(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 805 045**

(51) Int. CI.:

C07K 14/47 (2006.01)
C12N 15/86 (2006.01)
A61K 48/00 (2006.01)
C07K 14/755 (2006.01)
A61K 38/37 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **23.04.2015 PCT/GB2015/051201**

(87) Fecha y número de publicación internacional: **26.11.2015 WO15177501**

(96) Fecha de presentación y número de la solicitud europea: **23.04.2015 E 15720762 (2)**

(97) Fecha y número de publicación de la concesión europea: **08.04.2020 EP 3145949**

(54) Título: **Vectores lentivirales**

(30) Prioridad:

21.05.2014 GB 201409089

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
10.02.2021

(73) Titular/es:

IP2IPO INNOVATIONS LIMITED (50.0%)
Top Floor, The Walbrook Building, 25 Walbrook
London EC4N 8AF, GB y
ID PHARMA CO., LTD. (50.0%)

(72) Inventor/es:

ALTON, ERIC WALTER FREDERICK WOLFGANG;
GRIESENBACH, UTA;
PYTEL, KAMILA MALGORZATA;
PAUL-SMITH, MICHAEL CHRISTIAN;
PRINGLE, IAN ANDREW;
HYDE, STEPHEN CHARLES;
GILL, DEBORAH REBECCA;
DAVIES, LEE ADRIAN;
BOYD, ALAN CHRISTOPHER;
MCLACHLAN, GERARD y
INOUE, MAKOTO

(74) Agente/Representante:

SALVÀ FERRER, Joan

Observaciones:

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

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

Vectores lentivirales

[0001] La presente invención se refiere a vectores de transferencia de genes lentivirales pseudotipados con hemaglutinina-neuraminidasa (HN) y proteínas de fusión (F) de un paramixovirus respiratorio, que comprenden un promotor y un transgén; y procedimientos para elaborar los mismos. La presente invención también se refiere al uso de dichos vectores en terapia génica, particularmente para el tratamiento de enfermedades del tracto respiratorio tal como fibrosis quística (FQ).

ANTECEDENTES DE LA INVENCIÓN

[0002] Los lentivirus pertenecen a un género de virus de la familia Retroviridae, y se caracterizan por un largo periodo de incubación. Los lentivirus pueden entregar una cantidad significativa de ARN viral en el ADN de la célula huésped y tienen la capacidad única entre los retrovirus de poder infectar células no divisorias, por lo que son uno de los procedimientos más eficientes de un vector de entrega de genes.

[0003] Los vectores lentivirales, especialmente los derivados del VIH-1, son vectores ampliamente estudiados y de uso frecuente. La evolución del esqueleto de los vectores lentivirales y la capacidad de los virus para entregar moléculas de ADN recombinante (transgenes) en las células diana han conducido a su uso en muchas aplicaciones. Dos posibles aplicaciones de los vectores virales incluyen la restauración de genes funcionales en terapia genética y la producción de proteínas recombinantes *in vitro*.

[0004] El pseudotipado es el procedimiento de producción de virus o vectores virales en combinación con proteínas de envoltura viral externas. Como tal, las proteínas de envoltura viral externas se pueden usar para alterar el tropismo del huésped o una estabilidad aumentada/disminuida de las partículas del virus. Por ejemplo, el pseudotipado permite especificar el carácter de las proteínas de envoltura. Una proteína de uso frecuente es la glucoproteína G del virus de la estomatitis vesicular (VSV), VSV-G corto.

[0005] Se entiende que la expresión retroviral eficiente y controlable de un transgén requiere la presencia de secuencias intrónicas. Sin embargo, la incorporación de tales intrones en vectores retrovirales implica procedimientos elaborados y que requieren mucho tiempo debido a los procedimientos de múltiples etapas empleados.

[0006] Hasta la fecha, los agentes de transferencia de genes virales no han sido útiles para el tratamiento de enfermedades, sin la transducción de poblaciones de células madre, debido a la respuesta inmunitaria adaptativa del huésped, lo que impide una administración repetida exitosa.

[0007] Además, la transferencia génica al epitelio de las vías respiratorias ha resultado ser más difícil de lo que se había previsto originalmente. Por ejemplo, el uso de pseudotipos lentivirales que requieren la alteración de la integridad epitelial para transducir las vías respiratorias, por ejemplo mediante el uso de detergentes tales como lisofosfatidilcolina o ácido bis(2-aminoetil éter)-*N, N,N'N'*-tetraacético de etilenglicol, se ha relacionado con un mayor riesgo de sepsis.

[0008] Un ejemplo de un entorno clínico que se beneficiaría de la transferencia génica al epitelio de las vías respiratorias es el tratamiento de la fibrosis quística (FQ). La FQ es un trastorno genético mortal causado por mutaciones en el gen regulador de la conductancia transmembrana de la FQ (CFTR), que actúa como un canal de cloruro en las células epiteliales de las vías respiratorias. La FQ se caracteriza por infecciones recurrentes en el pecho, aumento de las secreciones de las vías respiratorias y, finalmente, insuficiencia respiratoria. En el Reino Unido, la edad media actual al morir es de ~25 años. Para la mayoría de los genotipos, no hay tratamientos dirigidos al defecto básico; los tratamientos actuales para el alivio sintomático requieren horas de terapia autoadministrada diariamente. La terapia génica, a diferencia de los fármacos de molécula pequeña, es independiente de la clase mutacional de CFTR y, por lo tanto, es aplicable a todos los individuos con FQ afectados. Sin embargo, hasta la fecha, ningún vector viral ha cumplido los requisitos para su uso clínico, y lo mismo se aplica a otras enfermedades, particularmente a muchas otras enfermedades del tracto respiratorio.

[0009] En este sentido, se han encontrado al menos tres problemas principales. La eficiencia de la transferencia génica es generalmente deficiente, al menos en parte debido a que los receptores respectivos para muchos vectores virales parecen estar predominantemente localizados en la superficie basolateral del epitelio de las vías respiratorias. En segundo lugar, la penetración de la capa mucosa del tracto respiratorio es generalmente deficiente. Finalmente, la capacidad de administrar vectores virales repetidamente, obligatoria para el tratamiento de por vida de un epitelio autorrenovador, es limitada.

[0010] La administración de los vectores para aplicación clínica es otro factor pertinente. Por lo tanto, la estabilidad viral a través del uso de dispositivos clínicamente relevantes (por ejemplo, broncoscopio y nebulizador) debe mantenerse para la eficacia del tratamiento.

[0011] Otro ejemplo de un objetivo potencial para la terapia génica es la deficiencia de α1-antitripsina (A1AT). La deficiencia de A1AT es un trastorno hereditario que puede causar enfermedad pulmonar y enfermedad hepática. Los síntomas incluyen dificultad para respirar/sibilancias, disminución de la capacidad para hacer ejercicio, pérdida de peso, infecciones respiratorias recurrentes, fatiga y latidos cardíacos rápidos al ponerse de pie. Las personas afectadas a menudo desarrollan enfisema. Aproximadamente el 10-15 % por ciento de los pacientes con deficiencia de A1AT desarrollan enfermedad hepática. Las personas con deficiencia de A1AT también corren el riesgo de desarrollar un carcinoma hepatocelular.

[0012] La A1AT es una proteína secretada, producida principalmente en el hígado y después transportada al pulmón, produciéndose también cantidades más pequeñas en el pulmón. La función principal de A1AT es unir y neutralizar la elastasa de neutrófilos. Es probable que la terapia génica con A1AT sea de valor terapéutico en pacientes con deficiencia de A1AT, FQ y enfermedad pulmonar obstructiva crónica (EPOC), donde aumentar o introducir A1AT puede mejorar la función pulmonar.

[0013] La terapia con A1AT también es potencialmente valiosa para el tratamiento de enfermedades no respiratorias/no pulmonares, tales como diabetes tipo 1 y tipo 2, infarto agudo de miocardio, artritis reumatoide, enfermedad inflamatoria intestinal, rechazo de trasplante, enfermedad de injerto contra huésped (GvH), esclerosis múltiple e infecciones, particularmente infecciones virales, debido al efecto de la deficiencia de A1AT en otros tejidos/órganos, tal como el hígado y el páncreas (véase, por ejemplo, Lewis Mol. Med. 2012; 18:957-970).

[0014] La deficiencia de A1AT es una enfermedad objetivo atractiva para la terapia génica porque los niveles umbral terapéuticos están bien definidos. Una comparación de los niveles de A1AT en sujetos con riesgo de desarrollar enfisema/EPOC determinó un nivel umbral de protección de 11 µM en suero, usándose niveles inferiores a 11 µM como umbral para iniciar la terapia de aumento de proteínas cuando esté disponible. Los niveles de A1AT en el fluido de revestimiento de las vías respiratorias son solo ~10 % del nivel sérico, porque el epitelio pulmonar constituye una barrera y el umbral terapéutico en el fluido de revestimiento de la superficie de las vías respiratorias, por lo tanto, se considera de 1,1 µM (véanse Ferraroti et al. Thorax. agosto de 2012;67(8):669-74 y Abusriwil y Stockley 2006 Current Opinion in Pulmonary Medicine 12:125-131).

[0015] Seis formulaciones comerciales aprobadas por la FDA de proteína A1AT aislada de sangre humana agrupada se usan clínicamente en Estados Unidos para el tratamiento de pacientes con deficiencia grave de A1AT (a través de inyecciones intravenosas semanales). La terapia de reemplazo enzimático (ERT) es costosa (~$100.000/año) y aunque la eficacia bioquímica para la terapia de aumento de proteínas de ERT ha sido probada, la eficacia clínica ha sido más difícil de probar.

[0016] La ERT de A1AT no está actualmente disponible en todos los países y actualmente no está disponible en Reino Unido. Además, es difícil alcanzar niveles tisulares suficientemente sostenidos utilizando las terapias actuales, que, en parte, pueden ser responsables de la modesta eficacia clínica observada hasta ahora.

[0017] Otros objetivos atractivos para la terapia génica incluyen enfermedades cardiovasculares y trastornos sanguíneos, particularmente deficiencias en la coagulación sanguínea tales como hemofilia (A y B), enfermedad de Van Willebrand y deficiencia de Factor VII.

[0018] La hemofilia, particularmente la hemofilia A, es un objetivo atractivo para la terapia génica. La hemofilia A es un trastorno hemorrágico hereditario causado por una deficiencia o mutación del Factor VIII (FVIII). Su herencia está ligada al sexo, y casi todos los pacientes son hombres. La hemorragia es típicamente en las articulaciones. La hemorragia en el músculo, el tejido mucoso y el sistema nervioso central (SNC) es poco frecuente, pero puede ocurrir. La gravedad de la enfermedad es inversamente proporcional al nivel de FVIII: menos del 1 % (<0,01 UI/ml) da como resultado una enfermedad grave, con hemorragia después de una lesión mínima; entre el 1-5 % (0,01 UI/ml-0,05 UI/ml) causa enfermedad moderada, con hemorragia después de una lesión leve; y más del 5 % (>0,05 UI/ml) causa enfermedad leve, con hemorragia solo después de un traumatismo significativo o cirugía.

[0019] El documento EP 1950307 describe vectores lentivirales pseudotipados con proteínas espiculares de virus de ARN o ADN para la transferencia génica a las células madre epiteliales de las vías respiratorias.

[0020] El documento WO 2007/110628 describe promotores de hCEFI que comprenden un potenciador de CMV humano unido operativamente a un promotor de EF1a humano. Se descubrió que estos promotores tenían niveles de expresión altos y sostenidos y eran útiles para construcciones para la expresión génica.

[0021] Por lo tanto, existe la necesidad de un vector de terapia génica que pueda sortear uno o más de los problemas descritos anteriormente.

RESUMEN DE LA INVENCIÓN

[0022] Los presentes inventores han desarrollado un vector lentiviral que se ha pseudotipado con hemaglutinina-neuraminidasa (HN) y proteínas de fusión (F) de un paramixovirus respiratorio, que comprenden un promotor y un transgén. Típicamente, el esqueleto del vector proviene de un virus de inmunodeficiencia simia (VIS), tal como SIV1 o VIS de mono verde africano (SIV-AGM). Preferentemente, el esqueleto de un vector viral de la invención es de SIV-AGM. Las proteínas HN y F funcionan, respectivamente, para unirse a los ácidos siálicos y mediar la fusión celular para la entrada del vector a las células diana. Los presentes inventores han descubierto que este vector lentiviral pseudotipado específicamente con F/HN puede transducir eficientemente el epitelio de las vías respiratorias, lo que da como resultado una expresión transgénica sostenida durante periodos más allá de la vida propuesta de las células epiteliales de las vías respiratorias. Es importante destacar que los presentes inventores también han encontrado que la readministración no da como resultado una pérdida de eficacia. Estas características hacen que los vectores de la presente invención sean candidatos atractivos para tratar enfermedades a través de su uso en la expresión de proteínas terapéuticas: (i) dentro de las células del tracto respiratorio; (ii) secretados en el lumen del tracto respiratorio; y (iii) secretados en el sistema circulatorio.

[0023] La presente invención aborda una o más de las necesidades anteriores proporcionando vectores lentivirales pseudotipados con hemaglutinina-neuraminidasa (HN) y proteínas de fusión (F) de un paramixovirus respiratorio, que comprenden un promotor y un transgén. Como se reivindica, el promotor es un promotor de EF1a (hCEF)/potenciador híbrido de CMV humano. La presente invención también proporciona procedimientos para fabricar dichos vectores, composiciones que comprenden dichos vectores, y dichos vectores para su uso en terapia.

[0024] Los vectores de la presente invención permiten una expresión génica mayor y sostenida a través de una transferencia génica eficiente. Los problemas identificados anteriormente se abordan por la presente invención que proporciona vectores lentivirales pseudotipados con F/HN que son aptos para: (i) transducción de las vías respiratorias sin alteración de la integridad epitelial; (ii) expresión génica persistente; (iii) falta de toxicidad crónica; y (iv) administración repetida eficiente. La expresión génica estable a largo plazo/persistente, preferentemente a un nivel terapéuticamente eficaz, puede lograrse usando dosis repetidas de un vector de la presente invención. Como alternativa, se puede usar una dosis única para lograr la expresión deseada a largo plazo.

[0025] En contraste con los vectores lentivirales conocidos, los vectores lentivirales de la invención exhiben una absorción celular eficiente de las vías respiratorias, una expresión transgénica mejorada, y no sufren pérdida de eficacia tras la administración repetida.

[0026] Por lo tanto, ventajosamente, los vectores lentivirales de la presente invención pueden usarse en terapia génica. A modo de ejemplo, las propiedades eficientes de absorción de células de las vías respiratorias de los vectores de la invención los hacen muy adecuados para tratar enfermedades del tracto respiratorio. Los vectores lentivirales de la invención también se pueden usar en procedimientos de terapia génica para promover la secreción de proteínas terapéuticas. A modo de ejemplo adicional, los vectores de la invención proporcionan la secreción de proteínas terapéuticas en el lumen del tracto respiratorio o el sistema circulatorio. Por lo tanto, la administración de un vector de la invención y su absorción por las células de las vías respiratorias puede permitir el uso de los pulmones (o nariz o vías respiratorias) como una "fábrica" para producir una proteína terapéutica que después se secreta y entra en la circulación general a niveles terapéuticos, donde puede viajar a células/tejidos de interés para provocar un efecto terapéutico. A diferencia de las proteínas intracelulares o de membrana, la producción de dichas proteínas secretadas no depende de que se transduzcan células diana específicas de la enfermedad, lo cual es una ventaja significativa y logra altos niveles de expresión de proteínas. Por lo tanto, otras enfermedades que no son enfermedades del tracto respiratorio, tales como enfermedades cardiovasculares y trastornos sanguíneos, particularmente deficiencias en la coagulación sanguínea, también pueden tratarse mediante los vectores de la presente invención.

[0027] Como ejemplo, la alfa-1 antitripsina (A1AT) es una antiproteasa secretada que se produce principalmente en el hígado y después se transporta al pulmón, produciéndose también cantidades más pequeñas en el pulmón. La función principal de A1AT es unir y neutralizar/inhibir la elastasa de neutrófilos. La terapia génica con A1AT según la presente invención es relevante para pacientes con deficiencia de A1AT, así como en otras enfermedades pulmonares tal como fibrosis quística o enfermedad pulmonar obstructiva crónica (EPOC), y ofrece la oportunidad de superar algunos de los problemas encontrados por el reemplazo enzimático terapia.

[0028] Los presentes inventores han demostrado previamente que existe una correlación significativa entre la elastasa de neutrófilos (NE) y A1AT en muestras de esputo de pacientes con fibrosis quística, lo que demuestra que el cuerpo produce A1AT en respuesta a la exposición a NE. Los presentes inventores también han demostrado que existe una correlación estadísticamente significativa entre la NE y el índice de depuración pulmonar, un marcador de enfermedad de las vías respiratorias pequeñas, lo que implica que el aumento de la NE tiene un impacto negativo en la función pulmonar. Como se presenta en esta invención, los inventores ahora han demostrado sorprendentemente que los vectores lentivirales de la invención pueden alcanzar altas concentraciones de A1AT y expresión de A1AT a largo plazo (al menos 90 días) *in vivo*. Por lo tanto, la terapia génica con A1AT puede neutralizar la NE, mejorando la función pulmonar de los pacientes con fibrosis quística y/o EPOC (y tener un efecto terapéutico en otras indicaciones como se describe en esta invención). Por consiguiente, la presente invención se refiere al uso de un vector lentiviral como se describe en esta invención para la administración de un transgén A1AT y terapia génica de afecciones que

incluyen, pero sin limitación, deficiencia de A1AT, fibrosis quística y/o EPOC. La administración de A1AT lentiviral directamente al epitelio nasal y/o el pulmón puede superar algunas de las limitaciones a las que actualmente se enfrenta la terapia de reemplazo enzimático (A1AT aislada de sangre humana y administrada por vía intravenosa cada semana), proporcionando una expresión estable y duradera en el tejido objetivo (epitelio pulmonar/nasal), facilidad de administración, y disponibilidad ilimitada.

[0029] En algunas realizaciones, la transducción con un vector lentiviral de la invención conduce a la secreción de la proteína recombinante en el lumen del pulmón, así como en la circulación. Un beneficio de esto es que la proteína terapéutica alcanza el intersticio. En el caso de deficiencia de A1AT, esto es ventajoso porque también se requiere inhibición de NE en este sitio. Por lo tanto, la terapia génica con A1AT también puede ser beneficiosa en otras indicaciones de enfermedad, cuyos ejemplos no limitantes incluyen diabetes tipo 1 y tipo 2, infarto agudo de miocardio, cardiopatía isquémica, artritis reumatoide, enfermedad inflamatoria intestinal, rechazo de trasplante, enfermedad de injerto contra huésped (GvH), esclerosis múltiple, enfermedad hepática, cirrosis, vasculitis e infecciones, tales como infecciones bacterianas y/o virales.

[0030] A1AT tiene muchos otros efectos antiinflamatorios y protectores de los tejidos, por ejemplo en modelos preclínicos de diabetes, enfermedad de injerto contra huésped, y enfermedad inflamatoria intestinal. La producción de A1AT en el pulmón y/o la nariz después de la transducción según la presente invención puede, por lo tanto, ser más ampliamente aplicable, incluso a estas indicaciones.

[0031] Otros ejemplos de enfermedades que pueden tratarse con terapia génica de una proteína secretada según la presente invención incluyen enfermedades cardiovasculares y trastornos sanguíneos, particularmente deficiencias en la coagulación sanguínea tales como hemofilia (A y B), enfermedad de von Willebrand y deficiencia de Factor VII.

[0032] En algunas realizaciones, la hemofilia A puede tratarse según la presente invención. La gravedad de la enfermedad es inversamente proporcional al nivel de FVIII, y un aumento en el FVIII del 2-5 % (0,02-0,05 UI/ml) es suficiente para ser terapéuticamente eficaz.

[0033] En algunas realizaciones, la nariz es un sitio de producción preferido para una proteína terapéutica que usa un vector de terapia génica de la invención por al menos una de las siguientes razones: (i) las barreras extracelulares tales como las células inflamatorias y el esputo son menos pronunciadas en la nariz; (ii) facilidad de administración vectorial; (iii) se requieren cantidades más pequeñas de vector; y (iv) consideraciones éticas. Por lo tanto, la transducción de células epiteliales nasales con un vector lentiviral de la invención puede dar como resultado una expresión eficiente (de alto nivel) y duradera del transgén terapéutico de interés.

[0034] Los vectores de la presente invención permiten la expresión génica a largo plazo, dando como resultado la expresión a largo plazo de una proteína terapéutica. Como se describe en esta invención, las frases "expresión a largo plazo", "expresión sostenida" y "expresión persistente" se usan indistintamente. La expresión a largo plazo según la presente invención significa la expresión de un gen y/o proteína terapéuticos, preferentemente a niveles terapéuticos, durante al menos 45 días, al menos 60 días, al menos 90 días, al menos 120 días, al menos 180 días, al menos 250 días, al menos 360 días, al menos 450 días, al menos 730 días o más. Preferentemente, expresión a largo plazo significa expresión durante al menos 90 días, al menos 120 días, al menos 180 días, al menos 250 días, al menos 360 días, al menos 450 días, al menos 720 días o más, más preferentemente al menos 360 días, al menos 450 días, al menos 720 días o más. Esta expresión a largo plazo puede lograrse mediante dosis repetidas o mediante una dosis única.

[0035] Las dosis repetidas pueden administrarse dos veces al día, diariamente, dos veces a la semana, semanalmente, mensualmente, cada dos meses, cada tres meses, cada cuatro meses, cada seis meses, anualmente, cada dos años, o más. La dosificación puede continuar durante el tiempo que sea necesario, por ejemplo, durante al menos seis meses, al menos un año, dos años, tres años, cuatro años, cinco años, diez años, quince años, veinte años o más, durante toda la vida del paciente a tratar.

[0036] Los vectores lentivirales, como los de la invención, pueden integrarse en el genoma de las células transducidas y conducir a una expresión duradera, haciéndolos adecuados para la transducción de células madre/progenitoras. En el pulmón, se han identificado varios tipos de células con capacidad regenerativa como responsables del mantenimiento de linajes celulares específicos en las vías respiratorias conductoras y los alvéolos. Estas incluyen las células basales y las células del conducto de la glándula submucosa en las vías respiratorias superiores, las células Clara y las células neuroendocrinas en las vías respiratorias bronquiolares, las células madre bronquioalveolares en los bronquiolos terminales y los neumocitos tipo II en los alvéolos. Por lo tanto, y sin quedar ligados a la teoría, se cree que los vectores de la presente invención provocan la expresión génica a largo plazo del transgén de interés al introducir el transgén en una o más células epiteliales de vías respiratorias de larga duración o tipos de células, tales como células basales y células del conducto de la glándula submucosa en las vías respiratorias superiores, células Clara y células neuroendocrinas en las vías respiratorias bronquiolares, células madre bronquioalveolares en los bronquiolos terminales, y neumocitos tipo II en los alvéolos.

[0037] Por consiguiente, los vectores lentivirales de la invención pueden transducir una o más células o líneas celulares con potencial regenerativo dentro del pulmón (incluidas las vías respiratorias y el tracto respiratorio) para lograr una expresión génica a largo plazo. En una realización preferida, el vector lentiviral de la invención transduce células basales, tales como las de las vías respiratorias superiores/tracto respiratorio. Las células basales tienen un papel central en los procedimientos de mantenimiento y reparación epitelial después de una lesión. Además, las células basales están ampliamente distribuidas a lo largo del epitelio respiratorio humano, con una distribución relativa que varía del 30 % (vías respiratorias más grandes) al 6 % (vías respiratorias más pequeñas).

[0038] Los vectores lentivirales de la invención pueden usarse para transducir células madre/progenitoras aisladas y expandidas *ex vivo* antes de la administración a un paciente. Preferentemente, los vectores lentivirales de la invención se usan para transducir células dentro del pulmón (o vías respiratorias/tracto respiratorio) *in vivo*.

[0039] Los vectores de la presente invención permiten altos niveles de expresión génica, dando como resultado altos niveles (preferentemente niveles terapéuticos) de expresión de una proteína terapéutica. La expresión se puede medir por cualquier procedimiento apropiado (cualitativo o cuantitativo, preferentemente cuantitativo), y las concentraciones se dan en cualquier unidad de medida apropiada, por ejemplo ng/ml. Un alto nivel de expresión según la presente invención puede significar la expresión de un gen terapéutico y/o proteína a una concentración de al menos 10 ng/ml, al menos 20 ng/ml, al menos 30 ng/ml, al menos 40 ng/ml, al menos 50 ng/ml, al menos 60 ng/ml, al menos 70 ng/ml, al menos 80 ng/ml, al menos 90 ng/ml, al menos 100 ng/ml, al menos 200 ng/ml, al menos 300 ng/ml, al menos 400 ng/ml, al menos 500 ng/ml, al menos 600 ng/ml, al menos 700 ng/ml, al menos 800 ng/ml, al menos 900 ng/ml, al menos 1.000 ng/ml, al menos 2.000 ng/ml, al menos 3.000 ng/ml, al menos 4.000 ng/ml, al menos 5.000 ng/ml, al menos 10.000, al menos 15.000 ng/ml, al menos 20.000 ng/ml o más. La expresión terapéutica puede definirse usando estos mismos valores.

[0040] Los vectores lentivirales de la presente invención típicamente proporcionan altos niveles de expresión de un transgén cuando se administran a un paciente. Los términos alta expresión y expresión terapéutica se usan indistintamente en esta invención.

[0041] Un alto nivel de expresión según la presente invención puede significar la expresión de un gen terapéutico y/o proteína a una concentración de al menos aproximadamente 100 nM, al menos aproximadamente 200 nM, al menos aproximadamente 300 nM, al menos aproximadamente 400 nM, al menos aproximadamente 500 nM, al menos aproximadamente 600 nM, al menos aproximadamente 700 nM, al menos aproximadamente 800 nM, al menos aproximadamente 900 nM, al menos aproximadamente 1 mM, al menos aproximadamente 1,1 mM, al menos aproximadamente 1,2 mM, al menos aproximadamente 1,3 mM, al menos aproximadamente 1,4 mM, al menos aproximadamente 1,5 mM, al menos aproximadamente 2 mM, al menos aproximadamente 3 mM, al menos aproximadamente 4 mM, al menos aproximadamente 5 mM, al menos aproximadamente 6 mM, al menos aproximadamente 7 mM, al menos aproximadamente 8 mM, al menos aproximadamente 9 mM, al menos aproximadamente 10 mM, al menos aproximadamente 11 mM, al menos aproximadamente 12 mM, al menos aproximadamente 13 mM, al menos aproximadamente 14 mM, al menos aproximadamente 15 mM, al menos aproximadamente 20 mM, al menos aproximadamente 25 mM, al menos aproximadamente 30 mM, al menos aproximadamente 40 mM, al menos aproximadamente 50 mM, al menos aproximadamente 75 mM, o al menos aproximadamente 100 mM o más. La expresión terapéutica puede definirse usando estos mismos valores.

[0042] Un alto nivel de expresión según la presente invención puede significar la expresión de un gen terapéutico (típicamente medida por la expresión de ARNm) al menos aproximadamente al 1 %, al menos aproximadamente al 2 %, al menos aproximadamente al 3 %, al menos aproximadamente al 4 %, al menos aproximadamente al 5 %, al menos aproximadamente al 6 %, al menos aproximadamente al 7 %, al menos aproximadamente al 8 %, al menos aproximadamente al 9 %, al menos aproximadamente al 10 %, al menos aproximadamente al 15 %, al menos aproximadamente al 20 % o más en comparación con el nivel de expresión del ARNm endógeno (defectuoso) correspondiente. La expresión terapéutica puede definirse usando estos mismos valores. Por ejemplo, un nivel de expresión típico de ARNm de CFTR endógeno puede cuantificarse en términos del número de copias del ARNm por célula pulmonar, por ejemplo, una copia del ARNm de CFTR endógeno por célula pulmonar, dos copias del ARNm de CFTR endógeno por célula pulmonar, tres copias del ARNm de CFTR endógeno por célula pulmonar, cuatro copias del ARNm de CFTR endógeno por célula pulmonar, cinco copias del ARNm de CFTR endógeno por célula pulmonar, o más, preferentemente, dos copias del ARNm de CFTR endógeno por célula pulmonar. La expresión del gen terapéutico de la invención, tal como un gen CFTR funcional, puede cuantificarse en relación con el gen endógeno, tal como los genes CFTR endógenos (disfuncionales) en términos de copias de ARNm por célula o cualquier otra unidad apropiada.

[0043] Un alto nivel de expresión según la presente invención puede significar la expresión de un gen y/o proteína terapéuticos a una concentración de al menos aproximadamente el 0,5 %, al menos aproximadamente el 1 %, al menos aproximadamente el 2 %, al menos aproximadamente el 3 %, al menos aproximadamente el 4 %, al menos aproximadamente el 5 %, al menos aproximadamente el 6 %, al menos aproximadamente el 7 %, al menos aproximadamente el 8 %, al menos aproximadamente el 9 %, al menos aproximadamente el 10 %, al menos

aproximadamente el 15 %, al menos aproximadamente el 20 %, al menos aproximadamente el 25 %, al menos aproximadamente el 30 %, al menos aproximadamente el 40 %, al menos aproximadamente el 50 %, al menos aproximadamente el 60 %, al menos aproximadamente el 70 %, al menos aproximadamente el 80 %, al menos aproximadamente el 90 %, al menos el 95 %, al menos aproximadamente el 96 %, al menos aproximadamente el 97 %, al menos aproximadamente el 98 %, al menos aproximadamente el 99 %, o más en comparación con el nivel natural del gen y/o proteína terapéuticos, en el que el nivel natural es el nivel en un individuo normal sin la enfermedad. En algunas realizaciones, la expresión natural se da al 100 %, con cualquier mejora en la expresión génica medida con respecto esta. Como ejemplo no limitante, si en un individuo normal sin la enfermedad la expresión del gen funcional se da al 100 %, y en un individuo con la enfermedad, la expresión del gen funcional es del 0 %, un nivel terapéutico de expresión del gen o proteína puede ser de al menos aproximadamente el 0,5 %, al menos aproximadamente el 1 %, al menos aproximadamente el 2 %, al menos aproximadamente el 3 %, al menos aproximadamente el 4 %, al menos aproximadamente el 5 %, al menos aproximadamente el 6 %, al menos aproximadamente el 7 %, al menos aproximadamente el 8 %, al menos aproximadamente el 9 %, al menos aproximadamente el 10 %, al menos aproximadamente el 15 %, al menos aproximadamente el 20 %, al menos aproximadamente el 25 %, al menos aproximadamente el 30 %, al menos aproximadamente el 40 %, al menos aproximadamente el 50 %, o más en comparación con el nivel natural del gen y/o proteína terapéuticos. Como otro ejemplo no limitante, si en un individuo normal sin la enfermedad la expresión del gen funcional se da al 100 %, y en un individuo con la enfermedad, la expresión del gen funcional es del 50 %, un nivel terapéutico de expresión del gen o proteína puede ser de al menos aproximadamente el 55 %, al menos aproximadamente el 60 %, al menos aproximadamente el 70 %, al menos aproximadamente el 80 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 %, al menos aproximadamente el 96 %, al menos aproximadamente el 97 %, al menos aproximadamente el 98 %, al menos aproximadamente el 99 %, o más en comparación con el nivel natural del gen y/o proteína terapéuticos.

[0044] Para proteínas secretadas tal como A1AT, típicamente la concentración en el pulmón o el fluido de revestimiento epitelial (medido usando BAL) es aproximadamente diez veces mayor que en el suero. Como ejemplo no limitativo, si la concentración de proteína secretada en el pulmón o el fluido de revestimiento epitelial está en la región de 750 ng/ml, la concentración sérica de la proteína está en la región de 75 ng/ml.

[0045] Los niveles de expresión de un gen y/o proteína terapéuticos de la invención pueden medirse en el tejido pulmonar, el fluido de revestimiento epitelial y/o el suero/plasma según sea apropiado. Por lo tanto, un nivel de expresión alto y/o terapéutico puede referirse a la concentración en el pulmón, el fluido de revestimiento epitelial y/o el suero/plasma.

[0046] Como un ejemplo no limitativo, un nivel de expresión terapéutica de CFTR es típicamente del 1-5 % del ARNm de CFTR terapéutico en comparación con el nivel de expresión del ARNm de CFTR endógeno (defectuoso).

[0047] Como otro ejemplo no limitante, un nivel de expresión terapéutica de A1AT es típicamente al menos aproximadamente 1 µM en el fluido de revestimiento epitelial, y/o al menos aproximadamente 0,1 µM en el suero. En una realización preferida, un nivel de expresión terapéutica de A1AT en el fluido de revestimiento epitelial es de al menos aproximadamente 1,1 µM, y/o un nivel de expresión terapéutica en suero de A1AT según la presente invención es de al menos aproximadamente 11 µM. Como otro ejemplo no limitante, un nivel de expresión terapéutica de A1AT en el fluido de revestimiento epitelial (ELF, es decir, el fluido que recubre las vías respiratorias y los espacios aéreos en los pulmones) es de 70 µg/ml (en comparación con un nivel objetivo "normal" de ATT (A1AT) en el ELF de 200 µg/ml).

[0048] Como otro ejemplo no limitante, un nivel de expresión terapéutica de proteína FVIII es típicamente de al menos aproximadamente el 1-3 % o al menos aproximadamente el 1-6 % del nivel de expresión en un individuo normal que no padece hemofilia.

[0049] El gen terapéutico incluido en el vector de la invención puede modificarse para facilitar la expresión. Por ejemplo, la secuencia génica puede estar en forma agotada en CpG y/o optimizada por codones para facilitar la expresión génica. Las técnicas estándar para modificar la secuencia génica de esta manera son conocidas en la técnica.

[0050] El promotor incluido en un vector puede seleccionarse y/o modificarse específicamente para ajustar adicionalmente la regulación de la expresión del gen terapéutico. De nuevo, los promotores adecuados y las técnicas estándar para su modificación se conocen en la técnica. Como ejemplo no limitativo, se describen varios promotores adecuados (libres de CpG) en Pringle et al. (J. Mol. Med. Berl. 2012, 90(12): 1487-96). Como se reivindica, el promotor de la invención es un promotor de hCEF.

[0051] El vector de la invención puede modificarse para permitir la inhibición de la expresión génica. Las técnicas estándar para modificar el vector de esta manera se conocen en la técnica. Como ejemplo no limitativo, los promotores sensibles a Tet son ampliamente utilizados.

[0052] Los vectores de la presente invención también demuestran una resistencia notable a las fuerzas de cizallamiento con solo una reducción modesta en la capacidad de transducción cuando se pasan a través de dispositivos de administración clínicamente relevantes tales como broncoscopios, frascos de pulverización y nebulizadores.

[0053] En una realización, la invención proporciona vectores lentivirales F/HN que comprenden un promotor y un transgén, que no tienen un intrón posicionado entre el promotor y el transgén. En una realización, el vector de la presente invención se administra a las células del tracto respiratorio. En la realización, el lentivirus es VIS. Como se reivindica, el promotor es un promotor de EF1 a (hCEF)/potenciador híbrido de CMV humano. Típicamente, dicho promotor de la invención carece del intrón correspondiente a los nucleótidos 570-709 y el exón correspondiente a los nucleótidos 728-733 del promotor de hCEF. Un ejemplo preferido de una secuencia promotora de hCEF de la invención se proporciona por la SEQ ID NO: 6. También se describen promotores de CMV. La SEQ ID NO: 17 proporciona un ejemplo de una secuencia promotora de CMV. En la técnica se conocen otros promotores para la expresión transgénica y su idoneidad para los vectores lentivirales se determina usando técnicas de rutina conocidas en la técnica. Ejemplos no limitantes de otros promotores incluyen UbC y UCOE. Como se describe en el presente documento, el promotor puede modificarse para regular adicionalmente la expresión del transgén de la invención.

[0054] En una realización, el transgén codifica un CFTR. La SEQ ID NO: 7 proporciona un ejemplo de un ADNc de CFTR.

[0055] En una realización, el transgén codifica un A1AT. Se proporciona un ejemplo de un transgén *A1AT* por la SEQ ID NO: 15, o por la secuencia complementaria de la SEQ ID NO: 26. La SEQ ID NO: 15 es un transgén *A1AT* agotado en CpG optimizado por codones diseñado por los presentes inventores para mejorar la traducción en células humanas. Se ha demostrado que dicha optimización mejora la expresión génica hasta 15 veces. Por lo tanto, en una realización, la invención proporciona un vector que comprende un polinucleótido que comprende o que consiste en la secuencia de nucleótidos de SEQ ID NO: 15. También se incluyen variantes de la misma secuencia (como se define en esta invención) que poseen el mismo efecto técnico de mejorar la traducción en comparación con la secuencia del gen *A1AT* no modificada (natural). La presente descripción proporciona además un polipéptido codificado por dicho transgén *A1AT*, como se ilustra por el polipéptido de la SEQ ID NO: 27, plásmidos (particularmente plásmidos del genoma del vector como se define en esta invención). La invención proporciona además vectores lentivirales que comprenden dicho transgén *A1AT*. En una realización preferida, los aspectos de la invención relacionados con la terapia génica con A1AT según la presente invención usan la secuencia del transgén *A1AT* de SEQ ID NO: 15.

[0056] En una realización, el transgén codifica un FVIII. Los ejemplos de un transgén *FVIII* se proporcionan por las SEQ ID NOs: 16 y 30, o por las secuencias complementarias respectivas de las SEQ ID NO: 28 y 31.

[0057] Los vectores lentivirales adecuados para su uso en la presente invención incluyen el virus de la inmunodeficiencia humana (VIH), el virus de la inmunodeficiencia simia (VIS), el virus de la inmunodeficiencia felina (VIF), el virus de la anemia infecciosa equina (VAIE), y el virus Visna/maedi. En una realización de la invención, se usa un vector del VIS, preferentemente SIV-AGM. En otra realización, se usa un vector del VIH.

[0058] Los vectores lentivirales de la presente invención están pseudotipados con hemaglutinina-neuraminidasa (HN) y proteínas de fusión (F) de un paramixovirus respiratorio. En una realización, el paramixovirus respiratorio es un virus Sendai (virus de la parainfluenza murina tipo 1).

[0059] En una realización de la invención, el vector lentiviral es competente en integrasa (IC). En una realización alternativa, el vector lentiviral es deficiente en integrasa (ID).

[0060] En otra realización de la invención, el transgén de la invención es uno cualquiera o más de *DNAH5, DNAH11, DNAI1* y *DNAI2*, u otro gen relacionado conocido.

[0061] En una realización de la invención, el epitelio del tracto respiratorio está dirigido a la administración del vector. En esta realización, el transgén codifica la alfa-1 antitripsina (A1AT), la proteína tensioactiva B (SFTPB) o el factor estimulante de colonias de granulocitos-macrófagos (GM-CSF). En otra realización, el transgén codifica un anticuerpo monoclonal (mAb) contra un agente infeccioso. En una realización, el transgén codifica anti-TNF alfa. En una realización adicional, el transgén codifica una proteína terapéutica implicada en una afección inflamatoria, inmunitaria o metabólica.

[0062] En una realización de la invención, el vector se administra a las células del tracto respiratorio para permitir que la producción de proteínas se secrete al sistema circulatorio. En esta realización, el transgén codifica el Factor VII, Factor VIII, Factor IX, Factor X, Factor XI y/o factor de von Willebrand. Tal vector puede usarse en el tratamiento de enfermedades, particularmente enfermedades cardiovasculares y trastornos de la sangre, preferentemente deficiencias en la coagulación de la sangre, tal como la hemofilia. En otra realización, el transgén codifica un anticuerpo monoclonal (mAb) contra un agente infeccioso. En una realización, el transgén codifica una proteína implicada en una afección inflamatoria, inmunitaria o metabólica, tal como enfermedad de almacenamiento

lisosomal.

[0063] Según la invención, se proporciona un vector lentiviral F/HN-SIV que comprende un promotor de hCEF y un transgén *CFTR*, que no tiene ningún intrón posicionado entre el promotor y el transgén. De manera similar, no hay ningún intrón entre el promotor y el transgén en el plásmido del genoma del vector (pDNA1) (por ejemplo, pGM326 como se describe en esta invención, ilustrado en la figura 1A y con la secuencia de SEQ ID NO: 1).

[0064] La invención también proporciona un vector lentiviral F/HN-SIV que comprende un promotor de hCEF y un transgén *A1AT*, que no tiene ningún intrón posicionado entre el promotor y el transgén. Tal vector lentiviral se puede producir mediante el procedimiento descrito en esta invención, usando un plásmido que porta el transgén *A1AT* y un promotor. De manera similar, no hay intrón entre el promotor y el transgén *A1AT* en el plásmido del genoma del vector (pDNA1). Una secuencia ejemplar de dicho plásmido se proporciona en la SEQ ID NO: 9 (F/HN-SIV-hCEF-soA1AT, ilustrada en la figura 15A).

[0065] La invención también proporciona un vector lentiviral F/HN-SIV que comprende (i) un promotor de hCEF; y (ii) un transgén *FVIII*; en el que no se posiciona ningún intrón entre el promotor y el transgén. Tal vector lentiviral se puede producir mediante el procedimiento descrito en esta invención, usando un plásmido que porta el transgén *FVIII* y un promotor. De manera similar, no hay intrón entre el promotor y el transgén *FVIII* en el plásmido del genoma del vector (pDNA1). Se dan secuencias ejemplares de dichos plásmidos en las SEQ ID NO: 11 a 14 (ilustradas en las figuras 22A a E).

[0066] El vector lentiviral como se ha descrito anteriormente comprende un transgén. El transgén comprende una secuencia de ácido nucleico que codifica un producto génico, por ejemplo, una proteína.

[0067] Por ejemplo, en una realización, la secuencia de ácido nucleico que codifica un *CFTR, A1AT* o *FVIII* comprende (o consiste en) una secuencia de ácido nucleico que tiene al menos un 90 % (tal como al menos un 90, 92, 94, 95, 96, 97, 98, 99 o un 100 %) de identidad de secuencia con la secuencia de ácido nucleico de *CFTR, A1AT* o *FVIII* respectivamente. En una realización adicional, la secuencia de ácido nucleico que codifica un *CFTR, A1AT* o *FVIII* comprende (o consiste en) una secuencia de ácido nucleico que tiene al menos un 95 % (tal como al menos un 95, 96, 97, 98, 99 o un 100 %) de identidad de secuencia con la secuencia de ácido nucleico de *CFTR, A1AT* o *FVIII* respectivamente. En una realización, la secuencia de ácido nucleico que codifica *CFTR* se proporciona por la SEQ ID NO: 7, la secuencia de ácido nucleico que codifica A1AT se proporciona por la SEQ ID NO: 15, o por la secuencia complementaria de la SEQ ID NO: 26, y/o la secuencia de ácido nucleico que codifica FVIII se proporciona por la SEQ ID NO: 16 y 30, o por las secuencias complementarias respectivas de SEQ ID NO: 28 y 31, o sus variantes.

[0068] El término "polipéptido", como se usa en esta invención también incluye secuencias variantes. Por lo tanto, el polipéptido codificado por el transgén de la invención puede tener al menos un 90 % (tal como al menos un 90, 92, 94, 95, 96, 97, 98, 99 o un 100 %) de identidad de secuencia con una secuencia de polipéptidos funcional de CFTR, A1AT o FVIII respectivamente. En una realización adicional, la secuencia de aminoácidos del transgén CFTR, A1AT o FVIII comprende (o consiste en) una secuencia de aminoácidos que tiene al menos un 95 % (tal como al menos un 95, 96, 97, 98, 99 o un 100 %) de identidad de secuencia con la secuencia de polipéptidos funcional de CFTR, A1AT o FVIII respectivamente. En una realización, la secuencia de aminoácidos de la proteína A1AT codificada por el transgén de la invención comprende (o consiste en) la secuencia de aminoácidos de SEQ ID NO: 27, o variantes de la misma. Preferentemente, dichas proteínas A1AT variantes de la invención tienen al menos un 90 % (tal como al menos un 90, 92, 94, 95, 96, 97, 98, 99 o un 100 %), más preferentemente al menos un 95 % o más de identidad de secuencia con la SEQ ID NO: 27.

[0069] En una realización, la secuencia de ácido nucleico que codifica *CFTR, A1AT* o *FVIII* comprende (o consiste en) la secuencia de ADN complementaria de *CFTR, A1AT* o *FVII*, respectivamente. En una realización, el transgén *CFTR, A1AT* o *FVIII* es un *CFTR, A1AT* o *FVIII* con secuencia optimizada (*soCFTR2, soA1AT* o *FVIII*). Se proporciona un ejemplo por las SEQ ID NOS: 7, 15 y 16, respectivamente. Una secuencia complementaria ejemplar de *A1AT* con secuencia optimizada se proporciona mediante la SEQ ID NO: 26. Las secuencias de *FVIII* complementarias con secuencia optimizada ejemplares se proporcionan mediante la SEQ ID NO: 28 y 31.

[0070] En un caso de la descripción, la expresión transgénica del vector F/HN está dirigida por el promotor de citomegalovirus (CMV). En otro caso, la expresión transgénica del vector es conducida por el promotor del factor de alargamiento 1a (EF1a). Como se reivindica, la expresión transgénica del vector es conducida por el promotor de EF1a (hCEF)/potenciador híbrido de CMV humano. En una realización, el promotor de hCEF tiene todos los dinucleótidos CG reemplazados por cualquiera de AG, TG o GT. Por lo tanto, en una realización, el promotor de hCEF está libre de CpG.

[0071] En una realización, el vector lentiviral se puede producir usando el plásmido F/HN-SIV-hCEF-soCFTR2-IC. En esta realización, *CFTR* se expresa bajo el control del promotor de hCEF. Puede describirse que este vector lentiviral comprende F/HN-SIV-hCEF-soCFTR2-IC, ya que comprende los elementos SIV F/HN, así como un casete de expresión competente en integrasa que comprende *CFTR* bajo el control del promotor de hCEF. Este vector

lentiviral de la invención es capaz de producir una expresión duradera, repetible y de alto nivel en las células de las vías respiratorias sin inducir una respuesta inmunitaria excesiva. En consecuencia, la invención proporciona un medio eficaz de terapia génica *in vivo*, por ejemplo, transferencia de genes *CFTR* al pulmón de FQ para el tratamiento de la enfermedad pulmonar por FQ.

[0072] En una realización preferida, el vector lentiviral se puede producir usando el plásmido F/HN-SIV-hCEF-soA1AT. En esta realización, *A1AT* se expresa bajo el control del promotor de hCEF. Puede describirse que este vector lentiviral comprende F/HN-SIV-hCEF-soA1AT, ya que comprende los elementos SIV F/HN, así como un casete de expresión que comprende *A1AT* bajo el control del promotor de hCEF. Este vector lentiviral de la invención es capaz de producir una expresión duradera, repetible y de alto nivel en las células de las vías respiratorias sin inducir una respuesta inmunitaria excesiva. En consecuencia, la invención proporciona un medio eficaz de terapia génica *in vivo*, por ejemplo, la transferencia del gen *A1AT* al pulmón o la nariz de un paciente para la producción de A1AT que después se secreta al sistema circulatorio (como se describe en esta invención). Por lo tanto, este vector y otros vectores de la invención que comprenden el transgén *A1AT* pueden usarse en el tratamiento de la deficiencia de A1AT u otras indicaciones como se describe en esta invención.

[0073] En otro caso preferido de la presente descripción, los vectores lentivirales se pueden producir usando los plásmidos F/HN- SIV-CMV-HFVIII-V3, F/HN-SIV-hCEF-HFVIII-V3, F/HN-SIV-CMV-HFVIII-N6-co, y/o F/HN-SIV-hCEF-HFVIII-N6-co. La invención proporciona solo los plásmidos que contienen hCEF. HFVIII se refiere a FVIII humano. En este caso, el FVIII se expresa bajo el control del promotor de hCEF (como se reivindica) o de CMV. Puede describirse que estos vectores lentivirales comprenden F/HN-SIV-CMV-HFVIII-V3, F/HN-SIV-hCEF-HFVIII-V3, F/HN-SIV-CMV-HFVIII-N6-co y F/HN-SIV-hCEF- HFVIII-N6-co, respectivamente, ya que comprenden los elementos SIV F/HN, así como un casete de expresión que comprende *FVIII* bajo el control del promotor de hCEF/CMV. También se conocen productos de vectores virales producidos usando los plásmidos F/HN-SIV-CMV-HFVIII-V3, F/HN-SIV-hCEF-HFVIII-V3, F/HN-SIV-CMV-HFVIII-N6-co y/o F/HN-SIV-hCEF-HFVIII-N6-co como vGM126, vGM127, vGM142 y vGM129 (véase la figura 22). Estos vectores lentivirales son capaces de producir una expresión duradera, repetible y de alto nivel en las células de las vías respiratorias sin inducir una respuesta inmunitaria excesiva. En consecuencia, la invención proporciona un medio eficaz de terapia génica *in vivo*, por ejemplo, la transferencia del gen *FVIII* al pulmón o la nariz de un paciente para la producción de FVIII que después se secreta al sistema circulatorio (como se describe en esta invención). Por lo tanto, los vectores de la invención que comprenden el transgén *FVIII* pueden usarse en el tratamiento de la hemofilia u otras indicaciones como se describe en esta invención.

[0074] Los vectores lentivirales de la invención no contienen un intrón entre el promotor y el transgén. De manera similar, los plásmidos del genoma del vector de la invención (utilizados para generar dichos vectores lentivirales como se describe en esta invención) tampoco contienen un intrón entre el promotor y el transgén. Por lo tanto, la invención no proporciona, en una realización, ningún intrón entre el promotor de hCEF y las secuencias codificantes a expresar. En una realización preferida, la secuencia codificante a expresar es una secuencia de ácido nucleico de *CFTR, A1AT* y/o *FVIII*.

[0075] En una realización, los vectores de la invención comprenden el tracto central de polipurina (cPPT) y los elementos reguladores postranscripcionales del virus de la hepatitis Woodchuck (WPRE). En una realización, la secuencia de WPRE se proporciona por la SEQ ID NO: 8.

[0076] En una realización, el vector de la invención se usa para terapia génica. En una realización, la enfermedad a tratar es FQ. En otra realización de la invención, la enfermedad a tratar es la discinesia ciliar primaria (PCD). En una realización, el vector se usa para tratar la lesión pulmonar aguda. En una realización de la invención, la enfermedad a tratar es la deficiencia de proteína tensioactiva B (SP-B), deficiencia de alfa 1-antitripsina (A1AD), proteinosis alveolar pulmonar (PAP), enfermedad pulmonar obstructiva crónica (EPOC). En otra realización, la enfermedad es una afección inflamatoria, inmunitaria o metabólica.

[0077] La enfermedad a tratar puede ser una enfermedad cardiovascular o un trastorno sanguíneo, particularmente una deficiencia en la coagulación sanguínea. Por lo tanto, en algunas realizaciones, la enfermedad a tratar es hemofilia A, hemofilia B o hemofilia C, deficiencia del factor VII y/o enfermedad de von Willebrand. En otra realización más, la enfermedad a tratar es una enfermedad inflamatoria, una enfermedad infecciosa o una afección metabólica, tal como enfermedad de almacenamiento lisosomal.

[0078] Los ejemplos no limitantes de enfermedades que pueden tratarse usando la terapia con el gen A1AT según la presente invención incluyen diabetes tipo 1 y tipo 2, infarto agudo de miocardio, cardiopatía isquémica, artritis reumatoide, enfermedad inflamatoria intestinal, rechazo de trasplante, enfermedad de injerto contra huésped (GvH), esclerosis múltiple, enfermedad hepática, cirrosis, vasculitis e infecciones, tales como infecciones bacterianas y/o virales.

[0079] En un aspecto de la invención, el vector puede tratar eficazmente una enfermedad proporcionando un transgén para la corrección de la enfermedad. Por ejemplo, insertando una copia funcional del gen *CFTR* para mejorar o prevenir la enfermedad pulmonar en pacientes con FQ, independientemente de la mutación subyacente.

[0080] En otra realización de la invención, se proporciona un procedimiento de producción para los vectores lentivirales de la invención. En esta realización, el procedimiento de la invención es un procedimiento escalable compatible con GMP.

[0081] Por lo tanto, el procedimiento de la invención permite la generación de vectores F/HN purificados de alto título.

[0082] El procedimiento de la invención comprende las siguientes etapas:

(a) cultivar células en suspensión;
(b) transfectar las células con uno o más plásmidos;
(c) añadir una nucleasa;
(d) recoger los lentivirus;
(e) añadir tripsina; y
(f) purificación.

[0083] En una realización del procedimiento de la invención, el uno o más plásmidos proporcionan el genoma del vector, Gag-Pol, Rev, F y HN. Por lo tanto, puede haber cinco plásmidos para cada genoma del vector, Gag-Pol, Rev, F y HN, respectivamente. En el procedimiento de 5 plásmidos preferido de la invención, el plásmido del genoma del vector codifica todo el material genético que se empaqueta en el vector lentiviral final, incluido el transgén. Típicamente, solo una parte del material genético encontrado en el plásmido del genoma del vector termina en el virus. El plásmido del genoma del vector puede designarse en esta invención como "pDNA1". Los otros cuatro plásmidos están fabricando plásmidos que codifican las proteínas Gag-Pol, Rev, F y HN. Estos plásmidos se pueden designar "pDNA2a", "pDNA2b", "pDNA3a" y "pDNA3b", respectivamente.

[0084] En una realización de la invención, el lentivirus es VIS, tal como SIV1, preferentemente SIV-AGM. En una realización, las proteínas F y HN se derivan de un Paramixovirus, tal como el virus Sendai. En una realización, el plásmido del genoma del vector (pDNA1) comprende el transgén y el promotor del transgén.

[0085] En una realización específica relacionada con *CFTR*, los cinco plásmidos se caracterizan por las figuras 1A-1E, por lo tanto, pDNA1 es el plásmido pGM326 de la figura 1A, pDNA2a es el plásmido pGM299 de la figura 1B, pDNA2b es el plásmido pGM299 de la figura 1C, pDNA3a es el plásmido pGM301 de la figura 1D y pDNA3b es el plásmido pGM303 de la figura 1E. En esta realización, el vector lentiviral que contiene CFTR final puede denominarse vGM058 (véanse los Ejemplos). El vector vGM058 es una realización preferida de la invención.

[0086] En una realización relacionada con *A1AT*, los cinco plásmidos pueden caracterizarse por las figuras 15A (por lo tanto, el plásmido pDNA1 puede ser pGM407) y 1B-E (como anteriormente para la realización específica de CFTR).

[0087] En relación con *FVIII*, los cinco plásmidos pueden caracterizarse por las figuras 22C-F (por lo tanto, el plásmido pDNA1 puede ser pGM411, pGM412, pGM413 o pGM414) y 1BE. Según la invención, pDNA1 puede ser pGM413 o pGM414.

[0088] En estas realizaciones de la invención, el plásmido como se define en la figura 1A se representa por la SEQ ID NO: 1; el plásmido como se define en la figura 1B se representa por la SEQ ID NO: 2; el plásmido como se define en la figura 1C se representa por la SEQ ID NO: 3; el plásmido como se define en la figura 1D se representa por la SEQ ID NO: 4; el plásmido como se define en la figura 1E se representa por la SEQ ID NO: 5; el plásmido como se define en la figura 15A se representa por la SEQ ID NO: 9. Los plásmidos F/HN-SIV-CMV-HFVIII-V3, F/HN-SIV-hCEF-HFVIII-V3, F/HN-SIV-CMV-HFVIII-N6-co y/o F/HN-SIV-hCEF-HFVIII-N6-co como se definen en la figura 22B se representan por las SEQ ID NOs: 11 a 14, respectivamente.

[0089] En el procedimiento de 5 plásmidos de la invención, los cinco plásmidos contribuyen a la formación del vector lentiviral final. Durante la fabricación del vector lentiviral, el plásmido del genoma del vector (pDNA1) proporciona el potenciador/promotor, Psi, RRE, cPPT, mWPRE, SIN LTR, poliA de SV40 (véase la figura 1A), que son importantes para la fabricación viral. Usando pGM326 como un ejemplo no limitante de un pDNA1, el potenciador/promotor de CMV, poliA de SV40, colE1 Ori y KanR están involucrados en la fabricación del vector lentiviral de la invención (por ejemplo, vGM058), pero no se encuentran en el vector viral final. El RRE, cPPT (tracto central de polipurina), hCEF, soCFTR2 (transgén) y mWPRE de pGM326 se encuentran en el vector viral final. SIN LTR (repeticiones terminales largas, SIN/IN autoinactivante) y Psi (señal de empaquetamiento) se pueden encontrar en el vector viral final.

[0090] Para otros vectores lentivirales de la invención, los elementos correspondientes de los otros plásmidos del genoma del vector (pDNA1) son necesarios para la fabricación (pero no se encuentran en el vector final), o están presentes en el vector viral final.

[0091] Las proteínas F y HN de pDNA3a y pDNA3b (preferentemente proteínas Sendai F y HN) son importantes para la infección de células diana con el vector lentiviral final, es decir, para la entrada de células epiteliales de un paciente (típicamente células pulmonares o nasales como se describe en esta invención). Los productos de los plásmidos pDNA2a y pDNA2b son importantes para la transducción del virus, es decir, para insertar el ADN lentiviral en el genoma del huésped. El promotor, los elementos reguladores (tal como WPRE) y el transgén son importantes para la expresión del transgén dentro de la célula o células diana.

[0092] En una realización, las etapas (a)-(f) se realizan secuencialmente. En una realización, las células son células HEK293 o células 293T/17. En una realización, las células se cultivan en medios sin componentes animales (sin suero). En una realización, la transfección se realiza mediante el uso de PEIPro™. En una realización, la nucleasa es una endonucleasa, por ejemplo, Benzonase®. En una realización, la actividad de la tripsina se proporciona por una enzima recombinante libre de origen animal tal como TrypLE Select™.

[0093] En una realización de la invención, la adición de la nucleasa es en la fase previa a la recogida. En una realización alternativa, la adición de la nucleasa es en la fase posterior a la recogida. En otra realización, la adición de tripsina es en la fase previa a la recogida. En otra realización, la adición de la tripsina es en la fase posterior a la recogida.

[0094] En una realización, la etapa de purificación comprende una etapa de cromatografía. En esta realización, se usa la cromatografía de exclusión por tamaño (SEC) de modo mixto. En una realización, se usa la cromatografía de intercambio aniónico. En esta realización, no se usa gradiente de sal para la etapa de elución.

[0095] En una realización, este procedimiento se usa para producir los vectores lentivirales de la invención. En esta realización, el vector de la invención comprende el gen *CFTR, A1AT* y/o *FVIII*. En una realización alternativa, el vector de la invención comprende cualquiera de los genes mencionados anteriormente, o los genes que codifican las proteínas mencionadas anteriormente.

[0096] En una realización del procedimiento de la invención, se usa cualquier combinación de una o más de las construcciones plasmídicas específicas proporcionadas por las figuras 1A-1E, la figura 15A y/o la figura 22C-22F para proporcionar un vector de la invención, siempre que el vector resultante comprenda el promotor de hCEF.

[0097] La invención proporciona además los vectores lentivirales de la invención para su uso en un procedimiento de tratamiento de una enfermedad, comprendiendo el procedimiento administrar un vector lentiviral de la invención a un sujeto. En esta realización, la invención proporciona un vector lentiviral de la invención para su uso en el tratamiento de una enfermedad pulmonar. En una realización, la enfermedad es una enfermedad crónica. En una realización específica, se proporciona un procedimiento para tratar la FQ. En otras realizaciones, se proporciona un procedimiento para tratar la discinesia ciliar primaria (PCD), deficiencia de proteína tensioactiva B (SP-B), deficiencia de alfa 1-antitripsina (A1AD), proteinosis alveolar pulmonar (PAP), enfermedad pulmonar obstructiva crónica (EPOC). En otra realización, la enfermedad es una afección inflamatoria, inmunitaria o metabólica.

[0098] En otra realización, la enfermedad puede ser una enfermedad cardiovascular o un trastorno sanguíneo, particularmente una deficiencia en la coagulación de la sangre, tal como hemofilia A, hemofilia B, hemofilia C, deficiencia del factor VII y/o enfermedad de von Willebrand, una enfermedad inflamatoria, una enfermedad infecciosa o afección metabólica, tal como enfermedad de almacenamiento lisosomal.

[0099] La enfermedad puede ser diabetes tipo 1 y tipo 2, infarto agudo de miocardio, cardiopatía isquémica, artritis reumatoide, enfermedad inflamatoria intestinal, rechazo de trasplante, enfermedad de injerto contra huésped (GvH), esclerosis múltiple, enfermedad hepática, cirrosis, vasculitis e infecciones, tales como infecciones bacterianas y/o virales.

