

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 821 938**

51 Int. Cl.:

C07K 14/47 (2006.01)
C12N 15/861 (2006.01)
C12N 15/864 (2006.01)
C12N 9/10 (2006.01)
C12N 9/12 (2006.01)
C12N 9/78 (2006.01)
A61K 48/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **09.03.2015 PCT/US2015/019513**
- 87 Fecha y número de publicación internacional: **17.09.2015 WO15138348**
- 96 Fecha de presentación y número de la solicitud europea: **09.03.2015 E 15712226 (8)**
- 97 Fecha y número de publicación de la concesión europea: **08.07.2020 EP 3116900**

54 Título: **Composiciones útiles en el tratamiento de la deficiencia de ornitina transcarbamilasa (OTC)**

30 Prioridad:

09.03.2014 US 201461950157 P

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

28.04.2021

73 Titular/es:

**THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA (100.0%)
3600 Civic Center Boulevard, 9th Floor
Philadelphia, PA 19104, US**

72 Inventor/es:

**WANG, LILI y
WILSON, JAMES M.**

74 Agente/Representante:

TEMIÑO CENICEROS, Ignacio

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 821 938 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

Composiciones útiles en el tratamiento de la deficiencia de ornitina transcarbamilasa (OTC)

5 **Declaración con respecto a la investigación o desarrollo con fondos federales**

Este trabajo fue apoyado en parte por subvenciones de los Institutos Nacionales de Salud, n.º P01-HD057247, P01-HL059407 y P30-DK047757. El gobierno de los Estados Unidos puede tener determinados derechos sobre la presente invención.

10

Antecedentes de la invención

La deficiencia de ornitina transcarbamilasa (OTC) representa casi la mitad de todos los casos de errores innatos de síntesis de urea, con una prevalencia estimada en al menos 1 de cada 15.000. Los defectos del ciclo de la urea ponen a los pacientes en riesgo de una elevación del amoníaco potencialmente mortal que puede conducir a deterioro cognitivo irreversible, coma y muerte. Los varones recién nacidos con deficiencia completa desarrollan coma hiperamonémico en los primeros 3 días de vida, que, si no se trata, es letal.

15

Las terapias actuales para la deficiencia de OTC (OTCD) tienen numerosos desafíos. Los pacientes pueden tratarse con una dieta baja en proteínas en combinación con el uso de medicamentos que activan rutas alternativas de eliminación de nitrógeno, pero esto no evita las crisis hiperamonémicas. A pesar del uso de diálisis y terapia de ruta alternativa, hay una tasa de mortalidad de casi 50 % en los recién nacidos. El trasplante de hígado puede curar la OTCD, pero el hígado del donante es limitante, el procedimiento conlleva una morbilidad significativa y son necesarios fármacos inmunosupresores durante toda la vida del sujeto.

20

25

La terapia génica de una enfermedad metabólica tal como la OTCD presenta un modelo más difícil para la terapia de reemplazo génico que otras afecciones. Debido a que el gen actúa de manera autónoma celular (es decir, solo puede influir en la célula en la que se expresa), los efectos terapéuticos deberían estar directamente relacionados con el número de células diana que se transducen, en lugar de con el nivel neto de expresión en el hígado, tal como con una proteína secretada donde una alta expresión por célula puede superar la baja transducción. Asimismo, ha habido al menos un informe publicado de que el ARNm de hOTCts es inestable. [Wang, L., *et al.*, *Molecular Genetics and Metabolism*, 105 (2012) 203-211].

30

Se han publicado informes sobre el uso de vectores víricos para intentar tratar la deficiencia de OTC. Por ejemplo, varios grupos han intentado esto en modelos murinos de deficiencia de OTC, utilizando adenovirus recombinantes portadores de ADNc de OTC de rata, ratón o ser humano. Por ejemplo, Wang, L. *et al.* (*Mol Genet Metab.* febrero de 2012; 105 (2): 203-211) muestra que una evaluación preclínica de un vector AAV8 candidato clínico para la deficiencia de ornitina transcarbamilasa (OTC) revela una enzima funcional de cada genoma de vector persistente. Wang, L. *et al.* (*Molecular Therapy*, vol. 20, n.º supl. 1, S56) proporciona una terapia génica neonatal para la deficiencia de OTC. Se han informado de cierta cantidad de reconstitución exitosa de la actividad OTC hepática y corrección de alteraciones metabólicas en modelos animales con virus que portan ADNc de OTC de rata o ratón. Estudios previos utilizando vectores adenovíricos han ilustrado las dificultades de expresar niveles suficientes de OTC humana activa en ratones con OTCD.

35

40

45

Por lo tanto, existe la necesidad de otros enfoques para la terapia de OTCD.

Sumario de la invención

En un aspecto, la invención proporciona un vector vírico recombinante que comprende un casete de expresión que comprende una secuencia de ácido nucleico que codifica la ornitina transcarbamilasa humana (hOTC) y secuencias de control de la expresión que dirigen la expresión de hOTC en una célula hepática, en donde la secuencia de ácido nucleico de hOTC es menos del 80 % idéntica a la secuencia de hOTC de tipo silvestre sobre la secuencia madura o hOTC de longitud completa de la SEQ ID NO: 1 y expresa una hOTC funcional, en donde dicha secuencia de ácido nucleico de hOTC se selecciona de la secuencia de ácido nucleico que comprende la SEQ ID NO: 5 o una secuencia de ácido nucleico al menos de aproximadamente 96 a aproximadamente 99 % idéntica a la misma.

50

55

En un aspecto adicional, la invención proporciona un virus adenoasociado recombinante (rAAV) que tiene una cápside de AAV y empaquetado en el mismo casete de expresión que comprende al menos una secuencia de repetición terminal invertida (ITR) de AAV, una secuencia de ácido nucleico modificada por ingeniería genética que codifica al menos la ornitina transcarbamilasa humana madura (hOTC) y secuencias de control de la expresión que dirigen la expresión de la hOTC en una célula hepática, comprendiendo dichas secuencias de control de la expresión un promotor específico del hígado, en donde la secuencia de ácido nucleico de hOTC es menos del 80 % idéntica a la secuencia de hOTC de tipo silvestre sobre al menos la hOTC madura de la SEQ ID NO: 1 y comprende al menos la hOTC madura de la SEQ ID NO: 5 o una secuencia de ácido nucleico al menos de aproximadamente 96 a aproximadamente 99,9 % idéntica a la misma.

60

65

En un aspecto, la divulgación proporciona un vector vírico recombinante que tiene un casete de expresión que comprende una secuencia de ácido nucleico modificada por ingeniería genética que codifica la ornitina transcarbamilasa humana (hOTCasa) y secuencias de control de la expresión que dirigen la expresión de hOTC en una célula hepática, en donde la secuencia de ácido nucleico de hOTC es menos del 80 % idéntica a la secuencia de hOTC de tipo silvestre sobre la hOTC de longitud completa de la secuencia de tipo silvestre (p. ej., SEQ ID NO: 1) o un fragmento de la misma que comprende la hOTC madura pero que carece de al menos la secuencia líder nativa, u otro intermedio que comprende al menos la hOTC madura) y expresa una hOTCasa funcional. Convenientemente, se han optimizado los codones de la secuencia modificada por ingeniería genética y esta se ha mejorado adicionalmente de modo que mejore al menos una de transducción, transcripción y/o traducción de la enzima.

La secuencia de ácido nucleico puede comprender la hOTC madura de la SEQ ID NO: 5, o una secuencia de ácido nucleico al menos de aproximadamente 96 a aproximadamente 99 % idéntica a la misma, o una secuencia de ácido nucleico que comprende al menos la hOTC madura de la SEQ ID NO: 9, o una secuencia de ácido nucleico al menos de aproximadamente 96 a aproximadamente 99 % idéntica a la misma, que expresa una hOTC funcional. En una divulgación, la hOTC es la longitud completa de la SEQ ID NO: 5 o una secuencia de ácido nucleico al menos de aproximadamente 96 a aproximadamente 99 % idéntica a la misma o una secuencia de ácido nucleico de la longitud completa de la SEQ ID NO: 9, o una secuencia de ácido nucleico al menos de aproximadamente 96 a aproximadamente 99 % idéntica a la misma. La secuencia de hOTC puede ser la de los nucleótidos correspondientes de las SEQ ID NO: 3, 4, 8 o 9. Se incluyen dentro del alcance de la invención las cadenas complementarias a las del listado de secuencias. El vector vírico puede seleccionarse de un vector de virus adenoasociado (AAV), un vector adenovírico y un vector lentivírico.

En un aspecto adicional, la divulgación proporciona un virus adenoasociado recombinante (rAAV) que tiene una cápside de AAV y empaquetado en el mismo un casete de expresión que comprende al menos una secuencia de repetición terminal invertida (ITR) de AAV, una secuencia de ácido nucleico modificada por ingeniería genética que codifica la ornitina transcarbamilasa humana (hOTCasa) y secuencias de control de la expresión que dirigen la expresión de la hOTC en una célula hepática, comprendiendo dichas secuencias de control de la expresión un promotor específico del hígado. La secuencia de ácido nucleico de hOTC modificada por ingeniería genética es menos del 80 % idéntica a la secuencia de hOTC de tipo silvestre sobre la secuencia madura o hOTC de longitud completa de la secuencia de tipo silvestre (p. ej., SEQ ID NO: 1) y expresa una hOTCasa funcional. La secuencia de ácido nucleico de hOTC sintética comprende al menos la hOTC madura de la SEQ ID NO: 5 o una secuencia de ácido nucleico al menos de aproximadamente 96 a aproximadamente 99,9 % idéntica a la misma o una secuencia de ácido nucleico que comprende al menos la hOTC madura de la SEQ ID NO: 9 o una secuencia de ácido nucleico al menos de aproximadamente 96 a aproximadamente 99,9 % idéntica a la misma.

En otro aspecto adicional más, la invención proporciona un vector vírico que comprende al menos un gen de hOTC que codifica una ornitina transcarbamilasa quimérica que comprende al menos ornitina transcarbamilasa humana madura con un péptido de tránsito heterólogo, en donde la secuencia codificante de la ornitina transcarbamilasa humana madura es la secuencia de ácido nucleico de las SEQ ID NO: 3, 4 o 5. Opcionalmente, pueden seleccionarse las secuencias de longitud completa de cualquiera de estas secuencias, que incluyen la secuencia de tránsito. Como alternativa, pueden seleccionarse un gen de OTC quimérico que incluye una secuencia de tránsito heteróloga, como se describe en el presente documento, y estas hOTC maduras.

En otro aspecto, una composición farmacéutica que comprende un vehículo y una cantidad eficaz de un vector, se proporciona el rAAV y/o el vector vírico de la invención.

Otro aspecto más es un vector vírico como se describe en el presente documento usado en la preparación de un medicamento para administrar ornitina transcarbamilasa a un sujeto que lo necesite y/o para tratar la deficiencia de ornitina transcarbamilasa. En una realización particularmente deseable, el sujeto es un sujeto humano. El sujeto puede ser homocigoto o heterocigoto para la deficiencia de ornitina transcarbamilasa.

En otro aspecto más, se proporciona el uso de un vector vírico que comprende una secuencia de ácido nucleico que codifica la ornitina transcarbamilasa humana funcional para prevenir y/o tratar la fibrosis o la cirrosis relacionada con la deficiencia de ornitina transcarbamilasa (OTCD) en un sujeto con OTCD. En una realización, el sujeto es un paciente humano. En una realización adicional, el sujeto es heterocigoto y puede presentar síntomas de inicio tardío.

En un aspecto adicional más, se proporciona el uso de un vector vírico que comprende una secuencia de ácido nucleico que codifica la ornitina transcarbamilasa humana funcional para prevenir y/o tratar el carcinoma hepatocelular en un sujeto que tiene OTCD. En una realización, el sujeto es un paciente humano. En una realización adicional, el sujeto es heterocigoto para OTCD y puede presentar síntomas de inicio tardío.

Otros aspectos y ventajas de la invención resultarán fácilmente evidentes a partir de la siguiente descripción detallada de la invención.

Breve descripción de los dibujos

La fig. 1A proporciona un ADNc de hOTC de tipo silvestre, que tiene 324 A, 223 C, 246 G y 269 T [SEQ ID NO: 1].

La fig. 1B - 1C proporciona la secuencia de ornitina transcarbamilasa humana codificada por la secuencia de la fig. 1A [SEQ ID NO: 2].

La fig. 2 proporciona un ADNc de hOTC modificado por ingeniería genética, con una proporción de CG alterada. El recuento de bases en la secuencia es de 283 A, 285 C, 284 G y 216 T [SEQ ID NO: 3].

La fig. 3 proporciona un ADNc de hOTC modificado por ingeniería genética denominado LW3 que comprende la secuencia de ácido nucleico expuesta en la SEQ ID NO: 4 que codifica una proteína hOTC. El recuento de bases en esta secuencia es de 279 A, 303 C, 288 G y 220 T. El codón de inicio para el marco de lectura abierto (ORF) de hOTC está precedido por una secuencia de Kozak en esta figura. La secuencia codificante del líder comienza en el nucleótido 15 (primeros 96 nucleótidos), seguido de la secuencia codificante de la hOTCasa de 322 aminoácidos. En esta figura, el codón de terminación va seguido de un sitio de restricción NotI (GCGGCCGC) que es un remanente del vector.

La fig. 4 proporciona un ADNc de hOTC modificado por ingeniería genética denominado LW4 que comprende la secuencia de ácido nucleico expuesta en la SEQ ID NO: 5 que codifica una proteína hOTC. El recuento de bases en esta secuencia es de 278 A, 303 C, 289 G y 220 T. La secuencia codificante del líder comienza en el nucleótido 15 (primeros 96 nucleótidos), seguido de la secuencia codificante de la hOTCasa de 322 aminoácidos. En esta figura, el codón de terminación va seguido de un sitio de restricción NotI (GCGGCCGC) que es un remanente del vector.

Las fig. 5A - 5C proporcionan una alineación de las secuencias de ADNc de la hOTC de tipo silvestre y cinco secuencias modificadas por ingeniería genética: GS (que comprende la secuencia de ácido nucleico expuesta en la SEQ ID NO: 3), LW3 (que comprende la secuencia de ácido nucleico expuesta en la SEQ ID NO: 4), LW4 (que comprende la secuencia de ácido nucleico expuesta en la SEQ ID NO: 5), LW5 (que comprende la secuencia de ácido nucleico expuesta en la SEQ ID NO: 8) y LW6 (que comprende la secuencia de ácido nucleico expuesta en la SEQ ID NO: 9), cada una de las cuales codifica una proteína hOTC. Las secuencias alineadas contienen una secuencia Kozak (primeros 14 nucleótidos de LW3 y LW4) y un sitio de enzima de restricción (después del codón de terminación para LW3 y LW4), que no forman parte del marco abierto de lectura.

Descripción detallada de la invención

En el presente documento se proporciona un ADNc de ornitina transcarbamilasa (OTC) humana (h) modificado por ingeniería genética, que se diseñó para maximizar la traducción y mejorar la estabilidad del ARNm en comparación con el ADN y/o ARNm de hOTC de tipo silvestre. También se proporcionan en el presente documento secuencias de ARNm de hOTC modificadas por ingeniería genética. Estas composiciones pueden utilizarse en métodos terapéuticos y/o profilácticos como se describe en el presente documento. Opcionalmente, estas composiciones se utilizan en combinación con otras terapias acordes con el tratamiento de referencia para las afecciones para las que se ha diagnosticado al sujeto (p. ej., un sujeto humano).

Con fines de comparación, se ilustra una secuencia de ADNc de OTC humana de tipo silvestre en la fig. 1A. Esta secuencia codifica la ornitina transcarbamilasa humana de la secuencia de aminoácidos de las fig. 1A - 1C. Esta misma secuencia de aminoácidos está codificada por los genes hOTC modificados por ingeniería genética de las fig. 2A - fig. 5. La enzima hOTC, que puede denominarse hOTCasa para distinguirla del gen, se expresa a partir de esta secuencia en forma de una preproteína que tiene un péptido líder de 32 aminoácidos en su extremo N (codificado por nt 1-96 de la fig. 1, aproximadamente los aminoácidos 1 a aproximadamente 32 de la SEQ ID NO: 2) que se escinde después de dirigir la enzima a la mitocondria celular, dejando la proteína "madura" de 322 restos de aminoácidos (de aproximadamente el aminoácido 33 a aproximadamente el aminoácido 354 de la SEQ ID NO: 2. Esta hOTCasa "denominada madura" es una proteína homotrimérica con una secuencia de 322 restos de aminoácidos en cada cadena polipeptídica. Opcionalmente, como alternativa a la secuencia de tipo silvestre de la SEQ ID NO: 2, se puede seleccionar una secuencia que incluya una o más de las posiciones polimórficas de origen natural que no están implicadas en la enfermedad: F101, L111, W1193-194 de la SEQ ID NO: 2 (véase, p. ej., <http://www.uniprot.org/uniprot/P00480>).

Aunque todas las secuencias de ADNc modificadas por ingeniería genética son de aproximadamente 77 % a aproximadamente 78 % idénticas a la secuencia de ácido nucleico de hOTC ts de la fig. 1A [SEQ ID NO: 1], existen diferencias estructurales entre estas secuencias (véase alineación en la fig. 5 que las ilustra). En particular, existe aproximadamente 4 % de diferencia en las secuencias de ácido nucleico entre hOTCco de la fig. 2 [SEQ ID NO: 3] y la hOTCcoLW4 de la fig. 4 [SEQ ID NO: 5]. Existe una diferencia de solo un nt entre LW-3 [fig. 3, SEQ ID NO: 4] y LW-4 [fig. 4, SEQ ID NO: 5], es decir, 0,094 % (1/1062) de diferencia (una A en LW-3 se cambia a una G en LW-4 como se muestra en la fig. 5).

En una realización, se proporciona una secuencia codificante de hOTC modificada cuya secuencia tiene menos de aproximadamente 80 % de identidad, preferentemente aproximadamente 77 % de identidad o menos con la secuencia

codificante de hOTC de tipo silvestre de longitud completa (fig. 1A, SEQ ID NO: 1), que codifica hOTC funcional. En una realización, la secuencia codificante de hOTC modificada se caracteriza por una mejor estabilidad en comparación con hOTC ts después de la administración mediada por AAV (p. ej., rAAV). Adicionalmente o como alternativa, se proporciona una secuencia codificante de hOTC modificada que carece de marcos de lectura alternativos para proteínas de al menos aproximadamente 9 aminoácidos de longitud. Adicionalmente, o como alternativa, se proporciona una secuencia codificante de hOTC modificada que tiene niveles de expresión de hOTCasa al menos aproximadamente 25 veces, al menos aproximadamente 50 veces o al menos aproximadamente 100 veces cuando se miden después de la expresión de un vector vírico, en comparación con el tipo silvestre de hOTCasa. Adicionalmente, o como alternativa, se proporciona una secuencia codificante de hOTC modificada que tiene actividad hepática de hOTCasa que es al menos aproximadamente 10 veces mayor, al menos aproximadamente 20 veces mayor o al menos aproximadamente 30 veces mayor en comparación con el tipo silvestre de hOTCasa expresada a partir de un vector vírico.

En una realización, una secuencia codificante de hOTC modificada es de 96 % a 99,9 % idéntica a la secuencia que codifica la enzima madura (de aproximadamente nt 99 a aproximadamente 1068) o la longitud completa de la fig. 4 (hOTCco-LW4, SEQ ID NO: 5) o de 96,5 % a 99 % idéntica, o aproximadamente 97 % o aproximadamente 98 % idéntica a la SEQ ID NO: 5 (fig. 4).

En una realización, una secuencia codificante de hOTC modificada es de 96 % a 99,9 % idéntica a la secuencia que codifica la enzima madura (de aproximadamente nt 99 a aproximadamente 1068) de la fig. 3 (hOTCco-LW3, SEQ ID NO: 4) o de 96,5 % a 99 % idéntica, o aproximadamente 97 % o aproximadamente 98 % idéntica a la SEQ ID NO: 4 (fig. 3).

En otra realización, una secuencia codificante de hOTC modificada es de 96 % a 99,9 % idéntica a la secuencia que codifica la enzima madura (de aproximadamente nt 99 a aproximadamente 1068) o la longitud completa de la fig. 2 (hOTCco, SEQ ID NO: 3), o de 96,5 % a 99 % idéntica, o aproximadamente 97 % o 98 % idéntico a la SEQ ID NO: 3 (fig. 2).

En otra divulgación más, una secuencia codificante de hOTC modificada tiene la secuencia que codifica la proteína madura (de aproximadamente nt 99 a aproximadamente 1068) o la longitud completa de hOTCco-LW5 [SEQ ID NO: 8] o hOTCco-LW6 [SEQ ID NO: 9], o una secuencia de 96 % a 99,9 % idéntica a la misma. hOTCco-LW5 y hOTCco-LW6 son aproximadamente 97 % idénticas entre sí y cada una es aproximadamente 78 % idéntica a la secuencia de tipo silvestre [SEQ ID NO: 1].

Las secuencias de las fig. 2-5 se proporcionan como la hebra con sentido de las secuencias de ADNc. La presente divulgación también abarca las hebras antisentido correspondientes a estas secuencias de ADNc y las secuencias de ARN, p. ej., ARNm, correspondientes. Por ejemplo, el ARNm modificado por ingeniería genética de la SEQ ID NO: 10, corresponde al ADN de la SEQ ID NO: 4; el ARN modificado por ingeniería genética de la SEQ ID NO: 11, corresponde al ADN de la SEQ ID NO: 5; el ARN modificado por ingeniería genética de la SEQ ID NO: 12, corresponde al ADN de la SEQ ID NO: 8; y el ARN de la SEQ ID NO: 13 corresponde al ADN de la SEQ ID NO: 9. Estas secuencias de ARN, y secuencias que son de 95 % a 99 %, o aproximadamente 97 % o aproximadamente 98 % idénticas a una o más de estas secuencias, están abarcadas dentro del alcance de la presente divulgación. Se conocen en la técnica métodos para alinear y determinar la identidad del ARN y estos incluyen bases de datos y servicios publicados y disponibles públicamente basados en la web o disponibles comercialmente. Véase, p. ej., LocARNA, CARNA, así como otros programas identificados en otras partes de los mismos.

En una realización de la divulgación, la secuencia de ARNm se puede administrar usando un sistema de administración de ARN seleccionado, ejemplos del cual se proporcionan en el presente documento.

También están abarcados en el presente documento fragmentos, p. ej., las secuencias que codifican el péptido de tránsito (aminoácidos 1 a aproximadamente 32), aproximadamente los aminoácidos 332 a aproximadamente 354, una enzima hOTC intermedia, o la enzima madura, u otros fragmentos según se desee. Puede hacerse referencia a la SEQ ID NO: 2. Véase, p. ej., Ye *et al.* 2001, Hum Gene Ther 12: 1035-1046.

