420  
 cggcacataa cggaaatcac aatcctaaca gtacaactaa tcgtggagtt cgccaagggg 480  
 ctgccagcat ttaccaagat tccacaggag gaccagatca cgctgctgaa ggctgctca 540  
 agcgagggta tgatgttgcg aatggcccgc cgctacgacg ctgccaccga ttcgatcctg 600  
 ttcgcgaaca accggtccta cacgagggac tcctaccgga tggccggcat ggcggacacg 660  
 atagaggacc tgctgcactt ctgccggcag atgttctccc tcacggtaga caacgtcgag 720  
 tacgcactcc tcacggcgat agtcatcttc tcggatcggc ccggactgga gcaagccgaa 780  
 ctggtcagac acatccagag ctactacatc gacacgctgc ggatctacat cctgaatagg 840  
 cacgcgggcg atccgaagtg cagtgtgata ttocccaaac tgctgtcgat cctgacggag 900  
 ctccgaacgc tgggcaacca gaactcggag atgtgcttct cgctcaagct gaagaaccgc 960  
 aaactgccac ggttcctgga ggagatctgg gacgtccagg acataccgcc ctcgatgcag 1020  
 gccagatgc acagccatgg caccagtcct tcgtcctcat cgctcctccag tagtagtagt 1080  
 agtagtaacg gtagtagtaa cggtaacagt agtagtaata gtaatagttc acagcacggg 1140  
 ccacatccgc atccgcacgg gcagcaatta acgccaatc agcagcagca tcagcagcag 1200  
 cacagtcagt tacagcaagt tcacgccaac ggcagcggaa gtggtggcgg cagtaacaat 1260  
 aatagcagta gtgggggctg ag 1282

<210> 34

<211> 560

<212> PRT

5 <213> *Amblyomma americanum*

<400> 34

Met Leu Ser Glu Leu Gln Leu Gln Pro Leu Gly Arg Arg Ser Ala Ala  
 1 5 10 .. 15

Ala Ala Pro Ser Ser Asp Glu Val Ala Ser Met Lys Pro Met Leu Leu  
 20 25 30

ES 2 422 303 T3

Gln Ala Ala Gln Gly Gly Ala Gly Gly Leu Ala Ala Gly Ser Pro Pro  
 35 40 45

Ala Leu Ser Pro Asn Leu Pro Ser Val Val Lys Val Glu Pro Arg Leu  
 50 55 60

Pro Ser Pro Cys Val Gly Gly Ala Ala Ser Gly Asp Gly Gly Pro Val  
 65 70 75 80

Pro Pro Lys Arg Val Arg Gln Asp Asp Ala Gly Ala Trp Ile Ser Ser  
 85 90 95

Pro Ser Ser Gln Met Ser Val Gly Ser Leu Ser Pro Pro Pro Pro Leu  
 100 105 110

Leu Asn Gly Val Ala Asn Ser Ser Gly Leu Ser Pro Val Ser Asn Cys  
 115 120 125

Ser Ser Tyr Asp Thr Tyr Ser Pro Arg Gly Pro Cys Lys Glu Glu Met  
 130 135 140

Ser Pro Ser Ser Gly Gly Gly Gly Leu Asn Gly Tyr Phe Val Asp Ser  
 145 150 155 160

Phe Gly Asp Pro Lys Lys Lys Lys Gly Pro Ala Pro Arg Gln Gln Glu  
 165 170 175

Glu Leu Cys Leu Val Cys Gly Asp Arg Ala Ser Gly Tyr His Tyr Asn  
 180 185 190

Ala Leu Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Ile Thr  
 195 200 205

Lys Asn Ala Val Tyr Gln Cys Lys Tyr Gly Asn Asn Cys Asp Ile Asp  
 210 215 220

Met Tyr Met Arg Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu  
 225 230 235 240

Ser Val Gly Met Arg Pro Glu Cys Val Val Pro Glu Tyr Gln Cys Ala  
 245 250 255

Ile Lys Arg Glu Ser Lys Lys His Gln Lys Asp Arg Pro Asn Ser Thr  
 260 265 270

Thr Arg Glu Ser Pro Ser Ala Leu Met Ala Pro Ser Ser Val Gly Gly

ES 2 422 303 T3

275    280    285  
 Val Ser Pro Thr Ser Gln Pro Met Gly Gly Gly Gly Ser Ser Leu Gly  
 290    295    300  
  
 Ser Ser Asn His Glu Glu Asp Lys Lys Pro Val Val Leu Ser Pro Gly  
 305    310    315    320  
  
 Val Lys Pro Leu Ser Ser Ser Gln Glu Asp Leu Ile Asn Lys Leu Val  
 325    330    335  
  
 Tyr Tyr Gln Gln Glu Phe Glu Ser Pro Ser Glu Glu Asp Met Lys Lys  
 340    345    350  
  
 Thr Thr Pro Phe Pro Leu Gly Asp Ser Glu Glu Asp Asn Gln Arg Arg  
 355    360    365  
  
 Phe Gln His Ile Thr Glu Ile Thr Ile Leu Thr Val Gln Leu Ile Val  
 370    375    380  
  
 Glu Phe Ser Lys Arg Val Pro Gly Phe Asp Thr Leu Ala Arg Glu Asp  
 385    390    395    400  
  
 Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met Leu Arg  
 405    410    415  
  
 Gly Ala Arg Lys Tyr Asp Val Lys Thr Asp Ser Ile Val Phe Ala Asn  
 420    425    430  
  
 Asn Gln Pro Tyr Thr Arg Asp Asn Tyr Arg Ser Ala Ser Val Gly Asp  
 435    440    445  
  
 Ser Ala Asp Ala Leu Phe Arg Phe Cys Arg Lys Met Cys Gln Leu Arg  
 450    455    460  
  
 Val Asp Asn Ala Glu Tyr Ala Leu Leu Thr Ala Ile Val Ile Phe Ser  
 465    470    475    480  
  
 Glu Arg Pro Ser Leu Val Asp Pro His Lys Val Glu Arg Ile Gln Glu  
 485    490    495  
  
 Tyr Tyr Ile Glu Thr Leu Arg Met Tyr Ser Glu Asn His Arg Pro Pro  
 500    505    510  
  
 Gly Lys Asn Tyr Phe Ala Arg Leu Leu Ser Ile Leu Thr Glu Leu Arg  
 515    520    525  
  
 Thr Leu Gly Asn Met Asn Ala Glu Met Cys Phe Ser Leu Lys Val Gln  
 530    535    540  
  
 Asn Lys Lys Leu Pro Pro Phe Leu Ala Glu Ile Trp Asp Ile Gln Glu  
 545    550    555    560

