

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 802 524**

51 Int. Cl.:

G01N 33/566 (2006.01)

C07K 14/435 (2006.01)

C07K 14/72 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.06.2016 PCT/US2016/036777**

87 Fecha y número de publicación internacional: **15.12.2016 WO16201153**

96 Fecha de presentación y número de la solicitud europea: **09.06.2016 E 16731742 (9)**

97 Fecha y número de publicación de la concesión europea: **01.04.2020 EP 3308168**

54 Título: **Líneas celulares para el cribado de receptores de aroma y olor**

30 Prioridad:

10.06.2015 US 201562173762 P

21.09.2015 US 201562221580 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
20.01.2021

73 Titular/es:

FIRMENICH SA (100.0%)
7, Rue de la Bergère
1242 Satigny, CH

72 Inventor/es:

JEONG, HYO-YOUNG;
PFISTER, PATRICK y
ROGERS, MATTHEW

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 802 524 T3

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

DESCRIPCIÓN

Líneas celulares para el cribado de receptores de aroma y olor

Campo

- 5 El campo técnico se dirige a los receptores y ensayos de olor y aroma que se pueden usar para identificar compuestos o moduladores de olor y/o aroma. Los ensayos se dirigen más específicamente a líneas celulares diseñadas que exhiben una actividad mejorada del receptor de olor.

Antecedentes

- 10 Los olores se codifican inicialmente en el sistema olfativo periférico (es decir, la nariz) a través de interacciones entre compuestos de aroma y sabor volátiles y proteínas del receptor de olor (OR) que residen en las membranas de las neuronas receptoras olfativas del tejido epitelial olfativo. Dichas interacciones se producen de forma combinatoria específica de olor, donde cualquier OR se puede activar por múltiples olores y, por el contrario, la mayoría de los olores son capaces de activar varios OR diferentes. Para un compuesto o mezcla de olor/aroma dado, estas interacciones del receptor generan señales neurofisiológicas en el cerebro y, en última instancia, dan lugar a una percepción consciente del olor. Aproximadamente ~400 genes OR en el genoma humano se pueden activar por miles o más estímulos de olor y es la complejidad inherente de las interacciones combinatorias entre los olores y los receptores lo que permite la amplitud de sensaciones olfativas que podemos percibir. Elucidar estas interacciones puede conducir al descubrimiento de productos beneficiosos que incluyen, pero no está limitado a, contrarrestadores de malos olores que bloquean la percepción de olores desagradables, nuevos ingredientes de sabores y fragancias que reemplazan compuestos no biodegradables o tóxicos, y potenciadores de olores que limitarían nuestra dependencia en compuestos difíciles de obtener de fuentes naturales.

- 25 Existe la necesidad de, por ejemplo, pero no está limitado a, nuevos procedimientos que puedan expresar funcionalmente OR en la superficie celular para una decodificación confiable de los códigos OR. Existe una necesidad adicional de un procedimiento que exprese funcionalmente los OR en células no olfativas (por ejemplo, pero no está limitado a, líneas celulares heterólogas) que sean susceptibles a cribado de alto rendimiento con librerías de compuestos de fragancias y sabores volátiles para una caracterización completa de la actividad OR. Esto podría acelerar significativamente el descubrimiento de contrarrestadores de malos olores, moduladores de olores y nuevos compuestos de sabor o fragancia altamente deseables.

El documento WO 2015/020158 A1 describe procedimientos para el cribado de compuestos de perfume de tipo almizcle entre sustancias candidatas que utilizan un receptor olfativo específico para un perfume de almizcle.

- 30 Ciertas proteínas que se derivan de las neuronas sensoriales olfativas pueden mejorar la localización de la superficie celular de los receptores de olor en las líneas celulares no olfativas. Estas proteínas funcionan ayudando en el tráfico de los receptores de olor desde el retículo endoplásmico al aparato de Golgi y la membrana plasmática de la célula. Se ha informado que la proteína 1 que transporta el receptor (RTP1), la proteína 2 que transporta el receptor (RTP2) y la proteína 1 que aumenta la expresión del receptor (REEP1) mejoran la localización del receptor de olor en la membrana plasmática y, por lo tanto, funcionan en células no olfativas. RTP1 y RTP2 se han utilizado en células heterólogas recombinantes para promover la expresión en la superficie de los receptores de olor (WO 2013/082522 A1). Se ha informado que RTP1 es la chaperona receptora de olor más eficaz. Se ha demostrado que esta proteína actúa, en parte, al interactuar con los receptores de olor en el retículo endoplásmico. Por lo tanto, una línea celular no olfativa que sea susceptible de cribado de alto rendimiento y que contenga el gen de RTP1 se desea altamente para la decodificación integral de las interacciones combinatorias entre los olores y los receptores de olores.

- 45 Aún más deseable es una línea celular que produce consistentemente la proteína RTP1, que incluye, pero no está limitada a RTP1S, al expresar de manera estable el gen del locus endógeno RTP1. Sin embargo, las técnicas actuales para desarrollar líneas celulares que expresan establemente el gen de RTP1 endógeno implican enfoques de biología molecular ineficientes y engorrosos para la inserción de ADN en células cultivadas que de otra manera no expresarían el gen. Por lo tanto, se desea usar una técnica que evite tales enfoques para desarrollar una línea celular estable y que permita la expresión consistente de la RTP1 endógena sin la necesidad de usar procedimientos recombinantes.

- 50 CRISPR/Cas9 es una herramienta de edición genómica altamente eficiente que se utiliza para generar modificaciones genómicas precisas, como inserciones y deleciones. Por ejemplo, un uso particular de CRISPR/Cas9 permite la expresión de un gen que de otro modo podría estar silenciado en una línea celular (véase, por ejemplo, Cheng y otros, Cell Research (2013) 23: 1163-1171 y Maeder y otros, Nature Methods (2013) 10: 977-979). Al incorporar un promotor transcripcional corriente arriba del gen, una línea celular puede expresar un gen endógeno que de otro modo estaría inactivo.

Sumario

- 55 La invención se refiere a una célula que comprende un ADN donante que se introduce en el sitio genómico objetivo corriente arriba de un locus del gen de RTP1 endógeno, en el que el ADN donante comprende un promotor en el que

el promotor impulsa la expresión del gen de RTP1, y en el que la célula se deriva de una línea celular HEK293 o HEK293T.

También se divulga en la presente memoria una línea celular no olfativa con función mejorada del receptor de olor que comprende un gen de RTP1 endógeno activado.

- 5 También se divulga en la presente memoria una línea celular no olfativa que comprende un gen de RTP1 endógeno activado dentro de la célula que además expresa una proteína RTP1.

La invención se refiere además a un procedimiento para activar un gen de RTP1 endógeno en una célula eucariota, en el que la célula se deriva de una línea celular HEK293 o HEK293T, que comprende:

- a. introducir un ARN guía complementario a un sitio genómico objetivo corriente arriba del gen de RTP1;
- 10 b. introducir una proteína nucleasa Cas para formar un complejo con el ARN guía; y
- c. utilizar el complejo objetivo genómico ARN guía/Cas9 para administrar los elementos de activación del gen para el gen de RTP1 específicamente.

La invención se refiere además a un procedimiento para identificar un compuesto o mezcla de compuestos que activan, imitan, bloquean, inhiben, modulan y/o aumentan la actividad de un receptor olfativo en una célula no olfativa en la que la célula se deriva de una línea celular HEK293 o HEK293T y comprende un ADN donante que se introduce en el sitio objetivo genómico corriente arriba de un locus del gen de RTP1 endógeno, en el que el ADN donante comprende un promotor en el que el promotor impulsa la expresión del gen de RTP1, en el que el procedimiento comprende además:

- a. poner en contacto el receptor, o una quimera o fragmento del mismo con un compuesto o mezcla de compuestos que activa, imita, bloquea, inhibe, modula y/o mejora el receptor; y
- 20 b. determinar si el compuesto tiene un efecto sobre la actividad del receptor.

Descripción de las figuras

La **Figura 1** muestra un esquema del locus del gen de RTP1 endógeno y el sitio objetivo específico que se utilizó para la edición del genoma.

- 25 La **Figura 2** muestra un esquema del procedimiento de inserción del promotor CMV.

La **Figura 3** muestra un esquema de los alelos tipo salvaje y recombinados y las regiones de genoma correspondientes.

La **Figura 4** muestra la caracterización de la integración del promotor CMV en la línea celular que se diseñó.

La **Figura 5** muestra la caracterización de la expresión del ARNm de RTP1 en la línea celular que se diseñó.

- 30 La **Figura 6** muestra la caracterización de la expresión de la proteína RTP1 en la línea celular que se diseñó.

La **Figura 7** muestra una curva dosis-respuesta del receptor de ratón Olfr741 en presencia de concentraciones crecientes de indol.

La **Figura 8** muestra una curva dosis-respuesta del receptor de ratón Olfr742 en presencia de concentraciones crecientes de indol.

- 35 La **Figura 9** muestra una curva dosis-respuesta del receptor de ratón Olfr96 en presencia de concentraciones crecientes de vulcanólido.

La **Figura 10** muestra una curva dosis-respuesta del receptor humano OR11A1 en presencia de concentraciones crecientes de vulcanólido.

- 40 La **Figura 11** muestra una curva dosis-respuesta del receptor de ratón Olfr740 en presencia de concentraciones crecientes de indol.

La **Figura 12** muestra una curva dosis-respuesta del receptor humano OR1A1 en presencia de concentraciones crecientes de carvona- (-).

Descripción detallada

- 45 Para las descripciones en la presente memoria y las reivindicaciones adjuntas, el uso de "o" significa "y/o" a menos que se indique lo contrario. De forma similar, "comprende", "comprenden", "que comprende", "incluye", "incluyen" e "incluyendo" son intercambiables y no pretenden ser limitantes.

Se debe entender además que cuando las descripciones de diversas realizaciones usan el término "que comprende", los expertos en la técnica entenderán que, en algunos casos específicos, una realización se puede describir alternativamente usando el lenguaje "que consiste esencialmente en" o que "consiste en".

En una realización, se proporciona una célula que comprende un ácido nucleico que codifica un receptor de olor.

- 5 En una realización adicional, se proporciona una célula que comprende un ácido nucleico que codifica un receptor de olor que se selecciona del grupo que consiste en Olfr741, Olfr742, Olfr96, Olfr740, OR1A1 y OR11A1.

En otra realización adicional en la presente memoria se proporciona una célula que comprende un promotor constitutivo corriente arriba de un locus del gen de RTP1 endógeno de tal manera que el promotor impulsa la expresión del gen de RTP1 endógeno.

- 10 En las neuronas sensoriales olfativas de ratón, el transcripto RTP1 contiene dos sitios de inicio de traducción alternativos que pueden conducir a dos formas distintas de la proteína RTP1: una versión larga (RTP1L) y una versión corta (RTP1S). Sin embargo, es la proteína RTP1S que se expresa predominantemente en las neuronas olfativas del ratón. Además, las células no olfativas (por ejemplo, pero no está limitado a HEK293T) que expresan de forma heteróloga la secuencia de codificación completa de RTP1 expresan predominantemente RTP1L, aunque la secuencia de codificación de RTP1S está contenida dentro de RTP1L. Sin embargo, se prefiere RTP1S para el cribado del receptor de olor en células no olfativas, ya que se sabe que RTP1S supera fuertemente a RTP1L con respecto a la expresión de OR en la superficie celular. Sorprendentemente, hemos encontrado que la activación endógena del gen de RTP1 completo conduce preferentemente a la expresión de RTP1S.

- 20 En una realización adicional, en la presente memoria, se proporciona una célula que comprende un promotor constitutivo corriente arriba de un locus del gen de RTP1 endógeno que impulsa la expresión de la versión corta del gen de RTP1, el RTP1S.

En la presente memoria se proporciona una línea celular no olfativa que comprende un gen de RTP1 endógeno que se activa dentro de la célula que además expresa la versión corta de una proteína RTP1, la RTP1S.

En otra realización más, el promotor constitutivo se selecciona del grupo que consiste en CMV, PGK, EF1a y SV40.

- 25 En una realización, el promotor es CMV, que se origina a partir del Citomegalovirus.

En una realización, en la presente memoria se proporciona una célula que comprende un promotor inducible corriente arriba de un locus del gen de RTP1 endógeno de modo que el promotor impulsa la expresión del gen de RTP1 cuando está presente el operador correspondiente.

- 30 En una realización adicional, el promotor inducible es un promotor del Elemento de Respuesta a Tetraciclina (TRE) que se induce por administración de tetraciclina (o su análogo doxiciclina).

En una realización, la proteína nucleasa Cas es una proteína Cas9.

En una realización adicional, la proteína Cas9 se selecciona del grupo que consiste en Cas9, dCas9 (Cas9 desactivado) y Cas9 nicaa.

- 35 En una realización, la secuencia de ARN guía (ARNg) se puede diseñar en base a reglas de última generación (Doench y otros, Nat Biotech (2014)) y herramientas de diseño de ARN guía disponibles públicamente para una orientación genómica eficiente (por ejemplo, www.blueheronbio.com/external/tools/gRNASrc.jsp). Se pueden usar los brazos homólogos que tienen 800 pb de largo en cada lado de la ruptura de ADN de doble cadena específica que se genera por Cas9. Es útil revisar cuidadosamente el sitio de integración del promotor CMV para evitar un sitio de inicio de traducción no deseado antes del sitio de inicio RTP1 endógeno.

- 40 Preferentemente, la línea celular original que se utiliza para la ingeniería CRISPR/Cas9 debería ser de origen mamífero y portar el locus del gen de RTP1. Dichas líneas celulares incluyen, pero no se restringen a, HEK293, HEK293T, HeLa, CHO, OP6, HeLa-S3, HEKn, HEKa, PC-3, Calu1, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, epitelio de riñón de mono BS-C-1, fibroblastos de embriones de ratón BALB/3T3, 3T3 Swiss, 3T3-L1, fibroblastos fetales humanos 132-d5; 10.1 fibroblastos de ratón, 293-T, 3T3, BHK, BHK-21, BR 293, BxPC3, C3H-IOT1/2, C6/36, Cal-27, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr -/-, COS-7, HL-60, LNCap, MCF-7, MCF-10A, MDCK II, SkBr3, células Vero, células olfativas inmortalizadas, células gustativas inmortalizadas y variedades transgénicas de las mismas. Las líneas celulares están disponibles en una variedad de fuentes conocidas por los expertos en la técnica (véase, por ejemplo, la Colección Americana de Cultivos Tipo (Manassus, Virginia).

- 50 Por lo tanto, se selecciona una línea celular estable del grupo que consiste en HEK293, HEK293T, HeLa, CHO, OP6, HeLa-S3, HEKn, HEKa, PC-3, Calu1, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, epitelio de riñón de mono BS-C-1, fibroblastos de embriones de ratón BALB/3T3, 3T3 Swiss, 3T3-L1, fibroblastos fetales humanos 132-d5; fibroblastos de ratón 10.1, 293-T, 3T3, BHK, BHK-21, BR 293, BxPC3, C3H-IOT1/2, C6/36, Cal-27, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr -/-, COS-7, HL-60, LNCap, MCF-7, MCF-10A, MDCK II, SkBr3, células Vero,

células olfativas immortalizadas, células gustativas immortalizadas y variedades transgénicas de las mismas. Las células de la invención o que se usan según la invención se derivan de la línea celular HEK293 o HEK293T.