[0100] Los vectores lentivirales de la invención pueden administrarse en cualquier dosificación apropiada para lograr el efecto terapéutico deseado. Las dosis apropiadas pueden determinarse por un médico u otro profesional médico utilizando técnicas estándar y dentro del curso normal de su trabajo. Ejemplos no limitativos de dosificaciones adecuadas incluyen $1x10^8$ unidades de transducción (UT), $1x10^9$ UT, $1x10^{10}$ UT, $1x10^{11}$ UT o más.

[0101] La invención también proporciona composiciones que comprenden los vectores lentivirales descritos anteriormente, y un vehículo farmacéuticamente aceptable. Ejemplos no limitantes de vehículos farmacéuticamente aceptables incluyen agua, una solución salina y una solución salina tamponada con fosfato. Sin embargo, en algunas realizaciones, la composición está en forma liofilizada, en cuyo caso puede incluir un estabilizador, tal como albúmina sérica bovina (BSA). En algunas realizaciones, puede ser deseable formular la composición con un conservante, tal como tiomersal o azida de sodio, para facilitar el almacenamiento a largo plazo.

[0102] Los vectores de la invención pueden administrarse por cualquier ruta apropiada. Se puede desear dirigir las composiciones de la presente invención (como se ha descrito anteriormente) al sistema respiratorio de un sujeto. La transmisión eficiente de una composición terapéutica/profiláctica o medicamento al sitio de infección en el tracto

respiratorio puede lograrse mediante administración oral o intranasal, por ejemplo, como aerosoles (por ejemplo, aerosoles nasales), o por catéteres. Típicamente, los vectores lentivirales de la invención son estables en nebulizadores, catéteres y aerosoles, etc., clínicamente relevantes.

[0103] Las formulaciones para administración intranasal pueden estar en forma de gotas nasales o un aerosol nasal. Una formulación intranasal puede comprender gotas que tienen diámetros aproximados en el intervalo de 100-5000 µm, tales como 500-4000 µm, 1000-3000 µm o 100-1000 µm. Como alternativa, en términos de volumen, las gotas pueden estar en el intervalo de aproximadamente 0,001-100 µl, tal como 0,1-50 µl o 1,0-25 µl, o tal como 0,001-1 µl.

[0104] La formulación de aerosol puede adoptar la forma de un polvo, suspensión o solución. El tamaño de las partículas de aerosol es relevante para la capacidad de administración de un aerosol. Las partículas más pequeñas pueden viajar más abajo por las vías respiratorias hacia los alvéolos que las partículas más grandes. En una realización, las partículas de aerosol tienen una distribución de diámetro para facilitar la administración a lo largo de toda la longitud de los bronquios, bronquiolos y alvéolos. Como alternativa, la distribución del tamaño de partícula puede seleccionarse para dirigirse a una sección particular de las vías respiratorias, por ejemplo los alvéolos. En el caso de la administración en aerosol del medicamento, las partículas pueden tener diámetros en el intervalo aproximado de 0,1-50 µm, preferentemente 1-25 µm, más preferentemente 1-5 µm.

[0105] Las partículas de aerosol pueden ser para administrarse usando un nebulizador (por ejemplo, a través de la boca) o un pulverizador nasal. Una formulación en aerosol puede contener opcionalmente un propulsor y/o tensioactivo.

[0106] Como se usan en esta invención, los términos "secuencia de ácido nucleico" y "polinucleótido" se usan indistintamente y no implican ninguna restricción de longitud. Como se usan en esta invención, los términos "ácido nucleico" y "nucleótido" se usan indistintamente. Los términos "secuencia de ácido nucleico" y "polinucleótido" abarcan secuencias de ADN (incluyendo ADNc) y ARN. Los términos "transgén" y "gen" también se usan indistintamente y ambos términos abarcan fragmentos o variantes de los mismos que codifican la proteína diana.

[0107] Los transgenes de la presente invención incluyen secuencias de ácido nucleico que se han eliminado de su entorno natural, aislados de ADN recombinantes o clonados, y análogos sintetizados químicamente o análogos sintetizados biológicamente por sistemas heterólogos.

[0108] Los polinucleótidos de la presente invención pueden prepararse por cualquier medio conocido en la técnica. Por ejemplo, se pueden producir grandes cantidades de los polinucleótidos por replicación en una célula huésped adecuada. Los fragmentos de ADN naturales o sintéticos que codifican un fragmento deseado se incorporarán en construcciones de ácido nucleico recombinante, típicamente construcciones de ADN, capaces de introducirse y replicarse en una célula procariota o eucariota. Usualmente, las construcciones de ADN serán adecuadas para la replicación autónoma en un huésped unicelular, tal como levadura o bacterias, pero también pueden estar destinadas a la introducción e integración dentro del genoma de un insecto, mamífero, planta cultivados u otras líneas celulares eucariotas.

[0109] Los polinucleótidos de la presente invención también pueden producirse por síntesis química, por ejemplo, por el procedimiento de fosforamidita o el procedimiento de tri-éster, y pueden realizarse en sintetizadores de oligonucleótidos automatizados comerciales. Se puede obtener un fragmento bicatenario a partir del producto monocatenario de síntesis química ya sea sintetizando la hebra complementaria e hibridando la hebra en condiciones apropiadas, o añadiendo la hebra complementaria usando ADN polimerasa con una secuencia de cebador apropiada.

[0110] Cuando se aplica a una secuencia de ácido nucleico, el término "aislado" en el contexto de la presente invención representa que la secuencia de polinucleótidos se ha eliminado de su medio genético natural y, por lo tanto, está libre de otras secuencias codificantes externas o no deseadas (pero puede incluir regiones no traducidas de 5' y 3' de origen natural, tales como promotores y terminadores), y está en una forma adecuada para su uso dentro de sistemas de producción de proteínas genéticamente modificadas. Dichas moléculas aisladas son aquellas que están separadas de su entorno natural.

[0111] En vista de la degeneración del código genético, es posible una considerable variación de secuencia entre los polinucleótidos de la presente invención. Los codones degenerados que abarcan todos los codones posibles para un aminoácido dado se exponen a continuación:

Aminoácido	Codones	Codón degenerado
Cys	TGC TGT	TGY
Ser	AGC AGT TCA TCC TCG TCT	WSN
Thr	ACA ACC ACG ACT	ACN

(continuación)

Aminoácido	Codones	Codón degenerado
Pro	CCA CCC CCG CCT	CCN
Ala	GCA GCC GCG GCT	GCN
Gly	GGA GGC GGG GGT	GGN
Asn	AAC AAT	AAY
Asp	GAC GAT	GAY
Glu	GAA GAG	GAR
Gln	CAA CAG	CAR
His	CAC CAT	CAY
Arg	AGA AGG CGA CGC CGG CGT	MGN
Lys	AAA AAG	AAR
Met	ATG	ATG
Ile	ATA ATC ATT	ATH
Leu	CTA CTC CTG CTT TTA TTG	YTN
Val	GTA GTC GTG GTT	GTN
Phe	TTC TTT	TTY
Tyr	TAC TAT	TAY
Trp	TGG	TGG
Ter	TAA TAG TGA	TRR
Asn/Asp		RAY
Glu/Gln		SAR
Cualquiera		NNN

[0112] Un experto en la técnica apreciará que existe flexibilidad al determinar un codón degenerado, representativo de todos los codones posibles que codifican cada aminoácido. Por ejemplo, algunos polinucleótidos abarcados por la secuencia degenerada pueden codificar secuencias de aminoácidos variantes, pero un experto en la técnica puede identificar fácilmente dichas secuencias variantes haciendo referencia a las secuencias de aminoácidos de la presente invención.

[0113] Una secuencia de ácido nucleico "variante" tiene una homología sustancial o similitud sustancial con una secuencia de ácido nucleico de referencia (o un fragmento de la misma). Una secuencia de ácido nucleico o fragmento de la misma es "sustancialmente homóloga" (o "sustancialmente idéntica") a una secuencia de referencia si, cuando está alineada de manera óptima (con inserciones o deleciones de nucleótidos apropiadas) con el otro ácido nucleico (o su hebra complementaria), existe identidad de secuencia de nucleótidos en al menos aproximadamente el 70 %, 75 %, 80 %, 82, 84, 86, 88, 90, 92, 94, 96, 98 o el 99 % de las bases de nucleótidos. Los procedimientos para la determinación de homología de secuencias de ácido nucleico se conocen en la técnica.

[0114] Como alternativa, una secuencia de ácido nucleico "variante" es sustancialmente homóloga con (o sustancialmente idéntica a) una secuencia de referencia (o un fragmento de la misma) si la "variante" y la secuencia de referencia son capaces de hibridarse en condiciones de hibridación rigurosas (por ejemplo, altamente rigurosas). La hibridación de la secuencia de ácido nucleico se verá afectada por condiciones tales como la concentración de sal (por ejemplo, NaCl), la temperatura, o los disolventes orgánicos, además de la composición de la base, la longitud de las hebras complementarias, y el número de desajustes de bases de nucleótidos entre los ácidos nucleicos de hibridación, como apreciarán fácilmente los expertos en la técnica. Las condiciones de temperatura rigurosas se emplean preferentemente, y generalmente incluyen temperaturas superiores a 30 °C, típicamente en un exceso de 37 °C y preferentemente en un exceso de 45 °C. Las condiciones de sal rigurosas normalmente serán menores de 1000 mM, típicamente menores de 500 mM, y preferentemente menores de 200 mM. El pH es típicamente entre 7,0 y 8,3. La combinación de parámetros es mucho más importante que cualquier parámetro individual.

[0115] Los procedimientos para determinar la identidad de secuencia porcentual de ácido nucleico se conocen en la técnica. A modo de ejemplo, al evaluar la identidad de secuencia de ácido nucleico, una secuencia que tiene un número definido de nucleótidos contiguos puede alinearse con una secuencia de ácido nucleico (que tenga el mismo número de nucleótidos contiguos) de la porción correspondiente de una secuencia de ácido nucleico de la presente invención. Las herramientas conocidas en la técnica para determinar la identidad de secuencia porcentual de ácido nucleico incluyen Nucleotide BLAST.

[0116] Un experto en la materia aprecia que diferentes especies exhiben un "uso de codones preferenciales". Como se usa en esta invención, el término "uso de codones preferenciales" se refiere a los codones que se usan con mayor frecuencia en células de una especie determinada, favoreciendo así a uno o algunos representantes de los posibles codones que codifican cada aminoácido. Por ejemplo, el aminoácido treonina (Thr) puede estar codificado por ACA, ACC, ACG o ACT, pero en las células huésped de mamífero, ACC es el codón más comúnmente utilizado; en otras especies, pueden ser preferenciales diferentes codones. Los codones preferenciales para una especie de célula huésped particular se pueden introducir en los polinucleótidos de la presente invención mediante una diversidad de procedimientos conocidos en la técnica. La introducción de secuencias de codones preferenciales en el ADN recombinante puede mejorar, por ejemplo, la producción de la proteína al hacer que la traducción de proteínas sea más eficiente dentro de un tipo o especie celular particular.

[0117] Por lo tanto, en una realización de la invención, la secuencia de ácido nucleico es un codón optimizado para la expresión en una célula huésped.

[0118] Un "fragmento" de un polinucleótido de interés comprende una serie de nucleótidos consecutivos de la secuencia de dicho polinucleótido de longitud completa. A modo de ejemplo, un "fragmento" de un polinucleótido de interés puede comprender (o consistir en) al menos 30 nucleótidos consecutivos de la secuencia de dicho polinucleótido (por ejemplo, al menos 35, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800 850, 900, 950 o 1000 residuos de ácido nucleico consecutivos de dicho polinucleótido). Un fragmento puede incluir al menos un determinante antigénico y/o puede codificar al menos un epítopo antigénico del polipéptido de interés correspondiente. Típicamente, un fragmento, como se define en esta invención, conserva la misma función que el polinucleótido o polipéptido de longitud completa.

DESCRIPCIÓN DE LOS DIBUJOS

[0119] La presente invención se describirá ahora a modo de ejemplo solo con referencia a los dibujos adjuntos, en los que:

La figura 1 muestra plásmidos ejemplares utilizados en la invención. Las figuras 1A-1E muestran dibujos esquemáticos de plásmidos utilizados para la producción de los vectores de la invención. En una realización de la invención, la figura 1A proporciona una herramienta de la invención.

La figura 2 muestra la duración de la expresión del transgén F/HN-SIV en tejido nasal de ratón, que se perfundió con $4x10^8$ UT de F/HN-SIV-CMV-EGFP y expresión de EGFP se determinó en el número indicado de días después del tratamiento. Se muestra un control negativo en el que el tejido nasal se perfundió con diluyente de vector (PBS).

La figura 3 demuestra la consistencia de la expresión del transgén F/HN-SIV, mostrándola 1 después del tratamiento. La expresión de EGFP se determinó en 10 ratones independientes 360 días después del tratamiento.

La figura 4 muestra la distribución celular de la expresión del transgén F/HN-SIV. La expresión de EGFP se determinó en secciones histológicas de la cavidad nasal del ratón (2 mm desde la punta de la nariz) 30 días después del tratamiento. Las células EGFP positivas producen una señal punteada blanca.

La figura 5 muestra los tipos de células transducidas por el tratamiento con F/HN-SIV de la nariz del ratón. El 69 % de las células transducidas en la cavidad nasal del ratón eran células epiteliales respiratorias ciliadas. Otros tipos de células transducidas incluyeron células neuronales en el epitelio olfativo (21 %) y células escamosas (7 %).

La figura 6 muestra la administración repetida de F/HN-SIV a la nariz del ratón. Tejido nasal de ratón que se transdujo (como en la figura 1) con una dosis de F/HN-SIV-CMV-Lux o, dos dosis de F/HN-SIV-CMV-EGFP seguidas de una dosis de F/HN-SIV-CMV-Lux a intervalos de 28 días. Por lo tanto, la administración repetida de F/HN-SIV a la nariz del ratón no altera los niveles de expresión génica. La expresión transgénica se compara con una formulación líder de transferencia génica no viral (plásmido CMV-Lux complejado con GL67A).

La figura 7 muestra la transducción de cultivos de células respiratorias de la interfaz aire-líquido (ALI) humana. Los cultivos de ALI humana se transdujeron con F/HN-SIV-Lux a la multiplicidad de infección (MOI) indicada y se sometieron a formación de imágenes para determinar la expresión de Lux 5 días después del tratamiento.

La figura 8 demuestra que F/HN-SIV puede dirigir la expresión funcional de CFTR. Las células HEK293T se transdujeron con F/HN-SIV-CMV-EGFP-CFTR a 500 MOI y la actividad funcional de CFTR se determinó por eflujo de yoduro. F/HN-SIV-CMV-EGFP sirvió como control negativo.

La figura 9 muestra que F/HN-SIV transduce eficazmente células de pulmón primarias de oveja y ser humano y pulmón de ratón. La figura 9A muestra que la transducción de células de cepillado nasal humano (MOI 250) y cortes de pulmón humano y de oveja cultivados *ex vivo* (1E7 UT/corte) con F/HN-SIV-CMV-Lux da como resultado

una expresión sustancial de transgén luciferasa 24-48 horas después de la traducción. Las figuras 9B y 9C muestran la transducción de (~1x10^5) células de pulmón de FQ humana primarias cultivadas en la interfaz aire-líquido (FQ hALI) con vectores (3E7 UT) F/HN-SIV-soCFTR2 que contienen promotores transgénicos de CMV y hCEF. (B) Número de copias de vector (copias de ADN proviral por copia de ADN de CFTR endógeno) 6-8 días después de la transducción. (C) Nivel de expresión de ARNm de CFTR (% de VE: copias de ARNm de CFTR por copia de ARNm de CFTR endógeno x 100) 6-8 días después de la transducción. La línea horizontal en (C) representa un nivel de expresión objetivo del 5 % de VE, que se cree que representa el umbral terapéutico. Después de la administración *in vivo* de vectores F/HN-SIV-EGFPLux que contienen promotores de CMV, EF1aS y hCEF en forma de integrasa defectuosa (ID) o competente en integrasa (IC o sin etiqueta), se determinó la expresión transgénica de las células de las vías respiratorias en el epitelio murino nasal (D) y pulmonar (E) (n = 6-10/grupo). El curso temporal de la expresión transgénica de luciferasa se monitorizó mediante imágenes de bioluminiscencia repetidas *in vivo* y se normalizó a la dosis administrada. La figura 9F muestra imágenes de bioluminiscencia representativas después de la transducción murina *in vivo* el día 14 después de la transducción. La figura 9G muestra imágenes de bioluminiscencia representativas después de la transducción *in vitro* de hALI no FQ el día 5-6 después de la transducción. La figura 9H representa la expresión de EGFP 14 días después de la transducción en el epitelio nasal murino después de la administración de 1,6E8 UT de F/HN-SIV-hCEF-EGFPLux (vGM020). EGFP visualizada por inmunohistoquímica, núcleos visualizados por DAPI. La figura 9I muestra el curso temporal de la expresión trangénica de luciferasa en ALI no FQ que se monitorizó mediante imágenes de bioluminiscencia repetidas y se normalizó a la dosis administrada.

La figura 10 muestra que F/HN-SIV transduce eficientemente pulmón de oveja *in vivo*. La figura A muestra que, para modelar la administración de virus al pulmón de oveja, se instiló (3x100 µl de alícuotas durante ~5 minutos) acriflavina a una vía respiratoria proximal bajo visualización broncoscópica directa. La distribución de la acriflavina se puede apreciar por la coloración naranja de la vía respiratoria disecada en la autopsia. Cabe apreciar que la acriflavina se restringe en gran medida a las vías respiratorias conductoras y está ausente en las regiones alveolares. La flecha indica el sitio aproximado de instilación. Los números en la regla están en cm. La figura 10B es una representación esquemática del pulmón de oveja (centro/parte superior de la tráquea). El círculo verde representa la región en (A). En B, la línea indica el paso del broncoscopio para administrar 3x100 µl de alícuotas de 2,2E9 UT/ml (6,6E8 TU total) de F/HN SIV CMV-EGFPLux con respecto a n = 3 ovejas individuales (códigos de los animales T121, T156 y T251). Siete días después de la administración, se tomaron 5-6 bloques de muestras de tejido en la autopsia a intervalos de ~1 cm desde el sitio de instilación. Los bloques se dividieron en 2-3 muestras aproximadamente equivalentes y se analizaron para determinar la expresión transgénica mediante (C) ensayos de luciferasa normalizados con respecto al contenido de proteína; y (D) RT-PCR cuantitativa normalizada a niveles de ARNm de CFTR endógeno. La línea horizontal en (C) representa la actividad de luciferasa más alta observada en cualquier muestra tratada con un vector de transferencia génica no viral, y (D) el nivel de expresión objetivo del 5 % de VE (se cree que representa el umbral terapéutico).

La figura 11 representa la producción y purificación de los vectores F/HN SIV. Los vectores F/HN SIV se produjeron mediante transfección transitoria mediada por PEI de 5 plásmidos (pDNA) de células 293T cultivadas en suspensión a escala de 1 l en biorreactores WAVE controlados con pH, usando procedimientos escalables de la invención. Los vectores se clarificaron por filtración de profundidad/final (GE/Pall), los ácidos nucleicos contaminantes se eliminaron con Benzonase® (Merck), los vectores se activaron con TrypLE Select™ (Life Technology), se purificaron y se concentraron mediante cromatografía de membrana de intercambio aniónico (Pall) y filtración de flujo tangencial (Spectrum). Todos los recipientes, contenedores y columnas de procedimiento cumplían el cGMP de un solo uso. Todos los reactivos, excepto el ADN plasmídico, eran compatibles con cGMP sin origen animal. Se muestran datos de una diversidad de configuraciones vectoriales (promotor transgénico, transgén, estado de la integrasa). Las valoraciones físicas y funcionales se determinaron usando Q-PCR. (A) Valoración física de la recogida clarificada inicial y el producto purificado final. El rendimiento medio del procedimiento es ~44 %. (B) Valoración funcional del producto final. La valoración funcional media es ~2x10^9 UT/ml de producto y ~3x10^9 UT/l de volumen de biorreactor. Se excedieron los objetivos de productividad y rendimiento. Los vectores CFTR de menor rendimiento utilizan el promotor transgénico de CMV. Los vectores CFTR de mayor rendimiento utilizan promotores transgénicos de EF1aS y hCEF. (C) La relación partícula:infectividad del producto final está estrechamente agrupada y es similar a los valores de otros fabricantes de vectores de alta calidad (Oxford BioMedica, BlueBirdBio). La relación media de partícula:infectividad es de ~300. La consistencia del producto se ha logrado con el uso de procedimientos de "Diseño de experimentos" que respaldan la transición final a la aprobación del procedimiento de fabricación por parte de la agencia reguladora basada en QbD. (D) La valoración funcional del producto final está fuertemente correlacionada con la valoración física inicial, lo que indica condiciones de procedimiento no saturantes sin etapas de procedimiento limitantes de la concentración del vector, lo que sugiere que la ampliación de la purificación será eficiente.

La figura 12 muestra el perfil del sitio de inserción (IS) y la supervivencia de ratones transducidos. Las figuras 12A-F proporcionan una comparación de perfiles IS de pulmón de ratón transducido con F/HN-SIV y retina de ratón transducida con vector VSV-G-VIH (Bartholomae et al, Mol Ther (2011) 19:703-710). Los perfiles IS para (A, C, E) se derivaron de la secuenciación profunda del ADN pulmonar de dos ratones transducidos con F/HN-SIV, y para (B, D, F) de las secuencias IS retinianas de VSV-G-HIV (laboratorio Schmidt). (A, B) Sitios IS agregados (pulmón,

2862; retina, 262) representados en los cariogramas generados a través del buscador de genoma Ensembl. (C, D) Distancias IS a los sitios de inicio de la transcripción (TSS). Los números de IS en cada intervalo de distancia se muestran sobre barras. La suma excede el número total de IS analizados porque un IS típico se encuentra cerca de varios TSS. Los gráficos se generaron mediante el uso del buscador de genoma UCSC y el analizador de genoma GREAT (great.stanford.edu). (E, F) Comparación de QuickMap (www.gtsg.org) entre frecuencias de inserción aleatorias (rombo) y observadas (cuadrado) por cromosoma. (G) Supervivencia de ratones tratados con el vector progenitor F/HN-SIV (vector de virus fabricado según el procedimiento conocido usando células adherentes: línea negra, 24 meses de datos) y el vector de generación actual (GTC: línea gris oscuro, 8 meses de datos) en comparación con ratones tratados con tampón (línea gris claro, 24 meses de datos). Datos agregados de diversos experimentos con ratones tratados con tampón o ~1E7 UT de virus por inhalación nasal.

La figura 13 muestra la expresión transgénica respiratoria mediada por hCEF - usando la transferencia de genes lentivirus con un transgén rico en CpG - expresión transgénica (radiancia promedio p/s/cm^2/sr) frente días después de la dosis. Se observaron altos niveles de expresión del gen indicador de luciferasa de *Gaussia* en comparación con el control (pulmón sin tratar) tanto en el pulmón (cuadrado) como en la nariz (círculo) durante al menos 168 días después de la dosificación.

La figura 14 muestra el efecto de la transducción de organoides intestinales humanos con un vector lentiviral CFTR (vGM058) de la invención. El panel de la izquierda muestra que la inflamación inducida por forskolina se redujo significativamente (p <0,001) en los organoides transducidos con vGM058. En el panel de la derecha, las células A549 se transdujeron con vGM058 o un virus de control y la función de CFTR se cuantificó usando un ensayo de eflujo de yoduro radiactivo. Se detectaron niveles significativos (p <0,05) de eflujo de yoduro mediado por CFTR en células transducidas con vGM058.

La figura 15A muestra dibujos esquemáticos de un plásmido usado para la producción de los vectores *A1AT* de la invención. En una realización de la invención, la figura 15A proporciona una herramienta de la invención. La figura 15B muestra un plásmido de control que codifica el gen indicador de luciferasa de *Gaussia*.

La figura 16 muestra que F/HN-SIV transduce eficientemente células pulmonares primarias humanas en cultivo de ALI. En particular, la transducción de cultivos de ALI humanos da como resultado una expresión transgénica de luciferasa sustancial durante al menos 80 días después de la transducción. Cada punto representa el valor medio de URL/µl en los medios de n = 6 ALI en el punto temporal mostrado. Las barras verticales representan el error estándar de la media.

La figura 17 muestra la expresión génica después de la transducción de cortes de pulmón humano (n = 6 por grupo) con SIV1 hCEF-sogLux (figura 17A) y SIV1 hCEF-sohAAT (figura 17B). Se observaron altos niveles de expresión. Cada punto representa el valor medio de URL/µl en los medios de n = 6 cortes de pulmón en el punto temporal mostrado. Las barras verticales representan el error estándar de la media.

La figura 18 muestra la expresión a largo plazo (>12 meses) de la luciferasa de *Gaussia* después de la transferencia génica mediada por lentivirales (SIV1 hCEF-soGLux) *in vivo*. A: Homogeneizado de tejido pulmonar; B: fluido de lavado broncoalveolar (BAL); C: suero. Cada punto representa el valor medio de URL/µl en un grupo de animales (n = 5 o 6 por grupo) recogidos en el punto temporal mostrado. Las barras verticales representan el error estándar de la media.

La figura 19 muestra altos niveles de expresión de la expresión de A1AT después de la transferencia génica mediada por lentivirales (SIV1 hCEF-sohAAT) *in vivo*. Cada punto representa un animal. Las barras horizontales representan la mediana de cada grupo.

La figura 21 muestra el nivel de A1A1 en el fluido de revestimiento epitelial después de la transferencia génica mediada por lentivirales (SIV1 hCEF-sohAAT) *in vivo*. Transferencia génica mediada por lentivirales (SIV1 hCEF-sohAAT) *in vivo*.

La figura 22A muestra dibujos esquemáticos de construcciones de ADNc de *FVIII* usadas para la producción de los vectores FVIII. La figura 22B muestra vectores virales de la invención (con hCEF) y otros vectores FVIII. En una realización de la invención, la figura 22B proporciona una herramienta. Las figuras 22C-F muestran dibujos esquemáticos de plásmidos pDNA1 utilizados para la producción de los vectores *FVIII*. En una realización de la invención, las figuras 22C-F proporcionan herramientas. La invención se refiere a plásmidos/vectores/herramientas en los que FVIII se expresa bajo un promotor de hCEF.

La figura 23 muestra la eficiencia de transducción de HEK293T con vGM142 (SIV-F/HN-FVIII-N6-co) para el lote 1 (A) y el lote 2 (B). Los gráficos muestran la actividad de FVIII para cada MOI 48 y 72 horas después de la transducción. Cada símbolo representa un experimento independiente (n = 5-6 experimentos). La barra horizontal indica la media del grupo +/- DE. El análisis se realizó utilizando Anova unidireccional (GraphPad Prism) con múltiples comparaciones con el control no tratado ** p <0,01; *** p<0,001; **** p<0,0001.

La figura 24 muestra la evaluación de vGM142 en un sistema *in vivo* (modelo murino). A: pulmón; B: fluido BAL; C: plasma. Los gráficos muestran el nivel de hFVIII (como porcentaje del nivel normal) frente a los diferentes tratamientos para los Grupos 1 a 3 (Grupos 1 y 2 - 10 días de tratamiento, 3 dosis/semana de 100 µl de vector por ratón (la cantidad total de dosis fue igual a 3 por animal); Grupo 3 - 28 días de tratamiento, 3 dosis/semana de 100 µl de vector por ratón (la cantidad total de dosis fue igual a 12 por animal)). Cada símbolo representa un ratón individual (Grupo 1 (n = 4) tratado con una dosis de vector total de $1,4x10^6$ UTT/ratón; Grupo 2 (n = 3) tratado con una dosis de vector total de $1,57x10^8$ UTT/ratón y Grupo 3 (n = 4) tratado con una dosis de vector total de $3,36x10^8$ UTT/ratón). La barra horizontal indica los niveles medios de FVIII:Ag +/- DE. El análisis se realizó utilizando Anova unidireccional (GraphPad Prism) con comparaciones múltiples entre los grupos tratados **** p <0,0001.

EJEMPLOS

[0120] La invención se describe ahora con referencia a los ejemplos a continuación. Estos no limitan el alcance de la invención, y un experto en la técnica apreciará que pueden usarse equivalentes adecuados dentro del alcance de la presente invención. Por lo tanto, los Ejemplos pueden considerarse partes componentes de la invención, y los aspectos individuales descritos en los mismos pueden considerarse como descritos independientemente, o en cualquier combinación.

[0121] Solo los vectores que comprenden un promotor hCEF y plásmidos/herramientas para la producción de dichos vectores están incluidos por la invención. La descripción relacionada con vectores que comprenden otros promotores y plásmidos del genoma del vector/otras herramientas específicamente para la producción de vectores que no comprenden un promotor hCEF se incluye solo con fines comparativos.

Ejemplo 1: Cultivo celular

[0122] Las células HEK293T, Freestyle 293F (Life Technologies, Paisley, Reino Unido) y 293T/17 (CRL-11268; ATCC, Manassas, VA) se mantuvieron en medio de Eagle mínimo de Dulbecco (Invitrogen, Carlsbad, CA) que contenía suero bovino fetal al 10 % y complementado con penicilina (100 U/ml) y estreptomicina (100 µg/ml) o medio de expresión Freestyle™ 293 (Life Technologies).

Ejemplo 2: Construcción de plásmidos

[0123] pCAGGS-Fct4 y pCAGGS-SIVct+HN se construyeron de la siguiente manera:

(i) El plásmido SIVct/HN contiene el gen que codifica la cola citoplasmática de SIVagm TMP (invertido) fusionado con el ectodominio y las regiones transmembrana de la proteína SeV HN. Se sintetizaron tres oligonucleótidos: Par 1, 5'-TCGAGATGTGGTCTGAGTTAAAAATCAGGAGCAACGACGGAGGTGAAGGACCAGA CGCCAACGACCC-3' (SEQ ID NO: 18) y 5'-CCGGGGGTCGTTGGCGTCTGGTCCTTCACCTCCGTCGTTGCTCCTGATTTTTAACT CAGACCACATC-3' (SEQ ID NO: 19); par 2, 5'-CCGGGGAAAGGGGGTGCAACACATCCATATCCAGCCATCTCTACCTGTTTATGGAC AGA-3' (SEQ ID NO: 20) y 5'-ACCCTCTGTCCATAAACAGGTAGAGATGGCTGGATATGGATGTGTTGCACCCCTTT CC-3' (SEQ ID NO: 21); y par 3, 5'-GGGTTAGGTGGTTGCTGATTCTCTCATTCACCCAGTGGG-3' (SEQ ID NO: 22) y 5'-GATCCCCACTGGGTGAATGAGAGAATCAGCAACCACCTA-3' (SEQ ID NO: 23). Estos pares de oligonucleótidos se hibridaron y se clonaron en los sitios *Xho*I y *Bam*HI de pBluescript KS+ (Stratagene) para producir pKS+SIVct. pCAGGS-SIVct/HN se construyó clonando el fragmento *Xho*I-*Dra*III de 160 pb de pKS+SIVct y un fragmento *Dra*III-*Bsu*36I de 1,5 kbp de pCAGGS-HN, que porta el gen HN natural (HNwt), en el sitio *Xho*I del vector pCAGGS, a los sitios *Xho*I y *Bsu*36I de pCAGGS. Este plásmido se construyó de modo que la cola citoplasmática de la proteína HN se reemplazara con la cola citoplasmática de SIVagm TMP.

[0124] Para la construcción de pCAGGS-SIVct+HN, los genes que codificaban la cola citoplasmática de SIVagm TMP y el extremo N de la proteína HN se amplificaron primero por PCR a partir de pCAGGS-SIVct/HN con el par de cebadores 5'-GAGACTCGAGATGTGGTCTGAGTTAAAAATCAGG-3' (SEQ ID NO: 24) y 5'-AGAGGTAGACCAGTACGAGTCACGTTTGCCCCTATCACCATCCCTAACCCTCTGTC ATAAAC-3' (SEQ ID NO: 25). El fragmento de PCR resultante se clonó en los sitios *Xho*I y *Acc*I de pKS+SIVct para generar pKS+SIVct-H. Después, un fragmento *Xho*I-*Dra*III de pKS+SIVct-H y un fragmento *Dra*III-*Bsu*36I de pCAGGS-HN se clonaron en los sitios *Xho*I y *Bsu*36I de pCAGGS para proporcionar pCAGGS-SIVct+HN.

[0125] Las secuencias de cPPT y WPRE se insertaron en el plásmido de transferencia génica derivado de VIS. Un ejemplo de la secuencia de WPRE utilizada se proporciona en la SEQ ID NO: 8.

[0126] El plásmido pGM101 contiene el origen colE1 de replicación, gen de resistencia a kanamicina y promotor y se creó por síntesis génica sintética (GeneArt, Ratisbona, Alemania; ahora LifeTechnologies Ltd).

[0127] Las secuencias híbridas CMV/SIV R U5 LTR, Gag parcial, RRE, cPPT, SIN U3 y R de pBS/CG2-Rc/s-

CMV-D U (Nakajima et al. 2000 Human Gene Therapy 11:1863) se amplificaron mediante PCR y se insertaron junto con la secuencia potenciadora/promotora de hCEF amplificada mediante PCR a partir de pGM169 (Hyde et al. Nature Biotechnology 26:549) y el ADNc de soCFTR2 aislado de pGM169 en un fragmento de enzima de restricción NheI-ApaI en pGM101 para crear pGM326.

[0128] El potenciador de CMV/promotor de la beta actina de pollo junto con secuencias de exón/intrón asociadas, secuencias SIV GagPol y RRE y poliA/origen de replicación de SV40 se amplificaron mediante PCR de pCAGGS/Sagm-gtr (Nakajima et al. 2000 Human Gene Therapy 11:1863) para crear pGM297. El potenciador/promotor de CMV junto con secuencias de exón/intrón asociadas y la secuencia de poliA de SV40 de pCI (Promega, Madison, WI, EE.UU.) se aislaron en un fragmento de enzima de restricción BgIII-BamHI y la secuencia SIV Rev derivada de pCAGGS/Sagm-gtr amplificada por PCR se insertó en pGM101 para crear pGM299.

[0129] El potenciador de CMV/promotor de la beta actina de pollo junto con secuencias de exón/intrón asociadas, el ADNc de Fct4 y el poliA/origen de SV40 de pCAGGS-Fct4 se aislaron en un fragmento de enzima de restricción SalI-HindIII mediante una combinación de síntesis génica, PCR y recombinación de fragmentos de enzimas de restricción y se insertaron en pGM101 para crear pGM301.

[0130] El potenciador de CMV/promotor de la beta actina de pollo junto con secuencias de exón/intrón asociadas, el ADNc de SIVct+HN y el poliA/origen de SV40 de pCAGGS-SIVct+HN se aislaron en un fragmento de enzima de restricción SalI-HindIII mediante una combinación de síntesis génica, PCR y recombinación de fragmentos de enzimas de restricción y se insertaron en pGM101 para crear pGM303.

[0131] Otros plásmidos pGM de la invención se elaboraron utilizando técnicas estándar y según la descripción anterior.

[0132] A través de estos enfoques de ensamblaje de ADN plasmídico, las enzimas de restricción y las polimerasas de PCR se suministraron por New England Biolabs (Ipswich, MA, EE.UU.), y los reactivos de purificación de ADN se suministraron por Qiagen (Limburgo, Países Bajos).

Ejemplo 3: Producción del vector SIV

Sistema de cuatro plásmidos:

[0133] El vector SIV auto-inactivador defectuoso de replicación se construyó con modificaciones secundarias. En resumen, el vector SIV pseudotipado con SeV-F/HN se produjo mediante la transfección de células 293T/17 (placas de cultivo de 15 cm de diámetro) con cuatro plásmidos complejados con reactivos Lipofectamine/Plus (Invitrogen) según las recomendaciones del fabricante [Plásmido 1: 10 µg de plásmido de transferencia derivado de VIS que lleva una GFP, un gen indicador de luciferasa (lux), o una construcción de fusión GFP-CFTR, Plásmido 2: 3 µg de plásmido de empaquetamiento, Plásmido 3: 2 µg de pCAGGS-Fct4, Plásmido 4: 2 µg de pCAGGS-SIVct+HN; las figuras 1A-E muestran dibujos esquemáticos de plásmidos usados para la producción de vectores de la invención]. El vector SIV pseudotipado con VSV-G se produjo usando un protocolo similar, pero se usó un plásmido pVSV-G (2 µg; Clontech, Mountain View, CA) en lugar de pCAGGS-Fct4 y pCAGGS-SIVct+HN. 12 horas después de la transfección, el medio de cultivo se reemplazó con 30 ml de medio Eagle modificado por Dulbecco sin suero que contenía 5 mmol/l de butirato de sodio. El butirato de sodio estimula la producción del vector para inhibir la histona desacetilasa. El sobrenadante de cultivo que contenía el vector SIV se recogió 48 horas después de la transfección, se filtró a través de una membrana de filtro de 0,45 µm, y se concentró adicionalmente por centrifugación a alta velocidad (20.000 g durante 4 horas a 4°C, rotor Avanti JA18; Beckman Coulter, Brea, CA). Los sedimentos vectoriales se suspendieron en PBS (Invitrogen) a una concentración de 100 a 200 veces y se almacenaron a -80°C.

Sistema de cinco plásmidos (preferido):

[0134] El vector SIV pseudotipado con SeV-F/HN se produjo mediante la transfección de células HEK293T o 293T/17 cultivadas en medio de expresión FreeStyle™ 293 con una mezcla de cinco plásmidos con las siguientes características: pDNA1 (por ejemplo pGM326; figura 1A) codifica el ARNm del vector lentiviral; pDNA2a (por ejemplo pGM297; figura 1B) codifica las proteínas Gag y Pol de VIS; pDNA2b (por ejemplo, pGM299: figura 1C) codifica las proteínas Rev de VIS; pDNA3a (por ejemplo pGM301; figura 1D) codifica la proteína Fct4 derivada del virus Sendai [Kobayashi et al., 2003 J. Virol. 77:2607]; y pDNA3b (por ejemplo, pGM303; figura 1E) codifica el SIVct+HN derivado del virus Sendai [Kobayashi et al., 2003 J. Virol. 77:2607] complejado con PEIpro (Polyplus, Illkirch, Francia).

[0135] El medio de cultivo celular se complementó 12-24 horas después de la transfección con butirato de sodio. El butirato de sodio estimula la producción vectorial a través de la inhibición de la histona desacetilasa, dando como resultado un aumento de la expresión de los componentes de la proteína de fusión del virus VIS y Sendai codificados por los cinco plásmidos. El medio de cultivo celular se complementó 44-52 horas y/o 68-76 horas después de la transfección con 5 unidades/ml de Benzonasa Nucleasa (Merck Millipore, Nottingham, Reino Unido). El sobrenadante de cultivo que contenía el vector SIV se recogió 68-76,5 horas después de la transfección, y se clarificó

por filtración a través de una membrana de 0,45 µm. El vector SIV se trata mediante digestión con una proteasa que contiene actividad de tripsina, por ejemplo, una enzima recombinante libre de origen animal tal como TrypLE Select™. Posteriormente, el vector SIV típicamente se purifica y se concentra adicionalmente mediante cromatografía de intercambio aniónico y/o filtración/ultrafiltración/diafiltración de flujo tangencial.

[0136] Este mismo procedimiento se usó para generar vectores lentivirales que comprendían los transgenes *A1AT* y *FVIII*, reemplazando los plásmidos de la figura 15A y la figura 22B a los de la figura 1A para proporcionar el transgén apropiado (véanse los Ejemplos 15 y 20).

Ejemplo 4: Valoración vectorial

Procedimiento 1:

[0137] La valoración de partículas se determinó usando PCR de transcriptasa inversa en tiempo real. El ARN viral se purificó utilizando un mini-kit de ARN viral QIAamp (QIAGEN, Strasse, Alemania), y se sometió a transcripción inversa utilizando Superscript II (Invitrogen). El sistema de PCR de la sonda QuantiTect (QIAGEN) y los cebadores para amplificar 131 nucleótidos (pb) que abarcan la secuencia de WPRE (cebador directo: 5'-ggatacgctgctttaatgcc-3', cebador inverso: 5'-acgccacgttgcctgacaac-3') se usaron según el protocolo del fabricante en un sistema detector de secuencias ABI PRISM 7700 (PE Applied Biosystems, Foster City, CA). Se usó ADN plasmídico de transferencia génica de VIS (3×10^4 a 2×10^6 moléculas) como estándar.

[0138] Las unidades de transducción (UT/ml) se determinaron transduciendo células 293T con diluciones en serie de solución madre vectorial y cuantificación de células transducidas por fluorescencia de GFP (para F/HN-SIV-GFP y VSV-G-SIV-GFP) o tinción con anticuerpo anti-luciferasa (para F/HN-SIV-lux).

Procedimiento 2 (preferido):

[0139] La valoración de partículas (VP/ml) se determinó típicamente usando PCR de transcriptasa inversa en tiempo real. El ARN viral se purificó usando un mini-kit de ARN viral QIAamp (QIAGEN, Strasse, Alemania), y se sometió a transcripción inversa utilizando transcriptasa inversa (Life Technologies).

[0140] Sistema de PCR cuantitativa TaqMan (Life Technologies) que usa cebadores que amplifican una porción de la secuencia de WPRE en un sistema detector de secuencias ABI PRISM 7700 (Life Technologies). Se usaron moléculas de ARN de WPRE transcritas *in vitro* como estándares cuantitativos.

[0141] Las unidades de transducción (UT/ml) se determinaron transduciendo células 293T/17 o Freestyle 293F con diluciones en serie del vector SIV y cuantificación de ADN provirus que contenía WPRE por el sistema de PCR cuantitativa TaqMan (Life Technologies) usando cebadores que amplificaron una porción de la secuencia de WPRE en un sistema detector de secuencias ABI PRISM 7700 (Life Technologies). Se usaron moléculas de ADN plasmídico que contenían secuencias de WPRE como estándares cuantitativos.

Ejemplo 5: Generación de cultivos de tEC enriquecidos con células basales

[0142] Se aislaron células epiteliales traqueales murinas (tEC) de la siguiente manera. Se sacrificaron ratones C57BL/6N y se extirparon las tráqueas desde la laringe hasta las ramificaciones bronquiales principales usando instrumentos quirúrgicos estériles. Los tejidos se colocaron en un tubo que contenía medio F-12 de Ham frío con 100 U/ml de penicilina (P), 100 µg/ml de estreptomicina (S) y 2,5 mg/ml de anfotericina B (A) (medio F12 de Ham/PSA) y se mantuvieron en hielo. En una campana de cultivo de tejido estéril, se limpiaron las tráqueas de los músculos adherentes y el tejido conectivo, se cortaron longitudinalmente para exponer el epitelio respiratorio interno, y se colocaron en una solución de pronasa al 0,15 % en medio F-12 (~5 ml en un tubo de 15 ml). La digestión tisular se realizó durante una noche (15-18 h) a 4 °C. Para bloquear la reacción enzimática, se añadió suero bovino fetal al 10 % (FBS) a la digestión tisular. Después de invertir suavemente el tubo para separar más células, las tráqueas se colocaron en un tubo nuevo que contenía una solución de FBS al 10 %/F-12 de Ham/PS, y se invirtieron como anteriormente. Esta etapa se repitió dos veces más. El contenido de los cuatro tubos se combinó y se centrifugó a 500 g durante 10 minutos a 4 °C. El sedimento se resuspendió en una solución de DNasa (0,5 mg/ml de DNasa pancreática en bruto más 10 mg/ml de BSA en una solución de FBS/F-12 de Ham/PS, aproximadamente 200 µl/tráquea), se incubó en hielo durante 5 minutos, y se centrifugó como anteriormente.

[0143] Después de eliminar el sobrenadante, las tEC se resuspendieron en medio dirigido a células progenitoras (PCT) (CnT-17, CELLnTEC, Berna, Suiza), una formulación libre de antibióticos y antimicóticos diseñada específicamente para el aislamiento y la proliferación de células progenitoras de las vías respiratorias humanas y de ratón. A continuación, las tEC se colocaron en una placa de cultivo tisular Primaria (Becton Dickinson Labwere, Franklin Lakes, NJ, EE.UU.) y se incubaron durante 3-4 horas en CO_2 al 5 % a 37 °C. Las células no adherentes se recogieron y se centrifugaron a 500 g durante 5 minutos a 4 °C y se contaron en un hemocitómetro. Para generar cultivos enriquecidos con células basales, se suspendieron tEC en medio PCT y se sembraron en una placa Nunclon™Δ (Nunc

A/S, Roskilde, Dinamarca), recubierta con 50 µg/ml de colágeno de cola de rata tipo 1 a una densidad de siembra recomendada de $4x10^3$ células/cm^2. Las tEC también se cultivaron en un medio básico de control, que contenía DMEM/F12 de Ham complementado con L-glutamina (4 mM), HEPES (15 mM) y NaHCO3 (3,4 mM). Las placas se incubaron a 37 °C con CO_2 al 5 %. Para determinar la proporción de células basales en la población de tEC antes y después de la expansión en medio PCT, las preparaciones de cytospin se tiñeron con el anticuerpo anti-KRT5 y el anticuerpo secundario apropiado.

[0144] El conjunto de tEC producido comprendía tanto células completamente diferenciadas (células claras y ciliadas) como células basales. Una vez sembradas en membranas permeables, las tEC pudieron generar un sistema de cultivo de interfaz aire-líquido (ALI) como resultado de la proliferación y rediferenciación de células basales en células secretoras y ciliadas. Para establecer un protocolo alternativo y rápido para la expansión de células basales, se evaluaron cultivos de tEC bidimensionales (2D) utilizando un medio patentado dirigido a células progenitoras (PCT) disponible comercialmente, formulado específicamente para soportar la proliferación de células progenitoras de las vías respiratorias mientras las mantiene en un estado indiferenciado. Como control negativo, las tEC se expusieron a una formulación de medio básica sin adición de factores de crecimiento específicos. Las tEC sembradas en superficies de plástico recubiertas de colágeno ($4x10^3$ células/cm^2) y expuestas al medio PCT pudieron crecer rápidamente y confluir en 5-8 días, mientras que las tEC expuestas al medio de control deficiente en factor de crecimiento no pudieron adherirse ni propagarse. Para establecer si el uso del medio PCT dio como resultado un enriquecimiento sustancial de la población de células basales, se recogieron tEC a aproximadamente un 80 % de confluencia (n = 6 pocillos), se fijaron y trataron con un anticuerpo anti-queratina 5 (Krt5), un marcador específico de células basales. Se utilizaron tEC recién aisladas como controles (n = 3 preparaciones únicas). La proporción de células basales Krt5 positivas después de la expansión en medio PCT fue mayor (78 ± 1,4 %) que en los grupos recién aislados de tEC (33 ± 0,6 %), lo que demuestra que las células basales de las vías respiratorias murinas pueden expandirse selectiva y rápidamente desde un grupo mixto de tEC utilizando un medio comercial.

Ejemplo 6: Transducción *ex vivo* de cultivos de tEC enriquecidos con células basales con F/HN-SIV-GFP

[0145] Para determinar si el vector F/HN-SIV puede transducir eficazmente células basales *ex vivo*, las tEC preparadas en el Ejemplo 5 se cultivaron hasta aproximadamente el 70 % de confluencia durante 7 días en medio PCT y se transdujeron con F/HN-SIV que llevaba un gen indicador de proteína verde fluorescente (F/HN-SIV-GFP) a una MOI 100 y se incubaron a 37 °C con CO_2 al 5 % durante 3 días. Las tEC derivadas de animales genéticamente intactos y transgénicos GFP se cultivaron en las mismas condiciones y se usaron como grupos de control negativo (sin transducción viral) y positivo, respectivamente (n = 3-6 pocillos/grupo).

[0146] Para cuantificar la proporción de células GFP positivas, los cultivos de tEC enriquecidos con células basales se separaron con la enzima accutase (CELLnTEC), se resuspendieron en PBS/BSA al 1 % y se sometieron a análisis FACS, contando un promedio de 20,277 ± 2,478 células/grupo. El vector F/HN-SIV transdujo el 26 % ± 0,9 % de las tEC enriquecidas con células basales ($p < 0,0001$ en comparación con los controles no transducidos).

[0147] Para evaluar si las células que expresan GFP transducidas eran células basales, tres días después de la infección las células se tiñeron doblemente con anticuerpos contra Krt5 y GFP. La tinción de inmunofluorescencia de células cultivadas mostró que aproximadamente el 40 % de las células que expresaban Krt5 también expresaron el gen indicador GFP, lo que demuestra que el vector F/HN-SIV puede transducir células basales progenitoras *ex vivo*.

Ejemplo 7: Administración *in vivo* a la nariz del ratón

[0148] Se usaron ratones C57BL/6N (hembra, 6-8 semanas). Los ratones se anestesiaron, se colocaron horizontalmente sobre sus espaldas sobre una placa calentada, y se insertó un catéter delgado (<0,5 mm de diámetro externo) ~2,5 mm desde la punta de la nariz hasta la fosa nasal izquierda. Usando una bomba de jeringa (Cole-Parmer, Vernon Hills, IL), a continuación, el vector (100 µl) se perfundió lentamente sobre el epitelio nasal (1,3 µl/min) durante 75 minutos. A pesar de la perfusión del virus en la fosa nasal izquierda, se observó rutinariamente la transfección en las fosas nasales izquierda y derecha, lo que se debe a la dispersión de las soluciones en toda la cavidad nasal. Se usaron como controles ratones transducidos con PBS y VSV-G-SIV preacondicionados con lisofosfatidilcolina al 1 % como se describe por Limberis et al., 2002. En los puntos de tiempo indicados (3-360 días después de la transducción), los ratones fueron sacrificados para visualizar la expresión de GFP. Como se muestra en la figura 2, se observó la expresión de GFP durante al menos 449 días después de la transducción, mientras que el control negativo no mostró expresión de GFP. Como se muestra en la figura 3, la expresión transgénica con el vector F/NH-SIV fue consistente, con expresión observable de GFP al menos 360 días después de la transducción en 10 ratones ensayados independientemente.

[0149] De manera similar, como se muestra en la figura 13, la transducción con un vector F/HN-SIV de la invención que comprende un promotor hCEF dio como resultado la expresión a largo plazo de un gen indicador rico en CpG (luciferasa). Se observaron altos niveles de expresión en relación con el control tanto en el pulmón como en la nariz durante al menos 169 días después de la transducción.

[0150] En los experimentos de administración repetida, se transdujeron grupos de ratones con una dosis de F/HN-SIV-lux (grupo de dosis única), o con dos dosis de F/HN-SIV-GFP (día 0, día 28), seguido de F/HN-SIV-lux el día 56 (grupo de dosis repetidas). Es importante destacar que los ratones que recibieron F/HN-SIV-lux (grupo de dosis única) y F/HN-SIV-lux el día 56 (grupo de dosis repetidas) tenían una edad similar y se transdujeron al mismo tiempo. La expresión génica se analizó 30 días después de la administración de F/HN-SIV-lux. A modo de comparación, los ratones se transfectaron con el lípido catiónico GL67A complejado con un gen indicador de luciferasa como se ha descrito previamente (Griesenbach, U. et al., Methods Mol Biol. 2008; 433:229-42) y la expresión de luciferasa se midió 2 días después de la transfección.

[0151] Como se muestra en la figura 6, la administración repetida de F/HN-SIV a la nariz del ratón no altera los niveles de expresión génica. La expresión transgénica se compara con una formulación líder de transferencia génica no viral (plásmido CMV-Lux complejado con GL67A).

[0152] El perfil del sitio de inserción se realizó en ratones transducidos, y se investigó el tiempo de supervivencia como se expone en la descripción de la figura 12 anterior. No se observó que la transducción usando el vector F/HN-SIV de la invención tuviera ningún efecto adverso sobre la supervivencia del ratón en comparación con un vector F/HN-SIV existente o control negativo (solo tampón) (véase la figura 12G).

Ejemplo 8: Regeneración inducida de células epiteliales nasales por tratamiento con polidocanol

[0153] Las células epiteliales nasales se extirparon mediante tratamiento con polidocanol según el procedimiento descrito (Borthwick et al., Am J Respir Cell Mol Biol. junio de 2001; 24(6):662-70), con alguna modificación. En resumen, los ratones se anestesiaron y se les administró 10 µl de polidocanol (2 %) por la nariz en forma de bolo mediante "olfateo nasal". Para confirmar la extirpación y la regeneración de las células epiteliales nasales, se perfundió el tejido nasal con 10 µl de polidocanol al 2 % (vol/vol en PBS) (éter mono-dodecílico de nonaetilenglicol; SIGMA, St Louis, MO) y se realizó un análisis histológico 24 horas y 7 días después del tratamiento (n = 3/grupo).

[0154] Para analizar la transducción de posibles células progenitoras o madre, primero se administró el vector F/HN-SIV-GFP (4×10^8 UT/ratón) al epitelio nasal del ratón. Siete días después de la transducción, el tejido nasal se perfundió con 10 µl de polidocanol al 2 % (vol/vol en PBS), y este tratamiento se repitió de nuevo 3 semanas después. Se analizaron las secciones histológicas 58 días después de la administración del vector (30 días después del último tratamiento con polidocanol).

Ejemplo 9: Imágenes bioluminiscentes

[0155] Se inyectó por vía intraperitoneal a los ratones 150 mg/kg de D-luciferina (Xenogen, Alameda, CA) 10 minutos antes de la formación de imágenes y se anestesiaron con isoflurano. La bioluminiscencia (fotones/s/cm2/sr) de los ratones vivos se midió usando un sistema IVIS50 (Xenogen) a una agrupación de 4 durante 10 minutos, usando el programa de software Living Image (Xenogen). Para la localización anatómica, se generó una imagen de pseudocolor que representaba la intensidad de la luz (azul: menos intenso, rojo: más intenso) utilizando el software Living Image y se superpuso sobre la imagen de referencia en escala de grises. Para cuantificar la bioluminiscencia en la nariz, se midió la emisión de fotones en un área definida (recuadro rojo) marcando un área estandarizada para la cuantificación. El tamaño del recuadro rojo se mantuvo constante y se colocó sobre las cabezas de los animales como se indica en la figura. Es importante destacar que las áreas se marcaron usando la imagen de referencia en escala de grises para evitar sesgos.

Ejemplo 10: Preparación de tejidos para la evaluación histológica de la expresión de GFP y/o la detección de células basales

[0156] Se sacrificaron los ratones y se retiró la piel. Se cortó la cabeza a la altura de los ojos y se retiraron cuidadosamente la piel, la mandíbula, la lengua y la punta suave de la nariz. Para obtener imágenes *in situ* de la expresión de GFP en la cavidad nasal, se detectó fluorescencia de GFP utilizando microscopía estereoscópica de fluorescencia (Leica, Ernst Leitz Optische Werke, Alemania). Posteriormente, el tejido se fijó en paraformaldehído al 4 % (pH 7,4) durante una noche a temperatura ambiente y después se sumergió en EDTA al 20 % (pH 7,5 durante 5 días) para la descalcificación. La solución de EDTA se cambió al menos cada dos días. Después de la descalcificación, el tejido se incubó en sacarosa al 15 % durante una noche a temperatura ambiente y después se incorporó en Tissue Mount (Chiba Medical, Soka, Japón). Se cortaron secciones de diez micrómetros en seis posiciones diferentes en cada cabeza de ratón (~0-6 mm desde la punta del hueso nasal). Se observó la expresión de GFP usando un microscopio fluorescente (Leica). La cuantificación e identificación de los tipos de células se realizaron en seis niveles por ratón utilizando un objetivo x40 o x63. La exposición prolongada a la imagen fue necesaria para capturar la estructura del epitelio nasal mediante microscopía fluorescente. Esto condujo a la saturación de píxeles de las células GFP positivas y provocó que las células GFP positivas parecieran casi blancas en lugar del aspecto verde común que nosotros, y otros, observamos con un aumento mayor.

[0157] La distribución celular de la expresión del transgén F/NH-SIV se investigó en secciones histológicas. Específicamente, la expresión de EGFP se determinó en secciones histológicas de la cavidad nasal del ratón (2 mm desde la punta de la nariz) 30 días después del tratamiento. La figura 4 muestra la ubicación de la expresión de EGFP (figura 4, señal punteada blanca).

[0158] La figura 5 muestra los tipos de células transducidas por el tratamiento con el transgén NH-SIV de la nariz del ratón. El 69 % de las células transducidas en la cavidad nasal del ratón eran células epiteliales respiratorias ciliadas. Otros tipos de células transducidas incluyeron células neuronales en el epitelio olfativo (21 %) y células escamosas (7 %).

[0159] Para detectar células basales después del tratamiento con polidocanol, se realizó inmunotinción basada en peroxidasa de rábano picante (HRP) usando el kit Envision (Dako, Glostrup, Dinamarca). Brevemente, los portaobjetos se trataron con peróxido de hidrógeno al 0,6 % en metanol durante 15 minutos, se lavaron en agua corriente y se incubaron con suero de cabra normal al 1,5 % (Abcam) durante 30 minutos. A continuación, los portaobjetos se incubaron con un anticuerpo policlonal de conejo anti-Citoqueratina 5 (1:500) (Abcam) durante 1 h después de una IgG anti-conejo de cabra conjugada con HRP (proporcionada con el kit) durante 30 min. Después, las secciones se lavaron en PBS y se incubaron con el sustrato de peroxidasa 3-amino-9-etilcarbazol (AEC) (proporcionado con el kit) durante 5 min. Finalmente, los portaobjetos se lavaron en H_2O destilada, se contratiñeron con hematoxilina acuosa de Harris (BDH) durante 15 segundos, se lavaron en agua corriente y después en H_2O destilada.

[0160] La detección de inmunofluorescencia de células epiteliales nasales transducidas GFP positivas y células basales Krt5 positivas se realizó usando los siguientes anticuerpos primarios y secundarios: anticuerpo policlonal de cabra anti-GFP (1:250) (Abcam), anticuerpo monoclonal de conejo anti-KRT5 (1:500) (Abcam), IgG anti-cabra de burro Alexa Fluor 488 (1:200) (Invitrogen, Paisley, Reino Unido) e IgG anti-conejo de cabra Alexa Fluor 594 (1:200) (Invitrogen). Para mejorar el rendimiento del anticuerpo, las secciones se sometieron a recuperación de antígeno mediada por calor en tampón citrato (ácido cítrico 10 mM, Tween20 al 0,05 %, pH 6,0) durante 20 min en un baño de agua a 100 °C. Se montaron secciones teñidas en reactivo ProLong® Gold Antifade con DAPI (Invitrogen) y se analizaron con un microscopio confocal como anteriormente (todo Zeiss). Las células basales GFP positivas (identificación basada en la morfología y ubicación dentro de la capa epitelial) se cuantificaron en un total de 13 secciones/ratón. Las secciones que mostraron células basales putativas GFP positivas se seleccionaron para doble tinción con los anticuerpos anti-KRT5 y anti-GFP para confirmar el fenotipo de células basales.

Ejemplo 11: Transducción de cultivos de ALI

[0161] Las células epiteliales de las vías respiratorias completamente diferenciadas cultivadas como cultivos de ALI se adquirieron en Epithelix (Ginebra, Suiza). Las ALI se transfectaron con F/HN-SIV-lux en una multiplicidad de infección que variaba de ~25 a ~300 UT/célula. Después de 6 horas, se eliminó el virus y se incubaron las ALI durante 10-26 días. El medio basolateral se cambió cada 48 horas durante este periodo de incubación. En los puntos de tiempo especificados, las ALI se lisaron en 100 µl de tampón de lisis indicador y se cuantificó la expresión de luciferasa usando el sistema de ensayo de luciferasa (Promega, Southampton, Reino Unido) según las instrucciones del fabricante. El contenido proteico total de los cultivos se cuantificó utilizando el kit de ensayo de proteínas BioRad (BioRad, Hemel Hempstead, Reino Unido). Cada muestra se ensayó por duplicado. A continuación, la expresión de luciferasa se presentó como unidades relativas de luz/mg de proteína total. Para las imágenes de bioluminiscencia, se añadieron 100 µg de luciferina en PBS a la membrana apical.

[0162] Como se muestra en la figura 7, las células en cultivos de ALI se transdujeron con éxito con F/HN-SIV-Lux, como lo demuestra la expresión de luciferasa. La expresión de luciferasa fue mayor a MOI 250 en comparación con MOI 25.

Ejemplo 12: Ensayo de eflujo de yoduro

[0163] Las células HEK293T se transfectaron con F/HN-SIV-GFP-CFTR o un virus de control F/HN-SIV-GFP a una multiplicidad de infección de 500 UT/célula y se cultivaron durante 2 días. La actividad del canal de cloruro de CFTR se ensayó midiendo la tasa de eflujo de ^{125}yoduro como se ha descrito previamente (Derand, R., et al., 2003). Las tasas de flujo de ^{125}yoduro se normalizaron al momento de la adición de forskolina/IBMX (tiempo 0). Las curvas se construyeron trazando tasas de eflujo de ^{125}yoduro frente al tiempo. Para reflejar los niveles acumulativos de eflujo de ^{125}yoduro después de la estimulación con agonista, todas las comparaciones se basan en áreas bajo las curvas de tiempo-eflujo de ^{125}yoduro. El área bajo la curva se calculó por la regla del trapecio. Los experimentos se realizaron por duplicado (n = 6 pocillos/grupo/experimento).

