(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 390 070**

(51) Int. CI.:

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) Número de solicitud europea: **02717504 .1**

(96) Fecha de presentación: **20.02.2002**

(97) Número de publicación de la solicitud: **1534738**

(97) Fecha de publicación de la solicitud: **01.06.2005**

(54) Título: **Nuevos receptores mutantes de sustitución 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 313908 P

(45) Fecha de publicación de la mención BOPI:
06.11.2012

(45) Fecha de la publicación del folleto de la patente:
06.11.2012

(73) Titular/es:
INTREXON CORPORATION (100.0%)
1750 Kraft Drive, Suite 1400
Blacksburg VA 24060, US

(72) Inventor/es:
PALLI, SUBBA, REDDY y
KAPITSKAYA, MARIANNA, ZINOVJEVNA

(74) Agente/Representante:
CARPINTERO LÓPEZ, Mario

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

Nuevos receptores mutantes de sustitución y su uso en un sistema de expresión génica inducible basado en receptores nucleares

Campo de la invención

La presente invención se refiere al campo de la biotecnología o la ingeniería genética. Específicamente, la presente invención se refiere al campo de la expresión génica. Más específicamente, la presente invención se refiere a nuevos receptores nucleares que comprenden una mutación por sustitución y su uso en un sistema de expresión génica inducible a base de receptores nucleares, y a procedimientos de modular la expresión de un gen en una célula huésped usando este sistema de expresión génica inducible.

Antecedentes de la invención

En el presente documento se citan varias publicaciones. No obstante, la mención de cualquier referencia en el presente documento no debe interpretarse como admisión de que dicha referencia está disponible como “Técnica anterior” a la presente solicitud.

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

El enfoque transgénico tradicional usa un promotor específico de tipo celular para dirigir la expresión de un transgen diseñado. Primero, una construcción de ADN que contiene el transgen se incorpora en un genoma huésped. Una vez desencadenada por un activador de la transcripción, la expresión del transgen se produce en un tipo de célula dada.

Otro medio para regular la expresión de genes extraños en células es a través de promotores inducibles. Ejemplos del uso de dichos promotores inducibles incluyen el promotor PR1-a, los sistemas represor-operador procarióticas, los sistemas inmunosupresor-inmunofilina y los sistemas de activación de la transcripción de eucariotas superiores, tales como los sistemas de receptores de hormonas esteroides y se describen más adelante.

El promotor PR1-a del tabaco es inducido durante la respuesta de resistencia adquirida sistémica tras un ataque de un patógeno. El uso de PR1-a puede estar limitado porque 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 shock térmico, interferón y metales pesados (Wurn y col., 1986, Proc. Natl. Acad. Sci. USA 83: 5414-5418; Amheiter y col., 1990, Cell 62: 51-61; Filmus y col., 1992 Nucleic Acids Research 20: 27550-27560). No obstante, estos sistemas tienen limitaciones debido a su efecto sobre la expresión de genes no diana. Estos sistemas también tienen fugas.

Los sistemas de operador-represor procarióticas usan proteínas represoras bacterianas y las secuencias de ADN de operador único a las que se unen. Los sistemas represor-operador de tetraciclina ("Tet") y lactosa ("Lac") de la bacteria *Escherichia coli* se han usado en plantas y animales para controlar la expresión génica. En el sistema Tet, la tetraciclina se une a la proteína represora TetR, que tiene 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 análisis sintéticos tales como isopropilo-b-D-tiogalactósido. Por desgracia, el uso de dichos sistemas está restringido por la química inestable de los ligandos, es decir tetraciclina y lactosa, su toxicidad, su presencia natural o los niveles relativamente altos necesarios para inducción o represión. Por motivos similares, la utilidad de dichos sistemas en animales es limitada. Las moléculas inmunosupresoras, tales como FK506, rapamicina y ciclosporina A, se pueden unir a las inmunoflinas FKBP12, ciclofilina etc. Usando esta información, se ha ideado una estrategia general para juntar dos proteínas cualesquiera simplemente colocando FK506 en cada una de las dos proteínas o colocando FK506 en una y ciclosporina a en otra. Después, se puede usar 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 y col., 1993, Science 262: 1019-24; Belshaw y col., 1996, Proc Natl Acad Sci USA 93: 4604-7). El dominio de unión a ADN Gal4 condensado con el dominio activador FKBP12 y VP16 condensado con ciclofilina y el compuesto FKCsA se usaron para mostrar la heterodimerización y activación de un gen indicador bajo el control de un promotor que contiene los sitios de unión Gal4. Por desgracia, este sistema incluye inmunosupresores que pueden tener efectos secundarios indeseados y, por tanto, limita su uso para varias aplicaciones de cambio de genes de mamífero.

También se han empleado sistemas de activación de la transcripción de eucariotas superiores, tales como los sistemas del receptor de hormonas esteroides, Los receptores de hormonas esteroides son miembros de la superfamilia de receptores nucleares y se encuentran en células de vertebrados e invertebrados. Por desgracia, el uso de compuestos esteroideos que activan los receptores para la regulación de la expresión génica, en particular en plantas y mamíferos, está limitado por su implicación en muchas otras vías biológicas naturales de dichos organismos. Con el fin de superar dichas dificultades se ha desarrollado un sistema alternativo usando receptores de ecdisona de insectos (EcR).

El crecimiento, la muda y el desarrollo en insectos están regulados por la hormona esteroide ecdisona (hormona de la muda) y las hormonas juveniles (Dhadialla, et al.,1998, Annu. Rev. Entomol. 43: 545-569). La diana molecular de la ecdisona en insectos está constituida por al menos un receptor de la ecdisona (EcR) y la proteína ultraespiráculo (USP). El EcR es un miembro de la superfamilia de los receptores nucleares de esteroides que se caracteriza por dominios de unión a ADN firma y a ligando, y un dominio de activación (Koelle y col. 1991, Cell, 67:59-77). Los receptores EcR responden a una serie de compuestos esteroideos, como la ponasterona A y la muristerona A. Recientemente se han descrito compuestos no esteroideos con actividad agonista ecdiesteroide, incluidos los insecticidas disponibles comercialmente tebufenozida y metoxifenozida comercializados en todo el mundo por Rohmand Haas Company (véase la solicitud de patente internacional Nº PCT/EP96/00686 y la patente de EE.UU. 5,530,028). Ambos análogos tienen perfiles de seguridad excepcionales para otros organismos.

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

Los sistemas de expresión génica inducible estrechamente regulados o “interruptores génicos” son útiles para varias aplicaciones, tales como terapia génica, producción a gran escala de proteínas en células, ensayos de detección selectiva de alto rendimiento a base de células, genómica funcional y regulación de los rasgos en plantas y animales transgénicos.

La primera versión de interruptor génico basado en EcR usó EcR de *Drosophila melanogaster (*DmEcR) y RXR de *Mus musculus* (MmRXR) y mostró que estos receptores en presencia de esteroides, ponasterona A, transactivan los genes indicadores en líneas celulares de mamífero y ratones transgénicos (Christopherson y col., 1992; No y col., 1996). Later, Suhr y col., 1998 mostraron que el agonista no esteroideo de ecdisona, tebufenozida, inducía un nivel alto de transactivación de los genes indicadores en células de mamífero mediante el EcR de *Bombyx mori* (BmEcR) en ausencia de la pareja heterodimérica exógena.

Las solicitudes de patente internacional Nº PCT/US97/05330 (WO 97/38117) y PCT/US99/08381 (WO99/58155) divulgan procedimientos para modular la expresión de un gen exógeno en el que una construcción de ADN que comprende el gen exógeno y un elemento de respuesta a ecdisona es activada por una segunda construcción de ADN que comprende un receptor de ecdisona que, en presencia de un ligando del mismo, y opcionalmente en presencia de un receptor capaz de actuar como pareja silente, se une al elemento de respuesta a ecdisona para inducir expresión génica. El receptor de ecdisona de elección se aisló de *Drosophila melanogaster.* Normalmente, estos sistemas requieren la presencia de la pareja silente, preferentemente el receptor X de retinoide (RXR), con el fin de proporcionar una activación óptima. En células de mamífero, el receptor de ecdisona (EcR) de insectos heterodimeriza con el receptor X de retinoide (RXR) y regula la expresión de genes diana de un modo dependiente de ligando. La solicitud de patente internacional Nº PCT/US98/14215 (WO 99/02683) divulga que el receptor de ecdisona aislada del gusano de seda *Bombyx mori* es funcional en sistemas de mamífero sin la necesidad de una pareja dimérica exógena.

La patente de EE.UU. nº 6,265,173 B1 divulga que varios miembros de la superfamilia de esteroides/tiroides de receptores se pueden combinar con el receptor de ultraespiráculo (USP) de *Drosophila melanogaster* o fragmentos de los mismos que comprenden al menos el dominio de dimerización de USP para usar en un sistema de expresión génica. La patente de EE.UU. Nº 5,880,333 divulga un sistema heterodimético 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 colocados en dos proteínas híbridas diferentes. Por desgracia, 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, el dominio de transactivación y el dominio de unión a ADN (bien como EcR nativo como en la solicitud de patente internacional nº PCT/US98/14215 o como EcR modificado como en la solicitud de patente internacional nº PCT/US97/05330) se incorporaron en una única molécula y las otras parejas heterodiméricas, bien USP o RXR, se usaron en su estado nativo.

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 usar en plantas y en animales (véanse las patentes de EE.UU. nº 5,880,333 y 6,265,173 B1).

Para la mayoría de las aplicaciones que dependen de la modulación de la expresión génica, estos sistemas basados en EcR son indeseables. Por tanto, existe la necesidad en la técnica de mejores sistemas para modular con precisión la expresión de genes exógenos tanto en plantas como en animales. Dichos sistemas mejorados serían útiles para aplicaciones, tales como terapia génica, producción a gran escala de proteínas y anticuerpos, ensayos de detección selectiva de alto rendimiento a base de células, genómica funcional y regulación de los rasgos animales transgénicos. Los sistemas mejorados que son simples, compactos y dependientes de ligandos que son relativamente baratos, fácilmente disponibles y de toxicidad baja para el huésped demostrarían ser útiles para regular los sistemas biológicos.

Recientemente, los solicitantes han demostrado que un sistema de expresión génica inducible basado en el receptor de ecdisona en el que los dominios de transactivación de unión a ADN se han separado uno de otro colocándolos en dos proteínas diferentes tiene como resultado una actividad de fondo muy reducida en ausencia de un ligando y una actividad significativamente aumentado sobre el fondo en presencia de un ligando (solicitud pendiente de tramitación PCT/US01/09050). Este sistema híbrido doble es un sistema de modulación de la expresión génica inducible mejorado significativamente en comparación con los sistemas divulgados en las solicitudes PCT/US97/05330 y PGT/US98/14215. El sistema híbrido doble explota la capacidad de un par de proteínas de interacción para llevar 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 con más eficacia el promotor (véase, por ejemplo, la patente de EE.UU. nº 5,283,173). En resumen, el sistema de expresión génica híbrida doble comprende dos casetes de expresión génica; en el que el primero codifica un dominio de unión a ADN fusionado a un polipéptido del receptor nuclear y el segundo codifica un dominio de transactivación fusionado con un polipéptido del receptor nuclear diferente. En presencia de ligando, la interacción del primer polipéptido con el segundo polipéptido une con eficacia el dominio de unión a ADN al dominio de transactivación. Dado que los dominios de unión a ADN y de transactivación residen en dos moléculas diferentes, la actividad de fondo en ausencia de ligando se reduce considerablemente.

Un sistema híbrido doble también comprende una sensibilidad mejorada a los ligandos no esteroideos, por ejemplo diacilhidrazinas, cuando se comparan con ligandos esteroideos, por ejemplo ponasterona A ("PonA") o muristerona A ("MurA"). Es decir, en comparación con los esteroides, los ligandos no esteroideos proporcionan mayor actividad a una concentración menor. Además, dado que la transactivación basada en interruptores génicos del EcR suele depender de la línea celular, es más fácil adaptar los sistemas interruptores para obtener una capacidad máxima de transactivación para cada aplicación. Adicionalmente, el sistema híbrido doble evita algunos efectos secundarios por sobreexpresión de RXR que se suele producir cuando se usa el RXR sin modificar como pareja de interruptor. En un sistema híbrido doble preferido, los dominios nativos de unión a ADN y transactivación de EcR o RXR se eliminan y, como resultado, estas moléculas híbridas tienen menor probabilidad de interaccionar con otros receptores de hormonas esteroides presentes en la célula, lo que tiene como resultado menos efectos secundarios.

Recientemente, los solicitantes han realizado el sorprendente descubrimiento de que un RXR de invertebrado puede funcionar de forma similar o mejor que un RXR de vertebrado en un sistema de expresión génica inducible basado en el receptor de ecdisona (véase la solicitud de EE.UU. pendiente de tramitación 60/294,814).

El RXR es un miembro de la superfamilia de receptores nucleares y está clasificado en la subfamilia 2, Grupo B (denominado en el presente documento “receptores nucleares del grupo B”). Los miembros de cada grupo comparten una identidad de aminoácidos del 40-60 % en el dominio E (unión al ligando) (Laudet y col., A Unified Nomenclature System for the Nuclear Receptor Subfamily, 1999; Cell 97: 161-163). Además del receptor X de retinoides, otros miembros de esta subfamilia 2 de receptores nucleares, el Grupo B incluye: Proteína de unión II a la región H-2 (H-2RIIBP), co-regulador 1 del receptor nuclear (RCoR-1), ultraespiráculo (USP), receptor nuclear 2Cl y factor 1 del corion 1 (CF-1).

En un esfuerzo por proporcionar mejores dominios de unión al ligando del receptor nuclear mejorado, los solicitantes han identificado residuos de aminoácidos dentro de los receptores nucleares del Grupo B que afectan a la sensibilidad del ligando y a la magnitud de la inducción en un sistema de expresión génica inducible a base de receptores nucleares. En el presente documento, los solicitantes han descrito la construcción de receptores nucleares del Grupo B que comprenden mutaciones por sustitución (denominadas en el presente documento “mutantes por sustitución”) en estos residuos críticos y la demostración de que estos receptores nucleares mutantes por sustitución son útiles en los procedimientos de modulación de la expresión génica. Como se presenta en el presente documento, los nuevos receptores nucleares mutantes por sustitución de los solicitantes y su uso en un sistema de expresión génica inducible a base de receptores nucleares proporcionan un sistema de expresión génica inducible mejorado en células huésped tanto procarióticas como eucarióticas en las que la sensibilidad del ligando y la magnitud de la transactivación se pueden seleccionar según se desee, en función de la aplicación.

Breve descripción de las figuras

Figura 1: Transactivación del gen indicador de mutantes VP16/MmRXRα-EF, VP16/LmUSP-EF, o VP16/MmRXRα-

EF E265D (RXRmutE/D) o G293S (RXRmutG/S) transfeccionados en células NIH3T3 junto con GAL4/CfEcR-DEF y pFRLuc Las células se cultivaron en presencia de GS™-E , 0,2, 1 y 10 mM durante 48 horas y se analizó la actividad indicadora. Los números en la parte superior de las barras muestran la inducción máxima.

Figura 2: Transactivación del gen indicador de VP16/MmRXRα-EF (RXR-E), VP16/LmUSP-EF (LmUSP-E), VP16/vertebrados quiméricos RXR/invertebrados RXR (Quimera), o mutantes VP16/MmRXRα-EF E401D (MutE265D), G429S (MutG293S) o tres clones independientes del mutante doble E401D+G429S (DM1, DM2, y DM4) transfeccionados en células NIH3T3 junto con GAL4/CfEcR-DEF y pFRLuc Las células se cultivaron en presencia de GS™-E , 0,2, 1 y 10 mM durante 48 horas y se analizó la actividad indicadora. Los números en la parte superior de las barras muestran la inducción máxima.

Figura 3: Transactivación del gen indicador de GAL4CfEcRDEF, pFRLUC y VP16HsRXREFβ o lsu versión mutante de los ADN transfeccionados en células NIH3T3. Las células transfeccionadas se cultivaron en presencia de medio que contiene DMSO o GS™-E 0,04, 0,2, 1, 5, o 25 mM en DMSO. La actividad indicadora se analizó a las 48 horas de la adición de los ligandos. Los números en la parte superior de las barras muestran la inducción máxima.

Figura 4: Transactivación del gen indicador de GAL4CfEcRDEF, pFRLUC y VP16HsRXREFβ o lsu versión mutante de los ADN transfeccionados en células NIH3T3. Las células transfeccionadas se cultivaron en presencia de medio que contiene DMSO o GS™-E 0,04, 0,2, 1, 5, o 25 mM en DMSO. La actividad indicadora se analizó a las 48 horas de la adición de los ligandos. Los números en la parte superior de las barras muestran la inducción máxima.

Descripción detallada de la invención

Los solicitantes han desarrollado un nuevo sistema de expresión génica inducible a base de receptores nucleares que comprende el dominio de unión a ligando del receptor nuclear del Grupo B que comprende una mutación por sustitución. Los solicitantes han mostrado que el efecto de dicha mutación por sustitución puede aumentar la actividad de unión a ligando o la sensibilidad del ligando y puede ser un específico de esteroide o no específico de esteroide. Por tanto, la invención de los solicitantes proporciona un sistema de expresión génica inducible a base de receptores nucleares del Grupo B útil para modular la expresión de un gen de interés en una célula huésped. El nuevo sistema de expresión génica inducible y su uso en procedimientos de modulación de la expresión génica en una célula huésped supera las limitaciones de los sistemas de expresión inducibles actualmente disponibles y proporciona al experto un medio eficaz para controlar la expresión génica.

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

Definiciones

En la presente divulgación se usa una serie de términos y abreviaturas. Las definiciones siguientes se proporcionan y deberán ser útiles para la comprensión del ámbito y la práctica de la presente invención.

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

La expresión “sustancialmente libre” significa que una composición que comprende “A” (en la que “A” es una proteína sencilla, molécula de ADN, vector, célula huésped recombinante 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 (en función de la categoría de la especie a la que pertenecen A y B) en la composición es “A”. Preferentemente, “A” comprende al menos aproximadamente el 90 % en peso de la especie A+B en la composición, más preferentemente al menos aproximadamente el 99 % en peso. También se prefiere que una composición, que carece sustancialmente de contaminación, contenga sólo una especie de peso molecular único que tenga la actividad o característica de la especie de interés.

El término “aislado” para los fines de la presente invención designa un material biológico (ácido nucleico o proteína) que se ha eliminado de su ambiente original (el ambiente en el que está presente de forma natural). Por ejemplo, un polinucleótido presente en el estado natural en una planta o un animal no está aislado, aunque el mismo polinucleótido separado de los ácidos nucleicos adyacentes en los que está presente de forma natural, se considera “aislado”. El término “purificado” no requiere que el material esté presente en una forma que exhiba pureza absoluta, exclusivo de la presencia de otros compuestos. Es una definición bastante relativa.

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, preferentemente 2 o 3 y, preferentemente, 4 o 5 órdenes de magnitud.

Un “ácido nucleico” es un compuesto polimérico compuesto por subunidades unidas de forma covalente denominadas nucleótidos. El ácido nucleico incluye ácido polirribonucleico (ARN) y ácido polidesoxirribonucleico (ADN), ambos pueden ser monocatenarios o bicatenarios. El ADN incluye, entre otros, ADNc, ADN genómico, ADN plasmídico, ADN sintético y ADN semisintético. El ADN puede ser lineal, circular o superenrrollado.

Una “molécula de ácido nucleico” se refiere a la forma polimérica éster fosfato de ribonucleósidos (adenosina, guanosina, uridina o citidina; “moléculas de ARN”) o desoxirribonucleósidos (desoxiadenosina, desoxiguanosina, desoxitimidina o desoxicitidina; “moléculas de ADN”) o cualquier análogo fosfoéster de los mismos, tales como fosforotioatos y tioésteres, en forma monocatenaria, o una hélice bicatenaria. Son posibles las hélices bicatenarias de ADN-ADN, ADN-ARN y ARN-ARN. El término molécula de ácido nucleico y, en particular, molécula de ADN o ARN, sólo se refiere a la estructura primaria y secundaria de la molécula, y no se limita a una forma terciaria concreta. Por tanto, este término incluye ADN bicatenario que se encuentra, entre otros, en moléculas de ADN lineal o circular (p. ej., fragmentos de restricción), plásmidos y cromosomas. Al tratar la estructura de las moléculas de ADN bicatenarias concretas, las secuencias se pueden describir en el presente documento de acuerdo con la convención normal de proporcionar únicamente la secuencia en la dirección 5’ a 3’ a lo largo de la hebra no transcrita de ADN (es decir, la hebra tiene una secuencia homóloga al ARNm). Una “molécula de ADN recombinante” es una molécula de ADN que ha sufrido una manipulación biológica molecular.

Se entenderá que el término “fragmento” quiere decir una secuencia de nucleótidos de longitud reducida respecto al ácido nucleico de referencia y que comprende, sobre la porción común, una secuencia de nucleótidos idéntica al ácido nucleico de referencia. Dicho fragmento de ácido nucleico de acuerdo con la invención puede incluirse, cuando sea adecuado, en un polinucleótido más grande del cual es un constituyente. Dichos fragmentos comprenden o, como alternativa, consisten en oligonucleótidos de longitud variable de 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 o 1500 nucleótidos consecutivos de un ácido nucleico de acuerdo con la invención.

Como se usa en el presente documento, un “fragmento de ácido nucleico aislado” es un polímero de ARN o ADN que es monocatenario o bicatenario, que opcionalmente contiene bases nucleotídicas 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 ensamblaje de nucleótidos que codifican un polipéptido e incluye ADNc y ácidos nucleicos de ADN genómico. “Gen” también se refiere a un fragmento de ácido nucleico que expresa una proteína o polipéptido específico, incluidas secuencias reguladoras precedentes (secuencias no codificantes en 5’) y posteriores (secuencias no codificantes en 3’) a la secuencia de codificación. “Gen nativo” se refiere a un gen 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 reguladoras y/o de codificación que no se encuentran juntas en la naturaleza. En consecuencia, un gen quimérico puede comprender secuencias reguladoras y secuencias de codificación derivadas de la misma fuente, pero dispuestas de un modo distinto a como se encuentran en la naturaleza. Un gen quimérico puede comprender secuencias de codificación derivadas de diferentes fuentes y/o secuencias reguladoras derivadas de diferentes fuentes. “Gen endógeno” se refiere a un gen nativo en su localización natural en el genoma de un organismo. Un gen “extraño” o gen “heterólogo” se refiere a un gen que normalmente no se encuentra en el organismo huésped, pero que se introduce en el organismo huésped mediante transferencia génica. Los genes extraños pueden comprender genes nativos insertados en un organismo no nativo o genes quiméricos. Un “transgen” es un gen que se ha introducido en el genoma mediante un procedimiento de transformación.

ADN “heterólogo” se refiere a ADN que no está naturalmente en la célula o en un sitio cromosómico de la célula. Preferentemente, el ADN heterólogo incluye un gen extraño para la célula.

El término ”genoma” incluye ADN o ARN cromosómico, además de mitocondrial, de cloroplastos o viral.

Una molécula de ácido nucleico es “hibridable” con otra molécula de ácido nucleico, tal como ADNc, ADN genómico o ARN, cuando una forma monocatenaria de la molécula de ácido nucleico puede hibridar con la otra molécula de ácido nucleico en las condiciones adecuadas de temperatura y fuerza iónica en solución (véase Sambrook y col., 1989, más adelante). Las condiciones de hibridación y lavado son bien conocidas y se encuentran ejemplos en Sambrook, J., Fritsch, E. F. and 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 dicho documento. Las condiciones de temperatura y fuerza iónica determinan la "rigurosidad" de la hibridación.

Las condiciones de rigurosidad se pueden ajustar para detectar fragmentos moderadamente similares, tales como secuencias homólogas de organismos relacionados de forma distante, a fragmentos altamente similares, tal como genes que duplican las enzimas funcionales de organismos estrechamente relacionados. Para la detección selectiva preliminar para ácidos nucleicos homólogos, se pueden usar condiciones de hibridación de baja rigurosidad correspondientes a una Tm de 55 ºC, por ejemplo 5x SSC, 0,1 % de SDS, 0,25 % de leche y sin formamida, o 30 % de formamida, 5x SSC, 0,5 % de SDS). Las condiciones de hibridación de rigurosidad moderada corresponden a una Tm mayor, 40 5 de formamida, con 5x o 6x SCC. Las condiciones de hibridación de rigurosidad alta

corresponden a la Tm más alta, por ejemplo del 50 % de formamida, con 5x o 6x SCC.

La hibridación requiere que los dos ácidos nucleicos contengan secuencias complementarias, aunque en función de la rigurosidad de la hibridación es posible que se formen apareamientos erróneos entre bases. El término "complementario" se usa para describir la relación entre las bases nucleotídicas que son capaces de hibridar entre sí. Por ejemplo, con respecto al ADN, la adenosina es complementaria de la timina y la citosina es complementaria de la guanina. De acuerdo con esto, la presente invención también incluye fragmentos de ácido nucleico aislado que son complementarios a las secuencias completas como se divulgan o usan en el presente documento, así como las secuencias de ácido nucleico sustancialmente similares.

En una realización específica de la invención, los polinucleótidos se detectan usando condiciones de hibridación que comprenden una etapa de hibridación a Tm de 55 ºC y usando las condiciones indicadas anteriormente. En una realización preferida, la Tm es 60 ºC, en una realización más preferida, la Tm es 63 ºC, en una realización todavía más preferida, la Tm es 65 ºC.

Los lavados posthibridación también determinan las condiciones de rigurosidad. Un conjunto de condiciones preferidas usa una serie de lavados que comienzan con 6X SSC, 0,5 % de SDS a temperatura ambiente durante 15 minutos (min), después se repiten con 2X SSC, 0,5 % de SDS a 45˚C durante 30 minutos y, después, se reoiten dos vecex con 0,2X SSC, 0,5 % de SDS a 50˚durante 30 minutos . Un conjunto más preferido de condiciones de rigurosidad usa temperaturas más altas en las que los lavados son idénticos a los anteriores a excepción de que la temperatura de los dos lavados finales de 30 minutos en 0,2X SSC, 0,5 % de SDS se aumentaban a 60 ºC. Otro conjunto preferido de condiciones altamente rigurosas usa dos lavados finales en 0,1X SSC, 0,1 % de SDS a 65˚C.. La hibridación requiere que los dos ácidos nucleicos comprendan secuencias complementarias, aunque dependiendo de la rigurosidad de la hibridación, son posibles los apareamientos erróneos entre bases.

La rigurosidad adecuada para hibridar ácidos nucleicos depende de la longitud de los ácidos nucleicos y el grado de complementariedad, variables bien conocidas en la técnica. Cuanto mayor es el grado de similitud u homología entre dos secuencias de nucleótidos, mayor es el valor de T_m para los híbridos de los ácidos nucleicos que tienen dichas secuencias. La estabilidad relativa (correspondiente a una T_m más alta) de las hibridaciones de los ácidos nucleicos disminuye en el orden siguiente: ARN:ARN, ADN:ARN, ADN:ADN. Para híbridos de una longitud mayor de 100 nucleótidos se han derivado ecuaciones para calcular la T_m (véase, Sambrook y col., ant., 9.50-0,51). Para hibridación con ácidos nucleicos más cortos, es decir oligonucleótidos, la posición de las apareamientos erróneos se convierte en más importante, y la longitud del oligonucleótido determina su especificidad (véase, Sambrook y col., ant., 11.7-11.8).

En una realización específica de la invención, los polinucleótidos se detectan usando condiciones de hibridación que comprenden una etapa de hibridación a menos de 500 mM de sal y al menos 37 grados centígrados y una etapa de lavado en 2 x SSPE a al menos 63 ºC. En una realización preferida, las condiciones de hibridación comprenden sal de menos de 200 mM y al menos 37 grados centígrados para la etapa de hibridación. En una realización más preferida, las condiciones de hibridación comprenden 2XSSPE y 63 grados centígrados para las etapas de hibridación y de lavado.

En una realización, la longitud para un ácido nucleico hibridable es de al menos aproximadamente 10 nucleótidos. Preferentemente, una longitud mínima para un ácido nucleico hibridables es, de al menos, aproximadamente 15 nucleótidos; más preferentemente, de al menos aproximadamente 20 nucleótidos y más preferentemente la longitud es de al menos aproximadamente 30 nucleótidos. Además, el experto en la técnica reconocerá que la temperatura y la concentración de sales de la solución de lavado se pueden ajustar en caso necesario de acuerdo con factores tales como la longitud de la sonda.

El término "sonda" se refiere a una molécula de ácido nucleico monocatenario cuyas bases se pueden aparear con un ácido nucleico diana monocatenario complementario para formar una molécula bicatenaria.

Como se usa en el presente documento, el término "oligonucleótido" se refiere a un ácido nucleico, en general de al menos 18 nucleótidos, que es hibridable con una molécula de ADN genómico, una molécula de ADNc, un ADN de plásmido o una molécula de ARNm. Los oligonucleótidos se pueden marcar don, por ejemplo, ^{32}P-nucleótidos o nucleótidos a los que un marcador, tal como la biotina, se han unido de forma conjugada. Un oligonucleótido marcado se puede usar como sonda para detectar la presencia de un ácido nucleico. Los oligonucleótidos (uno o ambos pueden marcarse) se pueden usar como cebadores para PCR, bien para clonar la longitud completa o un fragmento de un ácido nucleico, o para detectar la presencia de un ácido nucleico. También se puede usar un oligonucleótido para formar una triple hélice con una molécula de ADN. En general, los oligonucleótidos se preparan de forma sintética, preferentemente en un sintetizador de ácido nucleico. De acuerdo con esto, los oligonucleótidos se pueden preparar con enlaces análogos de fosfoéster, tales como enlaces tioéster, etc.

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

La "reacción en cadena de la polimerasa" se abrevia a PCR y significar un procedimiento in vitro para amplificar

enzimáticamente secuencias de ácido nucleico específicas. La PCR implica una serie repetitiva de ciclos de temperatura, en la que cada ciclo comprende tres etapas: Desnaturalización del ácido nucleico molde para separar las hebras de la molécula diana, hibridar un cebador oligonucleotídico monocatenario para PCR con el ácido nucleico molde y extensión del (los) cebador(es) hibridados mediante la ADN polimerasa. La PCR proporciona un significado para detectar la presencia de la molécula diana y, en condiciones cuantitativas o semicuantitativas, para determinar la cantidad relativa de dicha molécula diana dentro de la combinación de partida de ácidos nucleicos.

"Reacción en cadena de la polimerasa-transcripción inversa" se abrevia a RT-PCR y significa un procedimiento *in vitro* para producir enzimáticamente una molécula o moléculas del ADNc diana de una molécula o moléculas de ARN, seguido de amplificación enzimática de una secuencia o secuencias de ácido nucleico específico dentro de la molécula o moléculas de ADNc diana como se ha descrito anteriormente. La RT-PCR también proporciona un significado para detectar la presencia de la molécula diana y, en condiciones cuantitativas o semicuantitativas, para determinar la cantidad relativa de dicha molécula diana dentro de la combinación de partida de ácidos nucleicos.

Una "secuencia de codificación" de ADN es una secuencia de ADN bicatenario que se transcribe y se traduce en un polipéptido en una célula *in vivo* cuando se introduce bajo el control de las secuencias reguladoras adecuadas. "Secuencias reguladoras adecuadas" se refieren a secuencias nucleotídicas localizadas cadena arriba (secuencias no codificantes en 5'), dentro, o cadena abajo (secuencias no codificantes en 3') de una secuencia de codificación y que influyen sobre la transcripción, el procesamiento del ARN o la estabilidad o la traducción de la secuencia de codificación asociada. Las secuencias reguladoras pueden incluir promotores, secuencias líder de la traducción, intrones, secuencias de reconocimiento de la poliadenilación, sitio de procesamiento del ARN, sitio de unión al efector y estructura tallo-bucle. Los límites de la secuencia de codificación están determinados por un codón de iniciación en el extremo 5' (amino) y un codón de finalización de la traducción en el extremo 3' (carboxi). Una secuencia de codificación puede incluir, entre otras, secuencias procarióticas, ADNc de ARNm, secuencias de ADN genómico e incluso secuencias de ADN sintéticas. Si la secuencia de codificación está destinada a la expresión en una célula eucariota, normalmente una señal de poliadenilación y secuencia de terminación de la transcripción se localizará en 3' de la secuencia de codificación.

"Marco de lectura abierto" se abrevia a ORF y quiere decir una longitud de secuencia de ácido nucleico, ADN, ADNc o 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 potencialmente se puede traducir en una secuencia polipeptídica.

El término "cabeza-cabeza" se usa en el presente documento para describir la orientación de las dos secuencias polinucleotídicas en relación una con otra. Dos polinucleótidos se colocan en una orientación cabeza—cabeza cuando el extremo 5' de la hebra de dosificación de un polinucleótido está adyacente al extremo 5' de la hebra de codificación del otro polinucleótido, de modo que la dirección de transcripción de cada polinucleótido procede alejándose del extremo 5' del otro polinucleótido. El término "cabeza-cabeza" se puede abreviar a (5')-a-(5') y también se puede indicar con los símbolos (←→) o (3'←5'5'→3').

El término "cola-cola" se usa en el presente documento para describir la orientación de las dos secuencias polinucleotídicas en relación una con otra. Dos polinucleótidos se colocan en una orientación cola-cola cuando el extremo 3' de la hebra de dosificación de un polinucleótido está adyacente al extremo 3' de la hebra de codificación del otro polinucleótido, de modo que la dirección de transcripción de cada polinucleótido procede acercándose al otro polinucleótido. El término "cola-cola" se puede abreviar a (3')-a-(3') y también se puede indicar con los símbolos (→←) o (5'→3'3'←5').

El término "cabeza-cola" se usa en el presente documento para describir la orientación de las dos secuencias polinucleotídicas en relación una con otra. Dos polinucleótidos se colocan en una orientación cabeza—cola cuando el extremo 5' de la hebra de dosificación de un polinucleótido está adyacente al extremo 3' de la hebra de codificación del otro polinucleótido, de modo que la dirección de transcripción de cada polinucleótido procede en la misma dirección que la de los otros polinucleótidos. El término "cabeza-cola" se puede abreviar a (5')-a-(3') y también se puede indicar con los símbolos (→→) o (5'→3'5'→3').

El término "cadena abajo" se refiere a una secuencia de nucleótidos que se localiza en 3' de la secuencia de nucleótidos de referencia. En concreto, las secuencias nucleotídicas cadena abajo se refieren, en general, a secuencias que siguen el punto de partida de la transcripción. Por ejemplo, el codón de iniciación de la traducción de un gen se localiza cadena abajo del punto de partida de la transcripción.

El término "cadena arriba" se refiere a una secuencia de nucleótidos que se localiza en 5' de la secuencia de nucleótidos de referencia. En concreto, las secuencias nucleotídicas cadena arriba se refieren, en general, a secuencias que se localizan en el lado 5' de una secuencia de codificación o punto de partida de la transcripción. Por ejemplo, la mayoría de los promotores se localizan cadena arriba del punto de partida de la transcripción.

Las expresiones "endonucleasas de restricción" y "enzimas de restricción" se refieren a una enzima que se une y corta una secuencia nucleotídica específica dentro de un ADN bicatenario.

"Recombinación homóloga" se refiere a la inserción de una secuencia de ADN extraño en otra molécula de ADN, por ejemplo la inserción de un vector en un cromosoma. Preferentemente, el vector está dirigido a un sitio cromosómico

específico para recombinación homóloga. Para la recombinación homóloga específica, el vector contendrá regiones suficientemente largas de homología con secuencias del cromosoma para permitir la unión complementaria y la incorporación del vector en el cromosoma. Las regiones de homología más largas y los grados mayores de similitud de secuencia pueden incrementar la eficiencia de la recombinación homóloga.