5 En otra realización más, el complejo permite la escisión de la secuencia de ácido nucleico objetivo adyacente a la secuencia de ARN guía y el complejo de proteína Cas9 que se entrega a la célula y en el que el procedimiento comprende además introducir un ADN donante que comprende un promotor CMV dentro la célula.

En una realización, la secuencia de ARN guía y el complejo de proteína Cas9 permite la escisión de la secuencia de ácido nucleico objetivo adyacente a la secuencia de ARN guía y el complejo de proteína Cas9 y en el que el procedimiento comprende además introducir un ADN donante que comprende un promotor CMV a través de la homología celular dirigida, mecanismo de reparación en el sitio de escisión.

10 En una realización adicional, una línea celular se modifica para expresar de manera estable el gen de RTP1 endógeno bajo el promotor CMV.

15 En una realización adicional, un ADN donante también puede comprender un casete de selección de antibióticos (por ejemplo, que contiene el gen de resistencia a la puromicina). El cultivo de células en medios de cultivo que contienen antibióticos después del suministro de ADN puede ser beneficioso, ya que elimina las células que no experimentaron una integración adecuada del ADN y, por lo tanto, permite seleccionar eficientemente las poblaciones de clones celulares recombinadas que adquirieron un marcador de resistencia, de ahí la integración que se desea del ADN del donante (por ejemplo, el promotor constitutivo CMV). Dicho gen de resistencia a antibióticos se puede eliminar posteriormente por ingeniería mediante sitios flanqueantes "frt" que se reconocen específicamente por la enzima Flipasa. Este casete luego se retira mediante la entrega de dicha enzima a las células.

20 En una realización, el complejo funciona apuntando a ubicaciones específicas en el genoma y reclutando más factores de transcripción que activan genes endógenos corriente abajo como RTP1 sin la necesidad de escindir el ADN. Esto se realiza fusionando la proteína Cas9 con un dominio de activación de la transcripción para formar un complejo en el que el complejo no es capaz de escindir la secuencia de ácido nucleico objetivo. El procedimiento proporciona un ADN objetivo que dirige al ARN guía (es decir, corriente arriba del gen de RTP1) del dominio de activación de la transcripción que se fusiona a Cas9. En lugar de escindir el ADN para permitir la integración del promotor, se une y recluta transitoriamente factores de transcripción que activan el gen sin la necesidad de diseñar o modificar el genoma. Esto se puede hacer mediante el uso de una versión desactivada de Cas9 que se llama dCas9, que se fusiona a factores transcripcionales específicos que reclutan elementos como VP64, VPR y SAM, o mediante el uso de un ARN guía modificado (por ejemplo, un ARN guía truncado). Esta proteína de fusión no escinde el ácido nucleico.

30 Una realización, que se proporciona en la presente memoria es un procedimiento que comprende introducir un ácido nucleico que codifica un receptor olfativo dentro la célula.

35 Una realización adicional, que se proporciona en la presente memoria es un procedimiento para identificar un compuesto que activa, imita, bloquea, inhibe, modula y/o aumenta la actividad de un receptor olfativo en una célula no olfativa, en el que la célula se deriva de una línea celular HEK293 o HEK293T y comprende un ADN donante que se introduce en el sitio objetivo genómico corriente arriba de un locus del gen de RTP1 endógeno, en el que el ADN donante comprende un promotor en el que el promotor impulsa la expresión del gen de RTP1, en el que el procedimiento comprende además:

- a. poner en contacto el receptor, o una quimera o fragmento del mismo con un compuesto que activa, imita, bloquea, inhibe, modula y/o mejora el receptor y
- 40 b. determinar si el compuesto tiene un efecto sobre la actividad del receptor.

En una realización, el receptor olfativo es del grupo que consiste en un receptor de almizcle y de mal olor.

En una realización particular, el receptor de mal olor se selecciona de un receptor de escatol o indol.

En una realización particular, el receptor de almizcle se selecciona de un receptor de almizcle policíclico y un receptor de almizcle nitro.

45 En una realización, los ácidos nucleicos que codifican un receptor de olor se introducen en ausencia sustancial de una proteína G_{off}.

En una realización, se llevan a cabo las siguientes etapas:

1. editar el genoma de una línea celular usando la tecnología CRISPR/Cas9, que incluye: (1) diseñar un ADN que codifique un 'ARN guía' específico para el sitio de integración de ADN genómico deseado ubicado cerca del locus del gen de RTP1 endógeno; y (2) un 'ADN donante' para que se integre en el locus genómico que comprende un promotor transcripcional constitutivamente activo;
- 50 2. introducir los ADN diseñados en la etapa 1 en una línea celular de mamífero;

3. seleccionar una línea celular que tiene integrado el ADN del donante en el locus genómico deseado y que produce un ARNm de RTP1 mediante la activación del gen de RTP1 endógeno.

4. introducir una secuencia de ADN del receptor de olor dentro de la línea celular seleccionada.

5. poner en contacto un receptor, quimera o fragmento con un compuesto y analizar si el compuesto tiene un efecto sobre la actividad del receptor de olor.

Los procedimientos proporcionados en la presente memoria permiten el uso de líneas celulares para descubrir ingredientes tales como potenciadores o bloqueadores de olor para uso cosmético e industrial (por ejemplo, perfumes, potenciadores de perfume, potenciadores de sabor, desodorantes para el hogar y el cuerpo). Los nuevos ingredientes pueden proporcionar perfiles de fragancia, toxicidad y biodegradación más favorables y/o exhibir una mayor potencia.

10 En consecuencia, se divulga en la presente memoria un compuesto o mezcla de compuestos que activa, imita, bloquea, inhiba, modula y/o aumenta la actividad de un receptor olfativo en una célula no olfativa que se obtiene por uno cualquiera de los procedimientos divulgados en la presente memoria.

En una realización, se proporciona un sistema de expresión heterólogo de expresión funcional mejorada del receptor de olor que usa CRISPR/Cas9 para activar específica y constitutivamente el gen de RTP1 que se silencia (inactiva) en células HEK293T regulares.

Definiciones

Los siguientes términos tienen los significados atribuidos a éstos, a menos que se especifique lo contrario.

"Gen endógeno" se refiere a un gen que se origina dentro de un organismo, tejido o célula.

20 La expresión "efectos funcionales" incluye la determinación de cualquier parámetro que esté indirectamente o directamente bajo la influencia del receptor, por ejemplo, efectos funcionales, físicos y químicos. Incluye, pero no está limitado a, la unión de ligandos, cambios en el flujo de iones, potencial de membrana, flujo de corriente, transcripción, unión a proteínas G, fosforilación o desfosforilación de GPCR, interacciones receptor-ligando de transducción de señales, concentraciones de segundo mensajero (por ejemplo, AMPc, GMPc, IP3 o Ca intracelular²⁺), in vitro, in vivo y ex vivo y también incluye otros efectos fisiológicos tales como aumentos o disminuciones de la liberación de neurotransmisores u hormonas.

30 La expresión "determinar si el compuesto tiene un efecto sobre la actividad" en el contexto de los ensayos significa ensayos para un compuesto que aumenta o disminuye un parámetro que está indirectamente o directamente bajo la influencia de un miembro de la familia OR, por ejemplo, efectos funcionales, físicos y químicos. Tales efectos funcionales se pueden medir por cualquier medio conocido por los expertos en la técnica, por ejemplo, pero no está limitado a, cambios en las características espectroscópicas (por ejemplo, fluorescencia, absorbancia, índice de refracción), hidrodinámica (por ejemplo, forma), cromatográfica o propiedades de solubilidad, pinzamiento de parche, colorantes sensibles al voltaje, corrientes de células enteras, eflujo de radioisótopos, marcadores inducibles, expresión génica de OR de ovocitos; cultivo celular de tejidos con expresión de OR; activación transcripcional de genes OR; ensayos de unión a ligandos; cambios de voltaje, potencial de membrana y conductancia; ensayos de flujo de iones; cambios en los mensajeros intracelulares secundarios como AMPc, GMPc e inositol trifosfato (IP3); cambios en los niveles de calcio intracelular; liberación de neurotransmisores, y similares.

40 El término "vector de expresión" se refiere a cualquier sistema de expresión recombinante con el propósito de expresar una secuencia de ácido nucleico de la invención *in vitro* o *in vivo*, de manera constitutiva o inducible, en cualquier célula, que incluye células procariotas, de levaduras, fúngicas, vegetales, de insectos o de mamíferos. El término incluye sistemas de expresión lineales o circulares. El término incluye sistemas de expresión que permanecen episomales o se integran en el genoma de la célula huésped. Los sistemas de expresión pueden tener la capacidad de autorreplicarse o no, es decir, conducir solo la expresión transitoria en una célula. El término incluye casetes de expresión recombinante que contienen solo los elementos mínimos necesarios para la transcripción del ácido nucleico recombinante.

45 Por "célula huésped" se entiende una célula que contiene un vector de expresión y soporta la replicación o expresión del vector de expresión. Las células huésped pueden ser células procariotas como *E. coli*, o células eucariotas como levaduras, insectos, anfibios o células de mamíferos como células CHO, HeLa, HEK-293 y similares, por ejemplo, células cultivadas, explantes y células *in vivo*.

50 "Inhibidores", "activadores", "contrarrestantes" y "moduladores" de genes o proteínas OR se usan indistintamente para referirse a moléculas inhibitoras, activadoras o moduladoras identificadas usando ensayos *in vitro* e *in vivo* para la transducción olfativa, por ejemplo, ligandos, agonistas, antagonistas, potenciadores y sus homólogos y miméticos. Los inhibidores son compuestos que, por ejemplo, se unen, bloquean parcial o totalmente la estimulación, disminuyen, evitan, retrasan la activación, inactivan, desensibilizan o regulan negativamente la transducción olfativa, por ejemplo, los antagonistas. Los activadores son compuestos que, por ejemplo, se unen, estimulan, aumentan, activan la apertura, facilitan, mejoran la activación, sensibilizan o regulan la transducción

olfativa, por ejemplo, agonistas. Los moduladores incluyen compuestos que, por ejemplo, alteran la interacción de un receptor con: proteínas extracelulares que se unen a activadores o inhibidores (por ejemplo, proteínas de unión a odorantes y otros miembros de la familia de vehículos hidrófobos); proteínas G; quinasas (por ejemplo, homólogos de rodopsina quinasa y beta quinasas de receptores adrenérgicos que están involucradas en la desactivación y desensibilización de un receptor); y arrestinas, que también desactivan y desensibilizan los receptores. Los moduladores pueden incluir versiones genéticamente modificadas de miembros de la familia OR, por ejemplo, con actividad alterada, así como ligandos de origen natural o sintéticos, antagonistas, agonistas, moléculas químicas pequeñas y similares. Dichos ensayos para inhibidores y activadores incluyen, por ejemplo, expresar miembros de la familia OR en células o membranas celulares, aplicar compuestos moduladores putativos, en presencia o ausencia de moléculas de aroma o gusto, por ejemplo, almizcles o malos olores, y luego determinar los efectos funcionales sobre la transducción olfativa, como se describió anteriormente. Las muestras o ensayos que comprenden miembros de la familia OR que se tratan con un posible activador, inhibidor o modulador se comparan con muestras de control sin el inhibidor, activador o modulador para examinar el grado de modulación.

La región del "dominio N terminal" comienza en el N-terminal y se extiende hasta una región cercana al comienzo de la primera región transmembrana. "Dominio transmembrana", que comprende las siete "regiones transmembrana", se refiere al dominio de polipéptidos OR que se encuentra dentro de la membrana plasmática, y también puede incluir los lazos citoplasmáticos (intracelulares) y extracelulares correspondientes. Las siete regiones transmembrana y los lazos extracelulares y citoplasmáticos se pueden identificar utilizando procedimientos estándares, como se describe en Kyte y Doolittle, J. Mol. Biol., 157:105-32 (1982), o en Stryer. La estructura secundaria y terciaria general de los dominios transmembrana, en particular los siete dominios transmembrana de los receptores acoplados a proteínas G, como los receptores olfativos, son conocidos en la técnica. Por lo tanto, la secuencia de estructura primaria se puede diseñar o predecir en base a secuencias de dominio transmembrana conocidas, como se describe en detalle a continuación. Estos dominios transmembrana son útiles para ensayos *in vitro* de unión a ligandos, tanto en fase soluble como sólida.

El término "ácido nucleico" o "secuencia de ácido nucleico" se refiere a un oligonucleótido de desoxirribonucleótidos o ribonucleótidos en forma de simple o doble cadena. El término abarca ácidos nucleicos, es decir, oligonucleótidos, que contienen análogos conocidos de nucleótidos naturales. El término abarca, además, estructuras similares a ácidos nucleicos con cadenas principales sintéticas. A menos que se indique lo contrario, una secuencia de ácido nucleico particular implícitamente abarca las variantes de estos modificadas de forma muy conservadora (por ejemplo, las sustituciones de codones degenerados) y las secuencias complementarias, así como la secuencia que se indica explícitamente. Específicamente, las sustituciones de codones degenerados se pueden lograr mediante la generación, por ejemplo, de secuencias en las que la tercera posición de uno o más codones seleccionados se sustituye con residuos con base mezclada y/o desoxiinosina.

Receptor de olor u "OR" se refiere a uno o más miembros de una familia de receptores acoplados a proteínas G que se expresan en células olfativas. Las células receptoras olfativas también se pueden identificar en base a la morfología o mediante la expresión de proteínas expresadas específicamente en células olfativas. Los miembros de la familia OR pueden tener la capacidad de actuar como receptores para la transducción olfativa.

El receptor de olor o los ácidos nucleicos "OR" codifican una familia de receptores acoplados a proteínas G con siete regiones transmembrana que tienen actividad de receptor acoplado a proteínas G, por ejemplo, se pueden unir a proteínas G en respuesta a estímulos extracelulares y promover la producción de mensajeros secundarios como IP₃, AMPc, GMPc y Ca²⁺ a través de la estimulación de enzimas como la fosfolipasa C y la adenilato ciclasa.

Los polipéptidos "OR" se consideran como tales si pertenecen a la superfamilia de receptores acoplados a la proteína G de los 7 dominios transmembrana codificada por un solo exón de ~ 1 kb de longitud y exhibe motivos característicos de aminoácidos específicos del receptor olfativo. Los siete dominios predichos se denominan dominios "transmembrana" o "TM", TM I a TM VII que se conectan por tres dominios predichos de "lazo celular interno" o "IC", IC I a IC III, y tres dominios predichos de "lazo celular externo" o "EC", EC I a EC III. Los motivos se definen como, pero no están limitados a, el motivo MAYDRYVAIC superpuesto a TM III e IC II, el motivo FSTCSSH superpuesto a IC III y TM VI, el motivo PMLNPFIY en TM VII, así como tres residuos C conservados en EC II, y la presencia de residuos altamente conservados de GN en TM I [Zhang y Firestein (2002), The Olfactory Receptor Gene Superfamily of the Mouse. Nature Neuroscience: 5 (2):124-33; Malnic y otros, The Human Olfactory Receptor Gene Family: PNAS: 101(8):2584-9].