En otra realización, se proporciona una OTC quimérica en el que la presecuencia N-terminal de OTC de tipo silvestre se reemplaza con una secuencia de tránsito de otra fuente que es compatible con el sistema del sujeto de manera que transporte eficazmente la hOTCasa madura codificada por el gen de OTC quimérico al orgánulo deseado. Véase, p. ej., Ye *et al.* 2001, Hum Gene Ther 12: 1035-1046. Dichas secuencias de tránsito codifican un péptido de tránsito (también denominado péptido señal, señal de direccionamiento o señal de localización) que se fusiona con la secuencia codificante de la hOTC madura de las SEQ ID NO: 1, 3, 4 y/o 5 9. Por ejemplo, la secuencia de tránsito de hOTC de tipo silvestre corresponde a aproximadamente los primeros 98 nucleótidos de la SEQ ID NO: 1. Para construir una OTC quimérica, estas secuencias N-terminales de tipo silvestre pueden eliminarse (de aproximadamente ácidos nucleicos 1 a aproximadamente nt 96 - nt 98) y reemplazarse con una secuencia de tránsito heteróloga. Los péptidos de tránsito adecuados son, preferentemente, aunque no necesariamente, de origen humano. Los péptidos de tránsito adecuados pueden elegirse de <http://proline.bic.nus.edu./spdb/zhang270.htm> o pueden determinarse usando diversos programas informáticos para determinar el péptido de tránsito en una proteína seleccionada. Aunque sin limitación,

dichas secuencias pueden tener de aproximadamente 15 a aproximadamente 50 aminoácidos de longitud o de aproximadamente 20 a aproximadamente 28 aminoácidos de longitud, o pueden ser mayores o menores según sea necesario.

5 La expresión "porcentaje (%) de identidad", "identidad de secuencia", "porcentaje de identidad de secuencia" o "porcentaje idéntico" en el contexto de las secuencias de ácido nucleico se refiere a los restos en las dos secuencias que son iguales cuando se alinean por correspondencia. La longitud de la comparación de identidad de secuencias puede abarcar la longitud completa del genoma, la longitud completa de una secuencia que codifica un gen o se desea un fragmento de al menos aproximadamente 500 a 5000 nucleótidos. Sin embargo, también se puede desear identidad
10 entre fragmentos más pequeños, p. ej., de al menos aproximadamente nueve nucleótidos, habitualmente al menos de aproximadamente 20 a 24 nucleótidos, al menos de aproximadamente 28 a 32 nucleótidos, al menos aproximadamente 36 o más nucleótidos.

15 El porcentaje de identidad se puede determinar fácilmente para las secuencias de aminoácidos sobre la longitud completa de una proteína, un polipéptido, aproximadamente 32 aminoácidos, aproximadamente 330 aminoácidos o un fragmento peptídico de los mismos o las correspondientes secuencias codificantes de secuencias de ácido nucleico. Un fragmento de aminoácido adecuado puede tener al menos aproximadamente 8 aminoácidos de longitud y puede tener hasta aproximadamente 700 aminoácidos. En general, al hacer referencia a "identidad", "homología" o "similitud" entre dos secuencias diferentes, la "identidad", "homología" o "similitud" se determina con referencia a
20 secuencias "alineadas". Las secuencias "alineadas" o "alineaciones" se refieren a múltiples secuencias de ácidos nucleicos o secuencias de proteínas (aminoácidos), que contienen con frecuencia correcciones para bases o aminoácidos ausentes o adicionales en comparación con una secuencia de referencia.

25 Las alineaciones se realizan utilizando cualquiera de diversos programas de alineación de múltiples secuencias disponibles pública o comercialmente. Están disponibles programas de alineación de secuencias para secuencias de aminoácidos, p. ej., los programas "Clustal X", "MAP", "PIMA", "MSA", "BLOCKMAKER", "MEME" y "Match-Box". En general, cualquiera de estos programas se utiliza con la configuración predeterminada, aunque un experto en la técnica puede modificar esta configuración según sea necesario. Como alternativa, un experto en la materia puede utilizar otro algoritmo o programa informático que proporcione al menos el nivel de identidad o alineación que proporcionan
30 los algoritmos y programas a los que se ha hecho referencia. Véase, p. ej., J. D. Thomson *et al.*, Nucl. Acids. Res., "A comprehensive comparison of multiple sequence alignments", 27 (13): 2682-2690 (1999).

También están disponibles múltiples programas de alineación de secuencias para secuencias de ácidos nucleicos. Los ejemplos de dichos programas incluyen, "Clustal W", "CAP Sequence Assembly", "BLAST", "MAP" y "MEME", a
35 los que se puede acceder a través de servidores web en Internet. Los expertos en la técnica conocen otras fuentes para dichos programas. Como alternativa, también se utilizan las utilidades Vector NTI. También hay varios algoritmos conocidos en la técnica que pueden utilizarse para medir la identidad de secuencias de nucleótidos, incluyendo los contenidos en los programas descritos anteriormente. Como ejemplo adicional, se pueden comparar secuencias de polinucleótidos utilizando Fasta™, un programa en GCG Versión 6.1. Fasta™ proporciona alineaciones y un
40 porcentaje de identidad de secuencia de las regiones de mejor solapamiento entre las secuencias de consulta y búsqueda. Por ejemplo, el porcentaje de identidad de secuencia entre secuencias de ácido nucleico se puede determinar utilizando Fasta™ con sus parámetros predeterminados (un tamaño de palabra de 6 y el factor NOPAM para la matriz de puntuación) como se proporciona en GCG Versión 6.1.

45 En una realización, los genes de hOTC modificados descritos en el presente documento se modifican por ingeniería genética en un elemento genético adecuado (vector) útil para generar vectores víricos y para su administración a una célula hospedadora, que transfiere las secuencias de hOTC portadas en el mismo. El vector seleccionado puede administrarse mediante cualquier método adecuado, incluyendo transfección, electroporación, suministro en liposomas, técnicas de fusión de membranas, microgránulos recubiertos con ADN de alta velocidad, infección vírica o
50 fusión de protoplastos. Los métodos utilizados para crear dichas construcciones son conocidos por los expertos en la manipulación de ácido nucleico e incluyen ingeniería genética, ingeniería recombinante y técnicas sintéticas. Véase, p. ej., Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY.

55 Como se usa en el presente documento, un "casete de expresión" se refiere a una molécula de ácido nucleico que comprende las secuencias de hOTC, promotor y puede incluir otras secuencias reguladoras para el mismo, pudiendo dicho casete empaquetarse en la cápside de un vector vírico (p. ej., una partícula vírica). Normalmente, dicho casete de expresión para generar un vector vírico contiene las secuencias de hOTC descritas en el presente documento flanqueadas por señales de empaquetamiento del genoma vírico y otras secuencias de control de la expresión tales como las descritas en el presente documento. Por ejemplo, para un vector vírico de AAV, las señales de
60 empaquetamiento son la repetición terminal invertida 5' (ITR) y la ITR 3'.

65 En una realización, las secuencias de ITR de AAV2, o la versión suprimida de las mismas (Δ ITR), se utilizan por conveniencia y para acelerar la aprobación reguladora. Sin embargo, pueden seleccionarse ITR de otras fuentes de AAV adecuadas. Cuando la fuente de las ITR sea de AAV2 y la cápside de AAV sea de otra fuente de AAV, el vector resultante puede denominarse pseudotipificado.

Normalmente, un casete de expresión para un vector de AAV comprende una ITR 5' de AAV, las secuencias codificantes de hOTC y cualquier secuencia reguladora, y una ITR 3' de AAV. Sin embargo, pueden ser adecuadas otras configuraciones de estos elementos. Se ha descrito una versión abreviada de la ITR 5', denominada Δ ITR, en la que se eliminan la secuencia D y el sitio de resolución terminal (trs). En otras realizaciones, se utilizan las ITR 5' y 3' de AAV de longitud completa.

En una realización, la construcción es una molécula de ADN (p. ej., un plásmido) útil para generar vectores víricos. Un plásmido ilustrativo que contiene elementos vectoriales deseables se ilustra con pAAVsc.TBG.hOTCco-LW4, cuya secuencia es la SEQ ID NO: 6. Este plásmido ilustrativo contiene un casete de expresión que comprende: scITR (nt 5 - 109 de la SEQ ID NO: 6), una señal TATA (nt 851-854 de la SEQ ID NO: 6), una secuencia codificante sintética de hOTC (nt 976-2037 de la SEQ ID NO: 6), un poli A (nt 2182-2046 en el complemento de la SEQ ID NO: 6), una scITR (nt 2378-2211 en el complemento de la SEQ ID NO: 6) y un promotor específico de hígado (TBG) (nt 4172-4760) de la SEQ ID NO: 6). Se pueden generar otros casetes de expresión utilizando las otras secuencias codificantes sintéticas de hOTC y otros elementos de control de la expresión, descritos en el presente documento.

La abreviatura "sc" en este contexto se refiere a autocomplementario. "AAV autocomplementario" se refiere a una construcción en la que se ha diseñado una región codificante portada por una secuencia de ácido nucleico de AAV recombinante para formar un molde de ADN bicatenario intramolecular. Tras la infección, en lugar de esperar a la síntesis mediada por células de la segunda hebra, las dos mitades complementarias de scAAV se asociarán para formar una unidad de ADN bicatenario (ADNbc) que está lista para la replicación y transcripción inmediatas. Véase, p. ej., D M McCarty *et al.*, "Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis", *Gene Therapy*, (agosto de 2001), vol 8, número 16, páginas 1248-1254. Se describen AAV autocomplementarios en, p. ej., las patentes de los Estados Unidos n.º 6.596.535; 7.125.717; y 7.456.683.

El casete de expresión normalmente contiene una secuencia promotora como parte de las secuencias de control de la expresión, p. ej., ubicada entre la secuencia de ITR 5' seleccionada y la secuencia codificante de hOTC. El plásmido y el vector ilustrativos descritos en el presente documento usan la globulina de unión a tiroxina (TBG) promotora específica del hígado. Como alternativa, se pueden usar otros promotores específicos del hígado [véase, p. ej., la base de datos de promotores de genes específicos del hígado, Cold Spring Harbor, <http://rulai.schl.edu/LSPD/>, tales como, p. ej., alfa 1 anti-tripsina (A1AT); albúmina humana Miyatake *et al.*, *J. Virol.*, 71: 5124-32 (1997), humAlb; y promotor central del virus de la hepatitis B, Sandig *et al.*, *Gene Ther.*, 3: 1002-9 (1996)]. Potenciador/promotor mínimo de TTR, promotor de alfa-antitripsina, LSP (845 nt) 25 (requiere scAAV sin intrones); o LSP1. Aunque son menos deseables, otros promotores, tales como promotores constitutivos, promotores regulables [véase, p. ej., documentos WO 2011/126808 y WO 2013/04943] o se puede utilizar un promotor sensible a señales fisiológicas en los vectores descritos en el presente documento.

Además de un promotor, un casete de expresión y/o un vector puede contener uno o más de otras secuencias adecuadas de inicio, terminación, potenciadoras de la transcripción, señales de procesamiento del ARN eficaces tales como señales de corte y empalme y de poliadenilación (poliA); secuencias que estabilizan el ARNm citoplásmico; secuencias que potencian la eficacia de traducción (es decir, secuencia consenso de Kozak); secuencias que mejoran la estabilidad de las proteínas; y, cuando se desee, secuencias que mejoran la secreción del producto codificado. Los ejemplos de secuencias de poliA adecuadas incluyen, p. ej., SV40, SV50, hormona del crecimiento bovina (bGH), hormona del crecimiento humana y poliA sintéticas. Los ejemplos de potenciadores adecuados incluyen, p. ej., el potenciador de alfa fetoproteína, el promotor/potenciador mínimo de TTR, LSP (promotor de globulina de unión a TH/alfa1-microglobulina/potenciador de bikunina), entre otros. En una realización, el casete de expresión comprende uno o más potenciadores de la expresión. En una realización, el casete de expresión contiene dos o más potenciadores de la expresión. Estos potenciadores pueden ser iguales o pueden diferir entre sí. Por ejemplo, un potenciador puede incluir un potenciador Alfa mic/bik. Este potenciador puede estar presente en dos copias que están ubicadas adyacentes entre sí. Como alternativa, las copias dobles del potenciador pueden estar separadas por una o más secuencias. En otra realización más, el casete de expresión contiene además un intrón, p. ej., el intrón Promega. Otros intrones adecuados incluyen los conocidos en la técnica, p. ej., tales como se describen en el documento WO 2011/126808. Opcionalmente, pueden seleccionarse una o más secuencias para estabilizar el ARNm. Un ejemplo de dicha secuencia es una secuencia de WPRE modificada, que puede modificarse por ingeniería genética cadena arriba de la secuencia de poliA y cadena abajo de la secuencia codificante [véase, p. ej., MA Zanta-Boussif, *et al.*, *Gene Therapy* (2009) 16: 605-619].

Estas secuencias de control están "unidas operativamente" a las secuencias de genes de hOTC. Como se usa en el presente documento, la expresión "unidas operativamente" se refiere tanto a las secuencias de control de la expresión que son contiguas con el gen de interés y las secuencias de control de la expresión que actúan en *trans* o a distancia para controlar el gen de interés.

Los vectores víricos de AAV recombinantes son muy adecuados para el suministro de las secuencias de expresión de hOTC descritas en el presente documento. Dichos vectores de AAV pueden contener ITR que proceden de la misma fuente de AAV que la cápside. Como alternativa, los ITR de AAV pueden ser de una fuente de AAV diferente a la que suministra la cápside.

5 Cuando se va a producir AAV pseudotipificado, las ITR en la expresión se seleccionan de una fuente que difiere de la fuente de AAV de la cápside. Por ejemplo, se pueden seleccionar ITR de AAV2 para su uso con una cápside de AAV que tenga una eficacia particular para dirigirse al hígado (p. ej., hepatocitos). Las cápsides de AAV pueden seleccionarse de AAV8 [patente de los Estados Unidos 7790449; patente de los Estados Unidos 7282199] y rh10 [documento WO 2003/042397] para las composiciones descritas en el presente documento. Sin embargo, otros AAV, incluyendo, p. ej., AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV9 y otros, tales como, p. ej., los descritos en los documentos WO 2003/042397; WO 2005/033321, WO 2006/110689; US 7588772 B2,] se pueden utilizar en sujetos humanos.

10 En una realización, se proporciona un AAV autocomplementario. Este vector vírico puede contener una ITR Δ5' y una ITR 3' de AAV. En un ejemplo, el vector vírico es scAAV2/8.TBG.hOTCco. En otro ejemplo, el vector vírico es scAAV2/rh10.TBG.hOTCco. Ambos de estos vectores contienen la ΔITR 5' de AAV2, el promotor de TBG específico del hígado, una secuencia codificante de hOTCco modificada por ingeniería genética de la invención, una poliA de SV40 y la ITR 3' de AAV2 en una cápside de AAV8 [véase, p. ej., patente de los Estados Unidos n.º 8.318480B2] o cápside de AAV rh10. La secuencia puede seleccionarse de hOTC modificada por ingeniería genética de una de las SEQ ID NO: 3, 4, 5, 8 o 9. Opcionalmente, la secuencia de tránsito de la hOTC modificada por ingeniería genética puede sustituirse por una secuencia de tránsito heteróloga para proporcionar una hOTC quimérica, que conserva la hOTCasa madura.

15 En otra realización, se proporciona un vector vírico de AAV monocatenario. Dicho vector puede contener una ITR 5' de AAV y una ITR 3'. Un ejemplo es AAV2/8.TBG.hOTCco, que contiene la ITR 5' de AAV2 de longitud completa, el promotor de TBG específico del hígado, la secuencia codificante de hOTC, una poliA de la hormona del crecimiento bovina y ITR 3' de AAV2. Otro ejemplo es AAV2/8.TBG.hOTCco-WPRE.bGH, que contiene los mismos elementos de vector y contiene además el elemento regulador postranscripcional del virus de la hepatitis de la marmota (WPRE). En otras realizaciones, WPRE está ausente de las construcciones que se van a utilizar *in vivo*. La secuencia de hOTC modificada por ingeniería genética (abreviada en el presente documento hOTCco) puede seleccionarse de hOTC modificada por ingeniería genética de una de las SEQ ID NO: 3, 4 o 5. Opcionalmente, la secuencia del péptido de tránsito de la hOTC modificada por ingeniería genética puede sustituirse por una secuencia de tránsito heteróloga para proporcionar una hOTC quimérica, que conserva la hOTCasa madura.

20 Pueden seleccionarse otros promotores más, incluyendo promotores específicos de tejido. Se conocen en la técnica métodos para generar y aislar vectores víricos de AAV adecuados para su administración a un sujeto. Véase, p. ej., solicitud de patente publicada de los Estados Unidos n.º 2007/0036760 (15 de febrero de 2007), patente de los Estados Unidos 7790449; patente de los Estados Unidos 7282199; documentos WO 2003/042397; WO 2005/033321, WO 2006/110689; y US 7588772 B2]. Las secuencias de AAV8 y métodos para generar vectores basados en la cápside de AAV8 se describen en la patente de los Estados Unidos 7.282.199 B2, US 7.790.449 y US 8.318.480. En un primer sistema, una línea celular productora se transfecta transitoriamente con una construcción que codifica el transgén flanqueado por ITR y una construcción o construcciones que codifican rep y cap. En un segundo sistema, una línea celular de empaquetamiento que proporciona de forma estable rep y cap se transfecta de forma transitoria con una construcción que codifica el transgén flanqueado por las ITR. En cada uno de estos sistemas, se producen viriones de AAV en respuesta a la infección por adenovirus auxiliar o herpesvirus, lo que requiere la separación de los rAAV del virus contaminante. Más recientemente, se han desarrollado sistemas que no requieren infección con virus auxiliares para recuperar el AAV, las funciones auxiliares necesarias (es decir, adenovirus E1, E2a, VA y E4 o herpesvirus UL5, UL8, UL52 y UL29 y polimerasa de herpesvirus) también son suministradas, en *trans*, por el sistema. En estos sistemas más nuevos, las funciones auxiliares se pueden suministrar mediante transfección transitoria de las células con construcciones que codifican las funciones auxiliares necesarias o las células se pueden modificar por ingeniería genética para contener de manera estable genes que codifican las funciones auxiliares, cuya expresión se puede controlar a nivel transcripcional o postranscripcional. En otro sistema más, el transgén flanqueado por ITR y los genes rep/cap se introducen en células de insecto mediante infección con vectores basados en baculovirus. Para revisiones de estos sistemas de producción, véase, en general, p. ej., Zhang *et al.*, 2009, "Adenovirus-adenovirus-associated virus hybrid for large-scale recombinant adeno-associated virus production", Human Gene Therapy 20: 922-929. También se describen métodos de fabricación y uso de estos y otros sistemas de producción de AAV en las siguientes patentes de los Estados Unidos: 5.139.941; 5.741.683; 6.057.152; 6.204.059; 6.268.213; 6.491.907; 6.660.514; 6.951.753; 7.094.604; 7.172.893; 7.201.898; 7.229.823; y 7.439.065.

25 El espacio disponible para el empaquetamiento se puede conservar combinando más de una unidad de transcripción en una sola construcción, reduciendo de este modo la cantidad de espacio de secuencia reguladora necesario. Por ejemplo, un solo promotor puede dirigir la expresión de un solo ARN que codifica dos o tres o más genes, y la traducción de los genes cadena abajo está conducida por secuencias de IRES. En otro ejemplo, un solo promotor puede dirigir la expresión de un ARN que contiene, en un solo marco abierto de lectura (ORF), dos o tres o más genes separados entre sí por secuencias que codifican un péptido de autoescisión (p. ej., T2A) o un sitio de reconocimiento de proteasa (p. ej., furina). Por tanto, el ORF codifica una sola poliproteína, que, ya sea durante (en el caso de T2A) o después de la traducción, se escinde en las proteínas individuales (tales como, p. ej., factor de transcripción transgénico y dimerizable). Debe observarse, sin embargo, que aunque estos sistemas de poliproteínas e IRES se pueden utilizar para ahorrar espacio de empaquetamiento de AAV, solo pueden utilizarse para la expresión de

componentes que pueden ser conducidos por el mismo promotor. En otra alternativa, la capacidad transgénica de AAV puede aumentarse proporcionando ITR de AAV de dos genomas que pueden hibridarse para formar concatámeros de cabeza a cola.

- 5 Opcionalmente, los genes de hOTC descritos en el presente documento pueden utilizarse para generar vectores víricos distintos de rAAV. Dichos otros vectores víricos pueden incluir el uso de cualquier virus adecuado para terapia génica, incluyendo, pero sin limitación, adenovirus; virus del herpes; lentivirus; retrovirus; etc. Convenientemente, cuando se genera uno de estos otros vectores, se produce como un vector vírico de replicación defectuosa.
- 10 Un "virus de replicación defectuosa" o "vector vírico" se refiere a una partícula vírica sintética o artificial en la que un casete de expresión que contiene un gen de interés está empaquetado en una cápside o envoltura vírica, donde cualquier secuencia genómica vírica empaquetada también dentro de la cápside o envoltura vírica es deficiente en replicación; es decir, no puede generar viriones descendientes pero conservan la capacidad de infectar células diana. En una realización, el genoma del vector vírico no incluye genes que codifican las enzimas necesarias para la replicación (el genoma puede modificarse para que sea "dependiente de auxiliar", que contiene solo el transgén de interés flanqueado por las señales necesarias para la amplificación y el empaquetamiento del genoma artificial), pero estos genes pueden suministrarse durante la producción. Por lo tanto, se considera seguro para su uso en terapia génica, ya que la replicación y la infección por viriones descendientes no pueden ocurrir excepto en presencia de la enzima vírica necesaria para la replicación. Dichos virus de replicación defectuosa pueden ser virus adenoasociados (AAV), adenovirus, lentivirus (integradores o no integradores) u otra fuente de virus adecuada.
- 15
- 20

Las composiciones farmacéuticas descritas en el presente documento están diseñadas para su administración a sujetos que lo necesiten mediante cualquier vía adecuada o una combinación de diferentes vías. Se desea la administración directa o intrahepática al hígado y, opcionalmente, se puede realizar mediante administración intravascular, p. ej., a través de la vena porta, vena hepática, conducto biliar o por trasplante. Como alternativa, se pueden seleccionar otras vías de administración (p. ej., oral, inhalación, intranasal, intratraqueal, intraarterial, intraocular, intravenosa, intramuscular y otras vías parentales). Las construcciones de administración de hOTC descritas en el presente documento pueden administrarse en una composición única o en múltiples composiciones. Opcionalmente, se pueden administrar dos o más AAV diferentes [véase, p. ej., documentos WO 2011/126808 y WO 25 2013/049493]. En otra realización, dichos múltiples virus pueden contener diferentes virus de replicación defectuosa (p. ej., AAV, adenovirus y/o lentivirus). Como alternativa, la administración puede estar mediada por construcciones no víricas, p. ej., "ADN desnudo", "ADN plasmídico desnudo", ARN y ARNm; junto con diversas composiciones de y nanopartículas administración, incluyendo, p. ej., micelas, liposomas, composiciones de lípido catiónico - ácido nucleico, composiciones de poliglucanos y otros polímeros, conjugados de ácidos nucleicos a base de lípidos y/o colesterol, y otras construcciones como las que se describen en el presente documento. Véase, p. ej., X. Su *et al.*, Mol. Pharmaceutics, 2011, 8 (3), págs. 774-787; publicación web: 21 de marzo de 2011; documentos WO2013/182683, WO 2010/053572 y WO 2012/170930. Dichas construcciones de administración de hOTC no víricas pueden administrarse mediante las vías descritas anteriormente.