<210> 35  
 <211> 40  
 <212> ADN

5

<213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 35  
 5 gtaccaggac gggtagcgc agccttctga tgaagattg 40  
 <210> 36  
 <211> 40  
 <212> ADN  
 <213> Secuencia artificial  
 10 <220>  
 <223> Cebador de PCR  
 <400> 36  
 caaatctca tcagaaggct ggcgtaccc gtctgttac 40  
 <210> 37  
 15 <211> 37  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 20 <400> 37  
 ccaggacggg tacgaggcgc ctctgatga agattg 37  
 <210> 38  
 <211> 37  
 <212> ADN  
 25 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 38  
 caaatctca tcagaaggcg cctcgtaccc gtctgg 37  
 30 <210> 39  
 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 35 <223> Cebador de PCR  
 <400> 39  
 gagtctgaca ctccgcccg ccagatcaca g 31  
 <210> 40  
 <211> 31  
 40 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 40  
 45 ctgtgatctg gcggcgagg ggtcagact c 31  
 <210> 41  
 <211> 30  
 <212> ADN  
 <213> Secuencia artificial  
 50 <220>  
 <223> Cebador de PCR  
 <400> 41  
 cactccctc cgccaggcca cagagatgac 30  
 <210> 42  
 55 <211> 30  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 60 <400> 42  
 gtcatctctg tggcctggcg gaaggagtg 30  
 <210> 43  
 <211> 30  
 <212> ADN  
 65 <213> Secuencia artificial  
 <220>

<223> Cebador de PCR  
 <400> 43  
 cactccctccgccagatcg cagagatgac 30  
 <210> 44  
 5 <211> 30  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 10 <400> 44  
 gtcatctctg cgatctggcg gaaggagtg 30  
 <210> 45  
 <211> 34  
 <212> ADN  
 15 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 45  
 cgccagatca cagagatggc tatcctcagc gtcc 34  
 20 <210> 46  
 <211> 34  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 25 <223> Cebador de PCR  
 <400> 46  
 ggaccgtgag gatagccatc tctgtgatct ggcg 34  
 <210> 47  
 <211> 33  
 30 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 47  
 35 gagatgacta tcctcgcggt ccaactatc gtc 33  
 <210> 48  
 <211> 33  
 <212> ADN  
 <213> Secuencia artificial  
 40 <220>  
 <223> Cebador de PCR  
 <400> 48  
 cacgataagt tggaccgca gtagatcatc ctc 33  
 <210> 49  
 45 <211> 32  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 50 <400> 49  
 gatgactatc ctcacggccc aactatcgt gg 32  
 <210> 50  
 <211> 32  
 <212> ADN  
 55 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 50  
 ccacgataag ttggccgtg aggatagtc tc 32  
 60 <210> 51  
 <211> 35  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 65 <223> Cebador de PCR  
 <400> 51



ctatcctcac ggtccaagct atcgtggagt tcgcg 35  
 <210> 52  
 <211> 35  
 <212> ADN  
 5 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 52  
 cgcgactcc acgatagctt ggaccgtgag gatag 35  
 10 <210> 53  
 <211> 35  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 15 <223> Cebador de PCR  
 <400> 53  
 ctatcctcac ggtccaactt gccgtggagt tcgcg 35  
 <210> 54  
 <211> 35  
 20 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 54  
 25 cgcgactcc acggcaagtt ggaccgtgag gatag 35  
 <210> 55  
 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 30 <220>  
 <223> Cebador de PCR  
 <400> 55  
 gctcaagtga ggtagcgatg ctccgagtcg c 31  
 <210> 56  
 35 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 56  
 40 gcgactcgga gcatcgctac ctacttgag c 31  
 <210> 57  
 <211> 31  
 <212> ADN  
 45 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 57  
 gctcaagtga ggtaatggcg ctccgagtcg c 31  
 50 <210> 58  
 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 55 <223> Cebador de PCR  
 <400> 58  
 gcgactcgga gcgcattac ctacttgag c 31  
 <210> 59  
 <211> 30  
 60 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 59  
 65 gtaatgatgc tccgagccgc gcgacgatac 30  
 <210> 60

<211> 30  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Cebador de PCR  
 <400> 60  
 gtatcgtcgc gcggtcggca gcatcattac 30  
 <210> 61  
 <211> 36  
 10 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 61  
 15 gtgaggtaat gatgctcgca gtcgctcgac gatacg 36  
 <210> 62  
 <211> 36  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 20 <223> Cebador de PCR  
 <400> 62  
 cgtatcgtcg cgcgactcgg agcatcatta cctcac 36  
 <210> 63  
 25 <211> 32  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 63  
 30 cagacagtgt tctggccgca aacaaccaag cg 32  
 <210> 64  
 <211> 32  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 64  
 cgcttggttg ttcgcgcca gaactctg tg 32  
 40 <210> 65  
 <211> 33  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 45 <223> Cebador de PCR  
 <400> 65  
 tcagacagtg ttctggccgc gaacaacaa gcg 33  
 <210> 66  
 <211> 33  
 50 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 66  
 55 cgcttggttg ttcgcgcca gaactctg tga 33  
 <210> 67  
 <211> 32  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 60 <223> Cebador de PCR  
 <400> 67  
 cagacagtgt tctgtcccg aacaaccaag cg 32  
 <210> 68  
 65 <211> 32  
 <212> ADN

<213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 68  
 5 cgcttggtg ttcggaaca gaacactgc tg 32  
 <210> 69  
 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 10 <220>  
 <223> Cebador de PCR  
 <400> 69  
 cactcgcgac aacgcccga aggctggcat g 31  
 <210> 70  
 15 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 20 <400> 70  
 catgccagcc ttgcggcgt tgtcgcgagt g 31  
 <210> 71  
 <211> 34  
 <212> ADN  
 25 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 71  
 cgacaactac cgcaagttg gcatggccta cgtc 34  
 30 <210> 72  
 <211> 34  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 35 <223> Cebador de PCR  
 <400> 72  
 gacgtaggcc atgccaaact tgcgtagtt gtcg 34  
 <210> 73  
 <211> 31  
 40 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 73  
 45 ctaccgcaag gctggcggg cctacgtcat c 31  
 <210> 74  
 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 50 <220>  
 <223> Cebador de PCR  
 <400> 74  
 gatgacgtag gccgcgccag cctgcggtta g 31  
 <210> 75  
 55 <211> 32  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 60 <400> 75  
 gctcaagaac agaaaggcgc cgccttctcg cg 32  
 <210> 76  
 <211> 32  
 <212> ADN  
 65 <213> Secuencia artificial  
 <220>