Los términos "polipéptido", "péptido" y "proteína" se usan indistintamente en la presente memoria para referirse a un polímero de residuos de aminoácidos. Los términos se aplican a polímeros de aminoácidos en los que uno o más residuos de aminoácido es un mimético químico artificial de un correspondiente aminoácido de origen natural, así como de polímeros de aminoácidos de origen natural y polímeros de aminoácidos de origen no natural. El término "heterólogo" cuando se usa con referencia a las porciones de un ácido nucleico indica que el ácido nucleico comprende dos o más subsecuencias que no se encuentran en la naturaleza en igual relación entre sí. Por ejemplo, el ácido nucleico típicamente se produce por vía recombinante, y tiene dos o más secuencias a partir de genes no relacionados dispuestos para preparar un nuevo ácido nucleico funcional, por ejemplo, un promotor a partir de una fuente y una región de codificación a partir de otra fuente. Del mismo modo, una proteína heteróloga indica que la

proteína comprende dos o más subsecuencias que no se encuentran en la naturaleza en igual relación entre sí (por ejemplo, una proteína de fusión).

Un "promotor" se define como una serie de secuencias de ácidos nucleicos que dirigen la transcripción de un ácido nucleico. Como se usa en la presente memoria, un promotor incluye las secuencias de ácidos nucleicos necesarias cerca del sitio de inicio de la transcripción, tal como, en el caso de un promotor de polimerasa tipo II, un elemento TATA. Un promotor además incluye opcionalmente los elementos potenciadores o represores distales, que se pueden localizar hasta diversos miles de pares de bases a partir del sitio de inicio de la transcripción. Un promotor "constitutivo" es un promotor que se activa bajo la mayoría de las condiciones ambientales y de desarrollo. Un promotor "inducible" es un promotor que se activa bajo la regulación ambiental o de desarrollo.

Como se usa en la presente memoria, el término "recombinante" se refiere a un polinucleótido que se sintetiza o manipula de cualquier otra manera in vitro (por ejemplo, "polinucleótido recombinante"), mediante procedimientos para el uso de polinucleótidos recombinantes para producir productos génicos en células u otros sistemas biológicos, o para un polipéptido ("proteína recombinante") codificado por un polinucleótido recombinante. "Medios recombinantes" también abarca la ligadura de ácidos nucleicos que tienen diversas regiones codificadoras o dominios o secuencias promotoras de diferentes fuentes en un casete de expresión o vector para expresión de, por ejemplo, expresión inducible o constitutiva de una proteína de fusión que comprende un dominio de translocación de la invención y una secuencia de ácido nucleico amplificada que usa un cebador de la invención.

El ácido nucleico y las secuencias de aminoácidos que se identifican y/o utilizan en la presente memoria se enumeran a continuación:

Secuencia objetivo de ARN guía (SEQ ID NO: 1 ADN)

SEQ ID NO: 1

ctgcaatctcagttcagggcc

ADN donante para la reparación homóloga dirigida (SEQ ID NO: 2 ADN)

SEQ ID NO: 2

ggggttttatggaagagtcttactctcttttcttctatctatatttctatgttttctagaataaacccatagatttttaaaaggaaaaataattta
ttaaaaaatagcagcagagggcatgtatagtaaaggctgtttgcctgtgggtggtgctcctcttctgcgttctataatcagcttgaaaataat
ctgtctgctcctgcctggctgatgcaatgctcctacctttgtgcacaggtggtgttcttgcaaacggccattgcagcatggatcctattg
cacagttattcagtcacagtcagctacaagcactgacatagagcttggcacatgtctgcaaacctaccacatgctcggatatgtttg
aaatgaatgaattaatgaaccggctctgggtcaacagcttgaattgtatacaggtcgcgccatttataggctagtgagtcctaggctcc

tgatctgtactgcagcaatagtaataacttaagagacctccaattgtgtttgaaatggcaaagtgtgtgcacaaagatggctgggg
 aagccgagagagagttattatttctcatctactaacaatttacatctccccatccctatttctcttggtgcctaaggcatcatgg
 ttaccgtagcagccagatgctgatgatgctccaggggacggcaaggtgaaactgagccagttccagtcctcacctccccactctt
 tccaggccaggggtgagatggtctgaagctcagctctgtgctcaggtccccactctgtcttgatcatttagacccgcgccgcgccgc
 gcctcggaaattcgattgaagttcctattccgaagttcctattcttagaaagtataaggaaactcgggtgtggaaagtccccaggctccccag
 caggcagaagtatgcaaagcatgcatctcaattagtcagcaaccatagtcgccccctaactccgcccattccgcccctaactccgcc
 cagttccgcccatttccgccccatggctgactaattttttattatgcagagggccgagggccctcgccctcgactattccagaagt
 agtgaggaggctttttggaggcctaggttttgaaaaagcttgcatgctgcaggtcggccgccagaccgggtccgcccacatcc
 cctgaccacgcccctgaccctcacaaggagacgaccttcacgagtagacaagcccacgggtcgcctcgccaccgcgacg
 acgtccccggggcgtacgcacctcgcgcggcggttcgccactaccccgccacgcgccacacgtgcaccggaccgccaat
 cgagcgggtcaccgagctgcaagaactcttctcacgcgcgtcgggtcgacatcggaaggtgtgggtcgggacgacggcgcc
 gcggtggcggtctggaccacgcccggagagcgtcgaagcggggcggtgttcgccgagatcgcccgcatggccgagttgagc
 ggttcccggtcggcgcgagcaacagatggaaggcctcctgcccgcaccggcccaaggagcccggtgttctctggccacc
 gtcggcgtctcggccaccaccagggaagggtctgggcagcggcgtcgtctccccggagtggaggcggccgagcgcgcgg
 ggtgcccgccttctggagacctcgcgccccgcaacctccccctctacgagcggctcggcttcaccgtcaccgcccagctcaggt
 gccgaaggaccgcgacgtggtgcatgacccgcaagcccgtgctgacgcccgcacgaccgagcggccgaccgaaa
 ggagcgcacgacccccatggctccgaccgaagccacccggggcgccccgcggaccccgacccgccccgagggccaccgac
 tctagaggatcataatcagccataccacattttagaggttttactgtcttataaaaaactccacacctccccctgaacctgaacataaa
 atgaatgcaattgtgtgttaactgtttattgcagcttataatggttacaataaagcaatagcatcacaaattcacaaataaagcattttt
 tcactgcgaagtctctattccgaagtctcttcttagaaagtataaggaaactcaatcactagtaattcacgcgttgacattgattattgac
 tagttatlaatagtaataallacggggcattagttcatagcccatatagggagttccggtlacataactlacgglaaalggcccgcctgg
 ctgaccgccaacgacccccgccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggaatttccattgacgtcaatg
 ggtggagtatttacggtaaactgcccacttggcagtagcatcaagtgatcatatgcaagtacgccccctattgacgtcaatgacggtaa
 atgccccgctggcattatgccagtagacattatgggactttctacttggcagtagatctacgtattatgcatcgctattaccatggt
 gatgcgggtttggcagtagcatcaatggcggtgtagcgggttgactcacggggatttccaagtctccaccccatgacgtcaatggga
 gttgttttggcaccaaaatcaacgggactttccaaatgtcgtacaactccgccccattgacgcaaatggcggttagcggtgtacggt
 gggaggctatataagcagagctcgttttagtaaccgtgtttaaactcttcagagactccctctccccaaagctctgtcttctggcaacct
 gcctggttgcggtggaacaggttccactcgggacaaaaggaggagctgggtcctgcttctctggtctgtgatgaggttttag
 accgtggagactgcgctgcctgcctgcacctaccctcactctccgtgttctcactaaggtggaaattgccttccctcactactacga
 gaccatgtgtaaaagcgtgaccacagatgagtggaagaaagtcttctatgagaagatggaggaggcaaagccggtgacagctggg
 acctcatcatagaccccaacctcaagcacaatgtgctgagccctgggtggaagcagtagcttgaattgcatgcttcaggcaggtgagta
 gccaggaaagtggatccctgcaggccgcttaggtccctagctctggggcaccttccaaggagaggaaattacgtagaacca
 agtgttttagcttcaatctcactattaggtgctgtagactggaaagtagagaaagagtccttaactgggaactacgacacttgagttgga
 tttagctcttctactgatcacctgtgtactcttctctctgagtcacaattttccgtctggaaaataaagacatagaatatacgtatgagtc
 ctacacactgacattttacatattttctattttaacagctctcttaaaaagtagtttaaaaccagagaagaagggttgaggccactgggggt
 cgagacgtccgtgctctggtcctgggaccgggttaaatctatttaa

Olf741 de ratón (SEQ ID NO: 3 ADN; SEQ ID NO: 4 PROTEÍNA)

SEQ ID NO: 3

atgaaaaccctcagcagccccagcaactccagcaccatcactggttcacctcttgggttcgcctacccaggaggaggcaaatc
tctctttgtgatcttctcattgtttacatactcattcttatgggcaacgcttccatcatctgtgctgtactgtgatcagagactccacacc
ccatgtactctctgctggccaacttctcttcatggagattggatatgtcacctccacagtccccaacatgttgccaacttcttccagac
accaaggtcatctcttctctggatgcttctgcagttctatttcttcttcttcttgggttctacagaatgcttttctggcagtcattg
atcgataccttgccatctgtaggccactacattatctctcatgactgggcgcctccgaaacacccttgaccagttgctgggtgctt
ggtttctctggttccctgtaccatcatcatctcccagatgcttctgtgggtccagaattatagaccacttctgtgtgaccaggc
cctctttggcccttgctgttccagagtccttgcattgatagaggtttctgggtccattataatgtctatgctcctggttattcttcttcatcat
gggaacttacatattggtcctaagagctgtgttagacttctcaagagaaggacaaaaaaggcttctccacttgcgggtctcatctc
acagtagtttactcttttattgctcagtgatgataatgtatctgagcccaacatctgagcatgaggccggaatgcagaagcttgaactct
attttattctgtgggtacaccactgcttaacctatgatatacagcttgaggaacaaagatatgaaaaatgccctacagaagatttga

SEQ ID NO: 4

mktlsspsnssttgfillgfaypregqillfviffvylilmgnasiicavycdqrhltpmylllanfsfmeigyvtstvpnmlanfls
dtkvisfsgclqfyfffsfgstecfflavmafdrylaicrplhypslmtgrlrlntlvscwvlgflwfpvpriiisqmsfcgsriidhflc
dpgpllalacsrvplievfwsiimsmllvipflimgtyilvrlavfrlpsregqkkafstcgshltvsvlfycsvmimylsptseha
gmqklvltfysvgtpllnmpiyslrnkdmknlqkilt

Olf742 de ratón (SEQ ID NO: 5 ADN; SEQ ID NO: 6 PROTEÍNA)

SEQ ID NO: 5

atgaaaaccctcagcagccccagcaactccagcaccatcactggttcacctcttgggttccccctgccccaggaggaggcaaatcc
tctctttgtgaccttctcattgtttacatactcattcttatgggcaatgcttccatcatctgtgctgtactgtgatcagacctccacacc
ccatgtacttctgctggccaacttctcttctctggagatctggatgtcacctccacagtccccaacatgttgccaacttcttccagac
accaaggtcatctcttctctggatgcttctgcagttctatttcttcttcttcttgggttctacagaatgcttttctggcagtcattg
atcgataccttgccatctgtaggccactacattatctctcatgactgggcacctctgcaacatccttgatcagttgctgggtgcttg
gttctctggttccctgtaccatcatcatctcccagatgcttctgtgggtccagaattatagaccacttctgtgtgaccaggcc
ctcttttggcccttgctgttccagagccccattgatggaggtttctggacaattataatgtctatgctcctggttattcttcttcatcat
gggaacttacatattggtcctaagagctgtgttagacttctcaagagatggacaaaaaaggccttctccacttgcgggtctcatctc
acagtagtttactcttttattgctcagtgatgaaaatgtattgagcccaacatctgagcatgaagctggaatgcagaagcttgaactcta
ttttattctgtgggtactccactacttaacctgtgatatacagcttgaggaacaaagatatgaaaaatgccctgcagaagattttaagaaca
taa

5

SEQ ID NO: 6

mktlsspsnssttgfillgfpcpregqillfvtfvylilmgnasiicavycdqlhtpmylllanfsfleiwvyvtstvpnmlanflsd
tkvisfsgclqfyfffsfgstecfflavmafdrylaicrplhypslmtghlenilviscwvlgflwfpvpriiisqmsfcgsriidhflc
dpgpllalacsraplmevfwtiimsmllvipflimgtyilvrlavfrlpsrdgqkkafstcgshltvsvlfycsvmkmyslptsehe
agmqklvltfysvgtpllnmpiyslrnkdmknlqkilt

Olf96 de ratón (SEQ ID NO: 7 ADN; SEQ ID NO: 8 PROTEÍNA)

SEQ ID NO: 7

atgggaatcctttccacaggaaatcaaactgtcactgagttgtacttcttggttccatgaagtcctgggctgcacctcctgttttttctgt
gttcaccatcctctatgcctccatcatcacagggaacatgctcattgagtggtggtgagctcccagaggcttcacacccatgta
tttctttctgggtaactgtccttcatagagattgtctatactccacagtggtgccccaaatgctggaaggcttctacaggaggccacat
atctgtggctggctgctgctccagttcttttttggctctctggccacagatgagtggtttctgctggctgtgatggcatatgatcgatc
tcgcaattgtcacctctacgataccacacctcatggggcctcaatggtgcctggggtggtgctcacagctggtgctgctggctcat
ggtagatggactagttgtgctgtgatggccagttgagattctgtggcccaacttagttgatcacttttactgtgattttcaccttgatgg
tctggcctgctcagatacccaagtggccagtgactacattgttctctgtggttctctgactgtccctttgggctggttctgatctc
ctatgctcagattgtagtactgtgctgagagttccttctgggaccagaagaaccaaggccttctccacatgctcctctcacctggctgtg
gtgtccacgttctatggaacactcatggtattgtacattgtgccctctgctgttcattctcagctcctctccaaggtcattgccctgctctaca
cagtggtcactcccatttcaacctgtcatctacacctgaggaaccaggaggtgcagcaggcactaagaaggcttctctactgcaaa
ccaactgaaatgtga

SEQ ID NO: 8

Mgilstgnqvtcfvllgfhevpglhllffsvfilyasiitgnmliavvvssqrlhtpmyfflvnlsfieivystvvpkmllegflqea
tisvagcllqffvfgslatdecfllavmaydrylaichplryphlmgrpqwcglvltvwlsghmvdglvvalmaqlrfcgpnlvdh
fycdfsplmvlacsdtqvaqvttfvlsvfltpfglvlisyaqivvtlrvpsgtrrtkafstcshlavvstfygtlmvlyivpsavhs
qlskviallytvvtpifnpiytlrnqevqqalrlllyckptem

OR11A1 humano (SEQ ID NO: 9 ADN; SEQ ID NO: 10 PROTEÍNA)

SEQ ID NO: 9

atggaaattgtctccacaggaaacgaaactattactgaattgtcctccttggttctatgacatccctgaactgcatttctgtttttattgtat
tactgctgtctatgtcttcatcatataggaatatgctgattattgtagcagtggttagctcccagaggctccacaaacccatgtatattt
cttgcggaatctgtccttctggatattctctacacctccgcagtgtatgcaaaaatgctggagggttctctgcaagaagcaactatctct
gtggctggttctgtcctcagttcttattcttggctctctagccacagctgaatgcttactgctggctgtcatggcatatgaccgtacctg
gcaattgtaccactccactaccactcctgatggggcccagacgggtacatggggctggtggtcacaacctggctctctggattgt
ggtagatggactggtgtggccctggtggccagctgaggttctgtggcccaaccacattgaccagtttactgtgactttatgcttttg
tgggctggttctgctcgatccagagtggctcaggtgacaactctcattctgtctgttctgctcactattccttttgactgattctga
catcttatgccagaattgtggtggcagtgctgagagttcctgctggggaagcaggagaagggttctccacatgctcctcccacta
gctgtagtgaccacattctatggaacgctcatgatctttatgttgaccctctgctgtccattccagctcctctccaaggtcttctccctgc
tctacactgtggtcacccctctcttaactctgtgatctataccatgaggaacaaggaggtgcatcaggcacttcggaagattctctgtatc
aaacaaactgaaacacttgattga