[0164] Como se muestra en la figura 8, F/HN-SIV puede dirigir la expresión funcional de CFTR, con un eflujo de salida de ^{125}yoduro relativo de 0,65. Por el contrario, el vector plasmídico GL67A alcanzó un valor de eflujo de ^{125}yoduro inferior de 0,33.

Ejemplo 13: Transducción de células de pulmón primarias de oveja y ser humano y pulmón de ratón

[0165] F/HN-SIV transduce eficazmente células de pulmón primarias de oveja y ser humano y pulmón de ratón.

[0166] Se usó F/HN-SIV-CMV-Lux para transducir cepillados nasales humanos (MOI 250) y cortes de pulmón humano y ovino cultivados *ex vivo* ($1x10^7$ UT/corte). Como se muestra en la figura 9A, la transducción tanto de las células de cepillado nasal humano como de los cortes de pulmón humano y de oveja dio como resultado una expresión transgénica de luciferasa sustancial (valores promedio en la región de $2x10^2$ URL/mg de proteína para los cepillados nasales humanos, $1x10^7$ URL/mg de proteína para los cortes de pulmón humano y 2×10^7 URL/mg de proteína para los cortes de pulmón de oveja) 24-48 horas después de la transducción.

[0167] Las células de pulmón de FQ humanas primarias cultivadas en la interfaz aire-líquido (FQ hALI, ~$1x10^5$) se transdujeron con (3×10^7 UT) vectores F/HN-SIV-soCFTR2 que contenían promotores transgénicos de CMV y hCEF. El número de copias de vector (copias de ADN proviral por copia de ADN de CFTR endógeno) se midió 6-8 días después de la transducción. Tanto los promotores de CMW como de hCEF pudieron lograr un número de copias de vector de al menos 1×10^1 (figura 9B).

[0168] También se midió el nivel de expresión de ARNm de CFTR (% de VE: copias de ARNm de CFTR por copia de ARNm de CFTR endógeno x 100) 6-8 días después de la transducción. La línea discontinua horizontal en la figura 9C representa un nivel de expresión objetivo del 5 % de VE, que se cree que representa el umbral terapéutico. Tanto F/HN-SIV-soCFTR-CMV como F/HN-SIV-soCFTR2-hCEF indujeron una expresión significativamente por encima de este objetivo (en la región de 40583 ± 10687 y 18509 ± 13588 respectivamente, media ± DE, n = 4).

[0169] Después de la administración *in vivo* de vectores F/HN-SIV-EGFPLux que contienen promotores de CMV, EF1a y hCEF en forma de integrasa defectuosa (ID) o competente en integrasa (IC o sin etiqueta), se determinó la expresión transgénica de las células de las vías respiratorias en el epitelio murino nasal (figura 9D) y pulmonar (figura 9E) (n = 6-10/grupo). El curso temporal de la expresión transgénica de luciferasa se monitorizó mediante imágenes de bioluminiscencia repetidas *in vivo* y se normalizó a la dosis administrada. Cuatro de los cinco vectores ensayados (vGM012 CMW, vGM014 EF1a, vGM020 hCEF y vGM076 hCEF ID) lograron una expresión en la nariz por encima del nivel objetivo durante todo el curso temporal del experimento. El quinto vector, vGM074 CMV ID, logró una expresión en la nariz por encima del nivel de expresión aceptado durante todo el curso temporal del experimento.

[0170] Dos de los cinco vectores ensayados (vGM014 EF1a y vGM020 hCEF) lograron una expresión en el pulmón por encima del nivel objetivo durante todo el curso temporal del experimento. Un vector, vGM012 CMW, logró una expresión en el pulmón por encima del nivel de expresión aceptado durante todo el curso temporal del experimento.

[0171] Se detectó bioluminiscencia después de la transducción murina *in vivo* el día 14 después de la transducción. Se muestran imágenes representativas en la figura 9F. El vector vGM020 hCEF logró el nivel más alto de expresión *in vivo* de los cinco vectores ensayados.

[0172] La bioluminiscencia también se detectó después de la transducción *in vitro* de hALI no FQ 5-6 días después de la transducción. Se muestran imágenes representativas en la figura 9G. De nuevo, el vector vGM020 hCEF logró el nivel más alto de expresión de los cinco vectores ensayados.

[0173] La figura 9H muestra la expresión de EGFP 14 días después de la transducción en el epitelio nasal murino tras la administración de $1,6 \times 10^8$ UT de F/HN-SIV-hCEF-EGFPLux (vGM020), tal como se visualiza por inmunohistoquímica (núcleos teñidos con DAPI).

[0174] El curso temporal de la expresión transgénica de luciferasa en ALI no FQ se monitorizó mediante imágenes de bioluminiscencia repetidas y se normalizó a la dosis administrada. Como se muestra en la figura 9I, los vectores vGM014 EF1a y vGM020 hCEF alcanzaron el nivel más alto de expresión.

[0175] F/HN-SIV también transduce eficazmente pulmón de oveja *in vivo*. Se instiló acriflavina (3x100 µl de alícuotas durante ~5 minutos) a una vía respiratoria proximal bajo visualización broncoscópica directa. La distribución de la acriflavina se puede apreciar por la coloración naranja de la vía respiratoria disecada en la autopsia (figura 10A).

[0176] La acriflavina se restringió en gran medida a las vías respiratorias conductoras y estaba ausente en las regiones alveolares. La flecha en la figura 10A indica el sitio aproximado de instilación. Los números en la regla están en cm.

[0177] La figura 10B es una representación esquemática del pulmón de oveja (centro/parte superior de la tráquea). El círculo representa la región en la (figura 10A). En la figura 10B, la flecha indica el paso del broncoscopio para administrar 3x100 µl de alícuotas de 2,2E9 UT/ml (6,6E8 TU total) de F/HN SIV CMV-EGFPLux con respecto a n = 3 ovejas individuales (códigos de los animales T121, T156 y T251). Siete días después de la administración, se

tomaron 5-6 bloques de muestras de tejido en la autopsia a intervalos de ~1 cm desde el sitio de instilación.

[0178] Los bloques de muestra se dividieron en 2-3 muestras aproximadamente equivalentes y se analizaron para determinar la expresión transgénica, cuyos resultados se muestran en la figura 10C como ensayos de luciferasa normalizados con respecto al contenido de proteínas; y en la figura 10D como RT-PCR cuantitativa normalizada con respecto a niveles de ARNm de CFTR endógeno. La línea horizontal en la figura 10C representa la actividad de luciferasa más alta observada en cualquier muestra tratada con un vector de transferencia génica no viral, y en la figura 10D, el nivel de expresión objetivo del 5 % de VE (se cree que representa el umbral terapéutico, véase anteriormente). Para cada grupo de tratamiento, el promedio fue más alto que la comparación de vectores no virales (figura 10C), y también logró una expresión por encima del umbral objetivo del 5 % de VE.

Ejemplo 14: Expresión y función de CFTR medida usando organoides intestinales de FQ humana

[0179] Se generaron organoides intestinales de FQ humana como se describe por Dekkers JF et al. (Nature Medicine 2013, 19(7): 939-945). En resumen, las biopsias intestinales se lavaron en soluciones que contenían EDTA para disociar las células de la cripta. A continuación, las células de la cripta se transdujeron con vGM058 (aproximadamente $1x10^7$ unidades de transducción) o un virus de control (n = 3 pocillos/condición) y se incluyeron en Matrigel y se les permitió formar organoides durante 3-4 días. La función de CFTR se evaluó exponiendo los organoides a forskolina (aproximadamente 5 µM de forskolina), lo que aumenta los niveles intracelulares de cAMP y, por lo tanto, activa el CFTR. En respuesta a la activación de CFTR, los organoides aumentan el transporte de cloruro, lo que conduce a la absorción de agua y la hinchazón. Los organoides (mínimo 10/pocillo) se analizaron directamente por microscopía confocal de células vivas (LSM710, Zeiss, objetivo x5). La hinchazón organoide estimulada por forskolina se cuantificó automáticamente utilizando el software de imagen Volocity (Improvision). Se calculó y se promedió el aumento total del área organoide (plano xy) con respecto a t = 0 del tratamiento con forskolina. La hinchazón inducida por forskolina aumentó significativamente ($p < 0,001$) en los organoides transducidos con vGM058 en comparación con los controles (véase la figura 14A). También se detectaron niveles significativos de eflujo de yoduro mediado por CFTR ($p < 0,05$) en células transducidas con vGM058 (figura 14B) usando el ensayo de eflujo de yoduro descrito en esta invención (véase el Ejemplo 12).

Ejemplo 15: Generación de vectores lentivirales para A1AT

[0180] Los vectores lentivirales se prepararon usando el esqueleto de VIS y el procedimiento de 5 plásmidos descrito anteriormente en los Ejemplos 2 y 3 (para el lentivirus CFTR) y usando el promotor de hCEF como se describe en esta invención. Se generaron dos construcciones lentivirales separadas: una con un transgén alfa-1-antitripsina humana (hAAT); uno con un transgén de luciferasa de *Gaussia* (Glux) (véase la figura 15A y B, SEQ ID NOs: 9 y 10 respectivamente). Los ADNc contenidos en estos vectores estaban optimizados con codones y agotados en CpG.

Ejemplo 16: Cultivo de interfaz aire-líquido (ALI) utilizando un vector lentiviral del gen indicador lux

[0181] Se adquirieron cultivos de ALI humanos naturales completamente diferenciados (MucilAir) en Epithelix SARL (Ginebra, CH). Las ALI se cultivaron a 37 °C y con CO^2 al 5 %, y el medio de cultivo basolateral se cambió cada 2-3 días. El medio de cultivo se almacenó a -20 °C hasta su posterior análisis.

[0182] Las ALI se transdujeron (el día 0) añadiendo por pipeteo 100 µl de virus ($1x10^7$ unidades de transfección Tagman (UTT) sobre la superficie apical. El virus se eliminó después de 4 horas de incubación a 37 °C, y el medio basolateral se reemplazó.

[0183] En los puntos de tiempo indicados después de la transducción, la superficie apical de las ALI se lavó mediante incubación con PBS estéril durante una hora. Los lavados se retiraron y se almacenaron a -20 °C hasta su posterior análisis.

[0184] Se realizó un ensayo de luciferasa de *Gaussia* (New England Biolabs, Ipswich, EE.UU.) según las recomendaciones del fabricante. Se analizaron 15 µl de muestra por duplicado y se determinó la luminiscencia en un lector de placas Appliskan. La expresión de Glux se expresó como URL/µl de fluido (URL = unidades relativas de luz).

[0185] La figura 16 proporciona los resultados, representando cada punto el valor medio de URL/µl en los medios de n = 6 ALI en el punto de tiempo mostrado (error estándar indicado por barras de error).

[0186] La transferencia génica mediada por lentivirales en interfaces aire-líquido humanas dio como resultado la expresión a largo plazo de la proteína indicadora secretada luciferasa de *Gaussia*.

Ejemplo 17: Expresión transgénica en cultivos de corte de pulmón usando vectores lentivirales del gen indicador A1AT y lux

[0187] Se prepararon cortes de pulmón humano cortados con precisión como se describe en Moreno L et al,

Respir Res 2006 21 de agosto;7:111. Los cortes de pulmón se colocaron en placas de cultivo tisular de 12 pocillos (1 corte por pocillo) en 1 ml de medio y se incubaron a 37 °C y CO^2 al 5 %. El medio se cambió a diario y se almacenó a -20 °C hasta su posterior análisis.

[0188] El día 0 se transdujeron cortes pulmonares (n = 6 por grupo) con un virus SIV hCEF-sogLux (1 x 10^6 UTT) o SIV1 hCEF-sohAAT (2 x 10^6 UTT) diluido en medio a un volumen final de 1000 µl y se incubaron durante 4 horas. Después de la incubación, el medio se reemplazó y se almacenó a -20 °C hasta su posterior análisis.

[0189] La expresión de luciferasa de *Gaussia* se determinó como se ha descrito anteriormente (Ejemplo 16). Como se muestra en la figura 17A, los altos niveles de expresión de la proteína indicadora secretada luciferasa de *Gaussia* siguieron a la transferencia génica mediada por lentivirales en cortes de pulmón humano. La expresión de AAT (también denominada en el presente documento A1AT) se determinó utilizando un ELISA tipo sandwich (Abcam, Cambridge, Reino Unido), realizado según las recomendaciones del fabricante. Se ensayaron 50 µl de muestra por duplicado y se midió en un lector de microplacas a 450 nm. Como se muestra en la figura 17B, los altos niveles de expresión de alfa-1-antitripsina (AAT/A1AT) siguieron a la transferencia génica mediada por lentivirales en cortes de pulmón humano.

Ejemplo 18: Administración *in vivo* de vectores lentivirales del gen indicador A1AT y lux a la nariz del ratón

Transducción de pulmón de ratón:

[0190] Se anestesiaron ratones hembra C57BL/6 (Charles River, Reino Unido) con isoflurano y se les dieron 100 ul de virus por instilación nasal como se describe en Xenariou S et al, Gene Ther mayo de 2007; 14(9): 768-75. Los animales recibieron entre 1 y 5 dosis y se observaron diariamente para detectar signos de toxicidad.

[0191] Para la luciferasa de *Gaussia*, se anestesiaron ratones hembra C57BL/6 el día 0 usando isoflurano y se les administró una dosis única de 100 ul del virus SIV1 hCEF-soGLux (1x10^6 UTT) mediante instilación nasal. Los animales de control se trataron con DMEM (medio de cultivo tisular), el componente principal de la preparación viral utilizada en el estudio.

[0192] Para A1AT, se trataron ratones hembra C57bl/6 (n = 5 por grupo) con 3 dosis de SIV1 hCEF-sohAAT a intervalos de 10 días (100 µl por dosis, 6,8x10^7 UTT; dosis total 2,4x10^8 UTT). Los animales de control se instilaron con 100 ul de PBS esterilizado (el componente principal del lote de producción de lentivirus utilizado en el estudio) en cada punto de dosificación.

[0193] 10 días después de la tercera dosis, los animales se sacrificaron y el homogeneizado de tejido pulmonar, el fluido de lavado broncoalveolar y el suero se analizaron para determinar la expresión de AAT.

[0194] Además, se investigó la expresión a largo plazo de A1AT. Los días 1 a 5 del experimento, los ratones C57bl/6 se trataron con 100 µl de SIV1 hCEF-sohAAT mediante instilación nasal (5 dosis de 4 x 10^5 UTT, es decir, 2x10e^6 UTT por animal en total). Los animales de control se instilaron con 100 ul de DMEM (medio de cultivo tisular), el componente principal del lote de producción de lentivirus utilizado en el estudio. Los animales se sacrificaron en diversos puntos de tiempo después de la transducción y se analizaron el homogeneizado de tejido pulmonar, el fluido de lavado broncoalveolar y el suero para determinar la expresión de AAT.

Recogida de tejido de ratón:

[0195] Los ratones se sacrificaron en los puntos de tiempo indicados después de la transducción. Se extrajo sangre perforando el ventrículo izquierdo y se centrifugó a 760 $g_{prom.}$ durante 10 minutos para preparar suero. El suero se congeló posteriormente a -80 °C.

[0196] Se realizó un lavado broncoalveolar (BAL) diseccionando el cuello, insertando una cánula en la tráquea y asegurándola en su lugar con hilo de sutura. Se instilaron 500 µl de PBS en el pulmón, y se aspiró tres veces para obtener un lavado completo del revestimiento epitelial. La muestra se congeló inmediatamente en nitrógeno líquido y se almacenó a -80 °C para su posterior análisis.

[0197] A continuación, los pulmones se diseccionaron y se congelaron inmediatamente en nitrógeno líquido, y posteriormente se homogeneizaron en tubos de matrix D de lisis (MP Biomedicals), se centrifugaron en una máquina FastPrep (ThermoFisher Scientific, Waltham, MA, EE.UU.) a 4 m/s durante 45 segundos, y se almacenaron a -80 °C para su posterior análisis.

[0198] La expresión de AAT (A1AT) y luciferasa de *Gaussia* (Glux) se determinó como se ha descrito en los Ejemplos 16 y 17 anteriores. La transducción *in vivo* de células de las vías respiratorias de ratón con una dosis única del vector lentiviral del gen indicador lux del Ejemplo 15 dio como resultado la expresión a largo plazo (al menos 12 meses) de la proteína indicadora secretada luciferasa de *Gaussia*, en homogeneizado de pulmón (figura 18A), fluido

de lavado broncoalveolar (BAL, figura 18B) y suero (figura 18C).

[0199] Se observaron altos niveles de expresión de A1AT en homogeneizado de pulmón, BAL y suero después de la transferencia mediada por lentivirales del gen AAT (A1AT) *in vivo* (figura 19), con un aumento de más de 100 veces en la expresión de ATT (A1AT) en el homogeneizado de pulmón y BAL observado en comparación con los controles negativos correspondientes (PBS). También se observó un aumento significativo (al menos un orden de magnitud) en la expresión de ATT (A1AT) en el suero.

[0200] Además, se observó expresión a largo plazo (al menos 90 días) de alfa-1-antitripsina en homogeneizado de pulmón (figura 20A), BAL (figura 20B) y suero (figura 20C) después de la transferencia génica mediada por lentivirales del gen AAT (A1AT) *in vivo*.

Ejemplo 19: Ensayo de urea

[0201] Se trataron ratones C57bl/6 (n = 5 por grupo) con 3 dosis de SIV1 hCEF-sohAAT a intervalos de 10 días (100 µl por dosis, $6{,}8x10^7$ UTT; dosis total $2{,}4x10^8$ UTT). Los animales de control se instilaron con 100 ul de PBS esterilizado (el componente principal del lote de producción de lentivirus utilizado en el estudio) en cada punto de dosificación.

[0202] 10 días después de la tercera dosis, los animales se sacrificaron y el homogeneizado de tejido pulmonar, el fluido de lavado broncoalveolar y el suero se analizaron para determinar la expresión de A1AT.

[0203] Se realizó un ensayo de urea (Abcam, Cambridge, Reino Unido) según las instrucciones del fabricante.

[0204] En primer lugar, se prepararon diluciones seriadas de suero murino y muestras de fluido BAL y se analizaron para determinar la dilución apropiada para usar en experimentos adicionales.

[0205] En segundo lugar, se analizaron muestras de suero y fluido BAL correspondientes de ratones individuales (n = 14) para calcular la diferencia múltiplo entre la concentración de urea en suero y fluido BAL, equivalente al efecto diluyente de BAL en el fluido de revestimiento epitelial (según Rennard SI et al., J. Appl Physiol (1985). 1986 Feb;60(2):532-8). La dilución media de BAL fue de 41 veces (intervalo 24-88).

[0206] Teniendo en cuenta este efecto de dilución, se calculó la concentración de ATT (A1AT) en el fluido de revestimiento epitelial. Específicamente, la concentración de AAT en el fluido de lavado broncoalveolar se multiplicó por el factor de dilución, para proporcionar una estimación de la concentración de AAT "verdadera" en el fluido de revestimiento epitelial.

[0207] Un nivel objetivo "protector" de ATT (A1AT) en el fluido de revestimiento epitelial (ELF, es decir, el fluido que recubre las vías respiratorias y los espacios aéreos en los pulmones) es 70 µg/ml (en comparación con un nivel objetivo "normal" de ATT (A1AT) en el ELF de 200 µg/ml). Como se muestra en la figura 21, los niveles terapéuticos de alfa-1-antitripsina en el fluido de revestimiento epitelial siguieron a la transferencia génica mediada por lentivirales de ATT (A1AT) *in vivo*.

Ejemplo 20: Generación de vectores lentivirales para FVIII

[0208] Se prepararon cuatro vectores lentivirales de FVIII diferentes usando el esqueleto de VIS y el procedimiento de 5 plásmidos descrito anteriormente en los Ejemplos 2 y 3 (para el lentivirus CFTR) y 15 (para el lentivirus A1AT). Los plásmidos promotor-transgén tienen las SEQ ID NOs: 11 a 14 respectivamente.

[0209] La secuencia del VIS era idéntica a las construcciones CFTR (Ejemplos 2 y 3) excepto por el promotor y el ADNc. El promotor de citomegalovirus humano (CMV) o el promotor/potenciador de hCEFI específico de tejido se usó como se indica (figura 22) para impulsar la expresión de transgenes de FVIII.

[0210] SIV-F/HN-FVIII-N6-co contenía el ADNc de FVIII humano natural del que se ha eliminado el dominio BDD y se ha reemplazado con un fragmento de glucosilación 6N de 226 aminoácidos con optimización de codones.

[0211] SIV-FVIII-V3 contiene el ADNc de FVIII humano natural del que se ha eliminado el sitio de glucosilación de 226 aminoácidos y se ha reemplazado con un péptido de 17 aminoácidos que expresa triples de glucosilación 6N dentro del dominio B (McIntosh et al., Blood 2013 121(17); 3335-3344).

Ejemplo 21: Cuantificación del antígeno hFVIII y niveles de actividad en modelos *in vivo* e *in vitro*

[0212] Los niveles de antígeno FVIII humano en un modelo murino se cuantificaron mediante un ensayo inmunoabsorbente ligado a enzimas (ELISA) según el protocolo del fabricante. Brevemente, se analizó la presencia de antígeno FVIII en plasma, BAL y pulmón usando Elisa Asserachrom (FVIII:Ag) (Stago Diagnostics, Francia).

[0213] Las muestras se diluyeron 1:2 y se incubaron en una placa de 96 pocillos recubierta con fragmento de factor VIII antihumano monoclonal de ratón durante 2 horas a temperatura ambiente. Después del lavado, se añadió a la placa un anticuerpo secundario anti-ratón junto con peroxidasa y la incubación se realizó durante 2 horas a TA. Los niveles de hFVIII:Ag se determinaron por espectrofotometría a 450 nm usando sustrato TMB (datos no mostrados).

[0214] Se usó otro ensayo ELISA para evaluar la actividad de FVIII en un modelo HEK293T *in vitro* (FVIII:C, Affinity Biological, Canadá). Los sobrenadantes se recogieron 48 y 72 horas después de la transducción HEK293T con SIV-F/HN-FVIII-N6 o SIV-F/HN-FVIII-N3. La actividad de FVIII se evaluó siguiendo las instrucciones del fabricante utilizando 50 µl de sobrenadantes ensayados por duplicado. Como control negativo se ensayó el sobrenadante de las células HEK293T no tratadas. La actividad de hFVIII se calculó a partir de una curva estándar generada utilizando una serie de diluciones de plasma combinado humano normal (13º Estándar Británico para concentrado de Factor VIII de coagulación sanguínea, Humano; NIBSC).

[0215] Las células HEK293T se transdujeron con dos lotes diferentes de vGM142 (lote 1 - $5{,}9x10^8$ UTT/ml y lote 2 - $2{,}8x10^8$ UTT/ml). Las células HEK293T se transdujeron con el vector del lote 1 de vGM142 (figura 23A) y el vector del lote 2 de vGM142 (figura 23B) a 3 MOI diferentes (MOI 1; 10; 100), y se recogieron 48 y 72 horas después de la transducción.

[0216] Como queda claro en la figura 23, se observó una actividad de FVIII creciente con MOI creciente tanto para el lote 1 como para el lote 2 de vGM142 tanto 48 como 72 horas después de la transducción. Además, la actividad de FVIII aumentó de 48 horas a 72 horas por cada MOI ensayada.

Ejemplo 22: Administración *in vivo* de vectores lentivirales del gen indicador FVIII y lux a la nariz del ratón

Transducción de pulmón de ratón:

[0217] Todos los procedimientos con animales se realizaron según las condiciones y la limitación del Proyecto de Ministerio del Interior del Reino Unido y las regulaciones de licencia personal bajo la Ley de Procedimiento Científico Animal (1986).

[0218] Ratones hembra C57BL/6 genéticamente intactos de 6 a 8 semanas de edad (Charles River, Reino Unido) se anestesiaron con isofluorano y se les dieron 100 µl de virus en solución salina tamponada con fosfato de Dulbecco (D-PBS), como se ha descrito anteriormente (Griesenbach et al., 2012) y se evaluó la presencia del antígeno FVIII.

[0219] En dos experimentos (Grupo 1 y 2) los ratones recibieron 3 dosis (cada dos días) de SIV-F/HN-FVIII-N6 (vGM142) y se sacrificaron 10 días después de la primera dosis. El Grupo 1 (n = 4) se trató con una dosis de vector total de $1{,}4x10^6$ UTT/ratón. El Grupo 2 (n = 3) se trató con una dosis de vector total de $1{,}57x10^8$ UTT/ratón.

[0220] En un experimento (Grupo 3) los ratones se trataron con 12 dosis (cada dos días) de SIV-F/HN-FVIII-N6 (vGM142) y se sacrificaron 28 días después de la primera dosis. El Grupo 3 (n = 4) se trató con una dosis de vector total de $3{,}36x10^8$ UTT/ratón).

[0221] Se extrajeron plasma, fluido BAL y pulmón (como se describe en el Ejemplo 14). En resumen, los ratones se sacrificaron en los puntos de tiempo indicados después de la transducción. A continuación, se extrajo sangre del corazón en los tubos de recogida de anticoagulante de citrato trisódico 3.2, antes de centrifugarse a 2000 - 2500 x g para obtener plasma. El fluido BAL se recogió aplicando 3 instilaciones consecutivas de PBS (500 µl) en el pulmón de ratón a temperatura ambiente. Los sobrenadantes se almacenaron a -80 °C. Los pulmones se extrajeron y se almacenaron a -80 °C antes de la homogeneización de los tejidos.

[0222] A continuación, se evaluó la presencia de expresión de FVIII. Los niveles de FVIII se evaluaron en homogeneizados de tejido pulmonar (figura 24A), fluido BAL (figura 24B) y plasma (figura 24C) recogidos por separado en 3 experimentos independientes 10 y/o 28 días después del tratamiento con SIV-F/HN-FVIII-N6. El análisis se realizó utilizando Anova unidireccional (GraphPad Prism) con comparaciones múltiples entre los grupos tratados (**** $p < 0{,}0001$).

[0223] Como resultante evidente a partir de la figura 24A, los tres grupos de tratamiento produjeron un aumento observable en los niveles de hFVIII dentro del tejido pulmonar en comparación con el control correspondiente (D-PBS). El tratamiento de 28 días del Grupo 3 dio como resultado un aumento significativo en la expresión de hFVIII en comparación con los tratamientos de 10 días de los Grupos 1 y 2. Se observaron resultados similares para las muestras de fluido BAL (figura 24B), aunque en estas muestras también hubo un aumento significativo en los niveles de hFVIII en el Grupo 2 en comparación con el Grupo 1. El tratamiento del Grupo 3 dio como resultado un aumento significativo en los niveles de hFVIII en plasma (figura 24C).

Ejemplo 23: Procedimiento de fabricación para vectores lentivirales según la invención

[0224] Las células HEK293 se cultivan en suspensión, en medio de expresión Freestyle (químicamente definido, sin origen animal y libre de proteínas), y se monitoriza el recuento celular. La concentración de glucosa se determina y se ajusta a ~35 mM. Se prepara una mezcla de transfección de pDNA/PEIPro™ y las células se transfectan a 0,33 mg de células pDNA/1E9.

[0225] El recuento celular se monitoriza de nuevo y se añade más cantidad de medio de expresión Freestyle. La concentración de glucosa se determina de nuevo y se ajusta a ~35 mM. Se pueden añadir 5 u/ml de Benzonase® y se realiza una clarificación en línea del virus en tres fases. Se añade Benzonase® seguido de TrypLE Select™. El virus se enfría a 0 °C y se mantiene en hielo húmedo en todas las etapas posteriores. Después de filtrar cualquier material particulado distinto del virus (filtro de mPES de 0,45 µm), el virus se carga en Mustang® Q XT (membrana clarificada de 3 ml/l de virus clarificado) seguido de lavado con NaCl Tris 0,15 M a pH 7,5 y elución con NaCl Tris 1,0 M pH 7,5. La fracción de virus se recoge y se diluye a un volumen inicial de 0,1-0,2 con medio Freestyle. Se puede añadir aquí TrypLE Select™ si no se añadió antes y también se puede añadir Benzonase® en esta fase además o en lugar de anteriormente.

[0226] Se realiza la filtración de flujo tangencial Spectrum (TFF) (UF a ~0,1-0,05 de volumen inicial = HV; retenido de DF x5 HV frente a tampón de formulación; UF a ~0,001 - 0,002 de volumen inicial) y se recoge el retenido. Se puede realizar una segunda etapa de TFF y se puede usar una unidad de TFF más pequeña para el DF y/o UF final. Las etapas adicionales pueden incluir modo mixto/SEC y filtración estéril de 0,45 µm o 0,2 µm.

[0227] La figura 11 representa la producción y purificación de los vectores F/HN SIV. Los vectores F/HN SIV se produjeron mediante transfección transitoria mediada por PEI de 5 plásmidos (pDNA) de células 293T cultivadas en suspensión a escala de 1 l en biorreactores WAVE controlados con pH, usando procedimientos escalables de la invención. Los vectores se clarificaron por filtración de profundidad/final (GE/Pall), los ácidos nucleicos contaminantes se eliminaron con Benzonase® (Merck), los vectores se activaron con TrypLE Select™ (Life Technology), se purificaron y se concentraron mediante cromatografía de membrana de intercambio aniónico (Pall) y filtración de flujo tangencial (Spectrum). Todos los recipientes, contenedores y columnas de procedimiento cumplían el cGMP de un solo uso. Todos los reactivos, excepto el ADN plasmídico, eran compatibles con cGMP sin origen animal. Se muestran datos de una diversidad de configuraciones vectoriales (promotor transgénico, transgén, estado de la integrasa). Las valoraciones físicas y funcionales se determinaron usando Q-PCR.

[0228] Véanse los resultados de este procedimiento ejemplar de la invención que se analizan en la descripción de la figura 11 anterior.

Clave para las SEQ ID Nos

SEQ ID NO: 1	Plásmido como se define en la figura 1A (pDNA1 pGM326)
SEQ ID NO: 2	Plásmido como se define en la figura 1B (pDNA2a pGM297)
SEQ ID NO: 3	Plásmido como se define en la figura 1C (pDNA2b pGM299)
SEQ ID NO: 4	Plásmido como se define en la figura 1D (pDNA3a pGM301)
SEQ ID NO: 5	Plásmido como se define en la figura 1E (pDNA3b pGM303)
SEQ ID NO: 6	Promotor de hCEF ilustrado
SEQ ID NO: 7	Transgén CFTR ilustrado (soCFTR2)
SEQ ID NO: 8	Componente de WPRE ilustrado (mWPRE)
SEQ ID NO: 9	Plásmido F/HN-SIV-hCEF-soA1AT como se define en la figura 15 (pDNA1 pGM407)
SEQ ID NO: 10	Plásmido F/HN-SIV-hCEF-sogLux como se define en la figura 15 (pDNA1 pGM358)
SEQ ID NO: 11	Plásmido F/HN-SIV-CMV-HFVIII-V3 como se define en la figura 22C (pDNA1 pGM411)
SEQ ID NO: 12	Plásmido F/HN-SIV-hCEF-HFVIII-V3 como se define en la figura 22D (pDNA1 pGM413)
SEQ ID NO: 13	Plásmido F/HN-SIV-CMV-HFVIII-N6-co como se define en la figura 22E (pDNA1 pGM412)
SEQ ID NO: 14	Plásmido F/HN-SIV-hCEF-HFVIII-N6-co como se define en la figura 22F (pDNA1 pGM414)
SEQ ID NO: 15	Transgén A1AT ilustrado
SEQ ID NO: 16	Transgén FVIII (N6) ilustrado
SEQ ID NO: 17	Promotor de CMV ilustrado
SEQ ID NO: 18	Cebador para la construcción de pCAGGS-Fct4
SEQ ID NO: 19	Cebador para la construcción de pCAGGS-Fct4
SEQ ID NO: 20	Cebador para la construcción de pCAGGS-Fct4
SEQ ID NO: 21	Cebador para la construcción de pCAGGS-Fct4
SEQ ID NO: 22	Cebador para la construcción de pCAGGS-Fct4
SEQ ID NO: 23	Cebador para la construcción de pCAGGS-Fct4
SEQ ID NO: 24	Cebador para la construcción de pCAGGS-SIVct+HN
SEQ ID NO: 25	Cebador para la construcción de pCAGGS-SIVct+HN

(continuación)

SEQ ID NO: 26	Hebra complementaria con respecto al transgén A1AT ilustrado
SEQ ID NO: 27	Péptido A1A1 ilustrado
SEQ ID NO: 28	Hebra complementaria con respecto al transgén FVIII (N6) ilustrado
SEQ ID NO: 29	Péptido FVIII (N6) ilustrado
SEQ ID NO: 30	Transgén FVIII (V3) ilustrado
SEQ ID NO: 31	Hebra complementaria con respecto al transgén FVIII (V3) ilustrado
SEQ ID NO: 32	Péptido FVIII (V3) ilustrado
SEQ ID NO: 33	Hembra complementaria con respecto al promotor de CMV ilustrado

Secuencias

[0229]

SEQ ID NO: 1

```
        1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT
       61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
      121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
      181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
      241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
TGACGTATGT
      301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
      361 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT
      421 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAC
GGGACTTTCC
      481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA
      541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT
      601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA
      661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGCGTGTACG
GTGGGAGGTC
      721 TATATAAGCA GAGCTCGCTG GCTTGTAACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC
      781 TGGGTGTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA
      841 GAAAGCTAAT AAACTTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCCT
CATTGACGCC
      901 TCACTCTCTT GAACGGGAAT CTTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA
      961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA
     1021 GGCGTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA
     1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC
     1141 CGTAACTACT CTTGGGCAAG TAGGGCAGGC GGTGGGTACG CAATGGGGGC
GGCTACCTCA
     1201 GCACTAAATA GGAGACAATT AGACCAATTT GAGAAAATAC GACTTCGCCC
GAACGGAAAG
     1261 AAAAAGTACC AAATTAAACA TTTAATATGG GCAGGCAAGG AGATGGAGCG
CTTCGGCCTC
     1321 CATGAGAGGT TGTTGGAGAC AGAGGAGGGG TGTAAAAGAA TCATAGAAGT
CCTCTACCCC
     1381 CTAGAACCAA CAGGATCGGA GGGCTTAAAA AGTCTGTTCA ATCTTGTGTG
CGTGCTATAT
     1441 TGCTTGCACA AGGAACAGAA AGTGAAAGAC ACAGAGGAAG CAGTAGCAAC
AGTAAGACAA
     1501 CACTGCCATC TAGTGGAAAA AGAAAAAAGT GCAACAGAGA CATCTAGTGG
ACAAAAGAAA
     1561 AATGACAAGG GAATAGCAGC GCCACCTGGT GGCAGTCAGA ATTTTCCAGC
GCAACAACAA
```

1621 GGAAATGCCT GGGTACATGT ACCCTTGTCA CCGCGCACCT TAAATGCGTG
GGTAAAAGCA
1681 GTAGAGGAGA AAAAATTTGG AGCAGAAATA GTACCCATTT TTTTGTTTCA
AGCCCTATCG
1741 AATTCCCGTT TGTGCTAGGG TTCTTAGGCT TCTTGGGGGC TGCTGGAACT
GCAATGGGAG
1801 CAGCGGCGAC AGCCCTGACG GTCCAGTCTC AGCATTTGCT TGCTGGGATA
CTGCAGCAGC
1861 AGAAGAATCT GCTGGCGGCT GTGGAGGCTC AACAGCAGAT GTTGAAGCTG
ACCATTTGGG
1921 GTGTTAAAAA CCTCAATGCC CGCGTCACAG CCCTTGAGAA GTACCTAGAG
GATCAGGCAC
1981 GACTAAACTC CTGGGGGTGC GCATGGAAAC AAGTATGTCA TACCACAGTG
GAGTGGCCCT
2041 GGACAAATCG GACTCCGGAT TGGCAAAATA TGACTTGGTT GGAGTGGGAA
AGACAAATAG
2101 CTGATTTGGA AAGCAACATT ACGAGACAAT TAGTGAAGGC TAGAGAACAA
GAGGAAAAGA
2161 ATCTAGATGC CTATCAGAAG TTAACTAGTT GGTCAGATTT CTGGTCTTGG
TTCGATTTCT
2221 CAAAATGGCT TAACATTTTA AAAATGGGAT TTTTAGTAAT AGTAGGAATA
ATAGGGTTAA
2281 GATTACTTTA CACAGTATAT GGATGTATAG TGAGGGTTAG GCAGGGATAT
GTTCCTCTAT
2341 CTCCACAGAT CCATATCCGC GGCAATTTTA AAAGAAAGGG AGGAATAGGG
GGACAGACTT
2401 CAGCAGAGAG ACTAATTAAT ATAATAACAA CACAATTAGA AATACAACAT
TTACAAACCA
2461 AAATTCAAAA AATTTTAAAT TTTAGAGCCG CGGAGATCTG TTACATAACT
TATGGTAAAT
2521 GGCCTGCCTG GCTGACTGCC CAATGACCCC TGCCCAATGA TGTCAATAAT
GATGTATGTT
2581 CCCATGTAAT GCCAATAGGG ACTTTCCATT GATGTCAATG GGTGGAGTAT
TTATGGTAAC
2641 TGCCCACTTG GCAGTACATC AAGTGTATCA TATGCCAAGT ATGCCCCCTA
TTGATGTCAA
2701 TGATGGTAAA TGGCCTGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
ACTTTCCTAC
2761 TTGGCAGTAC ATCTATGTAT TAGTCATTGC TATTACCATG GGAATTCACT
AGTGGAGAAG
2821 AGCATGCTTG AGGGCTGAGT GCCCCTCAGT GGGCAGAGAG CACATGGCCC
ACAGTCCCTG
2881 AGAAGTTGGG GGGAGGGGTG GGCAATTGAA CTGGTGCCTA GAGAAGGTGG
GGCTTGGGTA
2941 AACTGGGAAA GTGATGTGGT GTACTGGCTC CACCTTTTTC CCCAGGGTGG
GGGAGAACCA
3001 TATATAAGTG CAGTAGTCTC TGTGAACATT CAAGCTTCTG CCTTCTCCCT
CCTGTGAGTT
3061 TGCTAGCCAC CATGCAGAGA AGCCCTCTGG AGAAGGCCTC TGTGGTGAGC
AAGCTGTTCT
3121 TCAGCTGGAC CAGGCCCATC CTGAGGAAGG GCTACAGGCA GAGACTGGAG
CTGTCTGACA
3181 TCTACCAGAT CCCCTCTGTG GACTCTGCTG ACAACCTGTC TGAGAAGCTG
GAGAGGGAGT
3241 GGGATAGAGA GCTGGCCAGC AAGAAGAACC CCAAGCTGAT CAATGCCCTG
AGGAGATGCT
3301 TCTTCTGGAG ATTCATGTTC TATGGCATCT TCCTGTACCT GGGGGAAGTG
ACCAAGGCTG

3361 TGCAGCCTCT GCTGCTGGGC AGAATCATTG CCAGCTATGA CCCTGACAAC
AAGGAGGAGA
3421 GGAGCATTGC CATCTACCTG GGCATTGGCC TGTGCCTGCT GTTCATTGTG
AGGACCCTGC
3481 TGCTGCACCC TGCCATCTTT GGCCTGCACC ACATTGGCAT GCAGATGAGG
ATTGCCATGT
3541 TCAGCCTGAT CTACAAGAAA ACCCTGAAGC TGTCCAGCAG AGTGCTGGAC
AAGATCAGCA
3601 TTGGCCAGCT GGTGAGCCTG CTGAGCAACA ACCTGAACAA GTTTGATGAG
GGCCTGGCCC
3661 TGGCCCACTT TGTGTGGATT GCCCCTCTGC AGGTGGCCCT GCTGATGGGC
CTGATTTGGG
3721 AGCTGCTGCA GGCCTCTGCC TTTTGTGGCC TGGGCTTCCT GATTGTGCTG
GCCCTGTTTC
3781 AGGCTGGCCT GGGCAGGATG ATGATGAAGT ACAGGGACCA GAGGGCAGGC
AAGATCAGTG
3841 AGAGGCTGGT GATCACCTCT GAGATGATTG AGAACATCCA GTCTGTGAAG
GCCTACTGTT
3901 GGGAGGAAGC TATGGAGAAG ATGATTGAAA ACCTGAGGCA GACAGAGCTG
AAGCTGACCA
3961 GGAAGGCTGC CTATGTGAGA TACTTCAACA GCTCTGCCTT CTTCTTCTCT
GGCTTCTTTG
4021 TGGTGTTCCT GTCTGTGCTG CCCTATGCCC TGATCAAGGG GATCATCCTG
AGAAAGATTT
4081 TCACCACCAT CAGCTTCTGC ATTGTGCTGA GGATGGCTGT GACCAGACAG
TTCCCCTGGG
4141 CTGTGCAGAC CTGGTATGAC AGCCTGGGGG CCATCAACAA GATCCAGGAC
TTCCTGCAGA
4201 AGCAGGAGTA CAAGACCCTG GAGTACAACC TGACCACCAC AGAAGTGGTG
ATGGAGAATG
4261 TGACAGCCTT CTGGGAGGAG GGCTTTGGGG AGCTGTTTGA GAAGGCCAAG
CAGAACAACA
4321 ACAACAGAAA GACCAGCAAT GGGGATGACT CCCTGTTCTT CTCCAACTTC
TCCCTGCTGG
4381 GCACACCTGT GCTGAAGGAC ATCAACTTCA AGATTGAGAG GGGGCAGCTG
CTGGCTGTGG
4441 CTGGATCTAC AGGGGCTGGC AAGACCAGCC TGCTGATGAT GATCATGGGG
GAGCTGGAGC
4501 CTTCTGAGGG CAAGATCAAG CACTCTGGCA GGATCAGCTT TTGCAGCCAG
TTCAGCTGGA
4561 TCATGCCTGG CACCATCAAG GAGAACATCA TCTTTGGAGT GAGCTATGAT
GAGTACAGAT
4621 ACAGGAGTGT GATCAAGGCC TGCCAGCTGG AGGAGGACAT CAGCAAGTTT
GCTGAGAAGG
4681 ACAACATTGT GCTGGGGGAG GGAGGCATTA CACTGTCTGG GGGCCAGAGA
GCCAGAATCA
4741 GCCTGGCCAG GGCTGTGTAC AAGGATGCTG ACCTGTACCT GCTGGACTCC
CCCTTTGGCT
4801 ACCTGGATGT GCTGACAGAG AAGGAGATTT TTGAGAGCTG TGTGTGCAAG
CTGATGGCCA
4861 ACAAGACCAG AATCCTGGTG ACCAGCAAGA TGGAGCACCT GAAGAAGGCT
GACAAGATCC
4921 TGATCCTGCA TGAGGGCAGC AGCTACTTCT ATGGGACCTT CTCTGAGCTG
CAGAACCTGC
4981 AGCCTGACTT CAGCTCTAAG CTGATGGGCT GTGACAGCTT TGACCAGTTC
TCTGCTGAGA
5041 GGAGGAACAG CATCCTGACA GAGACCCTGC ACAGATTCAG CCTGGAGGGA
GATGCCCCTG

5101 TGAGCTGGAC AGAGACCAAG AAGCAGAGCT TCAAGCAGAC AGGGGAGTTT
GGGGAGAAGA
5161 GGAAGAACTC CATCCTGAAC CCCATCAACA GCATCAGGAA GTTCAGCATT
GTGCAGAAAA
5221 CCCCCCTGCA GATGAATGGC ATTGAGGAAG ATTCTGATGA GCCCCTGGAG
AGGAGACTGA
5281 GCCTGGTGCC TGATTCTGAG CAGGGAGAGG CCATCCTGCC TAGGATCTCT
GTGATCAGCA
5341 CAGGCCCTAC ACTGCAGGCC AGAAGGAGGC AGTCTGTGCT GAACCTGATG
ACCCACTCTG
5401 TGAACCAGGG CCAGAACATC CACAGGAAAA CCACAGCCTC CACCAGGAAA
GTGAGCCTGG
5461 CCCCTCAGGC CAATCTGACA GAGCTGGACA TCTACAGCAG GAGGCTGTCT
CAGGAGACAG
5521 GCCTGGAGAT TTCTGAGGAG ATCAATGAGG AGGACCTGAA AGAGTGCTTC
TTTGATGACA
5581 TGGAGAGCAT CCCTGCTGTG ACCACCTGGA ACACCTACCT GAGATACATC
ACAGTGCACA
5641 AGAGCCTGAT CTTTGTGCTG ATCTGGTGCC TGGTGATCTT CCTGGCTGAA
GTGGCTGCCT
5701 CTCTGGTGGT GCTGTGGCTG CTGGGAAACA CCCCACTGCA GGACAAGGGC
AACAGCACCC
5761 ACAGCAGGAA CAACAGCTAT GCTGTGATCA TCACCTCCAC CTCCAGCTAC
TATGTGTTCT
5821 ACATCTATGT GGGAGTGGCT GATACCCTGC TGGCTATGGG CTTCTTTAGA
GGCCTGCCCC
5881 TGGTGCACAC ACTGATCACA GTGAGCAAGA TCCTCCACCA CAAGATGCTG
CACTCTGTGC
5941 TGCAGGCTCC TATGAGCACC CTGAATACCC TGAAGGCTGG GGGCATCCTG
AACAGATTCT
6001 CCAAGGATAT TGCCATCCTG GATGACCTGC TGCCTCTCAC CATCTTTGAC
TTCATCCAGC
6061 TGCTGCTGAT TGTGATTGGG GCCATTGCTG TGGTGGCAGT GCTGCAGCCC
TACATCTTTG
6121 TGGCCACAGT GCCTGTGATT GTGGCCTTCA TCATGCTGAG GGCCTACTTT
CTGCAGACCT
6181 CCCAGCAGCT GAAGCAGCTG GAGTCTGAGG GCAGAAGCCC CATCTTCACC
CACCTGGTGA
6241 CAAGCCTGAA GGGCCTGTGG ACCCTGAGAG CCTTTGGCAG GCAGCCCTAC
TTTGAGACCC
6301 TGTTCCACAA GGCCCTGAAC CTGCACACAG CCAACTGGTT CCTCTACCTG
TCCACCCTGA
6361 GATGGTTCCA GATGAGAATT GAGATGATCT TTGTCATCTT CTTCATTGCT
GTGACCTTCA
6421 TCAGCATTCT GACCACAGGA GAGGGAGAGG GCAGAGTGGG CATTATCCTG
ACCCTGGCCA
6481 TGAACATCAT GAGCACACTG CAGTGGGCAG TGAACAGCAG CATTGATGTG
GACAGCCTGA
6541 TGAGGAGTGT GAGCAGAGTG TTCAAGTTCA TTGATATGCC CACAGAGGGC
AAGCCTACCA
6601 AGAGCACCAA GCCCTACAAG AATGGCCAGC TGAGCAAAGT GATGATCATT
GAGAACAGCC
6661 ATGTGAAGAA GGATGATATC TGGCCCAGTG GAGGCCAGAT GACAGTGAAG
GACCTGACAG
6721 CCAAGTACAC AGAGGGGGGC AATGCTATCC TGGAGAACAT CTCCTTCAGC
ATCTCCCCTG
6781 GCCAGAGAGT GGGACTGCTG GGAAGAACAG GCTCTGGCAA GTCTACCCTG
CTGTCTGCCT

6841 TCCTGAGGCT GCTGAACACA GAGGGAGAGA TCCAGATTGA TGGAGTGTCC
TGGGACAGCA
6901 TCACACTGCA GCAGTGGAGG AAGGCCTTTG GTGTGATCCC CCAGAAAGTG
TTCATCTTCA
6961 GTGGCACCTT CAGGAAGAAC CTGGACCCCT ATGAGCAGTG GTCTGACCAG
GAGATTTGGA
7021 AAGTGGCTGA TGAAGTGGGC CTGAGAAGTG TGATTGAGCA GTTCCCTGGC
AAGCTGGACT
7081 TTGTCCTGGT GGATGGGGGC TGTGTGCTGA GCCATGGCCA CAAGCAGCTG
ATGTGCCTGG
7141 CCAGATCAGT GCTGAGCAAG GCCAAGATCC TGCTGCTGGA TGAGCCTTCT
GCCCACCTGG
7201 ATCCTGTGAC CTACCAGATC ATCAGGAGGA CCCTCAAGCA GGCCTTTGCT
GACTGCACAG
7261 TCATCCTGTG TGAGCACAGG ATTGAGGCCA TGCTGGAGTG CCAGCAGTTC
CTGGTGATTG
7321 AGGAGAACAA AGTGAGGCAG TATGACAGCA TCCAGAAGCT GCTGAATGAG
AGGAGCCTGT
7381 TCAGGCAGGC CATCAGCCCC TCTGATAGAG TGAAGCTGTT CCCCCACAGG
AACAGCTCCA
7441 AGTGCAAGAG CAAGCCCCAG ATTGCTGCCC TGAAGGAGGA GACAGAGGAG
GAAGTGCAGG
7501 ACACCAGGCT GTGAGGGCCC AATCAACCTC TGGATTACAA AATTTGTGAA
AGATTGACTG
7561 GTATTCTTAA CTATGTTGCT CCTTTTACGC TATGTGGATA CGCTGCTTTA
ATGCCTTTGT
7621 ATCATGCTAT TGCTTCCCGT ATGGCTTTCA TTTTCTCCTC CTTGTATAAA
TCCTGGTTGC
7681 TGTCTCTTTA TGAGGAGTTG TGGCCCGTTG TCAGGCAACG TGGCGTGGTG
TGCACTGTGT
7741 TTGCTGACGC AACCCCCACT GGTTGGGGCA TTGCCACCAC CTGTCAGCTC
CTTTCCGGGA
7801 CTTTCGCTTT CCCCCTCCCT ATTGCCACGG CGGAACTCAT CGCCGCCTGC
CTTGCCCGCT
7861 GCTGGACAGG GGCTCGGCTG TTGGGCACTG ACAATTCCGT GGTGTTGTCG
GGGAAATCAT
7921 CGTCCTTTCC TTGGCTGCTC GCCTGTGTTG CCACCTGGAT TCTGCGCGGG
ACGTCCTTCT
7981 GCTACGTCCC TTCGGCCCTC AATCCAGCGG ACCTTCCTTC CCGCGGCCTG
CTGCCGGCTC
8041 TGCGGCCTCT TCCGCGTCTT CGCCTTCGCC CTCAGACGAG TCGGATCTCC
CTTTGGGCCG
8101 CCTCCCCGCA AGCTTCGCAC TTTTTAAAAG AAAAGGGAGG ACTGGATGGG
ATTTATTACT
8161 CCGATAGGAC GCTGGCTTGT AACTCAGTCT CTTACTAGGA GACCAGCTTG
AGCCTGGGTG
8221 TTCGCTGGTT AGCCTAACCT GGTTGGCCAC CAGGGGTAAG GACTCCTTGG
CTTAGAAAGC
8281 TAATAAACTT GCCTGCATTA GAGCTCTTAC GCGTCCCGGG CTCGAGATCC
GCATCTCAAT
8341 TAGTCAGCAA CCATAGTCCC GCCCCTAACT CCGCCCATCC CGCCCCTAAC
TCCGCCCAGT
8401 TCCGCCCATT CTCCGCCCCA TGGCTGACTA ATTTTTTTTA TTTATGCAGA
GGCCGAGGCC
8461 GCCTCGGCCT CTGAGCTATT CCAGAAGTAG TGAGGAGGCT TTTTTGGAGG
CCTAGGCTTT
8521 TGCAAAAAGC TAACTTGTTT ATTGCAGCTT ATAATGGTTA CAAATAAAGC
AATAGCATCA

8581 CAAATTTCAC AAATAAAGCA TTTTTTTCAC TGCATTCTAG TTGTGGTTTG
TCCAAACTCA
8641 TCAATGTATC TTATCATGTC TGTCCGCTTC CTCGCTCACT GACTCGCTGC
GCTCGGTCGT
8701 TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT
CCACAGAATC
8761 AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA
GGAACCGTAA
8821 AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC
ATCACAAAAA
8881 TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC
AGGCGTTTCC
8941 CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG
GATACCTGTC
9001 CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA
GGTATCTCAG
9061 TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG
TTCAGCCCGA
9121 CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC
ACGACTTATC
9181 GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG
GCGGTGCTAC
9241 AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA AGAACAGTAT
TTGGTATCTG
9301 CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT
CCGGCAAACA
9361 AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC
GCAGAAAAAA
9421 AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT
GGAACGAAAA
9481 CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ATCTTCACCT
AGATCCTTTT
9541 AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT
GGTCTGACAG
9601 TTAGAAAAAC TCATCGAGCA TCAAATGAAA CTGCAATTTA TTCATATCAG
GATTATCAAT
9661 ACCATATTTT TGAAAAAGCC GTTTCTGTAA TGAAGGAGAA AACTCACCGA
GGCAGTTCCA
9721 TAGGATGGCA AGATCCTGGT ATCGGTCTGC GATTCCGACT CGTCCAACAT
CAATACAACC
9781 TATTAATTTC CCCTCGTCAA AAATAAGGTT ATCAAGTGAG AAATCACCAT
GAGTGACGAC
9841 TGAATCCGGT GAGAATGGCA ACAGCTTATG CATTTCTTTC CAGACTTGTT
CAACAGGCCA
9901 GCCATTACGC TCGTCATCAA AATCACTCGC ATCAACCAAA CCGTTATTCA
TTCGTGATTG
9961 CGCCTGAGCG AGACGAAATA CGCGATCGCT GTTAAAAGGA CAATTACAAA
CAGGAATCGA
10021 ATGCAACCGG CGCAGGAACA CTGCCAGCGC ATCAACAATA TTTTCACCTG
AATCAGGATA
10081 TTCTTCTAAT ACCTGGAATG CTGTTTTTCC GGGGATCGCA GTGGTGAGTA
ACCATGCATC
10141 ATCAGGAGTA CGGATAAAAT GCTTGATGGT CGGAAGAGGC ATAAATTCCG
TCAGCCAGTT
10201 TAGTCTGACC ATCTCATCTG TAACATCATT GGCAACGCTA CCTTTGCCAT
GTTTCAGAAA
10261 CAACTCTGGC GCATCGGGCT TCCCATACAA TCGATAGATT GTCGCACCTG
ATTGCCCGAC

10321 ATTATCGCGA GCCCATTTAT ACCCATATAA ATCAGCATCC ATGTTGGAAT
TTAATCGCGG
10381 CCTAGAGCAA GACGTTTCCC GTTGAATATG GCTCATAACA CCCCTTGTAT
TACTGTTTAT
10441 GTAAGCAGAC AGTTTTATTG TTCATGATGA TATATTTTTA TCTTGTGCAA
TGTAACATCA
10501 GAGATTTTGA GACACAACAA TTGGTCGACG GATCC

SEQ ID NO: 2

1 GCTCGAGACT AGTGACTTGG TGAGTAGGCT TCGAGCCTAG TTAGAGGACT
AGGAGAGGCC
61 GTAGCCGTAA CTACTCTGGG CAAGTAGGGC AGGCGGTGGG TACGCAATGG
GGGCGGCTAC
121 CTCAGCACTA AATAGGAGAC AATTAGACCA ATTTGAGAAA ATACGACTTC
GCCCGAACGG
181 AAAGAAAAAG TACCAAATTA AACATTTAAT ATGGGCAGGC AAGGAGATGG
AGCGCTTCGG
241 CCTCCATGAG AGGTTGTTGG AGACAGAGGA GGGGTGTAAA AGAATCATAG
AAGTCCTCTA
301 CCCCCTAGAA CCAACAGGAT CGGAGGGCTT AAAAAGTCTG TTCAATCTTG
TGTGCGTACT
361 ATATTGCTTG CACAAGGAAC AGAAAGTGAA AGACACAGAG GAAGCAGTAG
CAACAGTAAG
421 ACAACACTGC CATCTAGTGG AAAAAGAAAA AAGTGCAACA GAGACATCTA
GTGGACAAAA
481 GAAAAATGAC AAGGGAATAG CAGCGCCACC TGGTGGCAGT CAGAATTTTC
CAGCGCAACA
541 ACAAGGAAAT GCCTGGGTAC ATGTACCCTT GTCACCGCGC ACCTTAAATG
CGTGGGTAAA
601 AGCAGTAGAG GAGAAAAAAT TTGGAGCAGA AATAGTACCC ATGTTTCAAG
CCCTATCAGA
661 AGGCTGCACA CCCTATGACA TTAATCAGAT GCTTAATGTG CTAGGAGATC
ATCAAGGGGC
721 ATTACAAATA GTGAAAGAGA TCATTAATGA AGAAGCAGCC CAGTGGGATG
TAACACACCC
781 ACTACCCGCA GGACCCCTAC CAGCAGGACA GCTCAGGGAC CCTCGCGGCT
CAGATATAGC
841 AGGGACCACC AGCTCAGTAC AAGAACAGTT AGAATGGATC TATACTGCTA
ACCCCCGGGT
901 AGATGTAGGT GCCATCTACC GGAGATGGAT TATTCTAGGA CTTCAAAAGT
GTGTCAAAAT
961 GTACAACCCA GTATCAGTCC TAGACATTAG GCAGGGACCT AAAGAGCCCT
TCAAGGATTA
1021 TGTGGACAGA TTTTACAAGG CAATTAGAGC AGAACAAGCC TCAGGGGAAG
TGAAACAATG
1081 GATGACAGAA TCATTACTCA TTCAAAATGC TAATCCAGAT TGTAAGGTCA
TCCTGAAGGG
1141 CCTAGGAATG CACCCCACCC TTGAAGAAAT GTTAACGGCT TGTCAGGGGG
TAGGAGGCCC
1201 AAGCTACAAA GCAAAAGTAA TGGCAGAAAT GATGCAGACC ATGCAAAATC
AAAACATGGT
1261 GCAGCAGGGA GGTCCAAAAA GACAAAGACC CCCACTAAGA TGTTATAATT
GTGGAAAATT
1321 TGGCCATATG CAAAGACAAT GTCCGGAACC AAGGAAAACA AAATGTCTAA
AGTGTGGAAA
1381 ATTGGGACAC CTAGCAAAAG ACTGCAGGGG ACAGGTGAAT TTTTTAGGGT
ATGGACGGTG

1441 GATGGGGGCA AAACCGAGAA ATTTTCCCGC CGCTACTCTT GGAGCGGAAC
CGAGTGCGCC
1501 TCCTCCACCG AGCGGCACCA CCCCATACGA CCCAGCAAAG AAGCTCCTGC
AGCAATATGC
1561 AGAGAAAGGG AAACAACTGA GGGAGCAAAA GAGGAATCCA CCGGCAATGA
ATCCGGATTG
1621 GACCGAGGGA TATTCTTTGA ACTCCCTCTT TGGAGAAGAC CAATAAAGAC
AGTGTATATA
1681 GAAGGGGTCC CCATTAAGGC ACTGCTAGAC ACAGGGGCAG ATGACACCAT
AATTAAAGAA
1741 AATGATTTAC AATTATCAGG TCCATGGAGA CCCAAAATTA TAGGGGGCAT
AGGAGGAGGC
1801 CTTAATGTAA AAGAATATAA CGACAGGGAA GTAAAAATAG AAGATAAAAT
TTTGAGAGGA
1861 ACAATATTGT TAGGAGCAAC TCCCATTAAT ATAATAGGTA GAAATTTGCT
GGCCCCGGCA
1921 GGTGCCCGGT TAGTAATGGG ACAATTATCA GAAAAAATTC CTGTCACACC
TGTCAAATTG
1981 AAGGAAGGGG CTCGGGGACC CTGTGTAAGA CAATGGCCTC TCTCTAAAGA
GAAGATTGAA
2041 GCTTTACAGG AAATATGTTC CCAATTAGAG CAGGAAGGAA AAATCAGTAG
AGTAGGAGGA
2101 GAAAATGCAT ACAATACCCC AATATTTTGC ATAAAGAAGA AGGACAAATC
CCAGTGGAGG
2161 ATGCTAGTAG ACTTTAGAGA GTTAAATAAG GCAACCCAAG ATTTCTTTGA
AGTGCAATTA
2221 GGGATACCCC ACCCAGCAGG ATTAAGAAAG ATGAGACAGA TAACAGTTTT
AGATGTAGGA
2281 GACGCCTATT ATTCCATACC ATTGGATCCA AATTTTAGGA AATATACTGC
TTTTACTATT
2341 CCCACAGTGA ATAATCAGGG ACCCGGGATT AGGTATCAAT TCAACTGTCT
CCCGCAAGGG
2401 TGGAAAGGAT CTCCTACAAT CTTCCAAAAT ACAGCAGCAT CCATTTTGGA
GGAGATAAAA
2461 AGAAACTTGC CAGCACTAAC CATTGTACAA TACATGGATG ATTTATGGGT
AGGTTCTCAA
2521 GAAAATGAAC ACACCCATGA CAAATTAGTA GAACAGTTAA GAACAAAATT
ACAAGCCTGG
2581 GGCTTAGAAA CCCCAGAAAA GAAGGTGCAA AAAGAACCAC CTTATGAGTG
GATGGGATAC
2641 AAACTTTGGC CTCACAAATG GGAACTAAGC AGAATACAAC TGGAGGAAAA
AGATGAATGG
2701 ACTGTCAATG ACATCCAGAA GTTAGTTGGG AAACTAAATT GGGCAGCACA
ATTGTATCCA
2761 GGTCTTAGGA CCAAGAATAT ATGCAAGTTA ATTAGAGGAA AGAAAAATCT
GTTAGAGCTA
2821 GTGACTTGGA CACCTGAGGC AGAAGCTGAA TATGCAGAAA ATGCAGAGAT
TCTTAAAACA
2881 GAACAGGAAG GAACCTATTA CAAACCAGGA ATACCTATTA GGGCAGCAGT
ACAGAAATTG
2941 GAAGGAGGAC AGTGGAGTTA CCAATTCAAA CAAGAAGGAC AAGTCTTGAA
AGTAGGAAAA
3001 TACACCAAGC AAAAGAACAC CCATACAAAT GAACTTCGCA CATTAGCTGG
TTTAGTGCAG
3061 AAGATTTGCA AAGAAGCTCT AGTTATTTGG GGGATATTAC CAGTTCTAGA
ACTCCCGATA
3121 GAAAGAGAGG TATGGGAACA ATGGTGGGCG GATTACTGGC AGGTAAGCTG
GATTCCCGAA