<223> Cebador de PCR  
 <400> 76  
 cgaggaaagg cggcgcttt ctgttctga gc 32  
 <210> 77  
 5 <211> 36  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 10 <400> 77  
 ctccaacatg tgcatctccg ccaagctca gaacag 36  
 <210> 78  
 <211> 36  
 <212> ADN  
 15 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 78  
 ctgttctga gcttggcgga gatgcacatg ttggag 36  
 20 <210> 79  
 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 25 <223> Cebador de PCR  
 <400> 79  
 gaaagctgcc gccttcgcc gaggagatct g 31  
 <210> 80  
 <211> 31  
 30 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 80  
 35 cagatctct cggcgaaagg cggcagctt c 31  
 <210> 81  
 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 40 <220>  
 <223> Cebador de PCR  
 <400> 81  
 ctttctcga ggagatcgcg gatgtggcag g 31  
 <210> 82  
 45 <211> 31  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 50 <400> 82  
 cctgccacat ccgcatctc ctcgaggaaa g 31  
 <210> 83  
 <211> 32  
 <212> ADN  
 55 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR degenerado  
 <220>  
 <221> misc\_feature  
 60 <222> (18)..(18)  
 <223> "n" es a, t, g, o c  
 <400> 83  
 cagacagtgt tctgtgncg aacaaccaag cg 32  
 <210> 84  
 65 <211> 32  
 <212> ADN

<213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR degenerado  
 <220>  
 5 <221> misc\_feature  
 <222> (15)..(15)  
 <223> "n" es a, t, g, o c  
 <400> 84  
 cgcttggttg ttcgnaaca gaactctg tg 32  
 10 <210> 85  
 <211> 32  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 15 <223> Cebador de PCR degenerado  
 <220>  
 <221> misc\_feature  
 <222> (18)..(19)  
 <223> "n" es a, t, g, o c  
 20 <400> 85  
 cagacagtgt tctgtgnng aacaaccaag cg 32  
 <210> 86  
 <211> 32  
 <212> ADN  
 25 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR degenerado  
 <220>  
 <221> misc\_feature  
 30 <222> (14)..(15)  
 <223> "n" es a, t, g, o c  
 <400> 86  
 cgcttggttg ttcnnaaca gaactctg tg 32  
 <210> 87  
 35 <211> 40  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR degenerado  
 40 <220>  
 <221> misc\_feature  
 <222> (20)..(22)  
 <223> "n" es a, t, g, o c  
 <400> 87  
 45 cactcccttc cgccagatcn nngagatgac taccctcag 40  
 <210> 88  
 <211> 39  
 <212> ADN  
 <213> Secuencia artificial  
 50 <220>  
 <223> Cebador de PCR degenerado  
 <220>  
 <221> misc\_feature  
 <222> (19)..(21)  
 55 <223> "n" es a, t, g, o c  
 <400> 88  
 cgtgaggata gtcactcnn ngatctggcg gaagggagt 39  
 <210> 89  
 <211> 38  
 60 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR degenerado  
 <220>  
 65 <221> misc\_feature  
 <222> (16)..(18)

<223> "n" es a, t, g, o c  
 <400> 89  
 gtaatgatgc tccgannngc gcgacgatac gatgcggc 38  
 <210> 90  
 5 <211> 38  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR degenerado  
 10 <220>  
 <221> misc\_feature  
 <222> (21)..(23)  
 <223> "n" es a, t, g, o c  
 <400> 90  
 15 gccgcatcgt atcgtcgcgc nnntcgggagc atcattac 38  
 <210> 91  
 <211> 38  
 <212> ADN  
 <213> Secuencia artificial  
 20 <220>  
 <223> Cebador de PCR  
 <400> 91  
 tcggactcaa tattctccc gaataataga tcatatac 38  
 <210> 92  
 25 <211> 38  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 30 <400> 92  
 gtatatgac tattattcgg gaagaatatt gagtccga 38  
 <210> 93  
 <211> 34  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 93  
 tcttcaaaaa tggcccgaat ggctgataac attg 34  
 40 <210> 94  
 <211> 34  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 45 <223> Cebador de PCR  
 <400> 94  
 caatgtatc agccattcgg gccatttgt aaga 34  
 <210> 95  
 <211> 34  
 50 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 95  
 55 tcttcaaaaa tggccctaat ggctgataac attg 34  
 <210> 96  
 <211> 34  
 <212> ADN  
 <213> Secuencia artificial  
 60 <220>  
 <223> Cebador de PCR  
 <400> 96  
 caatgtatc agccattagg gccatttgt aaga 34  
 <210> 97  
 65 <211> 33  
 <212> ADN

<213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 97  
 5 acgctgggca accaggccgc cgagatgtgt ttc 33  
 <210> 98  
 <211> 33  
 <212> ADN  
 <213> Secuencia artificial  
 10 <220>  
 <223> Cebador de PCR  
 <400> 98  
 gaaacacatc tcggcggcct ggtgcccag cgt 33  
 <210> 99  
 15 <211> 33  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 20 <400> 99  
 cagaacgccg agatggcttt ctactaaag ctc 33  
 <210> 100  
 <211> 33  
 <212> ADN  
 25 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 100  
 gagcttagt gagaaagcca tctcggcgtt ctg 33  
 30 <210> 101  
 <211> 33  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 35 <223> Cebador de PCR  
 <400> 101  
 cagaacgccg agatgtcttt ctactaaag ctc 33  
 <210> 102  
 <211> 33  
 40 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 <400> 102  
 45 gagcttagt gagaaagaca tctcggcgtt ctg 33  
 <210> 103  
 <211> 34  
 <212> ADN  
 <213> Secuencia artificial  
 50 <220>  
 <223> Cebador de PCR  
 <400> 103  
 gtgatgatgc tgagagctgc ccggaatat gatg 34  
 <210> 104  
 55 <211> 34  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 60 <400> 104  
 catcatatt ccggcagct ctacgatca tcac 34  
 <210> 105  
 <211> 38  
 <212> ADN  
 65 <213> Secuencia artificial  
 <220>

<223> Cebador de PCR  
 <400> 105  
 acagattcta tagtgttcc caataaccag ccgtacac 38  
 <210> 106  
 5 <211> 38  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador de PCR  
 10 <400> 106  
 ggtacggct ggtattggg aaactata gaatctg 38  
 <210> 107  
 <211> 30  
 <212> ADN  
 15 <213> Secuencia artificial  
 <220>  
 <223> cebador  
 <400> 107  
 gggcctcag acagtattct gtcgcaac 30  
 20 <210> 108  
 <211> 38  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 25 <223> cebador  
 <400> 108  
 ggtgaagaa atccaggagt actacctgaa tacgctcc 38  
 <210> 109  
 <211> 29  
 30 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> cebador  
 <400> 109  
 35 caaggctggc atggccgagg tcacgagg 29  
 <210> 110  
 <211> 35  
 <212> ADN  
 <213> Secuencia artificial  
 40 <220>  
 <223> cebador  
 <400> 110  
 ccctccgcc agatcgtaga gatgactatc ctac 35  
 <210> 111  
 45 <211> 32  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> cebador  
 50 <400> 111 32  
 ggtaatgatg ctccaaccg cgcgacgata cg 32  
 <210> 112  
 <211> 1586  
 <212> ADN  
 55 <213> *Bamecia argentifoli*  
 <400> 112