SEQ ID NO: 10

Meivstgnetitefvllgfydipelhflffivftavyvfiiignmliavvvssqrlhkpmiyiflanlsfldilytsavmpkmllegflqea
atisvagcllqffvfgslataeclllavmaydrylaicyplyphlmgrpmyglvvtwlsghmvdglvvalvaqlrfcgpnhidqf
ycdfmlfvglacsdpvraqvttlilsvfcltipfglitsyariivavvlrvpagasrrafstcshlavvttfygtlmifyvapsavhsqll
skvfllytvvtpifnpiytmrnkevhqalrkilcikqtetld

Olfr740 de ratón (SEQ ID NO: 11 ADN; SEQ ID NO: 12 PROTEÍNA)

SEQ ID NO: 11

atgaaaaccttcagcagccccatcaactccagcaccaccactggcttcattctcttgggttccctgccccaggaggaggcgaatcct
cctcttgtgtctcttctcattgtctacctgcttacctcatgggcaacacttgcacatcttgcagtatgctgggatcagagactccacac
acctatgtacctactgtctggccaacttctcctcctggagatctggatgttacctccacagtccccaacatgttggccaatttctctctga
caccaaggatcatctcttctctggatgcttctgcagttctatttcttctccttgggttctacagaatgccttttctggcagtcattggcatt
gatcgataccttggccatctgtaggccactacattatcctgctctcatgactgggagcctctgcaacatccttgtgatcagttgctgggtgct
tggtttctctggttccctgttcccatcatcatctcccagatgtccttctgtgggtccagaattatagaccacttctgtgtgacccaggc
cctctattggccctcacctgttccagagccccattaatggaggttttctggacaattataacatcttcttcttcttcttcttctc
atgggatcttatacattggctcctgagagctgtgttcagagttcctcaagatggacaaaaaggcttctcacttgcggatctcatct
cacagtagtttactcttttatggctcagtgatgataatgtatctaagcccagcctctgagcatgaagctggaatgcagaagcttgactc
tattttattctgtgggtactccactcattaatcctgtgatatacagcttgaggaacaaggatatgaacatgccctgcagaagatttaagaa
cataa

SEQ ID NO: 12

mktfsspinsstttgfillgfpcpregqillfvlsivylltmgntciifavcwdqrlhtpmylllanfsfleiwvytstvpnmflanflsd
tkvisfsgcflqfyffflsgsteclflavmafdrylaicrplhypalmtgslcnilviscwvlglwfpvpiiiisqmsfcgsriidhflcd
pgplaltcsraplmevfwtiitslilfvpflfimgsytlvlravfrvpsrdgqkkafstcgshltvllfygsvmimylsptscheag
mqklvtlfysvvtplinpviyslrnkdmkhalqkilt

OR1A1 humano (SEQ ID NO: 13 ADN; SEQ ID NO: 14 PROTEÍNA)

SEQ ID NO: 13

atgagggaataaaccagtcctctacactggaattcatcctcctgggagttactggtcagcaggaacaggaagatttctctacatcctct
tcttgtcatttaccatcacattgattggaaacctgctcatgctcctagccatttgcctgatgttcgcttcacacccccatgtattttctcc
ttgccaacctctccttgggtgacatcttctctcatcgtaacctccctaagatgctggccaacctcttgggcagcaaatccatctctt
tgggggatgcctaacgcagatgtatttcatgatagccttgggtaacacagacagctatatttggctgcaatggcatatgatcagctgtg
gccatcagccgccacttactacacaacaattatgagtcacggcttctgtatctggcttattgttgggtcttgggtgattggaaatgcca
atgccctccccacactctgctcacagctagctgtccttctgtggcaaccaggaagtggccaacttctactgtgacattcccccttgc
gaagtatcctgttctgacatccacttcatgtgaagatgatgtacctaggggttggcattttctctgtgccattactatgcacattgtctcct
atatcagagtcttctccacagtcttccagggtcctccaccaaggcgctgtcgaaggccttctccacctgtggttcccacctcacgggtgtc
tcttgtattatggtacagtcagggcacgtatttccgccctttgaccaattatagcctaaaagacgcagtgatcactgtaatgtacacggc
agtgacccaatgttaaatccttcatctacagcttgagaaatcgggacatgaaggctgccctgcggaaactcttcaacaagagaatctc
ctcgtga

SEQ ID NO: 14

mrennqsstlefillgvtgqqeqedffylflfiypitlignllivlaicsdvrlnpmylllanlsldiffssvtipkmlanhllgsksis
fggcltqmyfmialgnntsyilaamaydravaisrplhyttimsprsciwiagswvignanalphtlltaslsfcgnqevanfycdi
tpllklscsdihfhvkmmylgvgifsvpllciiivsyirvfstvfqvpstkgvlfkafstcgshltvvslyygtvmgtyfrpltnyslkda
vitvmytavtpmlnlpfiyslrnrdmkaalrklnkriss

Etiqueta Flag (SEQ ID NO: 15 ADN; SEQ ID NO: 16 PROTEÍNA)

SEQ ID NO: 15

gattacaaggacgacgacgataag

SEQ ID NO: 16

dykdddk

Etiqueta Rho (SEQ ID NO: 17 ADN; SEQ ID NO: 18 PROTEÍNA)

5 **SEQ ID NO: 17** atgaacgggaccgagggcccaactctacgtgcctttccaacaagacgggcgtggtg

SEQ ID NO: 18

mngtegnfnfyvpsnktgvv

Etiqueta Lucy (SEQ ID NO: 19 ADN; SEQ ID NO: 20 PROTEÍNA)

SEQ ID NO: 19

10 atgagacccagatcctgctgctcctggccctgctgaccctaggcctggct

SEQ ID NO: 20

mrpqillllalttlla

Ejemplos

Los siguientes ejemplos son ilustrativos para la invención.

15 Ejemplo 1

Estrategia de edición del genoma para inducir la activación constitutiva del gen de RTP1 endógeno en células HEK293T.

Se describe una estrategia para desarrollar una expresión funcional mejorada del receptor de olor en el sistema de expresión heterólogo. Aprovechando una nueva tecnología de edición del genoma llamada CRISPR/Cas9, un gen de RTP1 endógeno que está silenciado (inactivo) en las células HEK293T normales, se activa de manera específica y constitutiva mediante la introducción de un promotor constitutivo (CMV) corriente arriba de su secuencia de codificación. Figura 1) El gen de RTP1 se encuentra en el cromosoma 3 y se muestra la secuencia de ADN alrededor de su sitio de inicio. La endonucleasa Cas9 está dirigida por un ARN guía (gRNA) de 20 pares de bases (pb) homólogo al objetivo. Tras la entrega a las células (GeneArt CRISPR Nuclease (enriquecido con CD4) Vector Kit, cat # A21175), la molécula de ARN guía y la proteína Cas9 forman un complejo activo que induce la ruptura de ADN de doble cadena deseada (DSB) corriente arriba de la secuencia de codificación. Los cuadros indican el gen de RTP1 en el cromosoma 3 (cuadro lleno, secuencia de codificación (CDS); cuadro abierto, región no traducida (UTR) en el exón). La supuesta región promotora corriente arriba del gen de RTP1 está inactiva en las células HEK293T. Se muestra la secuencia objetivo de ARN guía (SEQ ID NO: 1) entre la posición -150 y -131 y un Motivo Adyacente Protospacer (PAM) desde -153 hasta -151 desde el codón de inicio (ATG) respectivamente. Figura 1) El sitio DSB para la nucleasa Cas9, 3pb lejos del motivo PAM, que se especifica por un triángulo, permite que se inserte un ADN donante (SEQ ID NO: 2). Figura 2) Se muestra un esquema del procedimiento de inserción del promotor CMV. La configuración superior muestra el locus del gen de RTP1 antes de la modificación y el esquema inferior muestra el locus RTP1 después de que un ADN donante se dirige al sitio DSB mediante Reparación por Homología Directa (HDR). El ADN del donante está compuesto por un brazo de homología 5', un casete de selección de puomicina flanqueado por FRT (Objetivo de Reconocimiento a Flipasa), el promotor CMV y un brazo de homología 3'. La integración del promotor CMV corriente arriba del gen de RTP1 se obtiene luego por el mecanismo celular HDR inherente a las células eucariotas. Un plásmido donante que contiene dos ADN se estira homólogo a las secuencias a cada lado del punto de entrada deseado, flanqueando el casete de selección de resistencia a la puomicina (Puro^r) y el ADN del CMV se cotransfecta en células HEK293T. El HDR resulta en la introducción de Puro^r y CMV corriente arriba de la secuencia de codificación de RTP1. El casete de selección de puomicina se puede retirar posteriormente mediante el uso de la enzima Flipasa.

Ejemplo 2

Selección de una línea celular modificada que expresa endógenamente el gen de RTP1.

Varias etapas de control ayudan a caracterizar la modificación de la línea celular y su integridad. Figura 3) Muestra un esquema de los alelos tipo salvaje y los recombinados. Las líneas grises indican las posiciones relativas del amplicón de los resultados experimentales de PCR y RT-PCR para el genotipado de ADN y para los controles de expresión de ARN, respectivamente (no a escala). Figura 4) Se extrae el ADN genómico de la línea celular resistente a la puomicina y se realiza una PCR que discrimina entre el tipo salvaje no recombinado (WT) y las líneas celulares modificadas (Mod.). La PCR 1 amplifica una banda de 2,0 kb solo en células HEK293T de tipo salvaje pero no en una línea celular modificada. La línea modificada debería producir una banda de 4,0 kb con la PCR 1 pero no probablemente debido a

la longitud y la complejidad de la estructura genómica. Los resultados de genotipado para la línea celular modificada no pudieron producir la banda de 2,0 kb que indica una integración homocigótica del promotor CMV. La integración adecuada del ADN del donante se probó adicionalmente con la PCR 2 y 3, como se indica. Figura 5) Después de la extracción del ARNm y la síntesis de ADNc, se realiza un experimento de RT-PCR para demostrar que el ARNm de RTP1 se expresa específicamente en la línea celular modificada pero no en las células HEK293T originales. Esto confirma que el promotor de CMV que se integró en el locus genómico objetivo conduce adecuadamente la expresión del gen de RTP1. La especificidad de las bandas de RT-PCR se confirmó por secuenciación directa de las bandas amplificadas. La transcripción reversa negativa (RT-) y las condiciones de PCR y GAPDH indican la ausencia de ADN genómico contaminante y la presencia de ADNc en todas las muestras, respectivamente.

Ejemplo 3

Caracterización de la expresión de la proteína RTP1.

La expresión de la proteína RTP1 en la línea celular modificada seleccionada se determinó mediante análisis de transferencia Western usando un anticuerpo específico para RTP1. Una forma de proteína larga (RTP1L) y una corta (RTP1S) se pueden originar a partir del gen de RTP1 endógeno. La estrategia de modificación del genoma que se describe en la presente memoria implicó la introducción del promotor CMV corriente arriba del codón de inicio de RTP1L para evitar cualquier modificación de la secuencia de codificación endógena; por lo tanto, se esperaba que RTP1L se expresara. Sin embargo, los resultados indican que la línea celular modificada estaba fuertemente sesgada hacia la expresión de RTP1S, lo que sugiere que el gen de RTP1 endógeno expresa preferentemente la versión corta sin edición del genoma adicional. Esta última es la versión preferida, ya que se sabe que promueve mejor la expresión del receptor de olor en la superficie celular. Figura 6) muestra las transferencias Western de la proteína RTP1. Las puntas de flecha indican los tamaños de proteína esperados para RTP1S - 25 kDa, RTP1L - 28 kDa y la proteína de control β -actina - 42 kDa. Se muestra la ausencia de la proteína RTP1 en una línea celular HEK293T de tipo salvaje (WT) y la presencia de RTP1S en la línea celular modificada (Mod.). Sorprendentemente, se puede ver una banda mucho más fuerte para RTP1S en comparación con RTP1L. La extracción de proteína de membrana se preparó de acuerdo con el kit de extracción de proteína de membrana Mem-Per Plus (Pierce, cat # 89842). Se utilizó como marcador de peso Chameleon Duo Pre-stained (LiCor, cat#92860000). El etiquetado se realizó con los siguientes anticuerpos primarios: Anti-RTP1 de conejo (Invitrogen, cat#PA5-24028) y anti- β -actina de ratón (Pierce, cat#PIMA515739). La detección se realizó con los siguientes anticuerpos secundarios: anti-conejo de cabra (LiCor cat#925-32211) y anti-ratón de cabra (LiCor cat#925-68070). Las imágenes se realizaron en un Odyssey CLx (LiCor).

Ejemplo 4

Caracterización funcional de varios receptores de olor en la línea celular modificada.

Se realizaron experimentos de dosis-respuesta funcional para evaluar el nivel de mejora funcional de la actividad de los receptores de olor en la línea celular modificada. Los receptores de olor se modificaron en su N-terminal con secuencias o etiquetas cortas de polipéptidos [por ejemplo, Flag (SEQ ID NO: 15), Rho (SEQ ID NO: 17; 20 primeros aminoácidos del receptor de rodopsina bovina) o Lucy (SEQ ID NO: 19)], expresado transitoriamente en células WT o HEK293T modificadas, y se estimuló con compuestos olorosos para determinar la actividad de los receptores. Figuras 7 y 8) Usando un ensayo de unión a olores basado en células, se probó la actividad de Olfr741 (SEQ ID NO: 4) y Olfr742 (SEQ ID NO: 6) para indol en el diseño de la línea celular RTP1 y se comparó con HEK293T que carece de la expresión de la proteína RTP1. Los receptores de olor se transfectaron en ambas líneas celulares y se expusieron a concentraciones crecientes de indol. La actividad inducida por olor se detectó midiendo el nivel de aumento de AMPc en el citosol usando un kit basado en HTRF (CisBio, kit cAMP dynamic 2, cat#62AM4PEJ). Figura 9 y 10) Usando el mismo ensayo de unión a olores basado en células, se probó la actividad de Olfr96 (SEQ ID NO: 8) y OR11A1 (SEQ ID NO: 10) para vulcanólido en el diseño de la línea celular de expresión de RTP1 y se comparó con HEK293T que carece de la expresión de la proteína RTP1. La actividad de Olfr740 (SEQ ID NO: 12) para indol también se probó en ambos fondos celulares. Se registra un aumento de la actividad del receptor dependiente de la dosis para todos los OR en la línea celular RTP1 modificada y no en la línea celular de control no modificada que carece de expresión de RTP1. Además, la actividad de OR1A1 (SEQ ID NO: 14) para carvona- (-) se probó en ambos fondos celulares. Aunque OR1A1 se puede expresar en HEK293T regular, se registra un aumento más potente de la actividad del receptor dependiente de la dosis en la línea celular RTP1 modificada y se compara con la línea celular de control no modificada que carece de expresión de RTP1.