3181 TGGGATTTTG TCAGCACCCC ACCTTTGCTC AAACTATGGT ACACATTAAC
AAAAGAACCC
3241 ATACCCAAGG AGGACGTTTA CTATGTAGAT GGAGCATGCA ACAGAAATTC
AAAAGAAGGA
3301 AAAGCAGGAT ACATCTCACA ATACGGAAAA CAGAGAGTAG AAACATTAGA
AAACACTACC
3361 AATCAGCAAG CAGAATTAAC AGCTATAAAA ATGGCTTTGG AAGACAGTGG
GCCTAATGTG
3421 AACATAGTAA CAGACTCTCA ATATGCAATG GGAATTTTGA CAGCACAACC
CACACAAAGT
3481 GATTCACCAT TAGTAGAGCA AATTATAGCC TTAATGATAC AAAAGCAACA
AATATATTTG
3541 CAGTGGGTAC CAGCACATAA AGGAATAGGA GGAAATGAGG AGATAGATAA
ATTAGTGAGT
3601 AAAGGCATTA GAAGAGTTTT ATTCTTAGAA AAAATAGAAG AAGCTCAAGA
AGAGCATGAA
3661 AGATATCATA ATAATTGGAA AAACCTAGCA GATACATATG GGCTTCCACA
AATAGTAGCA
3721 AAAGAGATAG TGGCCATGTG TCCAAAATGT CAGATAAAGG GAGAACCAGT
GCATGGACAA
3781 GTGGATGCCT CACCTGGAAC ATGGCAGATG GATTGTACTC ATCTAGAAGG
AAAAGTAGTC
3841 ATAGTTGCGG TCCATGTAGC CAGTGGATTC ATAGAAGCAG AAGTCATACC
TAGGGAAACA
3901 GGAAAAGAAA CGGCAAAGTT TCTATTAAAA ATACTGAGTA GATGGCCTAT
AACACAGTTA
3961 CACACAGACA ATGGGCCTAA CTTTACCTCC CAAGAAGTGG CAGCAATATG
TTGGTGGGGA
4021 AAAATTGAAC ATACAACAGG TATACCATAT AACCCCCAAT CTCAAGGATC
AATAGAAAGC
4081 ATGAACAAAC AATTAAAAGA GATAATTGGG AAAATAAGAG ATGATTGCCA
ATATACAGAG
4141 ACAGCAGTAC TGATGGCTTG CCATATTCAC AATTTTAAAA GAAAGGGAGG
AATAGGGGGA
4201 CAGACTTCAG CAGAGAGACT AATTAATATA ATAACAACAC AATTAGAAAT
ACAACATTTA
4261 CAAACCAAAA TTCAAAAAAT TTTAAATTTT AGAGTCTACT ACAGAGAAGG
GAGAGACCCT
4321 GTGTGGAAAG GACCAGCACA ATTAATCTGG AAAGGGGAAG GAGCAGTGGT
CCTCAAGGAC
4381 GGAAGTGACC TAAAGGTTGT ACCAAGAAGG AAAGCTAAAA TTATTAAGGA
TTATGAACCC
4441 AAACAAAGAG TGGGTAATGA GGGTGACGTG GAAGGTACCA GGGGATCTGA
TAACTAAATG
4501 GCAGGGAATA GTCAGATATT GGATGAGACA AAGAAATTTG AAATGGAACT
ATTATATGCA
4561 TCAGCTGGCG GCCGCGAATT CACTAGTGAT TCCCGTTTGT GCTAGGGTTC
TTAGGCTTCT
4621 TGGGGGCTGC TGGAACTGCA ATGGGAGCAG CGGCGACAGC CCTGACGGTC
CAGTCTCAGC
4681 ATTTGCTTGC TGGGATACTG CAGCAGCAGA AGAATCTGCT GGCGGCTGTG
GAGGCTCAAC
4741 AGCAGATGTT GAAGCTGACC ATTTGGGGTG TTAAAAACCT CAATGCCCGC
GTCACAGCCC
4801 TTGAGAAGTA CCTAGAGGAT CAGGCACGAC TAAACTCCTG GGGGTGCGCA
TGGAAACAAG
4861 TATGTCATAC CACAGTGGAG TGGCCCTGGA CAAATCGGAC TCCGGATTGG
CAAAATATGA

4921 CTTGGTTGGA GTGGGAAAGA CAAATAGCTG ATTTGGAAAG CAACATTACG
AGACAATTAG
4981 TGAAGGCTAG AGAACAAGAG GAAAAGAATC TAGATGCCTA TCAGAAGTTA
ACTAGTTGGT
5041 CAGATTTCTG GTCTTGGTTC GATTTCTCAA AATGGCTTAA CATTTTAAAA
ATGGGATTTT
5101 TAGTAATAGT AGGAATAATA GGGTTAAGAT TACTTTACAC AGTATATGGA
TGTATAGTGA
5161 GGGTTAGGCA GGGATATGTT CCTCTATCTC CACAGATCCA TATCCAATCG
AATTCCCGCG
5221 GCCGCAATTC ACTCCTCAGG TGCAGGCTGC CTATCAGAAG GTGGTGGCTG
GTGTGGCCAA
5281 TGCCCTGGCT CACAAATACC ACTGAGATCT TTTTCCCTCT GCCAAAAATT
ATGGGGACAT
5341 CATGAAGCCC CTTGAGCATC TGACTTCTGG CTAATAAAGG AAATTTATTT
TCATTGCAAT
5401 AGTGTGTTGG AATTTTTTGT GTCTCTCACT CGGAAGGACA TATGGGAGGG
CAAATCATTT
5461 AAAACATCAG AATGAGTATT TGGTTTAGAG TTTGGCAACA TATGCCCATA
TGCTGGCTGC
5521 CATGAACAAA GGTTGGCTAT AAAGAGGTCA TCAGTATATG AAACAGCCCC
CTGCTGTCCA
5581 TTCCTTATTC CATAGAAAAG CCTTGACTTG AGGTTAGATT TTTTTTATAT
TTTGTTTTGT
5641 GTTATTTTTT TCTTTAACAT CCCTAAAATT TTCCTTACAT GTTTTACTAG
CCAGATTTTTT
5701 CCTCCTCTCC TGACTACTCC CAGTCATAGC TGTCCCTCTT CTCTTATGGA
GATCCCTCGA
5761 CCTGCAGCCC AAGCTTGGCG TAATCATGGT CATAGCTGTT TCCTGTGTGA
AATTGTTATC
5821 CGCTCACAAT TCCACACAAC ATACGAGCCG GAAGCATAAA GTGTAAAGCC
TGGGGTGCCT
5881 AATGAGTGAG CTAACTCACA TTAATTGCGT TGCGCTCACT GCCCGCTTTC
CAGTCGGGAA
5941 ACCTGTCGTG CCAGCGGATC CGCATCTCAA TTAGTCAGCA ACCATAGTCC
CGCCCCTAAC
6001 TCCGCCCATC CCGCCCCTAA CTCCGCCCAG TTCCGCCCAT TCTCCGCCCC
ATGGCTGACT
6061 AATTTTTTTT ATTTATGCAG AGGCCGAGGC CGCCTCGGCC TCTGAGCTAT
TCCAGAAGTA
6121 GTGAGGAGGC TTTTTTGGAG GCCTAGGCTT TTGCAAAAAG CTAACTTGTT
TATTGCAGCT
6181 TATAATGGTT ACAAATAAAG CAATAGCATC ACAAATTTCA CAAATAAAGC
ATTTTTTTCA
6241 CTGCATTCTA GTTGTGGTTT GTCCAAACTC ATCAATGTAT CTTATCATGT
CTGTCCGCTT
6301 CCTCGCTCAC TGACTCGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA
TCAGCTCACT
6361 CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG
AACATGTGAG
6421 CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC GTTGCTGGCG
TTTTTCCATA
6481 GGCTCCGCCC CCCTGACGAG CATCACAAAA ATCGACGCTC AAGTCAGAGG
TGGCGAAACC
6541 CGACAGGACT ATAAAGATAC CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG
CGCTCTCCTG
6601 TTCCGACCCT GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA
AGCGTGGCGC

6661 TTTCTCATAG CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCGTTCGC
TCCAAGCTGG
6721 GCTGTGTGCA CGAACCCCCC GTTCAGCCCG ACCGCTGCGC CTTATCCGGT
AACTATCGTC
6781 TTGAGTCCAA CCCGGTAAGA CACGACTTAT CGCCACTGGC AGCAGCCACT
GGTAACAGGA
6841 TTAGCAGAGC GAGGTATGTA GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG
CCTAACTACG
6901 GCTACACTAG AAGAACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT
ACCTTCGGAA
6961 AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT
GGTTTTTTTG
7021 TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA AGAAGATCCT
TTGATCTTTT
7081 CTACGGGGTC TGACGCTCAG TGGAACGAAA ACTCACGTTA AGGGATTTTG
GTCATGAGAT
7141 TATCAAAAAG GATCTTCACC TAGATCCTTT TAAATTAAAA ATGAAGTTTT
AAATCAATCT
7201 AAAGTATATA TGAGTAAACT TGGTCTGACA GTTAGAAAAA CTCATCGAGC
ATCAAATGAA
7261 ACTGCAATTT ATTCATATCA GGATTATCAA TACCATATTT TTGAAAAAGC
CGTTTCTGTA
7321 ATGAAGGAGA AAACTCACCG AGGCAGTTCC ATAGGATGGC AAGATCCTGG
TATCGGTCTG
7381 CGATTCCGAC TCGTCCAACA TCAATACAAC CTATTAATTT CCCCTCGTCA
AAAATAAGGT
7441 TATCAAGTGA GAAATCACCA TGAGTGACGA CTGAATCCGG TGAGAATGGC
AACAGCTTAT
7501 GCATTTCTTT CCAGACTTGT TCAACAGGCC AGCCATTACG CTCGTCATCA
AAATCACTCG
7561 CATCAACCAA ACCGTTATTC ATTCGTGATT GCGCCTGAGC GAGACGAAAT
ACGCGATCGC
7621 TGTTAAAAGG ACAATTACAA ACAGGAATCG AATGCAACCG GCGCAGGAAC
ACTGCCAGCG
7681 CATCAACAAT ATTTTCACCT GAATCAGGAT ATTCTTCTAA TACCTGGAAT
GCTGTTTTTC
7741 CGGGGATCGC AGTGGTGAGT AACCATGCAT CATCAGGAGT ACGGATAAAA
TGCTTGATGG
7801 TCGGAAGAGG CATAAATTCC GTCAGCCAGT TTAGTCTGAC CATCTCATCT
GTAACATCAT
7861 TGGCAACGCT ACCTTTGCCA TGTTTCAGAA ACAACTCTGG CGCATCGGGC
TTCCCATACA
7921 ATCGATAGAT TGTCGCACCT GATTGCCCGA CATTATCGCG AGCCCATTTA
TACCCATATA
7981 AATCAGCATC CATGTTGGAA TTTAATCGCG GCCTAGAGCA AGACGTTTCC
CGTTGAATAT
8041 GGCTCATAAC ACCCCTTGTA TTACTGTTTA TGTAAGCAGA CAGTTTTATT
GTTCATGATG
8101 ATATATTTTT ATCTTGTGCA ATGTAACATC AGAGATTTTG AGACACAACA
ATTGTCGACA
8161 TTGATTATTG ACTAGTTATT AATAGTAATC AATTACGGGG TCATTAGTTC
ATAGCCCATA
8221 TATGGAGTTC CGCGTTACAT AACTTACGGT AAATGGCCCG CCTGGCTGAC
CGCCCAACGA
8281 CCCCCGCCCA TTGACGTCAA TAATGACGTA TGTTCCCATA GTAACGCCAA
TAGGGACTTT
8341 CCATTGACGT CAATGGGTGG AGTATTTACG GTAAACTGCC CACTTGGCAG
TACATCAAGT

8401 GTATCATATG CCAAGTACGC CCCCTATTGA CGTCAATGAC GGTAAATGGC
CCGCCTGGCA
8461 TTATGCCCAG TACATGACCT TATGGGACTT TCCTACTTGG CAGTACATCT
ACGTATTAGT
8521 CATCGCTATT ACCATGGTCG AGGTGAGCCC CACGTTCTGC TTCACTCTCC
CCATCTCCCC
8581 CCCCTCCCCA CCCCCAATTT TGTATTTATT TATTTTTTAA TTATTTTGTG
CAGCGATGGG
8641 GGCGGGGGGG GGGGGGGGGC GCGCGCCAGG CGGGGCGGGG CGGGGCGAGG
GGCGGGGCGG
8701 GGCGAGGCGG AGAGGTGCGG CGGCAGCCAA TCAGAGCGGC GCGCTCCGAA
AGTTTCCTTT
8761 TATGGCGAGG CGGCGGCGGC GGCGGCCCTA TAAAAAGCGA AGCGCGCGGC
GGGCGGGAGT
8821 CGCTGCGCGC TGCCTTCGCC CCGTGCCCCG CTCCGCCGCC GCCTCGCGCC
GCCCGCCCCG
8881 GCTCTGACTG ACCGCGTTAC TCCCACAGGT GAGCGGGCGG GACGGCCCTT
CTCCTCCGGG
8941 CTGTAATTAG CGCTTGGTTT AATGACGGCT TGTTTCTTTT CTGTGGCTGC
GTGAAAGCCT
9001 TGAGGGGCTC CGGGAGGGCC CTTTGTGCGG GGGGAGCGGC TCGGGGGGTG
CGTGCGTGTG
9061 TGTGTGCGTG GGGAGCGCCG CGTGCGGCTC CGCGCTGCCC GGCGGCTGTG
AGCGCTGCGG
9121 GCGCGGCGCG GGGCTTTGTG CGCTCCGCAG TGTGCGCGAG GGGAGCGCGG
CCGGGGGCGG
9181 TGCCCCGCGG TGCGGGGGGG GCTGCGAGGG GAACAAAGGC TGCGTGCGGG
GTGTGTGCGT
9241 GGGGGGGTGA GCAGGGGGTG TGGGCGCGTC GGTCGGGCTG CAACCCCCCC
TGCACCCCCC
9301 TCCCCGAGTT GCTGAGCACG GCCCGGCTTC GGGTGCGGGG CTCCGTACGG
GGCGTGGCGC
9361 GGGGCTCGCC GTGCCGGGCG GGGGGTGGCG GCAGGTGGGG GTGCCGGGCG
GGGCGGGGCC
9421 GCCTCGGGCC GGGGAGGGCT CGGGGGAGGG GCGCGGCGGC CCCCGGAGCG
CCGGCGGCTG
9481 TCGAGGCGCG GCGAGCCGCA GCCATTGCCT TTTATGGTAA TCGTGCGAGA
GGGCGCAGGG
9541 ACTTCCTTTG TCCCAAATCT GTGCGGAGCC GAAATCTGGG AGGCGCCGCC
GCACCCCCTC
9601 TAGCGGGCGC GGGGCGAAGC GGTGCGGCGC CGGCAGGAAG GAAATGGGCG
GGGAGGGCCT
9661 TCGTGCGTCG CCGCGCCGCC GTCCCCTTCT CCCTCTCCAG CCTCGGGGCT
GTCCGCGGGG
9721 GGACGGCTGC CTTCGGGGGG GACGGGGCAG GGCGGGGTTC GGCTTCTGGC
GTGTGACCGG
9781 CGGCTCTAGA GCCTCTGCTA ACCATGTTCA TGCCTTCTTC TTTTTCCTAC
AGCTCCTGGG
9841 CAACGTGCTG GTTATTGTGC TGTCTCATCA TTTTGGCAAA GAATT

SEQ ID NO: 3

1 TCAATATTGG CCATTAGCCA TATTATTCAT TGGTTATATA GCATAAATCA
ATATTGGCTA
61 TTGGCCATTG CATACGTTGT ATCTATATCA TAATATGTAC ATTTATATTG
GCTCATGTCC
121 AATATGACCG CCATGTTGGC ATTGATTATT GACTAGTTAT TAATAGTAAT
CAATTACGGG

181 GTCATTAGTT CATAGCCCAT ATATGGAGTT CCGCGTTACA TAACTTACGG
TAAATGGCCC
241 GCCTGGCTGA CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT
ATGTTCCCAT
301 AGTAACGCCA ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC
GGTAAACTGC
361 CCACTTGGCA GTACATCAAG TGTATCATAT GCCAAGTCCG CCCCCTATTG
ACGTCAATGA
421 CGGTAAATGG CCCGCCTGGC ATTATGCCCA GTACATGACC TTACGGGACT
TTCCTACTTG
481 GCAGTACATC TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT
GGCAGTACAC
541 CAATGGGCGT GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC
CCATTGACGT
601 CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC
GTAATAACCC
661 CGCCCCGTTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA
TAAGCAGAGC
721 TCGTTTAGTG AACCGTCAGA TCACTAGAAG CTTTATTGCG GTAGTTTATC
ACAGTTAAAT
781 TGCTAACGCA GTCAGTGCTT CTGACACAAC AGTCTCGAAC TTAAGCTGCA
GAAGTTGGTC
841 GTGAGGCACT GGGCAGGTAA GTATCAAGGT TACAAGACAG GTTTAAGGAG
ACCAATAGAA
901 ACTGGGCTTG TCGAGACAGA GAAGACTCTT GCGTTTCTGA TAGGCACCTA
TTGGTCTTAC
961 TGACATCCAC TTTGCCTTTC TCTCCACAGG TGTCCACTCC CAGTTCAATT
ACAGCTCTTA
1021 AGGCTAGAGT ACTTAATACG ACTCACTATA GGCTAGCCTC GAGAATTCGA
TTATGCCCCT
1081 AGGACCAGAA GAAAGAAGAT TGCTTCGCTT GATTTGGCTC CTTTACAGCA
CCAATCCATA
1141 TCCACCAAGT GGGGAAGGGA CGGCCAGACA ACGCCGACGA GCCAGGAGAA
GGTGGAGACA
1201 ACAGCAGGAT CAAATTAGAG TCTTGGTAGA AAGACTCCAA GAGCAGGTGT
ATGCAGTTGA
1261 CCGCCTGGCT GACGAGGCTC AACACTTGGC TATACAACAG TTGCCTGACC
CTCCTCATTC
1321 AGCTTAGAAT CACTAGTGAA TTCACGCGTG GTACCTCTAG AGTCGACCCG
GGCGGCCGCT
1381 TCGAGCAGAC ATGATAAGAT ACATTGATGA GTTTGGACAA ACCACAACTA
GAATGCAGTG
1441 AAAAAAATGC TTTATTTGTG AAATTTGTGA TGCTATTGCT TTATTTGTAA
CCATTATAAG
1501 CTGCAATAAA CAAGTTAACA ACAACAATTG CATTCATTTT ATGTTTCAGG
TTCAGGGGGA
1561 GATGTGGGAG GTTTTTTAAA GCAAGTAAAA CCTCTACAAA TGTGGTAAAA
TCGATAAGGA
1621 TCCGTCGACC AATTGTTGTG TCTCAAAATC TCTGATGTTA CATTGCACAA
GATAAAAATA
1681 TATCATCATG AACAATAAAA CTGTCTGCTT ACATAAACAG TAATACAAGG
GGTGTTATGA
1741 GCCATATTCA ACGGGAAACG TCTTGCTCTA GGCCGCGATT AAATTCCAAC
ATGGATGCTG
1801 ATTTATATGG GTATAAATGG GCTCGCGATA ATGTCGGGCA ATCAGGTGCG
ACAATCTATC
1861 GATTGTATGG GAAGCCCGAT GCGCCAGAGT TGTTTCTGAA ACATGGCAAA
GGTAGCGTTG

1921 CCAATGATGT TACAGATGAG ATGGTCAGAC TAAACTGGCT GACGGAATTT
ATGCCTCTTC
1981 CGACCATCAA GCATTTTATC CGTACTCCTG ATGATGCATG GTTACTCACC
ACTGCGATCC
2041 CCGGAAAAAC AGCATTCCAG GTATTAGAAG AATATCCTGA TTCAGGTGAA
AATATTGTTG
2101 ATGCGCTGGC AGTGTTCCTG CGCCGGTTGC ATTCGATTCC TGTTTGTAAT
TGTCCTTTTA
2161 ACAGCGATCG CGTATTTCGT CTCGCTCAGG CGCAATCACG AATGAATAAC
GGTTTGGTTG
2221 ATGCGAGTGA TTTTGATGAC GAGCGTAATG GCTGGCCTGT TGAACAAGTC
TGGAAAGAAA
2281 TGCATAAGCT GTTGCCATTC TCACCGGATT CAGTCGTCAC TCATGGTGAT
TTCTCACTTG
2341 ATAACCTTAT TTTTGACGAG GGGAAATTAA TAGGTTGTAT TGATGTTGGA
CGAGTCGGAA
2401 TCGCAGACCG ATACCAGGAT CTTGCCATCC TATGGAACTG CCTCGGTGAG
TTTTCTCCTT
2461 CATTACAGAA ACGGCTTTTT CAAAAATATG GTATTGATAA TCCTGATATG
AATAAATTGC
2521 AGTTTCATTT GATGCTCGAT GAGTTTTTCT AACTGTCAGA CCAAGTTTAC
TCATATATAC
2581 TTTAGATTGA TTTAAAACTT CATTTTTAAT TTAAAAGGAT CTAGGTGAAG
ATCCTTTTTG
2641 ATAATCTCAT GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG
TCAGACCCCG
2701 TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC
TGCTGCTTGC
2761 AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG
CTACCAACTC
2821 TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTT
CTTCTAGTGT
2881 AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC
CTCGCTCTGC
2941 TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC
GGGTTGGACT
3001 CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT
TCGTGCACAC
3061 AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT
GAGCTATGAG
3121 AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC
GGCAGGGTCG
3181 GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT
TATAGTCCTG
3241 TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA
GGGGGGCGGA
3301 GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT
TGCTGGCCTT
3361 TTGCTCACAT GGCTCGACAG ATCT

SEQ ID NO: 4

1 ATTGATTATT GACTAGTTAT TAATAGTAAT CAATTACGGG GTCATTAGTT
CATAGCCCAT
61 ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA
CCGCCCAACG

121 ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA
ATAGGGACTT
181 TCCATTGACG TCAATGGGTG GAGTATTTAC GGTAAACTGC CCACTTGGCA
GTACATCAAG
241 TGTATCATAT GCCAAGTACG CCCCCTATTG ACGTCAATGA CGGTAAATGG
CCCGCCTGGC
301 ATTATGCCCA GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC
TACGTATTAG
361 TCATCGCTAT TACCATGGTC GAGGTGAGCC CCACGTTCTG CTTCACTCTC
CCCATCTCCC
421 CCCCCTCCCC ACCCCCAATT TTGTATTTAT TTATTTTTTA ATTATTTTGT
GCAGCGATGG
481 GGGCGGGGGG GGGGGGGGGG CGCGCGCCAG GCGGGGCGGG GCGGGGCAAG
GGGCGGGGCG
541 GGGCGAGGCG GAAAGGTGCG GCGGCAGCCA ATCAAAGCGG CGCGCTCCGA
AAGTTTCCTT
601 TTATGGCGAG GCGGCGGCGG CGGCGGCCCT ATAAAAAGCG AAGCGCGCGG
CGGGCGGGAG
661 TCGCTGCGCG CTGCCTTCGC CCCGTGCCCC GCTCCGCCGC CGCCTCGCGC
CGCCCGCCCC
721 GGCTCTGACT GACCGCGTTA CTCCCACAGG TGAGCGGGCG GGACGGCCCT
TCTCCTCCGG
781 GCTGTAATTA GCGCTTGGTT TAATGACGGC TTGTTTCTTT TCTGTGGCTG
CGTGAAAGCC
841 TTGAGGGGCT CCGGGAGGGC CCTTTGTGCG GGGGGAGCGG CTCGGGGGGT
GCGTGCGTGT
901 GTGTGTGCGT GGGGAGCGCC GCGTGCGGCT CCGCGCTGCC CGGCGGCTGT
GAGCGCTGCG
961 GGCGCGGCGC GGGGCTTTGT GCGCTCCGCA GTGTGCGCGA GGGGAGCGCG
GCCGGGGGCG
1021 GTGCCCCGCG GTGCGGGGGG GGCTGCGAGG GGAACAAAGG CTGCGTGCGG
GGTGTGTGCG
1081 TGGGGGGGTG AGCAGGGGGT GTGGGCGCGT CGGTCGGGCT GCAACCCCCC
CTGCACCCCC
1141 CTCCCCGAGT TGCTGAGCAC GGCCCGGCTT CGGGTGCGGG GCTCCGTACG
GGGCGTGGCG
1201 CGGGGCTCGC CGTGCCGGGC GGGGGGTGGC GGCAGGTGGG GGTGCCGGGC
GCGGCGGGGC
1261 CGCCTCGGGC CGGGGAGGGC TCGGGGGAGG GGCGCGGCGG CCCCCGGAGC
GCCGGCGGCT
1321 GTCGAGGCGC GGCGAGCCGC AGCCATTGCC TTTTATGGTA ATCGTGCGAG
AGGGCGCAGG
1381 GACTTCCTTT GTCCCAAATC TGTGCGGAGC CGAAATCTGG GAGGCGCCGC
CGCACCCCCT
1441 CTAGCGGGCG CGGGGCGAAG CGGTGCGGCG CCGGCAGGAA GGAAATGGGC
GGGGAGGGCC
1501 TTCGTGCGTC GCCGCGCCGC CGTCCCCTTC TCCCTCTCCA GCCTCGGGGC
TGTCCGCGGG
1561 GGGACGGCTG CCTTCGGGGG GGACGGGGCA GGGCGGGGTT CGGCTTCTGG
CGTGTGACCG
1621 GCGGCTCTAG AGCCTCTGCT AACCATGTTC ATGCCTTCTT CTTTTTCCTA
CAGCTCCTGG
1681 GCAACGTGCT GGTTATTGTG CTGTCTCATC ATTTTGGCAA AGAATTCGAT
TGCCATGGCA
1741 ACATATATCC AGAGAGTACA GTGCATCTCA ACATCACTAC TGGTTGTTCT
CACCACATTG
1801 GTCTCGTGTC AGATTCCCAG GGATAGGCTC TCTAACATAG GGGTCATAGT
CGATGAAGGG

1861 AAATCACTGA AGATAGCTGG ATCCCACGAA TCGAGGTACA TAGTACTGAG
TCTAGTTCCG
1921 GGGGTAGACT TTGAGAATGG GTGCGGAACA GCCCAGGTTA TCCAGTACAA
GAGCCTACTG
1981 AACAGGCTGT TAATCCCATT GAGGGATGCC TTAGATCTTC AGGAGGCTCT
GATAACTGTC
2041 ACCAATGATA CGACACAAAA TGCCGGTGCT CCCCAGTCGA GATTCTTCGG
TGCTGTGATT
2101 GGTACTATCG CACTTGGAGT GGCGACATCA GCACAAATCA CCGCAGGGAT
TGCACTAGCC
2161 GAAGCGAGGG AGGCCAAAAG AGACATAGCG CTCATCAAAG AATCGATGAC
AAAAACACAC
2221 AAGTCTATAG AACTGCTGCA AAACGCTGTG GGGGAACAAA TTCTTGCTCT
AAAGACACTC
2281 CAGGATTTCG TGAATGATGA GATCAAACCC GCAATAAGCG AATTAGGCTG
TGAGACTGCT
2341 GCCTTAAGAC TGGGTATAAA ATTGACACAG CATTACTCCG AGCTGTTAAC
TGCGTTCGGC
2401 TCGAATTTCG GAACCATCGG AGAGAAGAGC CTCACGCTGC AGGCGCTGTC
TTCACTTTAC
2461 TCTGCTAACA TTACTGAGAT TATGACCACA ATCAGGACAG GGCAGTCTAA
CATCTATGAT
2521 GTCATTTATA CAGAACAGAT CAAAGGAACG GTGATAGATG TGGATCTAGA
GAGATACATG
2581 GTCACCCTGT CTGTGAAGAT CCCTATTCTT TCTGAAGTCC CAGGTGTGCT
CATACACAAG
2641 GCATCATCTA TTTCTTACAA CATAGACGGG GAGGAATGGT ATGTGACTGT
CCCCAGCCAT
2701 ATACTCAGTC GTGCTTCTTT CTTAGGGGGT GCAGACATAA CCGATTGTGT
TGAGTCCAGA
2761 TTGACCTATA TATGCCCCAG GGATCCCGCA CAACTGATAC CTGACAGCCA
GCAAAAGTGT
2821 ATCCTGGGGG ACACAACAAG GTGTCCTGTC ACAAAAGTTG TGGACAGCCT
TATCCCCAAG
2881 TTTGCTTTTG TGAATGGGGG CGTTGTTGCT AACTGCATAG CATCCACATG
TACCTGCGGG
2941 ACAGGCCGAA GACCAATCAG TCAGGATCGC TCTAAAGGTG TAGTATTCCT
AACCCATGAC
3001 AACTGTGGTC TTATAGGTGT CAATGGGGTA GAATTGTATG CTAACCGGAG
AGGGCACGAT
3061 GCCACTTGGG GGGTCCAGAA CTTGACAGTC GGTCCTGCAA TTGCTATCAG
ACCCGTTGAT
3121 ATTTCTCTCA ACCTTGCTGA TGCTACGAAT TTCTTGCAAG ACTCTAAGGC
TGAGCTTGAG
3181 AAAGCACGGA AAATCCTCTC GGAGGTAGGT AGATGGTACA ACTCAAGAGA
GACTGTGATT
3241 ACGATCATAG TAGTTATGGT CGTAATATTG GTGGTCATTA TAGTGATCAT
CATCGTGCTT
3301 TATAGACTCA GAAGGTGAAA TCACTAGTGA ATTCACTCCT CAGGTGCAGG
CTGCCTATCA
3361 GAAGGTGGTG GCTGGTGTGG CCAATGCCCT GGCTCACAAA TACCACTGAG
ATCTTTTTCC
3421 CTCTGCCAAA AATTATGGGG ACATCATGAA GCCCCTTGAG CATCTGACTT
CTGGCTAATA
3481 AAGGAAATTT ATTTTCATTG CAATAGTGTG TTGGAATTTT TTGTGTCTCT
CACTCGGAAG
3541 GACATATGGG AGGGCAAATC ATTTAAAACA TCAGAATGAG TATTTGGTTT
AGAGTTTGGC

3601 AACATATGCC CATATGCTGG CTGCCATGAA CAAAGGTTGG CTATAAAGAG
GTCATCAGTA
3661 TATGAAACAG CCCCCTGCTG TCCATTCCTT ATTCCATAGA AAAGCCTTGA
CTTGAGGTTA
3721 GATTTTTTTT ATATTTTGTT TTGTGTTATT TTTTTCTTTA ACATCCCTAA
AATTTTCCTT
3781 ACATGTTTTA CTAGCCAGAT TTTTCCTCCT CTCCTGACTA CTCCCAGTCA
TAGCTGTCCC
3841 TCTTCTCTTA TGGAGATCCC TCGACCTGCA GCCCAAGCTT GGCGTAATCA
TGGTCATAGC
3901 TGTTTCCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA CAACATACGA
GCCGGAAGCA
3961 TAAAGTGTAA AGCCTGGGGT GCCTAATGAG TGAGCTAACT CACATTAATT
GCGTTGCGCT
4021 CACTGCCCGC TTTCCAGTCG GGAAACCTGT CGTGCCAGCG GATCCGCATC
TCAATTAGTC
4081 AGCAACCATA GTCCCGCCCC TAACTCCGCC CATCCCGCCC CTAACTCCGC
CCAGTTCCGC
4141 CCATTCTCCG CCCCATGGCT GACTAATTTT TTTTATTTAT GCAGAGGCCG
AGGCCGCCTC
4201 GGCCTCTGAG CTATTCCAGA AGTAGTGAGG AGGCTTTTTT GGAGGCCTAG
GCTTTTGCAA
4261 AAAGCTAACT TGTTTATTGC AGCTTATAAT GGTTACAAAT AAAGCAATAG
CATCACAAAT
4321 TTCACAAATA AAGCATTTTT TTCACTGCAT TCTAGTTGTG GTTTGTCCAA
ACTCATCAAT
4381 GTATCTTATC ATGTCTGTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG
GTCGTTCGGC
4441 TGCGGCGAGC GGTATCAGCT CACTCAAAGG CGGTAATACG GTTATCCACA
GAATCAGGGG
4501 ATAACGCAGG AAAGAACATG TGAGCAAAAG GCCAGCAAAA GGCCAGGAAC
CGTAAAAAGG
4561 CCGCGTTGCT GGCGTTTTTC CATAGGCTCC GCCCCCCTGA CGAGCATCAC
AAAAATCGAC
4621 GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG
TTTCCCCCTG
4681 GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA CCCTGCCGCT TACCGGATAC
CTGTCCGCCT
4741 TTCTCCCTTC GGGAAGCGTG GCGCTTTCTC ATAGCTCACG CTGTAGGTAT
CTCAGTTCGG
4801 TGTAGGTCGT TCGCTCCAAG CTGGGCTGTG TGCACGAACC CCCCGTTCAG
CCCGACCGCT
4861 GCGCCTTATC CGGTAACTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC
TTATCGCCAC
4921 TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT
GCTACAGAGT
4981 TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGAAC AGTATTTGGT
ATCTGCGCTC
5041 TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC TTGATCCGGC
AAACAAACCA
5101 CCGCTGGTAG CGGTGGTTTT TTTGTTTGCA AGCAGCAGAT TACGCGCAGA
AAAAAAGGAT
5161 CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC
GAAAACTCAC
5221 GTTAAGGGAT TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC
CTTTTAAATT
5281 AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA AACTTGGTCT
GACAGTTAGA

```
      5341 AAAACTCATC GAGCATCAAA TGAAACTGCA ATTTATTCAT ATCAGGATTA
TCAATACCAT
      5401 ATTTTTGAAA AAGCCGTTTC TGTAATGAAG GAGAAAACTC ACCGAGGCAG
TTCCATAGGA
      5461 TGGCAAGATC CTGGTATCGG TCTGCGATTC CGACTCGTCC AACATCAATA
CAACCTATTA
      5521 ATTTCCCCTC GTCAAAAATA AGGTTATCAA GTGAGAAATC ACCATGAGTG
ACGACTGAAT
      5581 CCGGTGAGAA TGGCAACAGC TTATGCATTT CTTTCCAGAC TTGTTCAACA
GGCCAGCCAT
      5641 TACGCTCGTC ATCAAAATCA CTCGCATCAA CCAAACCGTT ATTCATTCGT
GATTGCGCCT
      5701 GAGCGAGACG AAATACGCGA TCGCTGTTAA AAGGACAATT ACAAACAGGA
ATCGAATGCA
      5761 ACCGGCGCAG GAACACTGCC AGCGCATCAA CAATATTTTC ACCTGAATCA
GGATATTCTT
      5821 CTAATACCTG GAATGCTGTT TTTCCGGGGA TCGCAGTGGT GAGTAACCAT
GCATCATCAG
      5881 GAGTACGGAT AAAATGCTTG ATGGTCGGAA GAGGCATAAA TTCCGTCAGC
CAGTTTAGTC
      5941 TGACCATCTC ATCTGTAACA TCATTGGCAA CGCTACCTTT GCCATGTTTC
AGAAACAACT
      6001 CTGGCGCATC GGGCTTCCCA TACAATCGAT AGATTGTCGC ACCTGATTGC
CCGACATTAT
      6061 CGCGAGCCCA TTTATACCCA TATAAATCAG CATCCATGTT GGAATTTAAT
CGCGGCCTAG
      6121 AGCAAGACGT TTCCCGTTGA ATATGGCTCA TAACACCCCT TGTATTACTG
TTTATGTAAG
      6181 CAGACAGTTT TATTGTTCAT GATGATATAT TTTTATCTTG TGCAATGTAA
CATCAGAGAT
      6241 TTTGAGACAC AACAATTGGT CGAC
```

SEQ ID NO: 5

```
         1 ATTGATTATT GACTAGTTAT TAATAGTAAT CAATTACGGG GTCATTAGTT
CATAGCCCAT
        61 ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA
CCGCCCAACG
       121 ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA
ATAGGGACTT
       181 TCCATTGACG TCAATGGGTG GAGTATTTAC GGTAAACTGC CCACTTGGCA
GTACATCAAG
       241 TGTATCATAT GCCAAGTACG CCCCCTATTG ACGTCAATGA CGGTAAATGG
CCCGCCTGGC
       301 ATTATGCCCA GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC
TACGTATTAG
       361 TCATCGCTAT TACCATGGTC GAGGTGAGCC CCACGTTCTG CTTCACTCTC
CCCATCTCCC
       421 CCCCCTCCCC ACCCCCAATT TTGTATTTAT TTATTTTTTA ATTATTTTGT
GCAGCGATGG
       481 GGGCGGGGGG GGGGGGGGGG CGCGCGCCAG GCGGGGCGGG GCGGGGCGAG
GGGCGGGGCG
       541 GGGCGAGGCG GAAAGGTGCG GCGGCAGCCA ATCAAAGCGG CGCGCTCCGA
AAGTTTCCTT
       601 TTATGGCGAG GCGGCGGCGG CGGCGGCCCT ATAAAAAGCG AAGCGCGCGG
CGGGCGGGAG
```

661 TCGCTGCGCG CTGCCTTCGC CCCGTGCCCC GCTCCGCCGC CGCCTCGCGC
CGCCCGCCCC
721 GGCTCTGACT GACCGCGTTA CTCCCACAGG TGAGCGGGCG GGACGGCCCT
TCTCCTCCGG
781 GCTGTAATTA GCGCTTGGTT TAATGACGGC TTGTTTCTTT TCTGTGGCTG
CGTGAAAGCC
841 TTGAGGGGCT CCGGGAGGGC CCTTTGTGCG GGGGGAGCGG CTCGGGGGGT
GCGTGCGTGT
901 GTGTGTGCGT GGGGAGCGCC GCGTGCGGCT CCGCGCTGCC CGGCGGCTGT
GAGCGCTGCG
961 GGCGCGGCGC GGGGCTTTGT GCGCTCCGCA GTGTGCGCGA GGGGAGCGCG
GCCGGGGGCG
1021 GTGCCCCGCG GTGCGGGGGG GGCTGCGAGG GGAACAAAGG CTGCGTGCGG
GGTGTGTGCG
1081 TGGGGGGGTG AGCAGGGGGT GTGGGCGCGT CGGTCGGGCT GCAACCCCCC
CTGCACCCCC
1141 CTCCCCGAGT TGCTGAGCAC GGCCCGGCTT CGGGTGCGGG GCTCCGTACG
GGGCGTGGCG
1201 CGGGGCTCGC CGTGCCGGGC GGGGGGTGGC GGCAGGTGGG GGTGCCGGGC
GGGGCGGGGC
1261 CGCCTCGGGC CGGGGAGGGC TCGGGGGAGG GGCGCGGCGG CCCCCGGAGC
GCCGGCGGCT
1321 GTCGAGGCGC GGCGAGCCGC AGCCATTGCC TTTTATGGTA ATCGTGCGAG
AGGGCGCAGG
1381 GACTTCCTTT GTCCCAAATC TGTGCGGAGC CGAAATCTGG GAGGCGCCGC
CGCACCCCCT
1441 CTAGCGGGCG CGGGGCGAAG CGGTGCGGCG CCGGCAGGAA GGAAATGGGC
GGGGAGGGCC
1501 TTCGTGCGTC GCCGCGCCGC CGTCCCCTTC TCCCTCTCCA GCCTCGGGGC
TGTCCGCGGG
1561 GGGACGGGGC AGGGCGGGGT TCGGCTTCTG GCGTGTGACC GGCGGCTCTA
GAGCCTCTGC
1621 TAACCATGTT CATGCCTTCT TCTTTTTCCT ACAGCTCCTG GGCAACGTGC
TGGTTATTGT
1681 GCTGTCTCAT CATTTTGGCA AAGAATTCCT CGAGCATGTG GTCTGAGTTA
AAAATCAGGA
1741 GCAACGACGG AGGTGAAGGA CCAGAGGACG CCAACGACCC CCGGGGAAAG
GGGGTGCAAC
1801 ACATCCATAT CCAGCCATCT CTACCTGTTT ATGGACAGAG GGTTAGGGAT
GGTGATAGGG
1861 GCAAACGTGA CTCGTACTGG TCTACTTCTC CTAGTGGTAG CACCACAAAA
CCAGCATCAG
1921 GTTGGGAGAG GTCAAGTAAA GCCGACACAT GGTTGCTGAT TCTCTCATTC
ACCCAGTGGG
1981 CTTTGTCAAT TGCCACAGTG ATCATCTGTA TCATAATTTC TGCTAGACAA
GGGTATAGTA
2041 TGAAAGAGTA CTCAATGACT GTAGAGGCAT TGAACATGAG CAGCAGGGAG
GTGAAAGAGT
2101 CACTTACCAG TCTAATAAGG CAAGAGGTTA TAGCAAGGGC TGTCAACATT
CAGAGCTCTG
2161 TGCAAACCGG AATCCCAGTC TTGTTGAACA AAAACAGCAG GGATGTCATC
CAGATGATTG
2221 ATAAGTCGTG CAGCAGACAA GAGCTCACTC AGCACTGTGA GAGTACGATC
GCAGTCCACC
2281 ATGCCGATGG AATTGCCCCA CTTGAGCCAC ATAGTTTCTG GAGATGCCCT
GTCGGAGAAC
2341 CGTATCTTAG CTCAGATCCT GAAATCTCAT TGCTGCCTGG TCCGAGCTTG
TTATCTGGTT

2401 CTACAACGAT CTCTGGATGT GTTAGGCTCC CTTCACTCTC AATTGGCGAG
GCAATCTATG
2461 CCTATTCATC AAATCTCATT ACACAAGGTT GTGCTGACAT AGGGAAATCA
TATCAGGTCC
2521 TGCAGCTAGG GTACATATCA CTCAATTCAG ATATGTTCCC TGATCTTAAC
CCCGTAGTGT
2581 CCCACACTTA TGACATCAAC GACAATCGGA AATCATGCTC TGTGGTGGCA
ACCGGGACTA
2641 GGGGTTATCA GCTTTGCTCC ATGCCGACTG TAGACGAAAG AACCGACTAC
TCTAGTGATG
2701 GTATTGAGGA TCTGGTCCTT GATGTCCTGG ATCTCAAAGG GAGAACTAAG
TCTCACCGGT
2761 ATCGCAACAG CGAGGTAGAT CTTGATCACC CGTTCTCTGC ACTATACCCC
AGTGTAGGCA
2821 ACGGCATTGC AACAGAAGGC TCATTGATAT TTCTTGGGTA TGGTGGACTA
ACCACCCCTC
2881 TGCAGGGTGA TACAAAATGT AGGACCCAAG GATGCCAACA GGTGTCGCAA
GACACATGCA
2941 ATGAGGCTCT GAAAATTACA TGGCTAGGAG GGAAACAGGT GGTCAGCGTG
ATCATCCAGG
3001 TCAATGACTA TCTCTCAGAG AGGCCAAAGA TAAGAGTCAC AACCATTCCA
ATCACTCAAA
3061 ACTATCTCGG GGCGGAAGGT AGATTATTAA AATTGGGTGA TCGGGTGTAC
ATCTATACAA
3121 GATCATCAGG CTGGCACTCT CAACTGCAGA TAGGAGTACT TGATGTCAGC
CACCCTTTGA
3181 CTATCAACTG GACACCTCAT GAAGCCTTGT CTAGACCAGG AAATAAAGAG
TGCAATTGGT
3241 ACAATAAGTG TCCGAAGGAA TGCATATCAG GCGTATACAC TGATGCTTAT
CCATTGTCCC
3301 CTGATGCAGC TAACGTCGCT ACCGTCACGC TATATGCCAA TACATCGCGT
GTCAACCCAA
3361 CAATCATGTA TTCTAACACT ACTAACATTA TAAATATGTT AAGGATAAAG
GATGTTCAAT
3421 TAGAGGCTGC ATATACCACG ACATCGTGTA TCACGCATTT TGGTAAAGGC
TACTGCTTTC
3481 ACATCATCGA GATCAATCAG AAGAGCCTGA ATACCTTACA GCCGATGCTC
TTTAAGACTA
3541 GCATCCCTAA ATTATGCAAG GCCGAGTCTT AAGCGGCCGC GCATGCGAAT
TCACTCCTCA
3601 GGTGCAGGCT GCCTATCAGA AGGTGGTGGC TGGTGTGGCC AATGCCCTGG
CTCACAAATA
3661 CCACTGAGAT CTTTTTCCCT CTGCCAAAAA TTATGGGGAC ATCATGAAGC
CCCTTGAGCA
3721 TCTGACTTCT GGCTAATAAA GGAAATTTAT TTTCATTGCA ATAGTGTGTT
GGAATTTTTT
3781 GTGTCTCTCA CTCGGAAGGA CATATGGGAG GGCAAATCAT TTAAAACATC
AGAATGAGTA
3841 TTTGGTTTAG AGTTTGGCAA CATATGCCCA TATGCTGGCT GCCATGAACA
AAGGTTGGCT
3901 ATAAAGAGGT CATCAGTATA TGAAACAGCC CCCTGCTGTC TATTCCTTAT
TCCATAGAAA
3961 AGCCTTGACT TGAGGTTAGA TTTTTTTTAT ATTTTGTTTT GTGTTATTTT
TTTCTTTAAC
4021 ATCCCTAAAA TTTTCCTTAC ATGTTTTACT AGCCAGATTT TTCCTCCTCT
CCTGACTACT
4081 CCCAGTCATA GCTGTCCCTC TTCTCTTATG GAGATCCCTC GACCTGCAGC
CCAAGCTTGG

4141 CGTAATCATG GTCATAGCTG TTTCCTGTGT GAAATTGTTA TCCGCTCACA
ATTCCACACA
4201 ACATACGAGC CGGAAGCATA AAGTGTAAAG CCTGGGGTGC CTAATGAGTG
AGCTAACTCA
4261 CATTAATTGC GTTGCGCTCA CTGCCCGCTT TCCAGTCGGG AAACCTGTCG
TGCCAGCGGA
4321 TCCGCATCTC AATTAGTCAG CAACCATAGT CCCGCCCCTA ACTCCGCCCA
TCCCGCCCCT
4381 AACTCCGCCC AGTTCCGCCC ATTCTCCGCC CCATGGCTGA CTAATTTTTT
TTATTTATGC
4441 AGAGGCCGAG GCCGCCTCGG CCTCTGAGCT ATTCCAGAAG TAGTGAGGAG
GCTTTTTTGG
4501 AGGCCTAGGC TTTTGCAAAA AGCTAACTTG TTTATTGCAG CTTATAATGG
TTACAAATAA
4561 AGCAATAGCA TCACAAATTT CACAAATAAA GCATTTTTTT CACTGCATTC
TAGTTGTGGT
4621 TTGTCCAAAC TCATCAATGT ATCTTATCAT GTCTGTCCGC TTCCTCGCTC
ACTGACTCGC
4681 TGCGCTCGGT CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG
GTAATACGGT
4741 TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC
CAGCAAAAGG
4801 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
CCCCCTGACG
4861 AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA
CTATAAAGAT
4921 ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC
CTGCCGCTTA
4981 CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAT
AGCTCACGCT
5041 GTAGGTATCT CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG
CACGAACCCC
5101 CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC
AACCCGGTAA
5161 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA
GCGAGGTATG
5221 TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT
AGAAGAACAG
5281 TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT
GGTAGCTCTT
5341 GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG
CAGCAGATTA
5401 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
TCTGACGCTC
5461 AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA
AGGATCTTCA
5521 CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA
TATGAGTAAA
5581 CTTGGTCTGA CAGTTAGAAA AACTCATCGA GCATCAAATG AAACTGCAAT
TTATTCATAT
5641 CAGGATTATC AATACCATAT TTTTGAAAAA GCCGTTTCTG TAATGAAGGA
GAAAACTCAC
5701 CGAGGCAGTT CCATAGGATG GCAAGATCCT GGTATCGGTC TGCGATTCCG
ACTCGTCCAA
5761 CATCAATACA ACCTATTAAT TTCCCCTCGT CAAAAATAAG GTTATCAAGT
GAGAAATCAC
5821 CATGAGTGAC GACTGAATCC GGTGAGAATG GCAACAGCTT ATGCATTTCT
TTCCAGACTT

5881 GTTCAACAGG CCAGCCATTA CGCTCGTCAT CAAAATCACT CGCATCAACC
AAACCGTTAT
5941 TCATTCGTGA TTGCGCCTGA GCGAGACGAA ATACGCGATC GCTGTTAAAA
GGACAATTAC
6001 AAACAGGAAT CGAATGCAAC CGGCGCAGGA ACACTGCCAG CGCATCAACA
ATATTTTCAC
6061 CTGAATCAGG ATATTCTTCT AATACCTGGA ATGCTGTTTT TCCGGGGATC
GCAGTGGTGA
6121 GTAACCATGC ATCATCAGGA GTACGGATAA AATGCTTGAT GGTCGGAAGA
GGCATAAATT
6181 CCGTCAGCCA GTTTAGTCTG ACCATCTCAT CTGTAACATC ATTGGCAACG
CTACCTTTGC
6241 CATGTTTCAG AAACAACTCT GGCGCATCGG GCTTCCCATA CAATCGATAG
ATTGTCGCAC
6301 CTGATTGCCC GACATTATCG CGAGCCCATT TATACCCATA TAAATCAGCA
TCCATGTTGG
6361 AATTTAATCG CGGCCTAGAG CAAGACGTTT CCCGTTGAAT ATGGCTCATA
ACACCCCTTG
6421 TATTACTGTT TATGTAAGCA GACAGTTTTA TTGTTCATGA TGATATATTT
TTATCTTGTG
6481 CAATGTAACA TCAGAGATTT TGAGACACAA CAATTGGTCG AC

SEQ ID NO: 6

1 AGATCTGTTA CATAACTTAT GGTAAATGGC CTGCCTGGCT GACTGCCCAA
TGACCCCTGC
61 CCAATGATGT CAATAATGAT GTATGTTCCC ATGTAATGCC AATAGGGACT
TTCCATTGAT
121 GTCAATGGGT GGAGTATTTA TGGTAACTGC CCACTTGGCA GTACATCAAG
TGTATCATAT
181 GCCAAGTATG CCCCCTATTG ATGTCAATGA TGGTAAATGG CCTGCCTGGC
ATTATGCCCA
241 GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC TATGTATTAG
TCATTGCTAT
301 TACCATGGGA ATTCACTAGT GGAGAAGAGC ATGCTTGAGG GCTGAGTGCC
CCTCAGTGGG
361 CAGAGAGCAC ATGGCCCACA GTCCCTGAGA AGTTGGGGGG AGGGGTGGGC
AATTGAACTG
421 GTGCCTAGAG AAGGTGGGGC TTGGGTAAAC TGGGAAAGTG ATGTGGTGTA
CTGGCTCCAC
481 CTTTTTCCCC AGGGTGGGGG AGAACCATAT ATAAGTGCAG TAGTCTCTGT
GAACATTCAA
541 GCTTCTGCCT TCTCCCTCCT GTGAGTTTGC TAGC

SEQ ID NO: 7

1 GCTAGCCACC ATGCAGAGAA GCCCTCTGGA GAAGGCCTCT GTGGTGAGCA
AGCTGTTCTT
61 CAGCTGGACC AGGCCCATCC TGAGGAAGGG CTACAGGCAG AGACTGGAGC
TGTCTGACAT
121 CTACCAGATC CCCTCTGTGG ACTCTGCTGA CAACCTGTCT GAGAAGCTGG
AGAGGGAGTG
181 GGATAGAGAG CTGGCCAGCA AGAAGAACCC CAAGCTGATC AATGCCCTGA
GGAGATGCTT
241 CTTCTGGAGA TTCATGTTCT ATGGCATCTT CCTGTACCTG GGGGAAGTGA
CCAAGGCTGT

301 GCAGCCTCTG CTGCTGGGCA GAATCATTGC CAGCTATGAC CCTGACAACA
AGGAGGAGAG
361 GAGCATTGCC ATCTACCTGG GCATTGGCCT GTGCCTGCTG TTCATTGTGA
GGACCCTGCT
421 GCTGCACCCT GCCATCTTTG GCCTGCACCA CATTGGCATG CAGATGAGGA
TTGCCATGTT
481 CAGCCTGATC TACAAGAAAA CCCTGAAGCT GTCCAGCAGA GTGCTGGACA
AGATCAGCAT
541 TGGCCAGCTG GTGAGCCTGC TGAGCAACAA CCTGAACAAG TTTGATGAGG
GCCTGGCCCT
601 GGCCCACTTT GTGTGGATTG CCCCTCTGCA GGTGGCCCTG CTGATGGGCC
TGATTTGGGA
661 GCTGCTGCAG GCCTCTGCCT TTTGTGGCCT GGGCTTCCTG ATTGTGCTGG
CCCTGTTTCA
721 GGCTGGCCTG GGCAGGATGA TGATGAAGTA CAGGGACCAG AGGGCAGGCA
AGATCAGTGA
781 GAGGCTGGTG ATCACCTCTG AGATGATTGA GAACATCCAG TCTGTGAAGG
CCTACTGTTG
841 GGAGGAAGCT ATGGAGAAGA TGATTGAAAA CCTGAGGCAG ACAGAGCTGA
AGCTGACCAG
901 GAAGGCTGCC TATGTGAGAT ACTTCAACAG CTCTGCCTTC TTCTTCTCTG
GCTTCTTTGT
961 GGTGTTCCTG TCTGTGCTGC CCTATGCCCT GATCAAGGGG ATCATCCTGA
GAAAGATTTT
1021 CACCACCATC AGCTTCTGCA TTGTGCTGAG GATGGCTGTG ACCAGACAGT
TCCCCTGGGC
1081 TGTGCAGACC TGGTATGACA GCCTGGGGGC CATCAACAAG ATCCAGGACT
TCCTGCAGAA
1141 GCAGGAGTAC AAGACCCTGG AGTACAACCT GACCACCACA GAAGTGGTGA
TGGAGAATGT
1201 GACAGCCTTC TGGGAGGAGG GCTTTGGGGA GCTGTTTGAG AAGGCCAAGC
AGAACAACAA
1261 CAACAGAAAG ACCAGCAATG GGGATGACTC CCTGTTCTTC TCCAACTTCT
CCCTGCTGGG
1321 CACACCTGTG CTGAAGGACA TCAACTTCAA GATTGAGAGG GGGCAGCTGC
TGGCTGTGGC
1381 TGGATCTACA GGGGCTGGCA AGACCAGCCT GCTGATGATG ATCATGGGGG
AGCTGGAGCC
1441 TTCTGAGGGC AAGATCAAGC ACTCTGGCAG GATCAGCTTT TGCAGCCAGT
TCAGCTGGAT
1501 CATGCCTGGC ACCATCAAGG AGAACATCAT CTTTGGAGTG AGCTATGATG
AGTACAGATA
1561 CAGGAGTGTG ATCAAGGCCT GCCAGCTGGA GGAGGACATC AGCAAGTTTG
CTGAGAAGGA
1621 CAACATTGTG CTGGGGGAGG GAGGCATTAC ACTGTCTGGG GGCCAGAGAG
CCAGAATCAG
1681 CCTGGCCAGG GCTGTGTACA AGGATGCTGA CCTGTACCTG CTGGACTCCC
CCTTTGGCTA
1741 CCTGGATGTG CTGACAGAGA AGGAGATTTT TGAGAGCTGT GTGTGCAAGC
TGATGGCCAA
1801 CAAGACCAGA ATCCTGGTGA CCAGCAAGAT GGAGCACCTG AAGAAGGCTG
ACAAGATCCT
1861 GATCCTGCAT GAGGGCAGCA GCTACTTCTA TGGGACCTTC TCTGAGCTGC
AGAACCTGCA
1921 GCCTGACTTC AGCTCTAAGC TGATGGGCTG TGACAGCTTT GACCAGTTCT
CTGCTGAGAG
1981 GAGGAACAGC ATCCTGACAG AGACCCTGCA CAGATTCAGC CTGGAGGGAG
ATGCCCCTGT

2041 GAGCTGGACA GAGACCAAGA AGCAGAGCTT CAAGCAGACA GGGGAGTTTG
GGGAGAAGAG
2101 GAAGAACTCC ATCCTGAACC CCATCAACAG CATCAGGAAG TTCAGCATTG
TGCAGAAAAC
2161 CCCCCTGCAG ATGAATGGCA TTGAGGAAGA TTCTGATGAG CCCCTGGAGA
GGAGACTGAG
2221 CCTGGTGCCT GATTCTGAGC AGGGAGAGGC CATCCTGCCT AGGATCTCTG
TGATCAGCAC
2281 AGGCCCTACA CTGCAGGCCA GAAGGAGGCA GTCTGTGCTG AACCTGATGA
CCCACTCTGT
2341 GAACCAGGGC CAGAACATCC ACAGGAAAAC CACAGCCTCC ACCAGGAAAG
TGAGCCTGGC
2401 CCCTCAGGCC AATCTGACAG AGCTGGACAT CTACAGCAGG AGGCTGTCTC
AGGAGACAGG
2461 CCTGGAGATT TCTGAGGAGA TCAATGAGGA GGACCTGAAA GAGTGCTTCT
TTGATGACAT
2521 GGAGAGCATC CCTGCTGTGA CCACCTGGAA CACCTACCTG AGATACATCA
CAGTGCACAA
2581 GAGCCTGATC TTTGTGCTGA TCTGGTGCCT GGTGATCTTC CTGGCTGAAG
TGGCTGCCTC
2641 TCTGGTGGTG CTGTGGCTGC TGGGAAACAC CCCACTGCAG GACAAGGGCA
ACAGCACCCA
2701 CAGCAGGAAC AACAGCTATG CTGTGATCAT CACCTCCACC TCCAGCTACT
ATGTGTTCTA
2761 CATCTATGTG GGAGTGGCTG ATACCCTGCT GGCTATGGGC TTCTTTAGAG
GCCTGCCCCT
2821 GGTGCACACA CTGATCACAG TGAGCAAGAT CCTCCACCAC AAGATGCTGC
ACTCTGTGCT
2881 GCAGGCTCCT ATGAGCACCC TGAATACCCT GAAGGCTGGG GGCATCCTGA
ACAGATTCTC
2941 CAAGGATATT GCCATCCTGG ATGACCTGCT GCCTCTCACC ATCTTTGACT
TCATCCAGCT
3001 GCTGCTGATT GTGATTGGGG CCATTGCTGT GGTGGCAGTG CTGCAGCCCT
ACATCTTTGT
3061 GGCCACAGTG CCTGTGATTG TGGCCTTCAT CATGCTGAGG GCCTACTTTC
TGCAGACCTC
3121 CCAGCAGCTG AAGCAGCTGG AGTCTGAGGG CAGAAGCCCC ATCTTCACCC
ACCTGGTGAC
3181 AAGCCTGAAG GGCCTGTGGA CCCTGAGAGC CTTTGGCAGG CAGCCCTACT
TTGAGACCCT
3241 GTTCCACAAG GCCCTGAACC TGCACACAGC CAACTGGTTC CTCTACCTGT
CCACCCTGAG
3301 ATGGTTCCAG ATGAGAATTG AGATGATCTT TGTCATCTTC TTCATTGCTG
TGACCTTCAT
3361 CAGCATTCTG ACCACAGGAG AGGGAGAGGG CAGAGTGGGC ATTATCCTGA
CCCTGGCCAT
3421 GAACATCATG AGCACACTGC AGTGGGCAGT GAACAGCAGC ATTGATGTGG
ACAGCCTGAT
3481 GAGGAGTGTG AGCAGAGTGT TCAAGTTCAT TGATATGCCC ACAGAGGGCA
AGCCTACCAA
3541 GAGCACCAAG CCCTACAAGA ATGGCCAGCT GAGCAAAGTG ATGATCATTG
AGAACAGCCA
3601 TGTGAAGAAG GATGATATCT GGCCCAGTGG AGGCCAGATG ACAGTGAAGG
ACCTGACAGC
3661 CAAGTACACA GAGGGGGGCA ATGCTATCCT GGAGAACATC TCCTTCAGCA
TCTCCCCTGG
3721 CCAGAGAGTG GGACTGCTGG GAAGAACAGG CTCTGGCAAG TCTACCCTGC
TGTCTGCCTT

```
     3781 CCTGAGGCTG CTGAACACAG AGGGAGAGAT CCAGATTGAT GGAGTGTCCT
GGGACAGCAT
     3841 CACACTGCAG CAGTGGAGGA AGGCCTTTGG TGTGATCCCC CAGAAAGTGT
TCATCTTCAG
     3901 TGGCACCTTC AGGAAGAACC TGGACCCCTA TGAGCAGTGG TCTGACCAGG
AGATTTGGAA
     3961 AGTGGCTGAT GAAGTGGGCC TGAGAAGTGT GATTGAGCAG TTCCCTGGCA
AGCTGGACTT
     4021 TGTCCTGGTG GATGGGGGCT GTGTGCTGAG CCATGGCCAC AAGCAGCTGA
TGTGCCTGGC
     4081 CAGATCAGTG CTGAGCAAGG CCAAGATCCT GCTGCTGGAT GAGCCTTCTG
CCCACCTGGA
     4141 TCCTGTGACC TACCAGATCA TCAGGAGGAC CCTCAAGCAG GCCTTTGCTG
ACTGCACAGT
     4201 CATCCTGTGT GAGCACAGGA TTGAGGCCAT GCTGGAGTGC CAGCAGTTCC
TGGTGATTGA
     4261 GGAGAACAAA GTGAGGCAGT ATGACAGCAT CCAGAAGCTG CTGAATGAGA
GGAGCCTGTT
     4321 CAGGCAGGCC ATCAGCCCCT CTGATAGAGT GAAGCTGTTC CCCCACAGGA
ACAGCTCCAA
     4381 GTGCAAGAGC AAGCCCCAGA TTGCTGCCCT GAAGGAGGAG ACAGAGGAGG
AAGTGCAGGA
     4441 CACCAGGCTG TGAGGGCCC
```