25

30

35

40 Los vectores víricos, o restos de transferencia de ADN o ARN no víricos, se pueden formular con un vehículo fisiológicamente aceptable para su uso en aplicaciones de transferencia génica y terapia génica. En el caso de los vectores víricos de AAV, la cuantificación de las copias del genoma ("CG") puede usarse como la medida de la dosis contenida en la formulación. Puede usarse cualquier método conocido en la técnica para determinar el número de copias del genoma (CG) de las composiciones de virus de replicación defectuosa de la invención. Un método para realizar la valoración numérica de CG de AAV es el siguiente: Las muestras de vector de AAV purificadas se tratan primero con DNasa para eliminar el ADN del genoma de AAV no encapsidado o el ADN plasmídico contaminante del proceso de producción. Las partículas resistentes a la DNasa se someten después a un tratamiento térmico para liberar el genoma de la cápside. Después, los genomas liberados se cuantifican mediante PCR en tiempo real utilizando conjuntos de cebadores/sondas que se dirigen a una región específica del genoma vírico (habitualmente señal de poli A). Las composiciones de virus de replicación defectuosa se pueden formular en unidades de dosificación para contener una cantidad de virus de replicación defectuosa que se encuentra en el intervalo de aproximadamente $1,0 \times 10^9$ CG a aproximadamente $1,0 \times 10^{15}$ CG (para tratar a un sujeto promedio de 70 kg de peso corporal) y, preferentemente, de $1,0 \times 10^{12}$ CG a $1,0 \times 10^{14}$ CG para un paciente humano. Preferentemente, la dosis de virus de replicación defectuosa en la formulación es de $1,0 \times 10^9$ CG, $5,0 \times 10^9$ CG, $1,0 \times 10^{10}$ CG, $5,0 \times 10^{10}$ CG, $1,0 \times 10^{11}$ CG, $5,0 \times 10^{11}$ CG, $1,0 \times 10^{12}$ CG, $5,0 \times 10^{12}$ CG, o $1,0 \times 10^{13}$ CG, $5,0 \times 10^{13}$ CG, $1,0 \times 10^{14}$ CG, $5,0 \times 10^{14}$ CG, o $1,0 \times 10^{15}$ CG.

45

50

55

El ADN y el ARN se miden en general en cantidades de nanogramos (ng) a microgramos (μ g) de los ácidos nucleicos. En general, para un tratamiento en un ser humano, preferentemente se formulan y administran dosificaciones del ARN en el intervalo de 1 ng a 700 μ g, de 1 ng a 500 μ g, de 1 ng a 300 μ g, de 1 ng a 200 μ g o de 1 ng a 100 μ g. Se pueden utilizar cantidades de dosificación similares de una molécula de ADN que contiene un casete de expresión y que no se administra a un sujeto mediante un vector vírico para las construcciones de administración de ADN de hOTC no víricas.

60

65 La producción de lentivirus se mide tal como se describe en el presente documento y se expresa como UI por volumen (p. ej., ml). UI es unidad infecciosa o, como alternativa, unidades de transducción (UT); UI y UT se pueden usar

indistintamente como una medida cuantitativa del título de una preparación de partículas de vectores víricos. El vector lentivírico normalmente no es integrador. La cantidad de partículas víricas es al menos aproximadamente 3×10^6 UI, y puede ser al menos aproximadamente 1×10^7 UI, al menos aproximadamente 3×10^7 UI, al menos aproximadamente 1×10^8 UI, al menos aproximadamente 3×10^8 UI, al menos aproximadamente 1×10^9 UI o al menos aproximadamente 3×10^9 UI.

Los vectores recombinantes descritos anteriormente pueden administrarse a células hospedadoras según métodos publicados. El rAAV, preferentemente suspendido en un vehículo fisiológicamente compatible, puede administrarse a un paciente mamífero humano o no humano. Un experto en la materia puede seleccionar fácilmente vehículos adecuados a la vista de la indicación para la que se dirige el virus de transferencia. Por ejemplo, un vehículo adecuado incluye solución salina, que puede formularse con diversas soluciones tamponantes (p. ej., solución salina tamponada con fosfato). Otros vehículos ilustrativos incluyen solución salina estéril, lactosa, sacarosa, fosfato de calcio, gelatina, dextrano, agar, pectina, aceite de cacahuete, aceite de sésamo y agua. La selección del vehículo no es una limitación de la presente invención.

Opcionalmente, las composiciones de la invención pueden contener, además del rAAV y el vehículo o vehículos, otros ingredientes farmacéuticos convencionales, tales como conservantes o estabilizantes químicos. Los conservantes ilustrativos adecuados incluyen clorobutanol, sorbato de potasio, ácido sórbico, dióxido de azufre, galato de propilo, los parabenos, etil vainillina, glicerina, fenol y paraclorofenol. Los estabilizantes químicos adecuados incluyen gelatina y albúmina.

Los vectores víricos descritos en el presente documento pueden utilizarse para preparar un medicamento para administrar ornitina transcarbamilasa a un sujeto (p. ej., un paciente humano) que lo necesite, suministrar hOTCasa funcional a un sujeto y/o para tratar la deficiencia de ornitina transcarbamilasa. Un ciclo de tratamiento puede implicar opcionalmente la administración repetida del mismo vector vírico (p. ej., un vector de AAV8) o un vector vírico diferente (p. ej., un AAV8 y un AAVrh10). Pueden seleccionarse otras combinaciones más utilizando los vectores víricos y sistemas de administración no víricos descritos en el presente documento.

En otra divulgación, las secuencias de ácido nucleico descritas en el presente documento pueden administrarse mediante una vía no vírica. Por ejemplo, una secuencia de hOTC puede ser a través de un sistema transportador para expresión o administración en forma de ARN (p. ej., ARNm) utilizando uno de varios sistemas transportadores que son conocidos en la técnica. Dichos sistemas transportadores incluyen los proporcionados por entidades comerciales, tales como la tecnología denominada "SMARTT" de PhaseRx. Estos sistemas utilizan copolímeros en bloque para la administración a una célula huésped diana. Véase, p. ej., documento US 2011/0286957 titulado, "Multiblock Polymers", publicado el 24 de noviembre de 2011; documento US 2011/0281354, publicado el 17 de noviembre de 2011; documento EP2620161, publicado el 31 de julio de 2013; y documento WO 2015/017519, publicado el 5 de febrero de 2015. Véase, también, S. Uchida *et al.*, (febrero de 2013) PLoS ONE 8 (2): e56220. Otros métodos más implican generar e inyectar ARNbc sintéticos [véase Soutschek *et al.* Nature (2004) 432 (7014): 173-8; véase, también, Morrissey *et al.* Hepatol. (2005) 41 (6): 1349-56], también se ha demostrado la administración local en el hígado mediante la inyección de ARN bicatenario directamente en el sistema circulatorio que rodea el hígado mediante cateterismo de la vena renal. [Véase Hamar *et al.* PNAS (2004) 101 (41): 14883-8]. Otros sistemas más implican la administración de ARNbc y, en particular ARNip utilizando complejos catiónicos o formulaciones liposómicas [véase, p. ej., Landen *et al.* Cancer Biol. Ther. (2006) 5 (12); véase, también, Khoury *et al.* Arthritis Rheumatol. (2006) 54 (6): 1867-77. También están disponibles otras tecnologías de administración de ARN, p. ej., de Veritas Bio [véase, p. ej., documento US 2013/0323001, 23 de diciembre de 2010, "In vivo delivery of double stranded RNA to a target cell" (contenido citosólico incluyendo ARN, p. ej., ARNm, ARNip/ARNhp/miARN expresado, así como ARNip/ARNhp/miARN inyectado/introducido, o posiblemente incluso ADN transfectado presente en el citosol empaquetado dentro de las exovesículas y transportado a sitios distales tales como el hígado)]. Se han descrito otros sistemas más para administración *in vivo* de secuencias de ARN. Véase, p. ej., documento US 2012/0195917 (2 de agosto de 2012) (análogos de ARN de caperuza 5' para mejorar la estabilidad y aumentar la expresión de ARN), documento WO 2013/143555A1, 3 de octubre de 2013 y/o están disponibles en el mercado (BioNTech, Alemania; Valera (Cambridge, MA); Zata Pharmaceuticals).

Por tanto, en una realización, la divulgación proporciona un ARNm de hOTC modificado por ingeniería genética de la secuencia madura (al menos aproximadamente nt 99 - 1098) o la longitud completa de las SEQ ID NO: 10, 11, 12, 13, o una secuencia que tiene al menos de 97 % a 99 % de identidad con la misma, en una composición para la administración de ARN bicatenario o monocatenario que da lugar a la expresión de la hOTCasa madura en una célula hospedadora diana, p. ej., una célula hepática.

La cinética de la composición descrita en el presente documento que contiene ARNm (administrado directamente, en comparación con transcrito a partir de una molécula de administración de ADN) es en particular adecuada para su uso en sujetos en crisis aguda, ya que la expresión de la hOTCasa a partir del ARNm puede verse en un periodo de varias horas. Para evitar una rápida eliminación del ARN, se modifica como se describe en el presente documento (p. ej., utilizando una caperuza o una base modificada), de modo que sus efectos puedan conservarse durante más de 24 horas, durante 48 horas o hasta aproximadamente 3 días (aproximadamente 72 horas). Puede ser deseable coadministrar un ARNm directamente como se describe en el presente documento y coadministrar al mismo tiempo o

5 sustancialmente al mismo tiempo, una composición de hOTC basada en ADN o vector vírico como se define en el presente documento. Por tanto, un sujeto puede recibir tratamiento inmediato y en el momento en que la expresión mediada por ARNm comienza a disminuir, la expresión de hOTC a largo plazo conferida por una expresión mediada por un vector vírico comienza a surtir efecto. Como alternativa, un sujeto puede recibir una segunda administración de una composición basada en ARNm como se define en el presente documento. Las composiciones de ARNm descritas en el presente documento pueden usarse en otros regímenes o métodos terapéuticos, incluyendo los que implican a pacientes con OTCD que no están en crisis aguda.

10 Las composiciones según la presente invención pueden comprender un vehículo farmacéuticamente aceptable, tal como se ha definido anteriormente. En una realización, la composición comprende un copolímero en bloque asociado con un polinucleótido de hOTC como se describe en el presente documento. Los copolímeros en bloque pueden formar una micela, de manera que la micela comprenda una pluralidad de copolímeros en bloque.

15 Normalmente, dicha composición contiene una molécula de ácido nucleico que comprende la secuencia de ARNm correspondiente a la secuencia de hOTC que codifica la hOTCasa madura (al menos aproximadamente nt 99 a 1098). Además, esta molécula de ácido nucleico puede incluir la región 5' no traducida (UTR), también conocida como la secuencia líder o ARN líder, y uno o más de un intrón o intrones opcionales, un exón o exones opcionales, una secuencia opcional de Kozak, un WPRE opcional y una poliA, y la 3' UTR que flanquea las secuencias codificantes. Las secuencias líder adecuadas incluyen las analizadas anteriormente en relación con las secuencias de ADN de hOTC. Se han analizado anteriormente ejemplos de fuentes de secuencias líder adecuadas, distintas de las secuencias líder de hOTC nativas, o las correspondientes a fig. 2, fig. 3, fig. 4 o fig. 5. De manera similar, se han analizado anteriormente fuentes de intrones adecuados, poliA y secuencias de Kozak y son aplicables al suministro de las secuencias de ARN correspondientes analizadas en el presente párrafo. Además, se pueden generar diversas modificaciones al ARN, p. ej., se puede obtener por ingeniería genética una estructura de caperuza 5' modificada en la construcción para evitar la eliminación rápida del ARNm *in vivo*, o por otra razón deseada. Los expertos en la técnica conocen métodos para generar tales estructuras de caperuza 5'. Véase, p. ej., documentos US 2012/0195917 y WO 2013/143555A1, 3 de octubre de 2013. Además, se pueden utilizar nucleótidos modificados para producir ARNm *in vitro*, como pseudouridina. También se puede dosificar ARN repetidamente, o se puede dosificar al sujeto primero con ARNm para tratar crisis neonatales seguido de administración mediada por vectores víricos (p. ej., AAV) para terapia a largo plazo y para prevenir fibrosis/cirrosis y/o carcinoma hepatocelular.

35 El ARNm se puede sintetizar a partir de las secuencias de ADN de hOTC descritas en el presente documento, usando técnicas que se conocen bien en la materia. Por ejemplo, Cazenave C, Uhlenbeck OC, RNA template-directed RNA synthesis by T7 RNA polymerase. Proc Natl Acad Sci USA. 19 de julio de 1994; 91 (15): 6972-6, describen el uso de la ARN polimerasa de T7 para generar ARN a partir de moldes de ADNc o ARN. Véase también, Wichlacz A1, Legiewicz M, Ciesiolka J., Generating *in vitro* transcripts with homogenous 3' ends using trans-acting antigenomic delta ribozyme., Nucleic Acids Res. 18 de febrero de 2004; 32 (3): e39; Krieg PA, Melton DA., Functional messenger RNAs are produced by SP6 *in vitro* transcription of cloned cDNAs, Nucleic Acids Res. 25 de septiembre de 1984; 12 (18): 7057-70; y Rio, D. C., *et al.* RNA: A Laboratory Manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2011, 205-220. Además, los equipos y protocolos para generar ARNm están disponibles comercialmente que incluyen, sin limitación, el sistema de transcripción *in vitro* Riboprobe® (Promega Corp.); sistemas de producción de ARN a gran escala RiboMAX™ (Promega Corp.); equipo MAXIscript (Ambion); equipo MEGIscript (Ambion); equipo de ARNa MessageAmp™ (Ambion); equipos mMESSAGING mMACHINE® (Ambion); y equipo de síntesis de ARN de alto rendimiento HiScribe™ T7 (New England Biolabs® Inc.). El ARN personalizado también se puede generar comercialmente a partir de empresas incluyendo, sin limitación, TriLink Biotechnologies; bioSYNTHESIS; GE Dharmacon; e IBA Lifesciences.

50 Las secuencias de ADN de hOTC descritas en el presente documento se pueden generar *in vitro* y sintéticamente, usando técnicas muy conocidas en este campo. Por ejemplo, se puede utilizar el método de síntesis precisa basada en PCR (PAS) de secuencias largas de ADN, como se describe en Xiong *et al.*, PCR-based accurate synthesis of long DNA sequences, Nature Protocols 1, 791 - 797 (2006). Se describe un método que combina los métodos de PCR asimétrica doble y PCR de extensión solapante en Young y Dong, Two-step total gene synthesis method, Nucleic Acids Res. 2004; 32 (7): e59. Véase también, Gordeeva *et al.*, J Microbiol Methods. Improved PCR-based gene synthesis method and its application to the Citrobacter freundii phytase gene codon modification. mayo de 2010; 81 (2): 147-52. Publicación electrónica 10 de marzo de 2010; véase, también, las siguientes patentes sobre síntesis de oligonucleótidos y síntesis de genes, Gene Seq. abril de 2012; 6 (1): 10-21; US 8008005; y US 7985565. Además, los equipos y protocolos para generar ADN mediante PCR están disponibles en el mercado. Estos incluyen el uso de polimerasas incluyendo, sin limitación, Taq polimerasa; OneTaq® (New England Biolabs); ADN polimerasa de alta fidelidad Q5® (New England Biolabs); y polimerasa GoTaq® G2 (Promega). También se puede generar ADN a partir de células transfectadas con plásmidos que contienen las secuencias de hOTC descritas en el presente documento. Los kits y protocolos son conocidos y están disponibles en el mercado e incluyen, sin limitación, equipos de plásmidos QIAGEN; equipos de plásmidos de filtro Clargeswicl® Pro (Invitrogen); y equipos de plásmidos GenElute™ (Sigma Aldrich). Otras técnicas útiles en el presente documento incluyen métodos de amplificación isotérmica de secuencia específica, que eliminan la necesidad de termociclado. En lugar de calor, estos métodos emplean habitualmente una ADN polimerasa de desplazamiento de hebra, como ADN polimerasa Bst, fragmento grande (New England Biolabs), para separar ADN bicatenario. También se puede generar ADN a partir de moléculas de ARN por amplificación

mediante el uso de transcriptasas inversas (RT), que son ADN polimerasas dependientes de ARN. Las RT polimerizan una hebra de ADN que es complementaria del molde de ARN original y se denomina ADNc. Este ADNc puede después amplificarse adicionalmente mediante PCR o métodos isotérmicos como se ha descrito anteriormente. El ADN personalizado también se puede generar comercialmente a partir de empresas incluyendo, sin limitación, GenScript; 5 GENEWIZ®; GeneArt® (Life Technologies); e Integrated DNA Technologies.

El término "expresión" se usa en el presente documento en su significado más amplio y comprende la producción de ARN o de ARN y proteína. Con respecto al ARN, el término "expresión" o "traducción" se refiere en particular a la producción de péptidos o proteínas. La expresión puede ser transitoria o estable. 10

El término "traducción" en el contexto de la presente invención se refiere a un proceso en el ribosoma, en donde una cadena de ARNm controla el ensamblaje de una secuencia de aminoácidos para generar una proteína o un péptido.

Según la presente invención, se administra una "cantidad terapéuticamente eficaz" de la hOTC como se describe en el presente documento para lograr el resultado deseado, es decir, tratamiento de la deficiencia de OTC o uno o más síntomas de la misma. Como se describe en el presente documento, un resultado deseado incluye reducir los niveles de ácido orótico, reducir la hiperamonemia y/o minimizar o eliminar una o más de las complicaciones neurofísicas, incluyendo el retraso del desarrollo, dificultades de aprendizaje, discapacidad intelectual, trastorno por déficit de atención con hiperactividad y déficits de la función ejecutiva. El tratamiento puede incluir el tratamiento de sujetos que 15 tienen una enfermedad grave de inicio neonatal (hombres o, con menos frecuencia, mujeres) y enfermedad de inicio tardío (parcial) en hombres y mujeres, que puede presentarse desde la primera infancia hasta la niñez posterior, adolescencia o adultez. En determinadas realizaciones, la invención proporciona vectores víricos rAAV para su uso en un método de tratamiento de la deficiencia de ornitina transcarbamilasa (OCTD) en un paciente humano. Métodos de tratamiento y/o prevención de la fibrosis y/o cirrosis en sujetos, en particular, se describen en el presente documento los sujetos heterocigotos de aparición tardía mediante la administración de hOTC. En una realización, los objetivos 20 terapéuticos para la deficiencia de OTC son mantener el amoníaco en plasma a menos de <80 $\mu\text{mol/l}$, glutamina en plasma <1000 $\mu\text{mol/l}$, argininemia 80-150 $\mu\text{mol/l}$ y aminoácidos de cadena ramificada dentro del intervalo normal. Sin embargo, el médico tratante puede seleccionar otros criterios de valoración terapéuticos.

En otro aspecto más, la divulgación proporciona vectores víricos o rAAV para su uso en el rescate y/o el tratamiento de OTCD de un sujeto neonatal que comprende la etapa de administrar un gen de hOTC al hígado de un sujeto recién nacido (p. ej., un paciente humano). Este método puede utilizar cualquier secuencia de ácido nucleico que codifique una hOTC funcional, ya sea una hOTC sintética como se describe en el presente documento o una hOTC de tipo silvestre, o una hOTC de otra fuente, o una combinación de las mismas. En una realización, el tratamiento neonatal se define como la administración de una hOTC como se describe en el presente documento en las primeras 8 horas, las primeras 12 horas, las primeras 24 horas o las primeras 48 horas después del parto. En otra realización, en particular para un primate, el parto neonatal se realiza dentro del periodo de aproximadamente 12 horas a aproximadamente 1 semana, 2 semanas, 3 semanas o aproximadamente 1 mes, o después de aproximadamente 24 horas a aproximadamente 48 horas. Para abordar la dilución debido a la rápida renovación de las células hepáticas en un mamífero en crecimiento (p. ej., un primate no humano o humano), es deseable que la terapia neonatal vaya seguida de readministración aproximadamente a los 3 meses de edad, aproximadamente 6 meses, aproximadamente 9 meses o aproximadamente 12 meses. Opcionalmente, se permite más de una readministración. Dicha readministración puede ser con el mismo tipo de vector o con un vector vírico diferente. 30

En un aspecto de la divulgación, se utiliza un sistema de administración basado en ARN para hOTC funcional para estabilizar a un sujeto (p. ej., un paciente humano) en crisis, seguido de la administración de una administración mediada por un vector vírico de una hOTC funcional. 35

En otro aspecto de la divulgación, la terapia inicial de la divulgación implica la coadministración de sistemas de administración de hOTC mediados y no mediados por virus. 40

En un aspecto adicional de la divulgación, la hOTC de las construcciones de ADN y ARN de la divulgación se puede utilizar sola o en combinación con el tratamiento de referencia para el diagnóstico y el estado del paciente. 45

Como se describe en los ejemplos funcionales del presente documento, los inventores han descubierto que los sujetos heterocigotos para OTCD, incluyendo los que tienen OTCD de aparición tardía, tiene un aumento de fibrosis y/o esteatosis microvesicular en todo el hígado. Dichas fibrosis hepática y/o esteatosis microvesicular puede conducir a cirrosis relacionada con OTCD. Por tanto, en otro aspecto, la divulgación proporciona vectores víricos o rAAV para su uso en la prevención de la fibrosis hepática y/o la afección médica asociada a la cirrosis relacionada con OTCD mediante la administración al sujeto (p. ej., un paciente humano) de una hOTC. Este aspecto de la divulgación puede utilizar un sistema de administración vírico. El casete de expresión de ácido nucleico puede contener un ADN o ARN de hOTC sintético como se proporciona en el presente documento u otra secuencia adecuada que expresa hOTC funcional. En un aspecto, se proporciona un uso de vectores víricos o rAAV para tratar y/o prevenir la fibrosis hepática, esteatosis microvesicular y/o cirrosis relacionada con OTCD que implica administrar OTCasa a un sujeto que tiene 50 OTCD. El sujeto puede ser un paciente humano. En una realización, el paciente es heterocigoto y tiene OTCD de aparición tardía. Es posible que el paciente no haya recibido tratamiento previo para OTCD o que haya recibido otros 55

tratamientos convencionales. En la actualidad, no existe un tratamiento de referencia para OTCD, sino que los síntomas se tratan, p. ej., mediante la interrupción de la ingesta de proteínas, aumentos compensatorios de los carbohidratos y lípidos en la dieta, hemodiálisis para pacientes comatosos con niveles sanguíneos extremadamente altos; y/o administración intravenosa de benzoato de sodio, arginina y/o fenilacetato de sodio. La FDA de los Estados Unidos ha aprobado el fenilbutirato de glicerol (Ravicti®) para el tratamiento a largo plazo de algunos trastornos del ciclo de la urea en pacientes de 2 años o más; este fármaco ayuda a eliminar el amoníaco del cuerpo y está destinado a pacientes que no pueden ser tratados únicamente con una dieta restringida en proteínas o suplementos de aminoácidos. En un aspecto, el tratamiento del paciente (p. ej., una primera inyección) se inicia antes del primer año de vida. En otro aspecto, el tratamiento se inicia después del primer año, o después de los primeros 2 a 3 años de edad, después de los 5 años de edad, después de los 11 años o en una edad mayor.