ES 2 422 303 T3

```

gaattcgcgg ccgctcgcaa acttccgtac ctctcacccc ctcgccagga ccccccgcca      60
accagttcac cgtcatctcc tccaatggat actcatcccc catgtcttcg ggcagctacg      120
acccttatag tcccaccaat ggaagaatag ggaagaaga gctttcgccg gcgaatagtc      180
tgaacgggta caacgtggat agctgcgatg cgtcgcggaa gaagaaggga ggaacgggtc      240
ggcagcagga ggagctgtgt ctcgctcgcg gggaccgcgc ctccggctac cactacaacg      300
ccctcacctg cgaaggctgc aagggtctct tccgtcggag catcaccaag aatgccgtct      360
accagtgtaa atatggaaat aattgtgaaa ttgacatgta catgaggcga aaatgccaaag      420
agtgtcgtct caagaagtgt ctgagcgttg gcatgaggcc agaatgtgta gttccccgaat      480
tccagtgtgc tgtgaagcga aaagagaaaa aagcgcaaaa ggacaaagat aaacctaact      540
caacgacgag ttgttctcca gatggaatca aacaagagat agatcctcaa aggctggata      600
cagattcgca gctattgtct gtaaattggag ttaaaccctat tactccagag caagaagagc      660
tcatccatag gctagtttat tttcaaaatg aatatgaaca tccatcccca gaggatatca      720
aaaggatagt taatgctgca ccagaagaag aaaatgtagc tgaagaaagg tttaggcata      780
ttacagaaat tacaattctc actgtacagt taattgtgga attttctaag cgattacctg      840
gttttgacaa actaattcgt gaagatcaaa tagctttatt aaaggcatgt agtagtgaag      900
taatgatggt tagaatggca aggaggtatg atgctgaaac agattcgata ttgtttgcaa      960
ctaaccagcc gtatacgaga gaatcataca ctgtagctgg catgggtgat actgtggagg     1020
atctgctccg attttgcga catatgtgtg ccatgaaagt cgataacgca gaatatgctc     1080
ttctcactgc cattgtaatt ttttcagaac gaccatctct aagtgaaggc tggaagggtg     1140
agaagattca agaaatttac atagaagcat taaaagcata tgttgaaaat cgaaggaaac     1200
catatgcaac aaccattttt gctaagttac tatctgtttt aactgaacta cgaacattag     1260
ggaatatgaa ttcagaaaca tgcttctcat tgaagctgaa gaatagaaag gtgccatcct     1320
tctcgcagga gatttgggat gttgtttcat aaacagtctt acctcaattc catgttactt     1380
ttcatatttg atttatctca gcagggtgct cagtacttat cctcacatta ctgagctcac     1440
ggtatgctca tacaattata acttgaata tcatatcggt gatgacaaat ttgttacaat     1500

attctttggt accttaacac aatgttgatc tcataatgat gtatgaattt ttctgttttt     1560
gcaaaaaaaaa aagcggccgc gaattc                                         1586

```

- <210> 113
- <211> 1109
- <212> ADN
- 5 <213> *Nephotetix cincticeps*
- <400> 113

ES 2 422 303 T3

caggaggagc tctgcctggt gtgcggagac cgagcgtcgg gataccacta caacgctctc	60
acctgcgaag gatgcaaggg cttctttcgg aggagtatca ccaaaaacgc agtgtaccag	120
tccaaatacg gcaccaattg tgaaatagac atgtatatgc ggcgcaagtg ccaggagtgc	180
cgactcaaga agtgcctcag tgtagggatg aggccagaat gtgtagtacc tgagtatcaa	240
tgtgccgtaa aaaggaaaga gaaaaaagct caaaaggaca aagataaacc tgtctcttca	300
accaatggct cgcctgaaat gagaatagac caggacaacc gttgtgtggt gttgcagagt	360
gaagacaaca ggtacaactc gagtacgcc agtttcggag tcaaaccctc cagtccagaa	420
caagaggagc tcatccacag gctcgtctac ttccagaacg agtacgaaca ccctgccgag	480
gaggatctca agcggatcga gaacctcccc tgtgacgacg atgacctgtg tgatgttcgc	540
tacaaacaca ttacggagat cacaatactc acagtccagc tcatcgtgga gtttgcgaaa	600
aaactgcctg gtttcgacaa actactgaga gaggaccaga tcgtgttgct caaggcgtgt	660
tcgagcggag tgatgatgct gcggatggcg cggaggtacg acgtccagac agactcgatc	720
ctgttcgcca acaaccagcc gtacacgcga gagtcgtaca cgatggcagg cgtgggggaa	780
gtcatcgaag atctgctgcg gttcggccga ctcatgtgct ccatgaaggt ggacaatgcc	840
gagtatgctc tgctcacggc catcgtcatc ttctccgagc ggccgaacct ggcggaagga	900
tggaaggttg agaagatcca ggagatctac ctggaggcgc tcaagtccta cgtggacaac	960
cgagtgaaac ctgcgagtc gaccatcttc gccaaactgc tctccgttct caccgagctg	1020
cgaacactcg gcaaccagaa ctccgagatg tgcttctcgt taaactacgc aaccgcaaac	1080
atgccaccgt tcctcgaaga aatctggga	1109

## REIVINDICACIONES

1. Sistema de modulación de la expresión génica que comprende un casete de expresión génica que puede expresarse en una célula huésped que comprende un polinucleótido que codifica para un polipéptido que comprende:

5 i) un dominio de transactivación,  
 10 ii) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión ha de modularse, y

15 iii) un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución;  
 en el que el dominio de unión a ligando de receptor nuclear de grupo H procede de un receptor nuclear de grupo H seleccionado del grupo que consiste en un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína 15 de interacción de receptor retinoide X, receptores hepáticos X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor, y un receptor de farnesol, en el que el dominio de unión a ligando de receptor nuclear de grupo H se codifica por un polinucleótido  
 20 que comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido, estando el residuo de aminoácido en una posición equivalente o análoga a:

25 a) residuo de aminoácido 107 de SEQ ID NO: 1,  
 b) residuos de aminoácido 107 y 175 de SEQ ID NO: 1,  
 c) residuos de aminoácido 107 y 127 de SEQ ID NO: 1,  
 30 d) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o  
 e) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1; y

35 en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

40 2. Sistema de modulación de la expresión génica según la reivindicación 1, que comprende además un segundo dominio de unión a ligando de receptores nucleares seleccionado del grupo que consiste en un dominio de unión a ligando de receptor retinoide X de vertebrado, un dominio de unión a ligando de receptor retinoide X de invertebrado, un dominio de unión a ligando de proteína ultraespiráculo, y un dominio de unión a ligando quimérico que comprende dos fragmentos de polipéptido, en el que el primer fragmento de polipéptido procede de un dominio de unión a ligando de receptor retinoide X de vertebrado, un dominio de unión a ligando de receptor retinoide X de invertebrado, o un dominio de unión a ligando de proteína ultraespiráculo, y el segundo fragmento de polipéptido  
 45 procede de un dominio de unión a ligando de receptor retinoide X de vertebrado, dominio de unión a ligando de receptor retinoide X de invertebrado, o dominio de unión a ligando de proteína ultraespiráculo diferente.

3. Sistema de modulación de la expresión génica que comprende:

50 a) un primer casete de expresión génica que puede expresarse en una célula huésped que comprende un polinucleótido que codifica para un primer polipéptido que comprende:

i) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión ha de modularse, y  
 55 ii) un dominio de unión a ligando de receptor nuclear; y

b) un segundo casete de expresión génica que puede expresarse en la célula huésped que comprende un polinucleótido que codifica para un segundo polipéptido que comprende:

60 i) un dominio de transactivación, y  
 ii) un dominio de unión a ligando de receptor nuclear;

65 en el que uno de los dominios de unión a ligando de receptor nuclear es un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución;

5 en el que el dominio de unión a ligando de receptor nuclear de grupo H procede de un receptor nuclear de grupo H seleccionado del grupo que consiste en un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína 15 de interacción de receptor retinoide X, un receptor hepático X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor, y un receptor de farnesol;

10 en el que el dominio de unión a ligando de receptor nuclear de grupo H se codifica por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido, estando el residuo de aminoácido en una posición equivalente o análoga a:

- 10 a) residuo de aminoácido 107 de SEQ ID NO: 1,
- b) residuos de aminoácido 107 y 175 de SEQ ID NO: 1,
- 15 c) residuos de aminoácido 107 y 127 de SEQ ID NO: 1,
- d) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o
- 20 e) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1; y

20 en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

25 4. Sistema de modulación de la expresión génica según una cualquiera de las reivindicaciones 1 a 3, en el que el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un dominio de unión a ligando de receptor de ecdisona.

30 5. Sistema de modulación de la expresión génica según una cualquiera de las reivindicaciones 1 a 4, en el que el dominio de unión a ADN se selecciona del grupo que consiste en un dominio de unión a ADN de receptor de ecdisona, un dominio de unión a ADN de GAL4 y un dominio de unión a ADN de LexA.

35 6. Sistema de modulación de la expresión génica según una cualquiera de las reivindicaciones 1 a 5, en el que el dominio de transactivación se selecciona del grupo que consiste en un dominio de transactivación de receptor de ecdisona, un dominio de transactivación de VP16, un dominio de transactivación de activador ácido B42 y un dominio de transactivación de p65.

40 7. Casete de expresión génica que comprende un polinucleótido que codifica para un polipéptido que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución;

45 en el que el dominio de unión a ligando de receptor nuclear de grupo H procede de un receptor nuclear de grupo H seleccionado del grupo que consiste en un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína 15 de interacción de receptor retinoide X, un receptor hepático X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor, y un receptor de farnesol;

50 en el que el dominio de unión a ligando de receptor nuclear de grupo H se codifica por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido, estando el residuo de aminoácido en una posición equivalente o análoga a:

- 50 a) residuo de aminoácido 107 de SEQ ID NO: 1,
- b) residuos de aminoácido 107 y 175 de SEQ ID NO: 1,
- 55 c) residuos de aminoácido 107 y 127 de SEQ ID NO: 1,
- d) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o
- 60 e) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1; y

60 en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

65 8. Polinucleótido aislado que codifica para un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, comprendiendo el polinucleótido aislado una mutación de codón que da

como resultado una sustitución de un residuo de aminoácido en una posición equivalente o análoga a:

a) residuo de aminoácido 107 de SEQ ID NO: 1,

5 b) residuos de aminoácido 107 y 175 de SEQ ID NO: 1,

c) residuos de aminoácido 107 y 127 de SEQ ID NO: 1,

10 d) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1,

e) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1, o

f) aminoácidos 107, 110 y 175 de SEQ ID NO: 1; y

15 en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

20 9. Polinucleótido aislado según la reivindicación 8, en el que la mutación por sustitución es V107I/A110P/R175E de SEQ ID NO: 1.

25 10. Polinucleótido aislado que codifica para un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución, en el que el dominio de unión a ligando de receptor de ecdisona carece de actividad de unión a esteroides y el polinucleótido que codifica para el dominio de unión a ligando de receptor de ecdisona se hibrida con un polinucleótido que comprende una mutación de codón que da como resultado la mutación por sustitución V107I/A110P/R175E de SEQ ID NO: 1, en condiciones de hibridación que comprenden una etapa de hibridación en menos de 500 mM de sal y al menos 37 grados Celsius, y una etapa de lavado en 2XSSPE a al menos 63 grados Celsius.

30 11. Vector de expresión que comprende el polinucleótido aislado según la reivindicación 8, operativamente unido a un elemento regulador de la transcripción.

35 12. Célula huésped aislada que comprende el vector de expresión según la reivindicación 11, en la que el elemento regulador de la transcripción es operativo en la célula huésped.

13. Polipéptido aislado codificado por el polinucleótido aislado según la reivindicación 8.

40 14. Polipéptido aislado que comprende un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución, en el que la mutación por sustitución está en una posición equivalente o análoga a:

a) residuo de aminoácido 107 de SEQ ID NO: 1,

45 b) residuos de aminoácido 107 y 175 de SEQ ID NO: 1,

c) residuos de aminoácido 107 y 127 de SEQ ID NO: 1,

d) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o

50 e) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1; y

55 en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de un receptor nuclear de grupo H en el que el residuo de aminoácido en el residuo de aminoácido 107 en SEQ ID NO: 1 es valina.