SEQ ID NO: 8

```
        1 GGGCCCAATC AACCTCTGGA TTACAAAATT TGTGAAAGAT TGACTGGTAT
TCTTAACTAT
       61 GTTGCTCCTT TTACGCTATG TGGATACGCT GCTTTAATGC CTTTGTATCA
TGCTATTGCT
      121 TCCCGTATGG CTTTCATTTT CTCCTCCTTG TATAAATCCT GGTTGCTGTC
TCTTTATGAG
      181 GAGTTGTGGC CCGTTGTCAG GCAACGTGGC GTGGTGTGCA CTGTGTTTGC
TGACGCAACC
      241 CCCACTGGTT GGGGCATTGC CACCACCTGT CAGCTCCTTT CCGGGACTTT
CGCTTTCCCC
      301 CTCCCTATTG CCACGGCGGA ACTCATCGCC GCCTGCCTTG CCCGCTGCTG
GACAGGGGCT
      361 CGGCTGTTGG GCACTGACAA TTCCGTGGTG TTGTCGGGGA AATCATCGTC
CTTTCCTTGG
      421 CTGCTCGCCT GTGTTGCCAC CTGGATTCTG CGCGGGACGT CCTTCTGCTA
CGTCCCTTCG
      481 GCCCTCAATC CAGCGGACCT TCCTTCCCGC GGCCTGCTGC CGGCTCTGCG
GCCTCTTCCG
      541 CGTCTTCGCC TTCGCCCTCA GACGAGTCGG ATCTCCCTTT GGGCCGCCTC
CCCGCAAGCT
```

SEQ ID NO: 9

```
        1  GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT
       61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
      121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
```

181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
TGACGTATGT
301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
361 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT
421 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAC
GGGACTTTCC
481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA
541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT
601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA
661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGCGTGTACG
GTGGGAGGTC
721 TATATAAGCA GAGCTCGCTG GCTTGTAACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC
781 TGGGTGTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA
841 GAAAGCTAAT AAACTTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCCT
CATTGACGCC
901 TCACTCTCTT GAACGGGAAT CTTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA
961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA
1021 GGCGTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA
1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC
1141 CGTAACTACT CTTGGGCAAG TAGGGCAGGC GGTGGGTACG CAATGGGGGC
GGCTACCTCA
1201 GCACTAAATA GGAGACAATT AGACCAATTT GAGAAAATAC GACTTCGCCC
GAACGGAAAG
1261 AAAAAGTACC AAATTAAACA TTTAATATGG GCAGGCAAGG AGATGGAGCG
CTTCGGCCTC
1321 CATGAGAGGT TGTTGGAGAC AGAGGAGGGG TGTAAAAGAA TCATAGAAGT
CCTCTACCCC
1381 CTAGAACCAA CAGGATCGGA GGGCTTAAAA AGTCTGTTCA ATCTTGTGTG
CGTGCTATAT
1441 TGCTTGCACA AGGAACAGAA AGTGAAAGAC ACAGAGGAAG CAGTAGCAAC
AGTAAGACAA
1501 CACTGCCATC TAGTGGAAAA AGAAAAAAGT GCAACAGAGA CATCTAGTGG
ACAAAAGAAA
1561 AATGACAAGG GAATAGCAGC GCCACCTGGT GGCAGTCAGA ATTTTCCAGC
GCAACAACAA
1621 GGAAATGCCT GGGTACATGT ACCCTTGTCA CCGCGCACCT TAAATGCGTG
GGTAAAAGCA
1681 GTAGAGGAGA AAAAATTTGG AGCAGAAATA GTACCCATTT TTTTGTTTCA
AGCCCTATCG
1741 AATTCCCGTT TGTGCTAGGG TTCTTAGGCT TCTTGGGGGC TGCTGGAACT
GCAATGGGAG
1801 CAGCGGCGAC AGCCCTGACG GTCCAGTCTC AGCATTTGCT TGCTGGGATA
CTGCAGCAGC
1861 AGAAGAATCT GCTGGCGGCT GTGGAGGCTC AACAGCAGAT GTTGAAGCTG
ACCATTTGGG

1921 GTGTTAAAAA CCTCAATGCC CGCGTCACAG CCCTTGAGAA GTACCTAGAG
GATCAGGCAC
1981 GACTAAACTC CTGGGGGTGC GCATGGAAAC AAGTATGTCA TACCACAGTG
GAGTGGCCCT
2041 GGACAAATCG GACTCCGGAT TGGCAAAATA TGACTTGGTT GGAGTGGGAA
AGACAAATAG
2101 CTGATTTGGA AAGCAACATT ACGAGACAAT TAGTGAAGGC TAGAGAACAA
GAGGAAAAGA
2161 ATCTAGATGC CTATCAGAAG TTAACTAGTT GGTCAGATTT CTGGTCTTGG
TTCGATTTCT
2221 CAAAATGGCT TAACATTTTA AAAATGGGAT TTTTAGTAAT AGTAGGAATA
ATAGGGTTAA
2281 GATTACTTTA CACAGTATAT GGATGTATAG TGAGGGTTAG GCAGGGATAT
GTTCCTCTAT
2341 CTCCACAGAT CCATATCCGC GGCAATTTTA AAAGAAAGGG AGGAATAGGG
GGACAGACTT
2401 CAGCAGAGAG ACTAATTAAT ATAATAACAA CACAATTAGA AATACAACAT
TTACAAACCA
2461 AAATTCAAAA AATTTTAAAT TTTAGAGCCG CGGAGATCTG TTACATAACT
TATGGTAAAT
2521 GGCCTGCCTG GCTGACTGCC CAATGACCCC TGCCCAATGA TGTCAATAAT
GATGTATGTT
2581 CCCATGTAAT GCCAATAGGG ACTTTCCATT GATGTCAATG GGTGGAGTAT
TTATGGTAAC
2641 TGCCCACTTG GCAGTACATC AAGTGTATCA TATGCCAAGT ATGCCCCCTA
TTGATGTCAA
2701 TGATGGTAAA TGGCCTGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
ACTTTCCTAC
2761 TTGGCAGTAC ATCTATGTAT TAGTCATTGC TATTACCATG GGAATTCACT
AGTGGAGAAG
2821 AGCATGCTTG AGGGCTGAGT GCCCCTCAGT GGGCAGAGAG CACATGGCCC
ACAGTCCCTG
2881 AGAAGTTGGG GGGAGGGGTG GGCAATTGAA CTGGTGCCTA GAGAAGGTGG
GGCTTGGGTA
2941 AACTGGGAAA GTGATGTGGT GTACTGGCTC CACCTTTTTC CCCAGGGTGG
GGGAGAACCA
3001 TATATAAGTG CAGTAGTCTC TGTGAACATT CAAGCTTCTG CCTTCTCCCT
CCTGTGAGTT
3061 TGCTAGCCAC CATGCCCAGC TCTGTGTCCT GGGGCATTCT GCTGCTGGCT
GGCCTGTGCT
3121 GTCTGGTGCC TGTGTCCCTG GCTGAGGACC CTCAGGGGGA TGCTGCCCAG
AAAACAGACA
3181 CCTCCCACCA TGACCAGGAC CACCCCACCT TCAACAAGAT CACCCCCAAC
CTGGCAGAGT
3241 TTGCCTTCAG CCTGTACAGA CAGCTGGCCC ACCAGAGCAA CAGCACCAAC
ATCTTTTTCA
3301 GCCCTGTGTC CATTGCCACA GCCTTTGCCA TGCTGAGCCT GGGCACCAAG
GCTGACACCC
3361 ATGATGAGAT CCTGGAAGGC CTGAACTTCA ACCTGACAGA GATCCCTGAG
GCCCAGATCC
3421 ATGAGGGCTT CCAGGAACTG CTGAGAACCC TGAACCAGCC AGACAGCCAG
CTGCAGCTGA
3481 CAACAGGCAA TGGGCTGTTC CTGTCTGAGG GCCTGAAGCT GGTGGACAAG
TTTCTGGAAG
3541 ATGTGAAGAA GCTGTACCAC TCTGAGGCCT TCACAGTGAA CTTTGGGGAC
ACAGAAGAGG
3601 CCAAGAAACA GATCAATGAC TATGTGGAAA AGGGCACCCA GGGCAAGATT
GTGGACCTTG

3661 TGAAAGAGCT GGACAGGGAC ACTGTGTTTG CCCTTGTGAA CTACATCTTC
TTCAAGGGCA
3721 AGTGGGAGAG GCCCTTTGAA GTGAAGGACA CTGAGGAAGA GGACTTCCAT
GTGGACCAAG
3781 TGACCACAGT GAAGGTGCCA ATGATGAAGA GACTGGGGAT GTTCAATATC
CAGCACTGCA
3841 AGAAACTGAG CAGCTGGGTG CTGCTGATGA AGTACCTGGG CAATGCTACA
GCCATATTCT
3901 TTCTGCCTGA TGAGGGCAAG CTGCAGCACC TGGAAAATGA GCTGACCCAT
GACATCATCA
3961 CCAAATTTCT GGAAAATGAG GACAGAAGAT CTGCCAGCCT GCATCTGCCC
AAGCTGAGCA
4021 TCACAGGCAC ATATGACCTG AAGTCTGTGC TGGGACAGCT GGGAATCACC
AAGGTGTTCA
4081 GCAATGGGGC AGACCTGAGT GGAGTGACAG AGGAAGCCCC TCTGAAGCTG
TCCAAGGCTG
4141 TGCACAAGGC AGTGCTGACC ATTGATGAGA AGGGCACAGA GGCTGCTGGG
GCCATGTTTC
4201 TGGAAGCCAT CCCCATGTCC ATCCCCCCAG AAGTGAAGTT CAACAAGCCC
TTTGTGTTCC
4261 TGATGATTGA GCAGAACACC AAGAGCCCCC TGTTCATGGG CAAGGTTGTG
AACCCCACCC
4321 AGAAATGAGG GCCCAATCAA CCTCTGGATT ACAAAATTTG TGAAAGATTG
ACTGGTATTC
4381 TTAACTATGT TGCTCCTTTT ACGCTATGTG GATACGCTGC TTTAATGCCT
TTGTATCATG
4441 CTATTGCTTC CCGTATGGCT TTCATTTTCT CCTCCTTGTA TAAATCCTGG
TTGCTGTCTC
4501 TTTATGAGGA GTTGTGGCCC GTTGTCAGGC AACGTGGCGT GGTGTGCACT
GTGTTTGCTG
4561 ACGCAACCCC CACTGGTTGG GGCATTGCCA CCACCTGTCA GCTCCTTTCC
GGGACTTTCG
4621 CTTTCCCCCT CCCTATTGCC ACGGCGGAAC TCATCGCCGC CTGCCTTGCC
CGCTGCTGGA
4681 CAGGGGCTCG GCTGTTGGGC ACTGACAATT CCGTGGTGTT GTCGGGGAAA
TCATCGTCCT
4741 TTCCTTGGCT GCTCGCCTGT GTTGCCACCT GGATTCTGCG CGGGACGTCC
TTCTGCTACG
4801 TCCCTTCGGC CCTCAATCCA GCGGACCTTC CTTCCCGCGG CCTGCTGCCG
GCTCTGCGGC
4861 CTCTTCCGCG TCTTCGCCTT CGCCCTCAGA CGAGTCGGAT CTCCCTTTGG
GCCGCCTCCC
4921 CGCAAGCTTC GCACTTTTTA AAAGAAAAGG GAGGACTGGA TGGGATTTAT
TACTCCGATA
4981 GGACGCTGGC TTGTAACTCA GTCTCTTACT AGGAGACCAG CTTGAGCCTG
GGTGTTCGCT
5041 GGTTAGCCTA ACCTGGTTGG CCACCAGGGG TAAGGACTCC TTGGCTTAGA
AAGCTAATAA
5101 ACTTGCCTGC ATTAGAGCTC TTACGCGTCC CGGGCTCGAG ATCCGCATCT
CAATTAGTCA
5161 GCAACCATAG TCCCGCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC
CAGTTCCGCC
5221 CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA
GGCCGCCTCG
5281 GCCTCTGAGC TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG
CTTTTGCAAA
5341 AAGCTAACTT GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC
ATCACAAATT

5401 TCACAAATAA AGCATTTTTT TCACTGCATT CTAGTTGTGG TTTGTCCAAA
CTCATCAATG
5461 TATCTTATCA TGTCTGTCCG CTTCCTCGCT CACTGACTCG CTGCGCTCGG
TCGTTCGGCT
5521 GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG
AATCAGGGGA
5581 TAACGCAGGA AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC
GTAAAAAGGC
5641 CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA
AAAATCGACG
5701 CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT
TTCCCCCTGG
5761 AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC
TGTCCGCCTT
5821 TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC
TCAGTTCGGT
5881 GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAGC
CCGACCGCTG
5941 CGCCTTATCC GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT
TATCGCCACT
6001 GGCAGCAGCC ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG
CTACAGAGTT
6061 CTTGAAGTGG TGGCCTAACT ACGGCTACAC TAGAAGAACA GTATTTGGTA
TCTGCGCTCT
6121 GCTGAAGCCA GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA
AACAAACCAC
6181 CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA
AAAAAGGATC
6241 TCAAGAAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG
AAAACTCACG
6301 TTAAGGGATT TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC
TTTTAAATTA
6361 AAAATGAAGT TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG
ACAGTTAGAA
6421 AAACTCATCG AGCATCAAAT GAAACTGCAA TTTATTCATA TCAGGATTAT
CAATACCATA
6481 TTTTTGAAAA AGCCGTTTCT GTAATGAAGG AGAAAACTCA CCGAGGCAGT
TCCATAGGAT
6541 GGCAAGATCC TGGTATCGGT CTGCGATTCC GACTCGTCCA ACATCAATAC
AACCTATTAA
6601 TTTCCCCTCG TCAAAAATAA GGTTATCAAG TGAGAAATCA CCATGAGTGA
CGACTGAATC
6661 CGGTGAGAAT GGCAACAGCT TATGCATTTC TTTCCAGACT TGTTCAACAG
GCCAGCCATT
6721 ACGCTCGTCA TCAAAATCAC TCGCATCAAC CAAACCGTTA TTCATTCGTG
ATTGCGCCTG
6781 AGCGAGACGA AATACGCGAT CGCTGTTAAA AGGACAATTA CAAACAGGAA
TCGAATGCAA
6841 CCGGCGCAGG AACACTGCCA GCGCATCAAC AATATTTTCA CCTGAATCAG
GATATTCTTC
6901 TAATACCTGG AATGCTGTTT TTCCGGGGAT CGCAGTGGTG AGTAACCATG
CATCATCAGG
6961 AGTACGGATA AAATGCTTGA TGGTCGGAAG AGGCATAAAT TCCGTCAGCC
AGTTTAGTCT
7021 GACCATCTCA TCTGTAACAT CATTGGCAAC GCTACCTTTG CCATGTTTCA
GAAACAACTC
7081 TGGCGCATCG GGCTTCCCAT ACAATCGATA GATTGTCGCA CCTGATTGCC
CGACATTATC

```
      7141 GCGAGCCCAT TTATACCCAT ATAAATCAGC ATCCATGTTG GAATTTAATC
GCGGCCTAGA
      7201 GCAAGACGTT TCCCGTTGAA TATGGCTCAT AACACCCCTT GTATTACTGT
TTATGTAAGC
      7261 AGACAGTTTT ATTGTTCATG ATGATATATT TTTATCTTGT GCAATGTAAC
ATCAGAGATT
      7321 TTGAGACACA ACAATTGGTC GACGGATCC
```

SEQ ID NO: 10

```
       1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT
      61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
     121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
     181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
     241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
TGACGTATGT
     301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
     361 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT
     421 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAC
GGGACTTTCC
     481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA
     541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT
     601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA
     661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGCGTGTACG
GTGGGAGGTC
     721 TATATAAGCA GAGCTCGCTG GCTTGTAACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC
     781 TGGGTGTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA
     841 GAAAGCTAAT AAACTTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCCT
CATTGACGCC
     901 TCACTCTCTT GAACGGGAAT CTTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA
     961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA
    1021 GGCGTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA
    1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC
    1141 CGTAACTACT CTTGGGCAAG TAGGGCAGGC GGTGGGTACG CAATGGGGGC
GGCTACCTCA
    1201 GCACTAAATA GGAGACAATT AGACCAATTT GAGAAAATAC GACTTCGCCC
GAACGGAAAG
    1261 AAAAAGTACC AAATTAAACA TTTAATATGG GCAGGCAAGG AGATGGAGCG
CTTCGGCCTC
    1321 CATGAGAGGT TGTTGGAGAC AGAGGAGGGG TGTAAAAGAA TCATAGAAGT
CCTCTACCCC
```

1381 CTAGAACCAA CAGGATCGGA GGGCTTAAAA AGTCTGTTCA ATCTTGTGTG
CGTGCTATAT
1441 TGCTTGCACA AGGAACAGAA AGTGAAAGAC ACAGAGGAAG CAGTAGCAAC
AGTAAGACAA
1501 CACTGCCATC TAGTGGAAAA AGAAAAAAGT GCAACAGAGA CATCTAGTGG
ACAAAAGAAA
1561 AATGACAAGG GAATAGCAGC GCCACCTGGT GGCAGTCAGA ATTTTCCAGC
GCAACAACAA
1621 GGAAATGCCT GGGTACATGT ACCCTTGTCA CCGCGCACCT TAAATGCGTG
GGTAAAAGCA
1681 GTAGAGGAGA AAAAATTTGG AGCAGAAATA GTACCCATTT TTTTGTTTCA
AGCCCTATCG
1741 AATTCCCGTT TGTGCTAGGG TTCTTAGGCT TCTTGGGGGC TGCTGGAACT
GCAATGGGAG
1801 CAGCGGCGAC AGCCCTGACG GTCCAGTCTC AGCATTTGCT TGCTGGGATA
CTGCAGCAGC
1861 AGAAGAATCT GCTGGCGGCT GTGGAGGCTC AACAGCAGAT GTTGAAGCTG
ACCATTTGGG
1921 GTGTTAAAAA CCTCAATGCC CGCGTCACAG CCCTTGAGAA GTACCTAGAG
GATCAGGCAC
1981 GACTAAACTC CTGGGGGTGC GCATGGAAAC AAGTATGTCA TACCACAGTG
GAGTGGCCCT
2041 GGACAAATCG GACTCCGGAT TGGCAAAATA TGACTTGGTT GGAGTGGGAA
AGACAAATAG
2101 CTGATTTGGA AAGCAACATT ACGAGACAAT TAGTGAAGGC TAGAGAACAA
GAGGAAAAGA
2161 ATCTAGATGC CTATCAGAAG TTAACTAGTT GGTCAGATTT CTGGTCTTGG
TTCGATTTCT
2221 CAAAATGGCT TAACATTTTA AAAATGGGAT TTTTAGTAAT AGTAGGAATA
ATAGGGTTAA
2281 GATTACTTTA CACAGTATAT GGATGTATAG TGAGGGTTAG GCAGGGATAT
GTTCCTCTAT
2341 CTCCACAGAT CCATATCCGC GGCAATTTTA AAAGAAAGGG AGGAATAGGG
GGACAGACTT
2401 CAGCAGAGAG ACTAATTAAT ATAATAACAA CACAATTAGA AATACAACAT
TTACAAACCA
2461 AAATTCAAAA AATTTTAAAT TTTAGAGCCG CGGAGATCTG TTACATAACT
TATGGTAAAT
2521 GGCCTGCCTG GCTGACTGCC CAATGACCCC TGCCCAATGA TGTCAATAAT
GATGTATGTT
2581 CCCATGTAAT GCCAATAGGG ACTTTCCATT GATGTCAATG GGTGGAGTAT
TTATGGTAAC
2641 TGCCCACTTG GCAGTACATC AAGTGTATCA TATGCCAAGT ATGCCCCCTA
TTGATGTCAA
2701 TGATGGTAAA TGGCCTGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
ACTTTCCTAC
2761 TTGGCAGTAC ATCTATGTAT TAGTCATTGC TATTACCATG GGAATTCACT
AGTGGAGAAG
2821 AGCATGCTTG AGGGCTGAGT GCCCCTCAGT GGGCAGAGAG CACATGGCCC
ACAGTCCCTG
2881 AGAAGTTGGG GGGAGGGGTG GGCAATTGAA CTGGTGCCTA GAGAAGGTGG
GGCTTGGGTA
2941 AACTGGGAAA GTGATGTGGT GTACTGGCTC CACCTTTTTC CCCAGGGTGG
GGGAGAACCA
3001 TATATAAGTG CAGTAGTCTC TGTGAACATT CAAGCTTCTG CCTTCTCCCT
CCTGTGAGTT
3061 TGCTAGCCAC CATGGGAGTG AAGGTGCTGT TTGCCCTGAT CTGCATTGCT
GTGGCTGAGG

3121 CCAAGCCCAC AGAGAACAAT GAGGACTTCA ACATTGTGGC TGTGGCCAGC
AACTTTGCCA
3181 CCACAGACCT GGATGCTGAC AGGGGCAAGC TGCCTGGCAA GAAGCTGCCC
CTGGAAGTCC
3241 TGAAAGAGAT GGAAGCCAAT GCCAGGAAGG CTGGCTGCAC AAGAGGCTGT
CTGATCTGCC
3301 TGAGCCACAT CAAGTGCACC CCCAAGATGA AGAAGTTCAT CCCTGGCAGG
TGCCACACCT
3361 ATGAAGGGGA CAAAGAGTCT GCCCAGGGGG GAATTGGAGA GGCCATTGTG
GACATCCCTG
3421 AGATCCCTGG CTTCAAGGAC CTGGAACCCA TGGAACAGTT CATTGCCCAG
GTGGACCTGT
3481 GTGTGGACTG CACTACAGGC TGTCTCAAGG GCCTGGCCAA TGTGCAGTGC
TCTGACCTGC
3541 TGAAGAAGTG GCTGCCCCAG AGATGTGCCA CCTTTGCCAG CAAGATCCAG
GGCCAGGTGG
3601 ACAAGATCAA GGGAGCTGGG GGAGATTGAT GAGGGCCCAA TCAACCTCTG
GATTACAAAA
3661 TTTGTGAAAG ATTGACTGGT ATTCTTAACT ATGTTGCTCC TTTTACGCTA
TGTGGATACG
3721 CTGCTTTAAT GCCTTTGTAT CATGCTATTG CTTCCCGTAT GGCTTTCATT
TTCTCCTCCT
3781 TGTATAAATC CTGGTTGCTG TCTCTTTATG AGGAGTTGTG GCCCGTTGTC
AGGCAACGTG
3841 GCGTGGTGTG CACTGTGTTT GCTGACGCAA CCCCCACTGG TTGGGGCATT
GCCACCACCT
3901 GTCAGCTCCT TTCCGGGACT TTCGCTTTCC CCCTCCCTAT TGCCACGGCG
GAACTCATCG
3961 CCGCCTGCCT TGCCCGCTGC TGGACAGGGG CTCGGCTGTT GGGCACTGAC
AATTCCGTGG
4021 TGTTGTCGGG GAAATCATCG TCCTTTCCTT GGCTGCTCGC CTGTGTTGCC
ACCTGGATTC
4081 TGCGCGGGAC GTCCTTCTGC TACGTCCCTT CGGCCCTCAA TCCAGCGGAC
CTTCCTTCCC
4141 GCGGCCTGCT GCCGGCTCTG CGGCCTCTTC CGCGTCTTCG CCTTCGCCCT
CAGACGAGTC
4201 GGATCTCCCT TTGGGCCGCC TCCCCGCAAG CTTCGCACTT TTTAAAAGAA
AAGGGAGGAC
4261 TGGATGGGAT TTATTACTCC GATAGGACGC TGGCTTGTAA CTCAGTCTCT
TACTAGGAGA
4321 CCAGCTTGAG CCTGGGTGTT CGCTGGTTAG CCTAACCTGG TTGGCCACCA
GGGGTAAGGA
4381 CTCCTTGGCT TAGAAAGCTA ATAAACTTGC CTGCATTAGA GCTCTTACGC
GTCCCGGGCT
4441 CGAGATCCGC ATCTCAATTA GTCAGCAACC ATAGTCCCGC CCCTAACTCC
GCCCATCCCG
4501 CCCCTAACTC CGCCCAGTTC CGCCCATTCT CCGCCCCATG GCTGACTAAT
TTTTTTTATT
4561 TATGCAGAGG CCGAGGCCGC CTCGGCCTCT GAGCTATTCC AGAAGTAGTG
AGGAGGCTTT
4621 TTTGGAGGCC TAGGCTTTTG CAAAAAGCTA ACTTGTTTAT TGCAGCTTAT
AATGGTTACA
4681 AATAAAGCAA TAGCATCACA AATTTCACAA ATAAAGCATT TTTTTCACTG
CATTCTAGTT
4741 GTGGTTTGTC CAAACTCATC AATGTATCTT ATCATGTCTG TCCGCTTCCT
CGCTCACTGA
4801 CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA GCTCACTCAA
AGGCGGTAAT

4861 ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA
AAGGCCAGCA
4921 AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT TTCCATAGGC
TCCGCCCCCC
4981 TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA
CAGGACTATA
5041 AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC
CGACCCTGCC
5101 GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT
CTCATAGCTC
5161 ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT
GTGTGCACGA
5221 ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC TATCGTCTTG
AGTCCAACCC
5281 GGTAAGACAC GACTTATCGC CACTGGCAGC AGCCACTGGT AACAGGATTA
GCAGAGCGAG
5341 GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT
ACACTAGAAG
5401 AACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA
GAGTTGGTAG
5461 CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT TTTTTTGTTT
GCAAGCAGCA
5521 GATTACGCGC AGAAAAAAAG GATCTCAAGA AGATCCTTTG ATCTTTTCTA
CGGGGTCTGA
5581 CGCTCAGTGG AACGAAAACT CACGTTAAGG GATTTTGGTC ATGAGATTAT
CAAAAAGGAT
5641 CTTCACCTAG ATCCTTTTAA ATTAAAAATG AAGTTTTAAA TCAATCTAAA
GTATATATGA
5701 GTAAACTTGG TCTGACAGTT AGAAAAACTC ATCGAGCATC AAATGAAACT
GCAATTTATT
5761 CATATCAGGA TTATCAATAC CATATTTTTG AAAAAGCCGT TTCTGTAATG
AAGGAGAAAA
5821 CTCACCGAGG CAGTTCCATA GGATGGCAAG ATCCTGGTAT CGGTCTGCGA
TTCCGACTCG
5881 TCCAACATCA ATACAACCTA TTAATTTCCC CTCGTCAAAA ATAAGGTTAT
CAAGTGAGAA
5941 ATCACCATGA GTGACGACTG AATCCGGTGA GAATGGCAAC AGCTTATGCA
TTTCTTTCCA
6001 GACTTGTTCA ACAGGCCAGC CATTACGCTC GTCATCAAAA TCACTCGCAT
CAACCAAACC
6061 GTTATTCATT CGTGATTGCG CCTGAGCGAG ACGAAATACG CGATCGCTGT
TAAAAGGACA
6121 ATTACAAACA GGAATCGAAT GCAACCGGCG CAGGAACACT GCCAGCGCAT
CAACAATATT
6181 TTCACCTGAA TCAGGATATT CTTCTAATAC CTGGAATGCT GTTTTTCCGG
GGATCGCAGT
6241 GGTGAGTAAC CATGCATCAT CAGGAGTACG GATAAAATGC TTGATGGTCG
GAAGAGGCAT
6301 AAATTCCGTC AGCCAGTTTA GTCTGACCAT CTCATCTGTA ACATCATTGG
CAACGCTACC
6361 TTTGCCATGT TTCAGAAACA ACTCTGGCGC ATCGGGCTTC CCATACAATC
GATAGATTGT
6421 CGCACCTGAT TGCCCGACAT TATCGCGAGC CCATTTATAC CCATATAAAT
CAGCATCCAT
6481 GTTGGAATTT AATCGCGGCC TAGAGCAAGA CGTTTCCCGT TGAATATGGC
TCATAACACC
6541 CCTTGTATTA CTGTTTATGT AAGCAGACAG TTTTATTGTT CATGATGATA
TATTTTTATC

6601 TTGTGCAATG TAACATCAGA GATTTTGAGA CACAACAATT GGTCGACGGA TCC

SEQ ID NO: 11

1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT
61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
TGACGTATGT
301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
361 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT
421 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAC
GGGACTTTCC
481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA
541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT
601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA
661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGCGTGTACG
GTGGGAGGTC
721 TATATAAGCA GAGCTCGCTG GCTTGTAACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC
781 TGGGTGTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA
841 GAAAGCTAAT AAACTTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCCT
CATTGACGCC
901 TCACTCTCTT GAACGGGAAT CTTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA
961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA
1021 GGCGTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA
1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC
1141 CGTAACTACT CTGGGCAAGT AGGGCAGGCG GTGGGTACGC AATGGGGGCG
GCTACCTCAG
1201 CACTAAATAG GAGACAATTA GACCAATTTG AGAAAATACG ACTTCGCCCG
AACGGAAAGA
1261 AAAAGTACCA AATTAAACAT TTAATATGGG CAGGCAAGGA GATGGAGCGC
TTCGGCCTCC
1321 ATGAGAGGTT GTTGGAGACA GAGGAGGGGT GTAAAAGAAT CATAGAAGTC
CTCTACCCCC
1381 TAGAACCAAC AGGATCGGAG GGCTTAAAAA GTCTGTTCAA TCTTGTGTGC
GTGCTATATT
1441 GCTTGCACAA GGAACAGAAA GTGAAAGACA CAGAGGAAGC AGTAGCAACA
GTAAGACAAC
1501 ACTGCCATCT AGTGGAAAAA GAAAAAAGTG CAACAGAGAC ATCTAGTGGA
CAAAAGAAAA

1561 ATGACAAGGG AATAGCAGCG CCACCTGGTG GCAGTCAGAA TTTTCCAGCG
CAACAACAAG
1621 GAAATGCCTG GGTACATGTA CCCTTGTCAC CGCGCACCTT AAATGCGTGG
GTAAAAGCAG
1681 TAGAGGAGAA AAAATTTGGA GCAGAAATAG TACCCATGTT TCAAGCCCTA
TCGAATTCCC
1741 GTTTGTGCTA GGGTTCTTAG GCTTCTTGGG GGCTGCTGGA ACTGCAATGG
GAGCAGCGGC
1801 GACAGCCCTG ACGGTCCAGT CTCAGCATTT GCTTGCTGGG ATACTGCAGC
AGCAGAAGAA
1861 TCTGCTGGCG GCTGTGGAGG CTCAACAGCA GATGTTGAAG CTGACCATTT
GGGGTGTTAA
1921 AAACCTCAAT GCCCGCGTCA CAGCCCTTGA GAAGTACCTA GAGGATCAGG
CACGACTAAA
1981 CTCCTGGGGG TGCGCATGGA AACAAGTATG TCATACCACA GTGGAGTGGC
CCTGGACAAA
2041 TCGGACTCCG GATTGGCAAA ATATGACTTG GTTGGAGTGG GAAAGACAAA
TAGCTGATTT
2101 GGAAAGCAAC ATTACGAGAC AATTAGTGAA GGCTAGAGAA CAAGAGGAAA
AGAATCTAGA
2161 TGCCTATCAG AAGTTAACTA GTTGGTCAGA TTTCTGGTCT TGGTTCGATT
TCTCAAAATG
2221 GCTTAACATT TTAAAAATGG GATTTTTAGT AATAGTAGGA ATAATAGGGT
TAAGATTACT
2281 TTACACAGTA TATGGATGTA TAGTGAGGGT TAGGCAGGGA TATGTTCCTC
TATCTCCACA
2341 GATCCATATC CGCGGCAATT TTAAAAGAAA GGGAGGAATA GGGGGACAGA
CTTCAGCAGA
2401 GAGACTAATT AATATAATAA CAACACAATT AGAAATACAA CATTTACAAA
CCAAAATTCA
2461 AAAAATTTTA AATTTTAGAG CCGCGGAGAT CTCAATATTG GCCATTAGCC
ATATTATTCA
2521 TTGGTTATAT AGCATAAATC AATATTGGCT ATTGGCCATT GCATACGTTG
TATCTATATC
2581 ATAATATGTA CATTTATATT GGCTCATGTC CAATATGACC GCCATGTTGG
CATTGATTAT
2641 TGACTAGTTA TTAATAGTAA TCAATTACGG GGTCATTAGT TCATAGCCCA
TATATGGAGT
2701 TCCGCGTTAC ATAACTTACG GTAAATGGCC CGCCTGGCTG ACCGCCCAAC
GACCCCCGCC
2761 CATTGACGTC AATAATGACG TATGTTCCCA TAGTAACGCC AATAGGGACT
TTCCATTGAC
2821 GTCAATGGGT GGAGTATTTA CGGTAAACTG CCCACTTGGC AGTACATCAA
GTGTATCATA
2881 TGCCAAGTCC GCCCCCTATT GACGTCAATG ACGGTAAATG GCCCGCCTGG
CATTATGCCC
2941 AGTACATGAC CTTACGGGAC TTTCCTACTT GGCAGTACAT CTACGTATTA
GTCATCGCTA
3001 TTACCATGGT GATGCGGTTT TGGCAGTACA CCAATGGGCG TGGATAGCGG
TTTGACTCAC
3061 GGGGATTTCC AAGTCTCCAC CCCATTGACG TCAATGGGAG TTTGTTTTGG
CACCAAAATC
3121 AACGGGACTT TCCAAAATGT CGTAATAACC CCGCCCCGTT GACGCAAATG
GGCGGTAGGC
3181 GTGTACGGTG GGAGGTCTAT ATAAGCAGAG CTCGTTTAGT GAACCGTCAG
ATCACTAGAA
3241 GCTTTATTGC GGTAGTTTAT CACAGTTAAA TTGCTAACGC AGTCAGTGCT
TCTGACACAA

3301 CAGTCTCGAA CTTAAGCTGC AGAAGTTGGT CGTGAGGCAC TGGGCAGGCT
AGCCACCAAT
3361 GCAGATTGAG CTGAGCACCT GCTTCTTCCT GTGCCTGCTG AGGTTCTGCT
TCTCTGCCAC
3421 CAGGAGATAC TACCTGGGGG CTGTGGAGCT GAGCTGGGAC TACATGCAGT
CTGACCTGGG
3481 GGAGCTGCCT GTGGATGCCA GGTTCCCCCC CAGAGTGCCC AAGAGCTTCC
CCTTCAACAC
3541 CTCTGTGGTG TACAAGAAGA CCCTGTTTGT GGAGTTCACT GACCACCTGT
TCAACATTGC
3601 CAAGCCCAGG CCCCCCTGGA TGGGCCTGCT GGGCCCCACC ATCCAGGCTG
AGGTGTATGA
3661 CACTGTGGTG ATCACCCTGA AGAACATGGC CAGCCACCCT GTGAGCCTGC
ATGCTGTGGG
3721 GGTGAGCTAC TGGAAGGCCT CTGAGGGGGC TGAGTATGAT GACCAGACCA
GCCAGAGGGA
3781 GAAGGAGGAT GACAAGGTGT TCCCTGGGGG CAGCCACACC TATGTGTGGC
AGGTGCTGAA
3841 GGAGAATGGC CCCATGGCCT CTGACCCCCT GTGCCTGACC TACAGCTACC
TGAGCCATGT
3901 GGACCTGGTG AAGGACCTGA ACTCTGGCCT GATTGGGGCC CTGCTGGTGT
GCAGGGAGGG
3961 CAGCCTGGCC AAGGAGAAGA CCCAGACCCT GCACAAGTTC ATCCTGCTGT
TTGCTGTGTT
4021 TGATGAGGGC AAGAGCTGGC ACTCTGAAAC CAAGAACAGC CTGATGCAGG
ACAGGGATGC
4081 TGCCTCTGCC AGGGCCTGGC CCAAGATGCA CACTGTGAAT GGCTATGTGA
ACAGGAGCCT
4141 GCCTGGCCTG ATTGGCTGCC ACAGGAAGTC TGTGTACTGG CATGTGATTG
GCATGGGCAC
4201 CACCCCTGAG GTGCACAGCA TCTTCCTGGA GGGCCACACC TTCCTGGTCA
GGAACCACAG
4261 GCAGGCCAGC CTGGAGATCA GCCCCATCAC CTTCCTGACT GCCCAGACCC
TGCTGATGGA
4321 CCTGGGCCAG TTCCTGCTGT TCTGCCACAT CAGCAGCCAC CAGCATGATG
GCATGGAGGC
4381 CTATGTGAAG GTGGACAGCT GCCCTGAGGA GCCCCAGCTG AGGATGAAGA
ACAATGAGGA
4441 GGCTGAGGAC TATGATGATG ACCTGACTGA CTCTGAGATG GATGTGGTGA
GGTTTGATGA
4501 TGACAACAGC CCCAGCTTCA TCCAGATCAG GTCTGTGGCC AAGAAGCACC
CCAAGACCTG
4561 GGTGCACTAC ATTGCTGCTG AGGAGGAGGA CTGGGACTAT GCCCCCCTGG
TGCTGGCCCC
4621 TGATGACAGG AGCTACAAGA GCCAGTACCT GAACAATGGC CCCCAGAGGA
TTGGCAGGAA
4681 GTACAAGAAG GTCAGGTTCA TGGCCTACAC TGATGAAACC TTCAAGACCA
GGGAGGCCAT
4741 CCAGCATGAG TCTGGCATCC TGGGCCCCCT GCTGTATGGG GAGGTGGGGG
ACACCCTGCT
4801 GATCATCTTC AAGAACCAGG CCAGCAGGCC CTACAACATC TACCCCCATG
GCATCACTGA
4861 TGTGAGGCCC CTGTACAGCA GGAGGCTGCC CAAGGGGGTG AAGCACCTGA
AGGACTTCCC
4921 CATCCTGCCT GGGGAGATCT TCAAGTACAA GTGGACTGTG ACTGTGGAGG
ATGGCCCCAC
4981 CAAGTCTGAC CCCAGGTGCC TGACCAGATA CTACAGCAGC TTTGTGAACA
TGGAGAGGGA

5041 CCTGGCCTCT GGCCTGATTG GCCCCCTGCT GATCTGCTAC AAGGAGTCTG
TGGACCAGAG
5101 GGGCAACCAG ATCATGTCTG ACAAGAGGAA TGTGATCCTG TTCTCTGTGT
TTGATGAGAA
5161 CAGGAGCTGG TACCTGACTG AGAACATCCA GAGGTTCCTG CCCAACCCTG
CTGGGGTGCA
5221 GCTGGAGGAC CCTGAGTTCC AGGCCAGCAA CATCATGCAC AGCATCAATG
GCTATGTGTT
5281 TGACAGCCTG CAGCTGTCTG TGTGCCTGCA TGAGGTGGCC TACTGGTACA
TCCTGAGCAT
5341 TGGGGCCCAG ACTGACTTCC TGTCTGTGTT CTTCTCTGGC TACACCTTCA
AGCACAAGAT
5401 GGTGTATGAG GACACCCTGA CCCTGTTCCC CTTCTCTGGG GAGACTGTGT
TCATGAGCAT
5461 GGAGAACCCT GGCCTGTGGA TTCTGGGCTG CCACAACTCT GACTTCAGGA
ACAGGGGCAT
5521 GACTGCCCTG CTGAAAGTCT CCAGCTGTGA CAAGAACACT GGGGACTACT
ATGAGGACAG
5581 CTATGAGGAC ATCTCTGCCT ACCTGCTGAG CAAGAACAAT GCCATTGAGC
CCAGGAGCTT
5641 CAGCCAGAAT GCCACTAATG TGTCTAACAA CAGCAACACC AGCAATGACA
GCAATGTGTC
5701 TCCCCCAGTG CTGAAGAGGC ACCAGAGGGA GATCACCAGG ACCACCCTGC
AGTCTGACCA
5761 GGAGGAGATT GACTATGATG ACACCATCTC TGTGGAGATG AAGAAGGAGG
ACTTTGACAT
5821 CTACGACGAG GACGAGAACC AGAGCCCCAG GAGCTTCCAG AAGAAGACCA
GGCACTACTT
5881 CATTGCTGCT GTGGAGAGGC TGTGGGACTA TGGCATGAGC AGCAGCCCCC
ATGTGCTGAG
5941 GAACAGGGCC CAGTCTGGCT CTGTGCCCCA GTTCAAGAAG GTGGTGTTCC
AGGAGTTCAC
6001 TGATGGCAGC TTCACCCAGC CCCTGTACAG AGGGGAGCTG AATGAGCACC
TGGGCCTGCT
6061 GGGCCCCTAC ATCAGGGCTG AGGTGGAGGA CAACATCATG GTGACCTTCA
GGAACCAGGC
6121 CAGCAGGCCC TACAGCTTCT ACAGCAGCCT GATCAGCTAT GAGGAGGACC
AGAGGCAGGG
6181 GGCTGAGCCC AGGAAGAACT TTGTGAAGCC CAATGAAACC AAGACCTACT
TCTGGAAGGT
6241 GCAGCACCAC ATGGCCCCCA CCAAGGATGA GTTTGACTGC AAGGCCTGGG
CCTACTTCTC
6301 TGATGTGGAC CTGGAGAAGG ATGTGCACTC TGGCCTGATT GGCCCCCTGC
TGGTGTGCCA
6361 CACCAACACC CTGAACCCTG CCCATGGCAG GCAGGTGACT GTGCAGGAGT
TTGCCCTGTT
6421 CTTCACCATC TTTGATGAAA CCAAGAGCTG GTACTTCACT GAGAACATGG
AGAGGAACTG
6481 CAGGGCCCCC TGCAACATCC AGATGGAGGA CCCCACCTTC AAGGAGAACT
ACAGGTTCCA
6541 TGCCATCAAT GGCTACATCA TGGACACCCT GCCTGGCCTG GTGATGGCCC
AGGACCAGAG
6601 GATCAGGTGG TACCTGCTGA GCATGGGCAG CAATGAGAAC ATCCACAGCA
TCCACTTCTC
6661 TGGCCATGTG TTCACTGTGA GGAAGAAGGA GGAGTACAAG ATGGCCCTGT
ACAACCTGTA
6721 CCCTGGGGTG TTTGAGACTG TGGAGATGCT GCCCAGCAAG GCTGGCATCT
GGAGGGTGGA

6781 GTGCCTGATT GGGGAGCACC TGCATGCTGG CATGAGCACC CTGTTCCTGG
TGTACAGCAA
6841 CAAGTGCCAG ACCCCCCTGG GCATGGCCTC TGGCCACATC AGGGACTTCC
AGATCACTGC
6901 CTCTGGCCAG TATGGCCAGT GGGCCCCCAA GCTGGCCAGG CTGCACTACT
CTGGCAGCAT
6961 CAATGCCTGG AGCACCAAGG AGCCCTTCAG CTGGATCAAG GTGGACCTGC
TGGCCCCCAT
7021 GATCATCCAT GGCATCAAGA CCCAGGGGGC CAGGCAGAAG TTCAGCAGCC
TGTACATCAG
7081 CCAGTTCATC ATCATGTACA GCCTGGATGG CAAGAAGTGG CAGACCTACA
GGGGCAACAG
7141 CACTGGCACC CTGATGGTGT TCTTTGGCAA TGTGGACAGC TCTGGCATCA
AGCACAACAT
7201 CTTCAACCCC CCCATCATTG CCAGATACAT CAGGCTGCAC CCCACCCACT
ACAGCATCAG
7261 GAGCACCCTG AGGATGGAGC TGATGGGCTG TGACCTGAAC AGCTGCAGCA
TGCCCCTGGG
7321 CATGGAGAGC AAGGCCATCT CTGATGCCCA GATCACTGCC AGCAGCTACT
TCACCAACAT
7381 GTTTGCCACC TGGAGCCCCA GCAAGGCCAG GCTGCACCTG CAGGGCAGGA
GCAATGCCTG
7441 GAGGCCCCAG GTCAACAACC CCAAGGAGTG GCTGCAGGTG GACTTCCAGA
AGACCATGAA
7501 GGTGACTGGG GTGACCACCC AGGGGGTGAA GAGCCTGCTG ACCAGCATGT
ATGTGAAGGA
7561 GTTCCTGATC AGCAGCAGCC AGGATGGCCA CCAGTGGACC CTGTTCTTCC
AGAATGGCAA
7621 GGTGAAGGTG TTCCAGGGCA ACCAGGACAG CTTCACCCCT GTGGTGAACA
GCCTGGACCC
7681 CCCCCTGCTG ACCAGATACC TGAGGATTCA CCCCCAGAGC TGGGTGCACC
AGATTGCCCT
7741 GAGGATGGAG GTGCTGGGCT GTGAGGCCCA GGACCTGTAC TGAGCGGCCG
CGGGCCCAAT
7801 CAACCTCTGG ATTACAAAAT TTGTGAAAGA TTGACTGGTA TTCTTAACTA
TGTTGCTCCT
7861 TTTACGCTAT GTGGATACGC TGCTTTAATG CCTTTGTATC ATGCTATTGC
TTCCCGTATG
7921 GCTTTCATTT TCTCCTCCTT GTATAAATCC TGGTTGCTGT CTCTTTATGA
GGAGTTGTGG
7981 CCCGTTGTCA GGCAACGTGG CGTGGTGTGC ACTGTGTTTG CTGACGCAAC
CCCCACTGGT
8041 TGGGGCATTG CCACCACCTG TCAGCTCCTT TCCGGGACTT TCGCTTTCCC
CCTCCCTATT
8101 GCCACGGCGG AACTCATCGC CGCCTGCCTT GCCCGCTGCT GGACAGGGGC
TCGGCTGTTG
8161 GGCACTGACA ATTCCGTGGT GTTGTCGGGG AAATCATCGT CCTTTCCTTG
GCTGCTCGCC
8221 TGTGTTGCCA CCTGGATTCT GCGCGGGACG TCCTTCTGCT ACGTCCCTTC
GGCCCTCAAT
8281 CCAGCGGACC TTCCTTCCCG CGGCCTGCTG CCGGCTCTGC GGCCTCTTCC
GCGTCTTCGC
8341 CTTCGCCCTC AGACGAGTCG GATCTCCCTT TGGGCCGCCT CCCCGCAAGC
TTCGCACTTT
8401 TTAAAAGAAA AGGGAGGACT GGATGGGATT TATTACTCCG ATAGGACGCT
GGCTTGTAAC
8461 TCAGTCTCTT ACTAGGAGAC CAGCTTGAGC CTGGGTGTTC GCTGGTTAGC
CTAACCTGGT

8521 TGGCCACCAG GGGTAAGGAC TCCTTGGCTT AGAAAGCTAA TAAACTTGCC
TGCATTAGAG
8581 CTCTTACGCG TCCCGGGCTC GAGATCCGCA TCTCAATTAG TCAGCAACCA
TAGTCCCGCC
8641 CCTAACTCCG CCCATCCCGC CCCTAACTCC GCCCAGTTCC GCCCATTCTC
CGCCCCATGG
8701 CTGACTAATT TTTTTTATTT ATGCAGAGGC CGAGGCCGCC TCGGCCTCTG
AGCTATTCCA
8761 GAAGTAGTGA GGAGGCTTTT TTGGAGGCCT AGGCTTTTGC AAAAAGCTAA
CTTGTTTATT
8821 GCAGCTTATA ATGGTTACAA ATAAAGCAAT AGCATCACAA ATTTCACAAA
TAAAGCATTT
8881 TTTTCACTGC ATTCTAGTTG TGGTTTGTCC AAACTCATCA ATGTATCTTA
TCATGTCTGT
8941 CCGCTTCCTC GCTCACTGAC TCGCTGCGCT CGGTCGTTCG GCTGCGGCGA
GCGGTATCAG
9001 CTCACTCAAA GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA
GGAAAGAACA
9061 TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG
CTGGCGTTTT
9121 TCCATAGGCT CCGCCCCCCT GACGAGCATC ACAAAAATCG ACGCTCAAGT
CAGAGGTGGC
9181 GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC
CTCGTGCGCT
9241 CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
TCGGGAAGCG
9301 TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC
GTTCGCTCCA
9361 AGCTGGGCTG TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA
TCCGGTAACT
9421 ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA
GCCACTGGTA
9481 ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG
TGGTGGCCTA
9541 ACTACGGCTA CACTAGAAGA ACAGTATTTG GTATCTGCGC TCTGCTGAAG
CCAGTTACCT
9601 TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT
AGCGGTGGTT
9661 TTTTTGTTTG CAAGCAGCAG ATTACGCGCA GAAAAAAAGG ATCTCAAGAA
GATCCTTTGA
9721 TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAACTC ACGTTAAGGG
ATTTTGGTCA
9781 TGAGATTATC AAAAAGGATC TTCACCTAGA TCCTTTTAAA TTAAAAATGA
AGTTTTAAAT
9841 CAATCTAAAG TATATATGAG TAAACTTGGT CTGACAGTTA GAAAAACTCA
TCGAGCATCA
9901 AATGAAACTG CAATTTATTC ATATCAGGAT TATCAATACC ATATTTTTGA
AAAAGCCGTT
9961 TCTGTAATGA AGGAGAAAAC TCACCGAGGC AGTTCCATAG GATGGCAAGA
TCCTGGTATC
10021 GGTCTGCGAT TCCGACTCGT CCAACATCAA TACAACCTAT TAATTTCCCC
TCGTCAAAAA
10081 TAAGGTTATC AAGTGAGAAA TCACCATGAG TGACGACTGA ATCCGGTGAG
AATGGCAACA
10141 GCTTATGCAT TTCTTTCCAG ACTTGTTCAA CAGGCCAGCC ATTACGCTCG
TCATCAAAAT
10201 CACTCGCATC AACCAAACCG TTATTCATTC GTGATTGCGC CTGAGCGAGA
CGAAATACGC

10261 GATCGCTGTT AAAAGGACAA TTACAAACAG GAATCGAATG CAACCGGCGC
AGGAACACTG
10321 CCAGCGCATC AACAATATTT TCACCTGAAT CAGGATATTC TTCTAATACC
TGGAATGCTG
10381 TTTTTCCGGG GATCGCAGTG GTGAGTAACC ATGCATCATC AGGAGTACGG
ATAAAATGCT
10441 TGATGGTCGG AAGAGGCATA AATTCCGTCA GCCAGTTTAG TCTGACCATC
TCATCTGTAA
10501 CATCATTGGC AACGCTACCT TTGCCATGTT TCAGAAACAA CTCTGGCGCA
TCGGGCTTCC
10561 CATACAATCG ATAGATTGTC GCACCTGATT GCCCGACATT ATCGCGAGCC
CATTTATACC
10621 CATATAAATC AGCATCCATG TTGGAATTTA ATCGCGGCCT AGAGCAAGAC
GTTTCCCGTT
10681 GAATATGGCT CATAACACCC CTTGTATTAC TGTTTATGTA AGCAGACAGT
TTTATTGTTC
10741 ATGATGATAT ATTTTTATCT TGTGCAATGT AACATCAGAG ATTTTGAGAC
ACAACAATTG
10801 GTCGACGGAT CC

SEQ ID NO: 12

1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT
61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
TGACGTATGT
301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
361 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT
421 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAC
GGGACTTTCC
481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA
541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT
601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA
661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGCGTGTACG
GTGGGAGGTC
721 TATATAAGCA GAGCTCGCTG GCTTGTAACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC
781 TGGGTGTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA
841 GAAAGCTAAT AAACTTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCCT
CATTGACGCC
901 TCACTCTCTT GAACGGGAAT CTTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA
961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA

1021 GGCGTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA
1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC
1141 CGTAACTACT CTGGGCAAGT AGGGCAGGCG GTGGGTACGC AATGGGGGCG
GCTACCTCAG
1201 CACTAAATAG GAGACAATTA GACCAATTTG AGAAAATACG ACTTCGCCCG
AACGGAAAGA
1261 AAAAGTACCA AATTAAACAT TTAATATGGG CAGGCAAGGA GATGGAGCGC
TTCGGCCTCC
1321 ATGAGAGGTT GTTGGAGACA GAGGAGGGGT GTAAAAGAAT CATAGAAGTC
CTCTACCCCC
1381 TAGAACCAAC AGGATCGGAG GGCTTAAAAA GTCTGTTCAA TCTTGTGTGC
GTGCTATATT
1441 GCTTGCACAA GGAACAGAAA GTGAAAGACA CAGAGGAAGC AGTAGCAACA
GTAAGACAAC
1501 ACTGCCATCT AGTGGAAAAA GAAAAAAGTG CAACAGAGAC ATCTAGTGGA
CAAAAGAAAA
1561 ATGACAAGGG AATAGCAGCG CCACCTGGTG GCAGTCAGAA TTTTCCAGCG
CAACAACAAG
1621 GAAATGCCTG GGTACATGTA CCCTTGTCAC CGCGCACCTT AAATGCGTGG
GTAAAAGCAG
1681 TAGAGGAGAA AAAATTTGGA GCAGAAATAG TACCCATGTT TCAAGCCCTA
TCGAATTCCC
1741 GTTTGTGCTA GGGTTCTTAG GCTTCTTGGG GGCTGCTGGA ACTGCAATGG
GAGCAGCGGC
1801 GACAGCCCTG ACGGTCCAGT CTCAGCATTT GCTTGCTGGG ATACTGCAGC
AGCAGAAGAA
1861 TCTGCTGGCG GCTGTGGAGG CTCAACAGCA GATGTTGAAG CTGACCATTT
GGGGTGTTAA
1921 AAACCTCAAT GCCCGCGTCA CAGCCCTTGA GAAGTACCTA GAGGATCAGG
CACGACTAAA
1981 CTCCTGGGGG TGCGCATGGA AACAAGTATG TCATACCACA GTGGAGTGGC
CCTGGACAAA
2041 TCGGACTCCG GATTGGCAAA ATATGACTTG GTTGGAGTGG GAAAGACAAA
TAGCTGATTT
2101 GGAAAGCAAC ATTACGAGAC AATTAGTGAA GGCTAGAGAA CAAGAGGAAA
AGAATCTAGA
2161 TGCCTATCAG AAGTTAACTA GTTGGTCAGA TTTCTGGTCT TGGTTCGATT
TCTCAAAATG
2221 GCTTAACATT TTAAAAATGG GATTTTTAGT AATAGTAGGA ATAATAGGGT
TAAGATTACT
2281 TTACACAGTA TATGGATGTA TAGTGAGGGT TAGGCAGGGA TATGTTCCTC
TATCTCCACA
2341 GATCCATATC CGCGGCAATT TTAAAAGAAA GGGAGGAATA GGGGGACAGA
CTTCAGCAGA
2401 GAGACTAATT AATATAATAA CAACACAATT AGAAATACAA CATTTACAAA
CCAAAATTCA
2461 AAAAATTTTA AATTTTAGAG CCGCGGAGAT CTGTTACATA ACTTATGGTA
AATGGCCTGC
2521 CTGGCTGACT GCCCAATGAC CCCTGCCCAA TGATGTCAAT AATGATGTAT
GTTCCCATGT
2581 AATGCCAATA GGGACTTTCC ATTGATGTCA ATGGGTGGAG TATTTATGGT
AACTGCCCAC
2641 TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTATGCCCC CTATTGATGT
CAATGATGGT
2701 AAATGGCCTG CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC
TACTTGGCAG

2761 TACATCTATG TATTAGTCAT TGCTATTACC ATGGGAATTC ACTAGTGGAG
AAGAGCATGC
2821 TTGAGGGCTG AGTGCCCCTC AGTGGGCAGA GAGCACATGG CCCACAGTCC
CTGAGAAGTT
2881 GGGGGGAGGG GTGGGCAATT GAACTGGTGC CTAGAGAAGG TGGGGCTTGG
GTAAACTGGG
2941 AAAGTGATGT GGTGTACTGG CTCCACCTTT TTCCCCAGGG TGGGGGAGAA
CCATATATAA
3001 GTGCAGTAGT CTCTGTGAAC ATTCAAGCTT CTGCCTTCTC CCTCCTGTGA
GTTTGCTAGC
3061 CACCAATGCA GATTGAGCTG AGCACCTGCT TCTTCCTGTG CCTGCTGAGG
TTCTGCTTCT
3121 CTGCCACCAG GAGATACTAC CTGGGGGCTG TGGAGCTGAG CTGGGACTAC
ATGCAGTCTG
3181 ACCTGGGGGA GCTGCCTGTG GATGCCAGGT TCCCCCCCAG AGTGCCCAAG
AGCTTCCCCT
3241 TCAACACCTC TGTGGTGTAC AAGAAGACCC TGTTTGTGGA GTTCACTGAC
CACCTGTTCA
3301 ACATTGCCAA GCCCAGGCCC CCCTGGATGG GCCTGCTGGG CCCCACCATC
CAGGCTGAGG
3361 TGTATGACAC TGTGGTGATC ACCCTGAAGA ACATGGCCAG CCACCCTGTG
AGCCTGCATG
3421 CTGTGGGGGT GAGCTACTGG AAGGCCTCTG AGGGGGCTGA GTATGATGAC
CAGACCAGCC
3481 AGAGGGAGAA GGAGGATGAC AAGGTGTTCC CTGGGGGCAG CCACACCTAT
GTGTGGCAGG
3541 TGCTGAAGGA GAATGGCCCC ATGGCCTCTG ACCCCCTGTG CCTGACCTAC
AGCTACCTGA
3601 GCCATGTGGA CCTGGTGAAG GACCTGAACT CTGGCCTGAT TGGGGCCCTG
CTGGTGTGCA
3661 GGGAGGGCAG CCTGGCCAAG GAGAAGACCC AGACCCTGCA CAAGTTCATC
CTGCTGTTTG
3721 CTGTGTTTGA TGAGGGCAAG AGCTGGCACT CTGAAACCAA GAACAGCCTG
ATGCAGGACA
3781 GGGATGCTGC CTCTGCCAGG GCCTGGCCCA AGATGCACAC TGTGAATGGC
TATGTGAACA
3841 GGAGCCTGCC TGGCCTGATT GGCTGCCACA GGAAGTCTGT GTACTGGCAT
GTGATTGGCA
3901 TGGGCACCAC CCCTGAGGTG CACAGCATCT TCCTGGAGGG CCACACCTTC
CTGGTCAGGA
3961 ACCACAGGCA GGCCAGCCTG GAGATCAGCC CCATCACCTT CCTGACTGCC
CAGACCCTGC
4021 TGATGGACCT GGGCCAGTTC CTGCTGTTCT GCCACATCAG CAGCCACCAG
CATGATGGCA
4081 TGGAGGCCTA TGTGAAGGTG GACAGCTGCC CTGAGGAGCC CCAGCTGAGG
ATGAAGAACA
4141 ATGAGGAGGC TGAGGACTAT GATGATGACC TGACTGACTC TGAGATGGAT
GTGGTGAGGT
4201 TTGATGATGA CAACAGCCCC AGCTTCATCC AGATCAGGTC TGTGGCCAAG
AAGCACCCCA
4261 AGACCTGGGT GCACTACATT GCTGCTGAGG AGGAGGACTG GGACTATGCC
CCCCTGGTGC
4321 TGGCCCCTGA TGACAGGAGC TACAAGAGCC AGTACCTGAA CAATGGCCCC
CAGAGGATTG
4381 GCAGGAAGTA CAAGAAGGTC AGGTTCATGG CCTACACTGA TGAAACCTTC
AAGACCAGGG
4441 AGGCCATCCA GCATGAGTCT GGCATCCTGG GCCCCCTGCT GTATGGGGAG
GTGGGGGACA