En un aspecto, los vectores víricos o rAAV de la divulgación proporcionan un uso para tratar y/o invertir la fibrosis hepática y/o la cirrosis relacionada con OTCD mediante la administración al sujeto de una OTCasa funcional que está codificada por un ADN modificado por ingeniería genética de las SEQ ID NO: 1, 3, 4 o 5, o un ADN quimérico como se define en el presente documento. La administración del ADN puede estar mediada por un vector vírico que contiene el modificado por ingeniería genética en un casete de expresión, que media en la expresión de OTCasa funcional en las células hepáticas del sujeto. En otra realización, se administra al sujeto un ARN quimérico como se define en el presente documento. La administración del ARN puede estar mediada por un vector vírico que contiene el ARN modificado por ingeniería genética en un casete de expresión, que media en la expresión de OTCasa funcional en las células hepáticas del sujeto.

Los sujetos con OTCD heterocigotos tienen un mayor riesgo de desarrollar carcinoma hepatocelular (HCC). Véase, p. ej., JM Wilson *et al.*, Molecular Genetics and Metabolism (2012), Mol Genet Metab. feb de 2012; 105 (2): 263-265, publicado en línea el 7 de noviembre de 2011. Por tanto, en otro aspecto, la divulgación proporciona vectores víricos o rAAV para su uso en el tratamiento y/o la prevención de HCC al administrar al sujeto (p. ej., un paciente humano) una hOTC. Este aspecto de la invención utiliza un sistema de administración vírico. El casete de expresión de ácido nucleico puede contener un ADN o ARN de hOTC sintético como se proporciona en el presente documento. Es posible que el paciente no haya recibido tratamiento previo para OTCD o que haya recibido otros tratamientos convencionales, es decir, tratamiento de referencia. En una realización, el tratamiento del paciente (p. ej., una primera inyección) se inicia antes del diagnóstico de HCC. En otra realización, el tratamiento del paciente se inicia tras el diagnóstico de HCC. Opcionalmente, el tratamiento implica la coadministración con sorafenib (disponible en el mercado como Nexavar®) o su uso junto con quimioembolización, radiación, ablación térmica, inyección de etanol precutáneo, terapia dirigida (p. ej., fármacos antiangiogénesis), infusión arterial hepática de fármacos antineoplásicos, inmunoterapia, o con opciones quirúrgicas incluyendo, p. ej., extirpación, criocirugía y trasplante de hígado. Cuando se usa para el tratamiento del HCC, puede ser deseable seleccionar un sistema de administración no integrador (p. ej., virus no integradores tales como adenovirus o lentivirus no integradores) para administración de un ADN o ARN de hOTC sintético como se describe en el presente documento.

Por "OTC funcional", se entiende un gen que codifica la OTCasa de tipo silvestre tal como se caracteriza por la SEQ ID NO: 2 u otra OTCasa que proporciona al menos aproximadamente 50 %, al menos aproximadamente 75 %, al menos aproximadamente 80 %, al menos aproximadamente 90 % o aproximadamente el mismo o más de 100 % del nivel de actividad biológica de la enzima ornitina transcarbamilasa humana de tipo silvestre, que se puede caracterizar por la secuencia de la SEQ ID NO: 2 o una variante natural o polimorfo de la misma que no está asociado con la enfermedad. Más en particular, ya que los pacientes heterocigotos pueden tener un nivel funcional de OTCasa tan bajo como aproximadamente 50 % o menos, el tratamiento eficaz puede no requerir el reemplazo de la actividad OTCasa a niveles dentro del rango de pacientes "normales" o sin OTCD. De manera similar, los pacientes que no tienen cantidades detectables de OTCasa pueden ser rescatados administrando la función OTCasa a niveles de actividad inferiores al 100 % y, opcionalmente, pueden someterse a un tratamiento adicional posteriormente. Como se describe en el presente documento, la terapia génica descrita en el presente documento, ya sea vírica o no vírica, se puede utilizar junto con otros tratamientos, es decir, el tratamiento de referencia para el diagnóstico del sujeto (paciente).

En una realización, dicha OTCasa funcional tiene una secuencia que tiene aproximadamente 95 % o más de identidad con la proteína madura (es decir, aproximadamente los últimos 322 aminoácidos) o la secuencia de longitud completa de la SEQ ID NO: 2, o aproximadamente 97 % de identidad o más, o aproximadamente 99 % o más con la SEQ ID NO: 2 al nivel de aminoácidos. Dicha OTCasa funcional también puede abarcar polimorfos naturales que no están asociados con ninguna enfermedad (p. ej., F101, L111 y/o W1193-194 de la SEQ ID NO: 2). La identidad se puede determinar preparando una alineación de las secuencias y mediante el uso de diversos algoritmos y/o programas informáticos conocidos en la técnica o disponibles en el mercado [p. ej., BLAST, ExPASy; ClustalO; FASTA; utilizando, p. ej., algoritmo de Needleman-Wunsch, algoritmo de Smith-Waterman].

Existen diversos ensayos para medir la expresión y los niveles de actividad de OTC *in vitro*. Véase, p. ej., X Ye, *et al.*, 1996 Prolonged metabolic correction in adult ornithine transcarbamilase-deficient mice with adenoviral vectors. J Biol Chem 271: 3639-3646) o *in vivo*. Por ejemplo, la actividad enzimática de OTC se puede medir utilizando un método de dilución de isótopos estables por espectrometría de masas de cromatografía líquida para detectar la formación de citrulina normalizada a [1,2,3,4,5-¹³C₅] citrulina (98 % ¹³C). El método está adaptado de un ensayo desarrollado

previamente para la detección de la actividad de la N-acetilglutamato sintasa [Morizono H, *et al.*, Mammalian N-acetylglutamate synthase. *Mol Genet Metab.* 2004; 81 (Supl. 1): S4-11.]. Se pesan cortes de hígado fresco congelado y se homogeneizan brevemente en tampón que contiene HEPES 10 mM, Triton X-100 al 0,5 %, EDTA 2,0 mM y DTT 0,5 mM. El volumen de tampón de homogeneización se ajusta para obtener 50 mg/ml de tejido. La actividad enzimática se mide utilizando 250 µg de tejido hepático en Tris-acetato 50 mM, ornitina 4 mM, fosfato de carbamilo 5 mM, pH 8,3. La actividad enzimática se inicia con la adición de carbamilo fosfato 50 mM recién preparado disuelto en Tris-acetato 50 mM pH 8,3, se dejó continuar durante 5 minutos a 25 °C y se inactivó mediante la adición de un volumen igual de 13C5-citrulina 5 mM en TCA al 30 %. Los desechos se separan mediante 5 minutos de microcentrifugación y los sobrenadantes se transfieren a frascos para espectroscopia de masas. Se inyectan diez µl de muestra en un LC-MS Agilent serie 1100 en condiciones isocráticas con una fase móvil de 93 % de disolvente A (1 ml de ácido trifluoroacético en 1 l de agua):7 % de disolvente B (1 ml de ácido trifluoroacético en 1 l de agua/acetoniitrilo 1:9). Se cuantifican los picos correspondientes a citrulina [176,1 de relación masa:carga (m/z)] y 13C5-citrulina (181,1 m/z), y sus relaciones se comparan con las relaciones obtenidas para una curva patrón de citrulina ejecutada con cada ensayo. Las muestras se normalizan con respecto al tejido hepático total o a la concentración de proteína determinada mediante un equipo de ensayo de proteínas Bio-Rad (Bio-Rad, Hercules, CA). También se pueden utilizar otros ensayos, que no requieren biopsia hepática. Uno de estos ensayos es un ensayo de aminoácidos plasmáticos en el que se evalúa la proporción de glutamina y citrulina y si la glutamina es alta (>800 microlitros/litro) y la citrulina baja (p. ej., menos de 10), se sospecha un defecto del ciclo de la urea. Se pueden medir los niveles de amoníaco en plasma y una concentración de aproximadamente 100 micromoles por litro es indicativa de OTCD. Se puede realizar una gasometría si un paciente está hiperventilando; la alcalosis respiratoria es frecuente en la OTCD. El ácido orótico en la orina, p. ej., más de aproximadamente 20 micromoles por milimol de creatina es indicativo de OTCD, al igual que el orotato urinario elevado después de la prueba de provocación con alopurinol. Los criterios de diagnóstico para OTCD se han expuesto en Tuchman *et al.*, 2008, Consorcio de Trastornos del Ciclo de la Urea (UCDC) de la Red de Investigación Clínica de Enfermedades Raras (RDCRN). Tuchman M, *et al.*, Consorcio de la Red de Investigación Clínica de Enfermedades Raras. Cross-sectional multicenter study of patients with urea cycle disorders in the United States. *Mol Genet Metab.* 2008; 94: 397-402. Véase, también, <http://www.ncbi.nlm.nih.gov/books/NBK.154378/>, que proporciona un análisis del tratamiento de referencia para OTCD.

Cabe señalar que el término "un" o "uno/a" se refiere a uno o más. Como tales, las expresiones "un" (o "uno/a"), "uno/a o más" y "al menos uno/a" se usan indistintamente en el presente documento.

Las palabras "comprender", "comprende" y "que comprende" deben interpretarse de manera inclusiva en lugar de exclusiva. Las palabras "consistir", "que consiste" y sus variantes, deben interpretarse de manera exclusiva, en lugar de inclusiva. Aunque se presentan diversas realizaciones en la memoria descriptiva que utilizan la expresión "que comprende", en otras circunstancias, también se pretende que una realización relacionada también sea interpretada y descrita usando la expresión "que consiste en" o "que consiste esencialmente en".

Como se usa en el presente documento, el término "aproximadamente" significa una variabilidad del 10 % (± 10 %) con respecto a la referencia dada, a menos que se especifique otra cosa.

El término "regulación" o variaciones del mismo como se usa en el presente documento se refiere a la capacidad de un compuesto de fórmula (I) para inhibir uno o más componentes de una ruta biológica.

Un "sujeto" es un mamífero, p. ej., un ser humano, ratón, rata, cobaya, perro, gato, caballo, vaca, cerdo o primate no humano, tal como un mono, chimpancé, babuino o gorila.

Como se usa en el presente documento, "enfermedad", "trastorno" y "afección" se usan indistintamente, para indicar un estado anómalo en un sujeto.

A menos que se definan de otro modo en la presente memoria descriptiva, los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que el que entiende habitualmente un experto habitual en la materia y por referencia a textos publicados, que proporcionan a un experto en la materia una guía general para muchos de los términos utilizados en la presente solicitud.

Los siguientes ejemplos son ilustrativos.

Ejemplo 1: vectores scAAV que contienen hOTC

pAAVsc.TBG.hOTCwt y pAAVsc.TBG.hOTCco-LW4 se construyeron reemplazando la secuenciación de codificación mOTC con ADNc de hOTC de tipo silvestre (TS) (hOTCts) o hOTCcoLW, respectivamente, en un plásmido procedente del pAAVsc.TBG.mOTC1.3 previamente descrito con el intrón alterado [Moscioni D, *et al.*, "Long-term correction of ammonia metabolism and prolonged survival in ornithine transcarbamylase-deficient mice following liver-directed treatment with adeno-associated viral vectors", *Mol Ther.* 2006; 14: 25-33; Cunningham SC, *et al.*, "AAV2/8-mediated correction of OTC deficiency is robust in adult but not neonatal Spf(ash) mice", *Mol Ther.* 2009; 17: 1340-1346; Wang L, *et al.*, "Sustained correction of OTC deficiency in spfashmice using optimized self-complementary AAV2/8 vectors", *Gene Ther.* abril de 2012; 19 (4): 404-10, publicación electrónica del 18 de agosto de 2011].

El scAAV2/8.TBG.hOTCco-LW4 contiene una ITR 3' y un ITR 5' de AAV2 con una supresión en la secuencia D y trs (sitio de resolución terminal), un promotor de TBG, el gen de hOTCco-LW4 y un poliA de SV40 de 137 pb. Las dos preparaciones de vectores (AAV2/8sc.TBG.hOTCwt y AAV2/8sc.TBG.hOTCco-LW4) utilizadas en el estudio de comparación inicial se purificaron mediante dos ciclos de centrifugación en gradiente de cloruro de cesio, como se ha descrito anteriormente [Wang L, *et al.*, Systematic evaluation of AAV vectors for liver directed gene transfer in murine models. *Mol Ther.* 2010; 18: 118-125]. Los vectores utilizados en el resto del estudio se produjeron mediante un método de producción a escala basado en la transfección de polietilenimina (PEI) y se purificaron a partir del sobrenadante o lisado total mediante centrifugación en gradiente de iodoxanol como se ha descrito [Lock M, *et al.*, *Hum Gene Ther.* 2010; 21: 1259-1271]. Se determinaron los títulos del genoma [copias del genoma (CG)/ml] de los vectores de AAV mediante PCR en tiempo real utilizando conjuntos de cebadores y sondas dirigidos al promotor de TBG (cebador directo 5'-AAACTGCCAATTCCACTGCTG-3' [SEQ ID NO: 14], cebador inverso 5'-CCATAGGCAAAAGCACCAAGA-3' [SEQ ID NO: 15], sonda 6FAM-TTGGCCCAATAGTGAGAACTTTTCCTGC [SEQ ID NO: 16] -TAMRA) y usando un plásmido linealizado como patrón. El cebador directo está ubicado 400 pb cadena abajo de la horquilla cerrada 5'. Fagone *et al.* [Systemic errors in quantitative polymerase chain reaction titration of self-complementary adeno-associated viral vectors and improved over alternative methods, *Hum Gene Ther Methods.* febrero de 2012; 23 (1): 1-7.] informaron recientemente de que el método de PCR cuantitativa (Q-PCR) podría subestimar significativamente el título de los vectores scAAV, especialmente cuando los cebadores de PCR estaban cerca de la horquilla cerrada del vector scAAV. El título de un lote de vector AAV2/8sc.TBG.hOTCco-LW4 utilizando un conjunto de cebadores y sondas dirigido a la región poliA (1900 pb cadena abajo de la horquilla cerrada 5'), y el título del genoma fue 1,1 veces el título original, que estaba dentro del error intraensayo de Q-PCR.

Se evaluaron los niveles de expresión de proteínas OTC y la actividad OTC en el hígado de ratones *spf^{ash}* 14 días después de la inyección i.v. de 1×10^{11} CG de los vectores AAV2/8sc.TBG.hOTCwt o AAV2/8sc.TBG.hOTCco-LW4. Los ratones *spf^{ash}* son un modelo de enfermedad de OTC de aparición tardía en seres humanos. Todos los procedimientos con animales se realizaron de acuerdo con los protocolos aprobados por el Comité Institucional de Uso y Cuidado de Animales (IACUC) de la Universidad de Pensilvania. Los ratones *spf^{ash}* se mantuvieron en las instalaciones para animales de los laboratorios de investigación traslacional de la Universidad de Pensilvania como se ha descrito anteriormente. En los estudios se utilizaron ratones *spf^{ash}* de tres a seis meses de edad y sus compañeros de camada normales. Los vectores se administraron mediante inyección intravenosa (i.v.) a través de la vena de la cola. El alcance de la transferencia de genes basada en genomas de vectores residentes no fue estadísticamente diferente entre los dos grupos. Se recogieron muestras de orina antes y en varios puntos de tiempo después del tratamiento con vector para el análisis de ácido orótico como se ha descrito anteriormente [Moscioni D, *et al.*, Long-term correction of ammonia metabolism and prolonged survival in ornithine transcarbamylase-deficient mice following liver-directed treatment with adeno-associated viral vectors. *Mol Ther.* 2006; 14: 25-33].

Se realizó análisis de transferencia de Western para detectar la expresión de hOTC en el lisado de hígado como se ha descrito anteriormente [Wang L, *et al.*, 2012, publicación electrónica 2011]]. El anticuerpo primario para detectar hOTC fue un anticuerpo policlonal de conejo personalizado proporcionado por el laboratorio de Hiroki Morizono en el Centro Médico Infantil Nacional. Los lisados de hígado (10 µg/carril) también se transfirieron y exploraron con un anticuerpo anti-tubulina (Abcam, Cambridge, MA). El análisis de Western demostró una expresión 100 veces mayor de hOTC del vector hOTCco-LW4 en comparación con el vector hOTCwt, alcanzando niveles ligeramente superiores a los observados en ratones TS.

La actividad enzimática de OTC se midió utilizando un método de dilución de isótopos estables por espectrometría de masas de cromatografía líquida para detectar la formación de citrulina normalizada a [1,2,3,4,5-¹³C₅] citrulina (98 % de ¹³C). El método está adaptado de un ensayo desarrollado previamente para la detección de la actividad de la N-acetilglutamato sintasa [Morizono H, *et al.*, Mammalian N-acetylglutamate synthase. *Mol Genet Metab.* 2004; 81 (Supl. 1): S4-11.]. Se pesaron cortes de hígado fresco congelado y se homogeneizaron brevemente en tampón que contenía HEPES 10 mM, Triton X-100 al 0,5 %, EDTA 2,0 mM y DTT 0,5 mM. El volumen de tampón de homogeneización se ajustó para obtener 50 mg/ml de tejido. La actividad enzimática se midió utilizando 250 µg de tejido hepático en Tris-acetato 50 mM, ornitina 4 mM, fosfato de carbamilo 5 mM, pH 8,3. La actividad enzimática se inició con la adición de carbamifosfato 50 mM recién preparado disuelto en Tris-acetato 50 mM pH 8,3, se dejó continuar durante 5 minutos a 25 °C y se inactivó mediante la adición de un volumen igual de ¹³C₅-citrulina 5 mM en TCA al 30 %. Los desechos se separaron mediante 5 minutos de microcentrifugación y los sobrenadantes se transfirieron a frascos para espectroscopia de masas. Se inyectaron diez µl de muestra en un LC-MS Agilent serie 1100 en condiciones isocráticas con una fase móvil de 93 % de disolvente A (1 ml de ácido trifluoroacético en 1 l de agua):7 % de disolvente B (1 ml de ácido trifluoroacético en 1 l de agua/acetronitrilo 1:9). Se cuantificaron los picos correspondientes a citrulina [176,1 de relación masa:carga (m/z)] y ¹³C₅-citrulina (181,1 m/z), y sus relaciones se compararon con las relaciones obtenidas para una curva patrón de citrulina ejecutada con cada ensayo. Las muestras se normalizaron con respecto al tejido hepático total o a la concentración de proteína determinada mediante un equipo de ensayo de proteínas Bio-Rad (Bio-Rad, Hercules, CA).

Se encontró que el vector que porta ADNc de hOTC modificado por ingeniería genética denominado en el presente documento LW4 (fig. 4) mejora los niveles de proteína hOTC expresada en 100 veces. Una evaluación de la actividad enzimática de OTC se correlacionó en general con los experimentos de transferencia de Western de OTC, aunque la

proteína OTC fue más elevada que la actividad enzimática de OTC en comparación con la OTC endógena. Al restar los niveles de actividad de fondo en los ratones *spf^{ash}*, el hOTCco-LW4 dio lugar a una actividad más de 33 veces mayor que la hOTCwt. Se observaron niveles de actividad y expresión de hOTC sostenidos y correlacionados con la dosis en los ratones *spf^{ash}* tratados. En comparación con un vector de OTC murino descrito anteriormente que se diferenciaba principalmente en el ADNc, el vector que portaba el vector hOTCco-LW4 era aproximadamente 10 veces más potente.

El vector ilustrativo que portaba el hOTCco-LW4 modificado (fig. 4) proporcionó un alto nivel de transducción, según lo medido por ensayos histológicos de OTC, a lo largo de una amplia gama de dosis. Entre las dosis 1×10^{11} CG y 3×10^9 CG, la eficacia de transducción, según lo medido por tinción histoquímica, varió entre 50-70 %. A la dosis más baja de 1×10^9 CG, el 40 % de las áreas hepáticas fueron positivas por tinción histoquímica de OTC. La falta de un efecto de dosis claro por histoquímica e inmunotinción podría deberse al hecho de que la optimización de codones mejoró significativamente la expresión de hOTC en los hepatocitos transducidos. Esto conduce a una sensibilidad mejorada para detectar células transducidas con bajas copias del genoma del vector. La transducción podría saturarse con altas dosis de vector (1×10^{11} - 1×10^{10} CG) y, por lo tanto, la eficacia de transducción medida por métodos de detección *in situ* no diferenciarían entre grupos de dosis baja y alta en contraste con la actividad enzimática de OTC en lisados hepáticos medida por espectrometría de masas.

Se realizó un estudio adicional en el que se evaluó la expresión neonatal de hOTC en ratones *spf^{ash}*, inyectado el día 1 de vida, utilizando scAAV2/8.TBG.hOTCco en una dosis de 5×10^{10} CG/cría inyectado a través de la vena temporal. Se detectó expresión robusta a las 24 y 48 horas. Se realizaron estudios adicionales utilizando dosis de 1×10^{11} , 3×10^{10} , y 1×10^{10} y se evaluaron durante 12 semanas. Durante el periodo de 16 semanas del estudio, se observó una reducción en los niveles iniciales de expresión robusta en cada una de las dosis. Se cree que esto se debe a la dilución, es decir, un resultado natural de la proliferación de células hepáticas en animales en crecimiento. Por tanto, aunque se observa restauración inicial de la actividad hepática de OTC después de la transferencia de genes neonatales en ratones *spf^{ash}*, este resultado es temporal, con un descenso de la actividad de OTC de aproximadamente 1000 % de los niveles de tipo silvestre (ts) en aproximadamente 1 semana, a aproximadamente 50 % de los niveles de ts a las 4 semanas, a aproximadamente 10 % de los niveles de ts a las 12 semanas (nivel de 1×10^{11} CG); o de aproximadamente 500 % de los niveles de ts en la semana 1, a aproximadamente 20 % de los niveles de ts o aproximadamente 10 % de los niveles de ts en la semana 1 (dosis de 3×10^{10} CG) o de niveles de aproximadamente 200 % de ts en la semana 1, a niveles de aproximadamente 10 % de ts en la semana 4 (dosis de 1×10^{10} CG). En un estudio, utilizando animales que reciben la primera inyección de 3×10^{10} CG el día 1, se inyectó a los animales un segundo vector AAV que portaba el gen de hOTCco (scAAVrh10.hOTCco; $1,8 \times 10^{10}$ CG) en la semana 4. Como control, un grupo de animales no recibió readministración y un grupo recibió solo el segundo vector a las 4 semanas. La readministración del AAV.hOTCco dio lugar a restauración de la actividad OTC del hígado.