Varios procedimientos conocidos en la técnica se pueden usar para propagar un polinucleótido de acuerdo con la invención. Una vez que se han establecido el sistema huésped adecuado y las condiciones de crecimiento se pueden propagar los vectores de expresión recombinantes y preparar en cantidad. Como se ha descrito en el presente documento, los vectores de expresión que se pueden usar incluyen, entre otros, los siguientes vectores o sus derivados. Virus de origen humano o animal, tal como virus vaccinia o adenovirus; virus de insectos, tal como baculovirus; vectores de levaduras; vectores bacteriófagos (p. ej., lambda) y vectores de ADN de plásmido y cósmido, por citar algunos.

Un “vector” es cualquier medio para la clonación y/o transferencia de un ácido nucleico en una célula huésped. Un vector puede ser un replicón al que se puede unir otro segmento de ADN de forma que se realiza la replicación del segmento unido. Un “replicón” es cualquier elemento genético (p. ej., plásmido, fago, cósmico, cromosoma, virus) que funciona como una unidad autónoma de replicación de ADN un vivo, es decir capaz de replicarse bajo su propio control. El término “vector” incluye medios virales y no virales para introducir el ácido nucleico en una célula *in vitro, ex vivo* o *in vivo.* Un gran número de vectores conocidos en la técnica se pueden usar para manipular ácidos nucleicos e incorporar elementos de respuesta y promotores en los genes etc. Posibles vectores incluyen, por ejemplo, plásmidos o virus modificados incluidos, por ejemplo, bacteriófagos, tales como los derivados de lambda, o plásmidos tales como los derivados plasmídicos pBR322 o pUC, o el vector Bluescript. Por ejemplo, la inserción de los fragmentos de ADN correspondientes a los elementos de respuesta y los promotores en un vector adecuado se puede conseguir ligando los fragmentos de ADN adecuados en un vector elegido que tenga términos complementarios cohesivos. Como alternativa, los extremos de las moléculas de ADN pueden modificarse enzimáticamente o se puede producir cualquier sitio ligando las secuencias nucleotídicas (ligadores) en los extremos de ADN. Dichos vectores se pueden someter a ingeniería para contener genes marcadores seleccionables que proporcionan la selección de células que tienen incorporado el marcador en el genoma celular. Dichos 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 liberación génica en células, así como sujetos animales vivos. Los vectores virales que se pueden usar incluyen, entre otros, vectores de retrovirus, virus adenoasociados, virus pox, baculovirus, vaccinia, herpes simple, Epstein-Barr, adenovirus, geminivirus y caulimovirus. Vectores no virales incluyen plásmidos, liposomas, lípidos con carga eléctrica (citofectinas), complejos ADN-proteína y biopolímeros. Además de un ácido nucleico, un vector puede comprender también una o más regiones reguladoras y/o marcadores seleccionables útiles en la selección, medición y monitorización de los resultados de transferencia de ácido nucleico (transferencia a qué tejidos, duración de la expresión etc.).

El término “plásmido” se refiere a un elemento cromosómico adicional que a menudo porta un gen que no es parte del metabolismo central de la célula y normalmente en forma de moléculas de ADN bicatenario circular. Dichos elementos pueden ser secuencias de replicación autónoma, secuencias de integración de genoma, secuencias de fagos o nucleótidos, lineales, circulares o superenrolladas, de un ADN o ARN monocatenario o bicatenario, derivadas de cualquier fuente, en las que una serie de secuencias nucleotídicas se han unido o recombinado en una única construcción que es capaz de introducir en una célula un fragmento promotor y la secuencia de ADN para un producto génico seleccionado junto con la secuencia no traducida en 3' adecuada.

Un “vector de clonación” es un “replicón”, que es una longitud de unidad de un ácido nucleico, preferentemente ADN, que se replica secuencialmente y que comprende un origen de replicación, tal como un plásmido, fago o cósmido, al que se puede unir otro segmento de ácido nucleico de forma que se realiza la replicación del segmento unido. Los vectores de clonación pueden realizar la replicación en un tipo celular y la expresión en otro (“vector lanzadera”).

Los vectores se pueden introducir en las células huésped deseadas mediante procedimientos conocidos en la técnica, por ejemplo transfección, electroporación, microinyección, transducción, fusión celular, dextrano DEAE, precipitación en fosfato cálcico, lipofección (fusión de lisosomas), uso de una pistola génica o uso de un transportador de vector de ADN (véase, p. ej., Wu y col., J. Biol. Chem. 267:963-967 (1992); Wu y Wu, 1998, J. Biol. Chem. 263:14621-14624 y Hartmut y col. solicitud de patente canadiense Nº 2,012,311, presentada el 15 de marzo de 1990).

También se puede introducir *in vivo* mediante lipofección un polinucleótido de acuerdo con la invención. Durante la última década se ha producir un uso creciente de liposomas para encapsulación y transfección de ácidos nucleicos *in vitro*. Los lípidos catiónicos sintéticos diseñados para limitar las dificultades y peligros encontrados con la transfección mediada por liposomas se pueden usar para preparar liposomas para la transfección *in vivo* de un gen que codifica un marcador (Felgner, y col., 1987, Proc. Natl. Acad. Sci. U.S.A. 84:7413-7417; Mackey y col., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:8027-8031 y Ulmer y col. 1993, Science 259:. 1745-1748). El uso de lípidos catiónicos puede estimular la encapsulación de ácidos nucleicos con carga negativa y también promueve la fusión

con membranas celulares cargadas negativamente [Felgner y Ringold, 1989, Science 337:387-388]. Compuestos y composiciones lipídicas particularmente útiles para transferencia de ácidos nucleicos se describen en las publicaciones de patente internacional WO95/18863 y WO96/17823, y la patente de EE.UU. Nº 5,459,127. El uso de la lipofección para introducir genes exógenos en los órganos específicos *in vivo* tiene ciertas ventajas prácticas. Dirigir los liposomas a células específicas representa un área de beneficio. Está claro que dirigir la transfección a tipos celulares concretos sería particularmente preferido en un tejido con heterogeneidad celular, tal como páncreas, hígado, riñón y cerebro. Los lípidos pueden acoplarse químicamente a otras moléculas con el fin de dirigir el tratamiento (Mackey, y col., 1988, ant.). Los péptidos dirigidos, por ejemplo hormonas o neurotransmisores, y proteínas, como anticuerpos o moléculas no peptídicas, podrían acoplarse a los liposomas químicamente.

Otras moléculas también son útiles para facilitar la transfección de un ácido nucleico *in vivo*, tal como un oligopéptido catiónico (p. ej., el documento WO95/21931), péptidos derivados de proteínas de unión a ADN (p. ej., el documento WO96/25508), o un polímero catiónico (p. ej., el documento WO95/21931).

También es posible introducir un vector *in vivo* como un plásmido de ADN desnudo (véanse las patentes de EE.UU. 5,693,622, 5,589,466 y 5,580,859). La liberación de ADN mediada por receptores también se puede usar (Curiel y col., 1992. Hum. Gene Ther. 3:147-154; y Wu y Wu, 1987. J. Biol. Chem 262: 4429-4432).

El término “transfección” significa la captación e ARN o ADN exógeno o heterólogo por una célula. Una célula se ha “transfecionado” con ADN o ARN exógeno o heterólogo cuando dicho ADN o ARN se ha introducido en el interior de la célula. Una célula se ha “transformado” con ADN o ARN exógeno o heterólogo cuando dicho ADN o ARN efectúa un cambio fenotípico. El ARN o ADN transformante puede integrarse (unido de forma covalente) en el ADN cromosómico que compone el genoma de la célula.

“Transformación” se refiere a la transferencia de un fragmento de ácido nucleico en el genoma de un organismo huésped, que tiene como resultado una herencia genéticamente estable. Organismos huésped que contienen los fragmentos de ácido nucleico transformados se denominan organismos “transgénicos” o “recombinantes” o “transformados”.

La expresión “región genética” hace referencia a una región de una molécula de ácido nucleico o una secuencia nucleotídica que comprende un gen que codifica un polipéptido.

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

La expresión “marcador seleccionable” quiere decir un factor de identificación, normalmente un gen de resistencia a antibiótico o química, que es capaz de seleccionarse en base al efecto del gen marcador, es decir la resistencia a un antibiótico, la resistencia a un herbicida, marcadores colorimétricos, enzimas, marcadores fluorescentes y similares, en los que el efecto se usa para rastrear la herencia de un ácido nucleico de interés y/o identificar una célula u organismo que ha heredado el ácido nucleico de interés. Ejemplos de genes marcadores seleccionables conocidos y usados en la técnica incluyen: genes que proporcionan resistencia a ampicilina, estreptomicina, gentamicina, kanamicina, higromicina, herbicida bialafos, sulfonamida y similares, y genes que se usan como marcadores fenotípicos, es decir genes reguladores de la antocianina, gen de la isopentaniltransferasa y similares.

La expresión “gen indicador” quiere decir un ácido nucleico que codifica un factor de identificación que es capaz de identificarse en base al efecto del gen indicador, en el que el efecto se usa para rastrear la herencia de un ácido nucleico de interés, para identificar una célula u organismo que ha heredado el ácido nucleico de interés y/o para medir la inducción de expresión génica o transcripción. Ejemplos de genes indicadores conocidos y usados en la técnica incluyen: luciferasa (Luc), proteína fluorescente verde (GFP), cloranfenicol acetiltransferasa (CAT), β-galactosidasa (LacZ), β-glucuronidasa (Gus), y similares. Los genes marcadores seleccionables también se pueden considerar genes indicadores.

“Promotor” se refiere a una secuencia de ADN capaz de controlar la expresión de una secuencia de codificación o ARN funcional. En general, una secuencia de codificación se localiza en 3’ de una secuencia promotora. Los promotores pueden derivar en su totalidad de un gen nativo o estar compuestos de diferentes elementos derivados de diferentes promotores encontrados en la naturaleza, o incluso comprenden 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 etapas de desarrollo, o en respuesta a diferentes condiciones ambientales o fisiológicas. Los promotores que hacen que un gen se exprese en la mayoría de los tipos celulares se denominan, en la mayoría de los casos, “promotores constitutivos”. Los promotores que hacen que un gen se exprese en un tipo celular específico normalmente se denominan “promotores específicos de células” o “promotores específicos de tejido”. Los promotores que hacen que un gen se exprese en una etapa específica de desarrollo o diferenciación celular normalmente se denominan “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 tratamiento de la célula con un agente, molécula biológica, sustancia química, ligando, luz o similar, que induce el promotor normalmente se denominan "promotores inducibles" o "promotores regulables". Además se reconoce que,

ya que en la mayoría de los casos los límites exactos de las secuencias reguladoras no se han definido completamente, los fragmentos de ADN de diferentes longitudes pueden tener una actividad promotora idéntica.

Una “secuencia promotora” es una región reguladora de ADN capaz de unir la ARN polimerasa en una célula e iniciar la transcripción de una secuencia de codificación (dirección 3’) posterior. Para los fines de definir la presente invención, la secuencia promotora se une a su extremo 3’ mediante el sitio de inicio de la transcripción y se extiende hacia arriba (en dirección 5’) para incluir el número mínimo de bases o elementos necesarios para iniciar la transcripción a niveles detectables por encima del normal. Dentro de la secuencia promotora se encontrará un sitio de inicio de la transcripción (convenientemente definido mediante, por ejemplo, mapeo con la nucleasa S1), así como los dominios de unión a proteína (secuencias consenso) responsables de la unión de la ARN polimerasa.

Una secuencia de codificación está “bajo el control” de las secuencias de control de la transcripción y la traducción en una célula cuando la ARN polimerasa transcribe la secuencia de codificación en ARNm, que a su vez es trans-ARN de corte y empalme (si la secuencia de codificación contiene intrones) y se traduce en la proteína codificada por la secuencia de codificación.

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

La expresión “elemento de respuesta” significa uno o más elementos de ADN de acción cis que confieren capacidad de respuesta en un promotor mediada por la interacción con los dominios de unión a ADN del primer gen quimérico. Este elemento de ADN puede ser de secuencia palindrómica (perfecto o imperfecto) o compuesto por motivos de secuencia o semisitios separados por un número variable de nucleótidos. Los semisitios pueden ser similares o idénticos y estar dispuestos como repeticiones directas o invertidas o como un semisitio único o multímeros de semisitios adyacentes en tándem. El elemento de respuesta puede comprender un promotor mínimo aislado de diferentes organismos en función 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 de gen(es) cadena abajo bajo la regulación de este elemento de respuesta. Ejemplos de secuencias de ADN para los elementos de respuesta del receptor de ecdisona natural incluyen: RRGG/TTCANTGAC/ACYY (véase Cherbas L., y col., (1991), Genes Dev. 5, 120-131); AGGTCAN(n)AGGTCA, en la que N(n) puede ser uno o más nucleótidos espaciadores (véase D’Avino PP., y col., (1995), Mol. Cell. Endocrinol, 113, 1-9); y GGGTTGAATGAATTT (véase Antoniewski C., y col., (1994). Mol. Cell Biol. 14, 4465-4474).

La expresión “operablemente unido” se refiere a la asociación de secuencias de ácido nucleico en un fragmento de ácido nucleico sencillo de modo que la función de una está afectada por la otra. Por ejemplo, un promotor está operablemente unido a una secuencia de codificación cuando es capaz de afectar a la expresión de dicha secuencia de codificación (es decir, que la secuencia de codificación está bajo el control transcripcional del promotor). Las secuencias de codificación pueden estar operablemente unidas a secuencias reguladoras en orientación sentido o antisentido.

El término “expresión”, como se usa en el presente documento, se refiere a la transcripción y acumulación estable de ARN sentido (ARNm) o antisentido derivado de un ácido nucleico o polinucleótido. La expresión también se puede denominar traducción de ARNm en una proteína o polipéptido.

Los términos “casete”, “casete de expresión” y “casete de expresión génica” hacen referencia a un segmento de ADN que se puede insertar en un ácido nucleico o polinucleótido en sitios de restricción específicos o mediante recombinación homóloga. El segmento de ADN comprende un polinucleótido que codifica un polipéptido de interés y el casete y los sitios de restricción se han diseñado para garantizar la inserción del casete en un marco de lectura adecuado para la transcripción y la traducción. “Casete de transformación” se refiere a un vector específico que comprende un polinucleótido que codifica un polipéptido de interés y que tiene elementos además del polinucleótido que facilita la transformación de una célula huésped concreta. Los casetes, casetes de expresión, casetes de expresión génica y casetes de transformación de la invención pueden también comprender elementos que permiten la expresión potenciada de un polinucleótido que codifica un polipéptido de interés en una célula huésped. Estos elementos pueden incluir, entre otros: un promotor, un promotor mínimo, un potenciador, un elemento de respuesta, una secuencia de terminación, una secuencia de poliadenilación y similares.

Para los fines de la presente invención, la expresión “interruptor génico” hace referencia 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 cual el elemento de respuesta y el promotor se incorporan.

Los términos “modula” y “modulan” quieren decir inducir, reducir o inhibir la expresión de ácido nucleico o génica, lo que tiene como resultado la correspondiente inducción, reducción o inhibición de la producción de proteínas o polipéptidos.

Los plásmidos o vectores de acuerdo con la invención pueden comprender además al menos un promotor adecuado para dirigir la expresión de un gen en una célula huésped. La expresión “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, normalmente está colocada bajo el control de los elementos de control, como un promotor, un promotor mínimo, un potenciador o similares. Las regiones control de la iniciación o promotores que son útiles para dirigir la expresión de un ácido nucleico en una célula huésped deseada son numerosos y familiares para los expertos en la técnica. Casi cualquier promotor capaz de dirigir estos genes es adecuado para la presente invención, incluidos, entre otros: promotores virales, promotores bacterianos, promotores animales, promotores de mamífero, promotores sintéticos, promotores constitutivos, promotores específicos de tejido, promotores específicos del desarrollo, promotores inducibles, promotores regulados por luz; CYCI, HIS3, GAL1, GAL4, GAL10, ADH1, PGK, PHO5, GAPDH, ADC1, TRP1, URA3, LEU2, ENO, TPI, promotores de la fosfatasa alcalina (útiles para expresión en *Saccharomyces*); promotor AOXI (útil para expresión en *Pichia*); β-lactamasa, lac, ara, tet, trp, IPL, IPR, T7, tac, y promotores trc (útiles para expresión en *Escherichia coli*); promotores regulados por luz, específicos de semillas, específicos de polen, específicos de ovarios, relacionados con la patogenia o con una enfermedad, 35 S del virus del mosaico de la coliflor, mínimo 35S de CMV, del virus del mosaico de las nervaduras (CsVMV), próteína de unión a clorofila alb, ribulosa 1, 5-bisfosfato carboxilasa, específicos de brotes, específicos de raíz, 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 de la antocianina (útiles para expresión en células vegetales); promotores animales y de mamífero conocidos en la técnica incluyen, entre otros, 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 mayores E1A (MLP) de adenovirus (Ad), el promotor temprano del citomegalovirus (CMV), el promotor de la timidina quinasa (PGK) del virus del herpes simple (HSV) (TK), un promotor del factor de elongación 1 alfa (EF1), un promotor de la fosfoglicerato quinasa (PGK), un promotor de la ubiquitina (Ubc), un promotor de la albúmina, las secuencias reguladoras del promotor de la L-metalotioneína de ratón y regiones de control de la transcripción, los promotores ubicuos (HPRT, vimentina, α-actina, tubulina y similares), los promotores de los filamentos intermedios (desmina, neurofilamentos, queratina, GFAP, y similares), los promotores de los genes terapéuticos (de MDR, CFTR o el factor VIII y similares), promotores relacionados co patogenia o enfermedad y promotores que exhiben especificidad de tejido y que se han usando en animales transgénicos, tales como la región control del gen I de la elastasa que es activo en células acinares pancreáticas; la región control del gen de la insulina activo en las células pancreáticas beta, la región control del gen de la inmunoglobulina activa en células linfoides, la región control del virus del tumor de mama de ratón activa en células testiculares, de mama, linfoides y mastocitos; regiones control del gen de la albúmina, Apo AI y Apo AII activas en el hígado, región control del gen de la alfa-fetoproteína activa en el hígado, in región control del gen de la alfa 1-antitripsina activa en el hígado, región control del gen de la beta-globina activa en células mieloides, región control del gen de la proteína básica de la mielina activa en células oligodendrocitos en el cerebro, región control del gen de la cadena-2 ligera de la miosina activa en el músculo esquelético y región control del gen de la hormona liberadora de la gonadotropina activa en el hipotálamo, promotor de la piruvato quinasa, promotor de la vilina, promotor de la proteína intestinal de unión a ácidos grasos, promotor de la α-actina de las células de músculo liso y similares. Además, estas secuencias de expresión se pueden modificar mediante la adición de secuencias potenciadoras o reguladoras y similares.

Los potenciadores que se pueden usar en las realizaciones de la invención incluyen, entre otros: un potenciador de SV40, un potenciador de citomegalovirus (CMV), un potenciador del factor 1 de elongación EF1), potenciadores de levaduras, potenciadores de genes virales y similares.

Las regiones control de la terminación, es decir las secuencias terminadoras o de poliadenilación, también pueden derivar de varios genes nativos de los huéspedes preferidos. Opcionalmente, un sitio de terminación puede ser innecesario, aunque se prefiere su inclusión. En una realización preferida de la invención, la región control de la terminación puede estar compuesta por o derivada de una secuencia sintética, señal de poliadenilación sintética, una señal de poliadenilación tardía de DV40, una señal de poliadenilación de SV40, una señal de poliadenilación de la hormona de crecimiento bovina (BGH), secuencias de terminación virales o similares.

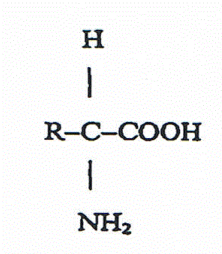
Los términos "secuencias no codificantes en 3'! o "región no traducida en 3' (UTR)" hacen referencia a secuencias de ADN localizadas cadena abajo (3') de una secuencia de codificación y pueden comprende secuencias de reconocimiento de poliadenilación [poli(A)] y otras secuencias que codifican señales reguladoras capaces de afectar al procesamiento del ARNm o a la expresión génica. La señal de poliadenilación normalmente se caracteriza porque afecta a la adición de tiras de ácido poliadenílico al extremo 3' del precursor de ARNm.

"Región reguladora" quiere decir una secuencia de ácido nucleico que regula la expresión de una segunda secuencia de ácido nucleico. Una región reguladora puede incluir secuencias responsables de forma natural para expresar un ácido nucleico concreto (una región homóloga) o puede incluir secuencias de un origen diferente que son responsables de la expresión de diferentes proteínas o incluso proteínas sintéticas (o región heteróloga). En concreto, las secuencias pueden ser secuencias de genes procariotas, eucariotas o virales, o secuencias derivadas que estimulan o reprimen la transcripción de un gen de un modo específico o inespecífico y de un modo 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 al polipéptido a las vías secretoras de la célula diana.

Una región reguladora de una “fuente heteróloga” es una región reguladora que no está asociada de forma natural con el ácido nucleico expresado. Incluidas entre las regiones reguladoras heterólogas están las regiones reguladoras de una especie diferente, las regiones reguladoras de un gen diferente, regiones reguladoras híbridas y regiones reguladoras que no se producen en la naturaleza pero que están diseñadas por un experto en la técnica.

“Tránscrito de ARN” hace referencia al producto resultante de la transcripción catalizada por la ARN polimerasa de una secuencia de ADN. Cuando el tránscrito de ARN es una copia complementaria perfecta de la secuencia de ADN, se denomina el tránscrito primario o puede ser una secuencia de ADN derivada del procesamiento postranscripcional del tránscrito primario y se denomina ARN maduro. “ARN mensajero (ARNm)” se refiere al ARN que no tiene intrones y que se puede traducir en proteínas en la célula. “ADNc” se refiere a un ADN bicatenario que es complementario y derivad el ARNm. ARN “sentido” se refiere al tránscrito de ARN que incluye ARNm y, de este modo, se puede traducir en proteínas en la célula. “ARN antisentido” se refiere a un tránscrito de ARN que es complementario a todo o parte de un tránscrito primario o ARNm y que bloquea la expresión de un gen diana. La complementariedad de un ARN antisentido puede ser cualquier parte del tránscrito génico específico, es decir en la secuencia no codificante en 5’, la secuencia no codificante en 3’ o la secuencia de codificación. “ARN funcional” se refiere a ARN antisentido, ARN de ribozima u otro ARN no traducido, auque tiene un efecto sobre los procesos celulares.

Un “polipéptido” es un compuesto polimérico que consta de residuos aminoácidos unidos covalentemente. Los aminoácidos presentan la estructura general siguiente:



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 encuentra fusionado con el grupo amino. Un polipéptido de la invención Preferiblemente comprende por lo menos aproximadamente 14 aminoácidos.

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

Un "polipéptido aislado" o una "proteína aislada" es un polipéptido o proteína que se encuentra sustancialmente libre de aquellos compuestos que normalmente se encuentran asociados a los mismos en su estado natural (por ejemplo otras proteínas o polipéptidos, ácidos nucleicos, carbohidratos y lípidos). El término "aislado" no pretende excluir mezclas artificiales o sintéticas con otros compuestos, o la presencia de impurezas que no interfieren con la actividad biológica, y que pueden encontrarse presentes, por ejemplo, debido a una purificación incompleta, la adición de estabilizadores o la formación de un compuesto en una preparación farmacéuticamente aceptable.

Un "polipéptido mutante por sustitución" o un "mutante por sustitución" se entiende que significa al menos un polipéptido mutante que comprende una sustitución de al menos un (1) aminoácido de tipo silvestre o natural por un aminoácido diferente respecto al polipéptido de tipo silvestre o natural. Un polipéptido mutante por sustitución puede comprender sólo una (1) sustitución de aminoácido de tipo silvestre o natural y puede denominarse como polipéptido "mutante puntual" o "mutante puntual único". Alternativamente, un polipéptido mutante por sustitución puede comprender una sustitución de dos (2) o más aminoácidos de tipo silvestre o natural por 2 o más aminoácidos respecto al polipéptido de tipo silvestre o natural. De acuerdo con la invención, un polipéptido de dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución comprende una sustitución de al menos un (1) aminoácido de tipo silvestre o natural por un aminoácido diferente respecto al polipéptido de dominio de unión a ligando de receptor nuclear de grupo B de tipo silvestre o natural.

Cuando el polipéptido mutante por sustitución comprende una sustitución de (2) o más aminoácidos de tipo silvestre o natural, esta sustitución puede comprender un número equivalente de aminoácidos de tipo silvestre o natural delecionados para la sustitución, es decir, 2 aminoácidos de tipo silvestre o natural sustituidos con 2 aminoácidos de tipo no silvestre o no natural o un número no equivalente de aminoácidos de tipo silvestre delecionados para la sustitución, es decir, 2 aminoácidos de tipo silvestre sustituidos con 1 aminoácido de tipo no silvestre (una mutación por sustitución + deleción) o 2 aminoácidos de tipo silvestre sustituidos con 3 aminoácidos de tipo no silvestre (una mutación por sustitución + inserción).

Los mutantes por sustitución se pueden describir usando un sistema de nomenclatura abreviada para indicar el

residuo aminoácido y el número sustituido en la secuencia polipeptídica de referencia y el nuevo residuo aminoácido sustituido. Por ejemplo, un mutante por sustitución en el cual el vigésimo (20) residuo aminoácido de un polipéptido está sustituido puede abreviarse como "x20z", en donde "x" es el residuo aminoácido a sustituir, "20" es la posición del residuo aminoácido o número dentro del polipéptido y "z" es el nuevo aminoácido sustituido. Por consiguiente, un mutante por sustitución abreviado indistintamente como "E20A" o "Glu20Ala" indica que el mutante comprende un residuo alanina (frecuentemente abreviado en la técnica como "A" o "Ala") en lugar del ácido glutámico (frecuentemente abreviado en la técnica como "E" o "Glu") en la posición 20 del polipéptido.

Una mutación por sustitución puede crearse mediante cualquier procedimiento de mutagénesis conocido en la técnica, incluyendo, aunque sin limitación, la mutagénesis dirigida al sito *in vitro* (Hutchinson, C., et al., 1978, J. Biol. Chem. 253:6551; Zoller and Smith, 1984, DNA 3:479-488; Oliphant et al., 1986, Gene 44:177; Hutchinson et al.,1986, Proc. Natl. Acad. Sci. U.S.A. 83: 710), el uso de enlazadores TAB® (Pharmacia), la digestión por endonucleasas de restricción/deleción de fragmentos y sustitución, la mutagénesis mediada por PCR/dirigida por oligonucleótidos y similares. Las técnicas basadas en la PCR son las preferidas para la mutagénesis dirigida al sitio (ver Higuchi, 1989, "Using PCR to Engineer DNA", in PCR Technology: Principles and Applications for DNA Amplification, H. Erlich, ed., Stockton Press, Capítulo 6, págs. 61-70).

Un "fragmento" de un polipéptido de acuerdo con la invención se entiende que hace referencia a un polipéptido cuya secuencia de aminoácidos es más corta que la del polipéptido de referencia y que comprende, a lo largo de la porción completa con dichos polipéptidos de referencia, una secuencia de aminoácidos idéntica. Dichos fragmentos pueden, en caso apropiado, incluirse en un polipéptido de mayor tamaño del que forman una parte. Dichos fragmentos de un polipéptido de acuerdo con la invención pueden presentar 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 o 300 aminoácidos.

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 por lo menos una propiedad biológica del polipéptido o proteína. Pueden existir diferentes variantes del polipéptido o proteína en la naturaleza. Estas variantes pueden ser variaciones alélicas caracterizadas por diferencias en las secuencias de nucleótidos del gen estructural codificante de la proteína, o pueden implicar una modificación diferente durante el procesamiento o después de la traducción. El experto en la materia puede producir variantes que presentan sustituciones, deleciones, adiciones o remplazamientos de un único o múltiples aminoácidos. Estas variantes pueden incluir, entre otros, (a) variantes en las que se sustituyen uno o más residuos aminoácidos por aminoácidos conservadores o no conservadores, (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 albúmina sérica. Las técnicas para obtener estas variantes, incluyendo las técnicas genéticas (supresiones, deleciones, mutaciones, etc.), químicas y enzimáticas, son conocidas por el experto ordinario en la materia. Un polipéptido variante Preferiblemente comprende por lo menos aproximadamente 14 aminoácidos.

Una "proteína heteróloga" se refiere a una proteína no producida naturalmente en la célula.

Una "proteína madura" se refiere a un polipéptido procesado post-traduccionalmente, es decir, uno del que cualquier prepéptido o propéptido presente en el producto primario de traducción ha sido eliminado. La proteína "precursora" se refiere al producto primario de la traducción del ARNm, es decir, con los prepéptidos y propéptidos todavía presentes. Los prepéptidos y propéptidos pueden encontrarse limitados, aunque no necesariamente, a señales de localización intracelular.

La expresión "péptido de señal" se refiere a un polipéptido aminoterminal que precede a la proteína madura secretada. El péptido de señal se escinde de la proteína madura y por lo tanto no se encuentra presente en la misma. Los péptidos de señal tienen la función de dirigir y traslocar las proteínas secretadas a través de las membranas celulares. El péptido de señal también se denomina proteína de señal.

Se incluye una "secuencia de señal" al inicio de la secuencia codificante de una proteína que debe expresarse sobre la superficie de una célula. Esta secuencia codifica un péptido de señal, en posición N-terminal respecto al polipéptido maduro, que dirige que la célula huésped trasloque el polipéptido. La expresión "secuencia de señal de traslocación" se usa en la presente memoria para referirse a este tipo de secuencia de señal. Las secuencias de señal de traslocación pueden encontrarse asociadas a una diversidad de proteínas nativas a los eucariotas y procariotas, y con frecuencia son funcionales en ambos tipos de organismo.

El término "homología" se refiere al porcentaje de identidad entre dos polinucleótidos o entre dos grupos polipeptídicos. La correspondencia entre la secuencia de un grupo y otro puede determinarse mediante técnicas conocidas de la técnica. Por ejemplo, la homología puede determinarse mediante una comparación directa entre la información de secuencia entre dos moléculas de polipéptido mediante la alineación de la información de secuencia y el uso de programas informáticos fácilmente disponibles. Alternativamente, la homología puede determinarse mediante hibridación de los polinucleótidos bajo condiciones que forman dúplex estables entre regiones homólogas, seguido de la digestión con una o más nucleasas específicas de la forma de cadena individual y la determinación de los tamaños de los fragmentos digeridos.

Tal como se usa en la presente memoria, el término "homólogo" en todas sus formas gramaticales y variaciones ortográficas se refiere a la relación entre proteínas que poseen un "origen evolutivo común", incluyendo proteínas de superfamilias (por ejemplo, la superfamilia de las inmunoglobulinas) y las proteínas homólogas de diferentes especies (por ejemplo la cadena ligera de la miosina, etc. (Reeck et al., Cell 50:667, 1987). Estas proteínas (y los genes codificantes de las mismas) presentan homología de secuencia, tal como se pone de manifiesto por su elevado grado de similitud de secuencias. Sin embargo, en su uso habitual y en la presente solicitud, el término "homólogo", en el caso de que se modifique con un adverbio tal como "altamente", puede referirse a la similitud de secuencias y no a un origen evolutivo común.

De acuerdo con lo anteriormente expuesto, la expresión "similitud de secuencias" en todas sus formas gramaticales se refiere al grado de identidad o correspondencia entre secuencias de ácidos nucleicos o de aminoácidos de proteínas que pueden compartir o no un origen evolutivo común (ver Reeck et al., Cell 50:667).

En una forma de realización específica, dos secuencias de ADN son "sustancialmente homólogas" o "sustancialmente similares" en el caso de que por lo menos aproximadamente el 50 % (Preferiblemente por lo menos aproximadamente el 75 %, y todavía más Preferiblemente por lo menos aproximadamente el 90 % ó el 95 %) de los nucleótidos son correspondientes a lo largo de la longitud definida de las secuencias de ADN. Las secuencias que son sustancialmente homólogas pueden identificarse mediante la comparación de las secuencias usando programas informáticos estándares disponibles en bancos de datos de secuencias, o en un experimento de hibridación Southern bajo, por ejemplo, condiciones astringentes definidas para ese sistema particular. La definición de las condiciones apropiadas de hibridación se encuentra comprendida dentro de los conocimientos del experto en la materia (ver, por ejemplo, Sambrook et al., supra, 1989).

Tal como se usa en la presente memoria, la expresión "sustancialmente similar" se refiere a fragmentos de ácidos nucleicos en los que los cambios en una o más bases nucleotídicas resulta en la sustitución de uno o más aminoácidos, pero que no afectan a las propiedades funcionales de la proteína codificada por la secuencia de ADN. La expresión "sustancialmente similar" también se refiere a fragmentos de ácidos nucleicos en los que los cambios de una o más bases nucleotídicas no afecta a la capacidad del fragmento de ácidos nucleicos de mediar en la alteración de la expresión génica mediante tecnología antisentido o de cosupresión. La expresión "sustancialmente similar" también se refiere a modificaciones de los fragmentos de ácidos nucleicos de la presente invención, tales como la deleción o la inserción de una o más bases nucleotídicas que no afecta sustancialmente a las propiedades funcionales del transcrito resultante. Por lo tanto, se entiende que la invención comprende más que las secuencias ejemplares específicas. Cada una de las modificaciones propuestas se encuentra perfectamente comprendida dentro de los conocimientos del experto rutinario en la materia, al igual que la determinación de la conservación de la actividad biológica de los productos codificados.

Además, el experto en la materia reconoce que secuencias sustancialmente similares comprendidas en la presente invención también se definen a partir de su capacidad para hibridarse, bajo condiciones astringentes (0,1X SSC, SDS al 0,1 %, 65ºC, y lavado con 2X SSC, SDS al 0,1 % seguido de 0,1X SSC, SDS al 0,1 %) con las secuencias ejemplificadas en la presente memoria. Los fragmentos de ácidos nucleicos sustancialmente similares de la descripción son aquellos fragmentos de ácidos nucleicos cuyas secuencias de ADN son por lo menos el 70 % idénticas a la secuencia de ADN de los fragmentos de ácidos nucleicos indicados en la presente memoria. Los fragmentos de ácidos nucleicos sustancialmente similares más preferidos son aquellos fragmentos de ácidos nucleicos cuyas secuencias de ADN son por lo menos el 80 % idénticas a la secuencia de ADN de los fragmentos de ácidos nucleicos indicados en la presente memoria. Los fragmentos de ácidos nucleicos aún más preferidos son por lo menos el 90 % idénticos a la secuencia de ADN de los fragmentos de ácidos nucleicos indicados en la presente memoria. Resultan todavía más preferidos los fragmentos de ácidos nucleicos que son por lo menos el 95 % idénticos a la secuencia de ADN de los fragmentos de ácidos nucleicos indicados en la presente memoria.

Dos secuencias de aminoácidos son "sustancialmente homólogas" o "sustancialmente similares" en el caso de que más de aproximadamente 40 % de los aminoácidos sean idénticos, o más de 60 % sean similares (funcionalmente idénticos). Preferiblemente, las secuencias similares u homólogas se identifican mediante alineamiento usando, por ejemplo, el programa Pileup del GCG (Genetics Computer Group, Program Manual for the GCG Package, versión 7, Madison, Wisconsin).

La expresión "correspondiente a" se usa en la presente memoria para referirse a secuencias similares u homólogas, sea o no la posición exacta idéntica o diferente de la molécula para la que se mide la similitud o la homología. Una alineación de secuencias de ácidos nucleicos o de aminoácidos puede incluir espacios. De esta manera, el término "correspondiente a" se refiere a la similitud de secuencias, y no a la numeración de los residuos aminoácidos o bases nucleotídicas.