15. Método para modular la expresión de un gen en una célula huésped que comprende las etapas de:

60 (a) introducir en la célula huésped un sistema de modulación de la expresión génica que comprende un casete de expresión génica que puede expresarse en una célula huésped que comprende un polinucleótido que codifica para un polipéptido que comprende:

i) un dominio de transactivación,

65 ii) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión ha de modularse, y

iii) un dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución;

en el que el dominio de unión a ligando de receptor nuclear de grupo H procede de un receptor nuclear de grupo H seleccionado del grupo que consiste en un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un  
 5 NER-1, un receptor nuclear de hormonas esteroideas 1, una proteína 15 de interacción de receptor retinoide X, receptores hepáticos X  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X  $\alpha$ , un receptor farnesoide X, una proteína 14 de interacción de receptor, y un receptor de farnesol;

10 en el que el dominio de unión a ligando de receptor nuclear de grupo H se codifica por un polinucleótido que comprende una mutación de codón que da como resultado una sustitución de un residuo de aminoácido, estando el residuo de aminoácido en una posición equivalente o análoga a:

a) residuo de aminoácido 107 de SEQ ID NO: 1,

15

b) residuos de aminoácido 107 y 175 de SEQ ID NO: 1,

c) residuos de aminoácido 107 y 127 de SEQ ID NO: 1,

20

d) residuos de aminoácido 107, 127 y 175 de SEQ ID NO: 1, o

e) residuos de aminoácido 52, 107 y 175 de SEQ ID NO: 1; y

25 en el que la(s) mutación/mutaciones por sustitución da(n) como resultado un aumento en la actividad de receptor nuclear de grupo H en respuesta a ligandos esteroideos y no esteroideos, con respecto a la actividad de receptor nuclear de grupo H no mutado, y

(b) introducir en la célula huésped un ligando;

30 en el que el gen que va a modularse es un componente de un casete de expresión génica que comprende:

i) un elemento de respuesta reconocido por el dominio de unión a ADN,

35

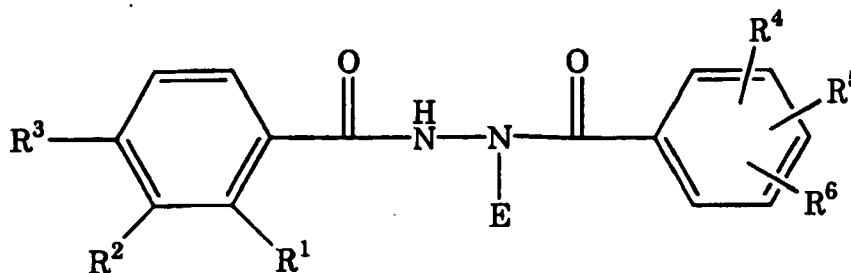
ii) un promotor que se activa por el dominio de transactivación, y

iii) un gen cuya expresión ha de modularse;

mediante lo cual tras la introducción del ligando en la célula huésped, se modula la expresión del gen de b) iii).

40 16. Método según la reivindicación 15, en el que el ligando es

a) un compuesto de fórmula:



en la que:

45

E es un alquilo (C<sub>4</sub>-C<sub>6</sub>) que contiene un carbono terciario o un cianoalquilo (C<sub>3</sub>-O<sub>5</sub>) que contiene un carbono terciario;

R<sup>1</sup> es H, Me, Et, i-Pr, F, formilo, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, OH, OMe, OEt, ciclopropilo, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, alilo, azido, SCN o SCHF<sub>2</sub>;

50

R<sup>2</sup> es H, Me, Et, n-Pr, i-Pr, formilo, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OK, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc, NMe<sub>2</sub>, NEt<sub>2</sub>, SMe, SEt, SOCF<sub>3</sub>, OCF<sub>2</sub>CF<sub>2</sub>H, COET, ciclopropilo, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, alilo, azido, OCF<sub>3</sub>, OCHF<sub>2</sub>, O-i-Pr, SCN, SCHF<sub>2</sub>, SOMe, NHCN, o está unido a R<sup>3</sup> y los carbonos de fenilo a los que están unidos R<sup>2</sup> y R<sup>3</sup> para formar un etilendioxi, un anillo dihidropirilo con el oxígeno adyacente a un carbono de fenilo, o un anillo dihidropirilo con el oxígeno adyacente a un carbono de fenilo;

55

R<sup>3</sup> es H, Et, o está unido a R<sup>2</sup> y los carbonos de fenilo a los que están unidos R<sup>2</sup> y R<sup>3</sup> para formar un etilendioxiolo, un anillo dihidrofurilo con el oxígeno adyacente a un carbono de fenilo, o un anillo dihidropirilo con el oxígeno adyacente a un carbono de fenilo;

5 R<sup>4</sup>, R<sup>5</sup> y R<sup>6</sup> son independientemente H, Me, Et, F, Cl, Br, formilo, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, OMe, OEt, SMe, o SEt; o

10 b) una ecdisona, 20-hidroxiectdisona, ponasterona A, muristerona A, un oxiestero, un 22(R)-hidroxicoestero, 24(S)-hidroxicoestero, 25-epoxicoestero, T0901317, 5-alfa-6-alfa-epoxicoestero-3-sulfato, 7-cetocolesterol-3-sulfato, farnesol, un ácido biliar, un éster de 1,1-bifosfonato, o una hormona juvenil III.

17. Método según la reivindicación 15 ó 16, que comprende además introducir en la célula huésped un segundo ligando, en el que el segundo ligando es ácido 9-cis-retinoico o un análogo sintético de un ácido retinoico.

15 18. Célula huésped aislada que comprende el sistema de modulación de la expresión génica según una cualquiera de las reivindicaciones 1 a 6.

19. Organismo no humano que comprende la célula huésped según la reivindicación 18.

20 20. Sistema de modulación de la expresión génica según la reivindicación 1, en el que la mutación por sustitución es una mutación por sustitución V107I, V107I/R175E, V107I/Y127E, V107I/Y127E/R175E o T52V/V107I/R175E de SEQ ID NO: 1.

25 21. Sistema de modulación de la expresión génica según la reivindicación 3, en el que la mutación por sustitución es una mutación por sustitución V107I, V107I/R175E, V107I/Y127E, V107I/Y127E/R175E o T52V/V107I/R175E de SEQ ID NO: 1.

30 22. Polinucleótido aislado según la reivindicación 8, en el que la mutación de codón da como resultado una mutación por sustitución, mutación por sustitución V107I, V107I/R175E, V107I/Y127E, V107I/Y127E/R175E o T52V/V107I/R175E de SEQ ID NO: 1.

23. Polipéptido aislado según la reivindicación 14, en el que la mutación por sustitución es una mutación por sustitución V107I, V107I/R175E, V107I/Y127E, V107I/Y127E/R175E o T52V/V107I/R175E de SEQ ID NO: 1.