Se diseñaron estudios adicionales para evaluar la capacidad de rescatar crías OTC-KO mediante terapia génica neonatal, tanto a corto como a largo plazo.

Ejemplo 2: Producción de vectores scAAV que tienen secuencias con codones optimizados

A. scAAV8.TBG.hOTC-co

Los plásmidos que contienen una secuencia de hOTC con codones optimizados de las SEQ ID NO: 3, 4, 5, 9 o 10, respectivamente, se clonan como se describe en el ejemplo 1 reemplazando la secuenciación codificante de mOTC con hOTCco en un plásmido derivado del pAAVsc.TBG.mOTC1.3 previamente descrito con el intrón. El plásmido pAAVsc.TBG.hOTCco resultante se clona en una cápside de AAV8 [Gao *et al.*, PNAS USA, 2002, 99: 11854-11859] utilizando técnicas convencionales.

B. scAAVrh10.TBG.hOTC-co

Los plásmidos que contienen una secuencia de hOTC con codones optimizados de las SEQ ID NO: 3, 4, 5, 9 o 10, respectivamente, se clonan como se describe en el ejemplo 1 reemplazando la secuenciación codificante de mOTC con hOTCco en un plásmido derivado del pAAVsc.TBG.mOTC1.3 previamente descrito con el intrón. Los plásmidos pAAVsc.TBG.hOTCco resultantes se clonan en una cápside de AAVrh10 [Gao *et al.*, PNAS USA, 2002, 99: 11854-11859] utilizando técnicas convencionales.

Ejemplo 3: Producción de vectores ssAAV que tienen secuencias con codones optimizados

ssAAV2/8.LSP1.hOTC-co

Los plásmidos que contienen las secuencias de hOTCco con codones optimizados se clonan como se describe reemplazando la secuenciación de codificación de mOTC del plásmido pLSP1mOTC [Cunningham *et al.*, Mol Ther, 2009, 17: 1340-1346] con la secuencia de ADNc correspondiente de las SEQ ID NO: 3, 4, 5, 9 o 10. Los plásmidos pAAVsc.LSP1.hOTCco resultantes se clonan en cápsides de AAV8 para formar los vectores ssAAV2/8.LSP1.hOTCco correspondientes utilizando técnicas descritas en el ejemplo 1.

B. ssAAV2/rh10.LSP1.hOTC-co

5 Los plásmidos que contienen las secuencias de hOTCco con codones optimizados se clonan como se describe reemplazando la secuenciación de codificación de mOTC del plásmido pLSP1mOTC [Cunningham *et al.*, Mol Ther, 2009, 17: 1340-1346] con la secuencia de ADNc correspondiente de las SEQ ID NO: 3, 4, 5, 9 o 10. Los plásmidos pAAVsc.LSP1.hOTCco resultantes se clonan en cápsides de AAV8 para formar los vectores ssAAV2/8.LSP1.hOTC-co correspondientes utilizando técnicas descritas en el ejemplo 1.

10 Los vectores generados según la Parte A o B pueden purificarse mediante dos ciclos de centrifugación en gradiente de cloruro de cesio, se intercambió el tampón con PBS y se concentraron utilizando dispositivos de filtro de centrifuga Amicon Ultra 15-100K (Millipore, Bedford, MA). El título del genoma (CG/ml) de los vectores de AAV se puede determinar mediante PCR en tiempo real utilizando un conjunto de cebador/sonda correspondiente al promotor de TBG y los patrones de plásmido linealizado. Los vectores pueden someterse a pruebas de control de calidad
15 adicionales, incluyendo análisis de electroforesis en gel de poliacrilamida-dodecil sulfato de sodio (SDS-PAGE) para determinar la pureza del vector y el lisado de amebocitos de *Limulus* (LAL) para la detección de endotoxinas (Cambrex Bio Science, Walkersville, MD, Estados Unidos).

Ejemplo 4: ssAAV8.TBG.hOTCco en el modelo de inicio tardío de OTCD

20 Se generó un vector de AAV8 utilizando los métodos descritos en el presente documento. El vector tiene empaquetado en el mismo una ITR 5' de AAV2, un promotor de TBG, un intrón, una hOTCco, un elemento WPRE, una poliA de hormona del crecimiento bovina y una ITR 3' de AAV2. La expresión y cinética de este vector se comparó con un vector de AAV8 autocomplementario con o sin el elemento WPRE. Los resultados muestran que las construcciones
25 monocatenarias con el elemento WPRE superaron los vectores que carecen del elemento WPRE; a dosis comparables, ambos vectores monocatenarios (con y sin WPRE) fueron menos robustos que el vector autocomplementario que carece de WPRE en los puntos temporales medidos.

30 Sin embargo, los vectores monocatenarios pueden tener otras propiedades deseables, p. ej., con respecto a cinética, dependiendo de la edad y condición del paciente.

Ejemplo 5: Producción de vectores de adenovirus que tienen secuencias con codones optimizados

35 El ADNc de hOTCco [SEQ ID NO: 3, 4, 5, 9 y 10] con conectores NotI se clona cadena abajo de un promotor de PEPCK de rata para generar pPEPCK-hOTC como se describe en A. Mian *et al.*, Molecular Therapy, 2004, 10: 492-499 (2004). Este plásmido se digiere con *Ascl* y el fragmento de PEPCK-hOTCco resultante se inserta en el plásmido de cadena principal adenovírica pC4HSU31 para generar los plásmidos parentales pC4HSU-PEPCK-hOTCco. El plásmido pWPRE se digiere con *Clal* para liberar el WPRE, que después se inserta en el sitio *Mlu*I de pPEPCK-hOTC, para generar el plásmido pPEPCK-hOTCco-WPRE con sus secuencias hOTCco respectivas. Las etapas restantes
40 para generar el plásmido adenovírico pC4HSU-PEPCK-hOTCco-WPRE son como se ha descrito anteriormente. Todos los sitios de clonación se confirman mediante análisis de secuencia de ADN. La identidad de los plásmidos adenovíricos recombinantes puede confirmarse mediante digestión con enzimas de restricción con *Hind*III y *Bam*HI. Los plásmidos adenovíricos se linealizan con *Pme*I antes de la transfección en células 293Cre4. Se rescatan y amplifican vectores adenovíricos con células 293Cre4 y el virus auxiliar *AdC8cluc*. Pueden utilizarse células 293N3Scre8 en suspensión en la etapa final de la producción del vector. Se realizan purificación, cuantificación
45 mediante DO260 y extracción de ADN vírico como se describe en detalle en otra parte [Brunetti-Pierri, N., *et al.* (2004). Acute toxicity after high-dose systemic injection of helper-dependent adenoviral vectors into nonhuman primates. Hum. Gene Ther. 15: 35-46; Ng, P., Parks, R. J. y Graham, F. L. (2002). Preparation of helper-dependent adenoviral vectors. Methods Mol. Med. 69: 371-388].

50 Ejemplo 6: Producción de vectores lentivíricos hOTCco

A. Los vectores lentivíricos de replicación defectuosa que contienen las secuencias de hOTCco proporcionadas en el presente documento pueden producirse reemplazando el inserto de secuencia del gen de OTC de rata del
55 plásmido pLenti-GIII-CMV-GFP-2A-Puro [disponible en el mercado de Applied Biological Materials (ABM) Inc.; Canadá] con la secuencia de hOTCco deseada [SEQ ID NO: 3, 4, 5, 9 y 10]. Los virus se generan según las instrucciones del fabricante. El sistema ABM incluye una supresión potenciadora en la región U3 de 3'ΔLTR para garantizar la auto-inactivación del vector lentivírico después de la transducción e integración en el ADN genómico de la célula diana; contiene genes lentivíricos mínimos necesarios para el empaquetamiento, la replicación y la transducción (*Gag/Pol/Rev*), procedentes de diferentes plásmidos que carecen todos de señales de empaquetamiento; además, no se incorpora ninguno de los genes *Gag*, *Pol* o *Rev* en el genoma vírico empaquetado, haciendo de este modo que el virus maduro sea incompetente para replicación.

60 B. Vectores lentivíricos de hOTC de replicación defectuosa, no integradores Una construcción de ADN que contiene un promotor específico del hígado y el ADN de hOTCco de las SEQ ID NO: 3, 4, 5, 9 y 10 se introducen por ingeniería genética en vectores de lentivirus que se pseudotipan en el virus sindbis E2 con envoltura producido como se describe en el documento US2011/0064763. Todos los vectores contienen donante de corte y empalme,

señal de empaquetamiento (psi), un elemento de respuesta a Rev (RRE), donante de corte y empalme, aceptor de corte y empalme, tramo central de poli-purina (cPPT). El elemento WPRE se elimina en determinados virus.

C. El ADN de hOTCco de las SEQ ID NO: 3, 4, 5, 9 y 10, se clona en un lentivirus pseudotipificado con un gen de la envoltura del virus de la estomatitis vesicular (VSV), obtenido de InvivoGen (SanDiego, CA) usando las instrucciones del fabricante.

Ejemplo 7: Sistemas de suministro de ARN de hOTCco de producción

El ARN puede prepararse mediante transcripción *in vitro* a partir de un molde de ADN o sintetizarse. Se prepara el casete de expresión de ARN que incluye una UTR 5', un intrón opcional con sitios donantes y aceptores de corte y empalme, una secuencia Kozak opcional, la secuencia codificante de hOTC proporcionada en el presente documento, una poliA y una UTR 3' usando técnicas conocidas.

A. Se incorpora una cantidad adecuada de ARNm en nanopartículas poliméricas sensibles al pH con envoltura lipídica generadas usando técnicas publicadas. [X. Su *et al.*, Mol. Pharmaceutics, 2011, 8 (3), págs. 774-787; publicación web: 21 de marzo de 2011].

B. Las formulaciones de nanopartículas poliméricas con polietilenimina (PEI) ramificada de 25 kDa se preparan de la siguiente manera. Cuando está presente PEI, esta puede ser PEI ramificada de un peso molecular que varía de 10 a 40 kDa, p. ej., PEI ramificada de 25 kDa (Sigma n.º 408727). Los polímeros ilustrativos adicionales adecuados para la presente invención incluyen los descritos en la publicación PCT WO2013182683. Se diluye la cantidad necesaria de ARNm justo antes de su aplicación en agua para inyección (Braun, Melsungen) hasta un volumen total de 4 ml y se añade rápidamente a 4 ml de una solución acuosa de PEI ramificada de 25 kDa utilizando una pipeta en una relación N/P de 10. La solución se mezcla pipeteando arriba y abajo.

C. Para una nanopartícula a base de lípidos, se crea una formulación de lípidos utilizando un casete de expresión que contiene el ARN de hOTCco en una formulación de cK - E12:DOPE:Col:PEG-DMG2K (cantidades relativas 50:25:20:5 (mg:mg:mg:mg)) para proporcionar una solución para la administración. Se utiliza el lípido catiónico cK -E12 (véase, p. ej., documento WO 2013/063468) y se combina con dioleoilfosfatidil-etanolamina o "DOPE", colesterol (col) y polietilenglicol (PEG) o un lípido PEGilado (PEG-DMG2K) utilizando los métodos de formulación descritos en las publicaciones de patente internacional WO 2010/053572 y WO 2012/170930.

Ejemplo 8: Sistemas de suministro de ADN de hOTCco

A. ADN plasmídico desnudo: las secuencias de hOTCco [SEQ ID NO: 3, 4 o 5] se obtienen por ingeniería genética como construcciones de ADN plasmídico desnudo que se administran a una célula hepática diana (p. ej., mediante administración intravascular) y expresan la proteína OTC humana en la célula diana.

B. Complejos de ADN-lípido catiónico - Se preparan complejos de ADN-lípido catiónico utilizando una cantidad adecuada de un casete de expresión que contiene al menos un promotor, un intrón opcional, una secuencia Kozak opcional, una hOTCco de las SEQ ID NO: 3, 4 o 5, una poliA y otras secuencias de control de la expresión opcionales. El promotor puede ser un promotor específico de hígado. Como alternativa, puede seleccionarse otro promotor no específico de tejido. Por ejemplo, se formula una cantidad adecuada de ADN con un lípido catiónico de cK - E12:DOPE:Col:PEG-DMG2K (cantidades relativas 50:25:20:5 (mg:mg:mg:mg)) para formar un complejo de ADN - lípido catiónico adecuado para su administración a un sujeto. Se utiliza el lípido catiónico cK -E12 (véase, p. ej., documento WO 2013/063468) y se combina con dioleoilfosfatidil-etanolamina o "DOPE", colesterol (col) y polietilenglicol (PEG) o un lípido PEGilado (PEG-DMG2K) utilizando los métodos de formulación descritos en las publicaciones de patente internacional WO 2010/053572 y WO 2012/170930.

Ejemplo 9 Corrección a largo plazo de una forma letal neonatal de deficiencia de OTC mediante múltiples tratamientos con vectores de AAV de diferentes serotipos

En el estudio actual, el scAAV8.TBG.hOTCcoLW4 preparado como se ha descrito en el ejemplo 1 se utilizó para rescatar animales en un modelo de ratón de OTCD de inicio neonatal (temprano). Se generaron ratones OTC KO mediante la supresión de los exones 2-3 y se caracterizaron las propiedades de este ratón en términos de similitud con pacientes humanos con mutaciones nulas de OTC. En resumen, el modelo de inactivación de OTC (KO) generado en el laboratorio de los inventores a través de la supresión de los exones 2-3 imita estrechamente la forma de aparición neonatal grave de OTCD en seres humanos. Las crías machos OTC KO neonatales tienen niveles elevados de amoníaco en plasma debido a la ausencia de expresión de OTC en el hígado e inevitablemente mueren en las primeras 24 horas después del nacimiento. Las hembras heterocigotas se reproducen normalmente, tienen niveles normales de amoníaco en plasma, reducción de la actividad enzimática de OTC hepática, niveles elevados de ácido orótico en orina y, en algunos casos, menor peso corporal en comparación con los compañeros de camada de tipo silvestre (TS). Una sola inyección de vector scAAV8-hOTCco preparado como se ha descrito en el ejemplo 1 a una dosis de 1-3x10¹⁰ CG/cría inmediatamente después del nacimiento es capaz de rescatar a las crías OTC KO y extender la vida a 6 semanas. Para lograr una corrección a largo plazo, un grupo de ratones OTC-KO de 4 semanas de edad recibió una segunda administración de vector del vector scAAVrh10-hOTCco, que se había preparado como se ha descrito en el ejemplo 1.

Se han rescatado con éxito más de 30 crías OTC-KO obtenidos por cesárea con la administración de genes. Las crías

rescatadas tenían un peso corporal más bajo que sus compañeros de camada normales y tenían un fenotipo transitorio de pelaje escaso y piel anómala. Lo que es más importante, sus niveles de amoníaco en plasma estaban en el intervalo normal. Sin embargo, la eficacia no puede mantenerse más allá de las 6 semanas debido a la pérdida del genoma del vector durante la rápida proliferación hepática en la etapa neonatal. Una segunda administración de vector del vector scAAVrh10-hOTCco en ratones OTC-KO de 4 semanas de edad puede prolongar adicionalmente sus vidas hasta la adultez. Los ratones más viejos han alcanzado los 18 meses de edad. Los ratones rescatados a largo plazo muestran niveles cercanos a los normales de amoníaco en plasma, aunque los niveles de ácido orótico en orina en un subconjunto de estos ratones estaban significativamente elevados. La tinción con rojo sirio en muestras de hígado de ratones heterocigotos de diferentes edades (6, 12 y 18 meses de edad) mostró fibrosis hepática en ratonas heterocigotas OTC-KO envejecidas (18 meses), similar a una muestra de hígado de un paciente con OTCD de 11 años de edad.

Ejemplo 10: Tratamiento de la deficiencia de OTC de inicio tardío (OTCD)

Ratones heterocigotos OTC-KO de dos meses de edad recibieron una única inyección en la vena de la cola de un vector de AAV8 autocomplementario que codifica un gen de OTC humano con codones optimizados (SEQ ID NO: 5) a 1×10^{10} , 3×10^{10} y 1×10^{11} copias del genoma del vector por ratón. Una semana después del tratamiento con vectores, los ratones de los tres grupos de dosis de vector tenían niveles normales de ácido orótico en orina que se mantuvieron durante todo el estudio (16 meses). Se recogieron muestras de hígado de ratones tratados de 18 meses de edad para el análisis patológico y se compararon con ratones heterocigotos sin tratar y compañeros de camada TS de la misma edad. Todos los ratones tratados mostraron una histología hepática normal similar a TS, en contraste con los animales heterocigotos sin tratar que tenían fibrosis en todo el hígado. En conclusión, una sola inyección del vector AAV8sc-hOTCco puede prevenir la fibrosis hepática en OTC-KO heterocigotos y tiene un gran potencial para la corrección de la fibrosis hepática en pacientes con OTCD.

Los vectores de terapia génica descritos en el presente documento tienen capacidad de expresión génica rápida, robusta y prolongada incluso en ratones con una falta total de OTC. Las hembras heterocigotas son capaces de reproducirse y dar a luz descendientes masculinos hemicigotos, pero estas crías mueren al día de su nacimiento si no se tratan. Las ratonas heterocigotas viejas sin tratar muestran pruebas de aumento de fibrosis y esteatosis microvesicular, un hallazgo que parece similar a las observaciones en pacientes humanos heterocigotos. Se ha desarrollado un régimen de transferencia de genes que puede rescatar a los machos afectados y los machos tratados han sobrevivido durante 72 semanas.

Por tanto, estos datos demuestran que la terapia génica específica del hígado con hOTC puede prevenir la fibrosis hepática. Estos datos se correlacionan con el tratamiento de seres humanos heterocigotos con deficiencia de OTC, p. ej., sujetos que tienen inicio tardío de OTCD.

(Texto independiente del listado de secuencias)

La siguiente información se proporciona para secuencias que contienen texto independiente con el identificador numérico <223>.

SEQ ID NO: (que contiene texto independiente)	Texto independiente bajo <223>
3	hOTC modificada por ingeniería genética
4	hOTC modificada por ingeniería genética
5	hOTC modificada por ingeniería genética
6	Plásmido pscAAVTBGhOTCLW
8	hOTC modificada por ingeniería genética
9	hOTC modificada por ingeniería genética
10	ARN de hOTC modificada por ingeniería genética
11	ARN de hOTC modificada por ingeniería genética
12	ARN de hOTC modificada por ingeniería genética
13	ARN de hOTC modificada por ingeniería genética
14	Cebador directo de PCR
15	Cebador inverso de PCR
16	Sonda

LISTADO DE SECUENCIAS

5 <110> Los Fideicomisarios de la Universidad de Pensilvania
 <120> COMPOSICIONES ÚTILES EN EL TRATAMIENTO DE LA DEFICIENCIA DE OTC
 <130> UPN-14-7037PCT
 10 <150> 61/950.157
 <151> 09/03/2014
 <160> 16
 15 <170> PatentIn versión 3.5
 <210> 1
 <211> 1062
 <212> ADN
 20 <213> *Homo sapiens*
 <220>
 <221> CDS
 <222> (1)..(1062)
 25 <400> 1

atg ctg ttt aat ctg agg atc ctg tta aac aat gca gct ttt aga aat
 48
 Met Leu Phe Asn Leu Arg Ile Leu Leu Asn Asn Ala Ala Phe Arg Asn
 1 5 10 15

ggt cac aac ttc atg gtt cga aat ttt cgg tgt gga caa cca cta caa
 96
 Gly His Asn Phe Met Val Arg Asn Phe Arg Cys Gly Gln Pro Leu Gln
 20 25 30

aat aaa gtg cag ctg aag ggc cgt gac ctt ctc act cta aaa aac ttt
 144
 Asn Lys Val Gln Leu Lys Gly Arg Asp Leu Leu Thr Leu Lys Asn Phe
 35 40 45

acc gga gaa gaa att aaa tat atg cta tgg cta tca gca gat ctg aaa
 192
 Thr Gly Glu Glu Ile Lys Tyr Met Leu Trp Leu Ser Ala Asp Leu Lys
 50 55 60

ttt agg ata aaa cag aaa gga gag tat ttg cct tta ttg caa ggg aag
 240
 Phe Arg Ile Lys Gln Lys Gly Glu Tyr Leu Pro Leu Leu Gln Gly Lys
 65 70 75 80

tcc tta ggc atg att ttt gag aaa aga agt act cga aca aga ttg tct
 288
 Ser Leu Gly Met Ile Phe Glu Lys Arg Ser Thr Arg Thr Arg Leu Ser

ES 2 821 938 T3

				85					90					95		
aca	gaa	aca	ggc	ttt	gca	ctt	ctg	gga	gga	cat	cct	tgt	ttt	ctt	acc	
336																
Thr	Glu	Thr	Gly	Phe	Ala	Leu	Leu	Gly	Gly	His	Pro	Cys	Phe	Leu	Thr	
			100					105					110			
aca	caa	gat	att	cat	ttg	ggg	gtg	aat	gaa	agt	ctc	acg	gac	acg	gcc	
384																
Thr	Gln	Asp	Ile	His	Leu	Gly	Val	Asn	Glu	Ser	Leu	Thr	Asp	Thr	Ala	
		115					120					125				
cgt	gta	ttg	tct	agc	atg	gca	gat	gca	gta	ttg	gct	cga	gtg	tat	aaa	
432																
Arg	Val	Leu	Ser	Ser	Met	Ala	Asp	Ala	Val	Leu	Ala	Arg	Val	Tyr	Lys	
	130					135					140					
caa	tca	gat	ttg	gac	acc	ctg	gct	aaa	gaa	gca	tcc	atc	cca	att	atc	
480																
Gln	Ser	Asp	Leu	Asp	Thr	Leu	Ala	Lys	Glu	Ala	Ser	Ile	Pro	Ile	Ile	
145					150					155					160	
aat	ggg	ctg	tca	gat	ttg	tac	cat	cct	atc	cag	atc	ctg	gct	gat	tac	
528																
Asn	Gly	Leu	Ser	Asp	Leu	Tyr	His	Pro	Ile	Gln	Ile	Leu	Ala	Asp	Tyr	
				165					170					175		
ctc	acg	ctc	cag	gaa	cac	tat	agc	tct	ctg	aaa	ggg	ctt	acc	ctc	agc	
576																
Leu	Thr	Leu	Gln	Glu	His	Tyr	Ser	Ser	Leu	Lys	Gly	Leu	Thr	Leu	Ser	
			180					185					190			
tgg	atc	ggg	gat	ggg	aac	aat	atc	ctg	cac	tcc	atc	atg	atg	agc	gca	
624																
Trp	Ile	Gly	Asp	Gly	Asn	Asn	Ile	Leu	His	Ser	Ile	Met	Met	Ser	Ala	
		195					200					205				
gcg	aaa	ttc	gga	atg	cac	ctt	cag	gca	gct	act	cca	aag	ggg	tat	gag	
672																
Ala	Lys	Phe	Gly	Met	His	Leu	Gln	Ala	Ala	Thr	Pro	Lys	Gly	Tyr	Glu	
	210					215					220					
ccg	gat	gct	agt	gta	acc	aag	ttg	gca	gag	cag	tat	gcc	aaa	gag	aat	
720																
Pro	Asp	Ala	Ser	Val	Thr	Lys	Leu	Ala	Glu	Gln	Tyr	Ala	Lys	Glu	Asn	
225					230					235					240	
ggg	acc	aag	ctg	ttg	ctg	aca	aat	gat	cca	ttg	gaa	gca	gcg	cat	gga	
768																
Gly	Thr	Lys	Leu	Leu	Leu	Thr	Asn	Asp	Pro	Leu	Glu	Ala	Ala	His	Gly	
			245						250					255		
ggc	aat	gta	tta	att	aca	gac	act	tgg	ata	agc	atg	gga	caa	gaa	gag	
816																