Una "parte sustancial" de una secuencia de aminoácidos o de nucleótidos comprende una parte suficiente de la secuencia de aminoácidos de un polipéptido o de la secuencia de nucleótidos de un gen para identificar putativamente el polipéptido o gen, mediante evaluación manual de la secuencia por parte del experto en la materia, o mediante la comparación automática computerizada de las secuencias y la identificación usando algoritmos tales como BLAST (Basic Local Alignment Search Tool; Altschul S.F. et al., J. Mol. Biol. 215:403-410; ver también www.ncbi.nlm.nih.gov/BLAST/). En general, una secuencia de diez o más aminoácidos contiguos o de treinta o más

nucleótidos resulta necesaria para identificar putativamente una secuencia polipeptídica o de ácidos nucleicos como homóloga respecto a una proteína o gen conocido. Además, con respecto a las secuencias de nucleótidos, pueden usarse sondas oligonucleótidas específicas de un gen que comprenden 20 a 30 nucleótidos contiguos, en procedimientos dependientes de secuencia para la identificación (por ejemplo la hibridación Southern) y aislamiento de genes (por ejemplo la hibridación in situ de colonias bacterianas o placas de bacteriófagos). Además, pueden usarse oligonucleótidos cortos de 12 a 15 bases a modo de cebadores de amplificación en la PCR con el fin de obtener un fragmento de ácidos nucleicos particular que comprende los cebadores. Por consiguiente, una "parte sustancial" de una secuencia de nucleótidos comprende una parte suficiente de la secuencia para identificar y/o aislar específicamente un fragmento de ácidos nucleicos que comprende la secuencia.

La expresión "porcentaje de identidad", tal como se conoce en la técnica, es una relación entre dos o más secuencias polipeptídicas o entre dos o más secuencias polinucleótidas, determinada mediante la comparación de las secuencias. En la técnica, "identidad" también se refiere al grado de relación entre secuencias de polipéptido o polinucleótida, según sea el caso, determinado como la correspondencia entre cadenas de dichas secuencias. La "identidad" o "similitud" puede calcularse fácilmente mediante procedimientos conocidos, incluyendo, aunque sin limitación, los descritos en: Computational Molecular Biology (Lesk A.M., editor), Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects (Smith D.W., editor), Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I (Griffin A.M. y Griffin H.G., editores), Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology (von Heinje G., editor), Academic Press, 1987; y Sequence Analysis Primer (Gribskov M. y Devereux J., editores), Stockton Press, New York, 1991. Los procedimientos preferidos para determinar la identidad están diseñados para proporcionar la correspondencia óptima entre las secuencias sometidas a ensayo. Los procedimientos para determinar la identidad y similitud se encuentran codificados en programas informáticos públicamente disponibles. Pueden llevarse a cabo cálculos de alineación de secuencias y de porcentaje de identidad usando el programa Megalign del conjunto de análisis bioinformático LASERGENE (DNASTAR Inc., Madison, WI). Pueden llevarse a cabo alineaciones múltiples de las secuencias usando el procedimiento Clustal de alineación (Higgins y Sharp, CABIOS 5:151-153, 1989) con los parámetros por defecto (PENALIZACIÓN POR HUECO=10, PENALIZACIÓN POR LONGITUD DE HUECO=10). Pueden seleccionarse los parámetros por defecto para las alineaciones por parejas usando el procedimiento Clustal: KTUPLO 1, PENALIZACIÓN POR HUECO=3, VENTANA=5 y DIAGONALES SALVADAS=5.

La expresión "programas de análisis de secuencias" se refiere a cualquier algoritmo de ordenador o programa informático que resulte útil para el análisis de secuencias de nucleótidos o de aminoácidos. Los "programas de análisis de secuencias" pueden encontrarse comercialmente disponibles o desarrollarse independientemente. Los programas de análisis de secuencias típicos incluirán, aunque sin limitación, el conjunto GCG de programas (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, USA). Dentro del contexto de la presente solicitud se entenderá que, en el caso de que se utilicen programas de análisis de secuencias para el análisis, los resultados del análisis se basarán en los "valores por defecto" del programa referenciado, a menos que se indique lo contrario. Tal como se usa en la presente memoria, los "valores por defecto" se refiere a cualquier conjunto de valores o parámetros que se cargan originalmente con los programas al inicializarlos.

Los "genes sintéticos" pueden construirse a partir de bloques constructivos de oligonucleótidos que se sintetizan químicamente usando procedimientos conocidos por el experto en la materia. Estos bloques constructivos se ligan y se hibridan formando segmentos génicos que seguidamente se ensamblan enzimáticamente para construir el gen completo. La expresión "sintetizado químicamente", referida a una secuencia de ADN, se refiere a que los nucleótidos componentes han sido ensamblados in vitro. La síntesis química manual de ADN puede llevarse a cabo usando procedimientos bien establecidos, o la síntesis química automática puede llevarse a cabo usando uno de entre varios aparatos disponibles comercialmente. De acuerdo con lo anterior, los genes pueden adaptarse para la expresión génica óptima basándose en la optimización de la secuencia de nucleótidos de manera que refleje el sesgo en el uso de codones de la célula huésped. El experto en la materia apreciará la probabilidad de que la expresión génica se produzca con éxito si el uso de codones se encuentra sesgado hacia aquellos codones favorecidos por el huésped. La determinación de los codones preferentes puede basarse en un estudio de los genes derivados de la célula huésped de la que se disponga de información de secuencia.

Tal como se usa en la presente memoria, dos o más sistemas de regulación génica individualmente operables son "ortogonales" en el caso de que: a) la modulación de cada uno de los sistemas dados por su ligando respectivo, a una concentración seleccionada, resulte en un cambio medible de la magnitud de expresión del gen de ese sistema, y b) el cambio es diferente en grado estadísticamente significativo respecto al cambio de expresión de todos los demás sistemas operables simultáneamente en la célula, el tejido o el organismo, con independencia de la simultaneidad o secuencialidad de la modulación producida. Preferiblemente, la modulación de cada sistema de regulación génica individualmente operable lleva a cabo un cambio en la expresión génica por lo menos 2 veces superior que todos los demás sistemas operables en la célula, tejido u organismo. Más preferiblemente, el cambio es por lo menos 5 veces superior. Todavía más preferiblemente, el cambio es por lo menos 10 veces superior. Todavía más preferiblemente, el cambio es por lo menos 100 veces superior. Todavía más preferiblemente, el cambio es por lo menos 500 veces superior. Idealmente, la modulación de cada uno de los sistemas dados por parte de su ligando respectivo a una concentración seleccionada resulta en un cambio medible de la magnitud de expresión del gen de

ese sistema y ningún cambio medible de la expresión de todos los demás sistemas operables en la célula, tejido u organismo. En estos casos, se dice que el sistema de regulación de múltiples genes inducibles es "totalmente ortogonal. La presente invención es útil para buscar ligandos ortogonales y sistemas de expresión génica basados en receptores ortogonales, tales como los descritos en la solicitud copendiente nº US60/237.446.

SISTEMA DE MODULACIÓN DE LA EXPRESIÓN GÉNICA DE LA INVENCIÓN

Por lo tanto, los solicitantes han identificado en la presente memoria residuos aminoácidos que afectan a la sensibilidad del ligando y a la magnitud de la inducción en un sistema de expresión génica inducible basado en el grupo B. Los solicitantes describen en la presente memoria la construcción de receptores nucleares de grupo B que comprenden mutaciones por sustitución (denominadas en la presente memoria "mutantes por sustitución") en estos residuos críticos y la demostración de que estos receptores nucleares mutantes por sustitución son útiles en los procedimientos de modulación de la expresión génica. Como se presenta en la presente memoria, los receptores nucleares mutantes por sustitución nuevos y su uso en un sistema de expresión génica inducible basado en un receptor nuclear proporciona un sistema de expresión génica inducible mejorado tanto en células procariotas como eucariotas, en el cual la sensibilidad del ligando y la magnitud de la transactivación se pueden seleccionar como se desee, dependiendo de la aplicación.

Por lo tanto, la presente invención se refiere a nuevos polinucleótidos y polipéptidos de receptor nuclear de grupo B mutantes por sustitución, a un sistema de expresión génica inducible basado en un receptor nuclear que comprende dichos polinucleótidos y polipéptidos del receptor nuclear de grupo B mutados y a procedimientos de modulación de la expresión de un gen dentro de una célula huésped usando dicho sistema de expresión génica inducible basado en un receptor nuclear.

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 es capaz de expresarse en una célula huésped que comprende un polinucleótido que codifica un polipéptido que comprende un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución. Preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución es del receptor X del retinoide α, receptor X del retinoide β, receptor X del retinoide γ, proteína de unión II a la región H2 (H-2RIIBP), corregulador 1 de receptor nuclear (RCoR-1), proteína Ultraespiráculo (USP), receptor nuclear 2C1 y factor coriónico 1 (CF-1). Más preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución es de un receptor X del retinoide α de vertebrados, receptor X del retinoide β de vertebrados, receptor X del retinoide γ de vertebrados o receptor X del retinoide de invertebrados.

[122] En una forma de realización específica, el sistema de modulación de la expresión génica comprende a) un primer casete de expresión génica que comprende un polinucleótido que codifica un primer 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 debe modularse y un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución y b) un segundo casete de expresión génica que comprende un polinucleótido que codifica un segundo polipéptido que comprende un dominio de unión a ligando de receptor nuclear. 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; ii) un promotor que es activado por el dominio de transactivación del primer polipéptido y iii) un gen cuya expresión debe modularse.

[123] En una forma de realización preferida, el segundo polipéptido comprende un dominio de unión a ligando de receptor nuclear de grupo H seleccionado del grupo que consiste en los dominios de unión a ligando de un receptor de ecdisona (EcR), un receptor ubicuo (UR), un receptor huérfano 1 (OR-1), un NER-1, una proteína 15 de interacción con receptor (RIP-15), el receptor X hepático β (LXRβ), la proteína de tipo receptor de hormona esteroidea (RLD-1), el receptor hepático X (LXR), el receptor X hepático α (LXRα), el receptor X de farnesoide (FXR), la proteína de interacción con receptores 14 (RIP-14) y el receptor de farnesol (HRR-1).

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 un primer polipéptido que comprende un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión debe modularse y un primer dominio de unión a ligando de receptor nuclear y un segundo casete de expresión génica que comprende un polinucleótido que codifica un segundo polipéptido que comprende un dominio de transactivación y un segundo dominio de unión a ligando de receptor nuclear, en donde uno de los dominios de unión a ligando de receptor nuclear es un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución. En una forma de 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 los fines de la invención, "substancialmente libre" significa que la proteína en cuestión no contiene una secuencia suficiente del dominio en cuestión para proporcionar la activación o la 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 es activado por el dominio de transactivación del segundo polipéptido del segundo casete

de expresión génica y iii) un gen cuya expresión debe modularse.

En una forma de realización preferida, cuando un dominio de unión a ligando de receptor nuclear es un dominio de unión a ligando de grupo B que comprende una mutación por sustitución, el otro dominio de unión a ligando de receptor nuclear puede ser cualquier otro receptor nuclear que forme un dímero con el dominio de unión a ligando de grupo B que comprende la mutación por sustitución. En una forma de realización específica, el dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución es un dominio de unión a ligando de receptor X del retinoide que comprende una mutación por sustitución y el otro dominio de unión a ligando de receptor nuclear ("pareja") es de un receptor nuclear de grupo H. En una forma de realización preferida, el dominio de unión a ligando de receptor nuclear de grupo H se selecciona del grupo que consiste en los dominios de unión a ligando de un receptor de ecdisona (EcR), un receptor ubicuo (UR), un receptor huérfano 1 (OR-1), un NER-1, una proteína 15 de interacción con receptor (RIP-15), el receptor X hepático β (LXRβ), la proteína de tipo receptor de hormona esteroidea (RLD-1), el receptor hepático X (LXR), el receptor X hepático α (LXRα), el receptor X de farnesoide (FXR), la proteína de interacción con receptores 14 (RIP-14) y el receptor de farnesol (HRR-1).

El dominio de unión a ligando (LBD) del receptor de ecdisona (EcR) puede ser un EcR de invertebrado seleccionado de entre el grupo constituido por un EcR de artrópodo. Preferiblemente, el EcR se selecciona del grupo que consiste en un EcR de lepidóptero, un EcR de díptero, un EcR de ortóptero, un EcR de homóptero y un EcR de hemíptero. Más preferiblemente, el dominio de unión a ligando de EcR para el uso en la presente invención procede del EcR del gusano del abeto *Choristoneura fumiferana* ("CfEcR"), un EcR del escarabajo amarillo de la harina, *Tenebrio molitor* ("TmEcR"),un EcR de la oruga del cuerno del tabaco *Manduca sexta* ("MsEcR"), un EcR del gusano del capullo del tabaco *Heliothies virescens* ("HvEcR"), un EcR de mosquito quironómido *Chironomus tentans* ("CtEcR"), un EcR de gusano de la seda *Bombyx mori* ("BmEcR"), un EcR de mariposa *Bicyclus anynana* ("BanEcR"), un EcR de mariposa color castaño común *Junonia coenia* ("JcEcR"), un EcR de la mosca de la fruta *Drosophila melanogaster* ("DmEcR"), un EcR del mosquito de la fiebre amarilla *Aedes aegypti* ("AeEcR"), un EcR de la moscarda *Lucilia capitata* ("LcEcR"), un EcR de la moscarda de las ovejas *Lucilia cuprina* ("LucEcR"), un EcR de la moscarda azul *Calliphora vicinia* ("CvEcR"), un EcR de la 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 *Celuca pugilator* ("CpEcR"), un EcR de mosca blanca *Bamecia argentifoli* ("BaEcR", SEC ID nº 74), un EcR de saltahojas verde *Nephotetix cincticeps* ("NcEcR", SEC ID nº 75) o un EcR de garrapata ixódida *Amblyomma americanum* ("AmaEcR"). Más Preferiblemente, el LBD es de un CfEcR, un DmEcR o un AmaEcR.

El dominio de unión a ligando de receptor nuclear de "pareja" puede comprender además un mutación por truncado, una mutación por deleción, una segunda mutación por sustitución u otra modificación

En una forma de realización específica, el LBD es de un polipéptido EcR truncado. El truncado del polipéptido EcR tiene como resultado una deleción de por lo 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 o 265 aminoácidos. Preferiblemente, el truncado del polipéptido EcR tiene como resultado una deleción de por lo menos un dominio parcial del polipéptido. Más preferiblemente, el truncado del polipéptido EcR tiene como resultado una deleción de por lo menos todo el dominio del polipéptido. En una forma de realización específica, el truncado del polipéptido EcR tiene como resultado una deleción de por lo menos un dominio A/B, un dominio C, un dominio D, un dominio F, un dominio AB/C, un dominio A/B/1/2-C, un dominio A/B/C/D, un dominio A/B/C/D/F, un dominio A/B/F, un dominio A/B/C/F un dominio E parcial, o un dominio F parcial. También se puede realizar una combinación de varias deleciones completas y/o parciales del dominio.

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

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

En un ejemplo específico, la unión del ligando al dominio de unión a ligando de un receptor nuclear grupo B y, opcionalmente, su dominio de unión a ligando de receptor nuclear de pareja permite la expresión o la inhibición del gen. Este mecanismo no excluye la posibilidad de que el ligando se una al receptor nuclear grupo B (GBNR) o su pareja y la formación resultante de complejos homodímeros activos (por ej. GBNR + GBNR o pareja + pareja). Preferiblemente, se modifica uno o más de los dominios de receptor, produciendo un interruptor génico híbrido. Normalmente, uno o más de los tres dominios, un dominio de unión a ADN (DBD), un dominio de unión a ligando (LBD) y un dominio de transactivación (AD), pueden seleccionarse de una fuente diferente de la fuente de los demás dominios, de manera que los genes híbridos y las proteínas híbridas resultantes se optimicen en la célula u organismo huésped seleccionado para su actividad transactivadora, la unión complementaria del ligando y el

reconocimiento de un elemento de respuesta específico. Además, el elemento de respuesta mismo 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 (ver Sadowski et al. (1988) Nature, 335:563-564) o la proteína LexA de *Escherichia coli* (ver Brent y Ptashne (1985), Cell 43:729-736) o los elementos de respuesta sintéticos específicos para interacciones diana con proteínas diseñadas, modificadas y seleccionadas para dichas interacciones específicas (ver, por ejemplo, Kim et al. (1997), Proc. Natl. Acad. Sci., USA 94:3616-3620) que incluyan receptores híbridos. Otra ventaja de los sistemas de dos híbridos es que permiten la elección de un promotor usado para controlar la expresión génica según un resultado final deseado. Este control doble puede resultar particularmente importante en áreas de terapia génica, especialmente en el caso de que se produzcan proteínas citotóxicas, debido a que puede controlarse tanto la temporización de la expresión como las células en que se produce la expresión. Al introducir genes, ligados funcionalmente a un promotor adecuado, en las células del sujeto, se controla la expresión de los genes exógenos a partir de la presencia del sistema de la presente invención. Los promotores puede estar regulados constitutiva o inducciblemente o pueden ser específicos de tejido (es decir, expresados únicamente en un tipo particular de células) o ser específicos de determinadas etapas del desarrollo del organismo.

El receptor X del retinoide es un miembro de la superfamilia de receptores nucleares y se clasifica en la subfamilia 2, grupo B (denominado en la presente memoria "receptores nucleares de grupo B"). Los miembros de cada grupo comparten una identidad de aminoácidos del 40 % al 60 % en el dominio E (de unión a ligando (Laudet et al., A Unified Nomenclature System for the Nuclear Receptor Subfamily, Cell 97:161-163, 1999). Además del receptor X del retinoide, entre otros miembros de esta subfamilia 2, grupo B de receptores nucleares se incluyen: proteína de unión II a la región H2 (H-2RIIBP), corregulador 1 de receptor nuclear (RCoR-1), proteína Ultraespiráculo (USP), receptor nuclear 2C1 y factor coriónico 1 (CF-1).

Los solicitantes han demostrado previamente que un RXR de vertebrado en pareja con un sistema de expresión génica basado en un receptor de ecdisona proporciona un sistema de expresión génica inducible en levaduras y células de mamífero que se caracteriza por una mayor sensibilidad del ligando y de la magnitud de la transactivación (véase solicitud copendiente PCT/US01/09050). Recientemente, los solicitantes han demostrado que un RXR de invertebrado puede funcionar también o mejor que un RXR de vertebrado en un sistema de expresión génica basado en un receptor de ecdisona aumentando la transactivación génica y la sensibilidad del ligando del sistema de expresión génica (véase solicitud copendiente US 60/294.814).

Como se describe en la presente memoria, los solicitantes han identificado ahora residuos aminoácidos críticos dentro del dominio de unión a ligando de los RXR que afectan a la transactivación y a la sensibilidad del ligando de un sistema de expresión basado en un receptor nuclear. En los Ejemplos 2-4 a continuación, los solicitantes han identificado aminoácidos de dominios de unión a ligando de receptores X del retinoide de invertebrados (ver Ejemplos 2 y 3) y proteínas Ultraspiráculo de invertebrados (ver Ejemplo 4) que difieren de los aminoácidos de los RXR de vertebrados. Los solicitantes han fabricado mutantes por sustitución dentro de los dominios de unión a ligando de RXR de vertebrados sustituyendo el aminoácido RXR de vertebrado de tipo silvestre por el RXR o USP de invertebrado en sus posición análoga y han ensayado estos mutantes por sustitución para determinar su capacidad de transactivar la expresión génica en un sistema de expresión génica inducible basado en un receptor nuclear. Como se presenta en los Ejemplos de la presente memoria, los solicitantes han identificado ahora varios mutantes por sustitución del dominio de unión a ligando de receptor X del retinoide nuevos que presentan unos niveles de actividad de transactivación y/o sensibilidad del ligando inesperados y sorprendentes.

Dada la estrecha relación entre el receptor X del retinoide y otros receptores nucleares grupo B, se espera que las mutaciones por sustitución del dominio de unión a ligando de receptor X del retinoide identificadas por los solicitantes actúen cuando se introduzcan en la posición análoga de los dominios de unión al ligando de los otros receptores nucleares de grupo B para modificar la unión del ligando o la sensibilidad del ligando en un sistema de expresión génica basado en un receptor nuclear de grupo B. Los polinucléotidos y polipéptidos de dominio de unión a ligando de receptor nuclear grupo B mutados nuevos de los solicitantes son útiles en un sistema de modulación génica inducible basado en un receptor nuclear 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 la presente memoria un novedoso sistema de modulación de la expresión génica que comprende un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución. Este sistema de expresión génica puede ser un sistema de expresión génica basado en único interruptor" en el que el dominio de transactivación, el dominio de unión a ADN y el dominio de unión a ligando se encuentran 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 un "doble interruptor" o en "dos híbridos", en el que el dominio de transactivación y el dominio de unión a ADN se encuentran situados en dos polipéptidos codificados diferentes. Los solicitantes han demostrado por primera vez que se puede usar un receptor nuclear mutado por sustitución como un componente de un sistema de expresión génica inducible basado en un receptor nuclear para modificar la actividad de unión al ligando y/o la especificada del ligando tanto en células procariotas como eucariotas. Como se explica en la presente memoria, los hallazgos de los solicitantes son tanto inesperados como sorprendentes.

Preferiblemente el sistema de modulación de la expresión génica basado en un receptor nuclear de grupo B de la

presente invención puede ser heterodimérico u homodimérico. En una forma de realización preferida, el dominio de unión a ligando de receptor nuclear de grupo B se heterodimeriza con un dominio de unión a ligando de receptor de ecdisona para formar un complejo EcR funcional. Un complejo de EcR funcional generalmente se refiere a un complejo heterodimérico de proteínas que consiste en dos miembros de la familia de los receptores esteroides, una proteína de receptor de ecdisona obtenido de diversos insectos, y una proteína Ultraespiráculo (USP) o el homólogo de vertebrado de USP, la proteína receptor X del retinoide (ver 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 indica posteriormente. El complejo funcional de receptores ecdisteroides también puede incluir una o más proteínas adicionales, tales como inmunofilinas. Algunos miembros adicionales de la familia de proteínas receptores de esteroides, conocidos como factores transcripcionales (tales como DHR38 o *betaFTZ-1*), también pueden ser parejas dependientes o independientes de ligando para EcR, USP y/o RXR. Además, pueden resultar necesarios otros cofactores, tales como proteínas generalmente conocidas como coactivadores (también denominados adaptadores o mediadores). Estas proteínas no se unen de manera específica de secuencia a ADN y no se encuentran implicadas en la transcripción basal. Pueden ejercer su efecto sobre la activación de la transcripción a través de diversos mecanismos, incluyendo la estimulación de la unión de ADN de los activadores, los efectos sobre la estructura de la cromatina o la mediación en las interacciones de activador-complejo de inicio. Entre los ejemplos de dichos coactivadores se 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 una revisión ver Glass et al., Curr. Opin. Cell Biol. 9:222-232, 1997). Además, los cofactores proteicos generalmente conocidos como correpresores (también conocidos como represores, silenciadores o mediadores silenciadores) pueden resultar necesarios para inhibir efectivamente la activación transcripcional en ausencia de ligando. Estos correpresores pueden interactuar con el receptor de ecdisona sin ligando silenciando la actividad en el elemento de respuesta. Los datos disponibles en la actualidad sugieren que la unión del ligando modifica la conformación del receptor, resultando en la liberación del correpresor y en el reclutamiento de los coactivadores anteriormente indicados, eliminando de esta manera su actividad silenciadora. Entre los ejemplos de correpresores se incluyen N-CoR y SMRT (para una revisión, ver Horwitz et al., Mol. Endocrinol. 10:1167-1177, 1996). Estos cofactores pueden ser endógenos de la célula u organismo, o pueden añadirse exógenamente en forma de transgenes que se expresan de un modo regulado o no regulado. Los complejos homodímeros de la proteína de receptor de ecdisona, USP o RXR también pueden ser funcionales bajo algunas circunstancias.

El complejo de receptor de ecdisona normalmente incluye proteínas que son miembros de la superfamilia de los receptores nucleares, en la que todos los miembros se caracterizan generalmente 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 por una región bisagra. Tal como se usa en la presente memoria, la expresión "dominio de unión a ADN" comprende una secuencia polipeptídica mínima de una proteína de unión a ADN, e incluso hasta la longitud completa de una proteína de unión a ADN, con la condición de que el dominio de unión a ADN funcione asociándose a 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, de F (ver la patente US nº 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 presentar 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 cinc-cisteína entre los que se encuentran dos motivos aminoácidos, la caja P y la caja D, que proporcionan especificidad a los elementos de respuesta a la ecdisona. Estos dominios pueden ser de origen natural, modificados o quimeras de diferentes dominios de proteínas receptores heterólogas. El receptor EcR, al igual que un subconjunto de la familia de receptores de esteroides, también presenta regiones peor definidas que son responsables de las propiedades de heterodimerización. Debido a que los dominios de receptores nucleares son de naturaleza modular, los dominios LBD, DBD y de transactivación son intercambiables.

Se sabe que los sistemas de interruptor génico incorporan componentes del complejo del receptor de ecdisona. Sin embargo, en estos sistemas conocidos, siempre que se use EcR, este está asociado a dominios de unión a ADN naturales o modificados y los dominios de transactivación en la misma molécula. USP o RXR se usan normalmente como parejas silenciosas. Los solicitantes sabían 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 elevada y que dicha actividad está drásticamente reducida cuando los dominios de unión a ADN y los dominios de transactivación están en diferentes moléculas, es decir, en cada una de las dos parejas de un complejo heterodímero u homodímero (ver solicitud PCT/US01/09050).

CASETES DE EXPRESIÓN GÉNICA DE LA INVENCIÓN

El nuevo sistema de expresión génica inducible basado en receptores nucleares de la invención comprende por lo menos un casete de expresión génica que es capaz de expresarse en una célula huésped, en la que los casetes de expresión génica comprenden un polinucleótido que codifica un polipéptido que comprende un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución. De esta manera, la invención de los solicitantes también proporciona casetes de expresión génica nuevos para el uso en el sistema de expresión

génica de la invención.

En una forma de realización específica, el casete de expresión génica que es capaz de expresarse en una célula huésped comprende un polinucleótido que codifica un polipéptido seleccionado de entre el grupo constituido por: 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 B 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 B, y c) un polipéptido que comprende un dominio de transactivación y un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución.

En otra forma de realización específica, la presente descripción proporciona un casete de expresión génica que es capaz de expresarse en una célula huésped, en el que el casete de expresión génica comprende un polinucleótido que codifica un polipéptido híbrido seleccionado de entre el grupo constituido por: 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 B 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 B 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 B que comprende una mutación por sustitución. Un polipéptido híbrido de acuerdo con la invención comprende por lo menos dos fragmentos polipeptídicos, en el que cada fragmento polipeptídico procede de una fuente diferente, es decir, un receptor nuclear diferente, una especie diferente, etc. El polipéptido híbrido de acuerdo con la invención puede comprender por lo menos dos dominios polipeptídicos, en el que cada dominio polipeptídico procede de una fuente diferente.

En una realización específica, el dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución es de un receptor X del retinoide α (RXRα), receptor X del retinoide β (RXRβ), receptor X del retinoide γ (RXRγ), proteína de unión II a la región H2 (H-2RIIBP), corregulador 1 de receptor nuclear (RCoR-1), proteína Ultraespiráculo (USP), receptor nuclear 2C1 o factor coriónico 1 (CF-1). En una forma de realización preferida, el dominio de unión a ligando de receptor nuclear de grupo B es de un receptor X del retinoide α de vertebrados, receptor X del retinoide β de vertebrados, receptor X del retinoide γ de vertebrados o receptor X del retinoide de invertebrados.

Por lo tanto, la presente invención también proporciona un casete de expresión génica que comprende un polinucleótido que codifica 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 X del retinoide 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 X del retinoide 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 X del retinoide que comprende una mutación por sustitución. Preferiblemente, el casete de expresión génica comprende un polinucleótido que codifica 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 X del retinoide 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 X del retinoide 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 X del retinoide que comprende una mutación por sustitución; en donde el polipéptido híbrido codificado comprende por lo menos dos fragmentos polipeptídicos, en donde cada fragmento polipeptídico procede de una fuente diferente.

El dominio de unión a ligando (LBD) de receptor X (RXR) puede ser de un RXR de invertebrado o vertebrado. Preferiblemente, el dominio de unión a ligando de RXR para su uso en la presente invención es de ser humano *Homo sapiens* (HsRXR), de ratón *Mus musculus* (MmRXR), de rata *Rattus norvegicus* (RnRXR), de pollo *Gallus gallus* (GgRXR), de cerdo *Sus scrofa domestica* (SsRXR), de rana *Xenopus laevis* (XlRXR), de pez cebra *Danio rerio* (DrRXR), un homólogo de RXR de escarabajo *Tenebrio molitor* ("TmRXR"), un homólogo USP/RXR de la langosta *Locusta migratoria* (denominado "LmUSP" o "LmRXR"), un homólogo de RXR de áfido *Myzus persicae* ("MpRXR"), un homólogo de RXR de abeja melífera *Apis mellifera* ("AmRXR"), un homólogo de RXR de cangrejo violinista *Celuca pugilator* ("CpRXR"), un homólogo 1 de RXR de garrapata ixódida *Amblyomma americanum* ("AmaRXR1"), un homólogo 2 de RXR de garrapata ixodida *Amblyomma americanum* ("AmaRXR2"), de un tunicado *Polyandrocarpa misakiensis* (PmRXR) o RXR de medusa *Tripedalia cysophora* (TcRXR). Más preferiblemente, el LBD es de un MmRXR o un HsRXR.

En una forma de realización específica, el dominio de unión a ligando de receptor nuclear de grupo B está codificado por un polinucleótido que comprende una mutación de codón que tiene como resultado una sustitución de un residuo de aminoácido, en donde el residuo aminoácido está en una posición equivalente o análoga a la de un residuo aminoácido seleccionado del grupo que consiste en a) 401 ó 429 de la SEC ID nº 1, b) 401 y 429 de la SEC ID nº 1, c) 337, 344, 355, 385, 431, 442, 462, 470, 472, 473, 495, 500, 511, 516 ó 528 de la SEC ID nº 2, d) 321, 322 y 323 de la SEC ID nº 2, e) 450, 451 y 452 de la SEC ID nº 2, f) 455, 456, 457 y 458 de la SEC ID nº 2, g) 470, 472 y 473 de la SEC ID nº 2, h) 475, 476, 477, 478 y 479 de la SEC ID nº 2 e i) 481, 482 y 483 de la SEC ID nº 2. En una forma de realización preferida, el dominio de unión a ligando de receptor nuclear de grupo B es del receptor X del retinoide.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo B está codificado por un polinucleótido que comprende una mutación de codón que tiene como resultado una sustitución de a) un residuo ácido aspártico en una posición equivalente o análoga a un residuo aminoácido 401 de la SEC ID nº 1, b) un residuo serina en una posición equivalente o análoga a un residuo aminoácido 429 de la SEC ID nº 1, c) un residuo ácido aspártico en una posición equivalente o análoga a un residuo aminoácido 401 de la SEC ID nº 1 y un residuo serina en una posición equivalente o análoga al aminoácido del residuo aminoácido 429 de la SEC ID nº 1, d) un residuo serina en una posición equivalente o análoga a los residuos aminoácidos 337, 495, 500 ó 528 de la SEC ID nº 2, e) un residuo asparagina en una posición equivalente o análoga a un residuo aminoácido 344 de la SEC ID nº 2, f) un residuo arginina en una posición equivalente o análoga al aminoácido de los residuos aminoácidos 355 ó 511 de la SEC ID nº 2, g) un residuo alanina en una posición equivalente análoga a los residuos aminoácidos 385 ó 470 de la SEC ID nº 2, h) un residuo leucina en una posición equivalente o análoga a un residuo aminoácido 431 ó 462 de la SEC ID nº 2, i) un residuo lisina en una posición equivalente o análoga al aminoácido del residuo aminoácido 442 de la SEC ID nº 2, j) un residuo ácido aspártico en una posición equivalente o análoga a un residuo aminoácido 472 de la SEC ID nº 2, k) un ácido glutámico en una posición equivalente o análoga al aminoácido del residuo aminoácido 473 de la SEC ID nº 2, 1) un residuo valina en una posición equivalente o análoga al aminoácido del residuo aminoácido 516 de la SEC ID nº 2, m) un residuo leucina en una posición equivalente o análoga al aminoácido del residuo aminoácido 321 de la SEC ID nº 2, un residuo arginina en una posición equivalente o análoga al aminoácido del residuo aminoácido 322 de la SEC ID nº 2 y un residuo valina en una posición equivalente o análoga al aminoácido del residuo aminoácido 323 de la SEC ID nº 2, n), un residuo ácido glutámico en una posición equivalente o análoga al aminoácido del residuo aminoácido 450 de la SEC ID nº 2, un residuo valina en una posición equivalente o análoga al aminoácido del residuo aminoácido 451 de la SEC ID nº 2 y un residuo arginina en una posición equivalente o análoga al aminoácido del residuo aminoácido 452 de la SEC ID nº 2, o) un residuo lisina en una posición equivalente o análoga al aminoácido del residuo aminoácido 455 de la SEC ID nº 2, un residuo serina en una posición equivalente o análoga al aminoácido del residuo aminoácido 456 de la SEC ID nº 2, un residuo alanina en una posición equivalente o análoga al aminoácido del residuo aminoácido 457 de la SEC ID nº 2 y un residuo glutamina en una posición equivalente o análoga al aminoácido del residuo aminoácido 458 de la SEC ID nº 2, p) un residuo alanina en una posición equivalente o análoga al aminoácido del residuo aminoácido 470 de la SEC ID nº 2, un residuo ácido aspártico en una posición equivalente o análoga al aminoácido del residuo aminoácido 472 de la SEC ID nº 2 y un residuo tirosina en una posición equivalente o análoga al aminoácido del residuo aminoácido 473 de la SEC ID nº 2, q) un residuo treonina en una posición equivalente o análoga al aminoácido de los residuos aminoácidos 475, 477 y 478 de la SEC ID nº 2, un residuo arginina en una posición equivalente o análoga al aminoácido del residuo aminoácido 476 de la SEC ID nº 2 y un residuo histidina en una posición equivalente o análoga al aminoácido del residuo aminoácido 479 de la SEC ID nº 2 o r) un residuo ácido aspártico en una posición equivalente o análoga al aminoácido del residuo aminoácido 481 de la SEC ID nº 2, un residuo ácido glutámico en una posición equivalente o análoga al aminoácido del residuo aminoácido 482 de la SEC ID nº 2 y un residuo prolina en una posición equivalente o análoga al aminoácido del residuo aminoácido 483 de la SEC ID nº 2. En una forma de realización preferida, el dominio de unión a ligando de receptor nuclear de grupo B es del receptor X del retinoide.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución es un dominio de unión a ligando de receptor X del retinoide que comprende una mutación por sustitución codificado por un polinucleótido que comprende una mutación de codón o mutaciones de codón que tiene como resultado una mutación por sustitución seleccionada del grupo que consiste en a) E401D o G429S de la SEC ID nº 1, b) E401D y G429S de la SEC ID nº 1, c) T337S, D344N, K355R, S385A, M431L, R442K, V462L, S470A, E471D, T473E, A495S, E500S, K511R, T516V o A528S de la SEC ID nº 2, d) G321L, P322R y G323V de la SEC ID nº 2, e) D450E, A451V y K452R de la SEC ID nº 2, f) S455K, N456S, P457A y S458Q de la SEC ID nº 2, g) S470A, E472D y T473Y de la SEC ID nº 2, h) C475T, K476R, Q477T, K478T y Y479H de la SEC ID nº 2 e i) E481D, Q482E y Q483P de la SEC ID nº 2.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo B comprende una mutación por sustitución en una posición equivalente o análoga a un residuo aminoácido a) 401 ó 429 de la SEC ID nº 1, b) 401 y 429 de la SEC ID nº 1, c) 337, 344, 355, 385, 431, 442, 462, 470, 472, 473, 495, 500, 511, 516 ó 528 de la SEC ID nº 2, d) 321, 322 y 323 de la SEC ID nº 2, e) 450, 451 y 452 de la SEC ID nº 2, f) 455, 456,457 y 458 de la SEC ID nº 2, g) 470, 472 y 473 de la SEC ID nº 2, h) 475, 476, 477, 478 y 479 de la SEC ID nº 2 e i) 481, 482 y 483 de la SEC ID nº 2. En una forma de realización preferida, el dominio de unión a ligando de receptor nuclear de grupo B es del receptor X del retinoide.

Preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo B comprende una sustitución de a) un residuo ácido aspártico en una posición equivalente o análoga a un residuo aminoácido 401 de la SEC ID nº 1, b) un residuo serina en una posición equivalente o análoga a un residuo aminoácido 429 de la SEC ID nº 1, c) un residuo ácido aspártico en una posición equivalente o análoga a un residuo aminoácido 401 de la SEC ID nº 1 y un residuo serina en una posición equivalente o análoga al aminoácido del residuo aminoácido 429 de la SEC ID nº 1, d) un residuo serina en una posición equivalente análoga a los residuos aminoácidos 337, 495, 500 ó 528 de la SEC ID nº 2, e) un residuo asparagina en una posición equivalente o análoga a un residuo aminoácido 344 de la SEC ID nº 2, f) un residuo arginina en una posición equivalente o análoga al aminoácido de los residuos aminoácidos 355 ó 511 de la SEC ID nº 2, g) un residuo alanina en una posición equivalente análoga a los residuos aminoácidos 385 ó 470 de

la SEC ID nº 2, h) un residuo leucina en una posición equivalente o análoga a un residuo aminoácido 431 ó 462 de la SEC ID nº 2, i) un residuo lisina en una posición equivalente o análoga al aminoácido del residuo aminoácido 442 de la SEC ID nº 2, j) un residuo ácido aspártico en una posición equivalente o análoga a un residuo aminoácido 472 de la SEC ID nº 2, k) un ácido glutámico en una posición equivalente o análoga al aminoácido del residuo aminoácido 473 de la SEC ID nº 2,1) un residuo valina en una posición equivalente o análoga al aminoácido del residuo aminoácido 516 de la SEC ID nº 2, m) un residuo leucina en una posición equivalente o análoga al aminoácido del residuo aminoácido 321 de la SEC ID nº 2, un residuo arginina en una posición equivalente o análoga al aminoácido del residuo aminoácido 322 de la SEC ID nº 2 y un residuo valina en una posición equivalente o análoga al aminoácido del residuo aminoácido 323 de la SEC ID nº 2, n) un residuo ácido glutámico en una posición equivalente o análoga al aminoácido del residuo aminoácido 450 de la SEC ID nº 2, un residuo valina en una posición equivalente o análoga al aminoácido del residuo aminoácido 451 de la SEC ID nº 2 y un residuo arginina en una posición equivalente o análoga al aminoácido del residuo aminoácido 452 de la SEC ID nº 2, o) un residuo lisina en una posición equivalente o análoga al aminoácido del residuo aminoácido 455 de la SEC ID nº 2, un residuo serina en una posición equivalente o análoga al aminoácido del residuo aminoácido 456 de la SEC ID nº 2, un residuo alanina en una posición equivalente o análoga al aminoácido del residuo aminoácido 457 de la SEC ID nº 2 y un residuo glutamina en una posición equivalente o análoga al aminoácido del residuo aminoácido 458 de la SEC ID nº 2, p) un residuo alanina en una posición equivalente o análoga al aminoácido del residuo aminoácido 470 de la SEC ID nº 2, un residuo ácido aspártico en una posición equivalente o análoga al aminoácido del residuo aminoácido 472 de la SEC ID nº 2 y un residuo tirosina en una posición equivalente o análoga al aminoácido del residuo aminoácido 473 de la SEC ID nº 2, q) un residuo treonina en una posición equivalente o análoga al aminoácido de los residuos aminoácidos 475, 477 y 478 de la SEC ID nº 2, un residuo arginina en una posición equivalente o análoga al aminoácido del residuo aminoácido 476 de la SEC ID nº 2 y un residuo histidina en una posición equivalente o análoga al aminoácido del residuo aminoácido 479 de la SEC ID nº 2 o r) un residuo ácido aspártico en una posición equivalente o análoga al aminoácido del residuo aminoácido 481 de la SEC ID nº 2, un residuo ácido glutámico en una posición equivalente o análoga al aminoácido del residuo aminoácido 482 de la SEC ID nº 2 y un residuo prolina en una posición equivalente o análoga al aminoácido del residuo aminoácido 483 de la SEC ID nº 2. En una forma de realización preferida, el dominio de unión a ligando de receptor nuclear de grupo B es del receptor X del retinoide.

En otra realización específica, el dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución es un dominio de unión a ligando de receptor X del retinoide que comprende una mutación por sustitución, en donde la mutación por sustitución se selecciona del grupo que consiste en a) E401D o G429S de la SEC ID nº 1, b) E401D y G429S de la SEC ID nº 1, c) T337S, D344N, K355R, S385A, M431L, R442K, V462L, S470A, E471D, T473E, A495S, E500S, K511R, T516V o A528S de la SEC ID nº 2, d) G321L, P322R y G323V de la SEC ID nº 2, e) D450E, A451V y K452R de la SEC ID nº 2, f) S455K, N456S, P457A y S458Q de la SEC ID nº 2, g) S470A, E472D y T473Y de la SEC ID nº 2, h) C475T, K476R, Q477T, K478T y Y479H de la SEC ID nº 2 y i) E481D, Q482E y Q483P de la SEC ID nº 2.

El dominio de unión a ADN (DBD) 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 receptores nucleares de grupo B, un DBD de miembro de la superfamilia de los receptores nucleares de hormonas esteroideas/tiroideas, un DBD de LacZ bacteriano o un DBD. Más Preferiblemente, el DBD es un DBD de receptor de ecdisona (EcR) [SEC ID Nº 3 (polinucleótido) o SEC ID Nº 4 (polipéptido)], un DBD de GAL4 [SEC ID Nº 5 (polinucleótido) o SEC ID Nº 6 (polipéptido)] o un DBD de LexA [(SEC ID Nº 7 (polinucleótido) o SEC ID Nº 8 (polipéptido)].

El dominio de transactivación (abreviado "AD" o "TA") puede ser cualquier AD de miembro de receptores nucleares de grupo B, AD de receptores nucleares de hormona esteroidea/tiroidea, AD sintético o quimérico, AD poliglutamina, AD aminoácido básico o ácido, un AD de VP16, un AD de GAL4, un AD de NF-κB, un AD de BP64, un AD de dominio de activación ácido de B42 (B42AD), un dominio de transactivación de p65 (p65AD), o un análogo, combinación, o su modificación.

En una forma de realización particular, el AD es un AD sintético o quimérico o se obtiene de un receptor de ecdisona, un receptor glucocorticoide, VP16, GAL4, NF-κB o un AD de dominio de activación ácido de B42. Preferiblemente, el AD es un AD de EcR [SEC ID Nº 9 (polinucleótido) o SEC ID Nº 10 (polipéptido)], un AD de VP16 [SEC ID Nº 11 (polinucleótido) o SEC ID Nº 12 (polipéptido)], un AD de B42 [SEC ID Nº 13 (polinucleótido) o SEC ID Nº 14 (polipéptido)] o un AD dep65 [SEC ID Nº 15 (polinucleótido) o SEC ID Nº 16 (polipéptido)].

En una forma de realización específica, el casete de expresión génica codifica un polipéptido híbrido que comprende a) un dominio de unión a ADN codificado por un polinucleótido que comprende una secuencia de ácidos nucleicos de la SEC ID nº 3, SEC ID Nº 5 o SEC ID Nº 7; un dominio de transactivación codificado por un polinucleótido que comprende una secuencia de ácidos nucleicos de la SEC ID nº 9, SEC ID Nº 11, SEC ID Nº 13 o SEC ID Nº 15 y un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución codificada por un polinucleótido de acuerdo con la invención; b) un dominio de unión a ADN codificado por un polinucleótido que comprende una secuencia de ácidos nucleicos de la SEC ID nº 3, SEC ID Nº 5 o SEC ID Nº 7 y un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución codificada por un

polinucleótido de acuerdo con la invención o c) un dominio de transactivación codificado por un polinucleótido que comprende una secuencia de ácidos nucleicos de la SEC ID nº 9, SEC ID Nº 11, SEC ID Nº 13 o SEC ID Nº 15 y un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución codificada por un polinucleótido de acuerdo con la invención. Preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo B 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 de acuerdo con la invención.

En otra realización específica, el casete de expresión génica codifica un polipéptido híbrido que comprende a) un dominio de unión a ADN que comprende una secuencia de aminoácidos de la SEC ID nº 4, SEC ID Nº 6 o SEC ID Nº 8; un dominio de transactivación que comprende una secuencia de aminoácidos de la SEC ID nº 10, SEC ID Nº 12, SEC ID Nº 14 o SEC ID Nº 16 y dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución de acuerdo con la invención; b) un dominio de unión a ADN que comprende una secuencia de aminoácidos de la SEC ID nº 4, SEC ID Nº 6 o SEC ID Nº 8 y un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución de acuerdo con la invención o c) un dominio de transactivación que comprende una secuencia de aminoácidos de la SEC ID nº 10, SEC ID Nº 12, SEC ID Nº 14 o SEC ID Nº 16 y un dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución de acuerdo con la invención. Preferiblemente, el dominio de unión a ligando de receptor nuclear de grupo B que comprende una mutación por sustitución es un dominio de unión a ligando de receptor de ecdisona que comprende una mutación por sustitución de acuerdo con la invención.

El elemento de respuesta ("ER") puede ser cualquier elemento de respuesta con un dominio de unión a ADN conocido, o un análogo, combinación, o modificación del mismo. Puede emplearse un único ER o múltiples RE, pueden usarse múltiples copias del mismo ER o dos o más diferentes RE, en la presente invención. En una realización específica, el ER es un ER de GAL4 ("GAL4ER"), LexA, un ER del receptor nuclear del Grupo B, un ER del receptor nuclear de hormonas esteroides/tiroideas, o un ER sintético que reconoce un dominio de unión a ADN sintético. Preferiblemente, el ER es un ER de ecdisona (EcER) que comprende una secuencia polinucleotídica de la SEC ID Nº 17, un GAL4ER que comprende una secuencia polinucleotídica de la SEC ID Nº 18, o un LexA ER (operón, "op") que comprende una secuencia polinucleotídica de la SEC ID Nº 19 ("2XLexAopER").

Un dominio de unión a ADN, dominio de activación o elemento de respuesta del receptor de hormonas esteroides/tiroideas de acuerdo con la invención puede obtenerse de un receptor de hormonas esteroides/tiroideas seleccionado entre el grupo compuesto por el receptor de la hormona tiroidea α (TRα), el receptor tiroideo 1 (c-erbA-1), receptor de la hormona tiroidea β (TRβ), receptor de ácido retinoico α (RARα), receptor de ácido retinoico β (RARβ, HAP), receptor de ácido retinoico γ (RARγ), receptor de ácido retinoico tipo gamma (RARD), receptor activado por el proliferador de peroxisomas α (PPARα), receptor activado por el proliferador de peroxisomas β (PPARβ), receptor activado por el proliferador de peroxisomas δ (PPARδ, NUC-1), receptor relacionado con el activador proliferado de peroxisomas (FFAR), receptor activado por el proliferador de peroxisomas γ (PPARγ), receptor huérfano codificado por la hebra no codificante del receptor de la hormona tiroidea α (REVERBα), receptor relacionado con v-erb A (EAR-1), receptor relacionado con v-erb (EAR-1A), γ), receptor huérfano codificado por la hebra no codificante del receptor de la hormona tiroidea β (REVERBβ), receptor relacionado con v-erb (EAR-1β), 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 78 de *Drosophila* (DR-78), receptor huérfano relacionado con retinoide α (RORα), receptor Z retinoide α (RZRα), receptor huérfano relacionado con retinoide β (RORβ), receptor Z retinoide β (RZRβ), receptor huérfano relacionado con retinoide γ (RORγ), receptor Z retinoide γ (RZRγ), receptor huérfano relacionado con retinoide (TOR), receptor hormonal 3 (HR-3), receptor hormonal 3 de *Drosophila* (DHR-3), receptor hormonal de manduca (MHR-3), receptor hormonal 3 de *Galleria* (GHR-3), receptor nuclear 3 de *C. elegans* (CNR-3), receptor hormonal 3 de *Choristoneura* (CHR-3), receptor nuclear 14 de *C. elegans* (CNR-14), receptor de ecdisona (ECR), receptor ubicuo (UR), receptor nuclear huérfano (OR-1), NER-1, proteína de interacción con receptores 15 (RIP-15), receptor X hepático β (LXRβ), proteína tipo receptor de hormonas esteroideas (RLD-1), receptor X hepático (LXR), receptor X hepático α (LXRα), receptor X farnesoide (FXR), proteína de interacción con receptores 14 (RIP-14), HRR-1, receptor de vitamina D (VDR), receptor nuclear huérfano (ONR-1), receptor X de pregnano (PXR), receptor esteroide y xenobiótico (SXR), receptor X de benzoato (BXR), receptor nuclear (MB-67), receptor constitutivo de androstano 1 (CAR-1), receptor constitutivo de androstano α (CARα), receptor constitutivo de androstano 2 (CAR-2), receptor constitutivo de androstano β (CARβ), receptor hormonal 96 de *Drosophila* (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), receptor nuclear de hepatocitos 4D (HNF-4D, DHNF-4), receptor X retinoide α (RXRα), receptor X retinoide β (RXRβ), proteína de unión a la región H-2 II (H-2RIIBP), co-regulador-1 de receptor nuclear (RCoR-1), receptor X retinoide γ (RXRγ), Ultraespiráculo (USP), receptor nuclear 2C1, factor coriónico 1 (CF-1), receptor testicular 2 (TR2), 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 cadena arriba de la ovalbúmina de pollo I (COUP-TFI), factor de transcripción del promotor cadena arriba de la ovalbúmina de pollo A (COUP-TFA), EAR-3, SVP-44, factor de transcripción del promotor cadena arriba de la ovalbúmina de pollo II (COUP-TFII), factor de transcripción del promotor cadena arriba de la ovalbúmina de pollo B (COUP-TFB), ARP-1, SVP-40, SVP, factor de transcripción del promotor cadena arriba de la ovalbúmina de pollo III (COUP-TFIII),

factor de transcripción del promotor cadena arriba de la ovalbúmina de pollo G (COUP-TFG), SVP-46, EAR-2, receptor de estrógenos α (ERα), receptor de estrógenos β (ERβ), receptor relacionado con estrógenos 1 (ERR1), receptor relacionado con estrógenos α (ERRα), receptor relacionado con estrógenos 2 (ERR2), receptor relacionado con estrógenos β (ERRβ), 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 1 relacionado con Nurr (NURR-1), un gen humano de respuesta temprana inmediata (NOT), receptor nuclear hepático regenerante 1 (RNR-1), dedos de zinc hematopoyético 3 (HZF-3), proteína relacionada con Nur-1 (TINOR), receptor huérfano nuclear 1 (NOR-1), receptor relacionado con NOR1 (MINOR), receptor hormonal 38 de *Drosophila* (DHR-38), receptor nuclear 8 de *C. elegans* (CNR-8), C48D5, factor esteroidogénico 1 (SF1), péptido tipo endozepina (ELP), factor fushi tarazu 1 (FTZ-F1), proteína de unión adrenal 4 (AD4BP), homólogo del receptor hepático (LRH-1), receptor huérfano A relacionado con Ftz-F1 (xFFrA), receptor huérfano B relacionado con Ftz-F1 (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 retinoides (RTR), knirps (KNI), relacionado 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).

Por tanto, la presente invención también proporciona una 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 tiene que modularse.

Los genes de interés para su uso en los casetes de expresión génica de los solicitantes pueden ser genes endógenos o genes heterólogos. La información de secuencia de ácido nucleico o aminoácidos para un gen deseado o proteína puede localizarse en una de muchas bases de datos de acceso público, por ejemplo, GENBANK, EMBL, Swiss-Prot, y PIR, o en muchas publicaciones de revistas relacionadas con la biología. Por tanto, los especialistas en la técnica tienen acceso a información de secuencia de ácido nucleico para casi todos los genes conocidos. Dicha información entonces puede usarse para construir las construcciones deseadas para la inserción del gen de interés dentro de los casetes de expresión génica usados en los procedimientos de los solicitantes descritos en este documento.

Ejemplos de genes de interés para su uso en casetes de expresión génica de los solicitantes incluyen, aunque sin limitación: genes que codifican polipéptidos o productos terapéuticamente deseables que pueden usarse para tratar una afección, una enfermedad, un trastorno, una disfunción, un defecto genético, tales como anticuerpos monoclonales, enzimas, proteasas, citoquinas, interferones, insulina, eritropoyetina, factores de coagulación, otros factores o componentes sanguíneos, vectores virales para terapia génica, virus para vacunas, dianas para descubrimiento de fármacos, análisis y aplicaciones de genómica funcional y proteómica, y similares.

Para propósitos de esta invención, los receptores nucleares y los receptores nucleares de Grupo B también incluyen receptores nucleares sintéticos y quiméricos y receptores nucleares de Grupo B y sus homólogos.

Polinucleótidos de la invención

El nuevo sistema de expresión génica inducible basado en receptores nucleares de la invención comprende al menos un casete de expresión génica que comprende un polinucléotido que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución. Estos casetes de expresión génica, los polinucléotidos 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 invención también proporciona un polinucléotido aislado que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución.

En una realización específica, el dominio de unión al ligando del receptor nuclear de Grupo B está codificado por un polinucléotido que comprende una mutación de codón que provoca una sustitución de un resto aminoacídico, donde el resto aminoacídico está en una posición equivalente o análoga a un resto aminoacídico seleccionado entre el grupo compuesto por a) 401 ó 429 de la SEC ID Nº 1, b) 401 y 429 de la SEC ID Nº 1, c) 337, 344, 355, 385, 431, 442, 462, 470, 472, 473, 495, 500, 511, 516, ó 528 de la SEC ID Nº 2, d) 321, 322, y 323 de la SEC ID Nº 2, e) 450, 451, y 452 de la SEC ID Nº 2, f) 455, 456, 457, y 458 de la SEC ID Nº 2, g) 470, 472, y 473 de la SEC ID Nº 2, h) 475, 476, 477, 478, y 479 de la SEC ID Nº 2, e i) 481, 482, y 483 de la SEC ID Nº 2. En una realización preferida, el dominio de unión al ligando del receptor nuclear de Grupo B es de un receptor X retinoide.

En otra realización específica, el dominio de unión al ligando del receptor nuclear de Grupo B está codificado por un polinucléotido que comprende una mutación de codón que provoca una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de ácido aspártico en una

posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, y un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, d) un resto de serina en una posición equivalente o análoga a los restos aminoacídicos 337, 495, 500, ó 528 de la SEC ID Nº 2, e) un resto de asparagina en una posición equivalente o análoga al resto aminoacídico 344 de la SEC ID Nº 2, f) un resto de arginina en una posición equivalente o análoga a los restos aminoacídicos 355 ó 511 de la SEC ID Nº 2, g) un resto de alanina en una posición equivalente o análoga a los resto aminoacídicos 385 ó 470 de la SEC ID Nº 2, h) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 431 ó 462 de la SEC ID Nº 2, i) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 442 de la SEC ID Nº 2, j) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, k) un ácido glutámico en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, l) un resto de valina en una posición equivalente o análoga al resto aminoacídico 516 de la SEC ID Nº 2, m) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 321 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 322 de la SEC ID Nº 2, y un resto de valina en una posición equivalente o análoga al resto aminoacídico 323 de la SEC ID Nº 2, n) un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 450 de la SEC ID Nº 2, un resto de valina en una posición equivalente o análoga al resto aminoacídico 451 de la SEC ID Nº 2, y un resto de arginina en una posición equivalente o análoga al resto aminoacídico 452 de la SEC ID Nº 2, o) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 455 de la SEC ID Nº 2, un resto de serina en una posición equivalente o análoga al resto aminoacídico 456 de la SEC ID Nº 2, un resto de alanina en una posición equivalente o análoga al resto aminoacídico 457 de la SEC ID Nº 2, y un resto de glutamina en una posición equivalente o análoga al resto aminoacídico 458 de la SEC ID Nº 2, p) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 470 de la SEC ID Nº 2, un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, y un resto de tirosina en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, q) un resto de treonina en una posición equivalente o análoga a los restos aminoacídicos 475, 477, y 478 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 476 de la SEC ID Nº 2, y un resto de histidina en una posición equivalente o análoga al resto aminoacídico 479 de la SEC ID Nº 2, o r) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 481 de la SEC ID Nº 2, un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 482 de la SEC ID Nº 2, y un resto de prolina en una posición equivalente o análoga al resto aminoacídico 483 de la SEC ID Nº 2. En una realización preferida, el dominio de unión al ligando del receptor nuclear de Grupo B es de un receptor X retinoide.

En otra realización específica, el dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución es un dominio de unión al ligando del receptor X retinoide que comprende una mutación de sustitución codificada por un polinucléotido que comprende una mutación de codón o mutaciones de codón que provocan una mutación de sustitución seleccionad entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, b) E401D y G429S de la SEC ID Nº 1, c) T337S, D344N, K355R, S385A, M431L, R442K, V462L, S470A, E471D, T473E, A495S, E500S, K511R, T516V, o A528S de la SEC ID Nº 2, d) G321L, P322R, y G323V de la SEC ID Nº 2, e) D450E, A451V, y K452R de la SEC ID Nº 2, f) S455K, N456S, P457A, y S458Q de la SEC ID Nº 2, g) S470A, E472D, y T473Y de la SEC ID Nº 2, h) C475T, K476R, Q477T, K478T, e Y479H de la SEC ID Nº 2, e i) E481D, Q482E, y Q483P de la SEC ID Nº 2.

La presente invención también proporciona un polinucléotido aislado que codifica un polipéptido seleccionado entre el grupo compuesto por a) un polipéptido que comprende un dominio de transactivación, un dominio de unión a ADN, y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención; b) un polipéptido que comprende un dominio de unión a ADN y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención; y c) un polipéptido que comprende un dominio de transactivación y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención.

En una realización específica, el polinucléotido aislado codifica un polipéptido híbrido seleccionado entre el grupo compuesto por a) un polipéptido híbrido que comprende un dominio de transactivación, un dominio de unión a ADN, y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención; b) un polipéptido híbrido que comprende un dominio de unión a ADN y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención; y c) un polipéptido híbrido que comprende un dominio de transactivación y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención.

La presente invención también se refiere a un polinucléotido aislado que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución afecta a la actividad de unión al ligando o la sensibilidad de ligando del dominio de unión al ligando del receptor nuclear de Grupo B.

En particular, la presente invención se refiere a un polinucléotido aislado que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución reduce la actividad de unión al ligando o la sensibilidad de ligando del dominio de unión al ligando del receptor nuclear de Grupo B.

En una realización específica, la presente invención se refiere a un polinucléotido aislado que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución reduces la actividad de unión a esteroides o la sensibilidad de esteroide del dominio de unión al ligando del receptor nuclear de Grupo B. Preferiblemente, el polinucleótido aislado comprende una mutación de codón que provoca una sustitución de un resto aminoacídico, donde el resto aminoacídico está en una posición equivalente o análoga a a) el resto aminoacídico 401 ó 429 de la SEC ID Nº 1, o b) el resto aminoacídico 344, 431, 442, 495, 511, ó 528 de la SEC ID Nº 2. Más preferiblemente, el polinucléotido aislado comprende una mutación de codón que provoca una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de asparagina en una posición equivalente o análoga al resto aminoacídico 344 de la SEC ID Nº 2, d) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 431 de la SEC ID Nº 2, e) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 442 de la SEC ID Nº 2, f) un resto de serina en una posición equivalente o análoga al resto aminoacídico 495 ó 528 de la SEC ID Nº 2, o g) un resto de arginina en una posición equivalente o análoga al resto aminoacídico 511 de la SEC ID Nº 2. Incluso más preferiblemente, el polinucléotido aislado comprende una mutación de codón que provoca una mutación de sustitución seleccionada entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, y b) D344N, M431L, R442K, A495S, K511R, o A528S de la SEC ID Nº 2.

En una realización específica, la presente invención se refiere a un polinucléotido aislado que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución reduce la actividad de unión al ligando no esteroideo o sensibilidad de ligando del dominio de unión al ligando del receptor nuclear de Grupo B. Preferiblemente, el polinucléotido aislado comprende una mutación de codón que provoca una sustitución de un resto aminoacídico, donde el resto aminoacídico está en una posición equivalente o análoga a a) el resto aminoacídico 401 ó 429 de la SEC ID Nº 1, o b) el resto aminoacídico 344, 431, 442, 495, 511, ó 528 de la SEC ID Nº 2. Más preferiblemente, el polinucléotido aislado comprende una mutación de codón que provoca una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de asparagina en una posición equivalente o análoga al resto aminoacídico 344 de la SEC ID Nº 2, d) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 431 de la SEC ID Nº 2, e) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 442 de la SEC ID Nº 2, f) un resto de serina en una posición equivalente o análoga al resto aminoacídico 495 ó 528 de la SEC ID Nº 2, o g) un resto de arginina en una posición equivalente o análoga al resto aminoacídico 511 de la SEC ID Nº 2. Incluso más preferiblemente, el polinucléotido aislado comprende una mutación de codón que provoca una mutación de sustitución seleccionada entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, y b) D344N, M431L, R442K, A495S, K511R, o A528S de la SEC ID Nº 2.

Además, la presente invención también se refiere a un polinucléotido aislado que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución potencia la actividad de unión al ligando o sensibilidad de ligando del dominio de unión al ligando del receptor nuclear de Grupo B.

En una realización específica, la presente invención se refiere a un polinucléotido aislado que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución potencia la actividad de unión a esteroides o sensibilidad de esteroide del dominio de unión al ligando del receptor nuclear de Grupo B. Preferiblemente, el polinucléotido aislado comprende una mutación de codón que provoca una sustitución de un resto aminoacídico en una posición equivalente o análoga al resto aminoacídico seleccionado entre el grupo compuesto por a) 401 ó 429 de la SEC ID Nº 1, b) 401 y 429 de la SEC ID Nº 1, c) 337, 355, 385, 462, 470, 472, 473, ó 500 de la SEC ID Nº 2, d) 321, 322, y 323 de la SEC ID Nº 2, e) 450, 451, y 452 de la SEC ID Nº 2, f) 455, 456, 457, y 458 de la SEC ID Nº 2, g) 470, 472, y 473 de la SEC ID Nº 2, h) 475, 476, 477, 478, y 479 de la SEC ID Nº 2, e i) 481, 482, y 483 de la SEC ID Nº 2. Más preferiblemente, el polinucléotido aislado comprende una mutación de codón que provoca una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, y un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, d) un resto de serina en una posición equivalente o análoga al resto aminoacídico 337 ó 500 de la SEC ID Nº 2, e) un resto de arginina en una posición equivalente o análoga al resto aminoacídico 355 de la SEC ID Nº 2, f) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 385 ó 470 de la SEC ID Nº 2, g) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 462 de la SEC ID Nº 2, h) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, i) un ácido glutámico en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, j) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 321 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 322 de la SEC ID Nº 2, y un resto de valina en una posición equivalente o análoga al resto aminoacídico 323 de la SEC ID Nº 2, k) un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 450 de la SEC ID Nº 2, un resto de valina en una posición equivalente o análoga al resto aminoacídico 451 de la SEC ID Nº 2, y un resto de arginina en una posición equivalente o análoga al resto aminoacídico 452 de la

SEC ID Nº 2, l) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 455 de la SEC ID Nº 2, un resto de serina en una posición equivalente o análoga al resto aminoacídico 456 de la SEC ID Nº 2, un resto de alanina en una posición equivalente o análoga al resto aminoacídico 457 de la SEC ID Nº 2, y un resto de glutamina en una posición equivalente o análoga al resto aminoacídico 458 de la SEC ID Nº 2, m) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 470 de la SEC ID Nº 2, un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, y un resto de tirosina en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, n) un resto de treonina en una posición equivalente o análoga a los restos aminoacídicos 475, 477, y 478 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 476 de la SEC ID Nº 2, y un resto de histidina en una posición equivalente o análoga al resto aminoacídico 479 de la SEC ID Nº 2, u o) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 481 de la SEC ID Nº 2, un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 482 de la SEC ID Nº 2, y un resto de prolina en una posición equivalente o análoga al resto aminoacídico 483 de la SEC ID Nº 2. Incluso más preferiblemente, el polinucléotido aislado comprende una mutación de codón o mutaciones de codón que provocan una mutación de sustitución seleccionada entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, b) E401D y G429S de la SEC ID Nº 1, c) T337S, K355R S385A, V462L, S470A, E471D, T473E, o E500S de la SEC ID Nº 2, d) G321L, P322R, y G323V de la SEC ID Nº 2, e) D450E, A451V, y K452R de la SEC ID Nº 2, f) S455K, N456S, P457A, y S458Q de la SEC ID Nº 2, g) S470A, E472D, y T473Y de la SEC ID Nº 2, h) C475T, K476R, Q477T, K478T, e Y479H de la SEC ID Nº 2, e i) E481D, Q482E, y Q483P de la SEC ID Nº 2.

En otra realización específica, la presente invención se refiere a un polinucléotido aislado que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución potencia la actividad de unión a no esteroides o sensibilidad de no esteroide del dominio de unión al ligando del receptor nuclear de Grupo B. Preferiblemente, el polinucléotido aislado comprende una mutación de codón que provoca una sustitución de un resto aminoacídico en una posición equivalente o análoga al resto aminoacídico seleccionado entre el grupo compuesto por a) 401 ó 429 de la SEC ID Nº 1, b) 401 y 429 de la SEC ID Nº 1, c) 337, 355, 385, 462, 470, 472, 473, ó 500 de la SEC ID Nº 2, d) 321, 322, y 323 de la SEC ID Nº 2, e) 450, 451, y 452 de la SEC ID Nº 2, f) 455, 456, 457, y 458 de la SEC ID Nº 2, g) 470, 472, y 473 de la SEC ID Nº 2, h) 475, 476, 477, 478, y 479 de la SEC ID Nº 2, e i) 481, 482, y 483 de la SEC ID Nº 2. Más preferiblemente, el polinucléotido aislado comprende una mutación de codón que provoca una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, y un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, d) un resto de serina en una posición equivalente o análoga al resto aminoacídico 337 ó 500 de la SEC ID Nº 2, e) un resto de arginina en una posición equivalente o análoga al resto aminoacídico 355 de la SEC ID Nº 2, f) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 385 ó 470 de la SEC ID Nº 2, g) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 462 de la SEC ID Nº 2, h) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, i) un ácido glutámico en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, j) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 321 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 322 de la SEC ID Nº 2, y un resto de valina en una posición equivalente o análoga al resto aminoacídico 323 de la SEC ID Nº 2, k) un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 450 de la SEC ID Nº 2, un resto de valina en una posición equivalente o análoga al resto aminoacídico 451 de la SEC ID Nº 2, y un resto de arginina en una posición equivalente o análoga al resto aminoacídico 452 de la SEC ID Nº 2, l) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 455 de la SEC ID Nº 2, un resto de serina en una posición equivalente o análoga al resto aminoacídico 456 de la SEC ID Nº 2, un resto de alanina en una posición equivalente o análoga al resto aminoacídico 457 de la SEC ID Nº 2, y un resto de glutamina en una posición equivalente o análoga al resto aminoacídico 458 de la SEC ID Nº 2, m) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 470 de la SEC ID Nº 2, un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, y un resto de tirosina en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, n) un resto de treonina en una posición equivalente o análoga a los restos aminoacídicos 475, 477, y 478 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 476 de la SEC ID Nº 2, y un resto de histidina en una posición equivalente o análoga al resto aminoacídico 479 de la SEC ID Nº 2, u o) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 481 de la SEC ID Nº 2, un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 482 de la SEC ID Nº 2, y un resto de prolina en una posición equivalente o análoga al resto aminoacídico 483 de la SEC ID Nº 2. Incluso más preferiblemente, el polinucléotido aislado comprende una mutación de codón o mutaciones de codón que provocan una mutación de sustitución seleccionada entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, b) E401D y G429S de la SEC ID Nº 1, c) T337S, K355R, S385A, V462L, S470A, E471D, T473E, o E500S de la SEC ID Nº 2, d) G321L, P322R, y G323V de la SEC ID Nº 2, e) D450E, A451V, y K452R de la SEC ID Nº 2, f) S455K, N456S, P457A, y S458Q de la SEC ID Nº 2, g) S470A, E472D, y T473Y de la SEC ID Nº 2, h) C475T, K476R, Q477T, K478T, e Y479H de la SEC ID Nº 2, e i) E481D, Q482E, y Q483P de la SEC ID Nº 2.

En otra realización específica, la presente invención se refiere a un polinucléotido aislado que codifica un dominio de

unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución potencia tanto la actividad de unión a esteroides o la sensibilidad de esteroide como la actividad de unión a no esteroides o sensibilidad de no esteroide del dominio de unión al ligando de Grupo B.

En otra realización específica, la presente invención también se refiere a un polinucléotido aislado que codifica un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución aumenta la sensibilidad de ligando de un heterodímero que comprende el dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución y una pareja de dimerización. En una realización específica, la pareja de dimerización es un polipéptido receptor de ecdisona. En otra realización específica, la pareja de dimerización es un polipéptido EcR truncado.

Además, la presente invención se refiere a un vector de expresión que comprende un polinucléotido de acuerdo con la invención, unido de forma funcional a un elemento regulador de la transcripción. Preferiblemente, el polinucléotido que codifica un dominio de unión al ligando de receptor nuclear que comprende una mutación de sustitución está unido de forma funcional con una secuencia de control de la expresión que permite la expresión del dominio de unión al 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 que se desea la expresión. El vector puede ser una molécula de ADN plasmídico o un vector viral. Los vectores virales preferidos incluyen retrovirus, adenovirus, virus adeno-asociados, herpesvirus, y virus vaccinia. La invención se refiere adicionalmente a un virus recombinante de replicación defectuosa que comprende en su genoma, el polinucléotido que codifica un dominio de unión al ligando de receptor nuclear que comprende una mutación de sustitución como se ha descrito anteriormente. Por tanto, la presente invención también se refiere a una célula huésped aislada que comprende dicho vector de expresión, donde el elemento regulador de la transcripción es funcional en la célula huésped.

La presente invención también se refiere a un polipéptido aislado codificado por un polinucléotido de acuerdo con la invención.

Polipéptidos de la invención

El nuevo sistema de expresión génica inducible basado en receptores nucleares de la invención comprende al menos un casete de expresión génica que comprende un polinucléotido que codifica un polipéptido que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución. Por tanto, la presente invención también proporciona un polipéptido aislado que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención.