35 24. Sistema de modulación de la expresión génica según la reivindicación 1, en el que la mutación por sustitución es una mutación por sustitución V107I/Y127E de SEQ ID NO: 1.

25. Sistema de modulación de la expresión génica según la reivindicación 3, en el que la mutación por sustitución es una mutación por sustitución V107I/Y127E de SEQ ID NO: 1.

40 26. Polinucleótido aislado según la reivindicación 8, en el que la mutación por sustitución es una mutación por sustitución V107I/Y127E de SEQ ID NO: 1.

45 27. Polipéptido aislado según la reivindicación 14, en el que la mutación por sustitución es una mutación por sustitución V107I/Y127E de SEQ ID NO: 1.

28. Casete de expresión génica según la reivindicación 7, en el que la mutación por sustitución es una mutación por sustitución V107I, V107I/R175E, V107I/Y127E, V107I/Y127E/R175E o T52V/V107I/R175E de SEQ ID NO: 1.

50 29. Casete de expresión génica según la reivindicación 28, en el que la mutación por sustitución es una mutación por sustitución V107I/Y127E de SEQ ID NO: 1.

55 30. Vector que comprende el sistema de modulación de la expresión génica según las reivindicaciones 1-6, 20, 21, 24 ó 25, el polinucleótido aislado según las reivindicaciones 8-10, 22 ó 26, o el casete de expresión génica según las reivindicaciones 7, 28 ó 29.

31. Célula huésped aislada que comprende el vector según la reivindicación 30.

60 32. Método según la reivindicación 15, que comprende además un segundo dominio de unión a ligando de receptor nuclear seleccionado del grupo que consiste en un dominio de unión a ligando de receptor retinoide X de vertebrado, un dominio de unión a ligando de receptor retinoide X de invertebrado, un dominio de unión a ligando de proteína ultraespiráculo, y un dominio de unión a ligando quimérico que comprende dos fragmentos de polipéptido, en el que el primer fragmento de polipéptido procede de un dominio de unión a ligando de receptor retinoide X de vertebrado, un dominio de unión a ligando de receptor retinoide X de invertebrado, o un dominio de unión a ligando de proteína ultraespiráculo, y el segundo fragmento de polipéptido procede de un dominio de unión a ligando de receptor retinoide X de vertebrado diferente, dominio de unión a ligando de receptor retinoide X de invertebrado, o dominio de

65

unión a ligando de proteína ultraespiráculo.

- 5 33. Método según la reivindicación 15, en el que el dominio de unión a ligando de receptor nuclear de grupo H que comprende una mutación por sustitución es un dominio de unión a ligando de receptor de ecdisona.
34. Método según la reivindicación 15, en el que el dominio de unión a ADN se selecciona del grupo que consiste en un dominio de unión a ADN de receptor de ecdisona, un dominio de unión a ADN de GAL4, y un dominio de unión a ADN de LexA.
- 10 35. Método según la reivindicación 15, en el que el dominio de transactivación se selecciona del grupo que consiste en un dominio de transactivación de receptor de ecdisona, un dominio de transactivación de VP16, un dominio de transactivación de activador ácido B42, y un dominio de transactivación de p65.