ES 2 821 938 T3

Gly Asn Val Leu Ile Thr Asp Thr Trp Ile Ser Met Gly Gln Glu Glu
 260 265 270

gag aag aaa aag cgg ctc cag gct ttc caa ggt tac cag gtt aca atg
 864

Glu Lys Lys Lys Arg Leu Gln Ala Phe Gln Gly Tyr Gln Val Thr Met
 275 280 285

aag act gct aaa gtt gct gcc tct gac tgg aca ttt tta cac tgc ttg
 912

Lys Thr Ala Lys Val Ala Ala Ser Asp Trp Thr Phe Leu His Cys Leu
 290 295 300

ccc aga aag cca gaa gaa gtg gat gat gaa gtc ttt tat tct cct cga
 960

Pro Arg Lys Pro Glu Glu Val Asp Asp Glu Val Phe Tyr Ser Pro Arg
 305 310 315 320

tca cta gtg ttc cca gag gca gaa aac aga aag tgg aca atc atg gct
 1008

Ser Leu Val Phe Pro Glu Ala Glu Asn Arg Lys Trp Thr Ile Met Ala
 325 330 335

gtc atg gtg tcc ctg ctg aca gat tac tca cct cag ctc cag aag cct
 1056

Val Met Val Ser Leu Leu Thr Asp Tyr Ser Pro Gln Leu Gln Lys Pro
 340 345 350

aaa ttt
 1062
 Lys Phe

<210> 2
 <211> 354
 <212> PRT
 <213> *Homo sapiens*

5

<400> 2

Met Leu Phe Asn Leu Arg Ile Leu Leu Asn Asn Ala Ala Phe Arg Asn
 1 5 10 15

Gly His Asn Phe Met Val Arg Asn Phe Arg Cys Gly Gln Pro Leu Gln
 20 25 30

Asn Lys Val Gln Leu Lys Gly Arg Asp Leu Leu Thr Leu Lys Asn Phe
 35 40 45

Thr Gly Glu Glu Ile Lys Tyr Met Leu Trp Leu Ser Ala Asp Leu Lys

10

ES 2 821 938 T3

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

ES 2 821 938 T3

Glu Lys Lys Lys Arg Leu Gln Ala Phe Gln Gly Tyr Gln Val Thr Met
275 280 285

Lys Thr Ala Lys Val Ala Ala Ser Asp Trp Thr Phe Leu His Cys Leu
290 295 300

Pro Arg Lys Pro Glu Glu Val Asp Asp Glu Val Phe Tyr Ser Pro Arg
305 310 315 320

Ser Leu Val Phe Pro Glu Ala Glu Asn Arg Lys Trp Thr Ile Met Ala
325 330 335

Val Met Val Ser Leu Leu Thr Asp Tyr Ser Pro Gln Leu Gln Lys Pro
340 345 350

Lys Phe

<210> 3
<211> 1068
<212> ADN
<213> Secuencia artificial

<220>
<223> hOTC modificada por ingeniería genética

<400> 3

atgctgttca acctgcgaat cctgctgaac aatgccgctt ttcggaacgg gcacaatttc
60

atggtgagga actttcgctg cggacagccc ctccagaaca aggtccagct gaagggcagg
120

gacctgctga ccctgaaaaa tttcacaggg gaggaaatca agtacatgct gtggctgtca
180

gccgatctga agttccggat caagcagaag ggcgaatatc tgcctctgct ccagggcaaa
240

agcctgggga tgatcttcga aaagcgcagt actcggacca gactgtcaac agagactgga
300

ttcgactgct tgggaggaca cccatgtttt ctgaccacac aggacattca tctgggagtg
360

aacgagtccc tgaccgacac agcacgcgct ctgagctcca tggctgatgc agtgcctggct
420

5

10

ES 2 821 938 T3

cgagtctaca aacagtctga cctggatacc ctggccaagg aagcttctat cccaatcatt
480

aatggcctga gtgacctgta tcaccccatc cagattctgg ccgattacct gaccctccag
540

gagcattatt ctagtctgaa agggctgaca ctgagctgga ttggggacgg aaacaatatac
600

ctgcaactcca ttatgatgag cgccgccaag tttggaatgc acctccagge tgcaacccca
660

aaaggctacg aacccgatgc ctccgtgaca aagctggcag aacagtatgc caaagagaac
720

ggcactaagc tgctgctgac caatgaccct ctggaggccg ctcacggagg caacgtgctg
780

atcactgata cctggattag tatgggacag gaggaagaga agaagaagcg gctccaggcc
840

ttccagggct accaggtgac aatgaaaact gctaaggctg cagccagcga ctggaccttt
900

ctgcattgcc tgcccagaaa gcctgaagag gtggacgatg aggtcttcta ctcaccaga
960

agcctggtgt ttctgaagc tgagaatagg aagtggacaa tcatggcagt gatggtcagc
1020

ctgctgactg attattcccc tcagctccag aaaccaaagt tctgataa
1068

<210> 4

<211> 1068

5 <212> ADN

<213> Secuencia artificial

<220>

10 <223> secuencia codificante de hOTC modificada por ingeniería genética

<400> 4

atgctgttca acctgcgaat cctgctgaac aacgccgctt ttcggaacgg gcacaacttt
60

atggtgagga actttcgctg cggacagccc ctccagaata aggtccagct gaagggcagg
120

gacctgctga ccctgaaaaa tttcacaggg gaggaaatca agtatatgct gtggctgtca
180

ES 2 821 938 T3

gctgatctga agttccggat caagcagaag ggcgaatata tgcctctgct ccagggcaaa
240

agcctgggga tgatcttcga aaagcgcagt actcggacca gactgtcaac cgagactgga
300

ttcgtctctgc tgggaggaca cccttgTTTT ctgaccactc aggacattca cctgggagtg
360

aacgagtccc tgaccgacac tgctcgcgtc ctgagctcta tggccgacgc tgtgctagct
420

cgagtctaca aacagtccga cctggatacc ctggccaagg aagcttctat cccaattatt
480

aacggcctgt cagacctgta tcaccccatc cagattctgg ccgattacct gaccctccag
540

gagcactatt ctagtctgaa agggctgaca ctgagttgga ttggggacgga aaacaatata
600

ctgcactcta ttatgatgac agccgccaag tttggaatgc acctccaggc tgcaacccca
660

aaaggctacg aacccgatgc ctcaagtaca aagctggctg aacagtacgc caagagaaac
720

ggcactaagc tgctgctgac caacgaccct ctggaggccg ctcacggagg caacgtgctg
780

atcaccgata cctggattag tatgggacag gaggaagaga agaagaagcg gctccaggcc
840

ttccagggtc accaggtgac aatgaaaacc gctaaggctc cagccagcga ttggaccttt
900

ctgcactgcc tgcccagaaa gcccgaaagag gtggacgacg aggtcttcta ctctcccaga
960

agcctggtgt ttcccgaagc tgagaatagg aagtggacaa ttatggcagt gatggtcagc
1020

ctgctgactg attattcacc tcagctccag aaaccaaagt tctgataa
1068

<210> 5
<211> 1068
<212> ADN
<213> Secuencia artificial

5

<220>
<223> hOTC modificada por ingeniería genética

10

<400> 5

ES 2 821 938 T3

atgctgttca acctgcgaat cctgctgaac aacgccgctt ttcggaacgg gcacaacttt
60

atggtgagga actttcgtcg cggacagccc ctccagaata aggtccagct gaagggcagg
120

gacctgctga ccctgaaaaa tttcacaggg gaggaaatca agtatatgct gtggctgtca
180

gctgatctga agttccggat caagcagaag ggcgaatata tgcctctgct ccagggcaaa
240

agcctgggga tgatcttcga aaagcgcagt actcggacca gactgtcaac cgagactgga
300

ttcgtctctg tgggaggaca cccttgtttt ctgaccactc aggacattca cctgggagtg
360

aacgagtccc tgaccgacac tgctcgcgtc ctgagctcta tggccgacgc tgtgctggct
420

cgagtctaca aacagtccga cctggatacc ctggccaagg aagcttctat cccaattatt
480

aacggcctgt cagacctgta tcaccccatc cagattctgg ccgattacct gaccctccag
540

gagcactatt ctagtctgaa agggctgaca ctgagttgga ttggggacgg aaacaatata
600

ctgcactcta ttatgatgtc agccgccaaag tttggaatgc acctccaggc tgcaacccca
660

aaaggctacg aaccgatgc ctcaagtaca aagctggctg aacagtacgc caaagagaac
720

ggcactaagc tgctgctgac caacgaccct ctggaggccg ctcacggagg caacgtgctg
780

atcacggata cctggattag tatgggacag gaggaagaga agaagaagcg gctccaggcc
840

ttccagggct accaggtgac aatgaaaacc gctaaggctg cagccagcga ttggaccttt
900

ctgcactgcc tgcccagaaa gcccgaaagag gtggacgacg aggtcttcta ctctcccaga
960

agcctgggtg ttcccgaagc tgagaatagg aagtggacaa ttatggcagt gatggctcagc
1020

ctgctgactg attattcacc tcagctccag aaaccaaagt tctgataa
1068

<210> 6
<211> 5195
<212> ADN

ES 2 821 938 T3

<213> Secuencia artificial

<220>
<223> Plásmido pscAAVTBGhOTCLW

5

<220>
<221> región de repetición
<222> (5)..(109)
<223> ITR

10

<220>
<221> señal TATA
<222> (851)..(854)

15

<220>
<221> CDS
<222> (976)..(2037)
<223> hOTCco-LW4

20

<220>
<221> poliA
<222> (2046)..(2182)
<223> ITR, ubicada en complemento

25

<220>
<221> región de repetición
<222> (2211)..(2378)
<223> ITR, ubicada en complemento

30

<220>
<221> misc_feature
<222> (4172)..(4760)
<223> TBG/promotor

35

<400> 6

taggctgcgc gctcgctcgc tcaactgaggc cgcccgggca aagcccgggc gtcgggcgac
60

ctttggctgc cggcctcag tgagcgagcg agcgcgcaga gagggagtgt agccatgctc
120

taggaagatc aattcaattc acgcgtggta cctagaacta tagctagaat tcgcccttaa
180

gctagcaggt taatttttaa aaagcagtca aaagtccaag tggcccttgg cagcatttac
240

ES 2 821 938 T3

tctctctggt tgctctgggt aataatctca ggagcacaaa cattccagat ccaggttaat
300

ttttaaaaag cagtcaaaaag tccaagtggc ccttggcagc atttactctc tctgtttgct
360

ctggttaata atctcaggag cacaaacatt ccagatccgg cgcgccaggg ctggaagcta
420

cotttgacat catttctctt gcgaatgcat gtataatttc tacagaacct attagaaagg
480

atcaccagc ctctgctttt gtacaacttt ccottaaaaa actgcccaatt ccaactgctgt
540

ttggccaat agtgagaact ttttctgct gcctcttggg gcttttgctt atggccccta
600

ttctgcctgc tgaagacact cttgccagca tggacttaaa ccctccagc tctgacaatc
660

ctctttctct tttgttttac atgaagggtc tggcagccaa agcaatcact caaagttcaa
720

accttatcat tttttgcttt gttcctcttg gccttggttt tgtacatcag ctttgaaaat
780

accatcccag ggtaaatgct ggggtaatt tataactaag agtgctctag ttttgcaata
840

caggacatgc tataaaaatg gaaagatggt gctttctgag agacagcttt attgcggtag
900

tttatcacag ttaaattgct aacgcagtca gtgcttctga cacaacagtc tcgaacttaa
960

gctgcagccg ccacc atg ctg ttc aac ctg cga atc ctg ctg aac aac gcc
1011

Met Leu Phe Asn Leu Arg Ile Leu Leu Asn Asn Ala
1 5 10

gct ttt cgg aac ggg cac aac ttt atg gtg agg aac ttt cgc tgc gga
1059

Ala Phe Arg Asn Gly His Asn Phe Met Val Arg Asn Phe Arg Cys Gly
15 20 25

cag ccc ctc cag aat aag gtc cag ctg aag ggc agg gac ctg ctg acc
1107

Gln Pro Leu Gln Asn Lys Val Gln Leu Lys Gly Arg Asp Leu Leu Thr
30 35 40

ctg aaa aat ttc aca ggg gag gaa atc aag tat atg ctg tgg ctg tca
1155

Leu Lys Asn Phe Thr Gly Glu Glu Ile Lys Tyr Met Leu Trp Leu Ser

ES 2 821 938 T3

45 50 55 60

gct gat ctg aag ttc cgg atc aag cag aag ggc gaa tat ctg cct ctg
1203
Ala Asp Leu Lys Phe Arg Ile Lys Gln Lys Gly Glu Tyr Leu Pro Leu
65 70 75

ctc cag ggc aaa agc ctg ggg atg atc ttc gaa aag cgc agt act cgg
1251
Leu Gln Gly Lys Ser Leu Gly Met Ile Phe Glu Lys Arg Ser Thr Arg
80 85 90

acc aga ctg tca acc gag act gga ttc gct ctg ctg gga gga cac cct
1299
Thr Arg Leu Ser Thr Glu Thr Gly Phe Ala Leu Leu Gly Gly His Pro
95 100 105

tgt ttt ctg acc act cag gac att cac ctg gga gtg aac gag tcc ctg
1347
Cys Phe Leu Thr Thr Gln Asp Ile His Leu Gly Val Asn Glu Ser Leu
110 115 120

acc gac act gct cgc gtc ctg agc tct atg gcc gac gct gtg ctg gct
1395
Thr Asp Thr Ala Arg Val Leu Ser Ser Met Ala Asp Ala Val Leu Ala
125 130 135 140

cga gtc tac aaa cag tcc gac ctg gat acc ctg gcc aag gaa gct tct
1443
Arg Val Tyr Lys Gln Ser Asp Leu Asp Thr Leu Ala Lys Glu Ala Ser
145 150 155

atc cca att att aac ggc ctg tca gac ctg tat cac ccc atc cag att
1491
Ile Pro Ile Ile Asn Gly Leu Ser Asp Leu Tyr His Pro Ile Gln Ile
160 165 170

ctg gcc gat tac ctg acc ctc cag gag cac tat tct agt ctg aaa ggg
1539
Leu Ala Asp Tyr Leu Thr Leu Gln Glu His Tyr Ser Ser Leu Lys Gly
175 180 185

ctg aca ctg agt tgg att ggg gac gga aac aat atc ctg cac tct att
1587
Leu Thr Leu Ser Trp Ile Gly Asp Gly Asn Asn Ile Leu His Ser Ile
190 195 200

atg atg tca gcc gcc aag ttt gga atg cac ctc cag gct gca acc cca
1635
Met Met Ser Ala Ala Lys Phe Gly Met His Leu Gln Ala Ala Thr Pro
205 210 215 220