En otra realización específica, el dominio de unión al ligando del receptor nuclear de Grupo B comprende una mutación de sustitución en una posición equivalente o análoga al resto aminoacídico a) 401 ó 429 de la SEC ID Nº 1, b) 401 y 429 de la SEC ID Nº 1, c) 337, 344, 355, 385, 431, 442, 462, 470, 472, 473, 495, 500, 511, 516, ó 528 de la SEC ID Nº 2, d) 321, 322, y 323 de la SEC ID Nº 2, e) 450, 451, y 452 de la SEC ID Nº 2, f) 455, 456, 457, y 458 de la SEC ID Nº 2, g) 470, 472, y 473 de la SEC ID Nº 2, h) 475, 476, 477, 478, y 479 de la SEC ID Nº 2, e i) 481, 482, y 483 de la SEC ID Nº 2.

Preferiblemente, el dominio de unión al ligando del receptor nuclear de Grupo B comprende una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, y un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, d) un resto de serina en una posición equivalente o análoga al resto aminoacídico 337, 495, 500, ó 528 de la SEC ID Nº 2, e) un resto de asparagina en una posición equivalente o análoga al resto aminoacídico 344 de la SEC ID Nº 2, f) un resto de arginina en una posición equivalente o análoga al resto aminoacídico 355 ó 511 de la SEC ID Nº 2, g) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 385 ó 470 de la SEC ID Nº 2, h) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 431 ó 462 de la SEC ID Nº 2, i) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 442 de la SEC ID Nº 2, j) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, k) un ácido glutámico en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, l) un resto de valina en una posición equivalente o análoga al resto aminoacídico 516 de la SEC ID Nº 2, m) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 321 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 322 de la SEC ID Nº 2, y un resto de valina en una posición equivalente o análoga al resto aminoacídico 323 de la SEC ID Nº 2, n) un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 450 de la SEC ID Nº 2, un resto de valina en una posición equivalente o análoga al resto aminoacídico 451 de la SEC ID Nº 2, y un resto de arginina en una posición equivalente o análoga al resto aminoacídico 452 de la SEC ID Nº 2, o) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 455 de la SEC ID Nº 2, un resto de serina en una posición equivalente o análoga al resto aminoacídico 456 de la SEC ID Nº 2, un resto de alanina en una posición equivalente o análoga al resto aminoacídico 457 de la SEC ID Nº 2, y un resto de glutamina en una posición equivalente o análoga al resto aminoacídico 458 de la SEC ID Nº 2, p) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 470 de la SEC ID Nº 2, un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, y un

resto de tirosina en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, q) un resto de treonina en una posición equivalente o análoga a los restos aminoacídicos 475, 477, y 478 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 476 de la SEC ID Nº 2, y un resto de histidina en una posición equivalente o análoga al resto aminoacídico 479 de la SEC ID Nº 2, o r) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 481 de la SEC ID Nº 2, un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 482 de la SEC ID Nº 2, y un resto de prolina en una posición equivalente o análoga al resto aminoacídico 483 de la SEC ID Nº 2. En una realización preferida, el dominio de unión al ligando del receptor nuclear de Grupo B es de un receptor X retinoide.

En otra realización específica, el dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución es un dominio de unión al ligando del receptor X retinoide que comprende una mutación de sustitución, donde la mutación de sustitución se selecciona entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, b) E401D y G429S de la SEC ID Nº 1, c) T337S, D344N, K355R, S385A, M431L, R442K, V462L, S470A, E471D, T473E, A495S, E500S, K511R, T516V, o A528S de la SEC ID Nº 2, d) G321L, P322R, y G323V de la SEC ID Nº 2, e) D450E, A451V, y K452R de la SEC ID Nº 2, f) S455K, N456S, P457A, y S458Q de la SEC ID Nº 2, g) S470A, E472D, y T473Y de la SEC ID Nº 2, h) C475T, K476R, Q477T, K478T, e Y479H de la SEC ID Nº 2, e i) E481D, Q482E, y Q483P de la SEC ID Nº 2.

La presente invención también proporciona un polipéptido aislado seleccionado entre el grupo compuesto por a) un polipéptido aislado que comprende un dominio de transactivación, un dominio de unión a ADN, y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención; b) un polipéptido aislado que comprende un dominio de unión a ADN y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención; y c) un polipéptido aislado que comprende un dominio de transactivación y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención. En una realización preferida, el dominio de unión al ligando del receptor nuclear de Grupo B es de un receptor X retinoide.

La presente invención también proporciona un polipéptido híbrido aislado seleccionado entre el grupo compuesto por a) un polipéptido híbrido aislado que comprende un dominio de transactivación, un dominio de unión a ADN, y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención; b) un polipéptido híbrido aislado que comprende un dominio de unión a ADN y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención; y c) un polipéptido híbrido aislado que comprende un dominio de transactivación y un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución de acuerdo con la invención. En una realización preferida, el dominio de unión al ligando del receptor nuclear de Grupo B es de un receptor X retinoide.

La presente invención también proporciona un polipéptido aislado que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución que afecta a la actividad de unión al ligando o sensibilidad de ligando del dominio de unión al ligando del receptor nuclear de Grupo B.

En particular, la presente invención se refiere a un polipéptido de receptor nuclear de Grupo B aislado que comprende un dominio de unión al ligando que comprende una mutación de sustitución que reduce la actividad de unión al ligando o sensibilidad de ligando del dominio de unión al ligando del receptor nuclear de Grupo B.

En una realización específica, la presente invención se refiere a un polipéptido aislado que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución reduce la actividad de unión a esteroides o sensibilidad de esteroide del dominio de unión al ligando del receptor nuclear de Grupo B. Preferiblemente, el polipéptido aislado comprende una sustitución de un resto aminoacídico en una posición equivalente o análoga a a) el resto aminoacídico 401 ó 429 de la SEC ID Nº 1, o b) el resto aminoacídico 344, 431, 442, 495, 511, ó 528 de la SEC ID Nº 2. Más preferiblemente, el polipéptido aislado comprende una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de asparagina en una posición equivalente o análoga al resto aminoacídico 344 de la SEC ID Nº 2, d) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 431 de la SEC ID Nº 2, e) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 442 de la SEC ID Nº 2, f) un resto de serina en una posición equivalente o análoga al resto aminoacídico 495 ó 528 de la SEC ID Nº 2, o g) un resto de arginina en una posición equivalente o análoga al resto aminoacídico 511 de la SEC ID Nº 2. Incluso más preferiblemente, el polipéptido aislado comprende una mutación de sustitución seleccionada entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, y b) D344N, M431L, R442K, A495S, K511R, o A528S de la SEC ID Nº 2.

En una realización específica, la presente invención se refiere a un polipéptido aislado que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución reduce la actividad de unión al ligando no esteroideo o sensibilidad de ligando del dominio de unión al ligando del receptor nuclear de Grupo B. Preferiblemente, el polipéptido aislado comprende una sustitución de un resto aminoacídico en una posición equivalente o análoga a a) el resto aminoacídico 401 ó 429 de la SEC ID Nº 1, o b) el resto aminoacídico 344, 431, 442, 495, 511, ó 528 de la SEC ID Nº 2. Más preferiblemente, el polipéptido

aislado comprende una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de asparagina en una posición equivalente o análoga al resto aminoacídico 344 de la SEC ID Nº 2, d) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 431 de la SEC ID Nº 2, e) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 442 de la SEC ID Nº 2, f) un resto de serina en una posición equivalente o análoga al resto aminoacídico 495 ó 528 de la SEC ID Nº 2, o g) un resto de arginina en una posición equivalente o análoga al resto aminoacídico 511 de la SEC ID Nº 2. Incluso más preferiblemente, el polipéptido aislado comprende una mutación de sustitución seleccionada entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, y b) D344N, M431L, R442K, A495S, K511R, o A528S de la SEC ID Nº 2.

Además, la presente invención se refiere a un polipéptido aislado que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución que potencia la actividad de unión al ligando o sensibilidad de ligando del dominio de unión al ligando del receptor nuclear de Grupo B.

En una realización específica, la presente invención se refiere a un polipéptido aislado que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución que potencia la actividad de unión a esteroides o sensibilidad de esteroide del dominio de unión al ligando del receptor nuclear de Grupo B. Preferiblemente, el polipéptido aislado comprende una sustitución de un resto aminoacídico en una posición equivalente o análoga al resto aminoacídico seleccionado entre el grupo compuesto por a) 401 ó 429 de la SEC ID Nº 1, b) 401 y 429 de la SEC ID Nº 1, c) 337, 355, 385, 462, 470, 472, 473, ó 500 de la SEC ID Nº 2, d) 321, 322, y 323 de la SEC ID Nº 2, e) 450, 451, y 452 de la SEC ID Nº 2, f) 455, 456, 457, y 458 de la SEC ID Nº 2, g) 470, 472, y 473 de la SEC ID Nº 2, h) 475, 476, 477, 478, y 479 de la SEC ID Nº 2, e i) 481, 482, y 483 de la SEC ID Nº 2. Más preferiblemente, el polipéptido aislado comprende una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, y un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, d) un resto de serina en una posición equivalente o análoga al resto aminoacídico 337 ó 500 de la SEC ID Nº 2, e) un resto de arginina en una posición equivalente o análoga al resto aminoacídico 355 de la SEC ID Nº 2, f) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 385 ó 470 de la SEC ID Nº 2, g) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 462 de la SEC ID Nº 2, h) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, i) un ácido glutámico en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, j) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 321 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 322 de la SEC ID Nº 2, y un resto de valina en una posición equivalente o análoga al resto aminoacídico 323 de la SEC ID Nº 2, k) un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 450 de la SEC ID Nº 2, un resto de valina en una posición equivalente o análoga al resto aminoacídico 451 de la SEC ID Nº 2, y un resto de arginina en una posición equivalente o análoga al resto aminoacídico 452 de la SEC ID Nº 2, l) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 455 de la SEC ID Nº 2, un resto de serina en una posición equivalente o análoga al resto aminoacídico 456 de la SEC ID Nº 2, un resto de alanina en una posición equivalente o análoga al resto aminoacídico 457 de la SEC ID Nº 2, y un resto de glutamina en una posición equivalente o análoga al resto aminoacídico 458 de la SEC ID Nº 2, m) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 470 de la SEC ID Nº 2, un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, y un resto de tirosina en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, n) un resto de treonina en una posición equivalente o análoga a los restos aminoacídicos 475, 477, y 478 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 476 de la SEC ID Nº 2, y un resto de histidina en una posición equivalente o análoga al resto aminoacídico 479 de la SEC ID Nº 2, u o) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 481 de la SEC ID Nº 2, un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 482 de la SEC ID Nº 2, y un resto de prolina en una posición equivalente o análoga al resto aminoacídico 483 de la SEC ID Nº 2. Incluso más preferiblemente, el polipéptido aislado comprende una mutación de sustitución seleccionada entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, b) E401D y G429S de la SEC ID Nº 1, c) T337S, K355R, S385A, V462L, S470A, E471D, T473E, o E500S de la SEC ID Nº 2, d) G321L, P322R, y G323V de la SEC ID Nº 2, e) D450E, A451V, y K452R de la SEC ID Nº 2, f) S455K, N456S, P457A, y S458Q de la SEC ID Nº 2, g) S470A, E472D, y T473Y de la SEC ID Nº 2, h) C475T, K476R, Q477T, K478T, e Y479H de la SEC ID Nº 2, e i) E481D, Q482E, y Q483P de la SEC ID Nº 2.

En otra realización específica, la presente invención se refiere a un polipéptido aislado que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución que potencia la actividad de unión a no esteroides o sensibilidad de no esteroide del dominio de unión al ligando del receptor nuclear de Grupo B. Preferiblemente, el polipéptido aislado comprende una sustitución de un resto aminoacídico en una posición equivalente o análoga al resto aminoacídico seleccionado entre el grupo compuesto por a) 401 ó 429 de la SEC ID Nº 1, b) 401 y 429 de la SEC ID Nº 1, c) 337, 355, 385, 462, 470, 472, 473, ó 500 de la SEC ID Nº 2, d) 321, 322, y 323 de la SEC ID Nº 2, e) 450, 451, y 452 de la SEC ID Nº 2, f) 455, 456, 457, y 458 de la SEC ID Nº 2, g) 470, 472, y 473 de la SEC ID Nº 2, h) 475, 476, 477, 478, y 479 de la SEC ID Nº 2, e i) 481, 482, y 483 de la SEC ID

Nº 2. Más preferiblemente, el polipéptido aislado comprende una sustitución de a) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, b) un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, c) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 401 de la SEC ID Nº 1, y un resto de serina en una posición equivalente o análoga al resto aminoacídico 429 de la SEC ID Nº 1, d) un resto de serina en una posición equivalente o análoga al resto aminoacídico 337 ó 500 de la SEC ID Nº 2, e) un resto de arginina en una posición equivalente o análoga al resto aminoacídico 355 de la SEC ID Nº 2, f) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 385 ó 470 de la SEC ID Nº 2, g) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 462 de la SEC ID Nº 2, h) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, i) un ácido glutámico en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, j) un resto de leucina en una posición equivalente o análoga al resto aminoacídico 321 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 322 de la SEC ID Nº 2, y un resto de valina en una posición equivalente o análoga al resto aminoacídico 323 de la SEC ID Nº 2, k) un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 450 de la SEC ID Nº 2, un resto de valina en una posición equivalente o análoga al resto aminoacídico 451 de la SEC ID Nº 2, y un resto de arginina en una posición equivalente o análoga al resto aminoacídico 452 de la SEC ID Nº 2, l) un resto de lisina en una posición equivalente o análoga al resto aminoacídico 455 de la SEC ID Nº 2, un resto de serina en una posición equivalente o análoga al resto aminoacídico 456 de la SEC ID Nº 2, un resto de alanina en una posición equivalente o análoga al resto aminoacídico 457 de la SEC ID Nº 2, y un resto de glutamina en una posición equivalente o análoga al resto aminoacídico 458 de la SEC ID Nº 2, m) un resto de alanina en una posición equivalente o análoga al resto aminoacídico 470 de la SEC ID Nº 2, un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 472 de la SEC ID Nº 2, y un resto de tirosina en una posición equivalente o análoga al resto aminoacídico 473 de la SEC ID Nº 2, n) un resto de treonina en una posición equivalente o análoga a los restos aminoacídicos 475, 477, y 478 de la SEC ID Nº 2, un resto de arginina en una posición equivalente o análoga al resto aminoacídico 476 de la SEC ID Nº 2, y un resto de histidina en una posición equivalente o análoga al resto aminoacídico 479 de la SEC ID Nº 2, u o) un resto de ácido aspártico en una posición equivalente o análoga al resto aminoacídico 481 de la SEC ID Nº 2, un resto de ácido glutámico en una posición equivalente o análoga al resto aminoacídico 482 de la SEC ID Nº 2, y un resto de prolina en una posición equivalente o análoga al resto aminoacídico 483 de la SEC ID Nº 2. Incluso más preferiblemente, el polipéptido aislado comprende una mutación de sustitución seleccionada entre el grupo compuesto por a) E401D o G429S de la SEC ID Nº 1, b) E401D y G429S de la SEC ID Nº 1, c) T337S, K355R, S385A, V462L, S470A, E471D, T473E, o E500S de la SEC ID Nº 2, d) G321L, P322R, y G323V de la SEC ID Nº 2, e) D450E, A451V, y K452R de la SEC ID Nº 2, f) S455K, N456S, P457A, y S458Q de la SEC ID Nº 2, g) S470A, E472D, y T473Y de la SEC ID Nº 2, h) C475T, K476R, Q477T, K478T, e Y479H de la SEC ID Nº 2, e i) E481D, Q482E, y Q483P de la SEC ID Nº 2.

En otra realización específica, la presente invención se refiere a un polipéptido aislado que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución potencia tanto la actividad de unión a esteroides o sensibilidad de esteroide como la actividad de unión a no esteroides o sensibilidad de no esteroide del dominio de unión al ligando de Grupo B.

En otra realización específica, la presente invención también se refiere a un polipéptido aislado que comprende un dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución, donde la mutación de sustitución aumenta la sensibilidad de ligando de un heterodímero que comprende el dominio de unión al ligando del receptor nuclear de Grupo B que comprende una mutación de sustitución y un pareja de dimerización. En una realización específica, la pareja de dimerización es un polipéptido receptor de ecdisona. En otra realización específica, la pareja de dimerización es un polipéptido EcR truncado.

La presente invención también se refiere a composiciones que comprenden un polipéptido aislado de acuerdo con la invención.

Procedimiento para modular la expresión génica de la invención

La invención de los solicitantes también se refiere a procedimientos para modular la expresión génica en una célula huésped usando un sistema de modulación de la expresión génica de acuerdo con la invención. Específicamente, la invención de los solicitantes proporciona un procedimiento 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 de acuerdo con la invención; y b) introducir en la célula huésped un ligando; donde el gen a modular es un componente de un casete de expresión génica que 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 tiene que modularse, por el cual, después de la introducción del ligando en la célula huésped, se modula la expresión del gen.

La invención también proporciona un procedimiento 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 de acuerdo con la invención; b) introducir en la célula huésped un casete de expresión génica de acuerdo con la invención, donde el casete de expresión génica comprende 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 tiene que modularse; y c) introducir en la célula huésped un ligando; por el cual, después de la introducción del ligando en la célula huésped, se modular la expresión del gen.

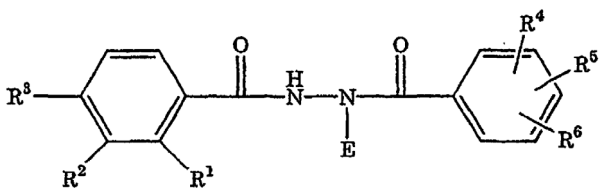
La invención de los solicitantes también proporciona un procedimiento para modular la expresión de un gen en una célula huésped que comprende un casete de expresión génica que comprende un elemento de respuesta que comprende un dominio al que se une el dominio de unión a ADN del primer polipéptido híbrido 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 tiene que modularse; donde el procedimiento comprende las etapas de: a) introducir en la célula huésped un sistema de modulación de la expresión génica de acuerdo con la invención; y b) introducir en la célula huésped un ligando; por el cual, después de la introducción del ligando en el huésped, se modular la expresión del gen.

Los genes de interés para su expresión en una célula huésped usando los procedimientos de los solicitantes pueden ser genes endógenos o genes heterólogos. La información de secuencia de ácido nucleico o de aminoácidos para un gen o proteína deseada puede localizarse en una de muchas bases de datos de acceso público, por ejemplo, GENBANK, EMBL, Swiss-Prot, y PIR, o en muchas publicaciones de revistas relacionadas con la biología. Por tanto, los especialistas en la técnica tienen acceso a información de secuencia de ácido nucleico para casi todos los genes conocidos. Dicha información entonces puede usarse para construir las construcciones deseadas para la inserción del gen de interés dentro de los casetes de expresión génica usados en los procedimientos de los solicitantes descritos en este documento.

Ejemplos de genes de interés para su expresión en una célula huésped usando los procedimientos de los solicitantes incluyen, aunque sin limitación: antígenos producidos en plantas como vacunas, enzimas como la alfa-amilasa, fitasa, glucanos, xilasa y xilanasa, genes para resistencia contra insectos, nematodos, hongos, bacterias, virus, y estrés abiótico, nutracéuticos, antígenos, agentes farmacéuticos, vitaminas, genes para modificar el contenido de aminoácidos, resistencia a herbicidas, frío, sequía, y tolerancia al calor, productos industriales, aceites, proteínas, carbohidratos, antioxidantes, plantas estériles masculinas, flores, combustibles, otros rasgos de producción, genes que codifican polipéptidos o productos terapéuticamente deseables que pueden usarse para tratar una afección, una enfermedad, un trastorno, una disfunción, un defecto genético, tales como anticuerpos monoclonales, enzimas, proteasas, citoquinas, interferones, insulina, eritropoyetina, factores de coagulación, otros factores o componentes sanguíneos, vectores virales para terapia génica, virus para vacunas, dianas para el descubrimiento de fármacos, análisis y aplicaciones de genómica funcional y proteómica, y similares.

Ligandos aceptables son cualquiera que module la expresión del gen cuando la unión del dominio de unión a ADN del sistema de expresión génica de acuerdo con la invención al elemento de respuesta en presencia del ligando provoca la activación o supresión de la expresión de los genes. Ligandos preferidos incluyen un ecdisteroide, tal como ecdisona, 20-hidroxiecdisona, ponasterona A, muristerona A, y similares, ácido 9-cis-retinoico, análogos sintéticos de ácido retinoico, N,N'-diacilhidrazinas tales como las descritas en las patentes de Estados Unidos Nº 6.013.836; 5.117.057; 5.530.028; y 5.378.726; dibenzoilalquil cianohidrazinas tales como las descritas en la solicitud europea Nº 461.809; N-alquil-N,N'-diaroilhidrazinas tales como las descritas en la patente de Estados Unidos Nº 5.225.443; N-acil-N-alquilcarbonilhidrazinas tales como las descritas en la solicitud europea Nº 234.994; N-aroil-N-alquil-N'-aroilhidrazinas tales como las descritas en la patente de Estados Unidos Nº 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 1,1-bifosfonato, hormona juvenil III, un ácido 9-cis-retinoide, ácido 4-(1-(3,5,5,8,8-pentametil-5,6,7,8-tetrahidro-2-naftil)-etenil) benzoico (3-metil-TTEB), (ácido (E)-2)2-(5,6,7,8-tetra-hidro-3,5,5,8,8-pentametil-2-naftil)propen-1-il)-4-tiofenocarboxílico), 2-(5,6,7,8-tetra-hidro-3,5,5,8,8-tetrametil-2-naftil)-2-(carboxifenil)-1,3-dioxolano, ácido 4-(5H-2,3-(2,5-dimetil-2,5-hexano)-5-metil-dibenzo (b,e) ácido (1,4)diazepin-11-il)-benzoico (HX600) o análogos tiadiazepina del mismo, ácido 3,7,11,15-tetrametil hexadecanoico (ácido fitánico), ácido 6-(1-(3,5,5,8,8-pentametil-5,6,7,8-tetrahidronaftalen-2-il)ciclopropil)nicotínico, 2(-4-carboxifenil)-2-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)-1,3-ditiano, o ácido 4-(2-metil)-1-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)propenil)benzoico, y similares.

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



en la que:

E es un alquilo (C_4-C_6) que contiene un carbono terciario o un cianoalquilo (C_3-C_5) que contiene un carbono

terciario;
R^1 es H, Me, Et, i-Pr, F, formilo, CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH CH_2OMe, CH_2CN, CN, CºCH, 1-propinilo, 2-propinilo, vinilo, OH, OMe, OEt, ciclopropilo, CF_2CF_3, CH=CHCN, alilo, azido, SCN o $SCHF_2$;
R^2 es H, Me, Et, n-Pr, i-Pr, formilo, CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CH_2OMe, CH_2CN, CN, CCH, 1-propinilo, 2-propinilo, vinilo, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc, NMe_2, NEt_2, SMe, SEt, $SOCF_3$, OCF_2CF_2H, COEt, ciclopropilo, CF_2CF_3, CH=CHCN, alilo, azido, OCF_3, $OCHF_2$, O-i-Pr, SCN, $SCHF_2$, SOMe, NH-CN, o unido con R^3 y los carbonos fenilo a los que están unidos R^2 y R^3 forma un etilenodioxi, un anillo dihidrofurilo con el oxígeno adyacentes a un carbono fenilo, o un anillo dihidropirilo con el oxígeno adyacentes a un carbono fenilo;
R^3 es H, Et, o unido con R^2 y los carbonos fenilo a los que están unidos R^2 y R^3 forma un etilenodioxi, un anillo dihidrofurilo con el oxígeno adyacentes a un carbono fenilo, o un anillo dihidropirilo con el oxígeno adyacente a un carbono fenilo;
R^4, R^5, y R^6 son independientemente H, Me, Et, F, Cl, Br, formilo, CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CN, CCH, 1-propinilo, 2-propinilo, vinilo, OMe, OEt, SMe, o SEt.

En otra realización preferida, el ligando para su uso en el procedimiento de los solicitantes para modular la expresión de un gen es una ecdisona, 20-hidroxiecdisona, ponasterona A, o muristerona A.

En otra realización preferida, el ligando para su uso en el procedimiento de los solicitantes para modular la expresión de un gen es un ácido 9-cis-retinoico, ácido 4-(1-(3,5,5,8,8-pentametil-5,6,7,8-tetrahidro-2-naftil)-etenil)benzoico (3-metil-TTEB), (ácido (E)-2)2-(5,6,7,8-tetra-hidro-3,5,5,8,8-pentametil-2-naftil)propen-1-il)-4-tiofenocarboxílico), 2-(5,6,7,8-tetra-hidro-3,5,5,8,8-tetrametil-2-naftil)-2-(carboxifenil)-1,3dioxolano, ácido 4-(5H-2,3-(2,5-dimetil-2,5-hexano)-2,5-metil-dibenzo (b,e) ácido (1,4)diazepin-11-il)-benzoico (HX600) o análogos tiadiazepina del mismo, ácido 3,7,11,15-tetrametil hexadecanoico (ácido fitánico), ácido 6-(1-(3,5,5,8,8-pentametil-5,6,7,8-tetrahidronaftalen-2-il)ciclopropil)nicotínico, 2(-4-carboxifenil)-2-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)-1,3-ditiano, o ácido 4-(2-metil)-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)propenil)benzoico.

En otra realización preferida, puede usarse un segundo ligando además del primer ligando analizado anteriormente en el procedimiento 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

Como se ha descrito anteriormente, el sistema de modulación de expresión génica de la presente invención puede usarse para modular 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 posibilita la modulación de expresión génica en células huésped procariotas y eucariotas. La expresión en células huésped transgénicas es útil para la expresión de diversos polipéptidos de interés incluyendo pero sin limitación 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, antígenos, compuestos farmacéuticos, vitaminas, genes para modificar el contenido de aminoácidos, resistencia a herbicida, tolerancia al frío, sequía y calor, productos industriales, aceites, proteína, carbohidratos, antioxidantes, plantas estériles masculinas, flores, combustibles, otros rasgos de rendimiento, polipéptidos terapéuticos, intermedios de ruta; para la modulación de rutas ya existentes en el huésped para la síntesis de nuevos productos no posibles hasta la fecha usando el huésped; ensayos basados en células; ensayos genómicos funcionales, 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.

Por lo tanto, la invención de los solicitantes proporciona una célula huésped aislada que comprende un sistema de expresión génica de acuerdo con la invención. La presente invención también proporciona una célula huésped aislada que comprende un casete de expresión génica de acuerdo con la invención. La invención de los solicitantes también proporciona una célula huésped aislada que comprende un polinucleótido o un polipéptido de acuerdo con la invención. En otra realización, la invención se refiere a una célula huésped transfectada con un vector de expresión de acuerdo con la invención. 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. En otra realización más, la invención se refiere a un procedimiento para producir un dominio de unión a ligando del receptor nuclear que comprende una mutación de sustitución, en la que el procedimiento comprende cultivar la célula huésped como se ha descrito anteriormente en medio de cultivo en condiciones que permiten la expresión de un polinucleótido que codifica el dominio de unión a ligando del receptor nuclear que comprende una mutación de sustitución, y aislar el dominio de unión a ligando del receptor nuclear que comprende una mutación de sustitución del cultivo.

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. Preferentemente, la célula huésped se selecciona en el 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 preferentemente, 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 sin limitación, especies fúngicas o de levadura 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, Methylomonas, 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, haba, maíz, melón, mijo, judía mung, avena, ocra, *Panicum,* papaya, cacahuete, guisante, pimiento, judía de la India, piña, *Phaseolus*, patata, calabaza gigante, arroz, sorgo, soja, calabaza, caña de azúcar, remolacha azucarera, girasol, boniato, té, tomate, tabaco, sandía y trigo, células huésped animales y de mamíferos.

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

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

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

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, haba, maíz, melón, mijo, judía mung, avena, ocra, *Panicum,* papaya, cacahuete, guisante, pimiento, judía de la India, piña, *Phaseolus,* patata, calabaza gigante, arroz, sorgo, soja, calabaza, caña de azúcar, remolacha azucarera, girasol, boniato, té, tomate, tabaco, sandía y trigo.

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.

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

La transformación de células huésped se conoce bien en la técnica y puede conseguirse por una diversidad de procedimientos incluyendo pero sin limitación electroporación, infección viral, transfección de plásmido/vector, transfección no mediada por vector 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 expresión del gen transformado. Se conocen bien en la técnica las condiciones de cultivo y protocolos de expresión génica en células procariotas y eucariotas (véase la sección de Procedimientos Generales de los Ejemplos). Las células pueden recogerse y los productos génicos aislarse de acuerdo con protocolos específicos para el producto génico.

Además, puede seleccionarse una célula huésped que module la expresión del polinucleótido insertado, o modifique y procese el producto polipeptídico de la manera específica deseada. Diferentes células huésped tienen mecanismos característicos y específicos para el procesamiento y modificación traduccional y postraduccional [por ejemplo, glicosilación, escisión (por ejemplo, de secuencia señal)] de proteínas. Pueden seleccionarse líneas celulares apropiadas o sistemas de huésped para asegurar la modificación y procesamiento deseados de la proteína ajena expresada. Por ejemplo, puede usarse expresión en un sistema bacteriano para producir un producto proteico central no glicosilado. Sin embargo, un polipéptido expresado en bacterias puede no estar plegado de forma apropiada. La expresión en levadura puede producir un producto glicosilado. La expresión en células eucariotas puede aumentar la probabilidad de glicosilación “nativa” y plegamiento de una proteína heteróloga. Además, la expresión en células de mamífero puede proporcionar una herramienta para reconstituir, o constituir, la actividad de polipéptido. Además, diferentes sistemas de expresión de huésped/vector pueden afectar a las reacciones de procesamiento, tales como escisiones proteolíticas, en un grado diferente.

La invención de los solicitantes también se refiere a un organismo no humano que comprende una célula huésped aislada de acuerdo con 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 invertebrado o un organismo vertebrado.

Preferentemente, 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 ave, un animal y un mamífero. Más preferentemente, 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, haba, maíz, melón, mijo, judía mung, avena, ocra, *Panicum,* papaya, cacahuete, guisante, pimiento, frijol de la India, piña, *Phaseolus,* patata, calabaza gigante, arroz, sorgo, soja, calabaza, 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 procedimientos 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 por cualquiera de varias tecnologías de expresión génica existentes.

La tecnología de matriz de ácido nucleico es una técnica útil para determinar expresión de ARNm diferencial. Dicha tecnología incluye, por ejemplo, microplacas 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 e hibridan con sondas preparadas a partir de grupos de ARNm total extraídos de células, tejidos u organismos completos y convertidos a ADNc. Las microplacas de oligonucleótidos son matrices de oligonucleótidos sintetizados en un sustrato usando técnicas fotolitográficas. Se han producido microplacas 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. Las micromatrices con hasta 10.000 genes se preparan de forma rutinaria en el mercado. La diferencia primaria entre estas dos técnicas es que las microplacas oligonucleotídicas normalmente usan oligonucleótidos de 25 unidades lo que permite el fraccionamiento de moléculas de ADN cortas mientras que las dianas de ADN mayores de micromatrices, aproximadamente 1000 pares de bases, pueden proporcionar más sensibilidad al fraccionar mezclas de ADN complejas.

Otra medición útil de los procedimientos de los solicitantes de la invención es la de determinar el estado de traducción de la célula midiendo las abundancias de la especie proteica constituyente presente en la célula usando procedimientos bien conocidos en la técnica.

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

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

Otros procedimientos para detectar los productos de expresión génica se conocen bien en la técnica e incluyen transferencias de Southern (detección de ADN), transferencias puntuales o por ranuras (ADN, ARN), transferencias de northern (ARN), RT-PCR (ARN), transferencias de 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 que hibrida.

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 procedimientos adecuados incluyendo, por ejemplo, reacción en cadena de la polimerasa (“PCR”), reacción en cadena de la ligasa (“LCR”), amplificación por desplazamiento de cadena (“SDA”), amplificación basada en transcripción y similares. La PCR se lleva a cabo de acuerdo con técnicas conocidas en las que, por ejemplo, una muestra de ácido nucleico se trata en presencia de una ADN polimerasa estable frente al calor, en condiciones de hibridación, con un par de cebadores oligonucleotídicos, hibridando un cebador con una hebra (molde) de la secuencia específica para detectar. Los cebadores son suficientemente complementarios a cada cadena de molde de la secuencia específica para hibridar con ella. Un producto de extensión de cada cebador se sintetiza y es complementario de la cadena molde de ácido nucleico con la que hibrida. El producto de extensión sintetizado a partir de cada cebador también puede actuar como un molde para síntesis adicional de productos de extensión usando los mismos cebadores. Después de un número suficiente de ciclos de síntesis de productos de

extensión, la muestra puede analizarse como se ha descrito anteriormente para evaluar si la secuencia o secuencias para detectar están presentes.

Ensayos de exploración de ligando

La presente invención también se refiere a procedimientos para explorar con respecto a un compuesto que induzca o reprima la transactivación de un dominio de unión a ligando del receptor nuclear que comprende una mutación de sustitución en una célula poniendo en contacto un dominio de unión a ligando del receptor nuclear con una molécula candidata y detectando la actividad de gen indicador en presencia del ligando. Los compuestos candidatos pueden ser agonistas o antagonistas del dominio de unión a ligando del receptor nuclear. En una realización preferida, el dominio de unión a ligando del 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.

En consecuencia, además del diseño racional de agonistas y antagonistas basándose en la estructura de un dominio de unión a ligando del receptor nuclear, la presente invención contempla un procedimiento alternativo para identificar ligandos específicos de un dominio de unión a ligando del receptor nuclear usando diversos ensayos de exploración conocidos en la técnica.

Cualquier técnica de exploración conocida en la materia puede usarse para explorar con respecto a agonistas o antagonistas de dominio de unión a ligando del receptor nuclear de Grupo B. Por ejemplo, una línea celular adecuada que comprende un sistema de expresión génica basado en receptor nuclear de acuerdo con la invención puede transfectarse con un casete de expresión génica que codifique un gen marcador unido operativamente con un promotor inducible o reprimible. Las células transfectadas se exponen después a una solución de ensayo que comprende un compuesto agonista o antagonista candidato, y después se ensayan con respecto a expresión o represión de gen marcador. La presencia de más expresión del gen marcador en relación con células de control no expuestas a la solución de ensayo es un indicio de la presencia de un compuesto agonista en la solución de ensayo. Por el contrario, la presencia de menos expresión del gen marcador en relación con células de control no expuestas a la solución de ensayo es un indicio de la presencia de un compuesto antagonista en la solución de ensayo.

La presente invención contempla exploraciones con respecto a ligandos moleculares pequeños o análogos y miméticos de ligando, así como exploraciones con respecto a ligandos naturales que se unen a y antagonizan o agonizan un dominio de unión a ligando del receptor nuclear de Grupo B de acuerdo con la invención *in vivo*. Por ejemplo, pueden explorarse bibliotecas de productos naturales usando ensayos de la invención con respecto a moléculas que agonizan o antagonizan la actividad del sistema de expresión génica basado en receptor nuclear.

La identificación y exploración de antagonistas se facilita adicionalmente determinando características estructurales de la proteína, por ejemplo, usando cristalografía de rayos X, difracción de neutrones, espectrometría de resonancia magnética nuclear y otras técnicas para determinación de estructura. Estas técnicas posibilitan el diseño racional o identificación de agonistas y antagonistas.