Listado de secuencias

<110> Firmenich SA

<120> Líneas celulares para el cribado de receptores de aroma y olor

<130> 81455-95501

<150> 62/173,762

<151> 2015-06-10

	<150> 62/221,580	
	<151> 2015-09-21	
	<160> 20	
	<170> PatentIn versión 3.5	
5	<210> 1	
	<211> 21	
	<212> ADN	
	<213> artificial	
10	<220>	
	<223> Secuencia objetivo del ARN guía	
	<400> 1	
	ctgcaatctc agttcagggc c 21	
15	<210> 2	
	<211> 3628	
	<212> ADN	
	<213> artificial	
	<220>	
	<223> ADN donante para reparación por homología directa	
	<400> 2	
	gggggttttat ggaagagtct tactttctctt ttctttcatc tatattttgt attttttcta 60	
	gaataaacc atagatgttt ttaaaaggaa aaataattta ttaaaaatag cagcagaggc 120	
	atgtatagta aaggctgttt tgcctgtggg tgggtgctcct cttctgogct tctataatca 180	
	gcttggaat aatcttgtct gctcctgcct ggctgatgca atgctcctac ctttgtgcac 240	
	aggtggctgt tcttgacaaa ggccattgca gcatggatcc tattgcacag ttattcagta 300	
	cacagtcagc tacaagcact gacatagagc ttggcacatg tctgcaaacc ctaccacat 360	
	gctcggatat gtttgaaatg aatgaattaa tgaaccggtc tggggtcaac agcttgaatt 420	
	tgtatacagg ctccgccatt tataggctag gtgagtccta ggctcctgat ctgtactgca 480	
	gcaatagtaa tcataactta agagacctcc aattgtgttt tgaaaatggc aaagtgctgg 540	
	tcacaagatg gctggggaag ccgagagaga gtttattatt attgctccat ctactaacia 600	
	atttacatct ccccatccct catctctcct tggctgccta aggcacatg gttaccgtag 660	
20	cagccagatg ctgatgatgc ctccagggga cggcaagggtg aaactgagcc agttccagat 720	

ES 2 802 524 T3

cctcacctcc	ccatactctt	tccaggccag	ggtgagatgg	tctgaagctc	agtctctggt	780
caggtccccc	actctgtctt	ggatcattta	gacccgcggc	cgccgcgcgc	ctcggaaattc	840
gattgaagtt	cctattccga	agttcctatt	ctctagaaag	tataggaact	tcggtgtgga	900
aagtccccag	gtccccagc	aggcagaagt	atgcaaagca	tgcattctca	ttagtcagca	960
accatagtcc	cgcccctaac	tccgcccata	ccgcccctaa	ctccgcccag	ttccgcccata	1020
tctccgcccc	atggctgact	aatTTTTTTT	atttatgcag	aggccgaggc	cgcctcggcc	1080
tctgagctat	tccagaagta	gtgaggaggc	TTTTTTggag	gcctaggctt	ttgcaaaaag	1140
cttgcatgcc	tgcaggctcg	ccgccacgac	cgggtccgcc	accatcccct	gacccacgcc	1200
cctgacccct	cacaaggaga	cgaccttcca	tgaccgagta	caagcccacg	gtgcgcctcg	1260
ccaccgcga	cgacgtcccc	cgggccgtac	gcaccctcgc	cgccgcgttc	gccgactacc	1320
ccgccacgcg	ccacaccgtc	gacccggacc	gccacatcga	gcgggtcacc	gagctgcaag	1380
aactcttctt	cacgcgcgtc	gggtctcgaca	tcggcaaggt	gtgggtcgcg	gacgacggcg	1440
ccgcggtggc	ggtctggacc	acgccggaga	gcgtcgaagc	gggggcgggtg	ttcgccgaga	1500
tcggcccgcg	catggccgag	ttgagcggtt	cccggctggc	cgccgagcaa	cagatggaag	1560
gcctcctggc	gcgcacccgg	cccaaggagc	ccgcgtgggt	cctggccacc	gtcggcgtct	1620
cgcccgacca	ccagggcaag	ggtctgggca	gcgcctcgt	gctccccgga	gtggaggcgg	1680
ccgagcgcgc	cggggtgccc	gccttctctg	agacctccgc	gccccgcaac	ctccccctct	1740
acgagcggct	cggcttcacc	gtcacccgcg	acgtcagaggt	gcccgaagga	ccgcgcacct	1800
ggtgcatgac	ccgcaagccc	ggtgcctgac	gcccgcacca	cgacccgcag	cgcccgaccg	1860
aaaggagcgc	acgaccccat	ggctccgacc	gaagccaccc	ggggcgggccc	cgccgacccc	1920
gcacccgccc	ccgaggccca	ccgactctag	aggatcataa	tcagccatac	cacatttgta	1980
gagggttttac	ttgctttaaa	aaacctccca	cacctccccc	tgaacctgaa	acataaaatg	2040
aatgcaattg	ttgttggtta	ctgttttatt	gcagcttata	atggttacia	ataaagcaat	2100
agcatcacia	atttcacaaa	taaagcattt	ttttcactgc	gaagttccta	ttccgaagtt	2160
cctattctct	agaaagtata	ggaacttcaa	tcactagtga	attcacgcgt	tgacattgat	2220
tattgactag	ttattaatag	taatcaatta	cggggtcatt	agttcatagc	ccatatatgg	2280
agttccgcgt	tacataactt	acggtaaagt	gcccgcctgg	ctgaccgccc	aacgaccccc	2340
gccattgac	gtcaataatg	acgtatgttc	ccatagtaac	gccaataggg	actttccatt	2400
gacgtcaatg	ggtggagtat	ttacggtaaa	ctgccactt	ggcagtacat	caagtgtatc	2460
atatgccaa	tacgccccct	attgacgtca	atgacggtaa	atggcccgcc	tggcattatg	2520
cccagtacat	gaccttatgg	gactttccta	cttggcagta	catctacgta	ttagtcatcg	2580
ctattaccat	ggtgatgcgg	ttttggcagt	acatcaatgg	gcgtggatag	cggtttgact	2640

ES 2 802 524 T3

cacggggatt	tccaagtctc	cacccattg	acgtcaatgg	gagtttgttt	tggcaccaaa	2700
atcaacggga	ctttccaaaa	tgtcgttaaca	actccgcccc	attgacgcaa	atgggcggtg	2760
ggcgtgtacg	gtgggaggtc	tatataagca	gagctcgttt	agtgaaccgt	gtttaaacct	2820
cttcagagac	tccctcctcc	ccaagctctg	tcttctggca	acctgcctgg	ttgccgtgga	2880
aacaggttcc	actgcggaca	aaggaggagg	ctgggtcctg	cttcctcctg	gtcttgtcga	2940
tgaggatttt	tagaccgtgg	agactgcgct	gccctgccct	gcacctacc	tcactctccg	3000
tgttctcact	aaggtggaaa	ttgccttccc	tcactactga	cgagaccatg	tgtaaaagcg	3060
tgaccacaga	tgagtggaa	aaagtcttct	atgagaagat	ggaggaggca	aagccggctg	3120
acagctggga	cctcatcata	gacccaacc	tcaagcaca	tgtgctgagc	cctgggttga	3180
agcagtacct	ggaattgcat	gcttcaggca	ggtgagtagc	ccaggaaagt	ggatccctgc	3240
agggcgctc	taggtcccta	gctctggggc	accttccaag	gagaggaaga	ttacgtagaa	3300
cccaagtgtt	tagcttcaat	ctcactatta	ggctggcgta	gactggaagt	cagagaaaga	3360
gtccctaact	gggaactacg	acacttgagt	tggatttcag	ctcttctact	gatcacctgt	3420
gttactcttc	ctctctgagt	cacaattttt	ccgtctggaa	aataaagaca	tagaatatac	3480
gtatgagtcc	tacacactga	cattttacat	attttctatt	ttaacagtct	cttaaaaagt	3540
agtttaaaac	cagagaagaa	gggtttgagg	cccactgggg	gtcgagacgt	ccgtgctctg	3600
gtcctgggac	cggtttaaat	ctatttaa				3628

<210> 3
 <211> 928
 <212> ADN
 <213> ratón

5

<400> 3

atgaaaaccc	tcagcagccc	cagcaactcc	agcaccatca	ctggcttcat	cctcttgggc	60
ttgcctacc	ccaggagggg	gcaaattctc	ctctttgtga	tcttcttcat	tgtttacata	120
ctcattctta	tgggcaacgc	ttccatcatc	tgtgctgtgt	actgtgatca	gagactccac	180
accccatgt	accttctgct	ggccaacttc	tccttcatgg	agattggata	tgtcacctcc	240
acagtcccca	acatgttggc	caacttcctt	tcagacacca	aggtcatctc	tttctctgga	300
tgcttctctg	agttctatct	cttcttctcc	tttggttcta	cagaatgctt	tttctctgga	360
gtcatggcat	ttgatcgata	ccttgccatc	tgtaggccac	tacattatcc	ttctctcatg	420
actgggcgcc	tccgaaacac	ccttgtgacc	agttgctggg	tgcttgggtt	cctctgggtc	480
cctgtaccca	tcacatcat	ctcccagatg	tccttctgtg	ggccagaat	tatagaccac	540
ttctgtgtg	accaggccc	tcttttggcc	cttgcctgtt	ccagagtccc	attgatagag	600
gttttctggt	ccattataat	gtctatgctc	ctggttatcc	cttctctctt	catcatggga	660

ES 2 802 524 T3

acttacatat	tggtcctaag	agctgtgttt	agacttcctt	caagagaagg	acaaaaaaag	720
gctttctcca	cttgccgggtc	tcatctcaca	gtagtttcac	tctttttattg	ctcagtgatg	780
ataatgtatc	tgagcccaac	atctgagcat	gaggccggaa	tgcagaagct	tgtaaactcta	840
ttttattctg	tgggtacacc	actgcttaat	cctatgatat	acagttctgag	gaacaaagat	900
atgaaaaatg	ccctacagaa	gattttga				928

<210> 4
<211> 311
<212> PRT
<213> ratón

<400> 4

Met 1	Lys	Thr	Leu	Ser 5	Ser	Pro	Ser	Asn 10	Ser	Thr	Ile	Thr	Gly 15	Phe	
Ile	Leu	Leu	Gly 20	Phe	Ala	Tyr	Pro	Arg 25	Glu	Gly	Gln	Ile	Leu 30	Leu	Phe
Val	Ile	Phe 35	Phe	Ile	Val	Tyr	Ile 40	Leu	Ile	Leu	Met	Gly 45	Asn	Ala	Ser
Ile	Ile 50	Cys	Ala	Val	Tyr	Cys 55	Asp	Gln	Arg	Leu	His 60	Thr	Pro	Met	Tyr
Leu 65	Leu	Leu	Ala	Asn	Phe 70	Ser	Phe	Met	Glu	Ile 75	Gly	Tyr	Val	Thr	Ser 80
Thr	Val	Pro	Asn	Met 85	Leu	Ala	Asn	Phe	Leu 90	Ser	Asp	Thr	Lys	Val 95	Ile
Ser	Phe	Ser	Gly 100	Cys	Phe	Leu	Gln	Phe 105	Tyr	Phe	Phe	Phe	Ser 110	Phe	Gly
Ser	Thr	Glu 115	Cys	Phe	Phe	Leu	Ala 120	Val	Met	Ala	Phe	Asp 125	Arg	Tyr	Leu
Ala	Ile 130	Cys	Arg	Pro	Leu	His 135	Tyr	Pro	Ser	Leu	Met 140	Thr	Gly	Arg	Leu
Arg 145	Asn	Thr	Leu	Val	Thr 150	Ser	Cys	Trp	Val	Leu 155	Gly	Phe	Leu	Trp	Phe 160
Pro	Val	Pro	Ile 165	Ile	Ile	Ile	Ser	Gln	Met 170	Ser	Phe	Cys	Gly	Ser 175	Arg

ES 2 802 524 T3

Ile Ile Asp His Phe Leu Cys Asp Pro Gly Pro Leu Leu Ala Leu Ala
180 185 190

Cys Ser Arg Val Pro Leu Ile Glu Val Phe Trp Ser Ile Ile Met Ser
195 200 205

Met Leu Leu Val Ile Pro Phe Leu Phe Ile Met Gly Thr Tyr Ile Leu
210 215 220

Val Leu Arg Ala Val Phe Arg Leu Pro Ser Arg Glu Gly Gln Lys Lys
225 230 235 240

Ala Phe Ser Thr Cys Gly Ser His Leu Thr Val Val Ser Leu Phe Tyr
245 250 255

Cys Ser Val Met Ile Met Tyr Leu Ser Pro Thr Ser Glu His Glu Ala
260 265 270

Gly Met Gln Lys Leu Val Thr Leu Phe Tyr Ser Val Gly Thr Pro Leu
275 280 285

Leu Asn Pro Met Ile Tyr Ser Leu Arg Asn Lys Asp Met Lys Asn Ala
290 295 300

Leu Gln Lys Ile Leu Arg Thr
305 310

<210> 5
<211> 936
<212> ADN
<213> ratón

5

<400> 5

atgaaaaccc tcagcagccc cagcaactcc agcaccatca ctggcttcat cctcttgggc	60
ttcccctgcc ccagggaggg gcaaactctc ctctttgtga ccttcttcat tgtttacata	120
ctcattctta tgggcaatgc ttccatcatc tgtgctgtgt actgtgatca ggcctccac	180
acccccatgt acttctgct ggccaacttc tccttctgga agatctggta tgtcacctcc	240
acagtcccca acatgttggc caacttctt tcagacacca aggtcatctc tttctctgga	300
tgcttctgct agttctatct cttcttctcc tttggttcta cagaatgctt tttctctgga	360
gtcatggcat ttgatcgata ccttgccatc tgtaggccac tacattatcc ttctctcatg	420
actgggcacc tctgcaacat ccttgtgatc agttgctggg tgcttggttt cctctgggtc	480
cctgtaccca tcatcatcat ctcccagatg tccttctgtg ggtccagaat tatagaccac	540
ttctgtgtg acccaggccc tcttttggcc cttgcctgtt ccagagcccc attgatggag	600
gttttctgga caattataat gtctatgctc ctggttatcc ctttctctt catcatggga	660

ES 2 802 524 T3

acttacatat	tggtcctaag	agctgtgttt	agacttcctt	caagagatgg	acaaaaaaag	720
gccttctcca	cttgccgggtc	tcatctcaca	gtagtttcac	tctttttattg	ctcagtgatg	780
aaaatgtatt	tgagcccaac	atctgagcat	gaagctggaa	tgcagaagct	tgtaactcta	840
ttttattctg	tgggtactcc	actacttaat	cctgtgatat	acagttctgag	gaacaaaagat	900
atgaaaaatg	ccctgcagaa	gattttaaga	acataa			936