aaa ggc tac gaa ccc gat gcc tca gtg aca aag ctg gct gaa cag tac
1683

ES 2 821 938 T3

Lys Gly Tyr Glu Pro Asp Ala Ser Val Thr Lys Leu Ala Glu Gln Tyr
 225 230 235

gcc aaa gag aac ggc act aag ctg ctg ctg acc aac gac cct ctg gag
 1731

Ala Lys Glu Asn Gly Thr Lys Leu Leu Leu Thr Asn Asp Pro Leu Glu
 240 245 250

gcc gct cac gga ggc aac gtg ctg atc acc gat acc tgg att agt atg
 1779

Ala Ala His Gly Gly Asn Val Leu Ile Thr Asp Thr Trp Ile Ser Met
 255 260 265

gga cag gag gaa gag aag aag aag cgg ctc cag gcc ttc cag ggc tac
 1827

Gly Gln Glu Glu Glu Lys Lys Lys Arg Leu Gln Ala Phe Gln Gly Tyr
 270 275 280

cag gtg aca atg aaa acc gct aag gtc gca gcc agc gat tgg acc ttt
 1875

Gln Val Thr Met Lys Thr Ala Lys Val Ala Ala Ser Asp Trp Thr Phe
 285 290 295 300

ctg cac tgc ctg ccc aga aag ccc gaa gag gtg gac gac gag gtc ttc
 1923

Leu His Cys Leu Pro Arg Lys Pro Glu Glu Val Asp Asp Glu Val Phe
 305 310 315

tac tct ccc aga agc ctg gtg ttt ccc gaa gct gag aat agg aag tgg
 1971

Tyr Ser Pro Arg Ser Leu Val Phe Pro Glu Ala Glu Asn Arg Lys Trp
 320 325 330

aca att atg gca gtg atg gtc agc ctg ctg act gat tat tca cct cag
 2019

Thr Ile Met Ala Val Met Val Ser Leu Leu Thr Asp Tyr Ser Pro Gln
 335 340 345

ctc cag aaa cca aag ttc tgataagcgg ccgctatttg tgaaatttgg
 2067

Leu Gln Lys Pro Lys Phe
 350

gatgctattg ctttatttgg aaccattata agctgcaata aacaagttaa caacaacaat
 2127

tgcattcatt ttatgtttca ggttcagggg gaggtgtggg aggtttttta ggcatcgata
 2187

aggatcttcc tagagcatgg ctacgtagat aagtagcatg gcgggttaat cattaactac
 2247

aaggaacccc tagtgatgga gttggcact cectctctgc gcgctcgcgc gctcactgag
 2307

ES 2 821 938 T3

gccgggacgac caaaggctgc ccgacgcccg ggctttgccc gggcggcctc agtgagcag
2367

cgagcgcgca gccttaatta acctaattca ctggccgtcg ttttacaacg tcgtgactgg
2427

gaaaaccctg gcggttaccca acttaatcgc cttgcagcac atcccccttt cgcagctgg
2487

cgtaatagcg aagaggcccg caccgatcgc ccttccaac agttgcgag cctgaatggc
2547

gaatgggacg cgccctgtag cggcgcatta agcgcggcgg gtgtgggtgt tacgcgcagc
2607

gtgaccgcta cacttgccag cgccctagcg cccgctcctt togetttctt cccttccttt
2667

ctcgccacgt togcggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc
2727

cgatttagtg ctttacggca cctcgacccc aaaaaacttg attagggatga tggttcacgt
2787

agtgggcat cgccctgata gacggttttt cgcccttga cgttggagtc cacgttcttt
2847

aatagtggac tcttgttcca aactggaaca aactcaacc ctatctcggc ctattctttt
2907

gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaacaa
2967

aaatttaacg cgaattttta caaatatta acgcttaca tttaggtggc acttttcggg
3027

gaaatgtgcg cggaaaccct atttgtttat ttttctaat acattcaaat atgtatccgc
3087

tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag agtatgagta
3147

ttcaacattt cgtgtcgc cttattcctt tttttgcggc attttgcctt cctgtttttg
3207

ctcaccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtg
3267

gttacatcga actggatctc aacagcggta agatccttga gagttttcgc cccgaagaac
3327

gttttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta tcccgattg
3387

ES 2 821 938 T3

acgccgggca agagcaactc ggtcgccgca tacactattc tcagaatgac ttggttgagt
3447

actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg
3507

ctgccataac catgagtgat aacactgcfg ccaacttact tctgacaacg atcggaggac
3567

cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt
3627

gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgtag
3687

caatggcaac aacgttgcfg aaactattaa ctggcgaact acttactcta gcttcccggc
3747

aacaattaat agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc
3807

ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta
3867

tcattgcagc actggggcca gatggttaagc cctcccgtat cgtagttatc tacacgacgg
3927

ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga
3987

ttaagcattg gtaactgtca gaccaagttt actcatatat actttagatt gatttaaaac
4047

ttcattttta atttaaaagg atctaggtga agatcctttt tgataatctc atgacaaaa
4107

tcccttaacg tgagttttcg ttccactgag cgtcagacc cgtagaaaag atcaaaggat
4167

cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc
4227

taccagcggg ggtttgtttg ccggatcaag agctaccaac tctttttccg aaggtaactg
4287

gcttcagcag agcgcagata ccaaatactg ttcttctagt gtagccgtag ttaggccacc
4347

acttcaagaa ctctgtagca ccgctacat acctcgtctt gctaatoctg ttaccagtgg
4407

ctgctgccag tggcgataag tcgtgtctta ccgggttga ctcaagacga tagttaccgg
4467

ES 2 821 938 T3

ataagggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa
4527

cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg
4587

aagggagaaa ggcggacagc tatccggtaa gcggcagggt cggaacagga gagcgcacga
4647

gggagcttcc agggggaaac gcctggtatc tttatagtcc tgtcggggtt cgccacctct
4707

gacttgagcg tcgatttttg tgatgctcgt cagggggcg gagcctatgg aaaaacgcca
4767

gcaacgcggc ctttttacgg ttcttgacct tttgctggcc ttttgctcac atgttctttc
4827

ctgcgttatc cctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg
4887

ctcgccgcag ccgaacgacc gagcgcagcg agtcagtgag cgaggaagcg gaagagcgcc
4947

caatacgcaa accgcctctc cccgcgcggt ggccgattca ttaatgcagc tggcacgaca
5007

ggtttccga ctggaaagcg ggcagtgagc gcaacgcaat taatgtgagt tagctcactc
5067

attaggcacc ccaggcttta cactttatgc ttccggctcg tatgttgtgt ggaattgtga
5127

gcggataaca atttcacaca ggaaacagct atgacatga ttacgccaga ttaattaag
5187

gccttaat
5195

<210> 7
<211> 354
<212> PRT
<213> Secuencia artificial

5

<220>
<223> hOTC modificada por ingeniería genética

10

<400> 7

Met Leu Phe Asn Leu Arg Ile Leu Leu Asn Asn Ala Ala Phe Arg Asn
1 5 10 15

ES 2 821 938 T3

Gly His Asn Phe Met Val Arg Asn Phe Arg Cys Gly Gln Pro Leu Gln
 20 25 30

Asn Lys Val Gln Leu Lys Gly Arg Asp Leu Leu Thr Leu Lys Asn Phe
 35 40 45

Thr Gly Glu Glu Ile Lys Tyr Met Leu Trp Leu Ser Ala Asp Leu Lys
 50 55 60

Phe Arg Ile Lys Gln Lys Gly Glu Tyr Leu Pro Leu Leu Gln Gly Lys
 65 70 75 80

Ser Leu Gly Met Ile Phe Glu Lys Arg Ser Thr Arg Thr Arg Leu Ser
 85 90 95

Thr Glu Thr Gly Phe Ala Leu Leu Gly Gly His Pro Cys Phe Leu Thr
 100 105 110

Thr Gln Asp Ile His Leu Gly Val Asn Glu Ser Leu Thr Asp Thr Ala
 115 120 125

Arg Val Leu Ser Ser Met Ala Asp Ala Val Leu Ala Arg Val Tyr Lys
 130 135 140

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

Asn Gly Leu Ser Asp Leu Tyr His Pro Ile Gln Ile Leu Ala Asp Tyr
 165 170 175

Leu Thr Leu Gln Glu His Tyr Ser Ser Leu Lys Gly Leu Thr Leu Ser
 180 185 190

Trp Ile Gly Asp Gly Asn Asn Ile Leu His Ser Ile Met Met Ser Ala
 195 200 205

Ala Lys Phe Gly Met His Leu Gln Ala Ala Thr Pro Lys Gly Tyr Glu
 210 215 220

Pro Asp Ala Ser Val Thr Lys Leu Ala Glu Gln Tyr Ala Lys Glu Asn

ES 2 821 938 T3

gccgacctga agttcagaat caagcagaag ggcgagtacc tgcccctgct gcagggcaag
240

agcctgggca tgatcttcga gaagagaagc accagaacca gactgagcac cgagaccggc
300

ctggccctgc tgggcggccca cccctgcttc ctgaccaccc aggacatcca cctgggctgt
360

aacgagagcc tgaccgacac cgccagagtg ctgagcagca tggccgacgc cgtgctggcc
420

agagtgtaca agcagagcga cctggacacc ctggccaagg aggccagcat ccccatcatc
480

aacggcctga gcgacctgta ccaocccatc cagatcctgg ccgactacct gaccctgcag
540

gagcactaca gcagcctgaa gggcctgacc ctgagctgga tcggcgacgg caacaacatc
600

ctgcacagca tcatgatgag cgccgccaag ttcggcatgc acctgcaggc cgccaccccc
660

aagggctacg agcccgacgc cagcgtgacc aagctggccg agcagtacgc caaggagaac
720

ggcaccaagc tgctgctgac caacgacccc ctggaggccg cccacggcgg caacgtgctg
780

atcaccgaca cctggatcag catgggcccag gaggaggaga agaagaagag actgcaggcc
840

ttccagggct accaggtgac catgaagacc gcccaaggtg cgcgacgca ctggaccttc
900

ctgcactgcc tgcccagaaa gcccgaggag gtggacgacg aggtgttcta cagccccaga
960

agcctggtgt tccccgaggc cgagaacaga aagtggacca tcatggccgt gatggtgagc
1020

ctgctgaccg actacagccc ccagctgcag aagcccaagt tctga
1065

<210> 9

<211> 1065

5 <212> ADN

<213> Secuencia artificial

<220>

10 <223> hOTC modificada por ingeniería genética

<400> 9

ES 2 821 938 T3

atgctgttca acctgcgcat cctgctgaac aacgccgcct tccgcaacgg ccacaacttc
60

atgggtgcgca acttccgctg cggccagccc ctgcagaaca aggtgcagct gaagggccgc
120

gacctgctga ccctgaagaa cttcaccggc gaggagatca agtacatgct gtggctgagc
180

gccgacctga agttccgcat caagcagaag ggcgagtacc tgcccctgct gcagggcaag
240

agcctgggca tgatcttcga gaagcgcagc acccgcacc gcctgagcac cgagaccggc
300

ctggccctgc tgggcggcca cccctgcttc ctgaccacc aggacatcca cctgggcgtg
360

aacgagagcc tgaccgacac cgcccgcgtg ctgagcagca tggccgacgc cgtgctggcc
420

cgcggtgata agcagagcga cctggacacc ctggccaag aggccagcat ccccatcatc
480

aacggcctga gcgacctgta ccaccccatc cagatcctgg ccgactacct gaccctgcag
540

gagcactaca gcagcctgaa gggcctgacc ctgagctgga tcggcgacgg caacaacatc
600

ctgcacagca tcatgatgag cgccgccaag ttcggcatgc acctgcaggc cgccaccccc
660

aagggctacg agcccgacgc cagcgtgacc aagctggccg agcagtacgc caaggagaac
720

ggcaccaagc tgctgctgac caacgacccc ctggaggccg cccacggcgg caacgtgctg
780

atcaccgaca cctggatcag catgggccag gaggaggaga agaagaagcg cctgcaggcc
840

ttccagggtc accaggtgac catgaagacc gcccaaggtg ccgccagcga ctggaccttc
900

ctgcaactgcc tgccccgcaa gcccgaggag gtggacgacg aggtgttcta cagccccgc
960

agcctggtgt tccccgaggc cgagaaccgc aagtggacca tcatggccgt gatggtgagc
1020

ctgctgaccg actacagccc ccagctgcag aagoccaaagt tctga
1065

<210> 10
<211> 1068
<212> RNA

ES 2 821 938 T3

<213> Secuencia artificial

<220>

<223> secuencia de ARN de hOTC modificada por ingeniería genética

5

<400> 10

augcuguuca accugcgaau ccugcugaac aacgccgcuu uucggaacgg gcacaacuuu
60

auggugagga acuuucgug cggacagccc succagaaua agguccagcu gaagggcagg
120

gaccugcuga cccugaaaaa uuucacaggg gaggaaaauca aguauaugcu guggcuguca
180

gcugaucuga aguuccggau caagcagaag ggcgaauauc ugccucugcu ccagggcaaa
240

agccugggga ugaucuucga aaagcgcagu acucggacca gacugucaac cgagacugga
300

uucgcucugc ugggaggaca cccuuguuuu cugaccacuc aggacauuca ccugggagug
360

aacgaguccc ugaccgacac ugcucgugc cugagcucua uggccgacgc ugugcuagcu
420

cgagucuaca aacaguccga ccuggauacc cuggccaagg aagcuucuau cccaauuau
480

aacggccugu cagaccugua ucacccauc cagauucugg ccgauuaccu gaccuccag
540

gagcacuauu cuagucugaa agggcugaca cugaguugga uuggggacgg aaacaauauc
600

cugcacucua uuauauguc agccgccaag uuuggaaugc accuccaggc ugcaaccca
660

aaaggcuacg aaccggaugc cucagugaca aagcuggcug aacaguacgc caaagagaac
720

ggcacuaagc ugcugcugac caacgaccu cuggaggccg cucacggagg caacgugcug
780

aucaccgaua ccuggauuag uauuggacag gaggaagaga agaagaagcg gcuccaggcc
840

ES 2 821 938 T3

uuccagggcu accaggugac aaugaaaacc gcuaaggucg cagccagcga uuggaccuuu
900

cugcacugcc ugcccagaaa gcccgaagag guggacgacg agguucuua cucuccaga
960

agccuggugu uucccgaagc ugagaauagg aaguggacaa uuauggcagu gauggucagc
1020

cugcugacug auuauucacc ucagcuccag aaaccaaagu ucugauaa
1068

<210> 11

<211> 1068

<212> RNA

<213> Secuencia artificial

<220>

<223> ARN de hOTC modificada por ingeniería genética

<400> 11

augcuguuca accugcgaau ccugcugaac aacgccgcuu uucggaacgg gcacaacuuu
60

auggugagga acuuucgcug cggacagccc cuccagaaua agguccagcu gaagggcagg
120

gaccugcuga cccugaaaaa uuucacaggg gaggaaauca aguauaugcu guggcuguca
180

gcugaucuga aguuccggau caagcagaag ggcgaauauc ugccucugcu ccagggcaaa
240

agccugggga ugaucuucga aaagcgcagu acucggacca gacugucaac cgagacugga
300

uucgcucugc ugggaggaca cccuuguuuu cugaccacuc aggacauuca ccugggagug
360

aacgaguccc ugaccgacac ugcucgcguc cugagcucua uggccgacgc ugugcuggcu
420

cgagucuaca aacaguccga ccuggauacc cuggccaagg aagcuucuau cccaauuuu
480

aacggccugu cagaccugua ucaccccuauc cagauucugg ccgauuaccu gaccuccag
540

gagcacuauu cuagucugaa agggcugaca cugaguugga uuggggacgg aaacaauauc
600

cugcacucua uuaugaugc agccgccaag uuuggaaugc accuccaggc ugcaaccca
660

5

10

ES 2 821 938 T3

aaaggcuacg aacccgaugc cucagugaca aagcuggcug aacaguacgc caaagagaac
720

ggcacuaagc ugcugcugac caacgacccu cuggaggccg cucacggagg caacgugcug
780

aucaccgaua ccuggauuag uaugggacag gaggaagaga agaagaagcg gcuccaggcc
840

uuccagggcu accaggugac aaugaaaacc gcuaaggucg cagccagcga uuggaccuuu
900

cugcacugcc ugcccagaaa gcccgaagag guggacgacg aggucuucua cucuccaga
960

agccuggugu uccccgaagc ugagaauagg aaguggacaa uuauggcagu gauggucagc
1020

cugcugacug auuauucacc ucagcuccag aaaccaagu ucugauaa
1068

<210> 12

<211> 1065

<212> RNA

<213> Secuencia artificial

5

<220>

<223> hOTC modificada por ingeniería genética

10

<400> 12

augcuguuca accugcgcau ccugcugaac aacgccgccu uccgcaacgg ccacaacuuc
60

auggugcgca acuuccgcug cggccagccc cugcagaaca aggugcagcu gaagggccgc
120

gaccugcuga ccugaagaa cuucaccggc gaggagauca aguacaugcu guggcugagc
180

gccgaccuga aguuccgcau caagcagaag ggcgaguacc ugccccugcu gcagggcaag
240

agccugggca ugaucuucga gaagcgcagc acccgacccc gccugagcac cgagaccggc
300

cuggcccugc ugggcggcca ccccugcuuc cugaccaccc aggacaacca ccugggcgug
360

aacgagagcc ugaccgacac cgcccgcgug cugagcagca uggccgacgc cgugcuggcc
420

ES 2 821 938 T3

cgcguguaca agcagagcga ccuggacacc cuggccaagg aggccagcau ccccaucauc
480

aacggccuga ggcaccugua ccaccccauc cagauccugg ccgacuaccu gaccucugcag
540

gagcacuaca gcagccugaa gggccugacc cugagcugga ucggcgacgg caacaacauc
600

cugcacagca ucaugaugag cgccgccaaag uucggcaugc accucgaggc cgccaccccc
660

aagggcuacg agcccgacgc cagcgugacc aagcuggccg agcaguacgc caaggagaac
720

ggaccaagc ugucugugac caacgacccc cuggaggccg cccacggcgg caacgugcug
780

aucaccgaca ccuggaucag caugggccag gaggaggaga agaagaagcg ccugcaggcc
840

uuccagggcu accaggugac caugaagacc gccaaggugg ccgccagcga cuggaccuuc
900

cugcacugcc ugccccgcaa gcccgaggag guggacgacg agguguucua cagcccccg
960

agccuggugu uccccgaggc cgagaaccgc aaguggacca ucauggccgu gauggugagc
1020

cugcugaccg acuacagccc ccagcugcag aagcccaagu ucuga
1065

<210> 13

<211> 1065

5 <212> RNA

<213> Secuencia artificial

<220>

10 <223> hOTC modificada por ingeniería genética

<400> 13

augcuguuca accugagaau ccugcugaac aacgcccgcuc ucagaaacgg ccacaacuuc
60

auggugagaa acuucagaug cggccagccc cugcagaaca aggugcagcu gaagggcaga
120

gaccugcuga ccugaagaa cuucaccggc gaggagauca aguacaugcu guggcugagc
180

ES 2 821 938 T3

gccgaccuga aguucagaau caagcagaag ggcgaguacc ugccccugcu gcagggcaag
240

agccugggca ugaucuucga gaagagaagc accagaacca gacugagcac cgagaccggc
300

cuggccccugc ugggcggccca cccugcuuc cugaccacc aggacaacca ccugggugug
360

aacgagagcc ugaccgacac cgccagagug cugagcagca uggccgacgc cgugcuggcc
420

agaguguaca agcagagcga ccuggacacc cuggccaagg aggccagcau ccccaucauc
480

aacggccuga gcgaccugua ccacccauc cagauccugg ccgacuaccu gaccucugag
540

gagcacuaca gcagccugaa gggccugacc cugagcugga ucggcgacgg caacaacauc
600

cugcacagca ucaugaugag cgccgccaag uucggcaugc accugcaggc cgccaccccc
660

aagggcuacg agcccgacgc cagcgugacc aagcuggccg agcaguacgc caaggagaac
720

ggcaccaagc ugcugcugac caacgacccc cuggaggccg cccacggcgg caacgugcug
780

aucaccgaca ccuggaucag caugggccag gaggaggaga agaagaagag acugcaggcc
840

uuccagggcu accaggugac caugaagacc gccaaaggugg ccgccagcga cuggaccuuc
900

cugcacugcc ugcccagaaa gcccgaggag guggacgacg agguguucua cagccccaga
960

agccuggugu uccccgaggc cgagaacaga aaguggacca ucauggccgu gauggugagc
1020

cugcugaccg acuacagccc ccagcugcag aagcccaagu ucuga
1065

<210> 14
<211> 21
5 <212> ADN
<213> Secuencia artificial

<220>
10 <223> cebador directo de PCR

<400> 14

aaactgccaa ttccactgct g
21

ES 2 821 938 T3

5 <210> 15
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador inverso de PCR

10 <400> 15

ccataggcaa aagcaccaag a
21

15 <210> 16
<211> 29
<212> ADN
<213> Secuencia artificial

<220>
20 <223> Sonda

<400> 16

ttggccaat agtgagaact tttcctgc
29

REIVINDICACIONES

1. Un vector vírico recombinante que comprende una secuencia de ácido nucleico que codifica la proteína ornitina transcarbamilasa humana (hOTC) y secuencias de control de la expresión que dirigen la expresión de hOTC en una célula hepática, en donde la secuencia de ácido nucleico de hOTC es menos de 80 % idéntica a la secuencia de hOTC de tipo silvestre de la SEQ ID NO: 1 que codifica la proteína hOTC madura o de longitud completa y expresa una hOTC funcional, en donde dicha secuencia de ácido nucleico de hOTC se selecciona de la secuencia de ácido nucleico que comprende la SEQ ID NO: 5 o una secuencia de ácido nucleico al menos de 96 % a aproximadamente 99 % idéntica a la misma.
2. El vector vírico recombinante según la reivindicación 1, en donde la secuencia de ácido nucleico de hOTC tiene la secuencia de la SEQ ID NO: 5.
3. El vector vírico recombinante según la reivindicación 1, en donde la secuencia de ácido nucleico de hOTC tiene la secuencia de la SEQ ID NO: 4 o de la SEQ ID NO: 3.
4. El vector vírico recombinante según una cualquiera de las reivindicaciones 1 a 3, en donde la hOTC es una OTC química que comprende una secuencia de tránsito heteróloga que sustituye la secuencia de tránsito nativa de la SEQ ID NO: 5.
5. El vector vírico recombinante según una cualquiera de las reivindicaciones 1 a 4, en donde el vector vírico se selecciona de un vector de virus adenoasociado (AAV), un vector adenovírico y un vector lentivírico.
6. El vector vírico recombinante según una cualquiera de las reivindicaciones 1 a 5, en donde las secuencias de control de la expresión comprenden un promotor específico del hígado, opcionalmente en donde el promotor específico del hígado es un promotor de globulina de unión a tiroxina (TBG).
7. El vector vírico recombinante según una cualquiera de las reivindicaciones 1 a 6, donde el casete de expresión comprende además uno o más de un intrón, una secuencia de Kozak, una poli A y elementos reguladores postranscripcionales.
8. El vector vírico recombinante de la reivindicación 1, en donde el vector vírico recombinante es un vector de AAV recombinante que comprende una cápside de AAV que tiene empaquetada en la misma una secuencia de ácido nucleico que comprende al menos una secuencia de ITR y el ácido nucleico que codifica hOTC, opcionalmente en donde la cápside de AAV se selecciona de AAV8, AAV9 y/o AAVrh10.
9. Un virus adenoasociado recombinante (rAAV) que tiene una cápside de AAV y empaquetado en el mismo un casete de expresión que comprende al menos una secuencia de repetición terminal invertida (ITR) de AAV, una secuencia de ácido nucleico que codifica al menos la ornitina transcarbamilasa humana madura (hOTC) y secuencias de control de la expresión que dirigen la expresión de la hOTC en una célula hepática, comprendiendo dichas secuencias de control de la expresión un promotor específico del hígado, en donde la secuencia de ácido nucleico de hOTC es menos del 80 % idéntica a la secuencia de hOTC de tipo silvestre de la SEQ ID NO: 1 que codifica la proteína hOTC madura y comprende al menos la secuencia de ácido nucleico que codifica la proteína hOTC madura de la SEQ ID NO: 5 o una secuencia de ácido nucleico al menos de 96 % a 99,9 % idéntica a la misma.
10. El rAAV según la reivindicación 9, en donde la secuencia de ácido nucleico de hOTC tiene la secuencia de la SEQ ID NO: 5.
11. El rAAV según una cualquiera de las reivindicaciones 8 a 10, en donde la cápside de AAV se selecciona de AAV8, AAV9 o AAVrh10.
12. El rAAV según la reivindicación 11, en donde el casete de expresión comprende además una secuencia de repetición terminal invertida (ITR) 5' de AAV y una secuencia de ITR 3' o en donde al menos una ITR de AAV comprende una ITR 5' en la que la secuencia D y el sitio de resolución terminal se suprimen o en donde las ITR 5' y 3' son de AAV2.
13. El rAAV según una cualquiera de las reivindicaciones 9 a 12, en donde el ácido nucleico que codifica hOTC tiene la secuencia codificante de la SEQ ID NO: 3 o, en donde el ácido nucleico que codifica hOTC tiene la secuencia codificante de la SEQ ID NO: 4.
14. Un vector vírico que comprende al menos un gen de hOTC que codifica una ornitina transcarbamilasa química que comprende al menos ornitina transcarbamilasa humana madura con una secuencia de tránsito heteróloga, en donde la secuencia de ácido nucleico que codifica la ornitina transcarbamilasa humana madura se selecciona de la de una secuencia de ácido nucleico de las SEQ ID NO: 3, 4 o 5.
15. El vector vírico según la reivindicación 14, en donde la secuencia codificante de hOTC es la secuencia de la SEQ

ID NO: 5.

5 16. Una composición farmacéutica que comprende un vehículo y una cantidad eficaz del vector según cualquiera de las reivindicaciones 1 a 8, el rAAV según una cualquiera de las reivindicaciones 9 a 13 y/o el vector vírico según la reivindicación 14 o 15.

17. Un vector vírico según una cualquiera de las reivindicaciones 1 a 8, el rAAV según una cualquiera de las reivindicaciones 9 a 13 y/o el vector según la reivindicación 14 o 15 para su uso en un método de tratamiento de la deficiencia de ornitina transcarbamilasa (OTCD) en un paciente humano.