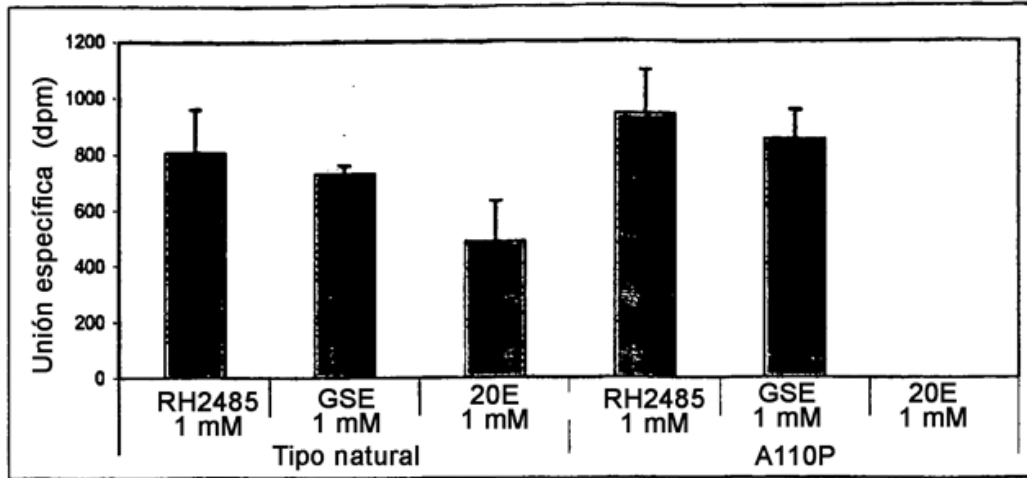


Figura 1

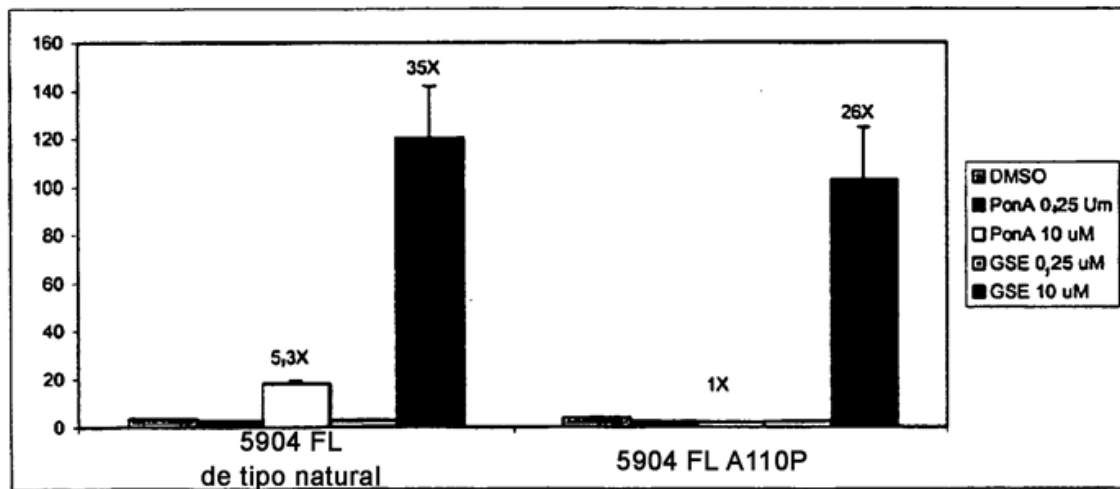


Figura 2

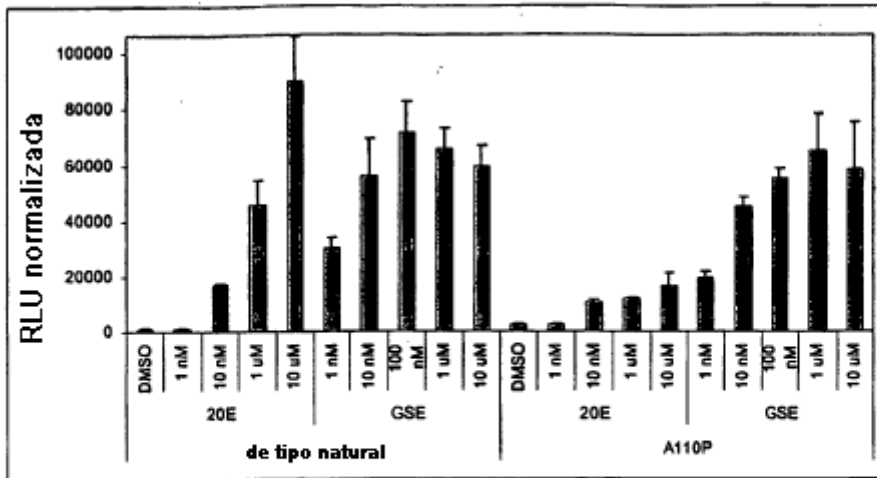


Figura 3

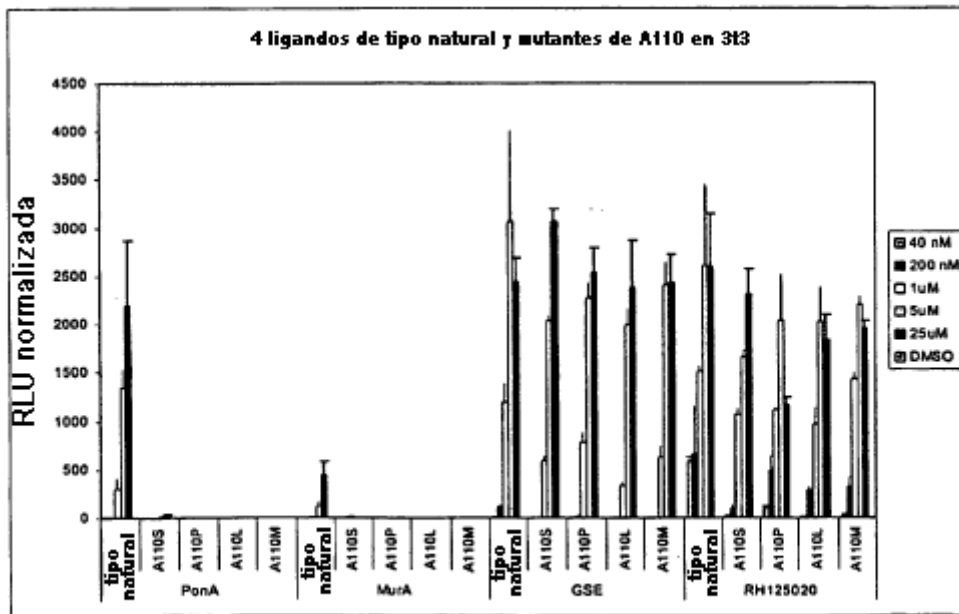


Figura 4

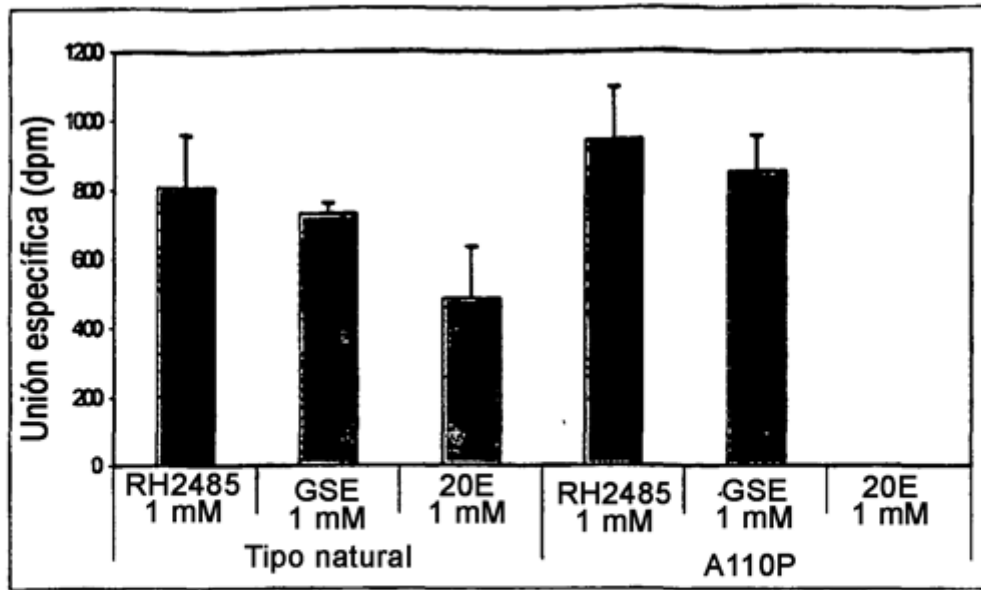


Figura 5

Unión no esteroidea (RH2485) mediante mutantes A110

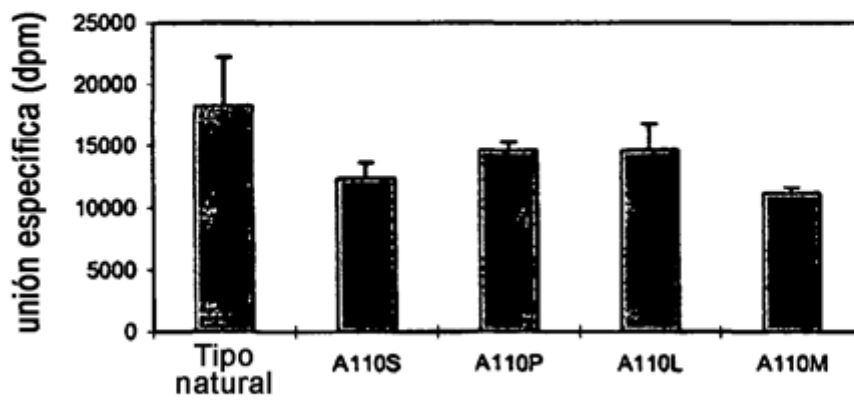


Figura 6