<210> 6
<211> 311
<212> PRT
<213> ratón

<400> 6

Met 1	Lys	Thr	Leu	Ser 5	Ser	Pro	Ser	Asn 10	Ser	Thr	Ile	Thr	Gly 15	Phe	
Ile	Leu	Leu	Gly 20	Phe	Pro	Cys	Pro	Arg 25	Glu	Gly	Gln	Ile	Leu 30	Leu	Phe
Val	Thr	Phe 35	Phe	Ile	Val	Tyr	Ile 40	Leu	Ile	Leu	Met	Gly 45	Asn	Ala	Ser
Ile	Ile 50	Cys	Ala	Val	Tyr	Cys 55	Asp	Gln	Ser	Leu	His 60	Thr	Pro	Met	Tyr
Phe 65	Leu	Leu	Ala	Asn	Phe 70	Ser	Phe	Leu	Glu	Ile 75	Trp	Tyr	Val	Thr	Ser 80
Thr	Val	Pro	Asn	Met 85	Leu	Ala	Asn	Phe	Leu 90	Ser	Asp	Thr	Lys	Val 95	Ile
Ser	Phe	Ser	Gly 100	Cys	Phe	Leu	Gln	Phe 105	Tyr	Phe	Phe	Phe	Ser 110	Phe	Gly
Ser	Thr	Glu 115	Cys	Phe	Phe	Leu	Ala 120	Val	Met	Ala	Phe	Asp 125	Arg	Tyr	Leu
Ala	Ile 130	Cys	Arg	Pro	Leu	His 135	Tyr	Pro	Ser	Leu	Met 140	Thr	Gly	His	Leu
Cys 145	Asn	Ile	Leu	Val	Ile 150	Ser	Cys	Trp	Val	Leu 155	Gly	Phe	Leu	Trp	Phe 160
Pro	Val	Pro	Ile 165	Ile	Ile	Ile	Ser	Gln	Met 170	Ser	Phe	Cys	Gly	Ser 175	Arg

ES 2 802 524 T3

Ile Ile Asp His Phe Leu Cys Asp Pro Gly Pro Leu Leu Ala Leu Ala
180 185 190

Cys Ser Arg Ala Pro Leu Met Glu Val Phe Trp Thr Ile Ile Met Ser
195 200 205

Met Leu Leu Val Ile Pro Phe Leu Phe Ile Met Gly Thr Tyr Ile Leu
210 215 220

Val Leu Arg Ala Val Phe Arg Leu Pro Ser Arg Asp Gly Gln Lys Lys
225 230 235 240

Ala Phe Ser Thr Cys Gly Ser His Leu Thr Val Val Ser Leu Phe Tyr
245 250 255

Cys Ser Val Met Lys Met Tyr Leu Ser Pro Thr Ser Glu His Glu Ala
260 265 270

Gly Met Gln Lys Leu Val Thr Leu Phe Tyr Ser Val Gly Thr Pro Leu
275 280 285

Leu Asn Pro Val Ile Tyr Ser Leu Arg Asn Lys Asp Met Lys Asn Ala
290 295 300

Leu Gln Lys Ile Leu Arg Thr
305 310

<210> 7
<211> 942
<212> ADN
<213> ratón

5

<400> 7

atgggaatcc tttccacagg aaatcaaact gtcactgagt ttgtacttct tggtttccat	60
gaagtccttg ggctgcacct cctgtttttt tctgtgttca ccatcctcta tgcctccatc	120
atcacaggga acatgctcat tgcagtggtg gtggtgagct cccagaggct tcacacaccc	180
atgtatttct ttctggtgaa tctgtccttc atagagattg tctataacct cacagtgggtg	240
cccaaaatgc tggaaggctt cttacaggag gccaccatat ctgtggctgg ctgcttgctc	300
cagttctttg tttttggctc tctggccaca gatgagtgtt ttctgctggc tgtgatggca	360
tatgatcgat atctcgcaat ttgtcaccct ctacgatacc cacacctcat ggggcctcaa	420
tgggtgcctgg ggttggtgct cacagtctgg ctgtctggct tcatggtaga tggactagtt	480
gttgctctga tggcccagtt gagattctgt ggcccaact tagttgatca cttttactgt	540
gatttttcac ctttgatggt cctggcttgc tcagataccc aagtggcca ggtgactaca	600
tttgttctct ctgtggtcct cctgactgtc ccctttgggc tggttctgat ctccatgct	660

ES 2 802 524 T3

cagattgtag	tgactgtgct	gagagtctct	tctgggacca	gaagaaccaa	ggccttctcc	720
acatgctcct	ctcacctggc	tgtggtgtcc	acgttctatg	gaacactcat	ggtattgtac	780
attgtgccct	ctgctgttca	ttctcagctc	ctctccaagg	tcattgccct	gctctacaca	840
gtggtcactc	ccatcttcaa	ccctgtcatc	tacaccttga	ggaaccagga	ggtgcagcag	900
gcactaagaa	ggcttctcta	ctgcaaacca	actgaaatgt	ga		942

<210> 8
<211> 313
<212> PRT
<213> ratón

<400> 8

Met	Gly	Ile	Leu	Ser	Thr	Gly	Asn	Gln	Thr	Val	Thr	Glu	Phe	Val	Leu
1				5				10						15	
Leu	Gly	Phe	His	Glu	Val	Pro	Gly	Leu	His	Leu	Leu	Phe	Phe	Ser	Val
			20					25					30		
Phe	Thr	Ile	Leu	Tyr	Ala	Ser	Ile	Ile	Thr	Gly	Asn	Met	Leu	Ile	Ala
		35					40					45			
Val	Val	Val	Val	Ser	Ser	Gln	Arg	Leu	His	Thr	Pro	Met	Tyr	Phe	Phe
	50					55					60				
Leu	Val	Asn	Leu	Ser	Phe	Ile	Glu	Ile	Val	Tyr	Thr	Ser	Thr	Val	Val
65					70					75					80
Pro	Lys	Met	Leu	Glu	Gly	Phe	Leu	Gln	Glu	Ala	Thr	Ile	Ser	Val	Ala
				85					90					95	
Gly	Cys	Leu	Leu	Gln	Phe	Phe	Val	Phe	Gly	Ser	Leu	Ala	Thr	Asp	Glu
			100					105					110		
Cys	Phe	Leu	Leu	Ala	Val	Met	Ala	Tyr	Asp	Arg	Tyr	Leu	Ala	Ile	Cys
		115					120					125			
His	Pro	Leu	Arg	Tyr	Pro	His	Leu	Met	Gly	Pro	Gln	Trp	Cys	Leu	Gly
	130					135					140				
Leu	Val	Leu	Thr	Val	Trp	Leu	Ser	Gly	Phe	Met	Val	Asp	Gly	Leu	Val
145					150					155					160
Val	Ala	Leu	Met	Ala	Gln	Leu	Arg	Phe	Cys	Gly	Pro	Asn	Leu	Val	Asp
				165					170					175	

ES 2 802 524 T3

His Phe Tyr Cys Asp Phe Ser Pro Leu Met Val Leu Ala Cys Ser Asp
180 185 190

Thr Gln Val Ala Gln Val Thr Thr Phe Val Leu Ser Val Val Phe Leu
195 200 205

Thr Val Pro Phe Gly Leu Val Leu Ile Ser Tyr Ala Gln Ile Val Val
210 215 220

Thr Val Leu Arg Val Pro Ser Gly Thr Arg Arg Thr Lys Ala Phe Ser
225 230 235 240

Thr Cys Ser Ser His Leu Ala Val Val Ser Thr Phe Tyr Gly Thr Leu
245 250 255

Met Val Leu Tyr Ile Val Pro Ser Ala Val His Ser Gln Leu Leu Ser
260 265 270

Lys Val Ile Ala Leu Leu Tyr Thr Val Val Thr Pro Ile Phe Asn Pro
275 280 285

Val Ile Tyr Thr Leu Arg Asn Gln Glu Val Gln Gln Ala Leu Arg Arg
290 295 300

Leu Leu Tyr Cys Lys Pro Thr Glu Met
305 310

<210> 9
<211> 948
<212> ADN
<213> ser humano

5

<400> 9

atggaaattg tctccacagg aaacgaaact attactgaat ttgtcctcct tggcttctat	60
gacatccctg aactgcattt cttgtttttt attgtattca ctgctgtcta tgtcttcac	120
atcatagggg atatgctgat tattgttagca gtggtttagct cccagaggct ccacaaaccc	180
atgtatatatt tcttggcgaa tctgtccttc ctggatatatc tctacacctc cgcagtgatg	240
ccaaaaatgc tggagggtct cctgcaagaa gcaactatct ctgtggctgg ttgcttgctc	300
cagttcttta tcttcggctc tctagccaca gctgaatgct tactgctggc tgtcatggca	360
tatgaccgct acctggcaat ttgctaccca ctccactacc cactcctgat ggggccaga	420
cggtacatgg ggctggtggt cacaacctgg ctctctggat ttgtggtaga tggactggtt	480
gtggccctgg tggcccagct gaggttctgt ggccccaaacc acattgacca gttttactgt	540
gactttatgc ttttcgtggg cctggcttgc tcggatccca gagtggctca ggtgacaact	600
ctcattctgt ctgtgttctg cctcactatt ccttttggac tgattctgac atcttatgcc	660

ES 2 802 524 T3

```

agaattgtgg tggcagtgct gagagttcct gctggggcaa gcaggagaag ggcttttctcc      720
acatgctcct cccacctagc tgtagtgacc acattctatg gaacgctcat gatcttttat      780
gttgccacct ctgctgtcca ttcccagctc ctctccaagg tcttctccct gctctacact      840
gtggtcaccc ctctcttcaa tcctgtgatc tataccatga ggaacaagga ggtgcatcag      900
gcacttcgga agattctctg tatcaaacia actgaaacac ttgattga                      948

```

<210> 10
 <211> 315
 <212> PRT
 <213> ser humano

5

<400> 10

```

Met Glu Ile Val Ser Thr Gly Asn Glu Thr Ile Thr Glu Phe Val Leu
 1                               5                               10          15

Leu Gly Phe Tyr Asp Ile Pro Glu Leu His Phe Leu Phe Phe Ile Val
                20                               25          30

Phe Thr Ala Val Tyr Val Phe Ile Ile Ile Gly Asn Met Leu Ile Ile
                35                               40          45

Val Ala Val Val Ser Ser Gln Arg Leu His Lys Pro Met Tyr Ile Phe
 50                               55          60

Leu Ala Asn Leu Ser Phe Leu Asp Ile Leu Tyr Thr Ser Ala Val Met
65                               70          75          80

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

Gly Cys Leu Leu Gln Phe Phe Ile Phe Gly Ser Leu Ala Thr Ala Glu
                100                              105          110

Cys Leu Leu Leu Ala Val Met Ala Tyr Asp Arg Tyr Leu Ala Ile Cys
                115                              120          125

Tyr Pro Leu His Tyr Pro Leu Leu Met Gly Pro Arg Arg Tyr Met Gly
                130                              135          140

Leu Val Val Thr Thr Trp Leu Ser Gly Phe Val Val Asp Gly Leu Val
145                              150          155          160

Val Ala Leu Val Ala Gln Leu Arg Phe Cys Gly Pro Asn His Ile Asp
                165                              170          175

```


ES 2 802 524 T3

Gln Phe Tyr Cys Asp Phe Met Leu Phe Val Gly Leu Ala Cys Ser Asp
180 185 190

Pro Arg Val Ala Gln Val Thr Thr Leu Ile Leu Ser Val Phe Cys Leu
195 200 205

Thr Ile Pro Phe Gly Leu Ile Leu Thr Ser Tyr Ala Arg Ile Val Val
210 215 220

Ala Val Leu Arg Val Pro Ala Gly Ala Ser Arg Arg Arg Ala Phe Ser
225 230 235 240

Thr Cys Ser Ser His Leu Ala Val Val Thr Thr Phe Tyr Gly Thr Leu
245 250 255

Met Ile Phe Tyr Val Ala Pro Ser Ala Val His Ser Gln Leu Leu Ser
260 265 270

Lys Val Phe Ser Leu Leu Tyr Thr Val Val Thr Pro Leu Phe Asn Pro
275 280 285

Val Ile Tyr Thr Met Arg Asn Lys Glu Val His Gln Ala Leu Arg Lys
290 295 300

Ile Leu Cys Ile Lys Gln Thr Glu Thr Leu Asp
305 310 315

<210> 11
<211> 936
<212> ADN
<213> ratón

5

<400> 11

atgaaaacct tcagcagccc catcaactcc agcaccacca ctggcttcat tctcttgggc	60
ttcccctgcc ccaggaggagg gcaaatactc ctctttgtgc tcttctccat tgtctacctg	120
cttaccctca tgggcaaacac ttgcatcatc ttgagctat gctgggatca gagactccac	180
acacccatgt acctactgct ggccaacttc tccttctctg agatctggta tgttacctcc	240
acagtcccca acatgttggc caatttcctc tctgacacca aggtcatctc tttctctgga	300
tgcttctctgc agttctatct cttcttctcc ttgggttcta cagaatgcct tttctctgga	360
gtcatggcat ttgatcgata ccttgccatc tgtaggccac tacattatcc tgctctcatg	420
actgggagcc tctgcaacat ccttgtgatc agttgctggg tgcttggttt cctctgggtc	480
cctgttccca tcatcatcat ctcccagatg tccttctctg ggtccagaat tatagaccac	540
ttcctgtgtg acccaggccc tctattggcc ctcacctgtt ccagagcccc attaatggag	600
gttttctgga caattataac atctcttacc ctgttcgttc ctttctctt catcatggga	660

ES 2 802 524 T3

tcttatacat tggctcctgag agctgtgttc agagttcctt caagagatgg acaaaaaaag 720
gctttctcca cttgcggatc tcatctcaca gtagttttac tcttttatgg ctcaagtgatg 780
ataatgtatc taagcccgac ctctgagcat gaagctggaa tgcagaagct tgtgactcta 840
ttttattctg tggttactcc actcattaat cctgtgatat acagtctgag gaacaaggat 900
atgaaacatg ccctgcagaa gattttaaga acataa 936

<210> 12
<211> 311
<212> PRT
<213> ratón

5

<400> 12

Met	Lys	Thr	Phe	Ser	Ser	Pro	Ile	Asn	Ser	Ser	Thr	Thr	Thr	Gly	Phe	1	5	10	15
Ile	Leu	Leu	Gly	Phe	Pro	Cys	Pro	Arg	Glu	Gly	Gln	Ile	Leu	Leu	Phe	20	25	30	
Val	Leu	Phe	Ser	Ile	Val	Tyr	Leu	Leu	Thr	Leu	Met	Gly	Asn	Thr	Cys	35	40	45	
Ile	Ile	Phe	Ala	Val	Cys	Trp	Asp	Gln	Arg	Leu	His	Thr	Pro	Met	Tyr	50	55	60	
Leu	Leu	Leu	Ala	Asn	Phe	Ser	Phe	Leu	Glu	Ile	Trp	Tyr	Val	Thr	Ser	65	70	75	80
Thr	Val	Pro	Asn	Met	Leu	Ala	Asn	Phe	Leu	Ser	Asp	Thr	Lys	Val	Ile	85	90	95	
Ser	Phe	Ser	Gly	Cys	Phe	Leu	Gln	Phe	Tyr	Phe	Phe	Phe	Ser	Leu	Gly	100	105	110	
Ser	Thr	Glu	Cys	Leu	Phe	Leu	Ala	Val	Met	Ala	Phe	Asp	Arg	Tyr	Leu	115	120	125	
Ala	Ile	Cys	Arg	Pro	Leu	His	Tyr	Pro	Ala	Leu	Met	Thr	Gly	Ser	Leu	130	135	140	
Cys	Asn	Ile	Leu	Val	Ile	Ser	Cys	Trp	Val	Leu	Gly	Phe	Leu	Trp	Phe	145	150	155	160
Pro	Val	Pro	Ile	Ile	Ile	Ile	Ser	Gln	Met	Ser	Phe	Cys	Gly	Ser	Arg	165	170	175	