Otro enfoque usa bacteriófagos recombinantes para producir bibliotecas grandes. Usando el "procedimiento de fagos" [Scott y Smith, 1990, Science 249: 386-390 (1990); Cwirla, y col., Proc. Natl. Acad Sci., 87: 6378-6382 (1990); Devlin y col., Science, 249:404-406 (1990)], pueden construirse bibliotecas muy grandes (10^6-10^8 entidades químicas). Un segundo enfoque usa principalmente procedimientos químicos, de los que son ejemplos el procedimiento de Geysen [Geysen y col., Molecular Immunology 23: 709-715 (1986); Geysen y col. J. Immunologic Method 102:259-274 (1987)] y el procedimiento de Fodor y col. [Science 251: 767-773 (1991)]. Furka y col. [14th International Congress of Biochemistry, Volumen 5, Resumen FR:013 (1988); Furka, Int. J. Peptide Protein Res. 37: 487-493 (1991)], Houghton [Patente de Estados Unidos Nº: 4.631.211, expedida en diciembre de 1986] y Rutter y col. [Patente de Estados Unidos Nº: 5.010.175, expedida el 23 de abril de 1991] describen procedimientos para producir una mezcla de péptidos que pueden ensayarse como agonistas o antagonistas.

En otro aspecto, pueden usarse bibliotecas sintéticas [Needels y col., Proc. Natl. Acad Sci. USA 90:10700-4 (1993); Ohlmeyer y col., Proc. Natl. Acad Sci. USA 90:10922-10926 (1993); Lam y col., Publicación de Patente Internacional Nº WO 92/00252; Kocis y col., Publicación de Patente Internacional Nº WO 9428028], y similares para explorar con respecto a ligandos candidato de acuerdo con la presente invención.

La exploración puede realizarse con células recombinantes que expresan un dominio de unión a ligando del receptor nuclear de acuerdo con la invención, o como alternativa, usando proteína purificada, por ejemplo, producida de forma recombinante, como se ha descrito anteriormente. Por ejemplo, pueden usarse dominios de unión a ligando del receptor nuclear solubles marcados para explorar bibliotecas, como se ha descrito en las referencias anteriores.

En una realización, un dominio de unión a ligando del receptor nuclear del Grupo B de acuerdo con la invención puede marcarse directamente. En otra realización, puede usarse un reactivo secundario marcado para detectar la unión de un dominio de unión a ligando del receptor nuclear de la invención con una molécula de interés, por ejemplo, una molécula unida a un soporte de fase sólida. La unión puede detectarse mediante formación *in situ* de un cromóforo por un marcador enzimático. Las enzimas adecuadas incluyen, pero sin limitación, fosfatasa alcalina y peroxidasa de rábano rusticano. 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.

Otros marcadores para su uso en la invención 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 Eu^{3+}, por nombrar unos pocos 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. Pueden detectarse moléculas o células marcadas visualmente o mediante medios mecánicos/ópticos. Los medios mecánicos/ópticos incluyen separación activada por fluorescencia, es decir, análogos a FACS, y medios de retirada de micromanipulador.

La presente invención puede entenderse mejor por referencia a los siguientes ejemplos no limitantes, que se proporcionan como ejemplares de la invención.

Ejemplos

Los solicitantes han desarrollado un nuevo sistema de expresión génica inducible basado en receptor nuclear que comprende un dominio de unión a ligando de receptor nuclear de Grupo B que comprende una mutación de sustitución. Los solicitantes han mostrado que el efecto de una mutación de sustitución tal puede aumentar la actividad de unión a ligando o sensibilidad a ligando y puede ser específica de esteroides o no esteroides. Por lo tanto, la invención de los solicitantes proporciona un sistema de expresión génica inducible basado en receptor nuclear de Grupo B útil para modular la expresión de un gen de interés en una célula hospedadora. El nuevo sistema de expresión génica inducible de los solicitantes y su uso en procedimientos para modular la expresión génica en una célula huésped supera las limitaciones de sistemas de expresión inducibles actualmente disponibles y proporcionan al experto en la materia un medio eficaz para controlar la expresión génica.

Los nuevos polinucleótidos y polipéptidos de receptor nuclear mutado por sustitución de los solicitantes son útiles en un sistema de modulación génica inducible basado en receptor nuclear para diversas aplicaciones incluyendo pero sin limitación terapia génica, expresión de proteínas de interés en células huésped, producción de organismos transgénicos y ensayos basados en células.

Procedimientos generales

Se conocen bien en la materia técnicas de clonación molecular y ADN recombinante convencionales usadas en el presente documento y se describen en 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 en T. J. Silhavy, M. L. Bennan y L. W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1984) y en Ausubel, F. M. y col., Current Protocols in Molecular Biology, Greene Publishing Assoc. y Wiley-Interscience (1987).

Se conocen bien en la técnica materiales y procedimientos adecuados para el mantenimiento y crecimiento de cultivos bacterianos. Pueden encontrarse técnicas adecuadas para su uso en los siguientes ejemplos como se exponen 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 y 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 no ser que se especifique de otro modo.

Pueden conseguirse manipulaciones de secuencias genéticas usando el conjunto de programas disponible del Genetics Computer Group Inc. (Wisconsin Package Versión 9.0, Genetics Computer Group (GCG), Madison, WI). Cuando se usa el programa de GCG "Pileup" puede 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 CGC "Gap" o "Bestfit" puede 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 en el que no se pidan los parámetros del programa de GCG, en estos o cualquier otro programa de GCG, pueden usarse los valores por defecto.

El significado de las abreviaturas es como sigue: "h" significa hora u horas, "min" significa minuto o minutos, "s" significa segundo o segundos, "d" significa día o días, "μ" significa microlitro o microlitros, "ml" significa mililitro o mililitros, "l" significa litro o litros, "μM" significa micromolar, "mM" significa milimolar, "μg" significa microgramo o microgramos, "mg" significa miligramo o miligramos, "A" significa adenina o adenosina, "T" significa timina o timidina, “G” significa guanina o guanosina, "C" significa citidina o citosina, "xg" significa veces de gravedad, "nt" significa nucleótido o nucleótidos, "aa" significa aminoácido o aminoácidos, "pb" significa par o pares de bases, "kb" significa kilobase o kilobases, "k" significa kilo, "μ" significa micro y “ºC” significa grados Celsius.

Ejemplo 1

Este Ejemplo describe la construcción de varios casetes de expresión génica que comprenden nuevos polinucleótidos y polipéptidos de receptores nucleares del Grupo B mutados por sustitución de la invención para su uso en un sistema de expresión génica inducible basado en receptor nuclear. Los solicitantes construyeron varios casetes de expresión génica basándose en el EcR de la tortrix de las yemas de la pícea *Choristoneura fumiferana* EcR ("CtEcR"), receptor X retinoide α de ratón *Mus musculus* ("MmRXRα"), receptor X retinoide β humano de *Homo sapiens* ("HsRXRβ"), proteína Ultraespiráculo de langosta *Locusta migratoria* (denominada en el presente documento de forma intercambiable "LmUSP" o "LmRXR"), que es un homólogo de RXR de invertebrado de RXR de vertebrado, y USP de *C. fumiferana* ("CfUSP"). Las construcciones de receptor preparadas comprenden un dominio de unión a ligando de un EcR, un USP de invertebrado, un RXR de vertebrado, o un RXR de invertebrado; y un dominio de unión a ADN (DBD) de GAL4 o un dominio de transactivación (AD) de VP16. Las construcciones indicadoras incluyen un gen indicador, luciferasa o LacZ (β-galactosidasa), unido operativamente a una construcción promotora sintética que comprende un elemento de respuesta a GAL4 con el que se une el DBD de Gal4. Se cotransfectaron diversas combinaciones de estas construcciones de receptor e indicador en células de mamífero como se describe en los Ejemplos 2-4 posteriormente. Casetes de Expresión Génica: Los casetes de expresión génica (interruptores) se construyeron como sigue, usando procedimientos de clonación convencionales disponibles en la técnica. Lo siguiente es una breve descripción de la preparación y composición de cada interruptor usado en los Ejemplos descritos en el presente documento.

1.1 - GAL4CfEcR-DEF/VP16LmUSP-EF: Los dominios D, E y F de tipo silvestre de EcR de tortrix de las yemas de la pícea *Choristoneura fumiferana* ("CfEcR-DEF"; SEC ID Nº: 20) se fusionaron con un dominio de unión a ADN de GAL4 ("Gal4DNABD" o "Gal4DBD"; SEC ID Nº: 5) y se situaron bajo el control de un promotor de SV40e (SEC ID Nº: 21). Los dominios E y F de proteína Ultraespiráculo de langosta *Locusta migratoria* ("LmUSP-EF"; SEC ID Nº: 22) se fusionaron con el dominio de transactivación de VP16 ("VP16AD"; SEC ID Nº: 11) y se situaron bajo el control de un promotor de SV40e (SEC ID Nº: 21). Se fusionaron cinco sitios de unión a elemento de respuesta a GAL4 consenso ("5XGAL4RE"; que comprende 5 copias de un GAL4RE que comprende SEC ID Nº: 18) con un promotor mínimo de E1b sintético (SEC ID Nº: 23) y se situaron corriente arriba del gen de luciferasa (SEC ID Nº: 24).

1.2 - GAL4CfEcR-DEF/VP16MmRXRα-EF: Esta construcción se preparó de la misma manera que en el interruptor 1.1 anterior excepto que LmUSP-EF se reemplazó con los dominios E y F de MmRXRα ("MmRXRα-EF"; SEC ID Nº: 25).

1.3 - GAL4CfEcR-DEF/VP16mutanteMmRXRα-EF: Esta construcción se preparó de la misma manera que en el interruptor 1.2 anterior excepto que se reemplazó MmRXRα-EF de tipo silvestre con MmRXRα-EF mutante que comprende un dominio de unión a ligando que comprende una mutación de sustitución seleccionada de la Tabla 1 posterior.

Tabla 1. Mutantes de sustitución de Dominio de Unión a Ligando (LBD) de Receptor X retinoide α de *Mus musculus* ("MmRXRα").

Mutación de LBD de MmRXRα	**Sustitución de Aminoácidos "WT a Mutante" Resultante**	**Aminoácido correspondiente en MmRXRα de longitud completa (SEC ID Nº: 1)**
E401D	Ácido Glutámico (E) a Ácido Aspártico (D)	401
G429S	Glicina (G) a Serina (S)	429
E401D/G429S	Ácido Glutámico (E) a Ácido Aspártico (D) y Glicina (G) a Serina (S)	401 y 429 respectivamente

1.4-GAL4CfEcR-DEF/VP16MmRXRαH1-7-LmUSP-H8-12quimera: Esta construcción se preparó de la misma manera que en el interruptor 1.1 anterior excepto que se reemplazó LmUSP-EF con un dominio de unión a ligando quimérico que comprende las hélices (H) 1-7 de MmRXRα-EF y hélices 8-12 de LmUSP-EF (SEC ID Nº: 26).

1.5 - GAL4CfEcR-DEFNP16HsRXRβ-EF: Esta construcción se preparó de la misma manera que en el interruptor 1.1 anterior excepto que se reemplazó LmUSP-EF con los dominios E y F de HsRXRβ ("HsRXRβ-EF"; SEC ID Nº: 27).

1.6 - GAL4CfEcR-DEF/VP16mutanteHsflXRβ-EF: Esta construcción se preparó de la misma manera que en el interruptor 1.5 anterior excepto que se reemplazó HsRXRβ-EF de tipo silvestre con HsRXRβ-EF mutante que comprende un dominio de unión a ligando que comprende una mutación de sustitución seleccionada de la Tabla 2 posterior.

Tabla 2 Mutantes de Sustitución de Dominio de Unión a Ligando (LBD) de Receptor X Retinoide β de *Homo sapiens* ("HsRXRβ").

Mutación de LBD de HsRXRβ	**Sustitución de Aminoácidos "WT a Mutante" Resultante**	**Aminoácido correspondiente en HsRXRβ de longitud completa (SEC ID Nº: 2)**
G321L/P322R/ G323V	Glicina (G) a Leucina (L), Prolina (P) a Arginina (R), y Glicina (G) a Valina (V)	321, 322 y 323 respectivamente
T337S	Treonina (T) a Serina (S)	337
D344N	Ácido Aspártico (D) a Asparagina (N)	344
K355R	Lisina (K) a Arginina (R)	355
S385A	Serina (S) a Alanina (A)	385
M431L	Metionina (M) a Leucina (L)	431
R442K	Arginina (R) a Lisina (K)	442
D450E/A451V/ K452R	Ácido Aspártico (D) a Ácido Glutámico (E), Alanina (A) a Valina (V), y Lisina (K) a Arginina (R)	450,451 y 452 respectivamente
S455K/N456S/ P457A/S458Q	Serina (S) a Lisina (K), Asparagina (N) a Serina (S), Prolina (P) a Alanina (A), y Serina (S) a Glutamina (Q)	455, 456, 457 y 458 respectivamente
V462L	Valina (V) a Leucina (L)	462
S470A	Serina (S) a Alanina (A)	470
E472D	Ácido Glutámico (E) a Ácido Aspártico (D)	472
T473E	Treonina (T) a Ácido Glutámico (E)	473
S470A/E472D/ T473I	Serina (S) a Alanina (A), Ácido Glutámico (E) a Ácido Aspártico (D), y Treonina (T) a Tirosina (Y)	470, 472 y 473 respectivamente
C475T/K476R/ Q477T/K478T/ I479H	Cisteína (C) a Treonina (T),Lisina (K) a Arginina (R), Glutamina (Q) a Treonina (T), Lisina (K) a Treonina (T), y Tirosina (Y) a Histidina (H)	475, 476, 477, 478 y 479 respectivamente
E481D/Q482E/ Q483P	Ácido Glutámico (E) a Ácido Aspártico (D), Glutamina (Q) a Ácido Glutámico (E), y Glutamina (Q) a Prolina (P)	481, 482, y 483 respectivamente
A495S	Alanina (A) a Serina (S)	495
G500S	Glicina (G) a Serina (S)	500
K511R	Lisina (K) a Arginina (R)	511
T516V	Treonina (T) a Valina (V)	516
A528S	Alanina (A) a Serina (S)	528

Construcción de dominios de unión a ligando del receptor X retinoide que comprenden una mutación de sustitución:

En un intento de modificar la actividad de transactivación de RXR, se mutaron restos dentro de los dominios de unión a ligando de RXR que se había predicho que eran importantes basándose en comparaciones de secuencia en RXR de dos organismos diferentes. Las Tablas 1 y 2 indican los restos de aminoácidos dentro del dominio de unión al ligando de MmRxRα y HsRXRβ, respectivamente, que se mutaron y examinaron para modificación de la capacidad de transactivación.

Cada una de las mutaciones de sustitución de aminoácidos enumerada en las Tablas 1 y 2 se construyó en un ADNc de RXR por mutagénesis dirigida mediada por PCR. También se construyó un LBD de RXR mutante de dos puntos (véase Tabla 1), cuatro LBD de RXR mutantes de tres puntos diferentes (véase Tabla 2), un LBD de RXR mutante de cuatro puntos (véase Tabla 2) y un LBD de RXR mutante de cinco puntos (véase Tabla 2).

Se realizó mutagénesis dirigida por PCR usando el kit de mutagénesis dirigida de Quikchange (Stratagene, La Jolla, CA) usando las condiciones de reacción y parámetros de ciclación como sigue. Se realizó mutagénesis dirigida por PCR usando tampón de reacción 1 x (proporcionado 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 (proporcionado por el fabricante) en un volumen de reacción final de 50 μl. Los pares de cebador directo y cebador complementario inverso usados para producir cada mutación de sustitución de RXR se presentan en las Tablas 3 y 4. Los parámetros de ciclación usados consistieron en un ciclo de desnaturalización a 95 ºC durante 30 segundos, seguido de 16 ciclos de desnaturalización a 95 ºC durante 30 segundos, hibridación a 55 ºC durante 1 minuto y extensión a 68 ºC durante 22 minutos.

Tabla 3: Cebadores de PCR para Construcción de Dominio de Unión a Ligando de MmRXRα Mutante de Sustitución

MUTANTE	CEBADOR (SEC ID Nº:)	SECUENCIA DE NUCLEÓTIDOS DEL CEBADOR (5' A 3')
E401D	FP (SEC ID Nº: 28)	GTGTATGCGTCAGTAGATGCGTACTGCAAACAC
E401D	RCP (SEC ID Nº: 29)	GTGTTTGCAGTACGCATCTAGTGACGCATACAC
G429S	FP (SEC ID Nº: 30)	GCACTGCGTTCCATCAGCCTCAAGTGCCTGGAG
G429S	RCP (SEC ID Nº: 31)	CTCCAGGCACTTGAGGCTGATGGAACGCAGTGC

Tabla 4: Cebadores de PCR para Construcción de Dominio de Unión a Ligando de HsRXRβ Mutante de Sustitución

MUTANTE	CEBADOR (SEC ID Nº:)	SECUENCIA DE NUCLEÓTIDOS DEL CEBADOR (5' A 3')
G321L/ P322R/ G323V	FP (SEC ID Nº: 32)	**GACCAGGGCGTTGAGCGTCGTGTGGGAACCGGGGGT AGC**
G321L/ P322R/ G323V	RCP (SEC ID Nº: 33)	**GCTACCCCCGGTTCCCACACGACGCTCAACGCCCTGG TC**
T337S	FP (SEC ID Nº: 34)	CCAAATGACCCTGTGTCTAACATCTGTCAGGC
T337S	RCP (SEC ID Nº: 35)	GCCTGACAGATGTTAGACACAGGGTCATTTGG
D344N	FP (SEC ID Nº: 36)	ATCTGTCAGGCAGCTAACAAACAGCTATTCACG
D344N	RCP (SEC ID Nº: 37)	CGTGAATAGCTGTTTGTTAGCTGCCTGACAGAT
K355R	FP (SEC ID Nº: 38)	CTTGTTGAGTGGGCGAGGAGGATCCCACACTTTTC
K355R	RCP (SEC ID Nº: 39)	GAAAAGTGTGGGATCCTCCTCGCCCACTCAACAAG
S385A	FP (SEC ID Nº: 40)	CTCATTGCCTCCTTTGCACACCGATCCATTGATG
S385A	RCP (SEC ID Nº: 41)	CATCAATGGATCGGTGTGCAAAGGAGGCAATGAG
M431L	FP (SEC ID Nº: 42)	TCCAAAATGCGTGACCTGAGGATGGACAAGAC
431L	RCP (SEC ID Nº: 43)	GTCTTGTCCATCCTCAGGTCACGCATTTTGGA

R442K	FP (SEC ID Nº: 44)	GAGCTTGGCTGCCTGAAGGCAATCATTCTGTTTAATC
R442K	RCP (SEC ID Nº: 45)	GATTAAACAGAATGATTGCCTTCAGGCAGCCAAGCTC
UTANTE	CEBADOR (SEC ID Nº:)	SECUENCIA DE NUCLEÓTIDOS DEL CEBADOR (5' A 3')
D450E/ A451V/ K452R	FP (SEC ID Nº: 46)	**CATTCTGTTTAATCCAGAGGTCAGGGGCCTCTCCAAC CC**
D450E/ A451V/ K452R	RCP (SEC ID Nº: 47)	**GGGTTGGAGAGGCCCCTGACCTCTGGATTAAACAGAA TG**
S455K/N456S/P457A/ S458Q	FP (SEC ID Nº: 48)	**GATGCCAAGGGCCTCAAGTCCGCGCAGGAGGTGGAG GTCCTG**
S455K/N456S/P457A/ S458Q	RCP (SEC ID Nº: 49)	**CAGGACCTCCACCTCCTGCGCGGACTTGAGGCCCTTG GCATC**
V462L	FP (SEC ID Nº: 50)	CCTAGTGAGGTGGAGCTCCTGCGGGAGAAAGTGTATG
V462L	RCP (SEC ID Nº: 51)	CATACACTTTCTCCCGCAGGAGCTCCACCTCACTAGG
S470A	FP (SEC ID Nº: 52)	GAGAAAGTGTATGCAGCACTGGAGACCTACTGC
S470A	RCP (SEC ID Nº: 53)	GCAGTAGGTCTCCAGTGCTGCATACACTTTCTC
E472D	FP (SEC ID Nº: 54)	GTGTATGCATCACTGGATACCTACTGCAAACAG
E472D	RCP (SEC ID Nº: 55)	CTGTTTGCAGTAGGTATCCAGTGATGCATACAC
T473E	FP (SEC ID Nº: 56)	GTATGCATCACTGGAGGAGTACTGCAAACAGAAG
T473E	RCP (SEC ID Nº: 57)	CTTCTGTTTGCAGTACTCCTCCAGTGATGCATAC
S470A/ E472D/ T473Y	FP (SEC ID Nº: 58)	**GAGAAAGTGTATGCAGCACTGGATGAGTACTGCAAAC AGAAG**
S470A/ E472D/ T473Y	RCP (SEC ID Nº: 59)	**CTTCTGTTTGCAGTACTCATCCAGTGCTGCATACACTT TCTC**
C475T/K476R/Q477T/ K478T/ Y479H	FP (SEC ID Nº: 60)	**CATCACTGGAGACCTACACCAGAACGACGCACCCTGA GCAGCAGGGAC**
C475T/K476R/Q477T/ K478T/ Y479H	RCP (SEC ID Nº: 61)	**GTCCCTGCTGCTCAGGGTGCGTCGTTCTGGTGTAGGT CTCCAGTGATG**
E481D/ Q482E/ Q483P	FP (SEC ID Nº: 62)	**CAAACAGAAGTACCCTGACGAGCCGGGACGGTTTGCC AAG**
E481D/ O482F/ Q482E/ Q483P	RCP (SEC ID Nº: 63)	**CTTGGCAAACCGTCCCGGCTCGTCAGGGTACTTCTGT TTG**
A495S	FP (SEC ID Nº: 64)	CTGCTACGTCTTCCTTCCCTCCGGTCCATTGGC
A495S	RCP (SEC ID Nº: 65)	GCCAATGGACCGGAGGGAAGGAAGACGTAGCAG
G500S	FP (SEC ID Nº: 66)	GCCCTCCGGTCCATTAGCCTTAAGTGTCTAGAG
G500S	RCP (SEC ID Nº: 67)	CTCTAGACACTTAAGGCTAATGGACCGGAGGGC
K511R	FP (SEC ID Nº: 68)	CATCTGTTTTTCTTCAGGCTCATTGGTGACACC
K511R	RCP (SEC ID Nº: 69)	GGTGTCACCAATGAGCCTGAAGAAAAACAGATG
T516V	FP (SEC ID Nº: 70)	AGCTCATTGGTGACGTCCCCATCGACACCTTCC
T516V	RCP (SEC ID Nº: 71)	GGAAGGTGTCGATGGGGACGTCACCAATGAGCT
A528S	FP (SEC ID Nº: 72)	ATGGAGATGCTTGAGTCTCCCCATCAACTGGCC

A528S	RCP (SEC ID Nº: 73)	GGCCAGTTGATGGGGAGACTCAAGCATCTCCAT

Los productos de ácido nucleico de PCR resultantes que codificaban los dominios de unión a ligando de RXR mutante se fusionaron después cada uno con un dominio de transactivación de VP16 como se ha descrito en los Ejemplos 1.3 y 1.6 anteriores. Las construcciones de receptor RXR mutante/VP16 se ensayaron con respecto a actividad transfectando con ellas células NIH3T3 junto con GAL4/CfEcR-DEF y pFRLuc en presencia de ligando no esteroideo.

Ligando: El ligando no esteroideo N-(2-etil-3-metoxibenzoil)-N'-(3,5-dimetilbenzoil)-N'-terc-butilhidrazina (GS™-E) es un ligando ecdisteroideo estable sintético sintetizado en Rohm y Haas Company. El ligando se disolvió en DMSO y la concentración final de DMSO se mantuvo al 0,1 % tanto en los controles como en los tratamientos.

Transfecciones: ADN correspondientes a las diversas construcciones de interruptor perfiladas en el Ejemplo 1.1-1.6 se transfectaron a células NIH3T3 de ratón (ATCC) como sigue. Se siguieron procedimientos convencionales para cultivo y mantenimiento de las células. Las células se recogieron cuando alcanzaron el 50 % de confluencia y se sembraron en placas de 6-, 12- o 24- pocillos a 125.000, 50.000 o 25.000 células, respectivamente en 2,5, 1,0 o 0,5 ml de medio de crecimiento que contenía suero bovino fetal al 10 % (FBS), respectivamente. Al día siguiente, las células se aclararon con medio de crecimiento y se transfectaron durante cuatro horas. Se usó Superfect™ (Qiagen Inc.) como el reactivo de transfección. Para placas de 12 pocillos, se mezclaron 4 μl de Superfect™ con 100 μl de medio de crecimiento. Se añadieron 1 μg de construcción indicadora y 0,25 μg de cada construcción de receptor del par de receptores para analizar a la mezcla de transfección. Se añadió una segunda construcción indicadora [pTKRL (Promega), 0,1 μg/mezcla de transfección] que comprende un gen de luciferasa de *Renilla* unido operativamente y situado bajo el control de un promotor constitutivo de timidina quinasa (TK) y se usó para normalización. Los contenidos de la mezcla de transfección se mezclaron en un mezclador vorticial y se dejaron en reposo a temperatura ambiente durante 30 minutos. Al final de la incubación, la mezcla de transfección se añadió a las células mantenidas en 400 μl de medio de crecimiento. Las células se mantuvieron a 37 ºC y CO_2 5 % durante 4 horas. Al final de la incubación, se añadieron 500 μl de medio de crecimiento que contenía FBS 20 % y dimetilsulfóxido (DMSO; control) o una solución de DMSO de ligando no esteroideo y las células se mantuvieron a 37 ºC y CO_2 5 % durante 48 horas. Las células se recogieron y se ensayó la actividad indicadora. Se siguió el mismo procedimiento también para placas de 6 y 24 pocillos excepto que todos los reactivos se doblaron para placas de 6 pocillos y se redujeron a la mitad para placas de 24 pocillos.

Ensayos Indicadores: Las células se recogieron 40 horas después de añadir ligando. Se añadieron 125 μl de tampón de lisis pasivo (parte del sistema de ensayo indicador Dual-luciferase™ de Promega Corporation) a cada pocillo de la placa de 24 pocillos. Las placas se situaron en un agitador rotatorio durante 15 minutos. Se ensayaron 20 μl de lisado. La actividad luciferasa se midió usando un sistema de ensayo indicador Dual-luciferase™ de Promega Corporation siguiendo las instrucciones del fabricante. Se midió β-Galactosidasa usando kit de ensayo Galacto-Star™ de TROPIX siguiendo las instrucciones del fabricante. Todas las actividades luciferasa y β-galactosidasa se normalizaron usando luciferasa de *Renilla* como un patrón. Se calcularon las actividades en veces dividiendo las unidades de luz relativas ("ULR") normalizadas en células tratadas con ligando con ULR normalizadas en células tratadas con DMSO (control no tratado).

Ejemplo 2

En células de mamífero, el receptor de ecdisona de insecto (*Choristoneura fumiferana*) (CfEcR) heterodimeriza con Ultraespiráculo (USP) o su homólogo receptor X retinoide (RXR) y genes transactivados que se sitúan bajo el control de elementos de respuesta afines. La inducibilidad de ligando de esta transactivación basada en EcR depende de su compañero de heterodimerización. Como se ha mostrado previamente por los solicitantes, la transactivación a través de CfEcR en asociación con USP de insecto es independiente de ligando, mientras que la transactivación a través de CfEcR en asociación con RXR de invertebrados o vertebrados es dependiente de ligando (véase solicitud de patente de Estados Unidos en trámite 60/294.814). Los solicitantes han descubierto ahora que la secuencia de RXR puede modificarse por mutación de sustitución del dominio de unión a ligando para influir en la magnitud de la transactivación así como la sensibilidad al ligando de un sistema de expresión génica inducible basado en EcR.

Este ejemplo describe la identificación de tres mutantes de sustitución de dominio de unión a ligando MmRXRα que muestran sensibilidad a ligando mejorada en respuesta a ligando no esteroideo. Brevemente, los solicitantes construyeron tres mutantes de sustitución de dominio de unión a ligando de isoforma α de RXR de ratón y crearon casetes de expresión génica de ADNc VP16/mutante MmRXRα-EF como se ha descrito en el Ejemplo 1 anterior usando el kit de mutagénesis dirigida mediada por PCR Quikchange (Stratagene, La Jolla, CA). Los ADNc mutados se ensayaron junto con GAL4/CfEcR-DEF en ensayos indicadores de luciferasa conducidos por GAL4 y los resultados se compararon con interruptores de tipo silvestre VP16/MmRXRα-EF y VP16/MmRXRα-LmUSPquimera-EF en células NIH3T3 de ratón.

Transfecciones: Se transfectaron ADN correspondientes a las diversas construcciones de interruptor perfiladas en el Ejemplo 1, específicamente interruptores 1.1-1.4 en células NIH3T3 de ratón (ATCC) como sigue. Las células se recogieron cuando alcanzaron el 50 % de confluencia y se sembraron en placas de 24 pocillos a 12.500 células/pocillo en 0,5 ml de medio de crecimiento que contenía suero bovino fetal al 10 % (FBS). Al día siguiente, las

células se aclararon con medio de crecimiento y se transfectaron durante 4 horas. Se descubrió que Superfect™ (Qiagen Inc.) era el mejor reactivo de transfección para células 3T3. Se mezclaron dos μl de Superfect™ con 100 μl de medio de crecimiento y 50 ng de casete GAL4/CfEcR-DEF, 50 ng de: VP16/LmUSP-EF, VP16/tipo silvestre MmRXRα-EF, VP16/MmRXRα-LmUSPquimera-EF, o VP16/mutanteMmRXRα-EF, y se añadieron 200 ng de pFRLuc a la mezcla de transfección. Se añadió una segunda construcción indicadora [pTKRL (Promega), 0,05 μg/mezcla de transfección] que comprende un gen de luciferasa de *Renilla* unido operativamente y situado bajo el control de un promotor constitutivo de timidina quinasa (TK) y se usó para normalización. Los contenidos de la mezcla de transfección se mezclaron en un mezclador vorticial y se dejaron reposar a temperatura ambiente durante 30 minutos. Al final de la incubación, la mezcla de transfección se añadió a las células mantenidas en 200 μl de medio de crecimiento. Las células se mantuvieron a 37 ºC y CO_2 5 % durante cuatro horas. Al final de la incubación, se añadieron 250 μl de medio de crecimiento que contenía FBS 20 % y dimetilsulfóxido (DMSO; control) o una solución de DMSO de ligando no esteroideo GS™-E 0 [N-(2-etil-3-metoxibenzoil)-N'-(3,5-dimetilbenzoil)-N'-terc-butilhidrazina] 0,2, 1 o 10 μM y las células se mantuvieron a 37 ºC y CO_2 5 % durante 40 horas. Las células se recogieron y se ensayó la actividad indicadora como se ha descrito anteriormente. Se calcularon las actividades en veces dividiendo las unidades de luz relativas ("URL") normalizadas en células tratadas con ligando con URL normalizadas en células tratadas con DMSO (control no tratado).

Como se ha analizado anteriormente, los solicitantes han mostrado previamente que los RXR y homólogos de RXR (USP) de invertebrados no lepidópteros/no dípteros, denominados de forma colectiva en el presente documento RXR de invertebrados, se unen a EcR y transactivan indicadores a niveles más altos que los conseguidos por RXR de vertebrados en células de mamífero. Sin embargo, los RXR de vertebrados proporcionan niveles de fondo más bajos en ausencia de ligando. Los solicitantes compararon las secuencias de aminoácidos de dominios de unión a ligando de RXR de vertebrados y RXR de invertebrados y han identificado dos aminoácidos particulares que se conservan en todos los RXR de vertebrados pero son diferentes en RXR de invertebrados. Los solicitantes determinaron si reemplazar un aminoácido de RXR de vertebrado con un aminoácido de RXR de invertebrado daría como resultado mayor actividad tras inducción, pero proporcionaría aún bajo fondo en ausencia de ligando.

Los solicitantes han identificado ahora dos restos de aminoácidos dentro de los dominios EF de un RXR de ratón vertebrado, isoforma α ("MmRXRα-EF") que, cuando se sustituye, produce un RXR mutante que muestra sensibilidad aumentada a un ligando no esteroideo. El efecto de estas dos mutaciones de sustitución: una sustitución de ácido aspártico en el resto de aminoácido 401 (mutante E401D) y una sustitución de serina en el resto de aminoácido 429 (mutante G429S) de SEC ID Nº: 1, en la actividad del receptor mutado MmRXRα-EF se presenta en la Figura 1.

En ensayos de transactivación en células NIH3T3, el mutante E401D se comportó más como un RXR de invertebrado que un RXR de vertebrado demostrando tanto niveles de inducción mayores como de fondo mayores (véase Figura 1). El mutante G429S se comportó más como un RXR de invertebrado, demostrando sensibilidad aumentada al ligando pero menos fondo (véase Figura 1). Por lo tanto, los solicitantes han demostrado un resultado sorprendente e inesperado de que un cambio de aminoácido sencillo puede alterar drásticamente el comportamiento del MmRXRα-EF en presencia y ausencia de ligando.

Para determinar si la combinación de estas dos mutaciones proporcionaría un RXR más mejorado que cada mutación puntual sencilla por separado, los solicitantes realizaron un mutante de sustitución doble MmRXRα-EF que comprendía una mutación puntual en ambas posiciones E401D y G429S. Se analizaron tres clones independientes de estos dobles mutantes (DM) en células NIH3T3 y se compararon con cada mutante puntual sencillo por separado, MmRXRα-EF, y un RXR de vertebrado quimérico/RXR de invertebrado (véase Figura 2). Los resultados de los solicitantes muestran que estos dobles mutantes actúan también como el RXR quimérico, lo que demuestra niveles de fondo bajos en ausencia de ligando y niveles de inducción y sensibilidad al ligando aumentados. Estos nuevos mutantes de sustitución dobles proporcionan un RXR ventajoso para su uso en aplicaciones *in vivo*.

Los solicitantes también han determinado que estos mutantes de sustitución MmRXRα-EF responden mejor que HsRXRβ-EF de tipo silvestre en presencia de un ligando esteroideo, ponasterona A (Invitrogen), similar a su respuesta a GS™-E presentada en las Figuras 1 y 2 (datos no mostrados).

Ejemplo 3

Como se ha analizado anteriormente, el uso de un RXR de invertebrado, LmUSP (también denominado en el presente documento LmRXR) como un compañero heterodimérico de CfEcR mejoró tanto la sensibilidad como la magnitud de inducción de un sistema de expresión génica inducible basado en CfEcR. Los solicitantes han alienado las secuencias polipeptídicas de los dominios de unión a ligando de LmRXR y HsRXRβ (dominios EF) e identificado restos de aminoácidos que son diferentes entre estas dos proteínas. Este Ejemplo describe el análisis de los solicitantes de los mutantes de sustitución de HsRXRβ en los que se sustituyeron aminoácidos de LmRXR en lugar de los restos de tipo silvestre en HsRXRβ. Los solicitantes han identificado ahora varios mutantes de sustitución que modifican tanto la sensibilidad como la magnitud de la inducción de un sistema de expresión génica inducible basado en CfEcR en células de mamífero.

Brevemente, los solicitantes construyeron y analizaron mutantes de sustitución en RXR humano isoforma β-EF (HsRXRβ-EF), en los que se mutaron aminoácidos que son diferentes en HsRXRβ-EF en comparación con el homólogo de RXR de invertebrado LmUSP-EF al resto de aminoácido de LmUSP-EF y se ensayaron con respecto a su efecto en la actividad de transactivación en células NIH3T3 como se ha descrito en el Ejemplo 2 anterior en presencia de ligando no esteroideo GS™-E [N-(2-etil-3-metoxibenzoil)-N'-(3,5-dimetilbenzoil)-N'-terc-butilhidrazina] 0, 0,04, 0,2, 1, 5 o 25 μM. Las células se recogieron y se ensayó actividad indicadora como se ha descrito anteriormente. Se calcularon las actividades en veces dividiendo las unidades de luz relativas ("ULR") normalizadas en células tratadas con ligando con ULR normalizadas en células tratadas con DMSO (control no tratado).