4501 CCCTGCTGAT CATCTTCAAG AACCAGGCCA GCAGGCCCTA CAACATCTAC
CCCCATGGCA
4561 TCACTGATGT GAGGCCCCTG TACAGCAGGA GGCTGCCCAA GGGGGTGAAG
CACCTGAAGG
4621 ACTTCCCCAT CCTGCCTGGG GAGATCTTCA AGTACAAGTG GACTGTGACT
GTGGAGGATG
4681 GCCCCACCAA GTCTGACCCC AGGTGCCTGA CCAGATACTA CAGCAGCTTT
GTGAACATGG
4741 AGAGGGACCT GGCCTCTGGC CTGATTGGCC CCCTGCTGAT CTGCTACAAG
GAGTCTGTGG
4801 ACCAGAGGGG CAACCAGATC ATGTCTGACA AGAGGAATGT GATCCTGTTC
TCTGTGTTTG
4861 ATGAGAACAG GAGCTGGTAC CTGACTGAGA ACATCCAGAG GTTCCTGCCC
AACCCTGCTG
4921 GGGTGCAGCT GGAGGACCCT GAGTTCCAGG CCAGCAACAT CATGCACAGC
ATCAATGGCT
4981 ATGTGTTTGA CAGCCTGCAG CTGTCTGTGT GCCTGCATGA GGTGGCCTAC
TGGTACATCC
5041 TGAGCATTGG GGCCCAGACT GACTTCCTGT CTGTGTTCTT CTCTGGCTAC
ACCTTCAAGC
5101 ACAAGATGGT GTATGAGGAC ACCCTGACCC TGTTCCCCTT CTCTGGGGAG
ACTGTGTTCA
5161 TGAGCATGGA GAACCCTGGC CTGTGGATTC TGGGCTGCCA CAACTCTGAC
TTCAGGAACA
5221 GGGGCATGAC TGCCCTGCTG AAAGTCTCCA GCTGTGACAA GAACACTGGG
GACTACTATG
5281 AGGACAGCTA TGAGGACATC TCTGCCTACC TGCTGAGCAA GAACAATGCC
ATTGAGCCCA
5341 GGAGCTTCAG CCAGAATGCC ACTAATGTGT CTAACAACAG CAACACCAGC
AATGACAGCA
5401 ATGTGTCTCC CCCAGTGCTG AAGAGGCACC AGAGGGAGAT CACCAGGACC
ACCCTGCAGT
5461 CTGACCAGGA GGAGATTGAC TATGATGACA CCATCTCTGT GGAGATGAAG
AAGGAGGACT
5521 TTGACATCTA CGACGAGGAC GAGAACCAGA GCCCCAGGAG CTTCCAGAAG
AAGACCAGGC
5581 ACTACTTCAT TGCTGCTGTG GAGAGGCTGT GGGACTATGG CATGAGCAGC
AGCCCCCATG
5641 TGCTGAGGAA CAGGGCCCAG TCTGGCTCTG TGCCCCAGTT CAAGAAGGTG
GTGTTCCAGG
5701 AGTTCACTGA TGGCAGCTTC ACCCAGCCCC TGTACAGAGG GGAGCTGAAT
GAGCACCTGG
5761 GCCTGCTGGG CCCCTACATC AGGGCTGAGG TGGAGGACAA CATCATGGTG
ACCTTCAGGA
5821 ACCAGGCCAG CAGGCCCTAC AGCTTCTACA GCAGCCTGAT CAGCTATGAG
GAGGACCAGA
5881 GGCAGGGGGC TGAGCCCAGG AAGAACTTTG TGAAGCCCAA TGAAACCAAG
ACCTACTTCT
5941 GGAAGGTGCA GCACCACATG GCCCCCACCA AGGATGAGTT TGACTGCAAG
GCCTGGGCCT
6001 ACTTCTCTGA TGTGGACCTG GAGAAGGATG TGCACTCTGG CCTGATTGGC
CCCCTGCTGG
6061 TGTGCCACAC CAACACCCTG AACCCTGCCC ATGGCAGGCA GGTGACTGTG
CAGGAGTTTG
6121 CCCTGTTCTT CACCATCTTT GATGAAACCA AGAGCTGGTA CTTCACTGAG
AACATGGAGA
6181 GGAACTGCAG GGCCCCCTGC AACATCCAGA TGGAGGACCC CACCTTCAAG
GAGAACTACA

6241 GGTTCCATGC CATCAATGGC TACATCATGG ACACCCTGCC TGGCCTGGTG
ATGGCCCAGG
6301 ACCAGAGGAT CAGGTGGTAC CTGCTGAGCA TGGGCAGCAA TGAGAACATC
CACAGCATCC
6361 ACTTCTCTGG CCATGTGTTC ACTGTGAGGA AGAAGGAGGA GTACAAGATG
GCCCTGTACA
6421 ACCTGTACCC TGGGGTGTTT GAGACTGTGG AGATGCTGCC CAGCAAGGCT
GGCATCTGGA
6481 GGGTGGAGTG CCTGATTGGG GAGCACCTGC ATGCTGGCAT GAGCACCCTG
TTCCTGGTGT
6541 ACAGCAACAA GTGCCAGACC CCCCTGGGCA TGGCCTCTGG CCACATCAGG
GACTTCCAGA
6601 TCACTGCCTC TGGCCAGTAT GGCCAGTGGG CCCCCAAGCT GGCCAGGCTG
CACTACTCTG
6661 GCAGCATCAA TGCCTGGAGC ACCAAGGAGC CCTTCAGCTG GATCAAGGTG
GACCTGCTGG
6721 CCCCCATGAT CATCCATGGC ATCAAGACCC AGGGGGCCAG GCAGAAGTTC
AGCAGCCTGT
6781 ACATCAGCCA GTTCATCATC ATGTACAGCC TGGATGGCAA GAAGTGGCAG
ACCTACAGGG
6841 GCAACAGCAC TGGCACCCTG ATGGTGTTCT TTGGCAATGT GGACAGCTCT
GGCATCAAGC
6901 ACAACATCTT CAACCCCCCC ATCATTGCCA GATACATCAG GCTGCACCCC
ACCCACTACA
6961 GCATCAGGAG CACCCTGAGG ATGGAGCTGA TGGGCTGTGA CCTGAACAGC
TGCAGCATGC
7021 CCCTGGGCAT GGAGAGCAAG GCCATCTCTG ATGCCCAGAT CACTGCCAGC
AGCTACTTCA
7081 CCAACATGTT TGCCACCTGG AGCCCCAGCA AGGCCAGGCT GCACCTGCAG
GGCAGGAGCA
7141 ATGCCTGGAG GCCCCAGGTC AACAACCCCA AGGAGTGGCT GCAGGTGGAC
TTCCAGAAGA
7201 CCATGAAGGT GACTGGGGTG ACCACCCAGG GGGTGAAGAG CCTGCTGACC
AGCATGTATG
7261 TGAAGGAGTT CCTGATCAGC AGCAGCCAGG ATGGCCACCA GTGGACCCTG
TTCTTCCAGA
7321 ATGGCAAGGT GAAGGTGTTC CAGGGCAACC AGGACAGCTT CACCCCTGTG
GTGAACAGCC
7381 TGGACCCCCC CCTGCTGACC AGATACCTGA GGATTCACCC CCAGAGCTGG
GTGCACCAGA
7441 TTGCCCTGAG GATGGAGGTG CTGGGCTGTG AGGCCCAGGA CCTGTACTGA
GCGGCCGCGG
7501 GCCCAATCAA CCTCTGGATT ACAAAATTTG TGAAAGATTG ACTGGTATTC
TTAACTATGT
7561 TGCTCCTTTT ACGCTATGTG GATACGCTGC TTTAATGCCT TTGTATCATG
CTATTGCTTC
7621 CCGTATGGCT TTCATTTTCT CCTCCTTGTA TAAATCCTGG TTGCTGTCTC
TTTATGAGGA
7681 GTTGTGGCCC GTTGTCAGGC AACGTGGCGT GGTGTGCACT GTGTTTGCTG
ACGCAACCCC
7741 CACTGGTTGG GGCATTGCCA CCACCTGTCA GCTCCTTTCC GGGACTTTCG
CTTTCCCCCT
7801 CCCTATTGCC ACGGCGGAAC TCATCGCCGC CTGCCTTGCC CGCTGCTGGA
CAGGGGCTCG
7861 GCTGTTGGGC ACTGACAATT CCGTGGTGTT GTCGGGGAAA TCATCGTCCT
TTCCTTGGCT
7921 GCTCGCCTGT GTTGCCACCT GGATTCTGCG CGGGACGTCC TTCTGCTACG
TCCCTTCGGC

7981 CCTCAATCCA GCGGACCTTC CTTCCCGCGG CCTGCTGCCG GCTCTGCGGC
CTCTTCCGCG
8041 TCTTCGCCTT CGCCCTCAGA CGAGTCGGAT CTCCCTTTGG GCCGCCTCCC
CGCAAGCTTC
8101 GCACTTTTTA AAAGAAAAGG GAGGACTGGA TGGGATTTAT TACTCCGATA
GGACGCTGGC
8161 TTGTAACTCA GTCTCTTACT AGGAGACCAG CTTGAGCCTG GGTGTTCGCT
GGTTAGCCTA
8221 ACCTGGTTGG CCACCAGGGG TAAGGACTCC TTGGCTTAGA AAGCTAATAA
ACTTGCCTGC
8281 ATTAGAGCTC TTACGCGTCC CGGGCTCGAG ATCCGCATCT CAATTAGTCA
GCAACCATAG
8341 TCCCGCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC CAGTTCCGCC
CATTCTCCGC
8401 CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA GGCCGCCTCG
GCCTCTGAGC
8461 TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG CTTTTGCAAA
AAGCTAACTT
8521 GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT
TCACAAATAA
8581 AGCATTTTTT TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATG
TATCTTATCA
8641 TGTCTGTCCG CTTCCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT
GCGGCGAGCG
8701 GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA
TAACGCAGGA
8761 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC
CGCGTTGCTG
8821 GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG
CTCAAGTCAG
8881 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG
AAGCTCCCTC
8941 GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT
TCTCCCTTCG
9001 GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT
GTAGGTCGTT
9061 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAGC CCGACCGCTG
CGCCTTATCC
9121 GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT
GGCAGCAGCC
9181 ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT
CTTGAAGTGG
9241 TGGCCTAACT ACGGCTACAC TAGAAGAACA GTATTTGGTA TCTGCGCTCT
GCTGAAGCCA
9301 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC
CGCTGGTAGC
9361 GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC
TCAAGAAGAT
9421 CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG AAAACTCACG
TTAAGGGATT
9481 TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA
AAAATGAAGT
9541 TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTAGAA
AAACTCATCG
9601 AGCATCAAAT GAAACTGCAA TTTATTCATA TCAGGATTAT CAATACCATA
TTTTTGAAAA
9661 AGCCGTTTCT GTAATGAAGG AGAAAACTCA CCGAGGCAGT TCCATAGGAT
GGCAAGATCC

```
     9721 TGGTATCGGT CTGCGATTCC GACTCGTCCA ACATCAATAC AACCTATTAA
TTTCCCCTCG
     9781 TCAAAAATAA GGTTATCAAG TGAGAAATCA CCATGAGTGA CGACTGAATC
CGGTGAGAAT
     9841 GGCAACAGCT TATGCATTTC TTTCCAGACT TGTTCAACAG GCCAGCCATT
ACGCTCGTCA
     9901 TCAAAATCAC TCGCATCAAC CAAACCGTTA TTCATTCGTG ATTGCGCCTG
AGCGAGACGA
     9961 AATACGCGAT CGCTGTTAAA AGGACAATTA CAAACAGGAA TCGAATGCAA
CCGGCGCAGG
    10021 AACACTGCCA GCGCATCAAC AATATTTTCA CCTGAATCAG GATATTCTTC
TAATACCTGG
    10081 AATGCTGTTT TTCCGGGGAT CGCAGTGGTG AGTAACCATG CATCATCAGG
AGTACGGATA
    10141 AAATGCTTGA TGGTCGGAAG AGGCATAAAT TCCGTCAGCC AGTTTAGTCT
GACCATCTCA
    10201 TCTGTAACAT CATTGGCAAC GCTACCTTTG CCATGTTTCA GAAACAACTC
TGGCGCATCG
    10261 GGCTTCCCAT ACAATCGATA GATTGTCGCA CCTGATTGCC CGACATTATC
GCGAGCCCAT
    10321 TTATACCCAT ATAAATCAGC ATCCATGTTG GAATTTAATC GCGGCCTAGA
GCAAGACGTT
    10381 TCCCGTTGAA TATGGCTCAT AACACCCCTT GTATTACTGT TTATGTAAGC
AGACAGTTTT
    10441 ATTGTTCATG ATGATATATT TTTATCTTGT GCAATGTAAC ATCAGAGATT
TTGAGACACA
    10501 ACAATTGGTC GACGGATCC
```

SEQ ID NO: 13

```
        1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT
       61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
      121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
      181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
      241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
TGACGTATGT
      301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
      361 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT
      421 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAC
GGGACTTTCC
      481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA
      541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT
      601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA
      661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGCGTGTACG
GTGGGAGGTC
      721 TATATAAGCA GAGCTCGCTG GCTTGTAACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC
```

781 TGGGTGTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA
841 GAAAGCTAAT AAACTTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCCT
CATTGACGCC
901 TCACTCTCTT GAACGGGAAT CTTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA
961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA
1021 GGCGTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA
1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC
1141 CGTAACTACT CTGGGCAAGT AGGGCAGGCG GTGGGTACGC AATGGGGGCG
GCTACCTCAG
1201 CACTAAATAG GAGACAATTA GACCAATTTG AGAAAATACG ACTTCGCCCG
AACGGAAAGA
1261 AAAAGTACCA AATTAAACAT TTAATATGGG CAGGCAAGGA GATGGAGCGC
TTCGGCCTCC
1321 ATGAGAGGTT GTTGGAGACA GAGGAGGGGT GTAAAAGAAT CATAGAAGTC
CTCTACCCCC
1381 TAGAACCAAC AGGATCGGAG GGCTTAAAAA GTCTGTTCAA TCTTGTGTGC
GTGCTATATT
1441 GCTTGCACAA GGAACAGAAA GTGAAAGACA CAGAGGAAGC AGTAGCAACA
GTAAGACAAC
1501 ACTGCCATCT AGTGGAAAAA GAAAAAAGTG CAACAGAGAC ATCTAGTGGA
CAAAAGAAAA
1561 ATGACAAGGG AATAGCAGCG CCACCTGGTG GCAGTCAGAA TTTTCCAGCG
CAACAACAAG
1621 GAAATGCCTG GGTACATGTA CCCTTGTCAC CGCGCACCTT AAATGCGTGG
GTAAAAGCAG
1681 TAGAGGAGAA AAAATTTGGA GCAGAAATAG TACCCATGTT TCAAGCCCTA
TCGAATTCCC
1741 GTTTGTGCTA GGGTTCTTAG GCTTCTTGGG GGCTGCTGGA ACTGCAATGG
GAGCAGCGGC
1801 GACAGCCCTG ACGGTCCAGT CTCAGCATTT GCTTGCTGGG ATACTGCAGC
AGCAGAAGAA
1861 TCTGCTGGCG GCTGTGGAGG CTCAACAGCA GATGTTGAAG CTGACCATTT
GGGGTGTTAA
1921 AAACCTCAAT GCCCGCGTCA CAGCCCTTGA GAAGTACCTA GAGGATCAGG
CACGACTAAA
1981 CTCCTGGGGG TGCGCATGGA AACAAGTATG TCATACCACA GTGGAGTGGC
CCTGGACAAA
2041 TCGGACTCCG GATTGGCAAA ATATGACTTG GTTGGAGTGG GAAAGACAAA
TAGCTGATTT
2101 GGAAAGCAAC ATTACGAGAC AATTAGTGAA GGCTAGAGAA CAAGAGGAAA
AGAATCTAGA
2161 TGCCTATCAG AAGTTAACTA GTTGGTCAGA TTTCTGGTCT TGGTTCGATT
TCTCAAAATG
2221 GCTTAACATT TTAAAAATGG GATTTTTAGT AATAGTAGGA ATAATAGGGT
TAAGATTACT
2281 TTACACAGTA TATGGATGTA TAGTGAGGGT TAGGCAGGGA TATGTTCCTC
TATCTCCACA
2341 GATCCATATC CGCGGCAATT TTAAAAGAAA GGGAGGAATA GGGGGACAGA
CTTCAGCAGA
2401 GAGACTAATT AATATAATAA CAACACAATT AGAAATACAA CATTTACAAA
CCAAAATTCA
2461 AAAAATTTTA AATTTTAGAG CCGCGGAGAT CTCAATATTG GCCATTAGCC
ATATTATTCA

2521 TTGGTTATAT AGCATAAATC AATATTGGCT ATTGGCCATT GCATACGTTG
TATCTATATC
2581 ATAATATGTA CATTTATATT GGCTCATGTC CAATATGACC GCCATGTTGG
CATTGATTAT
2641 TGACTAGTTA TTAATAGTAA TCAATTACGG GGTCATTAGT TCATAGCCCA
TATATGGAGT
2701 TCCGCGTTAC ATAACTTACG GTAAATGGCC CGCCTGGCTG ACCGCCCAAC
GACCCCCGCC
2761 CATTGACGTC AATAATGACG TATGTTCCCA TAGTAACGCC AATAGGGACT
TTCCATTGAC
2821 GTCAATGGGT GGAGTATTTA CGGTAAACTG CCCACTTGGC AGTACATCAA
GTGTATCATA
2881 TGCCAAGTCC GCCCCCTATT GACGTCAATG ACGGTAAATG GCCCGCCTGG
CATTATGCCC
2941 AGTACATGAC CTTACGGGAC TTTCCTACTT GGCAGTACAT CTACGTATTA
GTCATCGCTA
3001 TTACCATGGT GATGCGGTTT TGGCAGTACA CCAATGGGCG TGGATAGCGG
TTTGACTCAC
3061 GGGGATTTCC AAGTCTCCAC CCCATTGACG TCAATGGGAG TTTGTTTTGG
CACCAAAATC
3121 AACGGGACTT TCCAAAATGT CGTAATAACC CCGCCCCGTT GACGCAAATG
GGCGGTAGGC
3181 GTGTACGGTG GGAGGTCTAT ATAAGCAGAG CTCGTTTAGT GAACCGTCAG
ATCACTAGAA
3241 GCTTTATTGC GGTAGTTTAT CACAGTTAAA TTGCTAACGC AGTCAGTGCT
TCTGACACAA
3301 CAGTCTCGAA CTTAAGCTGC AGAAGTTGGT CGTGAGGCAC TGGGCAGGCT
AGCCACCAAT
3361 GCAGATTGAG CTGAGCACCT GCTTCTTCCT GTGCCTGCTG AGGTTCTGCT
TCTCTGCCAC
3421 CAGGAGATAC TACCTGGGGG CTGTGGAGCT GAGCTGGGAC TACATGCAGT
CTGACCTGGG
3481 GGAGCTGCCT GTGGATGCCA GGTTCCCCCC CAGAGTGCCC AAGAGCTTCC
CCTTCAACAC
3541 CTCTGTGGTG TACAAGAAGA CCCTGTTTGT GGAGTTCACT GACCACCTGT
TCAACATTGC
3601 CAAGCCCAGG CCCCCCTGGA TGGGCCTGCT GGGCCCCACC ATCCAGGCTG
AGGTGTATGA
3661 CACTGTGGTG ATCACCCTGA AGAACATGGC CAGCCACCCT GTGAGCCTGC
ATGCTGTGGG
3721 GGTGAGCTAC TGGAAGGCCT CTGAGGGGGC TGAGTATGAT GACCAGACCA
GCCAGAGGGA
3781 GAAGGAGGAT GACAAGGTGT TCCCTGGGGG CAGCCACACC TATGTGTGGC
AGGTGCTGAA
3841 GGAGAATGGC CCCATGGCCT CTGACCCCCT GTGCCTGACC TACAGCTACC
TGAGCCATGT
3901 GGACCTGGTG AAGGACCTGA ACTCTGGCCT GATTGGGGCC CTGCTGGTGT
GCAGGGAGGG
3961 CAGCCTGGCC AAGGAGAAGA CCCAGACCCT GCACAAGTTC ATCCTGCTGT
TTGCTGTGTT
4021 TGATGAGGGC AAGAGCTGGC ACTCTGAAAC CAAGAACAGC CTGATGCAGG
ACAGGGATGC
4081 TGCCTCTGCC AGGGCCTGGC CCAAGATGCA CACTGTGAAT GGCTATGTGA
ACAGGAGCCT
4141 GCCTGGCCTG ATTGGCTGCC ACAGGAAGTC TGTGTACTGG CATGTGATTG
GCATGGGCAC
4201 CACCCCTGAG GTGCACAGCA TCTTCCTGGA GGGCCACACC TTCCTGGTCA
GGAACCACAG

4261 GCAGGCCAGC CTGGAGATCA GCCCCATCAC CTTCCTGACT GCCCAGACCC
TGCTGATGGA
4321 CCTGGGCCAG TTCCTGCTGT TCTGCCACAT CAGCAGCCAC CAGCATGATG
GCATGGAGGC
4381 CTATGTGAAG GTGGACAGCT GCCCTGAGGA GCCCCAGCTG AGGATGAAGA
ACAATGAGGA
4441 GGCTGAGGAC TATGATGATG ACCTGACTGA CTCTGAGATG GATGTGGTGA
GGTTTGATGA
4501 TGACAACAGC CCCAGCTTCA TCCAGATCAG GTCTGTGGCC AAGAAGCACC
CCAAGACCTG
4561 GGTGCACTAC ATTGCTGCTG AGGAGGAGGA CTGGGACTAT GCCCCCCTGG
TGCTGGCCCC
4621 TGATGACAGG AGCTACAAGA GCCAGTACCT GAACAATGGC CCCCAGAGGA
TTGGCAGGAA
4681 GTACAAGAAG GTCAGGTTCA TGGCCTACAC TGATGAAACC TTCAAGACCA
GGGAGGCCAT
4741 CCAGCATGAG TCTGGCATCC TGGGCCCCCT GCTGTATGGG GAGGTGGGGG
ACACCCTGCT
4801 GATCATCTTC AAGAACCAGG CCAGCAGGCC CTACAACATC TACCCCCATG
GCATCACTGA
4861 TGTGAGGCCC CTGTACAGCA GGAGGCTGCC CAAGGGGGTG AAGCACCTGA
AGGACTTCCC
4921 CATCCTGCCT GGGGAGATCT TCAAGTACAA GTGGACTGTG ACTGTGGAGG
ATGGCCCCAC
4981 CAAGTCTGAC CCCAGGTGCC TGACCAGATA CTACAGCAGC TTTGTGAACA
TGGAGAGGGA
5041 CCTGGCCTCT GGCCTGATTG GCCCCCTGCT GATCTGCTAC AAGGAGTCTG
TGGACCAGAG
5101 GGGCAACCAG ATCATGTCTG ACAAGAGGAA TGTGATCCTG TTCTCTGTGT
TTGATGAGAA
5161 CAGGAGCTGG TACCTGACTG AGAACATCCA GAGGTTCCTG CCCAACCCTG
CTGGGGTGCA
5221 GCTGGAGGAC CCTGAGTTCC AGGCCAGCAA CATCATGCAC AGCATCAATG
GCTATGTGTT
5281 TGACAGCCTG CAGCTGTCTG TGTGCCTGCA TGAGGTGGCC TACTGGTACA
TCCTGAGCAT
5341 TGGGGCCCAG ACTGACTTCC TGTCTGTGTT CTTCTCTGGC TACACCTTCA
AGCACAAGAT
5401 GGTGTATGAG GACACCCTGA CCCTGTTCCC CTTCTCTGGG GAGACTGTGT
TCATGAGCAT
5461 GGAGAACCCT GGCCTGTGGA TTCTGGGCTG CCACAACTCT GACTTCAGGA
ACAGGGGCAT
5521 GACTGCCCTG CTGAAAGTCT CCAGCTGTGA CAAGAACACT GGGGACTACT
ATGAGGACAG
5581 CTATGAGGAC ATCTCTGCCT ACCTGCTGAG CAAGAACAAT GCCATTGAGC
CCAGGAGCTT
5641 CAGCCAGAAC AGCAGGCACC CCAGCACCAG GCAGAAGCAG TTCAATGCCA
CCACCATCCC
5701 TGAGAATGAC ATAGAGAAGA CAGACCCATG GTTTGCCCAC CGGACCCCCA
TGCCCAAGAT
5761 CCAGAATGTG AGCAGCTCTG ACCTGCTGAT GCTGCTGAGG CAGAGCCCCA
CCCCCCATGG
5821 CCTGAGCCTG TCTGACCTGC AGGAGGCCAA GTATGAAACC TTCTCTGATG
ACCCCAGCCC
5881 TGGGGCCATT GACAGCAACA ACAGCCTGTC TGAGATGACC CACTTCAGGC
CCCAGCTGCA
5941 CCACTCTGGG GACATGGTGT TCACCCCTGA GTCTGGCCTG CAGCTGAGGC
TGAATGAGAA

6001 GCTGGGCACC ACTGCTGCCA CTGAGCTGAA GAAGCTGGAC TTCAAAGTCT
CCAGCACCAG
6061 CAACAACCTG ATCAGCACCA TCCCCTCTGA CAACCTGGCT GCTGGCACTG
ACAACACCAG
6121 CAGCCTGGGC CCCCCCAGCA TGCCTGTGCA CTATGACAGC CAGCTGGACA
CCACCCTGTT
6181 TGGCAAGAAG AGCAGCCCCC TGACTGAGTC TGGGGGCCCC CTGAGCCTGT
CTGAGGAGAA
6241 CAATGACAGC AAGCTGCTGG AGTCTGGCCT GATGAACAGC CAGGAGAGCA
GCTGGGGCAA
6301 GAATGTGAGC AGCAGGGAGA TCACCAGGAC CACCCTGCAG TCTGACCAGG
AGGAGATTGA
6361 CTATGATGAC ACCATCTCTG TGGAGATGAA GAAGGAGGAC TTTGACATCT
ACGACGAGGA
6421 CGAGAACCAG AGCCCCAGGA GCTTCCAGAA GAAGACCAGG CACTACTTCA
TTGCTGCTGT
6481 GGAGAGGCTG TGGGACTATG GCATGAGCAG CAGCCCCCAT GTGCTGAGGA
ACAGGGCCCA
6541 GTCTGGCTCT GTGCCCCAGT TCAAGAAGGT GGTGTTCCAG GAGTTCACTG
ATGGCAGCTT
6601 CACCCAGCCC CTGTACAGAG GGGAGCTGAA TGAGCACCTG GGCCTGCTGG
GCCCCTACAT
6661 CAGGGCTGAG GTGGAGGACA ACATCATGGT GACCTTCAGG AACCAGGCCA
GCAGGCCCTA
6721 CAGCTTCTAC AGCAGCCTGA TCAGCTATGA GGAGGACCAG AGGCAGGGGG
CTGAGCCCAG
6781 GAAGAACTTT GTGAAGCCCA ATGAAACCAA GACCTACTTC TGGAAGGTGC
AGCACCACAT
6841 GGCCCCCACC AAGGATGAGT TTGACTGCAA GGCCTGGGCC TACTTCTCTG
ATGTGGACCT
6901 GGAGAAGGAT GTGCACTCTG GCCTGATTGG CCCCCTGCTG GTGTGCCACA
CCAACACCCT
6961 GAACCCTGCC CATGGCAGGC AGGTGACTGT GCAGGAGTTT GCCCTGTTCT
TCACCATCTT
7021 TGATGAAACC AAGAGCTGGT ACTTCACTGA GAACATGGAG AGGAACTGCA
GGGCCCCCTG
7081 CAACATCCAG ATGGAGGACC CCACCTTCAA GGAGAACTAC AGGTTCCATG
CCATCAATGG
7141 CTACATCATG GACACCCTGC CTGGCCTGGT GATGGCCCAG GACCAGAGGA
TCAGGTGGTA
7201 CCTGCTGAGC ATGGGCAGCA ATGAGAACAT CCACAGCATC CACTTCTCTG
GCCATGTGTT
7261 CACTGTGAGG AAGAAGGAGG AGTACAAGAT GGCCCTGTAC AACCTGTACC
CTGGGGTGTT
7321 TGAGACTGTG GAGATGCTGC CCAGCAAGGC TGGCATCTGG AGGGTGGAGT
GCCTGATTGG
7381 GGAGCACCTG CATGCTGGCA TGAGCACCCT GTTCCTGGTG TACAGCAACA
AGTGCCAGAC
7441 CCCCCTGGGC ATGGCCTCTG GCCACATCAG GGACTTCCAG ATCACTGCCT
CTGGCCAGTA
7501 TGGCCAGTGG GCCCCCAAGC TGGCCAGGCT GCACTACTCT GGCAGCATCA
ATGCCTGGAG
7561 CACCAAGGAG CCCTTCAGCT GGATCAAGGT GGACCTGCTG GCCCCCATGA
TCATCCATGG
7621 CATCAAGACC CAGGGGGCCA GGCAGAAGTT CAGCAGCCTG TACATCAGCC
AGTTCATCAT
7681 CATGTACAGC CTGGATGGCA AGAAGTGGCA GACCTACAGG GGCAACAGCA
CTGGCACCCT

7741 GATGGTGTTC TTTGGCAATG TGGACAGCTC TGGCATCAAG CACAACATCT
TCAACCCCCC
7801 CATCATTGCC AGATACATCA GGCTGCACCC CACCCACTAC AGCATCAGGA
GCACCCTGAG
7861 GATGGAGCTG ATGGGCTGTG ACCTGAACAG CTGCAGCATG CCCCTGGGCA
TGGAGAGCAA
7921 GGCCATCTCT GATGCCCAGA TCACTGCCAG CAGCTACTTC ACCAACATGT
TTGCCACCTG
7981 GAGCCCCAGC AAGGCCAGGC TGCACCTGCA GGGCAGGAGC AATGCCTGGA
GGCCCCAGGT
8041 CAACAACCCC AAGGAGTGGC TGCAGGTGGA CTTCCAGAAG ACCATGAAGG
TGACTGGGGT
8101 GACCACCCAG GGGGTGAAGA GCCTGCTGAC CAGCATGTAT GTGAAGGAGT
TCCTGATCAG
8161 CAGCAGCCAG GATGGCCACC AGTGGACCCT GTTCTTCCAG AATGGCAAGG
TGAAGGTGTT
8221 CCAGGGCAAC CAGGACAGCT TCACCCCTGT GGTGAACAGC CTGGACCCCC
CCCTGCTGAC
8281 CAGATACCTG AGGATTCACC CCCAGAGCTG GGTGCACCAG ATTGCCCTGA
GGATGGAGGT
8341 GCTGGGCTGT GAGGCCCAGG ACCTGTACTG AGCGGCCGCG GGCCCAATCA
ACCTCTGGAT
8401 TACAAAATTT GTGAAAGATT GACTGGTATT CTTAACTATG TTGCTCCTTT
TACGCTATGT
8461 GGATACGCTG CTTTAATGCC TTTGTATCAT GCTATTGCTT CCCGTATGGC
TTTCATTTTC
8521 TCCTCCTTGT ATAAATCCTG GTTGCTGTCT CTTTATGAGG AGTTGTGGCC
CGTTGTCAGG
8581 CAACGTGGCG TGGTGTGCAC TGTGTTTGCT GACGCAACCC CCACTGGTTG
GGGCATTGCC
8641 ACCACCTGTC AGCTCCTTTC CGGGACTTTC GCTTTCCCCC TCCCTATTGC
CACGGCGGAA
8701 CTCATCGCCG CCTGCCTTGC CCGCTGCTGG ACAGGGGCTC GGCTGTTGGG
CACTGACAAT
8761 TCCGTGGTGT TGTCGGGGAA ATCATCGTCC TTTCCTTGGC TGCTCGCCTG
TGTTGCCACC
8821 TGGATTCTGC GCGGGACGTC CTTCTGCTAC GTCCCTTCGG CCCTCAATCC
AGCGGACCTT
8881 CCTTCCCGCG GCCTGCTGCC GGCTCTGCGG CCTCTTCCGC GTCTTCGCCT
TCGCCCTCAG
8941 ACGAGTCGGA TCTCCCTTTG GGCCGCCTCC CCGCAAGCTT CGCACTTTTT
AAAAGAAAAG
9001 GGAGGACTGG ATGGGATTTA TTACTCCGAT AGGACGCTGG CTTGTAACTC
AGTCTCTTAC
9061 TAGGAGACCA GCTTGAGCCT GGGTGTTCGC TGGTTAGCCT AACCTGGTTG
GCCACCAGGG
9121 GTAAGGACTC CTTGGCTTAG AAAGCTAATA AACTTGCCTG CATTAGAGCT
CTTACGCGTC
9181 CCGGGCTCGA GATCCGCATC TCAATTAGTC AGCAACCATA GTCCCGCCCC
TAACTCCGCC
9241 CATCCCGCCC CTAACTCCGC CCAGTTCCGC CCATTCTCCG CCCCATGGCT
GACTAATTTT
9301 TTTTATTTAT GCAGAGGCCG AGGCCGCCTC GGCCTCTGAG CTATTCCAGA
AGTAGTGAGG
9361 AGGCTTTTTT GGAGGCCTAG GCTTTTGCAA AAAGCTAACT TGTTTATTGC
AGCTTATAAT
9421 GGTTACAAAT AAAGCAATAG CATCACAAAT TTCACAAATA AAGCATTTTT
TTCACTGCAT

9481 TCTAGTTGTG GTTTGTCCAA ACTCATCAAT GTATCTTATC ATGTCTGTCC
GCTTCCTCGC
9541 TCACTGACTC GCTGCGCTCG GTCGTTCGGC TGCGGCGAGC GGTATCAGCT
CACTCAAAGG
9601 CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGAACATG
TGAGCAAAAG
9661 GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GGCGTTTTTC
CATAGGCTCC
9721 GCCCCCCTGA CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA
AACCCGACAG
9781 GACTATAAAG ATACCAGGCG TTTCCCCCTG GAAGCTCCCT CGTGCGCTCT
CCTGTTCCGA
9841 CCCTGCCGCT TACCGGATAC CTGTCCGCCT TTCTCCCTTC GGGAAGCGTG
GCGCTTTCTC
9901 ATAGCTCACG CTGTAGGTAT CTCAGTTCGG TGTAGGTCGT TCGCTCCAAG
CTGGGCTGTG
9961 TGCACGAACC CCCCGTTCAG CCCGACCGCT GCGCCTTATC CGGTAACTAT
CGTCTTGAGT
10021 CCAACCCGGT AAGACACGAC TTATCGCCAC TGGCAGCAGC CACTGGTAAC
AGGATTAGCA
10081 GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC
TACGGCTACA
10141 CTAGAAGAAC AGTATTTGGT ATCTGCGCTC TGCTGAAGCC AGTTACCTTC
GGAAAAAGAG
10201 TTGGTAGCTC TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTTT
TTTGTTTGCA
10261 AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA TCCTTTGATC
TTTTCTACGG
10321 GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG
AGATTATCAA
10381 AAAGGATCTT CACCTAGATC CTTTTAAATT AAAAATGAAG TTTTAAATCA
ATCTAAAGTA
10441 TATATGAGTA AACTTGGTCT GACAGTTAGA AAAACTCATC GAGCATCAAA
TGAAACTGCA
10501 ATTTATTCAT ATCAGGATTA TCAATACCAT ATTTTTGAAA AAGCCGTTTC
TGTAATGAAG
10561 GAGAAAACTC ACCGAGGCAG TTCCATAGGA TGGCAAGATC CTGGTATCGG
TCTGCGATTC
10621 CGACTCGTCC AACATCAATA CAACCTATTA ATTTCCCCTC GTCAAAAATA
AGGTTATCAA
10681 GTGAGAAATC ACCATGAGTG ACGACTGAAT CCGGTGAGAA TGGCAACAGC
TTATGCATTT
10741 CTTTCCAGAC TTGTTCAACA GGCCAGCCAT TACGCTCGTC ATCAAAATCA
CTCGCATCAA
10801 CCAAACCGTT ATTCATTCGT GATTGCGCCT GAGCGAGACG AAATACGCGA
TCGCTGTTAA
10861 AAGGACAATT ACAAACAGGA ATCGAATGCA ACCGGCGCAG GAACACTGCC
AGCGCATCAA
10921 CAATATTTTC ACCTGAATCA GGATATTCTT CTAATACCTG GAATGCTGTT
TTTCCGGGGA
10981 TCGCAGTGGT GAGTAACCAT GCATCATCAG GAGTACGGAT AAAATGCTTG
ATGGTCGGAA
11041 GAGGCATAAA TTCCGTCAGC CAGTTTAGTC TGACCATCTC ATCTGTAACA
TCATTGGCAA
11101 CGCTACCTTT GCCATGTTTC AGAAACAACT CTGGCGCATC GGGCTTCCCA
TACAATCGAT
11161 AGATTGTCGC ACCTGATTGC CCGACATTAT CGCGAGCCCA TTTATACCCA
TATAAATCAG

11221 CATCCATGTT GGAATTTAAT CGCGGCCTAG AGCAAGACGT TTCCCGTTGA
ATATGGCTCA
11281 TAACACCCCT TGTATTACTG TTTATGTAAG CAGACAGTTT TATTGTTCAT
GATGATATAT
11341 TTTTATCTTG TGCAATGTAA CATCAGAGAT TTTGAGACAC AACAATTGGT
CGACGGATCC

SEQ ID NO: 14

```
         1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT
        61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
       121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
       181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
       241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
TGACGTATGT
       301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
       361 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT
       421 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAC
GGGACTTTCC
       481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA
       541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT
       601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA
       661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGCGTGTACG
GTGGGAGGTC
       721 TATATAAGCA GAGCTCGCTG GCTTGTAACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC
       781 TGGGTGTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA
       841 GAAAGCTAAT AAACTTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCCT
CATTGACGCC
       901 TCACTCTCTT GAACGGGAAT CTTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA
       961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA
      1021 GGCGTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA
      1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC
      1141 CGTAACTACT CTTGGGCAAG TAGGGCAGGC GGTGGGTACG CAATGGGGGC
GGCTACCTCA
      1201 GCACTAAATA GGAGACAATT AGACCAATTT GAGAAAATAC GACTTCGCCC
GAACGGAAAG
      1261 AAAAAGTACC AAATTAAACA TTTAATATGG GCAGGCAAGG AGATGGAGCG
CTTCGGCCTC
      1321 CATGAGAGGT TGTTGGAGAC AGAGGAGGGG TGTAAAAGAA TCATAGAAGT
CCTCTACCCC
      1381 CTAGAACCAA CAGGATCGGA GGGCTTAAAA AGTCTGTTCA ATCTTGTGTG
CGTGCTATAT
```

1441 TGCTTGCACA AGGAACAGAA AGTGAAAGAC ACAGAGGAAG CAGTAGCAAC
AGTAAGACAA
1501 CACTGCCATC TAGTGGAAAA AGAAAAAAGT GCAACAGAGA CATCTAGTGG
ACAAAAGAAA
1561 AATGACAAGG GAATAGCAGC GCCACCTGGT GGCAGTCAGA ATTTTCCAGC
GCAACAACAA
1621 GGAAATGCCT GGGTACATGT ACCCTTGTCA CCGCGCACCT TAAATGCGTG
GGTAAAAGCA
1681 GTAGAGGAGA AAAAATTTGG AGCAGAAATA GTACCCATGT TTCAAGCCCT
ATCGAATTCC
1741 CGTTTGTGCT AGGGTTCTTA GGCTTCTTGG GGGCTGCTGG AACTGCAATG
GGAGCAGCGG
1801 CGACAGCCCT GACGGTCCAG TCTCAGCATT TGCTTGCTGG GATACTGCAG
CAGCAGAAGA
1861 ATCTGCTGGC GGCTGTGGAG GCTCAACAGC AGATGTTGAA GCTGACCATT
TGGGGTGTTA
1921 AAAACCTCAA TGCCCGCGTC ACAGCCCTTG AGAAGTACCT AGAGGATCAG
GCACGACTAA
1981 ACTCCTGGGG GTGCGCATGG AAACAAGTAT GTCATACCAC AGTGGAGTGG
CCCTGGACAA
2041 ATCGGACTCC GGATTGGCAA AATATGACTT GGTTGGAGTG GGAAAGACAA
ATAGCTGATT
2101 TGGAAAGCAA CATTACGAGA CAATTAGTGA AGGCTAGAGA ACAAGAGGAA
AAGAATCTAG
2161 ATGCCTATCA GAAGTTAACT AGTTGGTCAG ATTTCTGGTC TTGGTTCGAT
TTCTCAAAAT
2221 GGCTTAACAT TTTAAAAATG GGATTTTTAG TAATAGTAGG AATAATAGGG
TTAAGATTAC
2281 TTTACACAGT ATATGGATGT ATAGTGAGGG TTAGGCAGGG ATATGTTCCT
CTATCTCCAC
2341 AGATCCATAT CCGCGGCAAT TTTAAAAGAA AGGGAGGAAT AGGGGGACAG
ACTTCAGCAG
2401 AGAGACTAAT TAATATAATA ACAACACAAT TAGAAATACA ACATTTACAA
ACCAAAATTC
2461 AAAAAATTTT AAATTTTAGA GCCGCGGAGA TCTGTTACAT AACTTATGGT
AAATGGCCTG
2521 CCTGGCTGAC TGCCCAATGA CCCCTGCCCA ATGATGTCAA TAATGATGTA
TGTTCCCATG
2581 TAATGCCAAT AGGGACTTTC CATTGATGTC AATGGGTGGA GTATTTATGG
TAACTGCCCA
2641 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTATGCCC CCTATTGATG
TCAATGATGG
2701 TAAATGGCCT GCCTGGCATT ATGCCCAGTA CATGACCTTA TGGGACTTTC
CTACTTGGCA
2761 GTACATCTAT GTATTAGTCA TTGCTATTAC CATGGGAATT CACTAGTGGA
GAAGAGCATG
2821 CTTGAGGGCT GAGTGCCCCT CAGTGGGCAG AGAGCACATG GCCCACAGTC
CCTGAGAAGT
2881 TGGGGGGAGG GGTGGGCAAT TGAACTGGTG CCTAGAGAAG GTGGGGCTTG
GGTAAACTGG
2941 GAAAGTGATG TGGTGTACTG GCTCCACCTT TTTCCCCAGG GTGGGGGAGA
ACCATATATA
3001 AGTGCAGTAG TCTCTGTGAA CATTCAAGCT TCTGCCTTCT CCCTCCTGTG
AGTTTGCTAG
3061 CCACCAATGC AGATTGAGCT GAGCACCTGC TTCTTCCTGT GCCTGCTGAG
GTTCTGCTTC
3121 TCTGCCACCA GGAGATACTA CCTGGGGGCT GTGGAGCTGA GCTGGGACTA
CATGCAGTCT

3181 GACCTGGGGG AGCTGCCTGT GGATGCCAGG TTCCCCCCCA GAGTGCCCAA
GAGCTTCCCC
3241 TTCAACACCT CTGTGGTGTA CAAGAAGACC CTGTTTGTGG AGTTCACTGA
CCACCTGTTC
3301 AACATTGCCA AGCCCAGGCC CCCCTGGATG GGCCTGCTGG GCCCCACCAT
CCAGGCTGAG
3361 GTGTATGACA CTGTGGTGAT CACCCTGAAG AACATGGCCA GCCACCCTGT
GAGCCTGCAT
3421 GCTGTGGGGG TGAGCTACTG GAAGGCCTCT GAGGGGGCTG AGTATGATGA
CCAGACCAGC
3481 CAGAGGGAGA AGGAGGATGA CAAGGTGTTC CCTGGGGGCA GCCACACCTA
TGTGTGGCAG
3541 GTGCTGAAGG AGAATGGCCC CATGGCCTCT GACCCCCTGT GCCTGACCTA
CAGCTACCTG
3601 AGCCATGTGG ACCTGGTGAA GGACCTGAAC TCTGGCCTGA TTGGGGCCCT
GCTGGTGTGC
3661 AGGGAGGGCA GCCTGGCCAA GGAGAAGACC CAGACCCTGC ACAAGTTCAT
CCTGCTGTTT
3721 GCTGTGTTTG ATGAGGGCAA GAGCTGGCAC TCTGAAACCA AGAACAGCCT
GATGCAGGAC
3781 AGGGATGCTG CCTCTGCCAG GGCCTGGCCC AAGATGCACA CTGTGAATGG
CTATGTGAAC
3841 AGGAGCCTGC CTGGCCTGAT TGGCTGCCAC AGGAAGTCTG TGTACTGGCA
TGTGATTGGC
3901 ATGGGCACCA CCCCTGAGGT GCACAGCATC TTCCTGGAGG GCCACACCTT
CCTGGTCAGG
3961 AACCACAGGC AGGCCAGCCT GGAGATCAGC CCCATCACCT TCCTGACTGC
CCAGACCCTG
4021 CTGATGGACC TGGGCCAGTT CCTGCTGTTC TGCCACATCA GCAGCCACCA
GCATGATGGC
4081 ATGGAGGCCT ATGTGAAGGT GGACAGCTGC CCTGAGGAGC CCCAGCTGAG
GATGAAGAAC
4141 AATGAGGAGG CTGAGGACTA TGATGATGAC CTGACTGACT CTGAGATGGA
TGTGGTGAGG
4201 TTTGATGATG ACAACAGCCC CAGCTTCATC CAGATCAGGT CTGTGGCCAA
GAAGCACCCC
4261 AAGACCTGGG TGCACTACAT TGCTGCTGAG GAGGAGGACT GGGACTATGC
CCCCCTGGTG
4321 CTGGCCCCTG ATGACAGGAG CTACAAGAGC CAGTACCTGA ACAATGGCCC
CCAGAGGATT
4381 GGCAGGAAGT ACAAGAAGGT CAGGTTCATG GCCTACACTG ATGAAACCTT
CAAGACCAGG
4441 GAGGCCATCC AGCATGAGTC TGGCATCCTG GGCCCCCTGC TGTATGGGGA
GGTGGGGGAC
4501 ACCCTGCTGA TCATCTTCAA GAACCAGGCC AGCAGGCCCT ACAACATCTA
CCCCCATGGC
4561 ATCACTGATG TGAGGCCCCT GTACAGCAGG AGGCTGCCCA AGGGGGTGAA
GCACCTGAAG
4621 GACTTCCCCA TCCTGCCTGG GGAGATCTTC AAGTACAAGT GGACTGTGAC
TGTGGAGGAT
4681 GGCCCCACCA AGTCTGACCC CAGGTGCCTG ACCAGATACT ACAGCAGCTT
TGTGAACATG
4741 GAGAGGGACC TGGCCTCTGG CCTGATTGGC CCCCTGCTGA TCTGCTACAA
GGAGTCTGTG
4801 GACCAGAGGG GCAACCAGAT CATGTCTGAC AAGAGGAATG TGATCCTGTT
CTCTGTGTTT
4861 GATGAGAACA GGAGCTGGTA CCTGACTGAG AACATCCAGA GGTTCCTGCC
CAACCCTGCT

4921 GGGGTGCAGC TGGAGGACCC TGAGTTCCAG GCCAGCAACA TCATGCACAG
CATCAATGGC
4981 TATGTGTTTG ACAGCCTGCA GCTGTCTGTG TGCCTGCATG AGGTGGCCTA
CTGGTACATC
5041 CTGAGCATTG GGGCCCAGAC TGACTTCCTG TCTGTGTTCT TCTCTGGCTA
CACCTTCAAG
5101 CACAAGATGG TGTATGAGGA CACCCTGACC CTGTTCCCCT TCTCTGGGGA
GACTGTGTTC
5161 ATGAGCATGG AGAACCCTGG CCTGTGGATT CTGGGCTGCC ACAACTCTGA
CTTCAGGAAC
5221 AGGGGCATGA CTGCCCTGCT GAAAGTCTCC AGCTGTGACA AGAACACTGG
GGACTACTAT
5281 GAGGACAGCT ATGAGGACAT CTCTGCCTAC CTGCTGAGCA AGAACAATGC
CATTGAGCCC
5341 AGGAGCTTCA GCCAGAACAG CAGGCACCCC AGCACCAGGC AGAAGCAGTT
CAATGCCACC
5401 ACCATCCCTG AGAATGACAT AGAGAAGACA GACCCATGGT TTGCCCACCG
GACCCCCATG
5461 CCCAAGATCC AGAATGTGAG CAGCTCTGAC CTGCTGATGC TGCTGAGGCA
GAGCCCCACC
5521 CCCCATGGCC TGAGCCTGTC TGACCTGCAG GAGGCCAAGT ATGAAACCTT
CTCTGATGAC
5581 CCCAGCCCTG GGGCCATTGA CAGCAACAAC AGCCTGTCTG AGATGACCCA
CTTCAGGCCC
5641 CAGCTGCACC ACTCTGGGGA CATGGTGTTC ACCCCTGAGT CTGGCCTGCA
GCTGAGGCTG
5701 AATGAGAAGC TGGGCACCAC TGCTGCCACT GAGCTGAAGA AGCTGGACTT
CAAAGTCTCC
5761 AGCACCAGCA ACAACCTGAT CAGCACCATC CCCTCTGACA ACCTGGCTGC
TGGCACTGAC
5821 AACACCAGCA GCCTGGGCCC CCCCAGCATG CCTGTGCACT ATGACAGCCA
GCTGGACACC
5881 ACCCTGTTTG GCAAGAAGAG CAGCCCCCTG ACTGAGTCTG GGGGCCCCCT
GAGCCTGTCT
5941 GAGGAGAACA ATGACAGCAA GCTGCTGGAG TCTGGCCTGA TGAACAGCCA
GGAGAGCAGC
6001 TGGGGCAAGA ATGTGAGCAG CAGGGAGATC ACCAGGACCA CCCTGCAGTC
TGACCAGGAG
6061 GAGATTGACT ATGATGACAC CATCTCTGTG GAGATGAAGA AGGAGGACTT
TGACATCTAC
6121 GACGAGGACG AGAACCAGAG CCCCAGGAGC TTCCAGAAGA AGACCAGGCA
CTACTTCATT
6181 GCTGCTGTGG AGAGGCTGTG GGACTATGGC ATGAGCAGCA GCCCCCATGT
GCTGAGGAAC
6241 AGGGCCCAGT CTGGCTCTGT GCCCCAGTTC AAGAAGGTGG TGTTCCAGGA
GTTCACTGAT
6301 GGCAGCTTCA CCCAGCCCCT GTACAGAGGG GAGCTGAATG AGCACCTGGG
CCTGCTGGGC
6361 CCCTACATCA GGGCTGAGGT GGAGGACAAC ATCATGGTGA CCTTCAGGAA
CCAGGCCAGC
6421 AGGCCCTACA GCTTCTACAG CAGCCTGATC AGCTATGAGG AGGACCAGAG
GCAGGGGGCT
6481 GAGCCCAGGA AGAACTTTGT GAAGCCCAAT GAAACCAAGA CCTACTTCTG
GAAGGTGCAG
6541 CACCACATGG CCCCCACCAA GGATGAGTTT GACTGCAAGG CCTGGGCCTA
CTTCTCTGAT
6601 GTGGACCTGG AGAAGGATGT GCACTCTGGC CTGATTGGCC CCCTGCTGGT
GTGCCACACC

6661 AACACCCTGA ACCCTGCCCA TGGCAGGCAG GTGACTGTGC AGGAGTTTGC
CCTGTTCTTC
6721 ACCATCTTTG ATGAAACCAA GAGCTGGTAC TTCACTGAGA ACATGGAGAG
GAACTGCAGG
6781 GCCCCCTGCA ACATCCAGAT GGAGGACCCC ACCTTCAAGG AGAACTACAG
GTTCCATGCC
6841 ATCAATGGCT ACATCATGGA CACCCTGCCT GGCCTGGTGA TGGCCCAGGA
CCAGAGGATC
6901 AGGTGGTACC TGCTGAGCAT GGGCAGCAAT GAGAACATCC ACAGCATCCA
CTTCTCTGGC
6961 CATGTGTTCA CTGTGAGGAA GAAGGAGGAG TACAAGATGG CCCTGTACAA
CCTGTACCCT
7021 GGGGTGTTTG AGACTGTGGA GATGCTGCCC AGCAAGGCTG GCATCTGGAG
GGTGGAGTGC
7081 CTGATTGGGG AGCACCTGCA TGCTGGCATG AGCACCCTGT TCCTGGTGTA
CAGCAACAAG
7141 TGCCAGACCC CCCTGGGCAT GGCCTCTGGC CACATCAGGG ACTTCCAGAT
CACTGCCTCT
7201 GGCCAGTATG GCCAGTGGGC CCCCAAGCTG GCCAGGCTGC ACTACTCTGG
CAGCATCAAT
7261 GCCTGGAGCA CCAAGGAGCC CTTCAGCTGG ATCAAGGTGG ACCTGCTGGC
CCCCATGATC
7321 ATCCATGGCA TCAAGACCCA GGGGGCCAGG CAGAAGTTCA GCAGCCTGTA
CATCAGCCAG
7381 TTCATCATCA TGTACAGCCT GGATGGCAAG AAGTGGCAGA CCTACAGGGG
CAACAGCACT
7441 GGCACCCTGA TGGTGTTCTT TGGCAATGTG GACAGCTCTG GCATCAAGCA
CAACATCTTC
7501 AACCCCCCCA TCATTGCCAG ATACATCAGG CTGCACCCCA CCCACTACAG
CATCAGGAGC
7561 ACCCTGAGGA TGGAGCTGAT GGGCTGTGAC CTGAACAGCT GCAGCATGCC
CCTGGGCATG
7621 GAGAGCAAGG CCATCTCTGA TGCCCAGATC ACTGCCAGCA GCTACTTCAC
CAACATGTTT
7681 GCCACCTGGA GCCCCAGCAA GGCCAGGCTG CACCTGCAGG GCAGGAGCAA
TGCCTGGAGG
7741 CCCCAGGTCA ACAACCCCAA GGAGTGGCTG CAGGTGGACT TCCAGAAGAC
CATGAAGGTG
7801 ACTGGGGTGA CCACCCAGGG GGTGAAGAGC CTGCTGACCA GCATGTATGT
GAAGGAGTTC
7861 CTGATCAGCA GCAGCCAGGA TGGCCACCAG TGGACCCTGT TCTTCCAGAA
TGGCAAGGTG
7921 AAGGTGTTCC AGGGCAACCA GGACAGCTTC ACCCCTGTGG TGAACAGCCT
GGACCCCCCC
7981 CTGCTGACCA GATACCTGAG GATTCACCCC CAGAGCTGGG TGCACCAGAT
TGCCCTGAGG
8041 ATGGAGGTGC TGGGCTGTGA GGCCCAGGAC CTGTACTGAG CGGCCGCGGG
CCCAATCAAC
8101 CTCTGGATTA CAAAATTTGT GAAAGATTGA CTGGTATTCT TAACTATGTT
GCTCCTTTTA
8161 CGCTATGTGG ATACGCTGCT TTAATGCCTT TGTATCATGC TATTGCTTCC
CGTATGGCTT
8221 TCATTTTCTC CTCCTTGTAT AAATCCTGGT TGCTGTCTCT TTATGAGGAG
TTGTGGCCCG
8281 TTGTCAGGCA ACGTGGCGTG GTGTGCACTG TGTTTGCTGA CGCAACCCCC
ACTGGTTGGG
8341 GCATTGCCAC CACCTGTCAG CTCCTTTCCG GGACTTTCGC TTTCCCCCTC
CCTATTGCCA

8401 CGGCGGAACT CATCGCCGCC TGCCTTGCCC GCTGCTGGAC AGGGGCTCGG
CTGTTGGGCA
8461 CTGACAATTC CGTGGTGTTG TCGGGGAAAT CATCGTCCTT TCCTTGGCTG
CTCGCCTGTG
8521 TTGCCACCTG GATTCTGCGC GGGACGTCCT TCTGCTACGT CCCTTCGGCC
CTCAATCCAG
8581 CGGACCTTCC TTCCCGCGGC CTGCTGCCGG CTCTGCGGCC TCTTCCGCGT
CTTCGCCTTC
8641 GCCCTCAGAC GAGTCGGATC TCCCTTTGGG CCGCCTCCCC GCAAGCTTCG
CACTTTTTAA
8701 AAGAAAAGGG AGGACTGGAT GGGATTTATT ACTCCGATAG GACGCTGGCT
TGTAACTCAG
8761 TCTCTTACTA GGAGACCAGC TTGAGCCTGG GTGTTCGCTG GTTAGCCTAA
CCTGGTTGGC
8821 CACCAGGGGT AAGGACTCCT TGGCTTAGAA AGCTAATAAA CTTGCCTGCA
TTAGAGCTCT
8881 TACGCGTCCC GGGCTCGAGA TCCGCATCTC AATTAGTCAG CAACCATAGT
CCCGCCCCTA
8941 ACTCCGCCCA TCCCGCCCCT AACTCCGCCC AGTTCCGCCC ATTCTCCGCC
CCATGGCTGA
9001 CTAATTTTTT TTATTTATGC AGAGGCCGAG GCCGCCTCGG CCTCTGAGCT
ATTCCAGAAG
9061 TAGTGAGGAG GCTTTTTTGG AGGCCTAGGC TTTTGCAAAA AGCTAACTTG
TTTATTGCAG
9121 CTTATAATGG TTACAAATAA AGCAATAGCA TCACAAATTT CACAAATAAA
GCATTTTTTT
9181 CACTGCATTC TAGTTGTGGT TTGTCCAAAC TCATCAATGT ATCTTATCAT
GTCTGTCCGC
9241 TTCCTCGCTC ACTGACTCGC TGCGCTCGGT CGTTCGGCTG CGGCGAGCGG
TATCAGCTCA
9301 CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA
AGAACATGTG
9361 AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG
CGTTTTTCCA
9421 TAGGCTCCGC CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA
GGTGGCGAAA
9481 CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG
TGCGCTCTCC
9541 TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG
GAAGCGTGGC
9601 GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTC
GCTCCAAGCT
9661 GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG
GTAACTATCG
9721 TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA
CTGGTAACAG
9781 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT
GGCCTAACTA
9841 CGGCTACACT AGAAGAACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG
TTACCTTCGG
9901 AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG
GTGGTTTTTT
9961 TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC
CTTTGATCTT
10021 TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT
TGGTCATGAG
10081 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT
TTAAATCAAT

10141 CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTAGAAA AACTCATCGA
GCATCAAATG
10201 AAACTGCAAT TTATTCATAT CAGGATTATC AATACCATAT TTTTGAAAAA
GCCGTTTCTG
10261 TAATGAAGGA GAAAACTCAC CGAGGCAGTT CCATAGGATG GCAAGATCCT
GGTATCGGTC
10321 TGCGATTCCG ACTCGTCCAA CATCAATACA ACCTATTAAT TTCCCCTCGT
CAAAAATAAG
10381 GTTATCAAGT GAGAAATCAC CATGAGTGAC GACTGAATCC GGTGAGAATG
GCAACAGCTT
10441 ATGCATTTCT TTCCAGACTT GTTCAACAGG CCAGCCATTA CGCTCGTCAT
CAAAATCACT
10501 CGCATCAACC AAACCGTTAT TCATTCGTGA TTGCGCCTGA GCGAGACGAA
ATACGCGATC
10561 GCTGTTAAAA GGACAATTAC AAACAGGAAT CGAATGCAAC CGGCGCAGGA
ACACTGCCAG
10621 CGCATCAACA ATATTTTCAC CTGAATCAGG ATATTCTTCT AATACCTGGA
ATGCTGTTTT
10681 TCCGGGGATC GCAGTGGTGA GTAACCATGC ATCATCAGGA GTACGGATAA
AATGCTTGAT
10741 GGTCGGAAGA GGCATAAATT CCGTCAGCCA GTTTAGTCTG ACCATCTCAT
CTGTAACATC
10801 ATTGGCAACG CTACCTTTGC CATGTTTCAG AAACAACTCT GGCGCATCGG
GCTTCCCATA
10861 CAATCGATAG ATTGTCGCAC CTGATTGCCC GACATTATCG CGAGCCCATT
TATACCCATA
10921 TAAATCAGCA TCCATGTTGG AATTTAATCG CGGCCTAGAG CAAGACGTTT
CCCGTTGAAT
10981 ATGGCTCATA ACACCCCTTG TATTACTGTT TATGTAAGCA GACAGTTTTA
TTGTTCATGA
11041 TGATATATTT TTATCTTGTG CAATGTAACA TCAGAGATTT TGAGACACAA
CAATTGGTCG
11101 ACGGATCC