ES 2 802 524 T3

Ile Ile Asp His Phe Leu Cys Asp Pro Gly Pro Leu Leu Ala Leu Thr
180 185 190

Cys Ser Arg Ala Pro Leu Met Glu Val Phe Trp Thr Ile Ile Thr Ser
195 200 205

Leu Ile Leu Phe Val Pro Phe Leu Phe Ile Met Gly Ser Tyr Thr Leu
210 215 220

Val Leu Arg Ala Val Phe Arg Val Pro Ser Arg Asp Gly Gln Lys Lys
225 230 235 240

Ala Phe Ser Thr Cys Gly Ser His Leu Thr Val Val Leu Leu Phe Tyr
245 250 255

Gly Ser Val Met Ile Met Tyr Leu Ser Pro Thr Ser Glu His Glu Ala
260 265 270

Gly Met Gln Lys Leu Val Thr Leu Phe Tyr Ser Val Val Thr Pro Leu
275 280 285

Ile Asn Pro Val Ile Tyr Ser Leu Arg Asn Lys Asp Met Lys His Ala
290 295 300

Leu Gln Lys Ile Leu Arg Thr
305 310

<210> 13
<211> 930
<212> ADN
<213> ser humano

5

<400> 13

atgagggaaa ataaccagtc ctctacactg gaattcatcc tcctgggagt tactggtcag	60
caggaacagg aagattttctt ctacatcctc ttcttgttca tttaccccat cacattgatt	120
ggaaacctgc tcatcgctct agccatttgc tctgatgttc gccttcacaa ccccatgtat	180
tttctccttg ccaacctctc ctgggttgac atcttcttct catcggtaac catccctaag	240
atgctggcca accatctctt gggcagcaaa tocatctctt ttgggggatg cctaacgcag	300
atgtatttca tgatagcctt gggtaacaca gacagctata ttttggctgc aatggcatat	360
gatcgagctg tggccatcag ccgcccactt cactacacaa caattatgag tccacggtct	420
tgtatctggc ttattgctgg gtcttgggtg attggaaatg ccaatgccct cccccacact	480
ctgctcacag ctagtctgtc cttctgtggc aaccaggaag tggccaactt ctactgtgac	540
attaccccct tgctgaagtt atcctgttct gacatccact ttcattgtgaa gatgatgtac	600
ctaggggttg gcattttctc tgtgccatta ctatgcatca ttgtctocta tattcgagtc	660

ES 2 802 524 T3

```

ttctccacag tcttcaggt tccttcacc aagggcgtgc tcaaggcctt ctccacctgt      720
ggttcccacc tcacggttgt ctctttgtat tatggtacag tcatgggcac gtatttcgcg      780
cctttgacca attatagcct aaaagacgca gtgatcactg taatgtacac ggcagtgacc      840
ccaatgttaa atcctttcat ctacagtctg agaaatcggg acatgaaggc tgcctcgcg      900
aaactcttca acaagagaat ctcctcgtga                                     930

```

<210> 14

<211> 309

<212> PRT

<213> ser humano

5

<400> 14

```

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

Val Thr Gly Gln Gln Glu Gln Glu Asp Phe Phe Tyr Ile Leu Phe Leu
          20           25           30

Phe Ile Tyr Pro Ile Thr Leu Ile Gly Asn Leu Leu Ile Val Leu Ala
          35           40           45

Ile Cys Ser Asp Val Arg Leu His Asn Pro Met Tyr Phe Leu Leu Ala
 50           55           60

Asn Leu Ser Leu Val Asp Ile Phe Phe Ser Ser Val Thr Ile Pro Lys
65           70           75           80

Met Leu Ala Asn His Leu Leu Gly Ser Lys Ser Ile Ser Phe Gly Gly
          85           90           95

Cys Leu Thr Gln Met Tyr Phe Met Ile Ala Leu Gly Asn Thr Asp Ser
          100          105          110

Tyr Ile Leu Ala Ala Met Ala Tyr Asp Arg Ala Val Ala Ile Ser Arg
          115          120          125

Pro Leu His Tyr Thr Thr Ile Met Ser Pro Arg Ser Cys Ile Trp Leu
          130          135          140

Ile Ala Gly Ser Trp Val Ile Gly Asn Ala Asn Ala Leu Pro His Thr
145          150          155          160

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

```

ES 2 802 524 T3

Phe Tyr Cys Asp Ile Thr Pro Leu Leu Lys Leu Ser Cys Ser Asp Ile
 180 185 190
 His Phe His Val Lys Met Met Tyr Leu Gly Val Gly Ile Phe Ser Val
 195 200 205
 Pro Leu Leu Cys Ile Ile Val Ser Tyr Ile Arg Val Phe Ser Thr Val
 210 215 220
 Phe Gln Val Pro Ser Thr Lys Gly Val Leu Lys Ala Phe Ser Thr Cys
 225 230 235 240
 Gly Ser His Leu Thr Val Val Ser Leu Tyr Tyr Gly Thr Val Met Gly
 245 250 255
 Thr Tyr Phe Arg Pro Leu Thr Asn Tyr Ser Leu Lys Asp Ala Val Ile
 260 265 270
 Thr Val Met Tyr Thr Ala Val Thr Pro Met Leu Asn Pro Phe Ile Tyr
 275 280 285
 Ser Leu Arg Asn Arg Asp Met Lys Ala Ala Leu Arg Lys Leu Phe Asn
 290 295 300
 Lys Arg Ile Ser Ser
 305

<210> 15
 <211> 24
 <212> ADN
 <213> artificial

5

<220>
 <223> etiqueta Flag

<400> 15
 gattacaagg acgacgacga taag 24

10

<210> 16
 <211> 8
 <212> PRT
 <213> artificial

<220>
 <223> etiqueta Flag

15

<400> 16

Asp Tyr Lys Asp Asp Asp Asp Lys
 1 5

<210> 17
 <211> 60
 <212> ADN
 <213> artificial

20

<220>
 <223> etiqueta Rho

<400> 17
 atgaacggga ccgagggccc aaactctac gtgcctttct ccaacaagac gggcgtggtg 60

25

<210> 18

<211> 20
 <212> PRT
 <213> artificial

 5 <220>
 <223> etiqueta Rho

 <400> 18

 Met Asn Gly Thr Glu Gly Pro Asn Phe Tyr Val Pro Phe Ser Asn Lys
 1 5 10 15

 Thr Gly Val Val
 20

 10 <210> 19
 <211> 51
 <212> ADN
 <213> artificial

 <220>
 <223> etiqueta Lucy

 <400> 19
 15 atgagacccc agatcctgct gctcctggcc ctgctgaccc taggcctggc t 51

 <210> 20
 <211> 17
 <212> PRT
 <213> artificial

 20 <220>
 <223> etiqueta Lucy

 <400> 20

 Met Arg Pro Gln Ile Leu Leu Leu Leu Ala Leu Leu Thr Leu Gly Leu
 1 5 10 15

 Ala

REIVINDICACIONES

1. Una célula que comprende un ADN donante introducido en el sitio objetivo genómico corriente arriba de un locus del gen de RTP1 endógeno, en la que el ADN donante comprende un promotor en el que el promotor impulsa la expresión del gen de RTP1, y en la que la célula se deriva de una línea celular HEK293 o HEK293T.
- 5 2. La célula según la reivindicación 1, que comprende además un ácido nucleico que codifica un receptor de olor.
3. La célula según la reivindicación 2, en la que el receptor de olor se selecciona del grupo que consiste en receptores de olor de indol, escatol y almizcle.
4. La célula según una cualquiera de las reivindicaciones 1-3, que comprende una proteína Cas o una proteína dCas fusionada a un dominio de activación transcripcional.
- 10 5. Un procedimiento para activar un gen de RTP1 endógeno en una célula eucariota, en el que la célula se deriva de una línea celular HEK293 o HEK293T, que comprende:
 - a. introducir un ARN guía complementario a un sitio genómico objetivo corriente arriba del gen de RTP1; y
 - b. introducir una proteína nucleasa Cas para producir un complejo con el ARN guía para formar un complejo de ARN guía-proteína Cas.
- 15 6. El procedimiento según la reivindicación 5, que comprende además la introducción de un ADN donante que comprende un promotor en el sitio genómico objetivo corriente arriba de un locus del gen de RTP1 endógeno en el que el promotor impulsa la expresión del gen de RTP1.
7. El procedimiento según la reivindicación 5, en el que el complejo comprende además una proteína dCas9 fusionada a un dominio de activación transcripcional, en el que el dominio de activación impulsa la expresión del gen de RTP1.
- 20 8. El procedimiento según las reivindicaciones 5-7, en el que la proteína Cas es una proteína Cas9 o una nicasa Cas9.
9. El procedimiento según la reivindicación 5, que comprende además la escisión de la secuencia de ácido nucleico objetivo adyacente a la secuencia de ARN guía con el complejo de ARN guía-proteína Cas.
- 25 10. El procedimiento según una cualquiera de las reivindicaciones 5-9, que comprende además la introducción de un ácido nucleico que codifica un receptor de olor en la célula.
11. Un procedimiento de identificación de un compuesto o mezcla de compuestos que activa, imita, bloquea, inhibe, modula y/o mejora la actividad de un receptor olfativo en una célula no olfativa de acuerdo con la reivindicación 2, en el que la célula se deriva de una línea celular HEK293 o HEK293T y comprende un ADN donante introducido en el sitio objetivo genómico corriente arriba de un locus del gen de RTP1 endógeno, en el que el ADN donante comprende un promotor en el que el promotor impulsa la expresión del gen de RTP1, en el que el procedimiento comprende además:
 - a. poner en contacto el receptor, o una quimera o fragmento del mismo con un compuesto o mezcla de compuestos que activa, imita, bloquea, inhibe, modula y/o mejora el receptor; y
 - b. determinar si el compuesto tiene un efecto sobre la actividad del receptor.
- 30 12. El procedimiento según la reivindicación 11, en el que el receptor olfativo es del grupo que consiste en un receptor de almizcle y de mal olor.
13. El procedimiento según la reivindicación 11, en el que el receptor de mal olor se selecciona de un receptor de escatol o indol, y en el que el receptor de almizcle se selecciona de un receptor de almizcle policíclico y de nitro almizcle.