FIG. 1A

```

1 atgctgttta atctgaggat cctgttaaac aatgcagctt ttagaaatgg tcacaacttc
61 atggttcgaa attttcggtg tggacaacca ctacaaaata aagtgcagct gaagggcctg
121 gaccttctca ctctaaaaaa ctttaccgga gaagaaatta aatatatgct atggctatca
181 gcagatctga aatttaggat aaaacagaaa ggagagtatt tgcctttatt gcaaggggaag
241 tccttaggca tgatTTTTga gaaaagaagt actcgaacaa gattgtctac agaaacaggc
301 tttgcacttc tgggaggaca tccttgtttt cttaccacac aagatattca tttgggtgtg
361 aatgaaagtc tcacggacac ggcccgtgta ttgtctagca tggcagatgc agtattggct
421 cgagtgtata aacaatcaga tttggacacc ctggctaaag aagcatccat cccaattatc
481 aatgggctgt cagatttcta ccatcctatc cagatcctgg ctgattacct cacgctccag
541 gaacactata gctctctgaa aggtcttacc ctcagctgga tcgggggatgg gaacaatata
601 ctgcactcca tcatgatgag cgcagcgaag ttcggaatgc accttcaggc agctactcca
661 aagggttatg agccggatgc tagtgtaacc aagttggcag agcagtatgc caaagagaat
721 ggtaccaagc tgttgctgac aaatgatcca ttggaagcag cgcattggagg caatgtatta
781 attacagaca cttggataag catgggacaa gaagaggaga agaaaaagcg gctccaggct
841 ttccaaggtt accaggttac aatgaagact gctaaagttg ctgcctctga ctggacattt
901 ttacactgct tgcccagaaa gccagaagaa gtggatgatg aagtctttta ttctctcga
961 tcaactagtgt tcccagaggc agaaaaacaga aagtggacaa tcatggctgt catggtgtcc
1021 ctgctgacag attactcacc tcagctccag aagcctaaat tt

```

FIG. 1B

Met Leu Phe Asn Leu Arg Ile Leu Leu Asn Asn Ala Ala Phe Arg Asn
 1 5 10 15

Gly His Asn Phe Met Val Arg Asn Phe Arg Cys Gly Gln Pro Leu Gln
 20 25 30

Asn Lys Val Gln Leu Lys Gly Arg Asp Leu Leu Thr Leu Lys Asn Phe
 35 40 45

Thr Gly Glu Glu Ile Lys Tyr Met Leu Trp Leu Ser Ala Asp Leu Lys
 50 55 60

Phe Arg Ile Lys Gln Lys Gly Glu Tyr Leu Pro Leu Leu Gln Gly Lys
 65 70 75 80

Ser Leu Gly Met Ile Phe Glu Lys Arg Ser Thr Arg Thr Arg Leu Ser
 85 90 95

Thr Glu Thr Gly Phe Ala Leu Leu Gly Gly His Pro Cys Phe Leu Thr
 100 105 110

Thr Gln Asp Ile His Leu Gly Val Asn Glu Ser Leu Thr Asp Thr Ala
 115 120 125

Arg Val Leu Ser Ser Met Ala Asp Ala Val Leu Ala Arg Val Tyr Lys
 130 135 140

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

Asn Gly Leu Ser Asp Leu Tyr His Pro Ile Gln Ile Leu Ala Asp Tyr
 165 170 175

Leu Thr Leu Gln Glu His Tyr Ser Ser Leu Lys Gly Leu Thr Leu Ser
 180 185 190

FIG. 1C

Trp Ile Gly Asp Gly Asn Asn Ile Leu His Ser Ile Met Met Ser Ala
 195 200 205

Ala Lys Phe Gly Met His Leu Gln Ala Ala Thr Pro Lys Gly Tyr Glu
 210 215 220

Pro Asp Ala Ser Val Thr Lys Leu Ala Glu Gln Tyr Ala Lys Glu Asn
 225 230 235 240

Gly Thr Lys Leu Leu Leu Thr Asn Asp Pro Leu Glu Ala Ala His Gly
 245 250 255

Gly Asn Val Leu Ile Thr Asp Thr Trp Ile Ser Met Gly Gln Glu Glu
 260 265 270

Glu Lys Lys Lys Arg Leu Gln Ala Phe Gln Gly Tyr Gln Val Thr Met
 275 280 285

Lys Thr Ala Lys Val Ala Ala Ser Asp Trp Thr Phe Leu His Cys Leu
 290 295 300

Pro Arg Lys Pro Glu Glu Val Asp Asp Glu Val Phe Tyr Ser Pro Arg
 305 310 315 320

Ser Leu Val Phe Pro Glu Ala Glu Asn Arg Lys Trp Thr Ile Met Ala
 325 330 335

Val Met Val Ser Leu Leu Thr Asp Tyr Ser Pro Gln Leu Gln Lys Pro
 340 345 350

Lys Phe

FIG. 2

1 atgctgttca acctgcgaat cctgctgaac aatgccgctt ttcggaacgg gcacaatttc
61 atggtgagga actttcgtg cggacagccc ctccagaaca aggtccagct gaagggcagg
121 gacctgctga ccctgaaaaa tttcacaggg gaggaaatca agtacatgct gtggctgtca
181 gccgatctga agttccggat caagcagaag ggcgaatata tgcctctgct ccagggcaaa
241 agcctgggga tgatcttga aaagcgcagt actcggacca gactgtcaac agagactgga
301 ttcgactgc tgggaggaca cccatgtttt ctgaccacac aggacattca tctgggagtg
361 aacgagtccc tgaccgacac agcacgcgtc ctgagctcca tggctgatgc agtgetgget
421 cgagtctaca aacagtctga cctggatacc ctggccaagg aagcttctat cccaatcatt
481 aatggcctga gtgacctgta tcacccatc cagattctgg ccgattacct gaccctccag
541 gagcattatt ctagtctgaa agggctgaca ctgagctgga ttggggacgg aaacaatata
601 ctgactcca ttatgatgag cgccgccaag tttggaatgc acctccaggc tgcaacccca
661 aaaggctacg aaccgatgc ctccgtgaca aagctggcag aacagtatgc caaagagaac
721 ggcactaagc tgctgctgac caatgaccct ctggaggccg ctcacggagg caacgtgctg
781 atcactgata cctggattag tatgggacag gaggaagaga agaagaagcg gctccaggcc
841 ttccagggct accaggtgac aatgaaaact gctaaggtcg cagccagcga ctggaccttt
901 ctgcattgcc tgcccagaaa gcctgaagag gtggacgatg aggtcttcta ctcaccaga
961 agcctggtgt ttctgaagc tgagaatagg aagtggacaa tcatggcagt gatggtcagc
1021 ctgctgactg attattcccc tcagctccag aaaccaaagt tctgataa

FIG. 3

```

1  ctgcagccgc caccatgctg ttcaacctgc gaatcctgct gaacaacgcc gcttttcgga
61  acgggcacaa ctttatgggt aggaactttc gctgcggaca gccctccag aataaggtcc
121 agctgaaggg cagggacctg ctgacctga aaaatttcac aggggaggaa atcaagtata
181 tgctgtggct gtcagctgat ctgaagtcc ggatcaagca gaaggcgaa tatctgcctc
241 tgctccaggg caaaagcctg gggatgatct tcgaaaagcg cagtactcgg accagactgt
301 caaccgagac tggattcgt ctgctgggag gacacccttg tttctgacc actcaggaca
361 ttcacctggg agtgaacgag tcctgaccg aactgctcg cgtcctgagc tctatggccg
421 acgctgtgct agctcgagtc taaaacagt ccgacctgga taccctggcc aaggaagctt
481 ctatcccaat tattaacggc ctgtcagacc tgtatcacc catccagatt ctggccgatt
541 acctgaccct ccaggagcac tattctagtc tgaaagggt gacactgagt tggattgggg
601 acggaaacaa tatcctgcac tctattatga tgtcagccgc caagtttgga atgcacctcc
661 aggctgcaac ccaaaaaggc tacgaaccgg atgcctcagt gacaaagctg gctgaacagt
721 acgccaaaga gaacggcact aagctgctgc tgaccaacga cctctggag gccgctcag
781 gaggcaacgt gctgatcacc gatacctgga ttagtatggg acaggaggaa gagaagaaga
841 agcggctcca ggcttccag ggctaccagg tgacaatgaa aaccgctaag gtcgcagcca
901 gcgattggac ctttctgcac tgctgcccga gaaagcccga agaggtggac gacgaggtct
961 tctactctcc cagaagcctg gtgttcccg aagctgagaa taggaagtgg acaattatgg
1021 cagtgatggt cagcctgctg actgattatt cacctcagct ccagaaacca aagttctgat
1081 aagcgccgc

```

FIG. 4

```

1  ctgcagccgc caccatgctg ttcaacctgc gaatcctgct gaacaacgcc gcttttcgga
61  acgggcacaa ctttatgggtg aggaactttc gctgcgggaca gcccctccag aataagggtcc
121 agctgaaggg cagggacctg ctgacctga aaaatttcac aggggaggaa atcaagtata
181 tgctgtggct gtcagctgat ctgaagtcc ggatcaagca gaagggcgaa tatctgctc
241 tgctccaggg caaaagcctg gggatgatct tcgaaaagcg cagtactcgg acccagactgt
301 caaccgagac tggattcgtc ctgctgggag gacacccttg tttctgacc actcaggaca
361 ttcacctggg agtgaacgag tccctgaccg aactgctcg cgtcctgagc tctatggccg
421 acgctgtgct ggctcgagtc tacaacagt ccgacctgga taccctggcc aaggaagctt
481 ctatcccaat tattaacggc ctgtcagacc tgtatcacc catccagatt ctggccgatt
541 acctgacctt ccaggagcac tattctagtc tgaaagggct gacctgagt tggattggg
601 acggaaacaa taccctgcac tctattatga tgtcagccgc caagtttgga atgcacctcc
661 aggctgcaac ccaaaaaggc tacgaaccgc atgcctcagt gacaaagctg gctgaacagt
721 acgccaaga gaacggcact aagctgctgc tgaccaacga ccctctggag gccgctcacg
781 gaggcaacgt gctgatcacc gatacctgga ttagtatggg acaggaggaa gagaagaaga
841 agcggctcca ggccttcag ggctaccagg tgacaatgaa aaccgctaag gtcgcagcca
901 gcgattggac ctttctgcac tgctgcccga gaaagcccga agagggtggac gacgaggtct
961 tctactctcc cagaagcctg gtgtttcccg aagctgagaa taggaagtgg acaattatgg
1021 cagtgatggt cagcctgctg actgattatt cacctcagct ccagaaacca aagttctgat
1081 aagcggccgc

```

FIG. 5A

```

hoTCLW5      -----ATGCTGTTCAACCTGAGAATCCTGCTGAACAACGCCGCTTCCAGAA 46
hoTCLW6      -----ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCTTCCGCA 46
hoTCco       -----ATGCTGTTCAACCTGCGAATCCTGCTGAACAATGCCGCTTTTCGGA 46
LW3          CTGCAGCCGCCACCATGCTGTTCAACCTGCGAATCCTGCTGAACAACGCCGCTTTTCGGA 60
LW4          CTGCAGCCGCCACCATGCTGTTCAACCTGCGAATCCTGCTGAACAACGCCGCTTTTCGGA 60
hoTts        -----ATGCTGTTTAACTCTGAGGATCCTGTTAAACAATGCAGCTTTTAGAA 46
              ***** ** *** * ***** * ***** ** ** * *
hoTCLW5      ACGGCCACAACCTTATGGTGAGAACTTCAAGATGCGGCCAGCCCCTGCAGAACAAGGTGC 106
hoTCLW6      ACGGCCACAACCTTATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAGGTGC 106
hoTCco       ACGGCCACAATTTTATGGTGAGGAACTTTTCGCTGCGGACAGCCCCTCCAGAACAAGGTCC 106
LW3          ACGGCCACAACCTTTATGGTGAGGAACTTTTCGCTGCGGACAGCCCCTCCAGAATAAGGTCC 120
LW4          ACGGCCACAACCTTTATGGTGAGGAACTTTTCGCTGCGGACAGCCCCTCCAGAATAAGGTCC 120
hoTts        ATGGTCAACAACCTTATGGTGAGAACTTTTCGCTGCGGACAGCCCCTCAAAAATAAGGTGC 106
              * ** ***** ** ***** * ** * * ** * ** * ** * ** * ** * ** *
hoTCLW5      AGCTGAAGGGCAGAGACCTGCTGACCTGAAGAACTTACCGGGCAGGAGATCAAGTACA 166
hoTCLW6      AGCTGAAGGGCCGCGACCTGCTGACCTGAAGAACTTACCGGGCAGGAGATCAAGTACA 166
hoTCco       AGCTGAAGGGCAGGGACCTGCTGACCTGAAAAATTTACAGGGGAGGAAATCAAGTACA 166
LW3          AGCTGAAGGGCAGGGACCTGCTGACCTGAAAAATTTACAGGGGAGGAAATCAAGTATA 180
LW4          AGCTGAAGGGCAGGGACCTGCTGACCTGAAAAATTTACAGGGGAGGAAATCAAGTATA 180
hoTts        AGCTGAAGGGCCGTGACCTTCTCACTCTAAAAAATTTACCGGAGAAGAAATTAATATA 166
              ***** * ***** ** ** * ** * ** * ** * ** * ** * ** *
hoTCLW5      TGCTGTGGCTGAGCGCCGACCTGAAGTTTCAAGATCAAGCAGAAGGGGCGAGTACCTGCC 226
hoTCLW6      TGCTGTGGCTGAGCGCCGACCTGAAGTTCCGATCAAGCAGAAGGGGCGAGTACCTGCC 226
hoTCco       TGCTGTGGCTGTAGCCGATCTGAAGTTCCGGATCAAGCAGAAGGGGCGAATATCTGCCTC 226
LW3          TGCTGTGGCTGTAGCTGATCTGAAGTTCCGGATCAAGCAGAAGGGGCGAATATCTGCCTC 240
LW4          TGCTGTGGCTGTAGCTGATCTGAAGTTCCGGATCAAGCAGAAGGGGCGAATATCTGCCTC 240
hoTts        TGCTATGGCTATCAGCAGATCTGAAATTTAGGATAAAACGAAAGGAGATTTGCCTT 226
              **** ***** ** ** ***** ** * ** * ** ***** ** ** * **
hoTCLW5      TGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGAGAAGCACCAGAACCAGACTGA 286
hoTCLW6      TGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCACC CGCACCCGCTGA 286
hoTCco       TGCTCCAGGGCAAAAGCCTGGGGATGATCTTCGAAAAGCGCAGTACTCGGACCAGACTGT 286
LW3          TGCTCCAGGGCAAAAGCCTGGGGATGATCTTCGAAAAGCGCAGTACTCGGACCAGACTGT 300
LW4          TGCTCCAGGGCAAAAGCCTGGGGATGATCTTCGAAAAGCGCAGTACTCGGACCAGACTGT 300
hoTts        TATTCAGGGGAAGTCTTAGGCATGATTTTTGAGAAAAGAAGTACTCGAACAAGATTGT 286
              * * ** ** * * * * ** ***** ** ** * * ** * ** * ** * **
hoTCLW5      GCACCGAGACCGCCTGGCCCTGCTGGGCGGCCACCCCTGCTTCCTGACCACCCAGGACA 346
hoTCLW6      GCACCGAGACCGCCTGGCCCTGCTGGGCGGCCACCCCTGCTTCCTGACCACCCAGGACA 346
hoTCco       CAACAGAGACTGGATTTCGACTGCTGGGAGGACACCCATGTTTTCTGACCACACAGGACA 346
LW3          CAACCGAGACTGGATTTCGCTCTGCTGGGAGGACACCCCTGTTTTCTGACCACTCAGGACA 360
LW4          CAACCGAGACTGGATTTCGCTCTGCTGGGAGGACACCCCTGTTTTCTGACCACTCAGGACA 360
hoTts        CTACAGAAAAGGCTTTGCACTTCTGGGAGGACATCCTTGTTTTCTTACCAACAAGATA 346
              ** ** ** * * * * ** ***** ** ** * ** * ** * ** ***** ** ** *
hoTCLW5      TCCACCTGGGCGTGAACGAGAGCCTGACCGACACCGCCAGAGTGCTGAGCAGCATGGCCG 406
hoTCLW6      TCCACCTGGGCGTGAACGAGAGCCTGACCGACACCGCCCGGTGCTGAGCAGCATGGCCG 406
hoTCco       TTCATCTGGGAGTGAACGAGTCCCTGACCGACACAGCAGCGTCTGAGCTCCATGGCTG 406
LW3          TTCACCTGGGAGTGAACGAGTCCCTGACCGACACTGCTCGCTCTGAGCTCTATGGCCG 420
LW4          TTCACCTGGGAGTGAACGAGTCCCTGACCGACACTGCTCGCTCTGAGCTCTATGGCCG 420
hoTts        TTCATTTGGGTGTGAATGAAAGTCTCACGGACACGGCCCGTATTGTCTAGCATGGCCG 406
              * ** ***** ***** ** ** * ** ***** ** * ** * ** ***** *

```

FIG. 5B

hoTCLW5	ACGCCGTGCTGGCCAGAGTGTACAAGCAGAGCGACCTGGACACCCTGGCCAAGGAGGCCA	466
hoTCLW6	ACGCCGTGCTGGCCCGCGTGTACAAGCAGAGCGACCTGGACACCCTGGCCAAGGAGGCCA	466
hoTCco	ATGCAGTGCTGGCTCGAGTCTACAAACAGTCTGACCTGGATACCCTGGCCAAGGAAGCTT	466
LW3	ACGCTGTGCTAGCTCGAGTCTACAAACAGTCCGACCTGGATACCCTGGCCAAGGAAGCTT	480
LW4	ACGCTGTGCTGGCTCGAGTCTACAAACAGTCCGACCTGGATACCCTGGCCAAGGAAGCTT	480
hoTts	ATGCAGTATTGGCTCGAGTGTATAAAACAATCAGATTTGGACACCCTGGCTAAAGAAGCAT	466
	* * * * *	
hoTCLW5	GCATCCCCATCATCAACGGCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACT	526
hoTCLW6	GCATCCCCATCATCAACGGCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACT	526
hoTCco	CTATCCCAATCATTAAATGGCCTGAGTGACCTGTATCACCCCATCCAGATTCTGGCCGATT	526
LW3	CTATCCCAATTATTAACGGCCTGTGAGACCTGTATCACCCCATCCAGATTCTGGCCGATT	540
LW4	CTATCCCAATTATTAACGGCCTGTGAGACCTGTATCACCCCATCCAGATTCTGGCCGATT	540
hoTts	CCATCCCAATTATCAATGGGCTGTGAGATTTGTACCATCCTATCCAGATCCTGGCTGAT	526
	* * * * *	
hoTCLW5	ACCTGACCCTGCAGGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCG	586
hoTCLW6	ACCTGACCCTGCAGGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCG	586
hoTCco	ACCTGACCCTCCAGGAGCATTATTCTAGTCTGAAAGGGCTGACACTGAGCTGGATTGGGG	586
LW3	ACCTGACCCTCCAGGAGCACTATTCTAGTCTGAAAGGGCTGACACTGAGTTGGATTGGGG	600
LW4	ACCTGACCCTCCAGGAGCACTATTCTAGTCTGAAAGGGCTGACACTGAGTTGGATTGGGG	600
hoTts	ACCTCAGCTCCAGGAACACTATAGCTCTCTGAAAGGTCTTACCCTCAGCTGGATCGGGG	586
	* * * * *	
hoTCLW5	ACGGCAACAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCCGGCATGCACCTGC	646
hoTCLW6	ACGGCAACAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCCGGCATGCACCTGC	646
hoTCco	ACGGAAACAATATCCTGCACTCCATTATGATGAGCGCCGCCAAGTTTGGAAATGCACCTCC	646
LW3	ACGGAAACAATATCCTGCACTCTATTATGATGTGAGCCGCCAAGTTTGGAAATGCACCTCC	660
LW4	ACGGAAACAATATCCTGCACTCTATTATGATGTGAGCCGCCAAGTTTGGAAATGCACCTCC	660
hoTts	ATGGAAACAATATCCTGCACTCCATCATGATGAGCGCAGCGAAATTCGGAATGCACCTTC	646
	* * * * *	
hoTCLW5	AGGCCGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCAGT	706
hoTCLW6	AGGCCGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCAGT	706
hoTCco	AGGCTGCAACCCCAAAGGCTACGAACCCGATGCCTCCGTGACAAAGCTGGCAGAACAGT	706
LW3	AGGCTGCAACCCCAAAGGCTACGAACCCGATGCCTCAGTGACAAAGCTGGCTGAACAGT	720
LW4	AGGCTGCAACCCCAAAGGCTACGAACCCGATGCCTCAGTGACAAAGCTGGCTGAACAGT	720
hoTts	AGGCAGCTACTCAAAGGGTTATGAGCCGGATGCTAGTGTAAACCAAGTTGGCAGAGCAGT	706
	* * * * *	
hoTCLW5	ACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGCCGCCACG	766
hoTCLW6	ACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGCCGCCACG	766
hoTCco	ATGCCAAAGAGAAACGGCACTAAGCTGCTGCTGACCAATGACCCCTCTGGAGGCCGCTCAG	766
LW3	ACGCCAAGGAGAACGGCACTAAGCTGCTGCTGACCAACGACCCCTCTGGAGGCCGCTCAG	780
LW4	ACGCCAAGGAGAACGGCACTAAGCTGCTGCTGACCAACGACCCCTCTGGAGGCCGCTCAG	780
hoTts	ATGCCAAGGAGAATGGTACCAAGCTGTTGCTGACAAATGATTCATTGGAAGCAGCGCATG	766
	* * * * *	
hoTCLW5	GCGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGAGGAGGAGAAGAAGA	826
hoTCLW6	GCGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGAGGAGGAGAAGAAGA	826
hoTCco	GAGGCAACGTGCTGATCACTGATACCTGGATTAGTATGGGACAGGAGGAAGAGAAGAAGA	826
LW3	GAGGCAACGTGCTGATCACCGATACTGGATTAGTATGGGACAGGAGGAAGAGAAGAAGA	840
LW4	GAGGCAACGTGCTGATCACCGATACTGGATTAGTATGGGACAGGAGGAAGAGAAGAAGA	840
hoTts	GAGGCAATGTATTAATTACAGCACTTGGATAAGCATGGGACAAGAAGGAGGAGAAGAAA	826
	* * * * *	

FIG. 5C

```

hotCLW5      AGAGACTGCAGGCCTTCCAGGGCTACCAGGTGACCATGAAGACCGCCAAGGTGGCCGCCA 886
hotCLW6      AGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGAAGACCGCCAAGGTGGCCGCCA 886
hotCco       AGCGGCTCCAGGCCTTCCAGGGCTACCAGGTGACAATGAAAACCGCTAAGGTTCGCAGCCA 886
LW3          AGCGGCTCCAGGCCTTCCAGGGCTACCAGGTGACAATGAAAACCGCTAAGGTTCGCAGCCA 900
LW4          AGCGGCTCCAGGCCTTCCAGGGCTACCAGGTGACAATGAAAACCGCTAAGGTTCGCAGCCA 900
hotTts       AGCGGCTCCAGGCCTTCCAAGGTTACCAGGTTACAATGAAGACTGCTAAAGTTGCTGCCT 886
              ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

hotCLW5      GCGACTGGACCTTCTGCACTGCCTGCCCAGAAAGCCCGAGGAGGTGGACGACGAGGTGT 946
hotCLW6      GCGACTGGACCTTCTGCACTGCCTGCCCCGAAAGCCCGAGGAGGTGGACGACGAGGTGT 946
hotCco       GCGACTGGACCTTCTGCACTGCCTGCCCAGAAAGCCCTGAAGAGGTGGACGATGAGGTCT 946
LW3          GCGATTGGACCTTCTGCACTGCCTGCCCAGAAAGCCCGAAGAGGTGGACGACGAGGTCT 960
LW4          GCGATTGGACCTTCTGCACTGCCTGCCCAGAAAGCCCGAAGAGGTGGACGACGAGGTCT 960
hotTts       CTGACTGGACATTTTACACTGCTTGCCCAGAAAGCCAGAAGAAGTGGATGATGAAGTCT 946
              ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

hotCLW5      TCTACAGCCCCAGAAGCCTGGTGTTCCTCCGAGGCCGAGAAACAGAAAGTGGACCATCATGG 1006
hotCLW6      TCTACAGCCCCGAGCCTGGTGTTCCTCCGAGGCCGAGAAACCGCAAGTGGACCATCATGG 1006
hotCco       TCTACTCAGCCAGAAGCCTGGTGTTCCTGAAGCTGAGAATAGGAAGTGGACAATCATGG 1006
LW3          TCTACTCTCCAGAAAGCCTGGTGTTCCTCCGAAAGCTGAGAATAGGAAGTGGACAATATGG 1020
LW4          TCTACTCTCCAGAAAGCCTGGTGTTCCTCCGAAAGCTGAGAATAGGAAGTGGACAATATGG 1020
hotTts       TTTATTCTCCTCGATCACTAGTGTTCCTCCAGAGGCAGAAAACAGAAAGTGGACAATCATGG 1006
              * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

hotCLW5      CCGTGATGGTGAGCCTGCTGACCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGA- 1065
hotCLW6      CCGTGATGGTGAGCCTGCTGACCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGA- 1065
hotCco       CAGTGATGGTCAGCCTGCTGACTGATTATTCCCCTCAGCTCCAGAAACCAAAGTTCTGAT 1066
LW3          CAGTGATGGTCAGCCTGCTGACTGATTATTACCTCAGCTCCAGAAACCAAAGTTCTGAT 1080
LW4          CAGTGATGGTCAGCCTGCTGACTGATTATTACCTCAGCTCCAGAAACCAAAGTTCTGAT 1080
hotTts       CTGTCATGGTGTCCCTGCTGACAGATTACTCACCTCAGCTCCAGAAGCCTAAATTT- --- 1062
              * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

hotCLW5      -----
hotCLW6      -----
hotCco       AA----- 1068
LW3          AAGCGGCCGC 1090
LW4          AAGCGGCCGC 1090
hotTts       -----
    
```