Como se muestra en la Figura 3, los mutantes de sustitución de HsRXRβ-EF T/S, DAK/EVR, SNPS/KSAQ, V/L, S/A, T/E, CKQKY/ TRTTH y EQQ/DEP mejoraron la inducción en veces y la magnitud de la inducción. Los mutantes de sustitución de HsRXRβ-EF de los solicitantes DAK/EVR, SNPS/KSAQ, V/L, S/A, T/E, CKQKY/TRTTH y EQQ/DEP también mejoraron la sensibilidad a ligando no esteroideo (véase Figura 3).

Los resultados presentados en la Figura 3 también muestran que los mutantes de sustitución de HsRXRβ-EF D344N, A495S (marcado como A/S con una inducción de 30 veces) y A528S (marcado como A/S con una inducción de 21 veces) muestran inducción en veces, magnitud de inducción y/o sensibilidad a ligando reducidas en comparación con el tipo silvestre.

Los solicitantes también han determinado que estos mutantes de sustitución de HsRXRβ-EF responden mejor que HsRXRβ-EF de tipo silvestre en presencia de un ligando esteroideo, ponasterona A (Invitrogen), similar a su respuesta a GS™-E presentada en la Figura 3 (datos no mostrados).

Ejemplo 4

Como se ha analizado anteriormente, un sistema de regulación génica basado en CfEcR es dependiente de ligando cuando se usa RXR como el compañero heterodimérico e independiente de ligando cuando se usan USP como el compañero heterodimérico. Sin embargo, se induce expresión génica indicadora a niveles muy altos cuando se usa CfUSP como el compañero heterodimérico. Para mejorar RXR como compañero heterodimérico, los solicitantes han alineado las secuencias polipeptídicas de varios dominios de unión a ligando de RXR y USP (dominios EF) e identificado restos de aminoácidos que son diferentes entre estos dos miembros de la familia de los receptores nucleares. En particular, los solicitantes identificaron restos que se conservan en todos los USP pero son diferentes en RXR y mutaron estos restos en HsRXRβ. Estos mutantes de sustitución se analizaron en células NIH3T3 como se ha descrito en el Ejemplo 2 anterior en presencia de ligando no esteroideo GS™-E [N-(2-etil-3-metoxibenzoil)-N'-(3,5-dimetilbenzoil)-N'-terc-butilhidrazina] 0, 0,04, 0,2, 1, 5 o 25 μM. Las células se recogieron y la actividad indicadora se ensayó como se ha descrito anteriormente. Las actividades en veces se calcularon dividiendo las unidades de luz relativas ("ULR") normalizadas en células tratadas con ligando con ULR normalizadas en células tratadas con DMSO (control no tratado).

Brevemente, este ejemplo describe la construcción y análisis de mutantes de sustitución en RXR humano isoforma β-EF (HsRXRβ-EF), en los que los aminoácidos que son diferentes en HsRxRβ-EF en comparación con restos conservados de proteínas Ultraespiráculo se mutaron al resto de aminoácido de USP y se ensayaron con respecto a su efecto en la actividad de transactivación en células NIH3T3.

Como se muestra en la Figura 4, los mutantes de sustitución de HsRXRβ-EF GPG/LRV, K/R, S/A y G/S mejoraron la inducción en veces y la magnitud de inducción del sistema de regulación génica basado en EcR. Los mutantes de sustitución de HsRXRβ-EF GPG/LRV, S/A, E/D, SET/ADY y G/S mejoraron la sensibilidad a ligando del sistema de regulación génica basado en EcR (véase Figura 4).

Los resultados presentados en la Figura 4 también muestran que los mutantes de sustitución de HsRXRβ-EF M431L, R442K y K511R muestran inducción en veces, magnitud de inducción y/o sensibilidad a ligando reducidas en comparación con el tipo silvestre. El mutante de sustitución de HsRXRβ-EF M/L esencialmente eliminó la actividad de transactivación (véase Figura 4). Este mutante M/L es útil en ensayos de exploración de ligando ortogonales.

Los solicitantes también han determinado que estos mutantes de sustitución de HsRXRβ-EF responden mejor que HsRXRβ-EF de tipo silvestre en presencia de un ligando esteroideo, ponasterona A (Invitrogen), similar a su respuesta a GS™-E presentada en la Figura 4 (datos no mostrados).

Listado de secuencias

<110> Rohm and Haas Company
Palli, Subba R.
Kapitskaya, Marianna Z.

<120> Nuevos receptores mutantes de sustitución y su uso en un sistema de expresión génica basado en receptores nucleares

<130> A01255

<150> US 60/313.908
<151> 21-08-2001

<160> 75

<170> PatentIn versión 3.1

<210> 1
<211> 467
<212> PRT
<213> Mus musculus

<400> 1

```
Met Asp Thr Lys His Phe Leu Pro Leu Asp Phe Ser Thr Gln Val Asn
1               5                  10                  15

Ser Ser Ser Leu Asn Ser Pro Thr Gly Arg Gly Ser Met Ala Val Pro
            20                  25                  30

Ser Leu His Pro Ser Leu Gly Pro Gly Ile Gly Ser Pro Leu Gly Ser
        35                  40                  45

Pro Gly Gln Leu His Ser Pro Ile Ser Thr Leu Ser Ser Pro Ile Asn
    50                  55                  60

Gly Met Gly Pro Pro Phe Ser Val Ile Ser Ser Pro Met Gly Pro His
65                  70                  75                  80

Ser Met Ser Val Pro Thr Thr Pro Thr Leu Gly Phe Gly Thr Gly Ser
                85                  90                  95

Pro Gln Leu Asn Ser Pro Met Asn Pro Val Ser Ser Thr Glu Asp Ile
            100                 105                 110

Lys Pro Pro Leu Gly Leu Asn Gly Val Leu Lys Val Pro Ala His Pro
        115                 120                 125

Ser Gly Asn Met Ala Ser Phe Thr Lys His Ile Cys Ala Ile Cys Gly
    130                 135                 140
```

Asp Arg Ser Ser Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly Cys
145 150 155 160

Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asp Leu Thr Tyr Thr Cys
165 170 175

Arg Asp Asn Lys Asp Cys Leu Ile Asp Lys Arg Gln Arg Asn Arg Cys
180 185 190

Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Ala Met Gly Met Lys Arg Glu
195 200 205

Ala Val Gln Glu Glu Arg Gln Arg Gly Lys Asp Arg Asn Glu Asn Glu
210 215 220

Val Glu Ser Thr Ser Ser Ala Asn Glu Asp Met Pro Val Glu Lys Ile
225 230 235 240

Leu Glu Ala Glu Leu Ala Val Glu Pro Lys Thr Glu Thr Tyr Val Glu
245 250 255

Ala Asn Met Gly Leu Asn Pro Ser Ser Pro Asn Asp Pro Val Thr Asn
260 265 270

Ile Cys Gln Ala Ala Asp Lys Gln Leu Phe Thr Leu Val Glu Trp Ala
275 280 285

Lys Arg Ile Pro His Phe Ser Glu Leu Pro Leu Asp Asp Gln Val Ile
290 295 300

Leu Leu Arg Ala Gly Trp Asn Glu Leu Leu Ile Ala Ser Phe Ser His
305 310 315 320

Arg Ser Ile Ala Val Lys Asp Gly Ile Leu Leu Ala Thr Gly Leu His
325 330 335

Val His Arg Asn Ser Ala His Ser Ala Gly Val Gly Ala Ile Phe Asp
340 345 350

Arg Val Leu Thr Glu Leu Val Ser Lys Met Arg Asp Met Gln Met Asp
355 360 365

Lys Thr Glu Leu Gly Cys Leu Arg Ala Ile Val Leu Phe Asn Pro Asp
370 375 380

Ser Lys Gly Leu Ser Asn Pro Ala Glu Val Glu Ala Leu Arg Glu Lys

385 390 395 400

Val Tyr Ala Ser Leu Glu Ala Tyr Cys Lys His Lys Tyr Pro Glu Gln
405 410 415

Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser
420 425 430

Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe Phe Lys Leu Ile Gly
435 440 445

Asp Thr Pro Ile Asp Thr Phe Leu Met Glu Met Leu Glu Ala Pro His
450 455 460

Gln Ala Thr
465

<210> 2
<211> 533
<212> PRT
<213> Homo sapiens

<400> 2

Met Ser Trp Ala Ala Arg Pro Pro Phe Leu Pro Gln Arg His Ala Ala
1 5 10 15

Gly Gln Cys Gly Pro Val Gly Val Arg Lys Glu Met His Cys Gly Val
20 25 30

Ala Ser Arg Trp Arg Arg Arg Arg Pro Trp Leu Asp Pro Ala Ala Ala
35 40 45

Ala Ala Ala Ala Val Ala Gly Gly Glu Gln Gln Thr Pro Glu Pro Glu
50 55 60

Pro Gly Glu Ala Gly Arg Asp Gly Met Gly Asp Ser Gly Arg Asp Ser
65 70 75 80

Arg Ser Pro Asp Ser Ser Ser Pro Asn Pro Leu Pro Gln Gly Val Pro
85 90 95

Pro Pro Ser Pro Pro Gly Pro Pro Leu Pro Pro Ser Thr Ala Pro Ser
100 105 110

Leu Gly Gly Ser Gly Ala Pro Pro Pro Pro Pro Met Pro Pro Pro Pro
115 120 125

```
Leu Gly Ser Pro Phe Pro Val Ile Ser Ser Ser Met Gly Ser Pro Gly
    130                 135                 140

Leu Pro Pro Pro Ala Pro Pro Gly Phe Ser Gly Pro Val Ser Ser Pro
145                 150                 155                 160

Gln Ile Asn Ser Thr Val Ser Leu Pro Gly Gly Gly Ser Gly Pro Pro
                165                 170                 175

Glu Asp Val Lys Pro Pro Val Leu Gly Val Arg Gly Leu His Cys Pro
            180                 185                 190

Pro Pro Pro Gly Gly Pro Gly Ala Gly Lys Arg Leu Cys Ala Ile Cys
        195                 200                 205

Gly Asp Arg Ser Ser Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly
    210                 215                 220

Cys Lys Gly Phe Phe Lys Arg Thr Ile Arg Lys Asp Leu Thr Tyr Ser
225                 230                 235                 240

Cys Arg Asp Asn Lys Asp Cys Thr Val Asp Lys Arg Gln Arg Asn Arg
                245                 250                 255

Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Ala Thr Gly Met Lys Arg
            260                 265                 270

Glu Ala Val Gln Glu Glu Arg Gln Arg Gly Lys Asp Lys Asp Gly Asp
        275                 280                 285

Gly Glu Gly Ala Gly Gly Ala Pro Glu Glu Met Pro Val Asp Arg Ile
    290                 295                 300

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

Gly Pro Gly Gly Thr Gly Gly Ser Gly Ser Ser Pro Asn Asp Pro Val
                325                 330                 335

Thr Asn Ile Cys Gln Ala Ala Asp Lys Gln Leu Phe Thr Leu Val Glu
            340                 345                 350

Trp Ala Lys Arg Ile Pro His Phe Ser Ser Leu Pro Leu Asp Asp Gln
        355                 360                 365
```

Val Ile Leu Leu Arg Ala Gly Trp Asn Glu Leu Leu Ile Ala Ser Phe
370 375 380

Ser His Arg Ser Ile Asp Val Arg Asp Gly Ile Leu Leu Ala Thr Gly
385 390 395 400

Leu His Val His Arg Asn Ser Ala His Ser Ala Gly Val Gly Ala Ile
405 410 415

Phe Asp Arg Val Leu Thr Glu Leu Val Ser Lys Met Arg Asp Met Arg
420 425 430

Met Asp Lys Thr Glu Leu Gly Cys Leu Arg Ala Ile Ile Leu Phe Asn
435 440 445

Pro Asp Ala Lys Gly Leu Ser Asn Pro Ser Glu Val Glu Val Leu Arg
450 455 460

Glu Lys Val Tyr Ala Ser Leu Glu Thr Tyr Cys Lys Gln Lys Tyr Pro
465 470 475 480

Glu Gln Gln Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ala Leu
485 490 495

Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe Phe Lys Leu
500 505 510

Ile Gly Asp Thr Pro Ile Asp Thr Phe Leu Met Glu Met Leu Glu Ala
515 520 525

Pro His Gln Leu Ala
530

<210> 3
<211> 198
<212> ADN
<213> Choristoneura fumiferana

<400> 3

tgtctggtat gcggggacag agcctccgga taccactaca atgcgctcac gtgtgaaggg 60

tgtaaagggt tcttcagacg gagtgttacc aaaaatgcgg tttatatttg taaattcggt 120

cacgcttgcg aaatggacat gtacatgcga cggaaatgcc aggagtgccg cctgaagaag 180

tgcttagctg taggcatg 198

<210> 4
<211> 66
<212> PRT

<213> Choristoneura fumiferana

<400> 4

```
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> 5
<211> 441
<212> ADN
<213> Saccharomyces cerevisiae

<400> 5

```
atgaagctac tgtcttctat cgaacaagca tgcgatattt gccgacttaa aaagctcaag     60

tgctccaaag aaaaaccgaa gtgcgccaag tgtctgaaga acaactggga gtgtcgctac    120

tctcccaaaa ccaaaaggtc tccgctgact agggcacatc tgacagaagt ggaatcaagg    180

ctagaaagac tggaacagct atttctactg atttttcctc 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> 6
<211> 147
<212> PRT
<213> Saccharomyces cerevisiae

<400> 6

```
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> 7
<211> 606
<212> ADN
<213> Escherichia coli

<400> 7

atgaaagcgt taacggccag gcaacaagag gtgtttgatc tcatccgtga tcacatcagc 60

cagacaggta tgccgccgac gcgtgcggaa atcgcgcagc gtttggggtt ccgttcccca 120

aacgcggctg aagaacatct gaaggcgctg gcacgcaaag gcgttattga aattgtttcc 180

ggcgcatcac gcgggattcg tctgttgcag gaagaggaag aagggttgcc gctggtaggt 240

cgtgtggctg ccggtgaacc acttctggcg caacagcata ttgaaggtca ttatcaggtc 300

gatccttcct tattcaagcc gaatgctgat ttcctgctgc gcgtcagcgg gatgtcgatg 360

aaagatatcg gcattatgga tggtgacttg ctggcagtgc ataaaactca ggatgtacgt 420

aacggtcagg tcgttgtcgc acgtattgat gacgaagtta ccgttaagcg cctgaaaaaa 480

cagggcaata aagtcgaact gttgccagaa aatagcgagt ttaaaccaat tgtcgtagat 540

cttcgtcagc agagcttcac cattgaaggg ctggcggttg gggttattcg caacggcgac 600

tggctg 606

<210> 8
<211> 202
<212> PRT
<213> Escherichia coli

<400> 8

```
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 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> 9
<211> 420
<212> ADN
<213> Choristoneura fumiferana

<400> 9

atgagacgcc gctggtccaa caacgggggc ttccagacgc tgcgaatgct cgaggagagc 60

tcgtccgaag tgacgtcgtc ctcagctctg ggtctgccgg ccgcgatggt tatgtctccg 120

gagtcgctcg cctcgccaga gtacggcggg ctcgagctct ggggatacga cgatgggttg 180

tcatacaaca cggcgcagtc cttgctgggc aatacttgca cgatgcagca gcagcaacag 240

acgcagccgc tgccgtcgat gccgttgcct atgccgccga ccacgccgaa gtctgaaaac 300

gagtctattt cctcaggccg tgaggaactg tcgccagctt caagtataaa tgggtgcagt 360

acagatggcg aggcacgacg tcagaagaag ggccctgcgc cccgtcagca agaggaactg 420

<210> 10
<211> 140
<212> PRT
<213> Choristoneura fumiferana

<400> 10

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> 11
<211> 271
<212> ADN
<213> virus 7 del herpex simple

<400> 11

atgggcccta 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 cgatgcccct 240

ggaattgacg agtacggtgg ggaattcccg g 271

<210> 12
<211> 90
<212> PRT
<213> virus 7 del herpex simple

<400> 12

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> 13
<211> 307
<212> ADN
<213> Saccharomyces cerevisiae

<400> 13

atgggtgctc ctccaaaaaa gaagagaaag gtagctggta tcaataaaga tatcgaggag 60

tgcaatgcca tcattgagca gtttatcgac tacctgcgca ccggacagga gatgccgatg 120

gaaatggcgg atcaggcgat taacgtggtg ccgggcatga cgccgaaaac cattcttcac 180

gccgggccgc cgatccagcc tgactggctg aaatcgaatg gttttcatga aattgaagcg 240

gatgttaacg ataccagcct cttgctgagt ggagatgcct cctacccctta tgatgtgcca 300

gattatg 307

<210> 14
<211> 102
<212> PRT
<213> Saccharomyces cerevisiae

<400> 14

```
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> 15
<211> 807
<212> ADN
<213> Homo sapiens

<400> 15

cccatggaat tccagtacct gccagataca gacgatcgtc accggattga ggagaaacgt 60

aaaaggacat atgagacctt caagagcatc atgaagaaga gtcctttcag cggacccacc 120

gacccccggc ctccacctcg acgcattgct gtgccttccc gcagctcagc ttctgtcccc 180

aagccagcac cccagcccta tcccttttacg tcatccctga gcaccatcaa ctatgatgag 240

tttcccacca tggtgtttcc ttctgggcag atcagccagg cctcggcctt ggccccggcc 300

cctccccaag tcctgcccca ggctccagcc cctgcccctg ctccagccat ggtatcagct 360

ctggcccagg ccccagcccc tgtcccagtc ctagccccag gccctcctca ggctgtggcc 420

ccacctgccc ccaagcccac ccaggctggg gaaggaacgc tgtcagaggc cctgctgcag 480

ctgcagtttg atgatgaaga cctgggggcc ttgcttggca acagcacaga cccagctgtg 540

ttcacagacc tggcatccgt cgacaactcc gagtttcagc agctgctgaa ccagggcata 600

cctgtggccc cccacacaac tgagcccatg ctgatggagt accctgaggc tataactcgc 660

ctagtgacag gggcccagag gccccccgac ccagctcctg ctccactggg ggccccgggg 720

ctccccaatg gcctcctttc aggagatgaa gacttctcct ccattgcgga catggacttc 780

tcagccctgc tgagtcagat cagctcc 807

<210> 16
<211> 269
<212> PRT
<213> Homo sapiens

<400> 16

```
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 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> 17
<211> 34
<212> ADN
<213> Secuencia artificial

<220>
<223> Elemento de respuesta a Ecdisona (EcRE) 1X

<400> 17
tcgagagaca agggttcaat gcacttgtcc aatg 34

<210> 18
<211> 19
<212> ADN
<213> Secuencia artificial

<220>
<223> Elemento de respuesta a GAL4

<400> 18
ggagtactgt cctccgagc 19

<210> 19
<211> 36
<212> ADN
<213> Secuencia artificial

<220>
<223> Elemento de respuesta a 2xLexAop

<400> 19
ctgctgtata taaaaccagt ggttatatgt acagta 36

<210> 20
<211> 1054
<212> ADN
<213> Choristoneura fumiferana

<400> 20

cctgagtgcg tagtacccga gactcagtgc gccatgaagc ggaaagagaa gaaagcacag 60

aaggagaagg acaaactgcc tgtcagcacg acgacggtgg acgaccacat gccgcccatt 120

atgcagtgtg aacctccacc tcctgaagca gcaaggattc acgaagtggt cccaaggttt 180

ctctccgaca agctgttgga gacaaaccgg cagaaaaaca tcccccagtt gacagccaac 240

cagcagttcc ttatcgccag gctcatctgg taccaggacg ggtacgagca gccttctgat 300

gaagatttga agaggattac gcagacgtgg cagcaagcgg acgatgaaaa cgaagagtct 360

gacactccct tccgccagat cacagagatg actatcctca cggtccaact tatcgtggag 420

ttcgcgaagg gattgccagg gttcgccaag atctcgcagc ctgatcaaat tacgctgctt 480

aaggcttgct caagtgaggt aatgatgctc cgagtcgcgc gacgatacga tgcggcctca 540

gacagtgttc tgttcgcgaa caaccaagcg tacactcgcg acaactaccg caaggctggc 600

atggcctacg tcatcgagga tctactgcac ttctgccggt gcatgtactc tatggcgttg 660

gacaacatcc attacgcgct gctcacggct gtcgtcatct tttctgaccg gccagggttg 720

gagcagccgc aactggtgga agaaatccag cggtactacc tgaatacgct ccgcatctat 780

atcctgaacc agctgagcgg gtcggcgcgt tcgtccgtca tatacggcaa gatcctctca 840

atcctctctg agctacgcac gctcggcatg caaaactcca acatgtgcat ctccctcaag 900

ctcaagaaca gaaagctgcc gcctttcctc gaggagatct gggatgtggc ggacatgtcg 960

cacacccaac cgccgcctat cctcgagtcc cccacgaatc tctagccccc gcgcgcacgc 1020

atcgccgatg ccgcgtccgg ccgcgctgct ctga 1054

<210> 21
<211> 309
<212> ADN
<213> Virus del simio 40

<400> 21

ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 60

agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca 120

tgcatctcaa ttagtcagca accatagtcc cgcccctaac tccgcccatc ccgcccctaa 180

ctccgcccag ttccgcccat tctccgcccc atggctgact aatttttttt atttatgcag 240

aggccgaggc cgcctcggcc tctgagctat tccagaagta gtgaggaggc ttttttggag 300

gcctaggct 309

<210> 22
<211> 635
<212> ADN
<213> Locusta migratoria

<400> 22

tgcatacaga catgcctgtt gaacgcatac ttgaagctga aaaacgagtg gagtgcaaag 60

cagaaaacca agtggaatat gagctggtgg agtgggctaa acacatcccg cacttcacat 120

ccctacctct ggaggaccag gttctcctcc tcagagcagg ttggaatgaa ctgctaattg 180

cagcattttc acatcgatct gtagatgtta aagatggcat agtacttgcc actggtctca 240

cagtgcatcg aaattctgcc catcaagctg gagtcggcac aatatttgac agagttttga 300

cagaactggt agcaaagatg agagaaatga aaatggataa aactgaactt ggctgcttgc 360

gatctgttat tcttttcaat ccagaggtga ggggtttgaa atccgcccag gaagttgaac 420

ttctacgtga aaaagtatat gccgctttgg aagaatatac tagaacaaca catcccgatg 480

aaccaggaag atttgcaaaa cttttgcttc gtctgccttc tttacgttcc ataggcctta 540

agtgtttgga gcatttgttt ttctttcgcc ttattggaga tgttccaatt gatacgttcc 600

tgatggagat gcttgaatca ccttctgatt cataa 635

<210> 23
<211> 24
<212> ADN
<213> Secuencia artificial

<220>
<223> promotor mínimo Elb sintético

<400> 23
tatataatgg atccccgggt accg 24

<210> 24
<211> 1653
<212> ADN
<213> Photinus pyralis

<400> 24

atggaagacg ccaaaaacat aaagaaaggc ccggcgccat tctatcctct agaggatgga 60

accgctggag agcaactgca taaggctatg aagagatacg ccctggttcc tggaacaatt 120

gcttttacag atgcacatat cgaggtgaac atcacgtacg cggaatactt cgaaatgtcc 180

gttcggttgg cagaagctat gaaacgatat gggctgaata caaatcacag aatcgtcgta 240

tgcagtgaaa actctcttca attctttatg ccggtgttgg gcgcgttatt tatcggagtt 300

gcagttgcgc ccgcgaacga catttataat gaacgtgaat tgctcaacag tatgaacatt 360

tcgcagccta ccgtagtgtt tgtttccaaa aaggggttgc aaaaaatttt gaacgtgcaa 420

aaaaaattac caataatcca gaaaattatt atcatggatt ctaaaacgga ttaccaggga 480

tttcagtcga tgtacacgtt cgtcacatct catctacctc ccggttttaa tgaatacgat 540

tttgtaccag agtcctttga tcgtgacaaa acaattgcac tgataatgaa ttcctctgga 600

tctactgggt tacctaaggg tgtggccctt ccgcatagaa ctgcctgcgt cagattctcg 660

catgccagag atcctatttt tggcaatcaa atcattccgg atactgcgat tttaagtgtt 720

gttccattcc atcacggttt tggaatgttt actacactcg gatatttgat atgtggattt 780

cgagtcgtct taatgtatag atttgaagaa gagctgtttt tacgatccct tcaggattac 840

aaaattcaaa gtgcgttgct agtaccaacc ctattttcat tcttcgccaa aagcactctg 900

attgacaaat acgatttatc taatttacac gaaattgctt ctgggggcgc acctctttcg 960

aaagaagtcg gggaagcggt tgcaaaacgc ttccatcttc cagggatacg acaaggatat 1020

gggctcactg agactacatc agctattctg attacacccg agggggatga taaaccgggc 1080

gcggtcggta aagttgttcc atttttgaa gcgaaggttg tggatctgga taccgggaaa 1140

acgctgggcg ttaatcagag aggcgaatta tgtgtcagag gacctatgat tatgtccggt 1200

tatgtaaaca atccggaagc gaccaacgcc ttgattgaca aggatggatg gctacattct 1260

ggagacatag cttactggga cgaagacgaa cacttcttca tagttgaccg cttgaagtct 1320

ttaattaaat acaaaggata tcaggtggcc cccgctgaat tggaatcgat attgttacaa 1380

caccccaaca tcttcgacgc gggcgtggca ggtcttcccg acgatgacgc cggtgaactt 1440

cccgccgccg ttgttgtttt ggagcacgga aagacgatga cggaaaaaga gatcgtggat 1500

tacgtcgcca gtcaagtaac aaccgcgaaa aagttgcgcg gaggagttgt gtttgtggac 1560

gaagtaccga aaggtcttac cggaaaactc gacgcaagaa aaatcagaga gatcctcata 1620

aaggccaaga agggcggaaa gtccaaattg taa 1653

<210> 25
<211> 714
<212> ADN
<213> Mus musculus

<400> 25

gccaacgagg acatgcctgt agagaagatt ctggaagccg agcttgctgt cgagcccaag 60

actgagacat acgtggaggc aaacatgggg ctgaacccca gctcaccaaa tgaccctgtt 120

accaacatct gtcaagcagc agacaagcag ctcttcactc ttgtggagtg ggccaagagg 180

atcccacact tttctgagct gcccctagac gaccaggtca tcctgctacg ggcaggctgg 240

aacgagctgc tgatcgcctc cttctcccac cgctccatag ctgtgaaaga tgggattctc 300

ctggccaccg gcctgcacgt acaccggaac agcgctcaca gtgctggggt gggcgccatc 360

tttgacaggg tgctaacaga gctggtgtct aagatgcgtg acatgcagat ggacaagacg 420

gagctgggct gcctgcgagc cattgtcctg ttcaaccctg actctaaggg gctctcaaac 480

cctgctgagg tggaggcgtt gagggagaag gtgtatgcgt cactagaagc gtactgcaaa 540

cacaagtacc ctgagcagcc gggcaggttt gccaagctgc tgctccgcct gcctgcactg 600

cgttccatcg ggctcaagtg cctggagcac ctgttcttct tcaagctcat cggggacacg 660

cccatcgaca ccttcctcat ggagatgctg gaggcaccac atcaagccac ctag 714

<210> 26
<211> 711
<212> ADN
<213> Secuencia artificial

<220>
<223>Dominios EF quiméricos MmRXRalpha hélices 1-7 - LmUSP hélices 8-12

<400> 26

gccaacgagg acatgcctgt agagaagatt ctggaagccg agcttgctgt cgagcccaag 60

actgagacat acgtggaggc aaacatgggg ctgaacccca gctcaccaaa tgaccctgtt 120

accaacatct gtcaagcagc agacaagcag ctcttcactc ttgtggagtg ggccaagagg 180

atcccacact tttctgagct gcccctagac gaccaggtca tcctgctacg ggcaggctgg 240

aacgagctgc tgatcgcctc cttctcccac cgctccatag ctgtgaaaga tgggattctc 300

ctggccaccg gcctgcacgt acaccggaac agcgctcaca gtgctggggt gggcgccatc 360

tttgacaggg tgctaacaga gctggtgtct aagatgcgtg acatgcagat ggacaagact 420

gaacttggct gcttgcgatc tgttattctt ttcaatccag aggtgagggg tttgaaatcc 480

gcccaggaag ttgaacttct acgtgaaaaa gtatatgccg ctttggaaga atatactaga 540

acaacacatc ccgatgaacc aggaagattt gcaaaacttt tgcttcgtct gccttcttta 600

cgttccatag gccttaagtg tttggagcat ttgtttttct ttcgccttat tggagatgtt 660

ccaattgata cgttcctgat ggagatgctt gaatcacctt ctgattcata a 711

<210> 27
<211> 720
<212> ADN
<213> Homo sapiens

<400> 27

gcccccgagg agatgcctgt ggacaggatc ctggaggcag agcttgctgt ggaacagaag 60

agtgaccagg gcgttgaggg tcctggggga accgggggta gcggcagcag cccaaatgac 120

cctgtgacta acatctgtca ggcagctgac aaacagctat tcacgcttgt tgagtgggcg 180

aagaggatcc cacacttttc ctccttgcct ctggatgatc aggtcatatt gctgcgggca 240

ggctggaatg aactcctcat tgcctccttt tcacaccgat ccattgatgt tcgagatggc 300

atcctccttg ccacaggtct tcacgtgcac cgcaactcag cccattcagc aggagtagga 360

gccatctttg atcgggtgct gacagagcta gtgtccaaaa tgcgtgacat gaggatggac 420

aagacagagc ttggctgcct gagggcaatc attctgttta atccagatgc caagggcctc 480

tccaacccta gtgaggtgga ggtcctgcgg gagaaagtgt atgcatcact ggagacctac 540

tgcaaacaga agtaccctga gcagcaggga cggtttgcca agctgctgct acgtcttcct 600

gccctccggt ccattggcct taagtgtcta gagcatctgt ttttcttcaa gctcattggt 660

gacaccccca tcgacacctt cctcatggag atgcttgagg ctccccatca actggcctga 720

<210> 28
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 28
gtgtatgcgt cactagatgc gtactgcaaa cac 33

<210> 29
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 29
gtgtttgcag tacgcatcta gtgacgcata cac 33

<210> 30
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 30
gcactgcgtt ccatcagcct caagtgcctg gag 33

<210> 31
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 31
ctccaggcac ttgaggctga tggaacgcag tgc 33

<210> 32
<211> 39
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 32
gaccagggcg ttgagcgtcg tgtgggaacc gggggtagc 39

<210> 33
<211> 39
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 33
gctacccccg gttcccacac gacgctcaac gccctggtc 39

<210> 34
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 34
ccaaatgacc ctgtgtctaa catctgtcag gc 32

<210> 35
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 35
gcctgacaga tgttagacac agggtcattt gg 32

<210> 36
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 36
atctgtcagg cagctaacaa acagctattc acg 33

<210> 37
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 37
cgtgaatagc tgtttgttag ctgcctgaca gat 33

<210> 38
<211> 35
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 38
cttgttgagt gggcgaggag gatcccacac ttttc 35

<210> 39
<211> 35
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 39
gaaaagtgtg ggatcctcct cgcccactca acaag 35

<210> 40
<211> 34
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 40
ctcattgcct cctttgcaca ccgatccatt gatg 34

<210> 41
<211> 34
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 46
cattctgttt aatccagagg tcaggggcct ctccaaccc 39

<210> 47
<211> 39
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 47
gggttggaga ggcccctgac ctctggatta aacagaatg 39

<210> 48
<211> 42
<212> ADN
<213> Secuencia artificial

<220>

<223> Cebador PCR

<400> 48
gatgccaagg gcctcaagtc cgcgcaggag gtggaggtcc tg 42

<210> 49
<211> 42
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 49
caggacctcc acctcctgcg cggacttgag gcccttggca tc 42

<210> 50
<211> 37
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 50
cctagtgagg tggagctcct gcgggagaaa gtgtatg 37

<210> 51
<211> 37
<212> ADN
<213> Secuencia artificial

<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 41
catcaatgga tcggtgtgca aaggaggcaa tgag 34

<210> 42
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 42
tccaaaatgc gtgacctgag gatggacaag ac 32

<210> 43
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 43
gtcttgtcca tcctcaggtc acgcattttg ga 32

<210> 44
<211> 37
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 44
gagcttggct gcctgaaggc aatcattctg tttaatc 37

<210> 45
<211> 37
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 45
gattaaacag aatgattgcc ttcaggcagc caagctc 37

<210> 46
<211> 39
<212> ADN

<220>
<223> Cebador PCR

<400> 51
catacacttt ctcccgcagg agctccacct cactagg 37

<210> 52
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 52
gagaaagtgt atgcagcact ggagacctac tgc 33

<210> 53
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 53
gcagtaggtc tccagtgctg catacacttt ctc 33

<210> 54
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 54
gtgtatgcat cactggatac ctactgcaaa cag 33

<210> 55
<211> 33
<212> ADN
<213> Secuencia artificial

<220>

<223> Cebador PCR

<400> 55
ctgtttgcag taggtatcca gtgatgcata cac 33

<210> 56
<211> 34
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 56
gtatgcatca ctggaggagt actgcaaaca gaag 34

<210> 57
<211> 34
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 57
cttctgtttg cagtactcct ccagtgatgc atac 34

<210> 58
<211> 42
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 58
gagaaagtgt atgcagcact ggatgagtac tgcaaacaga ag 42

<210> 59
<211> 42
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 59
cttctgtttg cagtactcat ccagtgctgc atacactttc tc 42

<210> 60
<211> 48
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 60
catcactgga gacctacacc agaacgacgc accctgagca gcagggac 48

<210> 61
<211> 48
<212> ADN
<213> Secuencia artificial

<220>

<223> Cebador PCR

<400> 61
gtccctgctg ctcagggtgc gtcgttctgg tgtaggtctc cagtgatg 48

<210> 62
<211> 40
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 62
caaacagaag taccctgacg agccgggacg gtttgccaag 40

<210> 63
<211> 40
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 63
cttggcaaac cgtcccggct cgtcagggta cttctgtttg 40

<210> 64
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 64
ctgctacgtc ttccttccct ccggtccatt ggc 33

<210> 65
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 65
gccaatggac cggagggaag gaagacgtag cag 33

<210> 66
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 66
gccctccggt ccattagcct taagtgtcta gag 33

<210> 67
<211> 33
<212> ADN
<213> Secuencia artificial

<220>

<223> Cebador PCR

<400> 67
ctctagacac ttaaggctaa tggaccggag ggc 33

<210> 68
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 68
catctgtttt tcttcaggct cattggtgac acc 33

<210> 69
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 69
ggtgtcacca atgagcctga agaaaaacag atg 33

<210> 70
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 70
agctcattgg tgacgtcccc atcgacacct tcc 33

<210> 71
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 71
ggaaggtgtc gatggggacg tcaccaatga gct 33

<210> 72
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador PCR

<400> 72
atggagatgc ttgagtctcc ccatcaactg gcc 33

<210> 73
<211> 33
<212> ADN
<213> Secuencia artificial

<220>

<223> Cebador PCR

<400> 73
ggccagttga tggggagact caagcatctc cat 33

<210> 74
<211> 1586
<212> ADN
<213> Bamecia argentifoli

<400> 74

gaattcgcgg ccgctcgcaa acttccgtac ctctcacccc ctcgccagga ccccccgcca 60
accagttcac cgtcatctcc tccaatggat actcatcccc catgtcttcg ggcagctacg 120
acccttatag tcccaccaat ggaagaatag ggaaagaaga gctttcgccg gcgaatagtc 180
tgaacgggta caacgtggat agctgcgatg cgtcgcggaa gaagaaggga ggaacgggtc 240
ggcagcagga ggagctgtgt ctcgtctgcg gggaccgcgc ctccggctac cactacaacg 300
ccctcacctg cgaaggctgc aagggcttct tccgtcggag catcaccaag aatgccgtct 360
accagtgtaa atatggaaat aattgtgaaa ttgacatgta catgaggcga aaatgccaag 420
agtgtcgtct caagaagtgt ctcagcgttg gcatgaggcc agaatgtgta gttcccgaat 480
tccagtgtgc tgtgaagcga aaagagaaaa aagcgcaaaa ggacaaagat aaacctaact 540
caacgacgag ttgttctcca gatggaatca aacaagagat agatcctcaa aggctggata 600
cagattcgca gctattgtct gtaaatggag ttaaacccat 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
taatgatgtt tagaatggca aggaggtatg atgctgaaac agattcgata ttgtttgcaa 960
ctaaccagcc gtatacgaga gaatcataca ctgtagctgg catgggtgat actgtggagg 1020
atctgctccg attttgtcga catatgtgtg ccatgaaagt cgataacgca gaatatgctc 1080
ttctcactgc cattgtaatt ttttcagaac gaccatctct aagtgaaggc tggaaggttg 1140
agaagattca agaaatttac atagaagcat taaaagcata tgttgaaaat cgaaggaaac 1200
catatgcaac aaccattttt gctaagttac tatctgtttt aactgaacta cgaacattag 1260
ggaatatgaa ttcagaaaca tgcttctcat tgaagctgaa gaatagaaag gtgccatcct 1320
tcctcgagga gatttgggat gttgtttcat aaacagtctt acctcaattc catgttactt 1380
ttcatatttg atttatctca gcaggtggct cagtacttat cctcacatta ctgagctcac 1440
ggtatgctca tacaattata acttgtaata tcatatcggt gatgacaaat ttgttacaat 1500
attctttgtt accttaacac aatgttgatc tcataatgat gtatgaattt ttctgttttt 1560
gcaaaaaaaa aagcggccgc gaattc 1586

<210> 75
<211> 1109
<212> ADN
<213> Nephotetix cincticeps

<400> 75

caggaggagc tctgcctgtt 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 gagtacgccc agtttcggag tcaaacccct cagtccagaa 420

caagaggagc tcatccacag gctcgtctac ttccagaacg agtacgaaca ccctgccgag 480

gaggatctca agcggatcga gaacctcccc tgtgacgacg atgacccgtg tgatgttcgc 540

tacaaacaca ttacggagat cacaatactc acagtccagc tcatcgtgga gtttgcgaaa 600

aaactgcctg gtttcgacaa actactgaga gaggaccaga tcgtgttgct caaggcgtgt 660

tcgagcgagg tgatgatgct gcggatggcg cggaggtacg acgtccagac agactcgatc 720

ctgttcgcca acaaccagcc gtacacgcga gagtcgtaca cgatggcagg cgtggggaa 780

gtcatcgaag atctgctgcg gttcggccga ctcatgtgct ccatgaaggt ggacaatgcc 840

gagtatgctc tgctcacggc catcgtcatc ttctccgagc ggccgaacct ggcggaagga 900

tggaaggttg agaagatcca ggagatctac ctggaggcgc tcaagtccta cgtggacaac 960

cgagtgaaac ctcgcagtcc gaccatcttc gccaaactgc tctccgttct caccgagctg 1020

cgaacactcg gcaaccagaa ctccgagatg tgcttctcgt taaactacgc aaccgcaaac 1080

atgccaccgt tcctcgaaga aatctggga 1109

REIVINDICACIONES

1. Un sistema de modulación de expresión de genes que comprende

a) un primer casete de expresión de genes que es capaz de expresarse en una célula huésped que comprende un polinucleótido que codifica un primer polipéptido que comprende:

i) un dominio de transactivación;
ii) un dominio de unión al ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse; y
iii) un domino de unión a ligandos de receptores nucleares del Grupo B, en el que el domino de unión a ligandos de receptores nucleares del Grupo B se selecciona del grupo que consiste en un receptor X retinoide, una proteína de unión a la región II de H-2 (H-2RIIBP), un co-regulador-1 de Receptores Nucleares (RCoR-1), una proteína ultra espiráculo, un receptor nuclear 2C1 y un domino de unión a ligando del factor 1 coriónico (CF-1), y en el que los restos de aminoácidos que se encuentran en una posición equivalente o análoga a la de los restos de aminoácidos 481, 482 y 483 de la SEC ID Nº: 2 hacen que el sistema de modulación de expresión de genes presente una inducción de plegamiento y una magnitud de inducción mejoradas, en comparación con un sistema de modulación de expresión de genes que contiene el dominio de unión a ligandos de receptores nucleares del Grupo B de tipo silvestre correspondiente y

b) un segundo casete de expresión de genes que es capaz de expresarse en una célula huésped que comprende un polinucleótido que codifica un segundo polipéptido que comprende un dominio de unión a ligandos de receptores nucleares seleccionado del grupo que consiste en un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un receptor 1 nuclear de hormonas esteroideas, una proteína 15 que interacciona con el receptor X retinoide, un receptor hepático X β, una proteína similar al receptor de hormonas esteroideas, un receptor hepático X, un receptor hepático X α, un receptor X farnesoide, una proteína 14 que interacciona con receptores y un dominio de unión a ligandos de receptores de farnesol.