SEQ ID NO: 15

ATGCCCAGCTCTGTGTCCTGGGGCATTCTGCTGCTGGCTGGCCTGTGCTGTCTGGTGCCTGTGTCCCTGG
CTGAGGACCCTCAGGGGGATGCTGCCCAGAAAACAGACACCTCCCACCATGACCAGGACCACCCCACCTT
CAACAAGATCACCCCCAACCTGGCAGAGTTTGCCTTCAGCCTGTACAGACAGCTGGCCCACCAGAGCAAC
AGCACCAACATCTTTTTCAGCCCTGTGTCCATTGCCACAGCCTTTGCCATGCTGAGCCTGGGCACCAAGG
CTGACACCCATGATGAGATCCTGGAAGGCCTGAACTTCAACCTGACAGAGATCCCTGAGGCCCAGATCCA
TGAGGGCTTCCAGGAACTGCTGAGAACCCTGAACCAGCCAGACAGCCAGCTGCAGCTGACAACAGGCAAT
GGGCTGTTCCTGTCTGAGGGCCTGAAGCTGGTGGACAAGTTTCTGGAAGATGTGAAGAAGCTGTACCACT
CTGAGGCCTTCACAGTGAACTTTGGGGACACAGAAGAGGCCAAGAAACAGATCAATGACTATGTGGAAAA
GGGCACCCAGGGCAAGATTGTGGACCTTGTGAAAGAGCTGGACAGGGACACTGTGTTTGCCCTTGTGAAC
TACATCTTCTTCAAGGGCAAGTGGGAGAGGCCCTTTGAAGTGAAGGACACTGAGGAAGAGGACTTCCATG
TGGACCAAGTGACCACAGTGAAGGTGCCAATGATGAAGAGACTGGGGATGTTCAATATCCAGCACTGCAA
GAAACTGAGCAGCTGGGTGCTGCTGATGAAGTACCTGGGCAATGCTACAGCCATATTCTTTCTGCCTGAT
GAGGGCAAGCTGCAGCACCTGGAAAATGAGCTGACCCATGACATCATCACCAAATTTCTGGAAAATGAGG
ACAGAAGATCTGCCAGCCTGCATCTGCCCAAGCTGAGCATCACAGGCACATATGACCTGAAGTCTGTGCT
GGGACAGCTGGGAATCACCAAGGTGTTCAGCAATGGGGCAGACCTGAGTGGAGTGACAGAGGAAGCCCCT
CTGAAGCTGTCCAAGGCTGTGCACAAGGCAGTGCTGACCATTGATGAGAAGGGCACAGAGGCTGCTGGGG
CCATGTTTCTGGAAGCCATCCCCATGTCCATCCCCCCAGAAGTGAAGTTCAACAAGCCCTTTGTGTTCCT
GATGATTGAGCAGAACACCAAGAGCCCCCTGTTCATGGGCAAGGTTGTGAACCCCACCCAGAAATGA

SEQ ID NO: 16

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGGAGAT
ACTACCTGGGGGGCTGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGGGGAGCTGCCTGTGGATGC
CAGGTTCCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCTGTGGTGTACAAGAAGACCCTGTTT
GTGGAGTTCACTGACCACCTGTTCAACATTGCCAAGCCCAGGCCCCCCTGGATGGGCCTGCTGGGCCCCA
CCATCCAGGCTGAGGTGTATGACACTGTGGTGATCACCCTGAAGAACATGGCCAGCCACCCTGTGAGCCT
GCATGCTGTGGGGGTGAGCTACTGGAAGGCCTCTGAGGGGGCTGAGTATGATGACCAGACCAGCCAGAGG
GAGAAGGAGGATGACAAGGTGTTCCCTGGGGGCAGCCACACCTATGTGTGGCAGGTGCTGAAGGAGAATG
GCCCCATGGCCTCTGACCCCCTGTGCCTGACCTACAGCTACCTGAGCCATGTGGACCTGGTGAAGGACCT
GAACTCTGGCCTGATTGGGGCCCTGCTGGTGTGCAGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACC
CTGCACAAGTTCATCCTGCTGTTTGCTGTGTTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACA
GCCTGATGCAGGACAGGGATGCTGCCTCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGT
GAACAGGAGCCTGCCTGGCCTGATTGGCTGCCACAGGAAGTCTGTGTACTGGCATGTGATTGGCATGGGC
ACCACCCCTGAGGTGCACAGCATCTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGGCAGGCCA
GCCTGGAGATCAGCCCCATCACCTTCCTGACTGCCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCT
GTTCTGCCACATCAGCAGCCACCAGCATGATGGCATGGAGGCCTATGTGAAGGTGGACAGCTGCCCTGAG
GAGCCCCAGCTGAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGA
TGGATGTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAGCA
CCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCCCTGGTGCTGGCC
CCTGATGACAGGAGCTACAAGAGCCAGTACCTGAACAATGGCCCCCAGAGGATTGGCAGGAAGTACAAGA
AGGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATCCAGCATGAGTCTGGCAT
CCTGGGCCCCCTGCTGTATGGGGAGGTGGGGGACACCCTGCTGATCATCTTCAAGAACCAGGCCAGCAGG
CCCTACAACATCTACCCCCATGGCATCACTGATGTGAGGCCCCCTGTACAGCAGGAGGCTGCCCAAGGGGG
TGAAGCACCTGAAGGACTTCCCCATCCTGCCTGGGGAGATCTTCAAGTACAAGTGGACTGTGACTGTGGA
GGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAGATACTACAGCAGCTTTGTGAACATGGAGAGG
GACCTGGCCTCTGGCCTGATTGGCCCCCTGCTGATCTGCTACAAGGAGTCTGTGGACCAGAGGGGCAACC
AGATCATGTCTGACAAGAGGAATGTGATCCTGTTCTCTGTGTTTGATGAGAACAGGAGCTGGTACCTGAC
TGAGAACATCCAGAGGTTCCTGCCCAACCCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGC
AACATCATGCACAGCATCAATGGCTATGTGTTTGACAGCCTGCAGCTGTCTGTGTGCCTGCATGAGGTGG
CCTACTGGTACATCCTGAGCATTGGGGCCCAGACTGACTTCCTGTCTGTGTTCTTCTCTGGCTACACCTT
CAAGCACAAGATGGTGTATGAGGACACCCTGACCCTGTTCCCCTTCTCTGGGGAGACTGTGTTCATGAGC
ATGGAGAACCCTGGCCTGTGGATTCTGGGCTGCCACAACTCTGACTTCAGGAACAGGGGCATGACTGCCC
TGCTGAAAGTCTCCAGCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGC
CTACCTGCTGAGCAAGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGAACAGCAGGCACCCCAGCACC
AGGCAGAAGCAGTTCAATGCCACCACCATCCCTGAGAATGACATAGAGAAGACAGACCCATGGTTTGCCC
ACCGGACCCCCATGCCCAAGATCCAGAATGTGAGCAGCTCTGACCTGCTGATGCTGCTGAGGCAGAGCCC
CACCCCCCATGGCCTGAGCCTGTCTGACCTGCAGGAGGCCAAGTATGAAACCTTCTCTGATGACCCCAGC
CCTGGGGCCATTGACAGCAACAACAGCCTGTCTGAGATGACCCACTTCAGGCCCCAGCTGCACCACTCTG
GGGACATGGTGTTCACCCCTGAGTCTGGCCTGCAGCTGAGGCTGAATGAGAAGCTGGGCACCACTGCTGC
CACTGAGCTGAAGAAGCTGGACTTCAAAGTCTCCAGCACCAGCAACAACCTGATCAGCACCATCCCCTCT
GACAACCTGGCTGCTGGCACTGACAACACCAGCAGCCTGGGCCCCCCCAGCATGCCTGTGCACTATGACA
GCCAGCTGGACACCACCCTGTTTGGCAAGAAGAGCAGCCCCCTGACTGAGTCTGGGGGCCCCCTGAGCCT
GTCTGAGGAGAACAATGACAGCAAGCTGCTGGAGTCTGGCCTGATGAACAGCCAGGAGAGCAGCTGGGGC
AAGAATGTGAGCAGCAGGGAGATCACCAGGACCACCCTGCAGTCTGACCAGGAGGAGATTGACTATGATG
ACACCATCTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTACGACGAGGACGAGAACCAGAGCCCCAG
GAGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGAGGCTGTGGGACTATGGCATGAGC
AGCAGCCCCCATGTGCTGAGGAACAGGGCCCAGTCTGGCTCTGTGCCCCAGTTCAAGAAGGTGGTGTTCC
AGGAGTTCACTGATGGCAGCTTCACCCAGCCCCTGTACAGAGGGGAGCTGAATGAGCACCTGGGCCTGCT
GGGCCCCTACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCAGGAACCAGGCCAGCAGGCCC
TACAGCTTCTACAGCAGCCTGATCAGCTATGAGGAGGACCAGAGGCAGGGGGCTGAGCCCAGGAAGAACT
TTGTGAAGCCCAATGAAACCAAGACCTACTTCTGGAAGGTGCAGCACCACATGGCCCCCACCAAGGATGA
GTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTGGAGAAGGATGTGCACTCTGGCCTGATT
GGCCCCCTGCTGGTGTGCCACACCAACACCCTGAACCCTGCCCATGGCAGGCAGGTGACTGTGCAGGAGT
TTGCCCTGTTCTTCACCATCTTTGATGAAACCAAGAGCTGGTACTTCACTGAGAACATGGAGAGGAACTG
CAGGGCCCCCTGCAACATCCAGATGGAGGACCCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAAT
GGCTACATCATGGACACCCTGCCTGGCCTGGTGATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTGA
GCATGGGCAGCAATGAGAACATCCACAGCATCCACTTCTCTGGCCATGTGTTCACTGTGAGGAAGAAGGA
GGAGTACAAGATGGCCCTGTACAACCTGTACCCTGGGGTGTTTGAGACTGTGGAGATGCTGCCCAGCAAG
GCTGGCATCTGGAGGGTGGAGTGCCTGATTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGG

TGTACAGCAACAAGTGCCAGACCCCCCTGGGCATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGC
CTCTGGCCAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACTCTGGCAGCATCAATGCCTGG
AGCACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAGA
CCCAGGGGGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGATGG
CAAGAAGTGGCAGACCTACAGGGGCAACAGCACTGGCACCCTGATGGTGTTCTTTGGCAATGTGGACAGC
TCTGGCATCAAGCACAACATCTTCAACCCCCCCATCATTGCCAGATACATCAGGCTGCACCCCACCCACT
ACAGCATCAGGAGCACCCTGAGGATGGAGCTGATGGGCTGTGACCTGAACAGCTGCAGCATGCCCCTGGG
CATGGAGAGCAAGGCCATCTCTGATGCCCAGATCACTGCCAGCAGCTACTTCACCAACATGTTTGCCACC
TGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGCAATGCCTGGAGGCCCCAGGTCAACAACC
CCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAGGTGACTGGGGTGACCACCCAGGGGGTGAA
GAGCCTGCTGACCAGCATGTATGTGAAGGAGTTCCTGATCAGCAGCAGCCAGGATGGCCACCAGTGGACC
CTGTTCTTCCAGAATGGCAAGGTGAAGGTGTTCCAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACA
GCCTGGACCCCCCCCTGCTGACCAGATACCTGAGGATTCACCCCCAGAGCTGGGTGCACCAGATTGCCCT
GAGGATGGAGGTGCTGGGCTGTGAGGCCCAGGACCTGTACTGA

SEQ ID NO: 17

CCGCGGAGATCTCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCT
ATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACC
GCCATGTTGGCATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCA
TATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCC
CATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGT
GGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATT
GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGG
CACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGC
GTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGC
GGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGC
AGAAGTTGGTCGTGAGGCACTGGGCAGGCTAGC

SEQ ID NO: 18
TCGAGATGTGGTCTGAGTTAAAAATCAGGAGCAACGACGGAGGTGAAGGACCAGACGCCAACGACCC
SEQ ID NO: 19
CCGGGGGTCGTTGGCGTCTGGTCCTTCACCTCCGTCGTTGCTCCTGATTTTTAACTCAGACCACATC
SEQ ID NO: 20
CCGGGGAAAGGGGGTGCAACACATCCATATCCAGCCATCTCTACCTGTTTATGGACA
SEQ ID NO: 21
ACCCTCTGTCCATAAACAGGTAGAGATGGCTGGATATGGATGTGTTGCACCCCTTTCC
SEQ ID NO: 22
GGGTTAGGTGGTTGCTGATTCTCTCATTCACCCAGTGGG
SEQ ID NO: 23
GATCCCCACTGGGTGAATGAGAGAATCAGCAACCACCTA
SEQ ID NO: 24
GAGACTCGAGATGTGGTCTGAGTTAAAAATCAGG
SEQ ID NO: 25
AGAGGTAGACCAGTACGAGTCACGTTTGCCCCTATCACCATCCCTAACCCTCTGTCATAAAC

SEQ ID NO: 26

TACGGGTCGAGACACAGGACCCCGTAAGACGACGACCGACCGGACACGACAGACCACGGACACAGGGACC
GACTCCTGGGAGTCCCCCTACGACGGGTCTTTTGTCTGTGGAGGGTGGTACTGGTCCTGGTGGGGTGGAA
GTTGTTCTAGTGGGGGTTGGACCGTCTCAAACGGAAGTCGGACATGTCTGTCGACCGGGTGGTCTCGTTG
TCGTGGTTGTAGAAAAAGTCGGGACACAGGTAACGGTGTCGGAAACGGTACGACTCGGACCCGTGGTTCC
GACTGTGGGTACTACTCTAGGACCTTCCGGACTTGAAGTTGGACTGTCTCTAGGGACTCCGGGTCTAGGT
ACTCCCGAAGGTCCTTGACGACTCTTGGGACTTGGTCGGTCTGTCGGTCGACGTCGACTGTTGTCCGTTA
CCCGACAAGGACAGACTCCCGGACTTCGACCACCTGTTCAAAGACCTTCTACACTTCTTCGACATGGTGA
GACTCCGGAAGTGTCACTTGAAACCCCTGTGTCTTCTCCGGTTCTTTGTCTAGTTACTGATACACCTTTT
CCCGTGGGTCCCGTTCTAACACCTGGAACACTTTCTCGACCTGTCCCTGTGACACAAACGGGAACACTTG
ATGTAGAAGAAGTTCCCGTTCACCCTCTCCGGGAAACTTCACTTCCTGTGACTCCTTCTCCTGAAGGTAC
ACCTGGTTCACTGGTGTCACTTCCACGGTTACTACTTCTCTGACCCCTACAAGTTATAGGTCGTGACGTT
CTTTGACTCGTCGACCCACGACGACTACTTCATGGACCCGTTACGATGTCGGTATAAGAAAGACGGACTA
CTCCCGTTCGACGTCGTGGACCTTTTACTCGACTGGGTACTGTAGTAGTGGTTTAAAGACCTTTTACTCC
TGTCTTCTAGACGGTCGGACGTAGACGGGTTCGACTCGTAGTGTCCGTGTATACTGGACTTCAGACACGA
CCCTGTCGACCCTTAGTGGTTCCACAAGTCGTTACCCCGTCTGGACTCACCTCACTGTCTCCTTCGGGGA
GACTTCGACAGGTTCCGACACGTGTTCCGTCACGACTGGTAACTACTCTTCCCGTGTCTCCGACGACCCC
GGTACAAAGACCTTCGGTAGGGGTACAGGTAGGGGGGTCTTCACTTCAAGTTGTTCGGGAAACACAAGGA
CTACTAACTCGTCTTGTGGTTCTCGGGGGACAAGTACCCGTTCCAACACTTGGGGTGGGTCTTTACT

SEQ ID NO: 27

AEDPQGDAAQKTDTSHHDQDHPTFAEDPQGDAAQKTDTSHHDQDHPTFNKITPNLAEFAFSLYRQLAHQSN
STNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGNG
LFLSEGLKLVDKFLEDVKKLYHSEAFTVNFGDTEEAKKQINDYVEKGTQGKIVDLVKELDRDTVFALVNYI
FFKGKWERPFEVKDTEEEDFHVDQVTTVKVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIFFLPDEGK
LQHLENELTHDIITKFLENEDRRSASLHLPKLSITGTYDLKSVLGQLGITKVFSNGADLSGVTEEAPLKLS
KAVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK

SEQ ID NO: 28

TACGTCTAACTCGACTCGTGGACGAAGAAGGACACGGACGACTCCAAGACGAAGAGACGGTGGTCCTCTA
TGATGGACCCCCGACACCTCGACTCGACCCTGATGTACGTCAGACTGGACCCCCTCGACGGACACCTACG
GTCCAAGGGGGGGTCTCACGGGTTCTCGAAGGGGAAGTTGTGGAGACACCACATGTTCTTCTGGGACAAA
CACCTCAAGTGACTGGTGGACAAGTTGTAACGGTTCGGGTCCGGGGGGACCTACCCGGACGACCCGGGGT
GGTAGGTCCGACTCCACATACTGTGACACCACTAGTGGGACTTCTTGTACCGGTCGGTGGGACACTCGGA
CGTACGACACCCCCACTCGATGACCTTCCGGAGACTCCCCCGACTCATACTACTGGTCTGGTCGGTCTCC
CTCTTCCTCCTACTGTTCCACAAGGGACCCCCGTCGGTGTGGATACACACCGTCCACGACTTCCTCTTAC
CGGGGTACCGGAGACTGGGGGACACGGACTGGATGTCGATGGACTCGGTACACCTGGACCACTTCCTGGA
CTTGAGACCGGACTAACCCCGGGACGACCACACGTCCCTCCCGTCGGACCGGTTCCTCTTCTGGGTCTGG
GACGTGTTCAAGTAGGACGACAAACGACACAAACTACTCCCGTTCTCGACCGTGAGACTTTGGTTCTTGT
CGGACTACGTCCTGTCCCTACGACGGAGACGGTCCCGGACCGGGTTCTACGTGTGACACTTACCGATACA
CTTGTCCTCGGACGGACCGGACTAACCGACGGTGTCCTTCAGACACATGACCGTACACTAACCGTACCCG
TGGTGGGGACTCCACGTGTCGTAGAAGGACCTCCCGGTGTGGAAGGACCAGTCCTTGGTGTCCGTCCGGT
CGGACCTCTAGTCGGGGTAGTGGAAGGACTGACGGGTCTGGGACGACTACCTGGACCCGGTCAAGGACGA
CAAGACGGTGTAGTCGTCGGTGGTCGTACTACCGTACCTCCGGATACACTTCCACCTGTCGACGGGACTC
CTCGGGGTCGACTCCTACTTCTTGTTACTCCTCCGACTCCTGATACTACTACTGGACTGACTGAGACTCT
ACCTACACCACTCCAAACTACTACTGTTGTCGGGGTCGAAGTAGGTCTAGTCCAGACACCGGTTCTTCGT

GGGGTTCTGGACCCACGTGATGTAACGACGACTCCTCCTCCTGACCCTGATACGGGGGGACCACGACCGG
GGACTACTGTCCTCGATGTTCTCGGTCATGGACTTGTTACCGGGGGTCTCCTAACCGTCCTTCATGTTCT
TCCAGTCCAAGTACCGGATGTGACTACTTTGGAAGTTCTGGTCCCTCCGGTAGGTCGTACTCAGACCGTA
GGACCCGGGGGACGACATACCCCTCCACCCCCTGTGGGACGACTAGTAGAAGTTCTTGGTCCGGTCGTCC
GGGATGTTGTAGATGGGGGTACCGTAGTGACTACACTCCGGGGACATGTCGTCCTCCGACGGGTTCCCCC
ACTTCGTGGACTTCCTGAAGGGGTAGGACGGACCCCTCTAGAAGTTCATGTTCACCTGACACTGACACCT
CCTACCGGGGTGGTTCAGACTGGGGTCCACGGACTGGTCTATGATGTCGTCGAAACACTTGTACCTCTCC
CTGGACCGGAGACCGGACTAACCGGGGGACGACTAGACGATGTTCCTCAGACACCTGGTCTCCCCGTTGG
TCTAGTACAGACTGTTCTCCTTACACTAGGACAAGAGACACAAACTACTCTTGTCCTCGACCATGGACTG
ACTCTTGTAGGTCTCCAAGGACGGGTTGGGACGACCCCACGTCGACCTCCTGGGACTCAAGGTCCGGTCG
TTGTAGTACGTGTCGTAGTTACCGATACACAAACTGTCGGACGTCGACAGACACACGGACGTACTCCACC
GGATGACCATGTAGGACTCGTAACCCCGGGTCTGACTGAAGGACAGACACAAGAAGAGACCGATGTGGAA
GTTCGTGTTCTACCACATACTCCTGTGGGACTGGGACAAGGGGAAGAGACCCCTCTGACACAAGTACTCG
TACCTCTTGGGACCGGACACCTAAGACCCGACGGTGTTGAGACTGAAGTCCTTGTCCCCGTACTGACGGG
ACGACTTTCAGAGGTCGACACTGTTCTTGTGACCCCTGATGATACTCCTGTCGATACTCCTGTAGAGACG
GATGGACGACTCGTTCTTGTTACGGTAACTCGGGTCCTCGAAGTCGGTCTTGTCGTCCGTGGGGTCGTGG
TCCGTCTTCGTCAAGTTACGGTGGTGGTAGGGACTCTTACTGTATCTCTTCTGTCTGGGTACCAAACGGG
TGGCCTGGGGGTACGGGTTCTAGGTCTTACACTCGTCGAGACTGGACGACTACGACGACTCCGTCTCGGG
GTGGGGGGTACCGGACTCGGACAGACTGGACGTCCTCCGGTTCATACTTTGGAAGAGACTACTGGGGTCG
GGACCCCGGTAACTGTCGTTGTTGTCGGACAGACTCTACTGGGTGAAGTCCGGGGTCGACGTGGTGAGAC
CCCTGTACCACAAGTGGGGACTCAGACCGGACGTCGACTCCGACTTACTCTTCGACCCGTGGTGACGACG
GTGACTCGACTTCTTCGACCTGAAGTTTCAGAGGTCGTGGTCGTTGTTGGACTAGTCGTGGTAGGGGAGA
CTGTTGGACCGACGACCGTGACTGTTGTGGTCGTCGGACCCGGGGGGGTCGTACGGACACGTGATACTGT
CGGTCGACCTGTGGTGGGACAAACCGTTCTTCTCGTCGGGGGACTGACTCAGACCCCCGGGGGACTCGGA
CAGACTCCTCTTGTTACTGTCGTTCGACGACCTCAGACCGGACTACTTGTCGGTCCTCTCGTCGACCCCG
TTCTTACACTCGTCGTCCCTCTAGTGGTCCTGGTGGGACGTCAGACTGGTCCTCCTCTAACTGATACTAC
TGTGGTAGAGACACCTCTACTTCTTCCTCCTGAAACTGTAGATGCTGCTCCTGCTCTTGGTCTCGGGGTC
CTCGAAGGTCTTCTTCTGGTCCGTGATGAAGTAACGACGACACCTCTCCGACACCCTGATACCGTACTCG
TCGTCGGGGGTACACGACTCCTTGTCCCGGGTCAGACCGAGACACGGGGTCAAGTTCTTCCACCACAAGG
TCCTCAAGTGACTACCGTCGAAGTGGGTCGGGGACATGTCTCCCCTCGACTTACTCGTGGACCCGGACGA
CCCGGGGATGTAGTCCCGACTCCACCTCCTGTTGTAGTACCACTGGAAGTCCTTGGTCCGGTCGTCCGGG
ATGTCGAAGATGTCGTCGGACTAGTCGATACTCCTCCTGGTCTCCGTCCCCCGACTCGGGTCCTTCTTGA
AACACTTCGGGTTACTTTGGTTCTGGATGAAGACCTTCCACGTCGTGGTGTACCGGGGGTGGTTCCTACT
CAAACTGACGTTCCGGACCCGGATGAAGAGACTACACCTGGACCTCTTCCTACACGTGAGACCGGACTAA
CCGGGGGACGACCACACGGTGTGGTTGTGGGACTTGGGACGGGTACCGTCCGTCCACTGACACGTCCTCA
AACGGGACAAGAAGTGGTAGAAACTACTTTGGTTCTCGACCATGAAGTGACTCTTGTACCTCTCCTTGAC
GTCCCGGGGGACGTTGTAGGTCTACCTCCTGGGGTGGAAGTTCCTCTTGATGTCCAAGGTACGGTAGTTA
CCGATGTAGTACCTGTGGGACGGACCGGACCACTACCGGGTCCTGGTCTCCTAGTCCACCATGGACGACT
CGTACCCGTCGTTACTCTTGTAGGTGTCGTAGGTGAAGAGACCGGTACACAAGTGACACTCCTTCTTCCT
CCTCATGTTCTACCGGGACATGTTGGACATGGGACCCCACAAACTCTGACACCTCTACGACGGGTCGTTC
CGACCGTAGACCTCCCACCTCACGGACTAACCCCTCGTGGACGTACGACCGTACTCGTGGGACAAGGACC
ACATGTCGTTGTTCACGGTCTGGGGGGACCCGTACCGGAGACCGGTGTAGTCCCTGAAGGTCTAGTGACG
GAGACCGGTCATACCGGTCACCCGGGGGTTCGACCGGTCCGACGTGATGAGACCGTCGTAGTTACGGACC
TCGTGGTTCCTCGGGAAGTCGACCTAGTTCCACCTGGACGACCGGGGGTACTAGTAGGTACCGTAGTTCT
GGGTCCCCCGGTCCGTCTTCAAGTCGTCGGACATGTAGTCGGTCAAGTAGTAGTACATGTCGGACCTACC
GTTCTTCACCGTCTGGATGTCCCCGTTGTCGTGACCGTGGGACTACCACAAGAAACCGTTACACCTGTCG
AGACCGTAGTTCGTGTTGTAGAAGTTGGGGGGGTAGTAACGGTCTATGTAGTCCGACGTGGGGTGGGTGA
TGTCGTAGTCCTCGTGGGACTCCTACCTCGACTACCCGACACTGGACTTGTCGACGTCGTACGGGGACCC
GTACCTCTCGTTCCGGTAGAGACTACGGGTCTAGTGACGGTCGTCGATGAAGTGGTTGTACAAACGGTGG
ACCTCGGGGTCGTTCCGGTCCGACGTGGACGTCCCGTCCTCGTTACGGACCTCCGGGGTCCAGTTGTTGG
GGTTCCTCACCGACGTCCACCTGAAGGTCTTCTGGTACTTCCACTGACCCCACTGGTGGGTCCCCCACTT
CTCGGACGACTGGTCGTACATACACTTCCTCAAGGACTAGTCGTCGTCGGTCCTACCGGTGGTCACCTGG
GACAAGAAGGTCTTACCGTTCCACTTCCACAAGGTCCCGTTGGTCCTGTCGAAGTGGGGACACCACTTGT
CGGACCTGGGGGGGGACGACTGGTCTATGGACTCCTAAGTGGGGGTCTCGACCCACGTGGTCTAACGGGA
CTCCTACCTCCACGACCCGACACTCCGGGTCCTGGACATGACT

SEQ ID NO: 29

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFV
EFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREK
EDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHK
FILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPE
VHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLR
MKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSY
KSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPH
GITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIG
PLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGY
VFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILG
CHNSDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNSRHPSTRQKQFNATTIP
ENDIEKTDPWFAHRTPMPKIQNVSSSDLLMLLRQSPTPHGLSLSDLQEAKYETFSDDPSPGAIDSNNSLSE
MTHFRPQLHHSGDMVFTPESGLQLRLNEKLGTTAATELKKLDFKVSSTSNNLISTIPSDNLAAGTDNTSSL
GPPSMPVHYDSQLDTTLFGKKSSPLTESGGPLSLSEENNDSKLLESGLMNSQESSWGKNVSSREITRTTLQ
SDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSV
PQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQ
GAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGR
QVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRI
RWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMS
TLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIH
GIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHP
THYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVN
NPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVN
SLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

SEQ ID NO: 30

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGGAGAT
ACTACCTGGGGGGCTGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGGGGAGCTGCCTGTGGATGC
CAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCTGTGGTGTACAAGAAGACCCTGTTT
GTGGAGTTCACTGACCACCTGTTCAACATTGCCAAGCCCAGGCCCCCCTGGATGGGCCTGCTGGGCCCCA
CCATCCAGGCTGAGGTGTATGACACTGTGGTGATCACCCTGAAGAACATGGCCAGCCACCCTGTGAGCCT
GCATGCTGTGGGGGTGAGCTACTGGAAGGCCTCTGAGGGGGCTGAGTATGATGACCAGACCAGCCAGAGG
GAGAAGGAGGATGACAAGGTGTTCCCTGGGGGCAGCCACACCTATGTGTGGCAGGTGCTGAAGGAGAATG
GCCCCATGGCCTCTGACCCCCTGTGCCTGACCTACAGCTACCTGAGCCATGTGGACCTGGTGAAGGACCT
GAACTCTGGCCTGATTGGGGCCCTGCTGGTGTGCAGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACC
CTGCACAAGTTCATCCTGCTGTTTGCTGTGTTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACA
GCCTGATGCAGGACAGGGATGCTGCCTCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGT
GAACAGGAGCCTGCCTGGCCTGATTGGCTGCCACAGGAAGTCTGTGTACTGGCATGTGATTGGCATGGGC
ACCACCCCTGAGGTGCACAGCATCTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGGCAGGCCA
GCCTGGAGATCAGCCCCATCACCTTCCTGACTGCCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCT
GTTCTGCCACATCAGCAGCCACCAGCATGATGGCATGGAGGCCTATGTGAAGGTGGACAGCTGCCCTGAG
GAGCCCCAGCTGAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGA
TGGATGTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAGCA
CCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCCCTGGTGCTGGCC
CCTGATGACAGGAGCTACAAGAGCCAGTACCTGAACAATGGCCCCCAGAGGATTGGCAGGAAGTACAAGA
AGGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATCCAGCATGAGTCTGGCAT
CCTGGGCCCCCTGCTGTATGGGGAGGTGGGGGACACCCTGCTGATCATCTTCAAGAACCAGGCCAGCAGG
CCCTACAACATCTACCCCCATGGCATCACTGATGTGAGGCCCCTGTACAGCAGGAGGCTGCCCAAGGGGG
TGAAGCACCTGAAGGACTTCCCCATCCTGCCTGGGGAGATCTTCAAGTACAAGTGGACTGTGACTGTGGA
GGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAGATACTACAGCAGCTTTGTGAACATGGAGAGG
GACCTGGCCTCTGGCCTGATTGGCCCCCTGCTGATCTGCTACAAGGAGTCTGTGGACCAGAGGGGCAACC
AGATCATGTCTGACAAGAGGAATGTGATCCTGTTCTCTGTGTTTGATGAGAACAGGAGCTGGTACCTGAC
TGAGAACATCCAGAGGTTCCTGCCCAACCCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGC
AACATCATGCACAGCATCAATGGCTATGTGTTTGACAGCCTGCAGCTGTCTGTGTGCCTGCATGAGGTGG
CCTACTGGTACATCCTGAGCATTGGGGCCCAGACTGACTTCCTGTCTGTGTTCTTCTCTGGCTACACCTT
CAAGCACAAGATGGTGTATGAGGACACCCTGACCCTGTTCCCCTTCTCTGGGGAGACTGTGTTCATGAGC
ATGGAGAACCCTGGCCTGTGGATTCTGGGCTGCCACAACTCTGACTTCAGGAACAGGGGCATGACTGCCC

TGCTGAAAGTCTCCAGCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGC
CTACCTGCTGAGCAAGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGAATGCCACTAATGTGTCTAAC
AACAGCAACACCAGCAATGACAGCAATGTGTCTCCCCCAGTGCTGAAGAGGCACCAGAGGGAGATCACCA
GGACCACCCTGCAGTCTGACCAGGAGGAGATTGACTATGATGACACCATCTCTGTGGAGATGAAGAAGGA
GGACTTTGACATCTACGACGAGGACGAGAACCAGAGCCCCAGGAGCTTCCAGAAGAAGACCAGGCACTAC
TTCATTGCTGCTGTGGAGAGGCTGTGGGACTATGGCATGAGCAGCAGCCCCCATGTGCTGAGGAACAGGG
CCCAGTCTGGCTCTGTGCCCCAGTTCAAGAAGGTGGTGTTCCAGGAGTTCACTGATGGCAGCTTCACCCA
GCCCCTGTACAGAGGGGAGCTGAATGAGCACCTGGGCCTGCTGGGCCCCTACATCAGGGCTGAGGTGGAG
GACAACATCATGGTGACCTTCAGGAACCAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCT
ATGAGGAGGACCAGAGGCAGGGGGCTGAGCCCAGGAAGAACTTTGTGAAGCCCAATGAAACCAAGACCTA
CTTCTGGAAGGTGCAGCACCACATGGCCCCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTC
TCTGATGTGGACCTGGAGAAGGATGTGCACTCTGGCCTGATTGGCCCCCTGCTGGTGTGCCACACCAACA
CCCTGAACCCTGCCCATGGCAGGCAGGTGACTGTGCAGGAGTTTGCCCTGTTCTTCACCATCTTTGATGA
AACCAAGAGCTGGTACTTCACTGAGAACATGGAGAGGAACTGCAGGGCCCCCTGCAACATCCAGATGGAG
GACCCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGCC
TGGTGATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTGAGCATGGGCAGCAATGAGAACATCCACAG
CATCCACTTCTCTGGCCATGTGTTCACTGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTGTACAACCTG
TACCCTGGGGTGTTTGAGACTGTGGAGATGCTGCCCAGCAAGGCTGGCATCTGGAGGGTGGAGTGCCTGA
TTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGGTGTACAGCAACAAGTGCCAGACCCCCCT
GGGCATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGCCAGTATGGCCAGTGGGCCCCC
AAGCTGGCCAGGCTGCACTACTCTGGCAGCATCAATGCCTGGAGCACCAAGGAGCCCTTCAGCTGGATCA
AGGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAGACCCAGGGGGCCAGGCAGAAGTTCAGCAG
CCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGATGGCAAGAAGTGGCAGACCTACAGGGGCAAC
AGCACTGGCACCCTGATGGTGTTCTTTGGCAATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAACC
CCCCCATCATTGCCAGATACATCAGGCTGCACCCCACCCACTACAGCATCAGGAGCACCCTGAGGATGGA
GCTGATGGGCTGTGACCTGAACAGCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATCTCTGATGCC
CAGATCACTGCCAGCAGCTACTTCACCAACATGTTTGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACC
TGCAGGGCAGGAGCAATGCCTGGAGGCCCCAGGTCAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCA
GAAGACCATGAAGGTGACTGGGGTGACCACCCAGGGGGTGAAGAGCCTGCTGACCAGCATGTATGTGAAG
GAGTTCCTGATCAGCAGCAGCCAGGATGGCCACCAGTGGACCCTGTTCTTCCAGAATGGCAAGGTGAAGG
TGTTCCAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCTGCTGACCAGATA
CCTGAGGATTCACCCCCAGAGCTGGGTGCACCAGATTGCCCTGAGGATGGAGGTGCTGGGCTGTGAGGCC
CAGGACCTGTACTGA

SEQ ID NO: 31

TACGTCTAACTCGACTCGTGGACGAAGAAGGACACGGACGACTCCAAGACGAAGAGACGGTGGTCCTCTA
TGATGGACCCCCGACACCTCGACTCGACCCTGATGTACGTCAGACTGGACCCCCTCGACGGACACCTACG
GTCCAAGGGGGGGTCTCACGGGTTCTCGAAGGGGAAGTTGTGGAGACACCACATGTTCTTCTGGGACAAA
CACCTCAAGTGACTGGTGGACAAGTTGTAACGGTTCGGGTCCGGGGGGACCTACCCGGACGACCCGGGGT
GGTAGGTCCGACTCCACATACTGTGACACCACTAGTGGGACTTCTTGTACCGGTCGGTGGGACACTCGGA
CGTACGACACCCCCACTCGATGACCTTCCGGAGACTCCCCCGACTCATACTACTGGTCTGGTCGGTCTCC
CTCTTCCTCCTACTGTTCCACAAGGGACCCCCGTCGGTGTGGATACACACCGTCCACGACTTCCTCTTAC
CGGGGTACCGGAGACTGGGGGACACGGACTGGATGTCGATGGACTCGGTACACCTGGACCACTTCCTGGA
CTTGAGACCGGACTAACCCCGGGACGACCACACGTCCCTCCCGTCGGACCGGTTCCTCTTCTGGGTCTGG
GACGTGTTCAAGTAGGACGACAAACGACACAAACTACTCCCGTTCTCGACCGTGAGACTTTGGTTCTTGT
CGGACTACGTCCTGTCCCTACGACGGAGACGGTCCCGGACCGGGTTCTACGTGTGACACTTACCGATACA
CTTGTCCTCGGACGGACCGGACTAACCGACGGTGTCCTTCAGACACATGACCGTACACTAACCGTACCCG
TGGTGGGGACTCCACGTGTCGTAGAAGGACCTCCCGGTGTGGAAGGACCAGTCCTTGGTGTCCGTCCGGT
CGGACCTCTAGTCGGGGTAGTGGAAGGACTGACGGGTCTGGGACGACTACCTGGACCCGGTCAAGGACGA
CAAGACGGTGTAGTCGTCGGTGGTCGTACTACCGTACCTCCGGATACACTTCCACCTGTCGACGGGACTC
CTCGGGGTCGACTCCTACTTCTTGTTACTCCTCCGACTCCTGATACTACTACTGGACTGACTGAGACTCT
ACCTACACCACTCCAAACTACTACTGTTGTCGGGGTCGAAGTAGGTCTAGTCCAGACACCGGTTCTTCGT
GGGGTTCTGGACCCACGTGATGTAACGACGACTCCTCCTCCTGACCCTGATACGGGGGGACCACGACCGG
GGACTACTGTCCTCGATGTTCTCGGTCATGGACTTGTTACCGGGGGTCTCCTAACCGTCCTTCATGTTCT
TCCAGTCCAAGTACCGGATGTGACTACTTTGGAAGTTCTGGTCCCTCCGGTAGGTCGTACTCAGACCGTA
GGACCCGGGGGACGACATACCCCTCCACCCCCTGTGGGACGACTAGTAGAAGTTCTTGGTCCGGTCGTCC
GGGATGTTGTAGATGGGGGTACCGTAGTGACTACACTCCGGGGACATGTCGTCCTCCGACGGGTTCCCCC

ACTTCGTGGACTTCCTGAAGGGGTAGGACGGACCCCTCTAGAAGTTCATGTTCACCTGACACTGACACCT
CCTACCGGGGTGGTTCAGACTGGGGTCCACGGACTGGTCTATGATGTCGTCGAAACACTTGTACCTCTCC
CTGGACCGGAGACCGGACTAACCGGGGGACGACTAGACGATGTTCCTCAGACACCTGGTCTCCCCGTTGG
TCTAGTACAGACTGTTCTCCTTACACTAGGACAAGAGACACAAACTACTCTTGTCCTCGACCATGGACTG
ACTCTTGTAGGTCTCCAAGGACGGGTTGGGACGACCCCACGTCGACCTCCTGGGACTCAAGGTCCGGTCG
TTGTAGTACGTGTCGTAGTTACCGATACACAAACTGTCGGACGTCGACAGACACACGGACGTACTCCACC
GGATGACCATGTAGGACTCGTAACCCCGGGTCTGACTGAAGGACAGACACAAGAAGAGACCGATGTGGAA
GTTCGTGTTCTACCACATACTCCTGTGGGACTGGGACAAGGGGAAGAGACCCCTCTGACACAAGTACTCG
TACCTCTTGGGACCGGACACCTAAGACCCGACGGTGTTGAGACTGAAGTCCTTGTCCCCGTACTGACGGG
ACGACTTTCAGAGGTCGACACTGTTCTTGTGACCCCTGATGATACTCCTGTCGATACTCCTGTAGAGACG
GATGGACGACTCGTTCTTGTTACGGTAACTCGGGTCCTCGAAGTCGGTCTTACGGTGATTACACAGATTG
TTGTCGTTGTGGTCGTTACTGTCGTTACACAGAGGGGGTCACGACTTCTCCGTGGTCTCCCTCTAGTGGT
CCTGGTGGGACGTCAGACTGGTCCTCCTCTAACTGATACTACTGTGGTAGAGACACCTCTACTTCTTCCT
CCTGAAACTGTAGATGCTGCTCCTGCTCTTGGTCTCGGGGTCCTCGAAGGTCTTCTTCTGGTCCGTGATG
AAGTAACGACGACACCTCTCCGACACCCTGATACCGTACTCGTCGTCGGGGGTACACGACTCCTTGTCCC
GGGTCAGACCGAGACACGGGGTCAAGTTCTTCCACCACAAGGTCCTCAAGTGACTACCGTCGAAGTGGGT
CGGGGACATGTCTCCCCTCGACTTACTCGTGGACCCGGACGACCCGGGGATGTAGTCCCGACTCCACCTC
CTGTTGTAGTACCACTGGAAGTCCTTGGTCCGGTCGTCCGGGATGTCGAAGATGTCGTCGGACTAGTCGA
TACTCCTCCTGGTCTCCGTCCCCCGACTCGGGTCCTTCTTGAAACACTTCGGGTTACTTTGGTTCTGGAT
GAAGACCTTCCACGTCGTGGTGTACCGGGGGTGGTTCCTACTCAAACTGACGTTCCGGACCCGGATGAAG
AGACTACACCTGGACCTCTTCCTACACGTGAGACCGGACTAACCGGGGGACGACCACACGGTGTGGTTGT
GGGACTTGGGACGGGTACCGTCCGTCCACTGACACGTCCTCAAACGGGACAAGAAGTGGTAGAAACTACT
TTGGTTCTCGACCATGAAGTGACTCTTGTACCTCTCCTTGACGTCCCGGGGGACGTTGTAGGTCTACCTC
CTGGGGTGGAAGTTCCTCTTGATGTCCAAGGTACGGTAGTTACCGATGTAGTACCTGTGGGACGGACCGG
ACCACTACCGGGTCCTGGTCTCCTAGTCCACCATGGACGACTCGTACCCGTCGTTACTCTTGTAGGTGTC
GTAGGTGAAGAGACCGGTACACAAGTGACACTCCTTCTTCCTCCTCATGTTCTACCGGGACATGTTGGAC
ATGGGACCCCACAAACTCTGACACCTCTACGACGGGTCGTTCCGACCGTAGACCTCCCACCTCACGGACT
AACCCCTCGTGGACGTACGACCGTACTCGTGGGACAAGGACCACATGTCGTTGTTCACGGTCTGGGGGGA
CCCGTACCGGAGACCGGTGTAGTCCCTGAAGGTCTAGTGACGGAGACCGGTCATACCGGTCACCCGGGGG
TTCGACCGGTCCGACGTGATGAGACCGTCGTAGTTACGGACCTCGTGGTTCCTCGGGAAGTCGACCTAGT
TCCACCTGGACGACCGGGGGTACTAGTAGGTACCGTAGTTCTGGGTCCCCCGGTCCGTCTTCAAGTCGTC
GGACATGTAGTCGGTCAAGTAGTAGTACATGTCGGACCTACCGTTCTTCACCGTCTGGATGTCCCCGTTG
TCGTGACCGTGGGACTACCACAAGAAACCGTTACACCTGTCGAGACCGTAGTTCGTGTTGTAGAAGTTGG
GGGGGTAGTAACGGTCTATGTAGTCCGACGTGGGGTGGGTGATGTCGTAGTCCTCGTGGGACTCCTACCT
CGACTACCCGACACTGGACTTGTCGACGTCGTACGGGGACCCGTACCTCTCGTTCCGGTAGAGACTACGG
GTCTAGTGACGGTCGTCGATGAAGTGGTTGTACAAACGGTGGACCTCGGGGTCGTTCCGGTCCGACGTGG
ACGTCCCGTCCTCGTTACGGACCTCCGGGGTCCAGTTGTTGGGGTTCCTCACCGACGTCCACCTGAAGGT
CTTCTGGTACTTCCACTGACCCCACTGGTGGGTCCCCCACTTCTCGGACGACTGGTCGTACATACACTTC
CTCAAGGACTAGTCGTCGTCGGTCCTACCGGTGGTCACCTGGGACAAGAAGGTCTTACCGTTCCACTTCC
ACAAGGTCCCGTTGGTCCTGTCGAAGTGGGGACACCACTTGTCGGACCTGGGGGGGGACGACTGGTCTAT
GGACTCCTAAGTGGGGGTCTCGACCCACGTGGTCTAACGGGACTCCTACCTCCACGACCCGACACTCCGG
GTCCTGGACATGACT

SEQ ID NO: 32

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLF
VEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQR
EKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQT
LHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMG
TTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMRAYVKVDSCPE
EPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLA
PDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASR
PYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMER
DLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQAS
NIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMS
MENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNATNVSN
NSNTSNDSNVSPPVLKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHY
FIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVE

DNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYF
SDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQME
DPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNL
YPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAP
KLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGN
STGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDA
QITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVK
EFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEA
QDLY

SEQ ID NO: 33

GGCGCCTCTAGAGTTATAACCGGTAATCGGTATAATAAGTAACCAATATATCGTATTTAGTTATAACCGA
TAACCGGTAACGTATGCAACATAGATATAGTATTATACATGTAAATATAACCGAGTACAGGTTATACTGG
CGGTACAACCGTAACTAATAACTGATCAATAATTATCATTAGTTAATGCCCCAGTAATCAAGTATCGGGT
ATATACCTCAAGGCGCAATGTATTGAATGCCATTTACCGGGCGGACCGACTGGCGGGTTGCTGGGGGCGG
GTAACTGCAGTTATTACTGCATACAAGGGTATCATTGCGGTTATCCCTGAAAGGTAACTGCAGTTACCCA
CCTCATAAATGCCATTTGACGGGTGAACCGTCATGTAGTTCACATAGTATACGGTTCAGGCGGGGGATAA
CTGCAGTTACTGCCATTTACCGGGCGGACCGTAATACGGGTCATGTACTGGAATGCCCTGAAAGGATGAA
CCGTCATGTAGATGCATAATCAGTAGCGATAATGGTACCACTACGCCAAAACCGTCATGTGGTTACCCGC
ACCTATCGCCAAACTGAGTGCCCCTAAAGGTTCAGAGGTGGGGTAACTGCAGTTACCCTCAAACAAAACC
GTGGTTTTAGTTGCCCTGAAAGGTTTTACAGCATTATTGGGGCGGGGCAACTGCGTTTACCCGCCATCCG
CACATGCCACCCTCCAGATATATTCGTCTCGAGCAAATCACTTGGCAGTCTAGTGATCTTCGAAATAACG
CCATCAAATAGTGTCAATTTAACGATTGCGTCAGTCACGAAGACTGTGTTGTCAGAGCTTGAATTCGACG
TCTTCAACCAGCACTCCGTGACCCGTCCGATCG

REIVINDICACIONES

1. Un vector lentiviral pseudotipado con hemaglutinina-neuraminidasa (HN) y proteínas de fusión (F) de un paramixovirus respiratorio, en el que dicho vector lentiviral comprende un promotor EF1a (hCEF)/potenciador híbrido de CMV humano y un transgén, y en el que el vector lentiviral carece de uno o cualquier intrón posicionado entre dicho promotor y dicho transgén.

2. El vector lentiviral, según la reivindicación 1, en el que el vector lentiviral se selecciona del grupo que consiste en un vector del virus de inmunodeficiencia humana (VIH), un vector del virus de inmunodeficiencia simia (VIS), un vector del virus de inmunodeficiencia felina (VIF), un vector del virus de la anemia infecciosa equina (VAIE), y un vector del virus Visna/maedi; en el que, preferentemente, el vector lentiviral es un vector del VIS.

3. El vector lentiviral, según la reivindicación 1 o 2, en el que el paramixovirus respiratorio es un virus Sendai.

4. El vector lentiviral, según cualquiera de las reivindicaciones anteriores, en el que el transgén se selecciona de:

(a) una proteína terapéutica secretada, opcionalmente, alfa-1 antitripsina (*A1AT*), Factor VIII, proteína tensioactiva B (*SFTPB*), Factor VII, Factor IX, Factor X, Factor XI, Factor von Willebrand, factor estimulante de colonias de granulocitos-macrófagos (*GM-CSF*), o un anticuerpo monoclonal contra un agente infeccioso; o
(b) *CFTR, DNAH5, DNAH11, DNAI1* o *DNAI2*.

5. El vector lentiviral, según cualquiera de las reivindicaciones 1 a 4, en el que:
(a) el transgén codifica CFTR;
(b) el transgén codifica A1AT; o
(c) el transgén codifica FVIII.

6. El vector lentiviral, según cualquiera de las reivindicaciones 1 a 5, en el que el transgén codifica CFTR.

7. Un procedimiento para producir lentivirus, tal como se define en cualquiera de las reivindicaciones 1 a 6, comprendiendo dicho procedimiento las siguientes etapas:
(a) cultivar células en suspensión;
(b) transfectar las células con uno o más plásmidos;
(c) añadir una nucleasa;
(d) recoger los lentivirus;
(e) añadir tripsina; y
(f) purificación.

8. El procedimiento, según la reivindicación 7, en el que:
(i) las etapas (a)-(f) se realizan secuencialmente;
(ii) las células son células HEK293T o 293T/17;
(iii) la adición de la nucleasa es en la fase previa a la recogida;
(iv) la adición de tripsina es en la fase posterior a la recogida; y/o
(v) la etapa de purificación comprende una etapa de cromatografía.

9. Un vector lentiviral, tal como se define en cualquiera de las reivindicaciones 1 a 6, para su uso en un procedimiento de tratamiento de una enfermedad.

10. El vector lentiviral para el uso, según la reivindicación 9, en el que la enfermedad es una enfermedad pulmonar seleccionada de entre: fibrosis quística (FQ); discinesia ciliar primaria (PCD); deficiencia de proteína tensioactiva B (SP-B); deficiencia de alfa-1-antitripsina (A1AD); proteinosis alveolar pulmonar (PAP); y enfermedad pulmonar obstructiva crónica (EPOC).

11. El vector lentiviral para el uso, según la reivindicación 9, en el que:
(a) la enfermedad a tratar es FQ;
(b) la enfermedad a tratar es deficiencia de A1AT; o
(c) la enfermedad a tratar es hemofilia, en la que, opcionalmente, el transgén es el Factor VIII.

12. Una célula huésped que comprende el vector, según cualquiera de las reivindicaciones 1-6.

13. Una composición que comprende un vector, según cualquiera de las reivindicaciones 1-6, y un vehículo farmacéuticamente aceptable.

Figura 1:

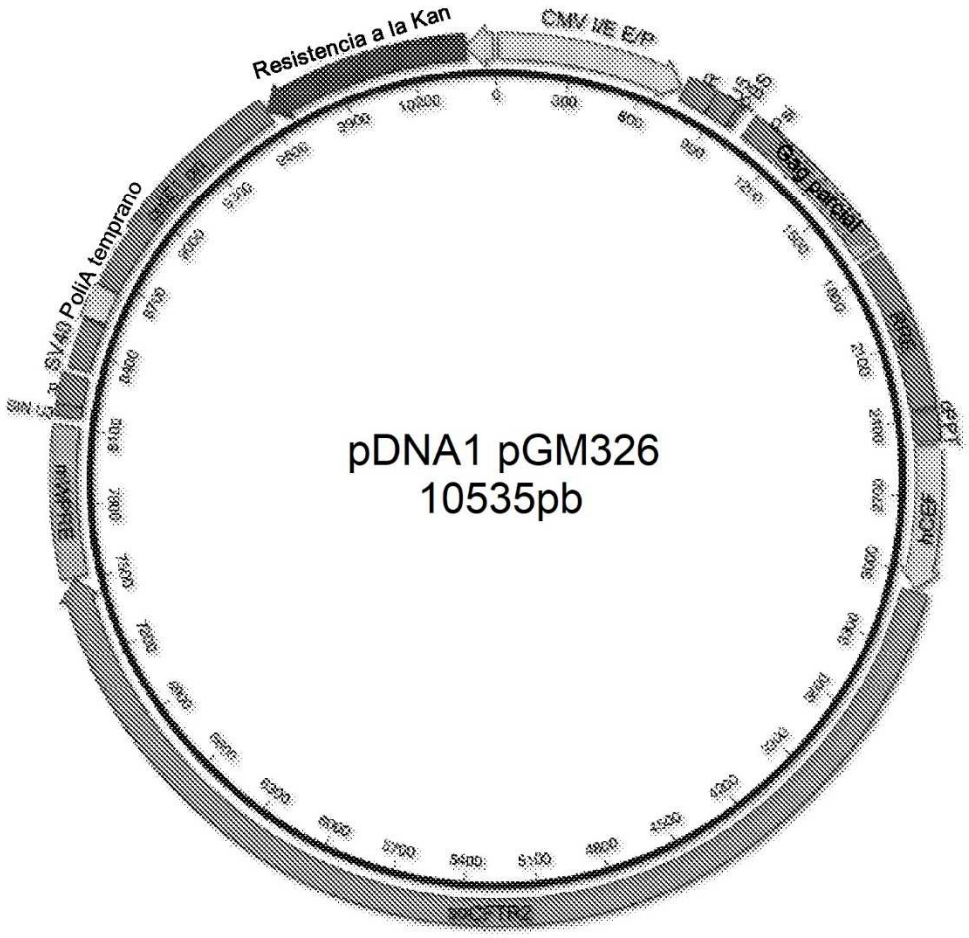


Figura 1A

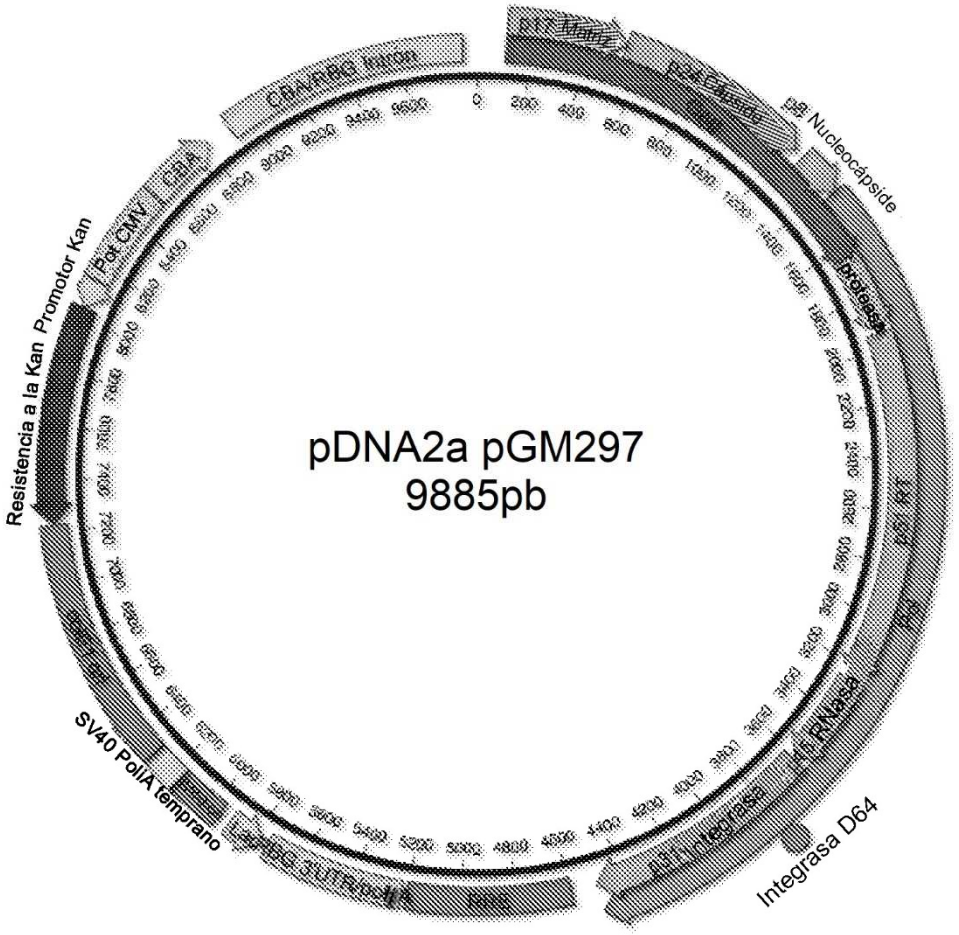

pDNA2a pGM297
9885pb
Resistencia a la Kan
Promotor Kan
Pol CMV
CBA
CRA/RRG kanR
Nucleocápside
proteasa
RT RNasa
Integrasa D64
SV40 PolyA temprano

Figura 1B

Figura 1C

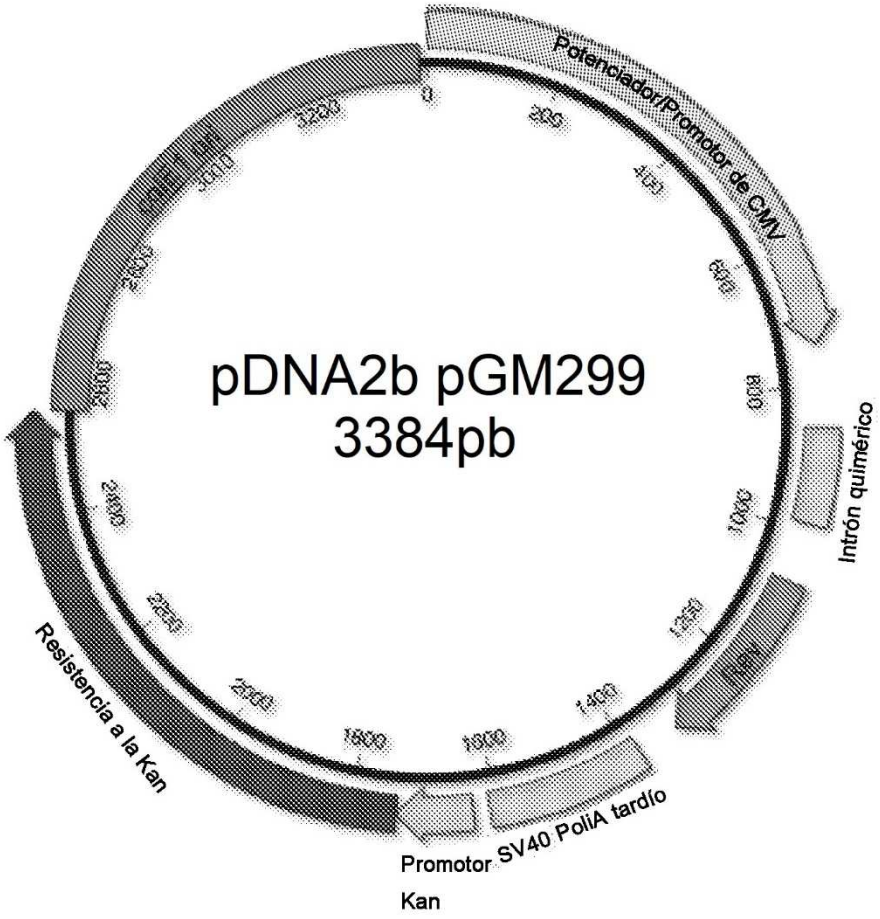
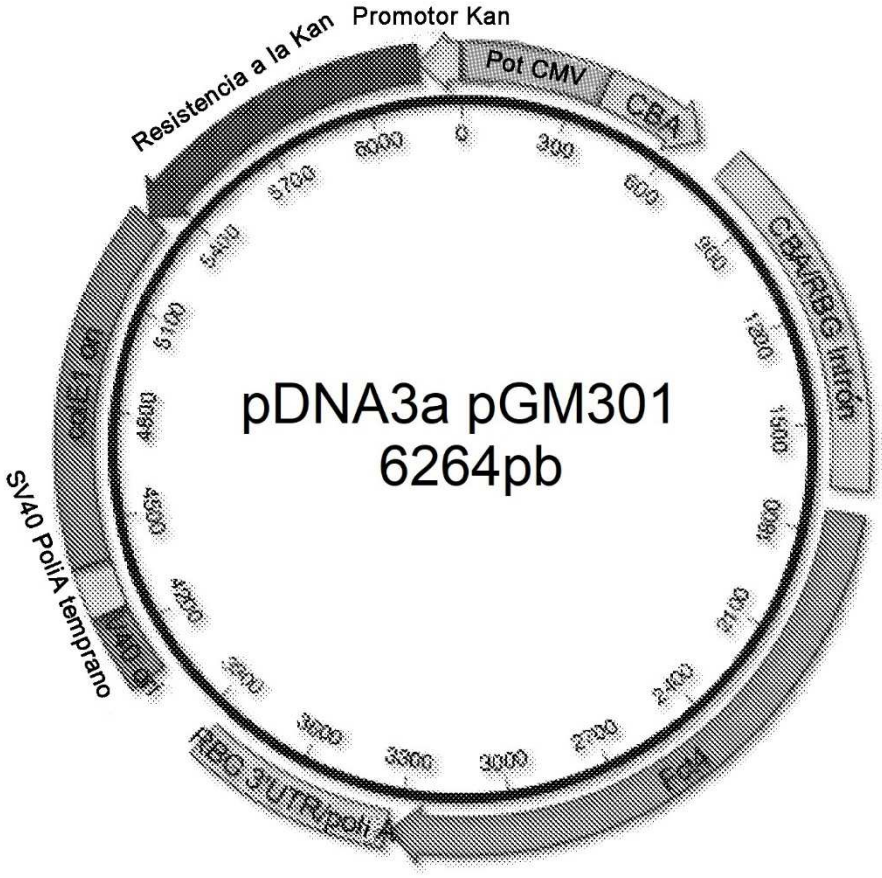

pDNA2b pGM299
3384pb
Potenciador/Promotor de CMV
Intrón quimérico
SV40 PoliA tardío
Promotor Kan
Resistencia a la Kan
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
3200