40

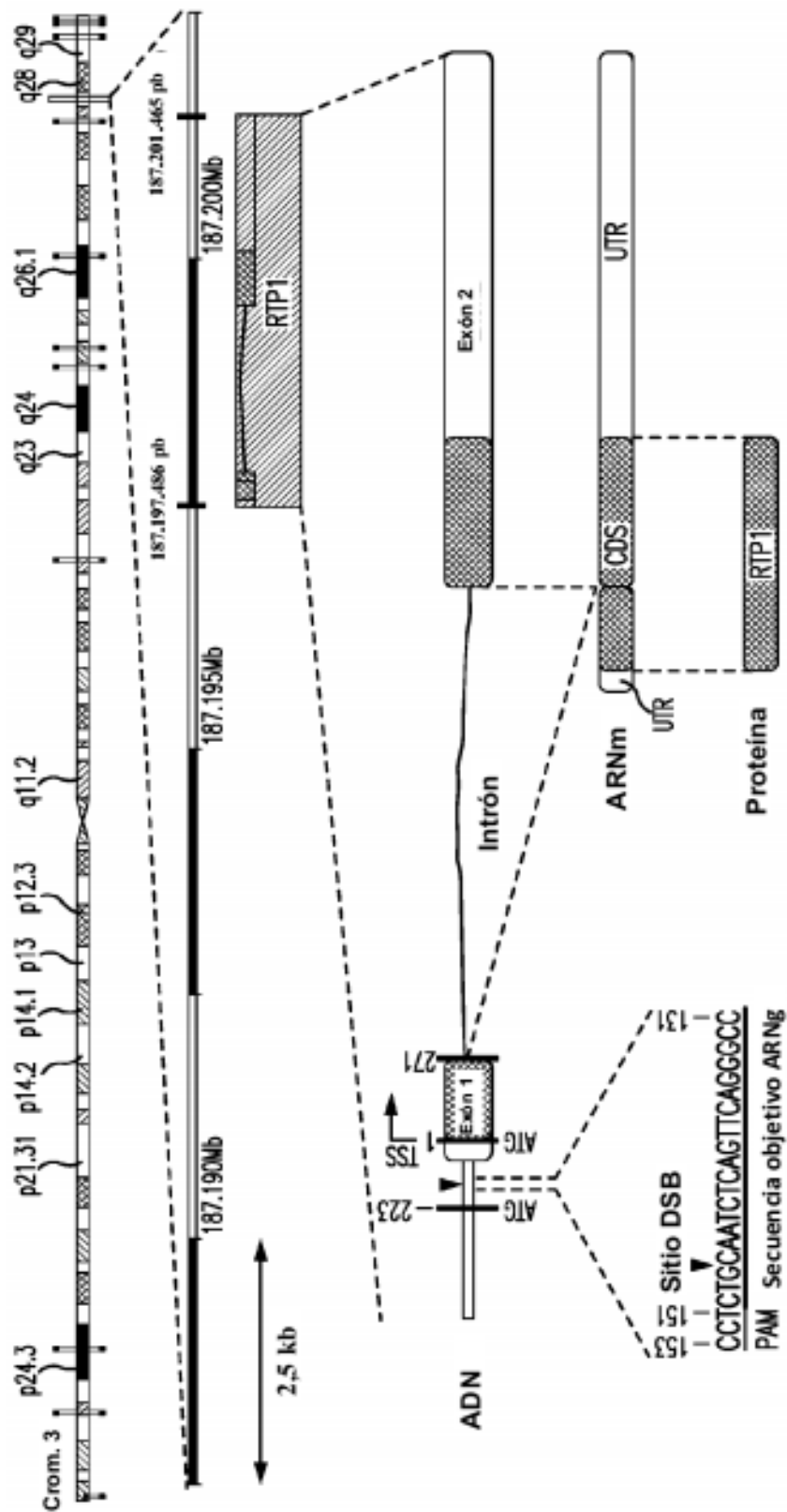


Figura 1

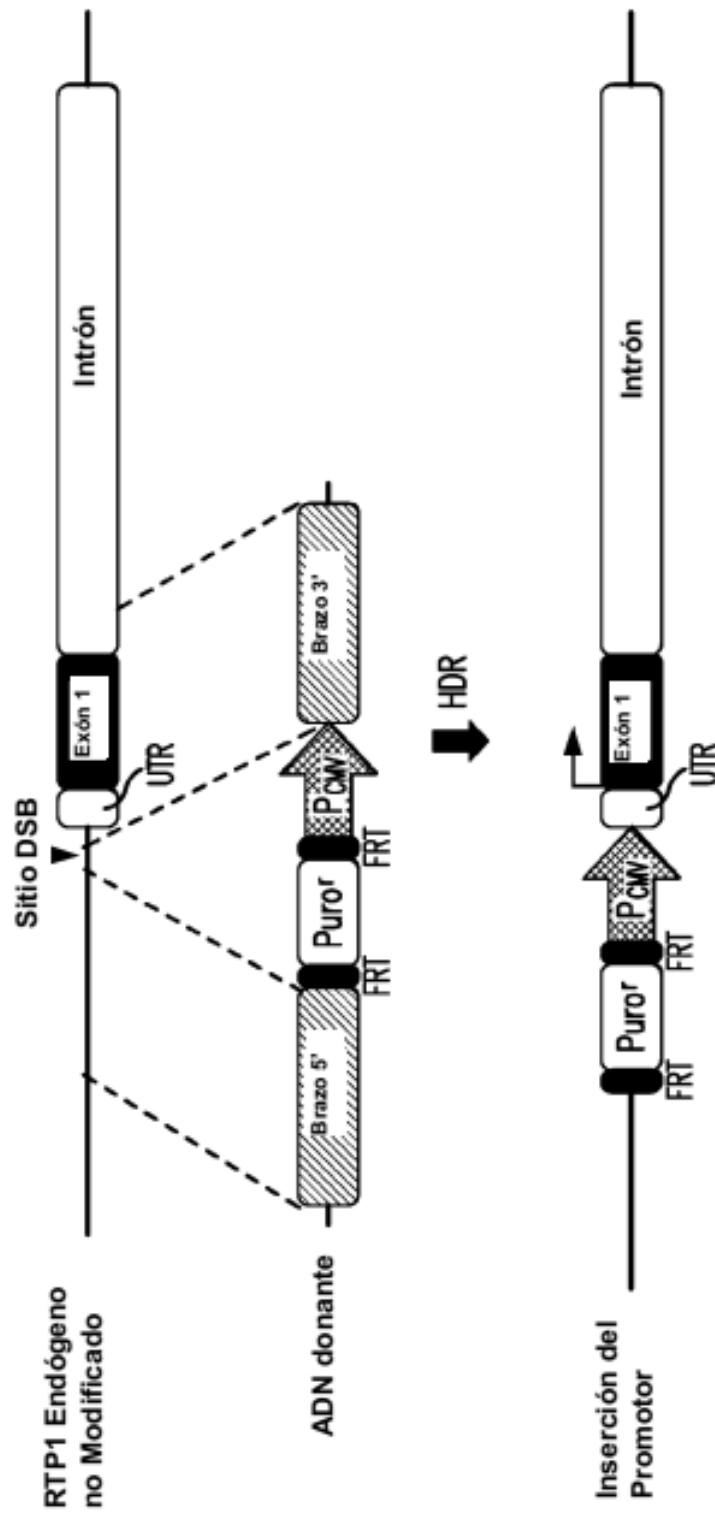


Figura 2

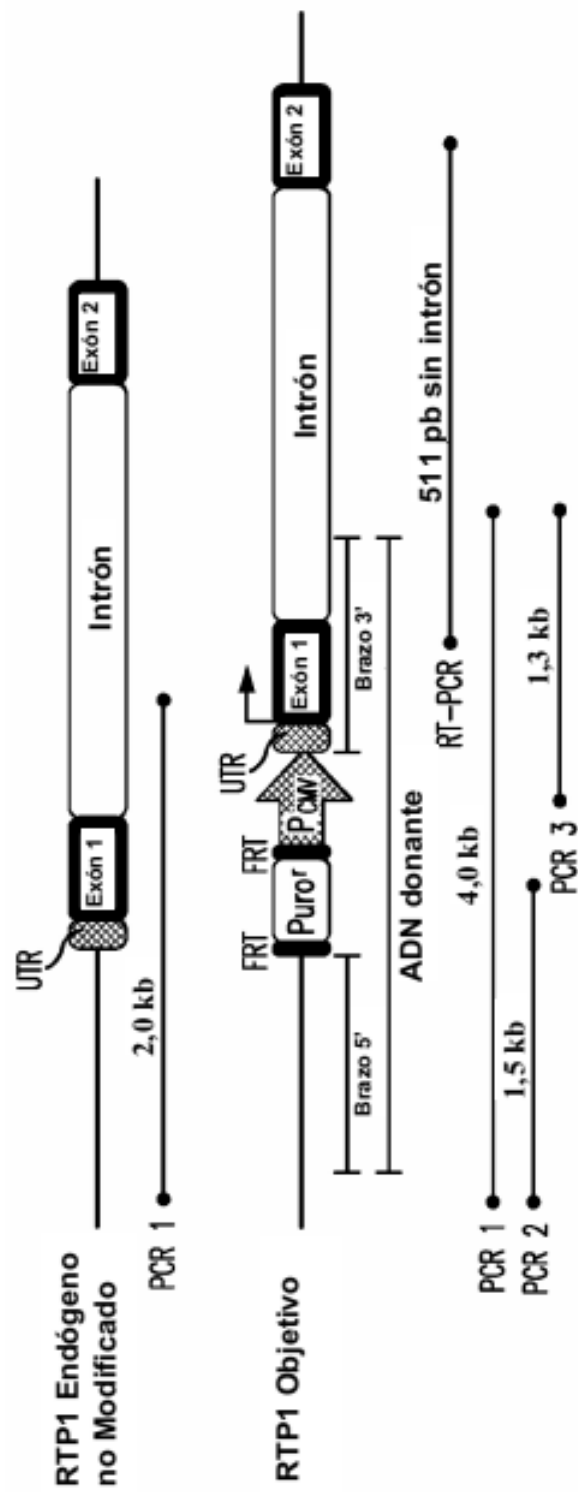


Figura 3

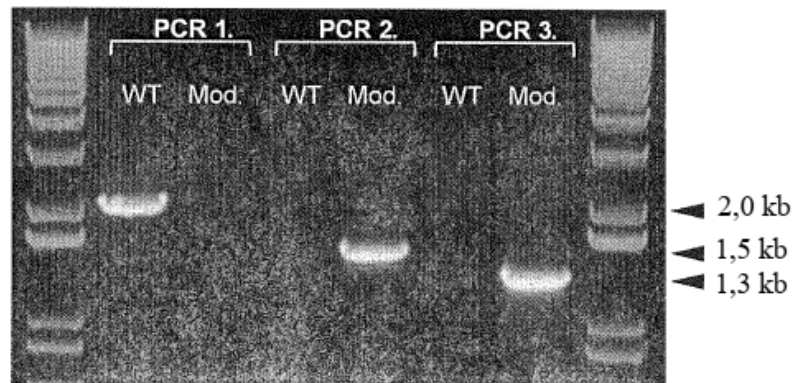


Figura 4

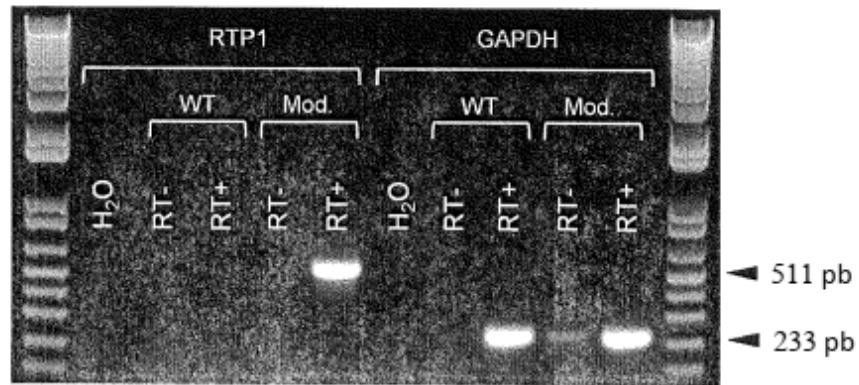


Figura 5

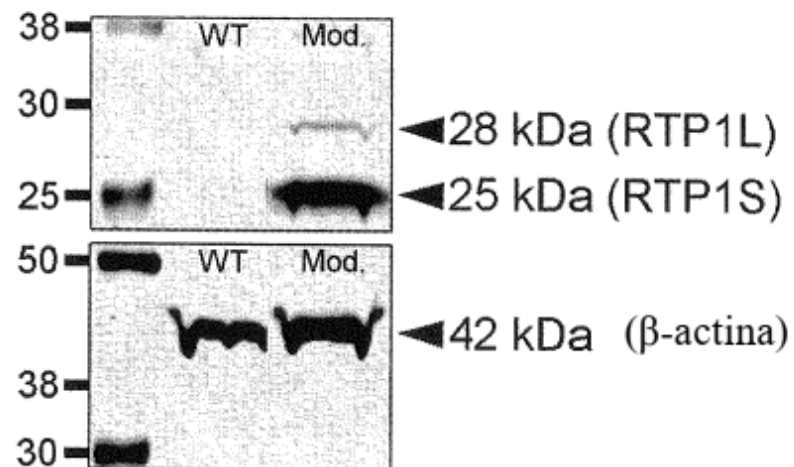


Figura 6

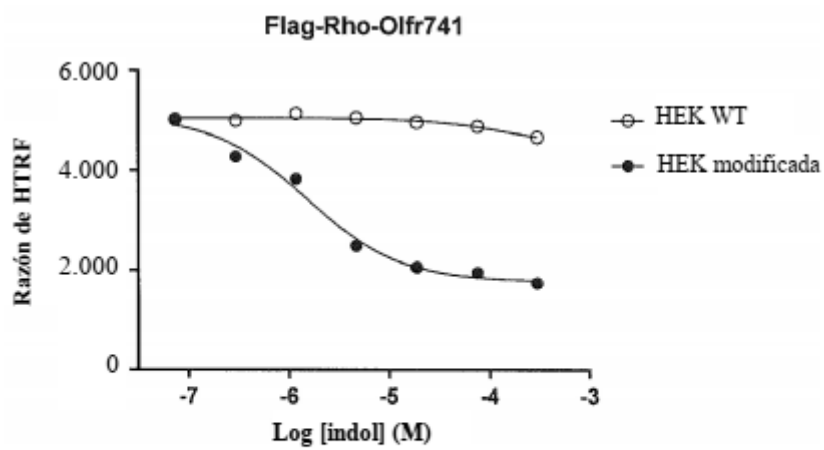


Figura 7

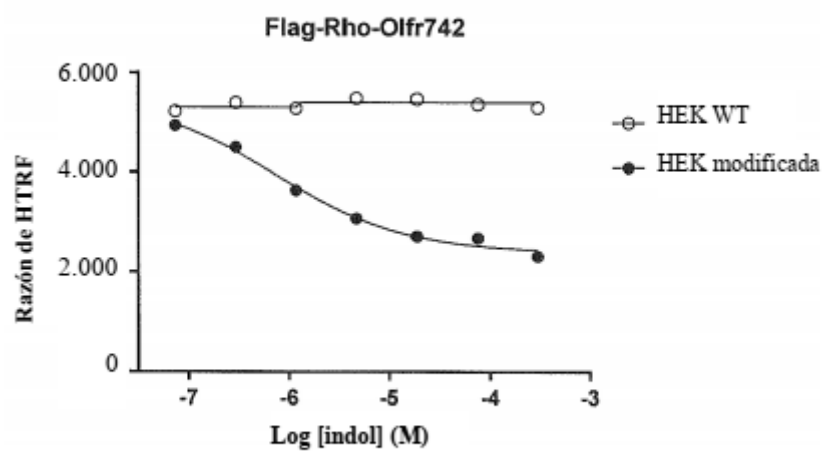


Figura 8

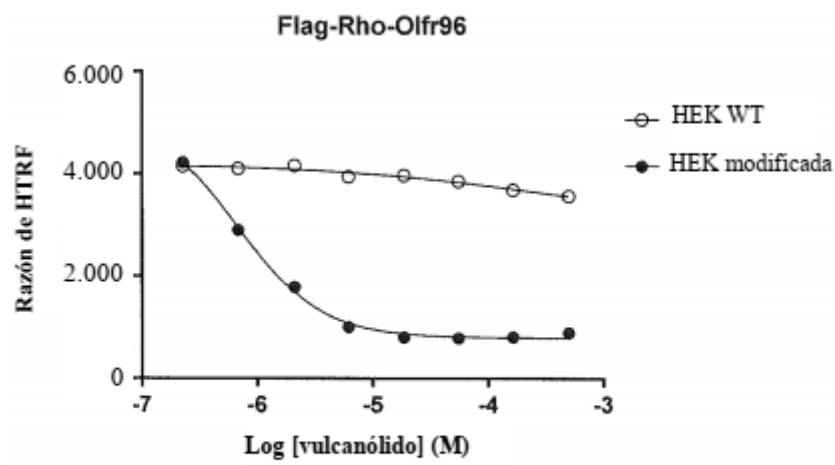


Figura 9

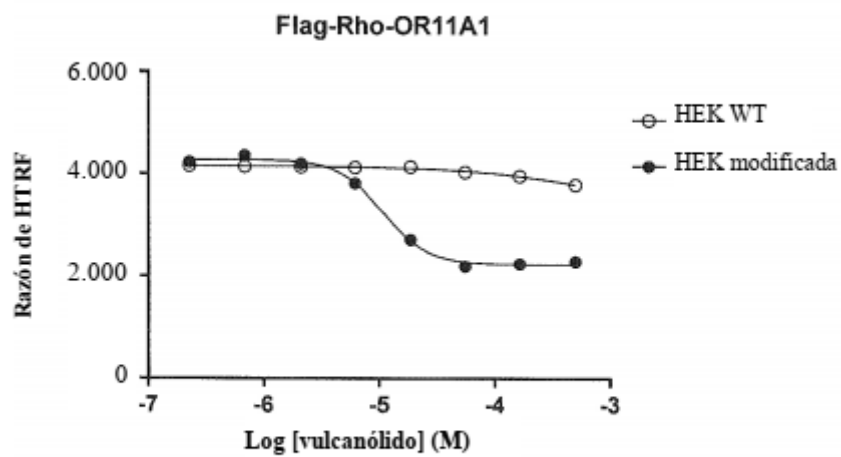


Figura 10

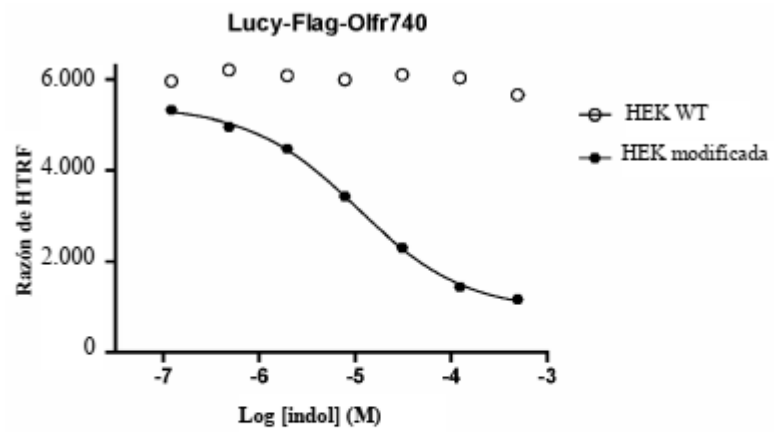


Figura 11

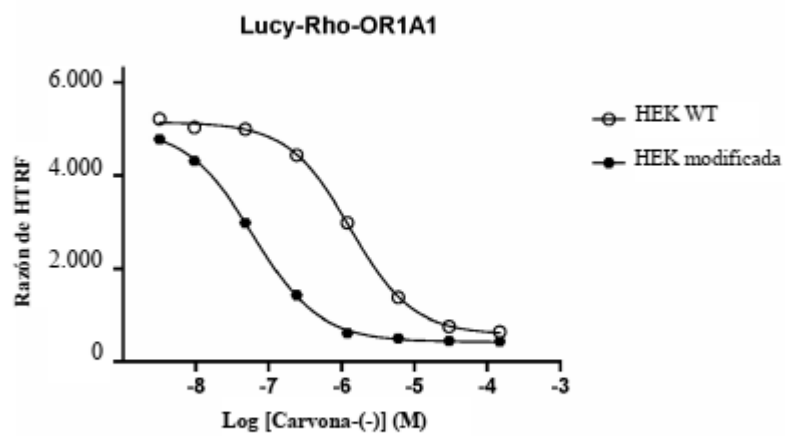


Figura 12