2. Un sistema de modulación de expresión de genes que comprende

a) un primer casete de expresión de genes que es capaz de expresarse en una célula huésped que comprende un polinucleótido que codifica un primer polipéptido que comprende:

i) un dominio de unión al ADN que reconoce un elemento de respuesta asociado a un gen cuya expresión va a modularse; y
ii) un dominio de unión a ligandos de receptores nucleares; y

b) un segundo casete de expresión de genes que es capaz de expresarse en la célula huésped que comprende un polinucleótido que codifica un segundo polipéptido que comprende:

i) un dominio de transactivación; y
ii) un domino de unión a ligandos de receptores nucleares, en el que uno de los dominios de unión a ligandos de receptores nucleares es un dominio de unión a ligandos de receptores nucleares del Grupo B seleccionado del grupo que consiste en un receptor X retinoide, una proteína de unión a la región II de H-2 (H-2RIIBP), un co-regulador-1 de Receptores Nucleares (RCoR-1), una proteína ultra-espiráculo, un receptor nuclear 2C1 y un domino de unión a ligando del factor 1 coriónico (CF-1) y en el que los restos de aminoácidos que se encuentran en una posición equivalente o análoga a la de los restos de aminoácidos 481, 482 y 483 de la SEC ID Nº: 2 hacen que el sistema de modulación de expresión de genes presente una inducción de plegamiento y una magnitud de inducción mejoradas en comparación con un sistema de modulación de expresión de genes que contiene el dominio de unión a ligandos de receptores nucleares del Grupo B de tipo silvestre correspondiente y

3. El sistema de modulación de expresión de genes de acuerdo con la reivindicación 1 o 2, en el que el dominio de unión a ligandos de receptores nucleares del Grupo B es un dominio de unión a ligandos de un receptor X retinoide.

4. El sistema de modulación de expresión de genes de acuerdo con la reivindicación 3, en el que el polinucleótido codifica un ácido aspártico en una posición equivalente o análoga a la del resto de aminoácido 481 en la SEC ID Nº: 2, un ácido glutámico en una posición equivalente o análoga a la del resto de aminoácido 482 en la SEC ID Nº: 2 y una prolina en una posición equivalente o análoga a la del resto de aminoácido 483 de la SEC ID Nº: 2.

5. El sistema de modulación de expresión de genes de acuerdo con la reivindicación 1 o la reivindicación 2, en el que el domino de unión al ADN se selecciona del grupo que consiste en un dominio de unión al ADN del receptor de ecdisona, un dominio de unión al ADN de GAL4 y un dominio de unión al ADN LexA.

6. El sistema de modulación de expresión de genes de acuerdo con la reivindicación 1 o la reivindicación 2, en el que el dominio de transactivación se selecciona del grupo que consiste en un dominio de transactivación del receptor de ecdisona, un dominio de transactivación de VP16, un domino de transactivación del activador ácido B42 y un dominio de transactivación de p65.

7. Un casete de expresión de genes que comprende un polinucleótido que codifica un polipéptido seleccionado del grupo que consiste en:

a) un polipéptido que comprende un dominio de transactivación, un dominio de unión al ADN y un dominio de unión a ligandos de receptores nucleares del Grupo B,
b) un polipéptido que comprende un dominio de unión al ADN y un dominio de unión a ligandos de receptores nucleares del Grupo B, y
c) un polipéptido que comprende un dominio de transactivación y un dominio de unión a ligandos de receptores nucleares del Grupo B; en el que el dominio de unión a ligandos de receptores nucleares del Grupo B se selecciona del grupo que consiste en un receptor X retinoide, una proteína de unión a la región II de H-2 (H-2RIIBP), un co-regulador-1 de Receptores Nucleares (RCoR-1), una proteína ultra-espiráculo, un receptor nuclear 2C1 y un domino de unión al ligando del factor 1 coriónico (CF-1) y en el que los restos de aminoácidos en dicho dominio de unión a receptores nucleares del Grupo B que se encuentran en una posición equivalente o análoga a la de los restos de aminoácidos 481, 482 y 483 de la SEC ID Nº: 2 hacen que el polipéptido presente una inducción de plegamiento y una magnitud de inducción mejoradas en comparación con un polipéptido que contiene el dominio de unión a ligandos de receptores nucleares del Grupo B de tipo silvestre correspondiente.

8. Un polinucleótido aislado que codifica un domino de unión a ligandos de receptores nucleares del Grupo B, en el que el polinucleótido aislado codifica los aminoácidos que se encuentran en una posición equivalente o análoga a la de los restos de aminoácidos 481, 482 y 483 de la SEC ID Nº: 2 que dan como resultado una inducción de plegamiento y una magnitud de inducción mejoradas en comparación con el dominio de unión a ligandos de receptores nucleares del Grupo B de tipo silvestre de la SEC ID Nº: 2.

9. El polinucleótido aislado de acuerdo con la reivindicación 8, en el que el polinucleótido aislado codifica un ácido aspártico en una posición equivalente o análoga a la del resto de aminoácido 481 en la SEC ID Nº: 2, un ácido glutámico en una posición equivalente o análoga a la del resto de aminoácido 482 en la SEC ID Nº: 2 y una prolina en una posición equivalente o análoga a la del resto de aminoácido 483 de la SEC ID Nº: 2.

10. Un vector de expresión que comprende el polinucleótido aislado de la reivindicación 8 o 9 unido operativamente a un elemento regulador de la transcripción.

11. Una célula huésped aislada que comprende el vector de expresión de la reivindicación 10, en la que el elemento regulador de la transcripción es operativo en la célula huésped.

12. Un polipéptido aislado codificado por el polipéptido aislado de acuerdo con la reivindicación 8 o 9.

13. Un polipéptido aislado que comprende un domino de unión a ligandos de receptores nucleares del Grupo B, en el que los restos de aminoácidos que se encuentran en una posición equivalente o análoga a 481, 482 y 483 de la SEC ID Nº: 2 hacen que el polipéptido presente una inducción de plegamiento y una magnitud de inducción mejoradas en comparación con un polipéptido que contiene el dominio de unión a ligandos de receptores nucleares del Grupo B de tipo silvestre correspondiente.

14. El polipéptido aislado de acuerdo con la reivindicación 13, en el que el dominio de unión a ligandos de receptores nucleares del Grupo B es un dominio de unión a ligandos de un receptor X retinoide.

15. El polipéptido aislado de acuerdo con la reivindicación 14, en el que el polipéptido contiene ácido aspártico en una posición equivalente o análoga a la del resto de aminoácido 481 en la SEC ID Nº: 2, ácido glutámico en una posición equivalente o análoga a la del resto de aminoácido 482 en la SEC ID Nº: 2 y prolina en una posición equivalente o análoga a la del resto de aminoácido 483 de la SEC ID Nº: 2.

16. Un procedimiento para modular la expresión de un gen en una célula huésped no animal que comprende las etapas de:

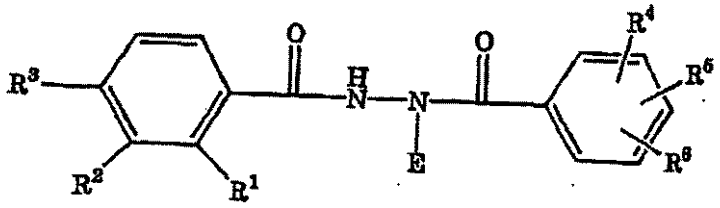
a) introducir en la célula huésped el sistema de modulación de expresión de genes de acuerdo con una cualquiera de las reivindicaciones 1 a 6; y
b) introducir un ligando en la célula huésped;

en el que el gen a modular es un componente de un casete de expresión de genes que comprende:

i) un elemento de respuesta reconocido por el dominio de unión al ADN
ii) un promotor que es activado por el dominio de transactivación; y
iii) un gen cuya expresión va a modularse; mediante el cual después de la introducción del ligando en la célula huésped, se modula la expresión del gen de b) iii).

17. El procedimiento de acuerdo con la reivindicación 16, en el que el ligando es

a) un compuesto de fórmula:



en la que:

E es un alquilo (C_4-C_6) que contiene un carbono terciario o un cianoalquilo (C_3-C_5) que contiene un carbono terciario;
R^1 es H, Me, Et, i-Pr, F, formilo, CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CH_2OMe, CH_2CN, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, OH, OMe, OEt, ciclopropilo, CF_2CF_3, CH=CHCN, alilo, azido, SCN, $SCHF_2$;
R_2 es H, Me, Et, n-Pr, i-Pr, formilo CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CH_2OMe, CH_2CN, CN, C˚CH, 1-propinilo, 2-propinilo, vinilo, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc, NMe_2, NEt_2, SMe, SEt, $SOCF_3$, OCF_2CF_2H, COEt, ciclopropilo, CF_2CF_3, CH=CHCN, alilo, azido, OCF_3, $OCHF_2$, O-i-Pr, SCN, $SCHF_2$, SOMe, NH-CN, o está unido a R_3 y a los carbonos fenilo a los cuales R_2 y R_3 se unen para formar un etilendioxi, un anillo dihidrofurilo con el oxígeno adyacente a un carbono fenilo o un anillo dihidropirilo con el oxígeno adyacente a un carbono fenilo;
R^3 es H, Et, o está unido a R^2 y a los carbonos fenilo a los cuales R^2 y R^3 se unen para formar un etilendioxi, un anillo dihidrofurilo con el oxígeno adyacente a un carbono fenilo o un anillo dihidropirilo con el oxígeno adyacente a un carbono fenilo;
R^4, R^5 y R^6 son independientemente H, Me, Et, F, Cl, Br, formilo, CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, OMe, OEt, SMe o SEt;

b) una ecdisona, 20-hidroxiecdisona, ponasterona A o muristerona A;
c) un oxisterol, un 22(R) hidroxicolesterol, 24(S) hidroxicolesterol, 25-epoxicolesterol, T0901317, 5-alfa-6-alfa-epoxicolesterol-3-sulfato, 7-cetocolesterol-3-sulfato, farnesol, un ácido biliar, un éster 1,1-bifosfonato o una hormona Juvenil III; o
d) un ácido 9-cis-retinoico, ácido 4-(1-(3, 5,5,8,8-pentametil-5,6,7,8-tetrahidro-2-naftil)-etenil) benzoico (3-metil-TTEB), ácido ((E)-2)2-(5,6,7,8-tetra-hidro-3,5,5,8,8-pentametil-2-naftil)propen-1-il)-4-tiofenocarboxílico), 2-(5,6,7,8 tetra-hidro-3,5,5,8,8-tetrametil-2-naftil)-2-(carboxifenil)-1,3-dioxolano, ácido 4-(5H-2,3-(2,5-dimetil-2,5-hexano)-5-metil-dibenzo (b,e) ácido (1,4)diazepin-11-il)-benzoico (HX600) o sus análogos tiadiazepínicos, ácido 3,7,11,15-tetrametil hexadecanoico (ácido fitánico), ácido 6-(1-(3,5,5,8,8-pentametil-5,6,7,8-tetrahidronaftalen-2-il)ciclopropil) nicotínico, 2-(4-carboxifenil)-2-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)-1,3-ditiano o ácido 4-(2-metil)-1-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)propenil) benzoico.

18. El procedimiento de acuerdo con la reivindicación 16 o 17, en el que el procedimiento comprende adicionalmente introducir en la célula huésped un segundo ligando, en el que el segundo ligando es el ácido 9-cis retinoico o un análogo sintético de un ácido retinoico.

19. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 16 o 18, en el que la célula huésped no animal es una célula huésped bacteriana, una célula huésped fúngica o una célula huésped vegetal.

20. Una célula huésped aislada que comprende el sistema de modulación de expresión de genes de acuerdo con una cualquiera de las reivindicaciones 1 a 6.

21. Un organismo no humano que comprende la célula huésped de la reivindicación 20.

22. El sistema de modulación de expresión de genes de acuerdo con la reivindicación 1 o 2, en el que dicho sistema presenta una sensibilidad mejorada a ligandos no esteroideos en comparación con un sistema de modulación de expresión de genes que contiene el dominio de unión a ligandos de receptores nucleares del Grupo B de tipo silvestre de la SEC ID Nº:2.

23. El casete de expresión de genes de acuerdo con la reivindicación 7, en el que dicho polipéptido presenta sensibilidad mejorada a ligandos no esteroideos en comparación con un polipéptido que contiene el dominio de unión a ligandos de receptores nucleares del Grupo B de tipo silvestre correspondiente.

24. El polipéptido aislado de acuerdo con la reivindicación 8, en el que el polinucleótido aislado codifica aminoácidos que se encuentran en una posición equivalente o análoga a la de los restos de aminoácidos 481, 482 y 483 de la SEC ID Nº: 2 que dan como resultado una sensibilidad mejorada a ligandos no esteroideos en comparación con el dominio de unión a ligandos de receptores nucleares del Grupo B de tipo silvestre correspondiente.

25. El polipéptido aislado de acuerdo con la reivindicación 13, en el que dicho polipéptido presenta una sensibilidad mejorada a ligandos no esteroideos en comparación con un polipéptido que contiene el dominio de unión a ligandos

de receptores nucleares del Grupo B de tipo silvestre correspondiente.

26. Un procedimiento de modulación *in vitro* de la expresión de un gen en una célula huésped que comprende las etapas de:

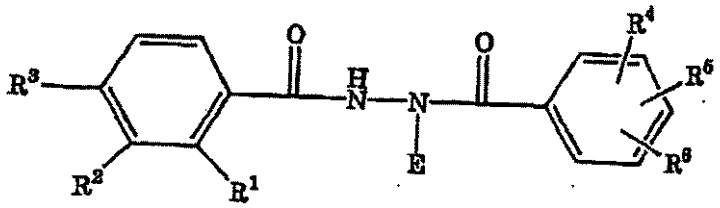
a) introducir en la célula huésped el sistema de modulación de expresión de genes de acuerdo con una cualquiera de las reivindicaciones 1 a 6 y
b) introducir un ligando en la célula huésped;

en el que el gen a modular es un componente de un casete de expresión de genes que comprende:

i) un elemento de respuesta reconocido por el dominio de unión al ADN;
ii) un promotor que es activado por el dominio de transactivación; y
iii) un gen cuya expresión va a modularse; mediante el cual después de la introducción del ligando en la célula huésped, se modula la expresión del gen de b) iii).

27. El procedimiento de acuerdo con la reivindicación 26, en el que el ligando es

a) un compuesto de fórmula:



en la que:

E es un alquilo (C_4-C_6) que contiene un carbono terciario o un cianoalquilo (C_3-C_5) que contiene un carbono terciario;
R^1 es H, Me, Et, i-Pr, F, formilo, CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CH_2OMe, CH_2CN, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, OH, OMe, OEt, ciclopropilo, CF_2CF_3, CH=CHCN, alilo, azido, SCN, $SCHF_2$;
R_2 es H, Me, Et, n-Pr, i-Pr, formilo CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CH_2OMe, CH_2CN, CN, C˚CH, 1-propinilo, 2-propinilo, vinilo, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc, NMe_2, NEt_2, SMe, SEt, $SOCF_3$, OCF_2CF_2H, COEt, ciclopropilo, CF_2CF_3, CH=CHCN, alilo, azido, OCF_3, $OCHF_2$, O-i-Pr, SCN, $SCHF_2$, SOMe, NH-CN, o está unido a R_3 y a los carbonos fenilo a los cuales R_2 y R_3 se unen para formar un etilendioxi, un anillo dihidrofurilo con el oxígeno adyacente a un carbono fenilo o un anillo dihidropirilo con el oxígeno adyacente a un carbono fenilo;
R^3 es H, Et, o está unido a R^2 y a los carbonos fenilo a los cuales R^2 y R^3 se unen para formar un etilendioxi, un anillo dihidrofurilo con el oxígeno adyacente a un carbono fenilo o un anillo dihidropirilo con el oxígeno adyacente a un carbono fenilo;
R^4, R^5 y R^6 son independientemente H, Me, Et, F, Cl, Br, formilo, CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, OMe, OEt, SMe, o SEt;

b) una ecdisona, 20-hidroxiecdisona, ponasterona A o muristerona A;
c) un oxisterol, un 22(R) hidroxicolesterol, 24(S) hidroxicolesterol, 25-epoxicolesterol, T0901317, 5-alfa-6-alfa-epoxicolesterol-3-sulfato, 7-cetocolesterol-3-sulfato, farnesol, un ácido biliar, un éster 1,1-bifosfonato o una hormona Juvenil III; o
d) un ácido 9-cis-retinoico, ácido 4-(1-(3, 5,5,8,8-pentametil-5,6,7,8-tetrahidro-2-naftil)-etenil) benzoico (3-metil-TTEB), ácido ((E)-2)2-(5,6,7,8-tetra-hidro-3,5,5,8,8-pentametil-2-naftil)propen-1-il)-4-tiofenocarboxílico), 2-(5,6,7,8 tetra-hidro-3,5,5,8,8-tetrametil-2-naftil)-2-(carboxifenil)-1,3-dioxolano, ácido 4-(5H-2,3-(2,5-dimetil-2,5-hexano)-5-metil-dibenzo (b,e) ácido (1,4)diazepin-11-il)-benzoico (HX600) o sus análogos tiadiazepínicos, ácido 3,7,11,15-tetrametil hexadecanoico (ácido fitánico), ácido 6-(1-(3,5,5,8,8-pentametil-5,6,7,8-tetrahidronaftalen-2-il)ciclopropil) nicotínico, 2-(4-carboxifenil)-2-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)-1,3-ditiano o ácido 4-(2-metil)-1-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)propenil) benzoico.

28. El procedimiento de acuerdo con la reivindicación 26 o 27, en el que el procedimiento comprende adicionalmente introducir en la célula huésped un segundo ligando, en el que el segundo ligando es el ácido 9-cis retinoico o un análogo sintético de un ácido retinoico.

29. Un sistema de expresión de genes de acuerdo con una cualquiera de las reivindicaciones 1 a 6 junto con un ligando para su uso en el tratamiento de una enfermedad o trastorno;
en el que el gen a modular es un componente de un casete de expresión de genes que comprende:

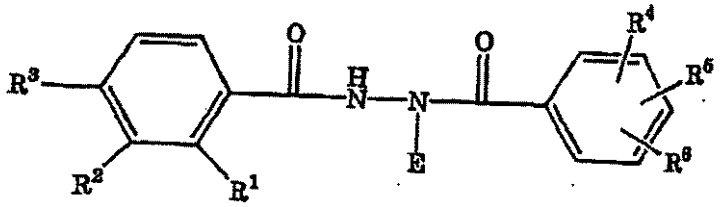
i) un elemento de respuesta reconocido por el dominio de unión al ADN;
ii) un promotor que es activado por el dominio de transactivación; y
iii) un gen cuya expresión va a modularse; mediante el cual después del contacto con el ligando, se modula la expresión del gen de iii).

30. Uso del sistema de modulación de expresión de genes de acuerdo con una cualquiera de las reivindicaciones 1 a 6 y de un ligando en la fabricación de un medicamento para el tratamiento de una enfermedad o trastorno;
en el que el gen a modular es un componente de un casete de expresión de genes que comprende:

i) un elemento de respuesta reconocido por el dominio de unión al ADN;
ii) un promotor que es activado por el dominio de transactivación; y
iii) un gen cuya expresión va a modularse; mediante el cual después del contacto con el ligando, se modula la expresión del gen de iii).

31. El sistema de expresión de genes junto con el ligando para su uso de acuerdo con la reivindicación 29, o el uso de acuerdo con la reivindicación 30 en el que el ligando es

a) un compuesto de fórmula:



en la que:
E es un alquilo (C_4-C_6) que contiene un carbono terciario o un cianoalquilo (C_3-C_5) que contiene un carbono terciario;
R^1 es H, Me, Et, i-Pr, F, formilo, CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CH_2OMe, CH_2CN, CN, C=CH, 1-propinilo, 2-propinilo, vinilo, OH, OMe, OEt, ciclopropilo, CF_2CF_3, CH=CHCN, alilo, azido, SCN, $SCHF_2$;
R_2 es H, Me, Et, n-Pr, i-Pr, formilo CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CH_2OMe, CH_2CN, CN, C˚CH, 1-propinilo, 2-propinilo, vinilo, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc, NMe_2, NEt_2, SMe, SEt, $SOCF_3$, OCF_2CF_2H, COEt, ciclopropilo, CF_2CF_3, CH=CHCN, alilo, azido, OCF_3, $OCHF_2$, O-i-Pr, SCN, $SCHF_2$, SOMe, NH-CN, o está unido a R_3 y a los carbonos fenilo a los cuales R_2 y R_3 se unen para formar un etilendioxi, un anillo dihidrofurilo con el oxígeno adyacente a un carbono fenilo o un anillo dihidropirilo con el oxígeno adyacente a un carbono fenilo;
R^3 es H, Et, o está unido a R^2 y a los carbonos fenilo a los cuales R^2 y R^3 se unen para formar un etilendioxi, un anillo dihidrofurilo con el oxígeno adyacente a un carbono fenilo o un anillo dihidropirilo con el oxígeno adyacente a un carbono fenilo;
R^4, R^5 y R^6 son independientemente H, Me, Et, F, Cl, Br, formilo, CF_3, CHF_2, $CHCl_2$, CH_2F, CH_2Cl, CH_2OH, CN, C≡CH, 1-propinilo, 2-propinilo, vinilo, OMe, OEt, SMe, o SEt;

b) una ecdisona, 20-hidroxiecdisona, ponasterona A o muristerona A;
c) un oxisterol, un 22(R) hidroxicolesterol, 24(S) hidroxicolesterol, 25-epoxicolesterol, T0901317, 5-alfa-6-alfa-epoxicolesterol-3-sulfato, 7-cetocolesterol-3-sulfato, farnesol, un ácido biliar, un éster 1,1-bifosfonato o una hormona Juvenil III; o
d) un ácido 9-cis-retinoico, ácido 4-(1-(3, 5,5,8,8-pentametil-5,6,7,8-tetrahidro-2-naftil)-etenil) benzoico (3-metil-TTEB), ácido ((E)-2)2-(5,6,7,8-tetra-hidro-3,5,5,8,8-pentametil-2-naftil)propen-1-il)-4-tiofenocarboxílico), 2-(5,6,7,8 tetra-hidro-3,5,5,8,8-tetrametil-2-naftil)-2-(carboxifenil)-1,3-dioxolano, ácido 4-(5H-2,3-(2,5-dimetil-2,5-hexano)-5-metil-dibenzo (b,e) ácido (1,4)diazepin-11-il)-benzoico (HX600) o sus análogos tiadiazepínicos, ácido 3,7,11,15-tetrametil hexadecanoico (ácido fitánico), ácido 6-(1-(3,5,5,8,8-pentametil-5,6,7,8-tetrahidronaftalen-2-il)ciclopropil) nicotínico, 2-(4-carboxifenil)-2-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)-1,3-ditiano o ácido 4-(2-metil)-1-(5,6,7,8-tetrahidro-5,5,8,8-tetrametil-2-naftalenil)propenil) benzoico.

Comparación de RXR, LmUSP, Quimera, RXRmutE265D, RXRmutG293S con CfEcRCDEF en células 3T3; respuesta GSE

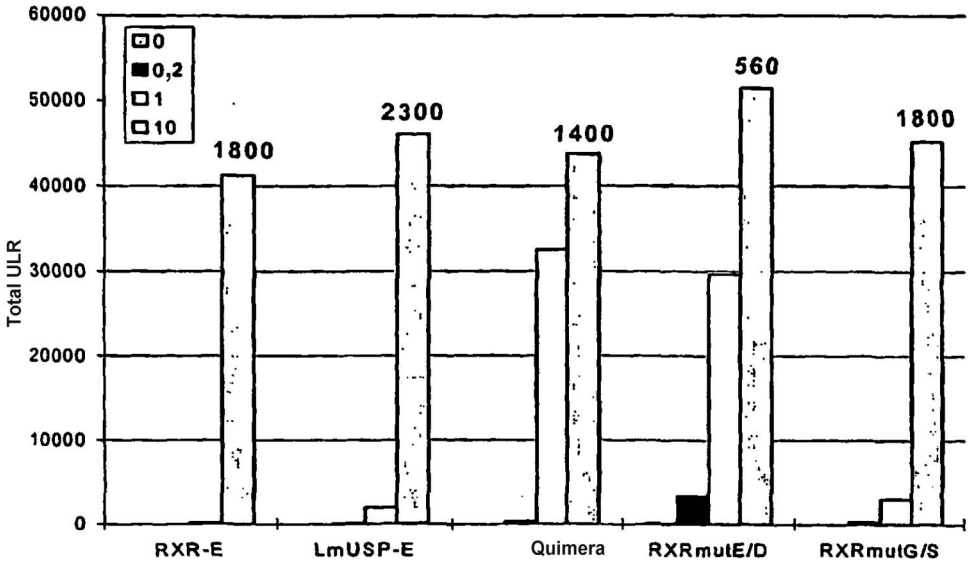


Figura 1

Comparación de respuestas de RXR-E, mutantes sencillos y dobles frente a GSE en células 3T3

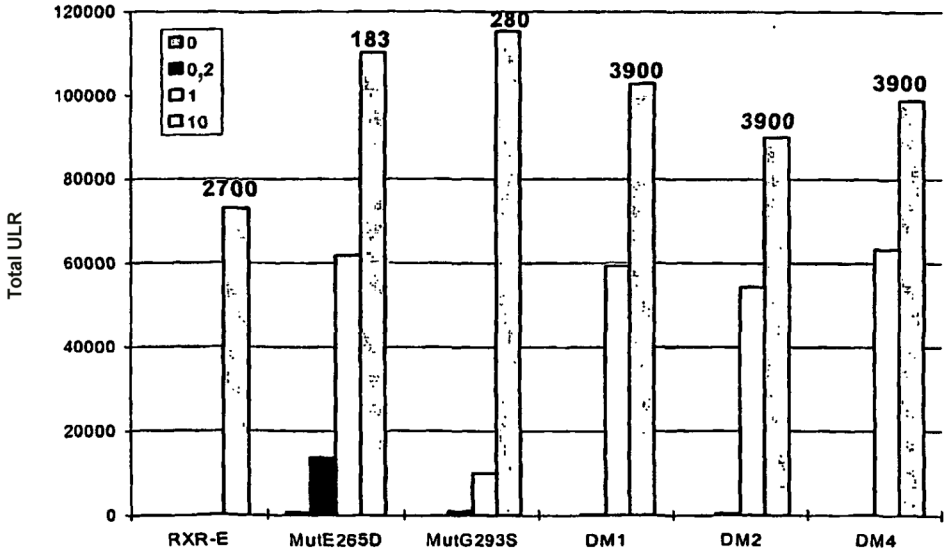


Figura 2

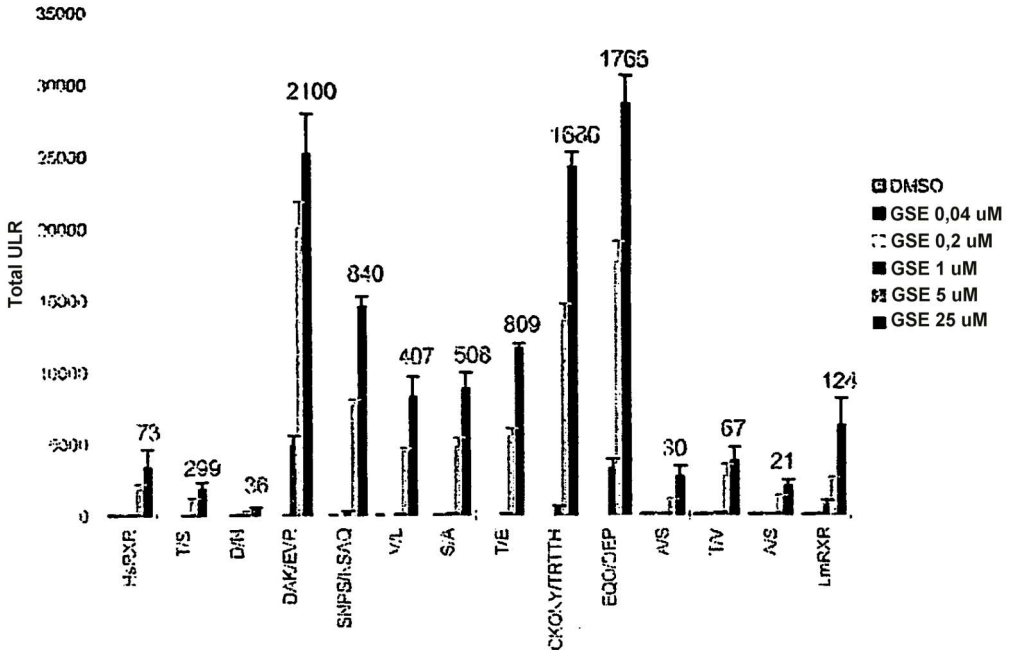


Figura 3

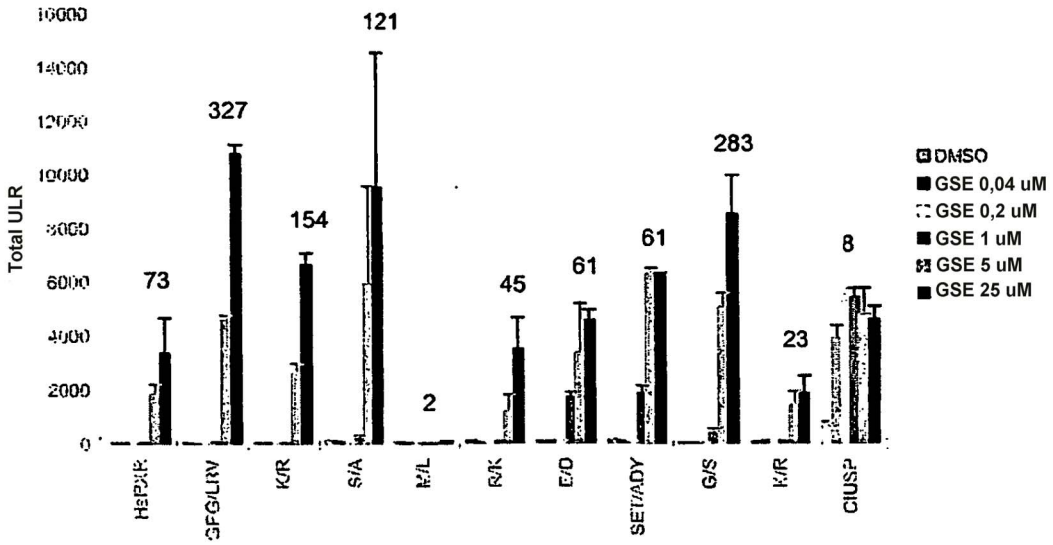


Figura 4