Figura 1D


Promotor Kan
Resistencia a la Kan
Pot CMV
CBA
CBA/RBG Intrón
pDNA3a pGM301
6264pb
SV40 PoliA temprano
RBG 3'UTR/poli A
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
3300
3600
3900
4200
4500
4800
5100
5400
5700
6000

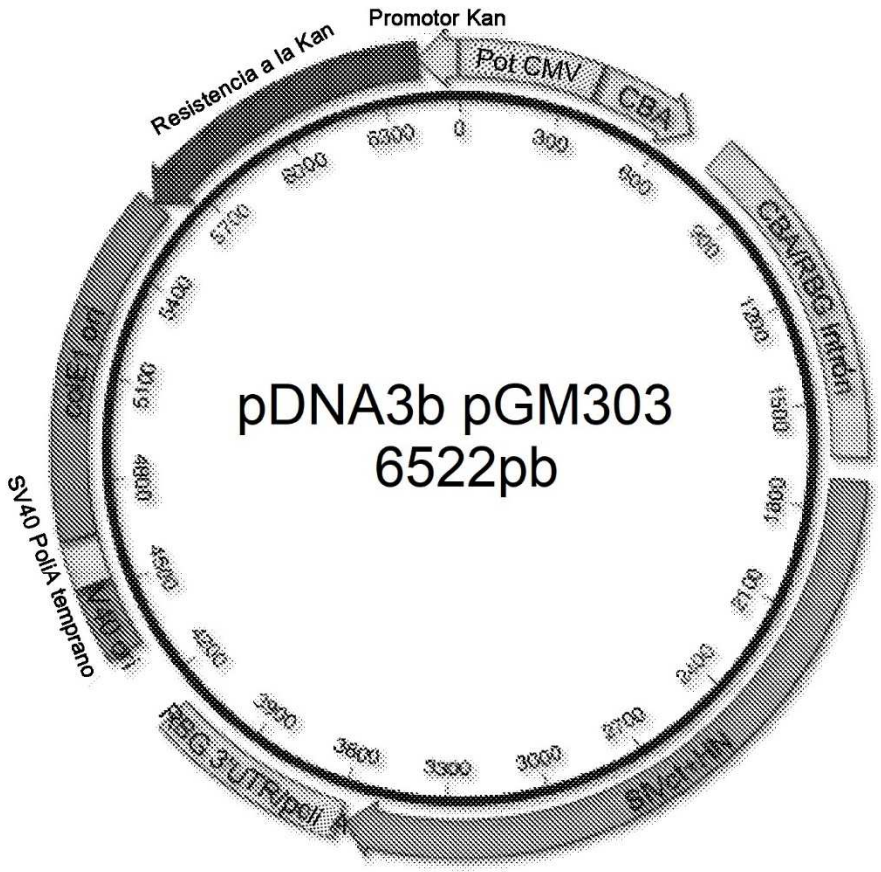

Promotor Kan
Resistencia a la Kan
Pot CMV
CBA
CBA/rBG Intron
pDNA3b pGM303
6522pb
SV40 PoliA temprano
rBG 3'UTR polI
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
3300
3600
3900
4200
4500
4800
5100
5400
5700
6000
6300

Figura 1E

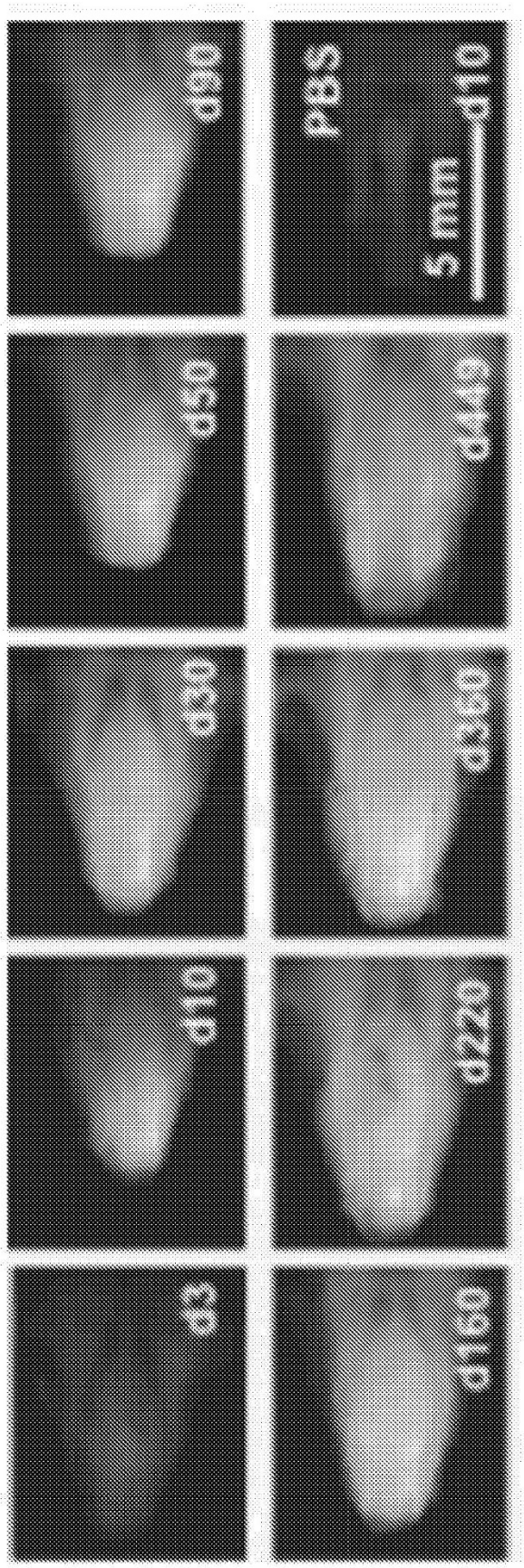

PBS
5 mm
d10
d60
d120
d30
d90
d150
d5
d15
d45
d75
d105
d135
d0
d10
d20
d30
d40
d50
d60
d70
d80
d90
d100
d110
d120
d130
d140
d150

Figura 2

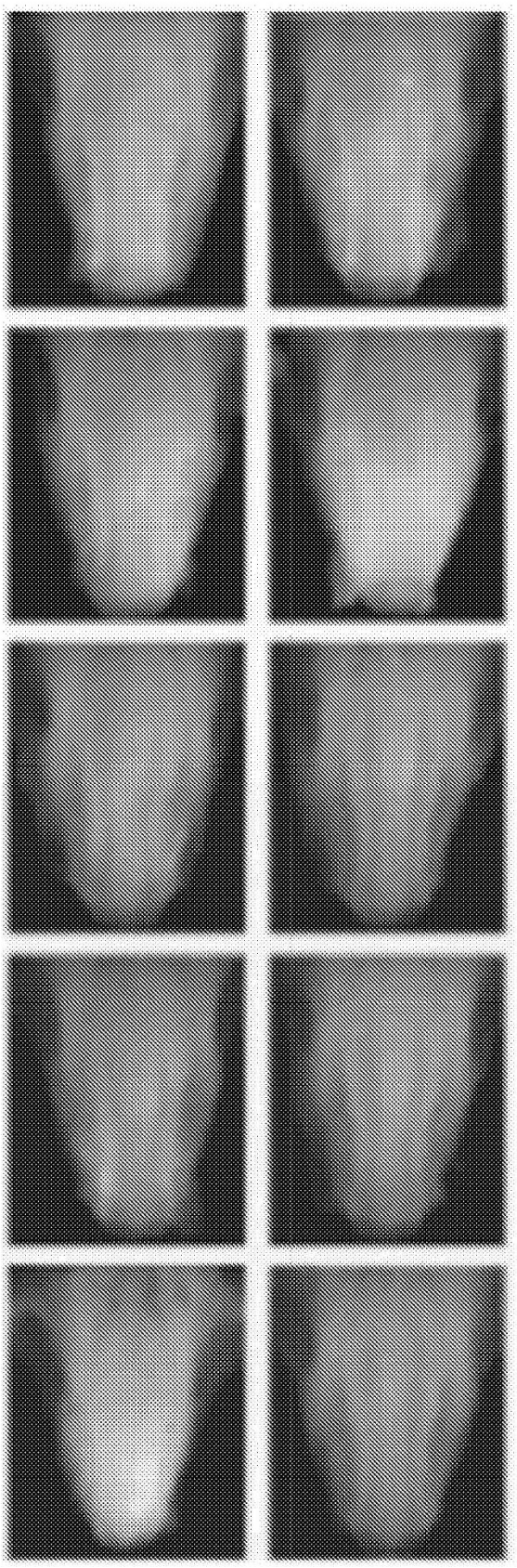


Figura 3

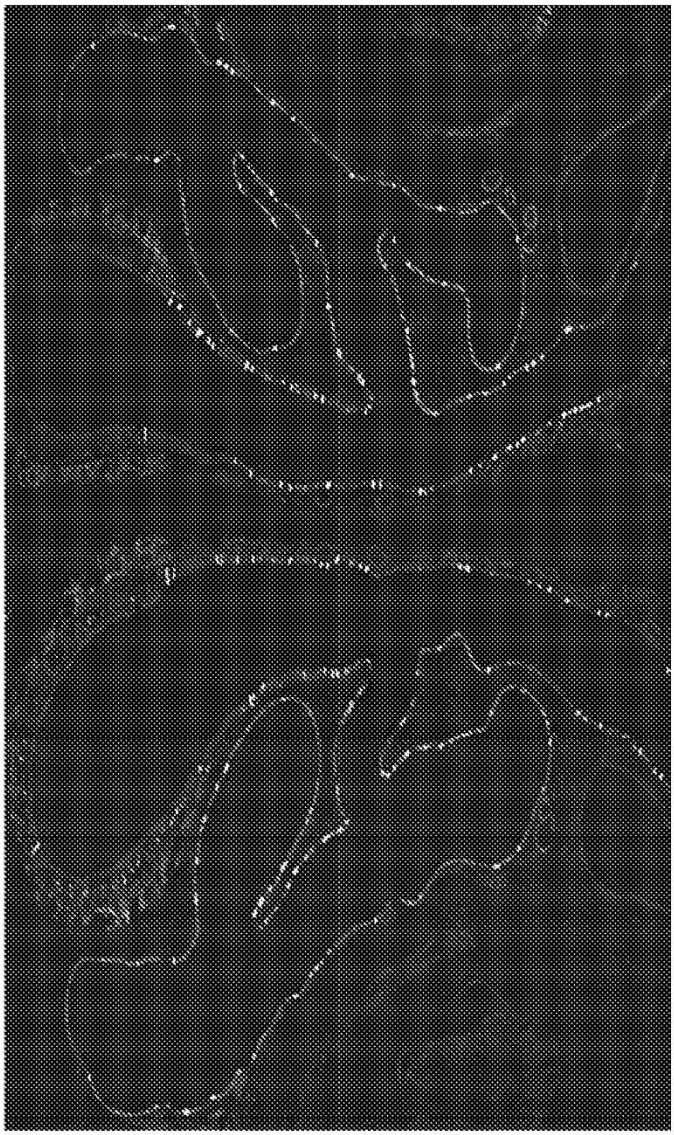


Figura 4



Figura 5

e

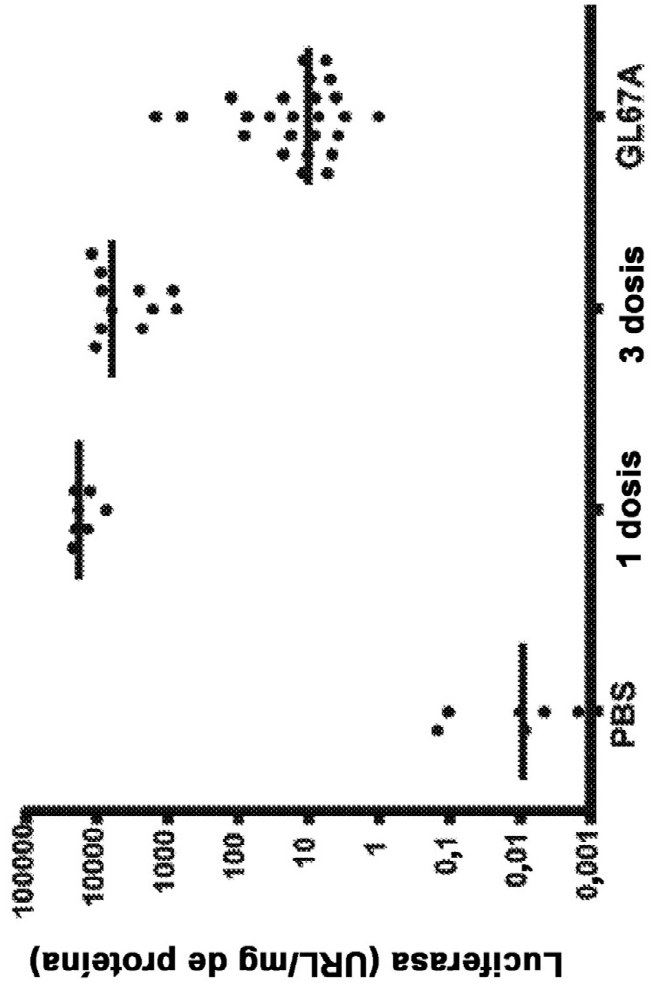

Luciferasa (URL/mg de proteína)
100000
10000
1000
100
10
1
0,1
0,01
0,001
PBS
1 dosis
3 dosis
GL67A

Figura 6

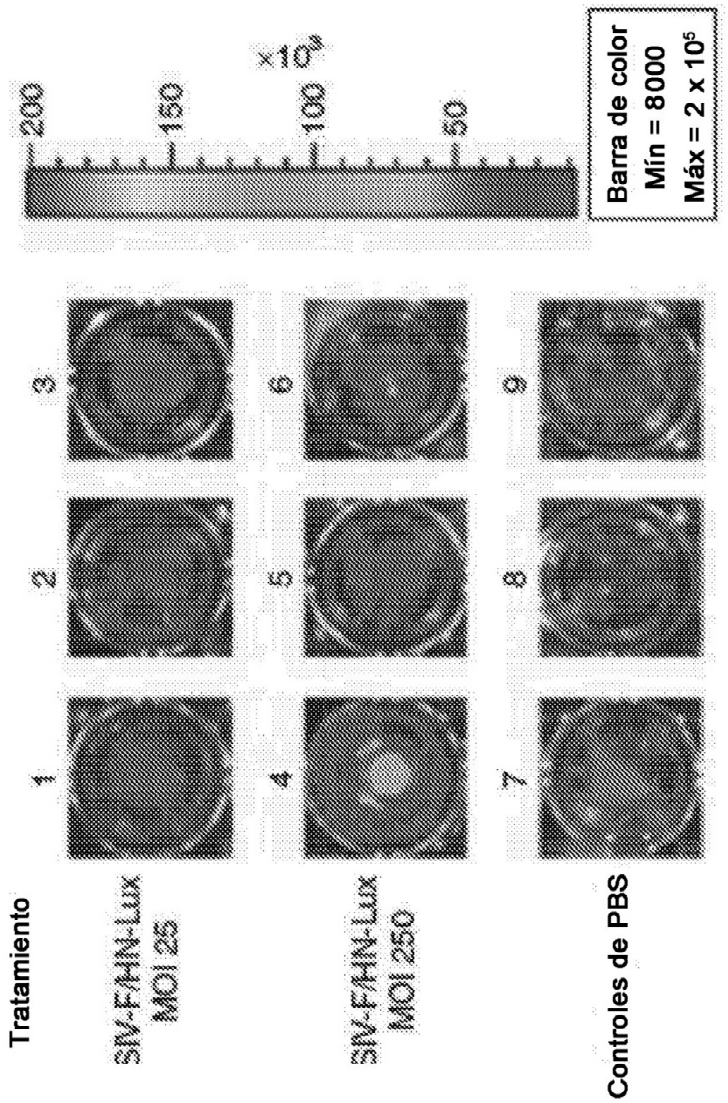

Tratamiento
SIV-F/HN-Lux MOI 25
SIV-F/HN-Lux MOI 250
Controles de PBS
1
2
3
4
5
6
7
8
9
200
150
100
50
×10³
Barra de color
Mín = 8000
Máx = 2 x 10⁵

Figura 7

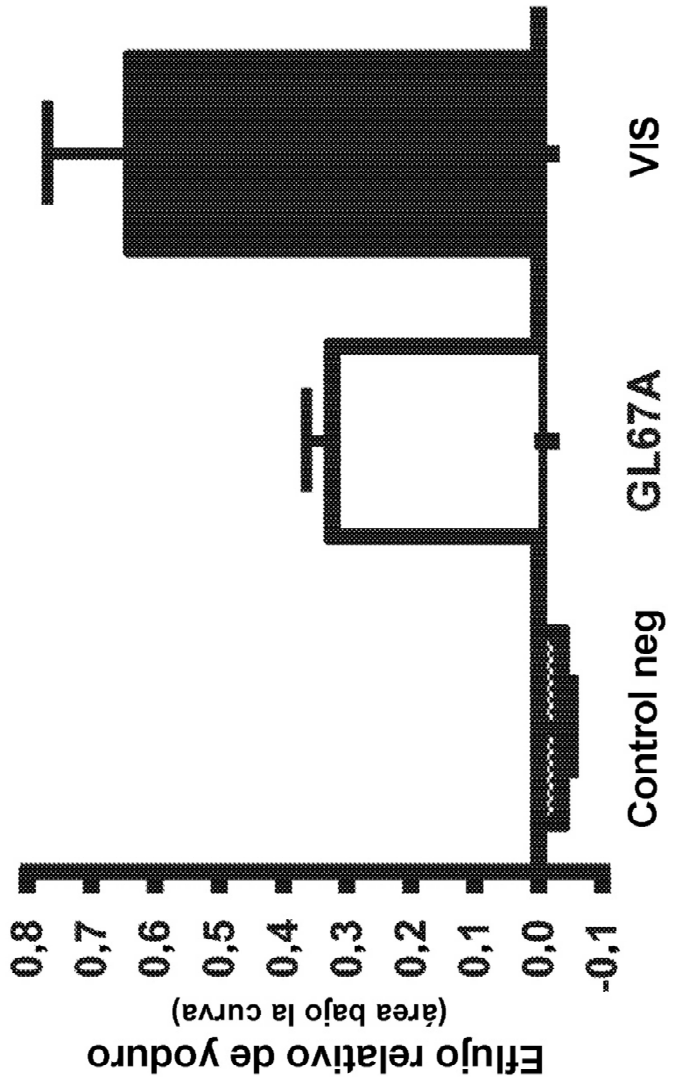
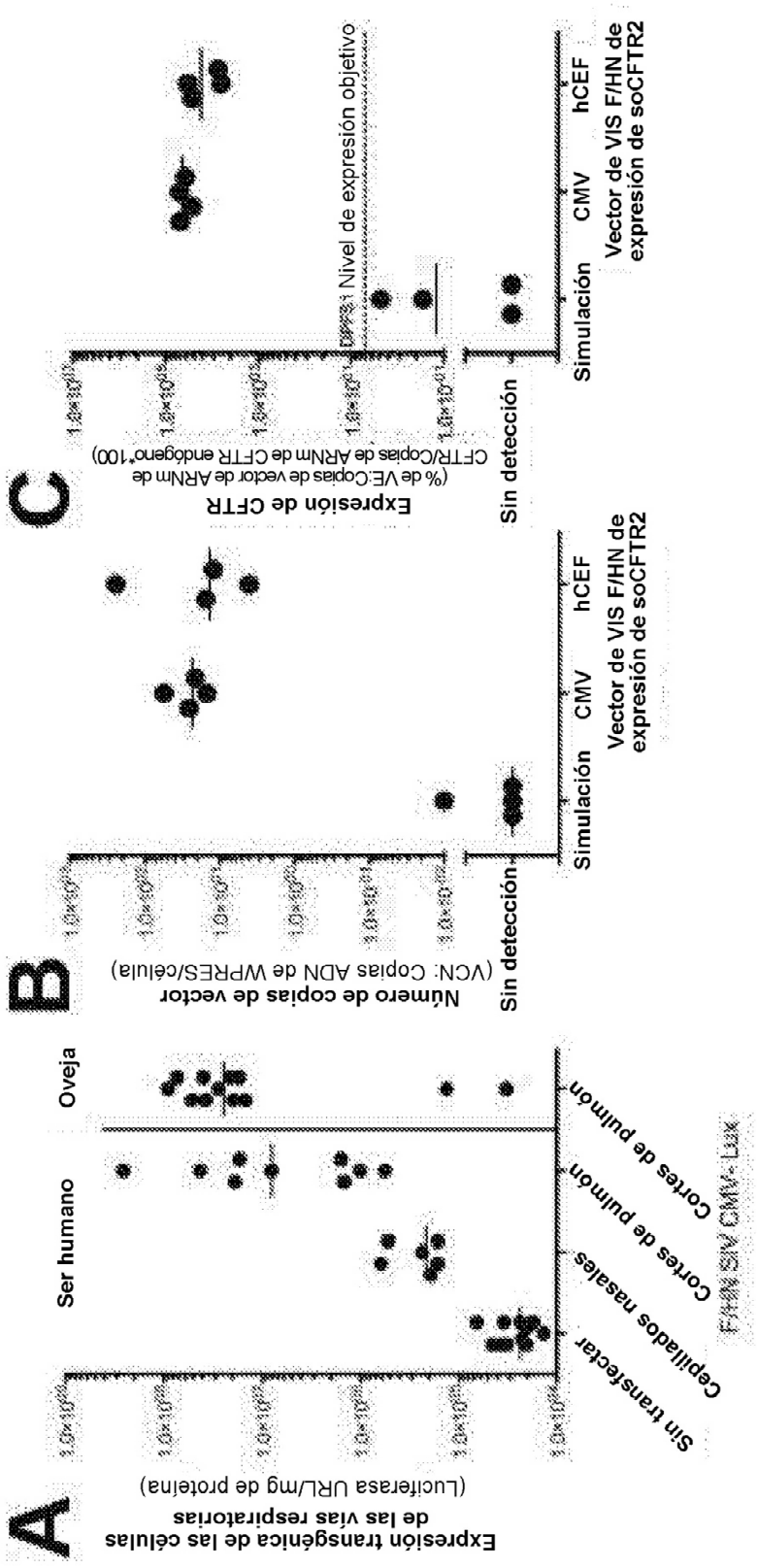

Eflujo relativo de yoduro
(área bajo la curva)
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0,0
-0,1
Control neg
GL67A
VIS

Figura 8

Figura 9


A
Expresión transgénica de las células de las vías respiratorias
(Luciferasa URL/mg de proteína)
Ser humano
Oveja
Sin transfectar
Cepillados nasales
Cortes de pulmón
Cortes de pulmón
F/HN SIV CMV- Lux
B
Número de copias de vector
(VCN: Copias ADN de WPRES/célula)
Sin detección
Simulación
CMV
hCEF
Vector de VIS F/HN de expresión de soCFTR2
C
Expresión de CFTR
(% de VE:Copias de vector de ARNm de CFTR/Copias de ARNm de CFTR endógeno*100)
Nivel de expresión objetivo
Sin detección
Simulación
CMV
hCEF
Vector de VIS F/HN de expresión de soCFTR2

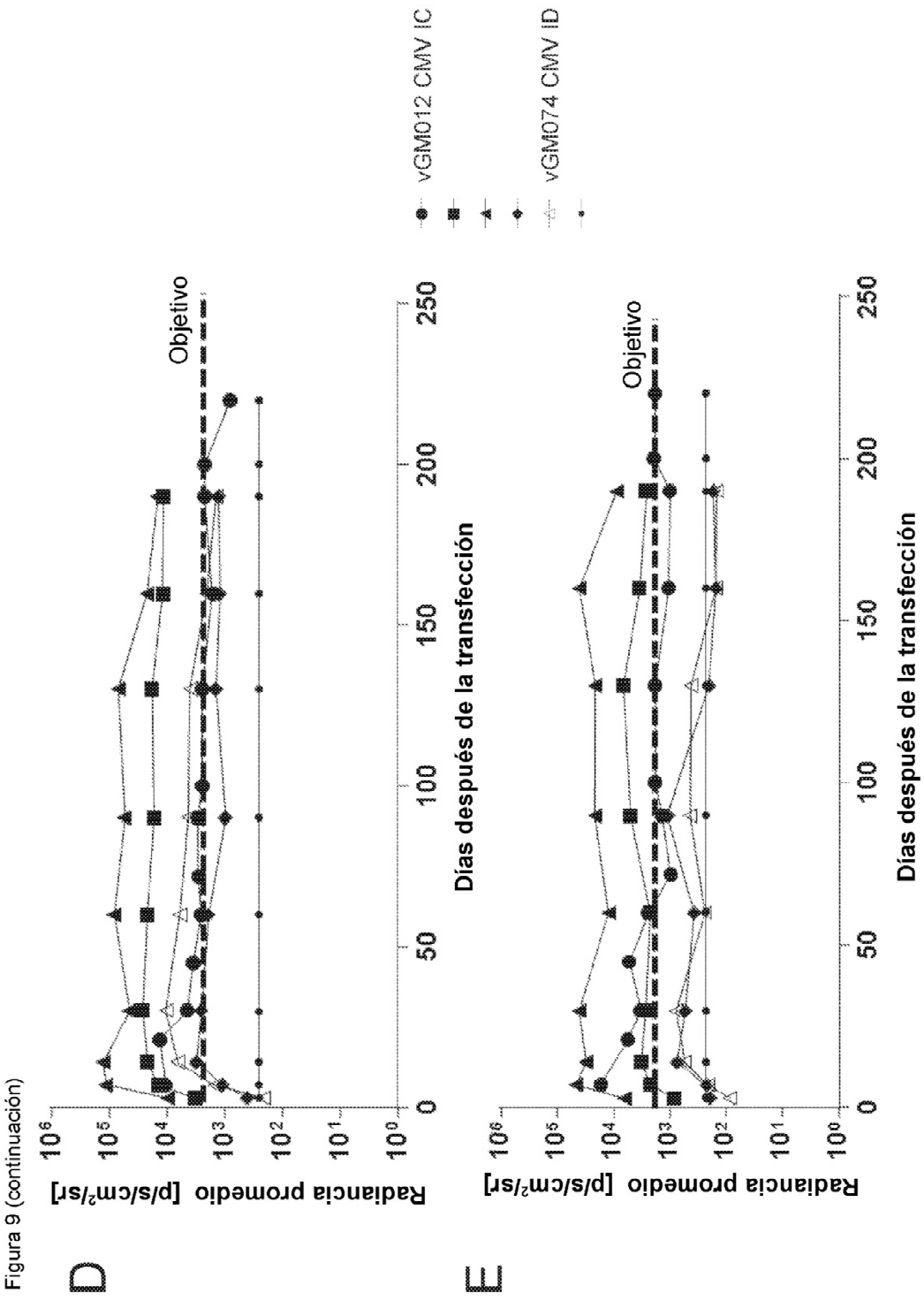
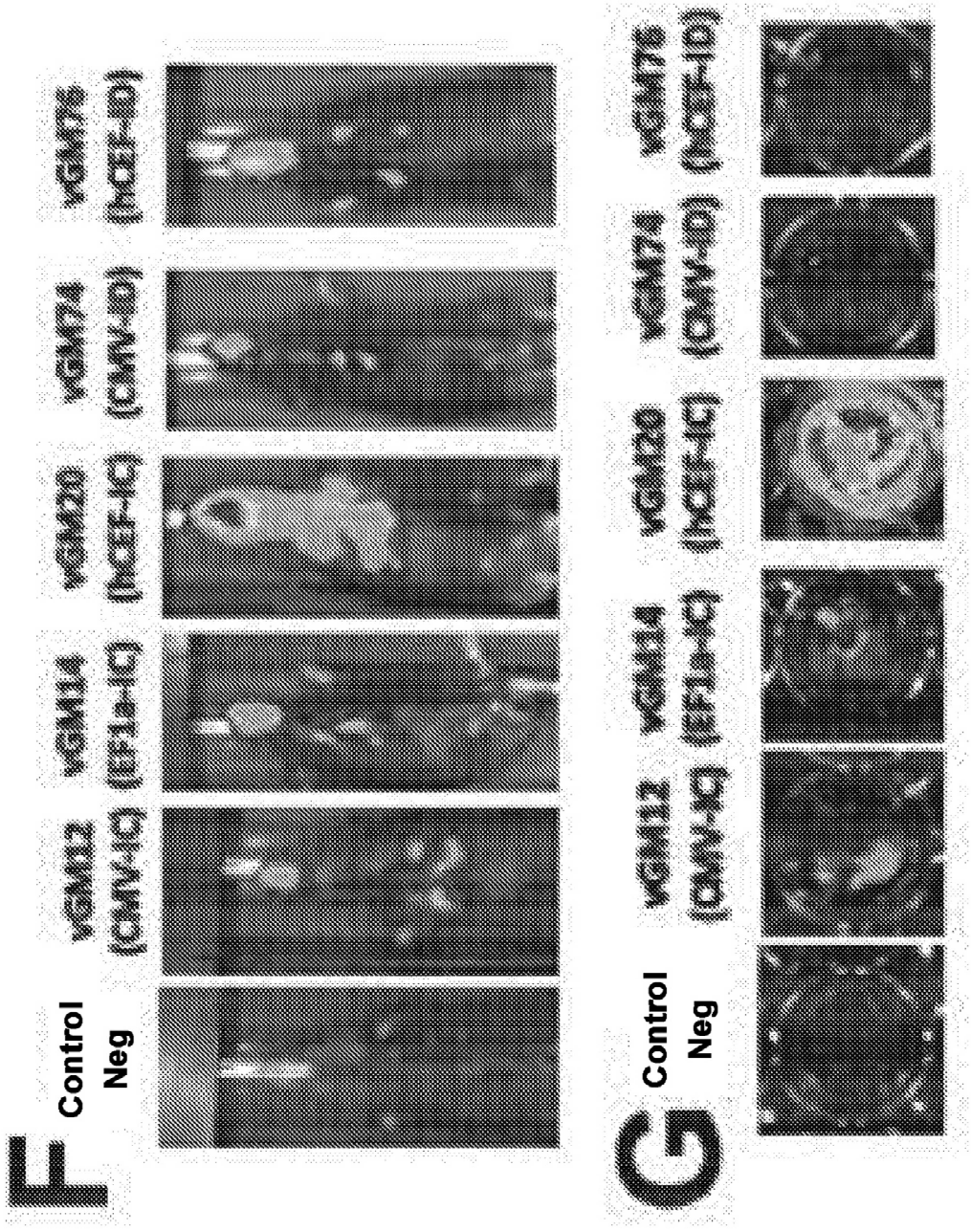


Figura 9 (continuación)

Figura 9 (continuación)



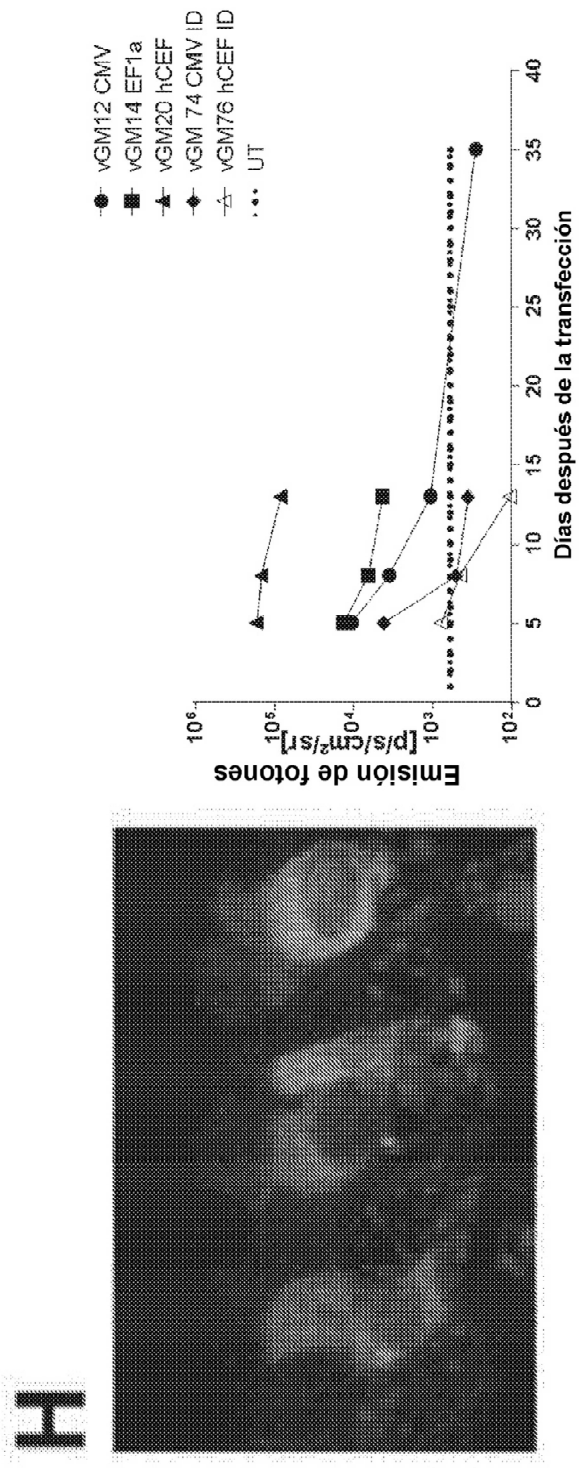


Figura 9 (continuación)

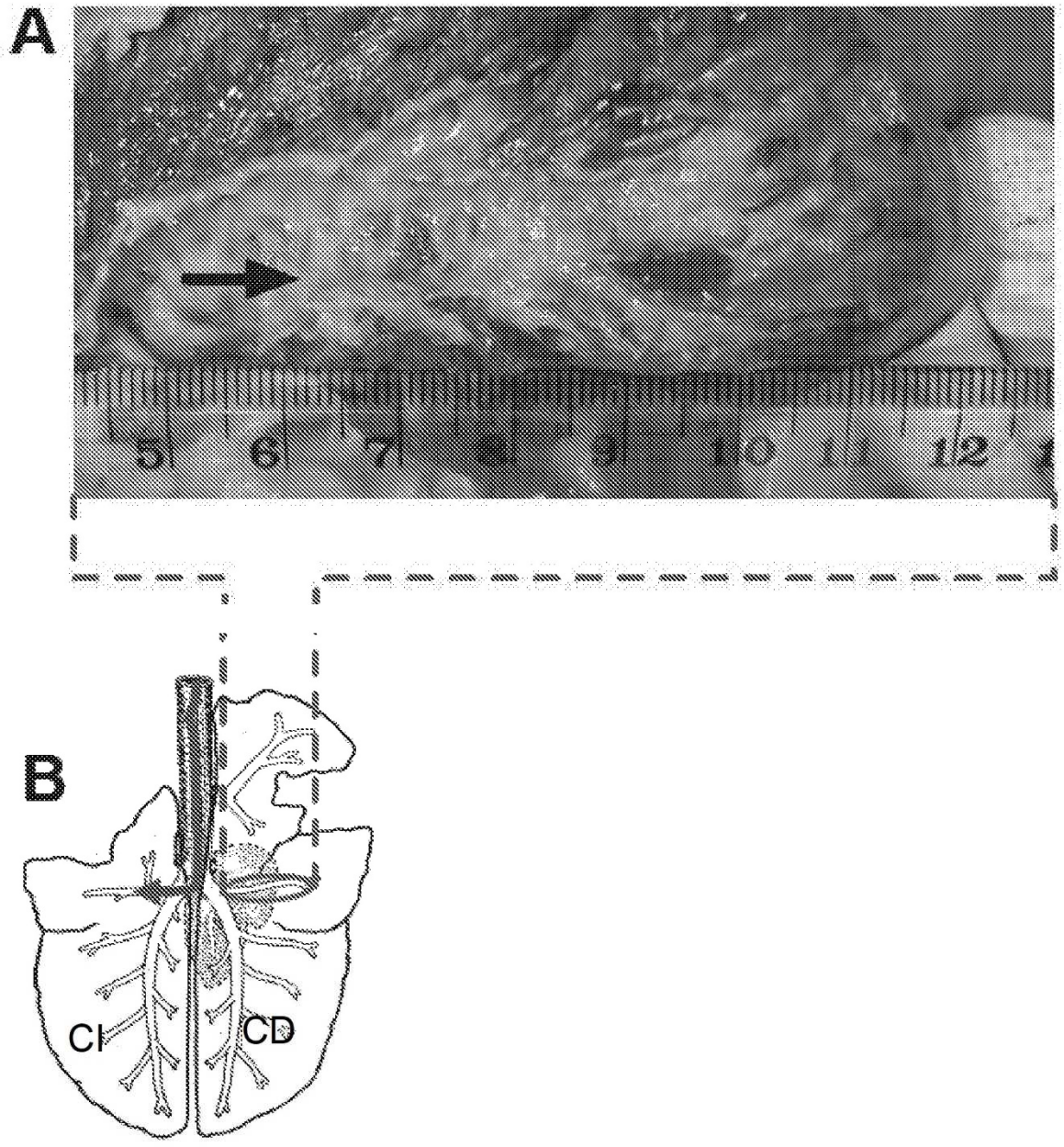


Figura 10

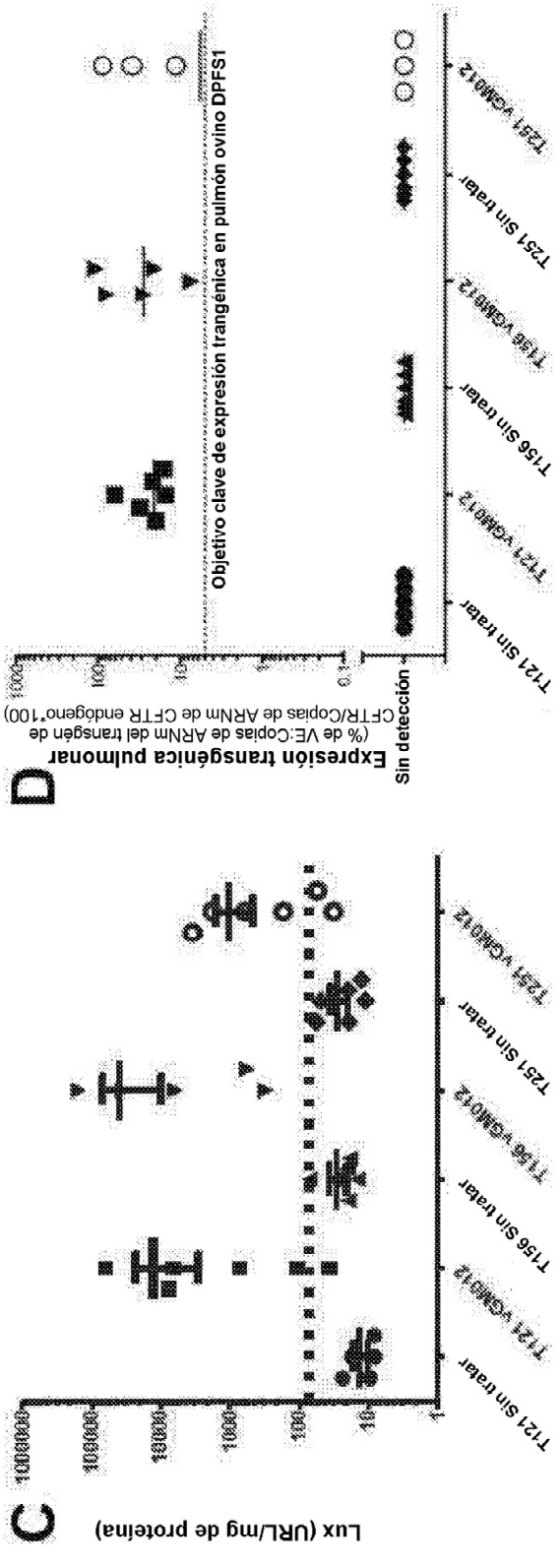


Figura 10 (continuación)

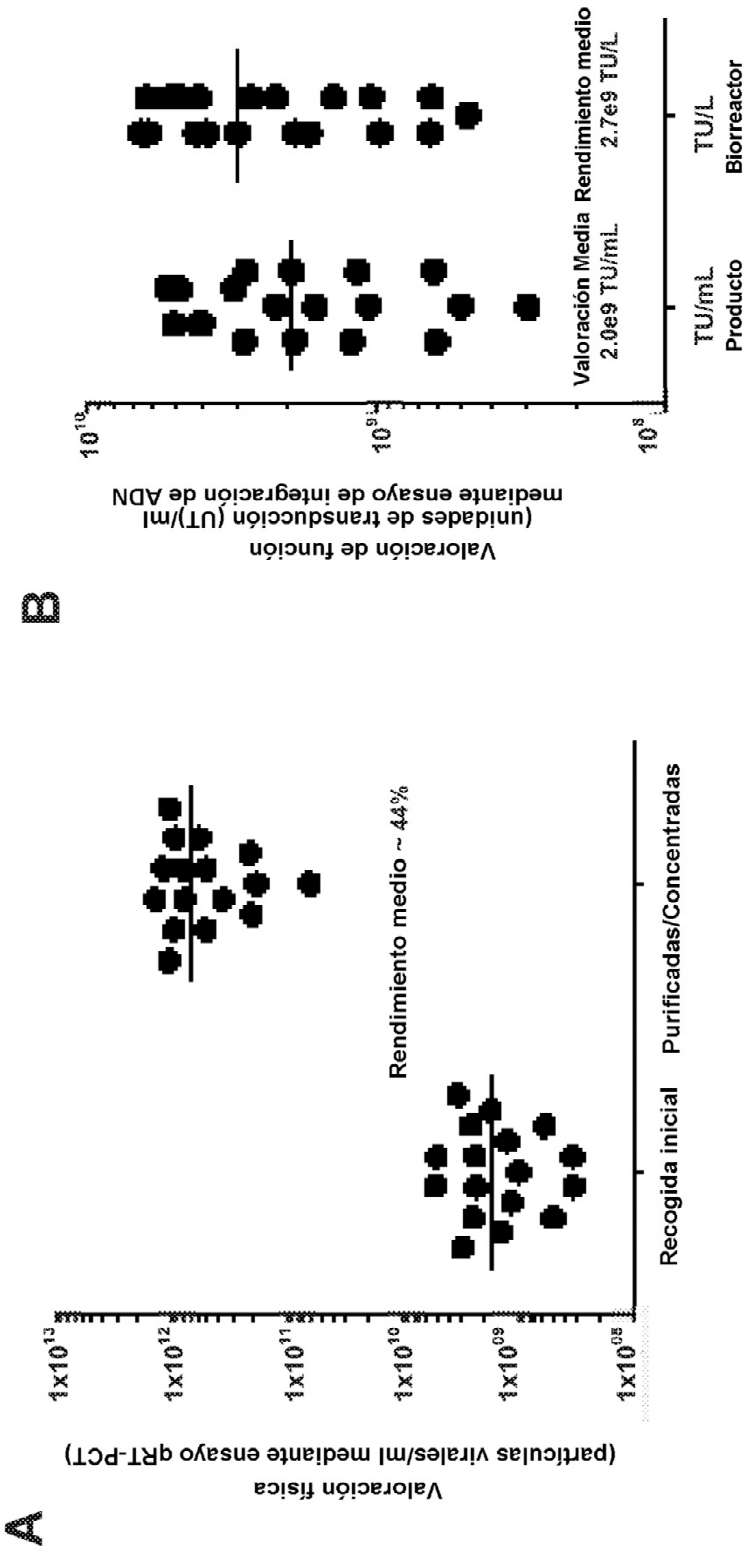

A
Valoración física
(partículas víricas/ml mediante ensayo qRT-PCT)
1x10^13
1x10^12
1x10^11
1x10^10
1x10^09
1x10^08
Rendimiento medio ~ 44%
Recogida inicial
Purificadas/Concentradas
B
Valoración de función
(unidades de transducción (UT)/ml
mediante ensayo de integración de ADN
10^10
10^9
10^8
Valoración Media
2.0e9 TU/mL
Rendimiento medio
2.7e9 TU/L
TU/mL
Producto
TU/L
Biorreactor

Figura 11

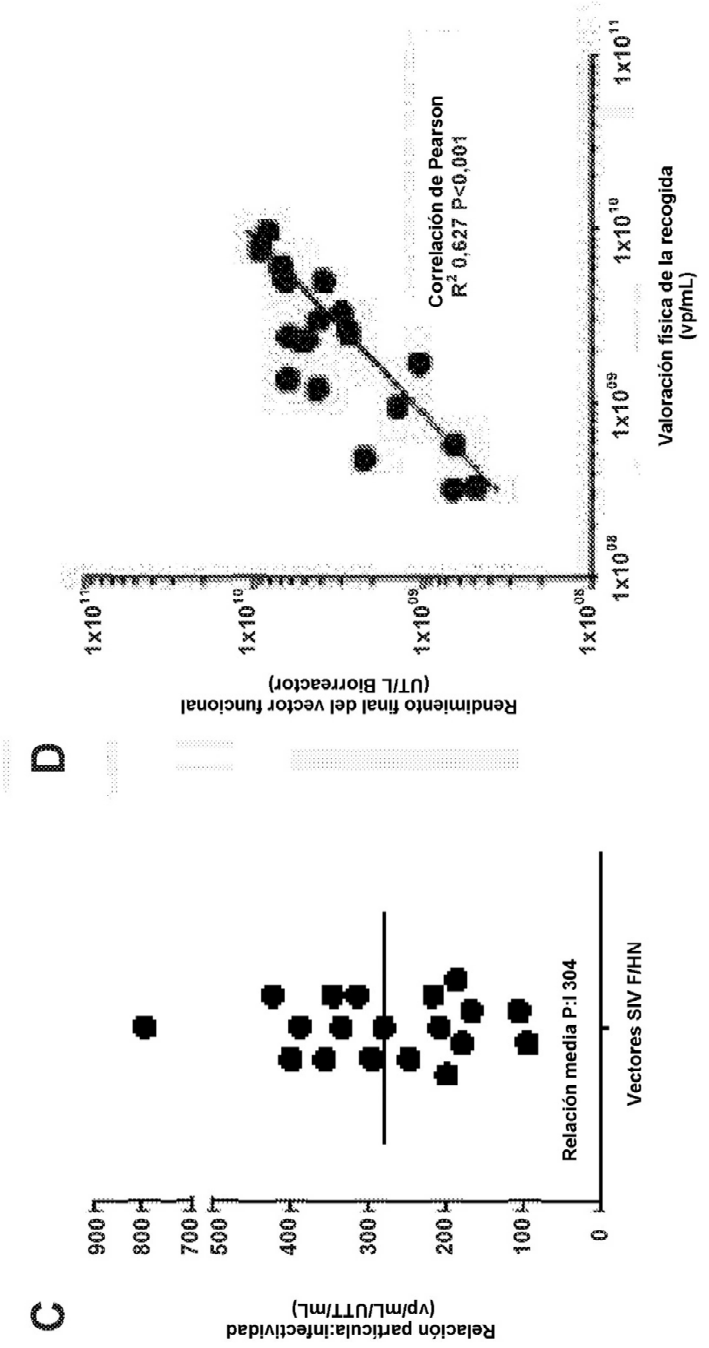


Figura 11 (continuación)

Figura 12

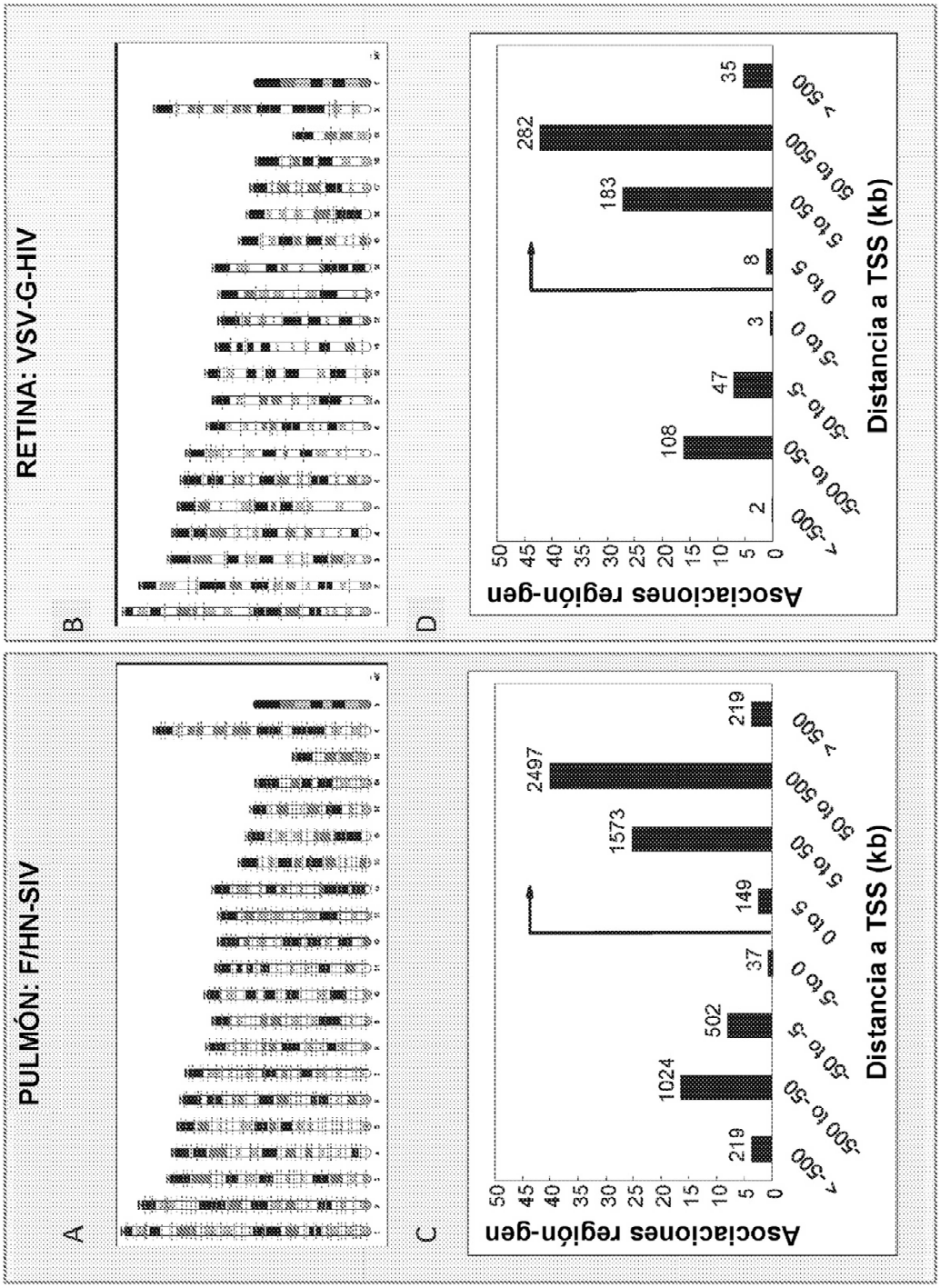
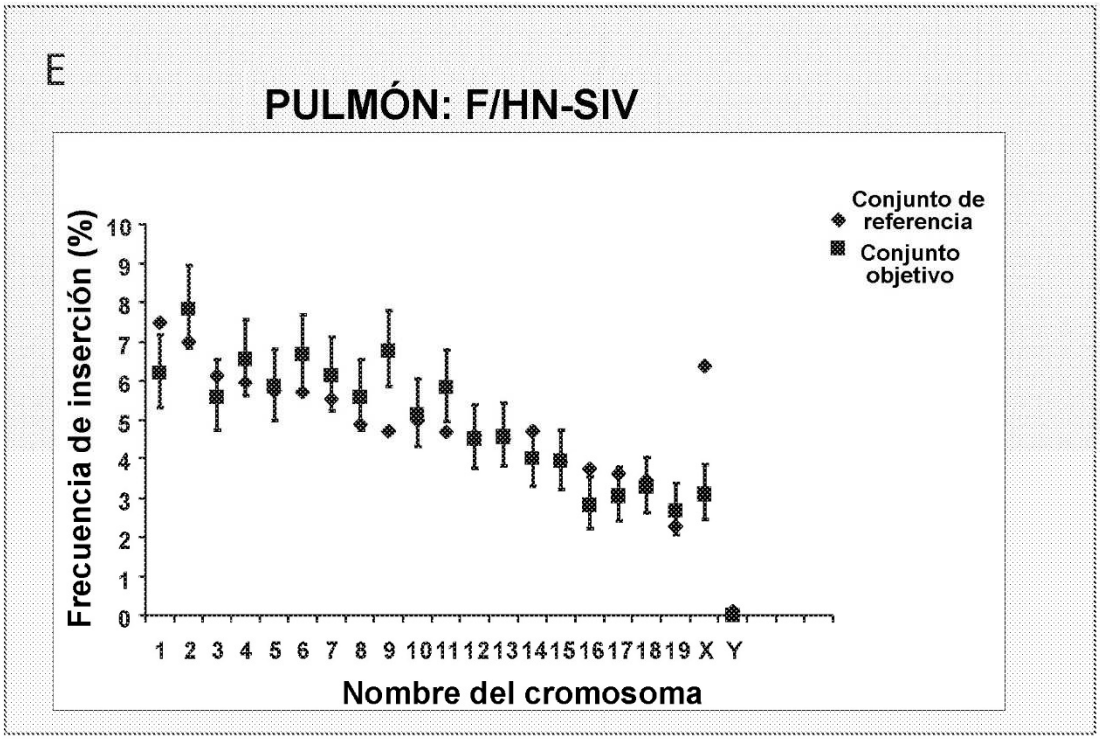
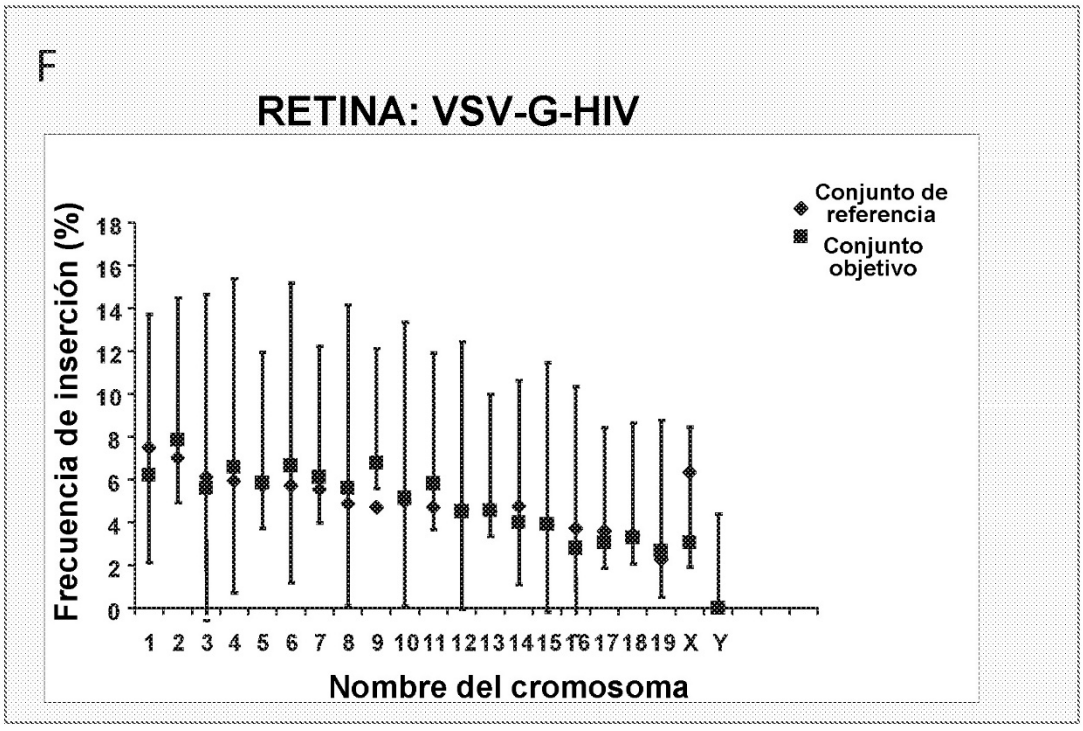

PULMÓN: F/HN-SIV
RETINA: VSV-G-HIV
A
B
C
D
Asociaciones región-gen
Distancia a TSS (kb)
<-500
-500 to -50
-50 to -5
-5 to 0
0 to 5
5 to 50
50 to 500
>500
219
1024
502
37
149
1573
2497
219
2
108
47
3
8
183
262
35

Figura 12 (continuación)





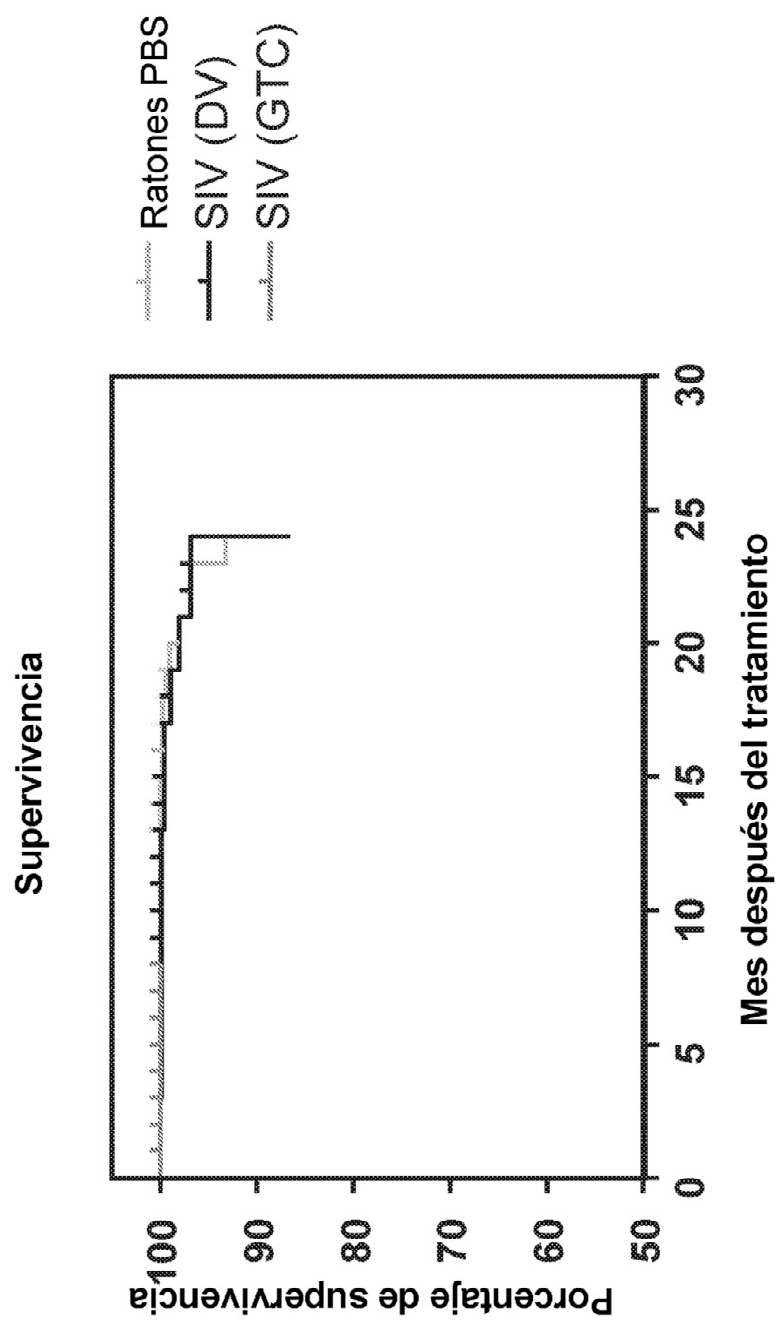


Figura 12 (continuación)

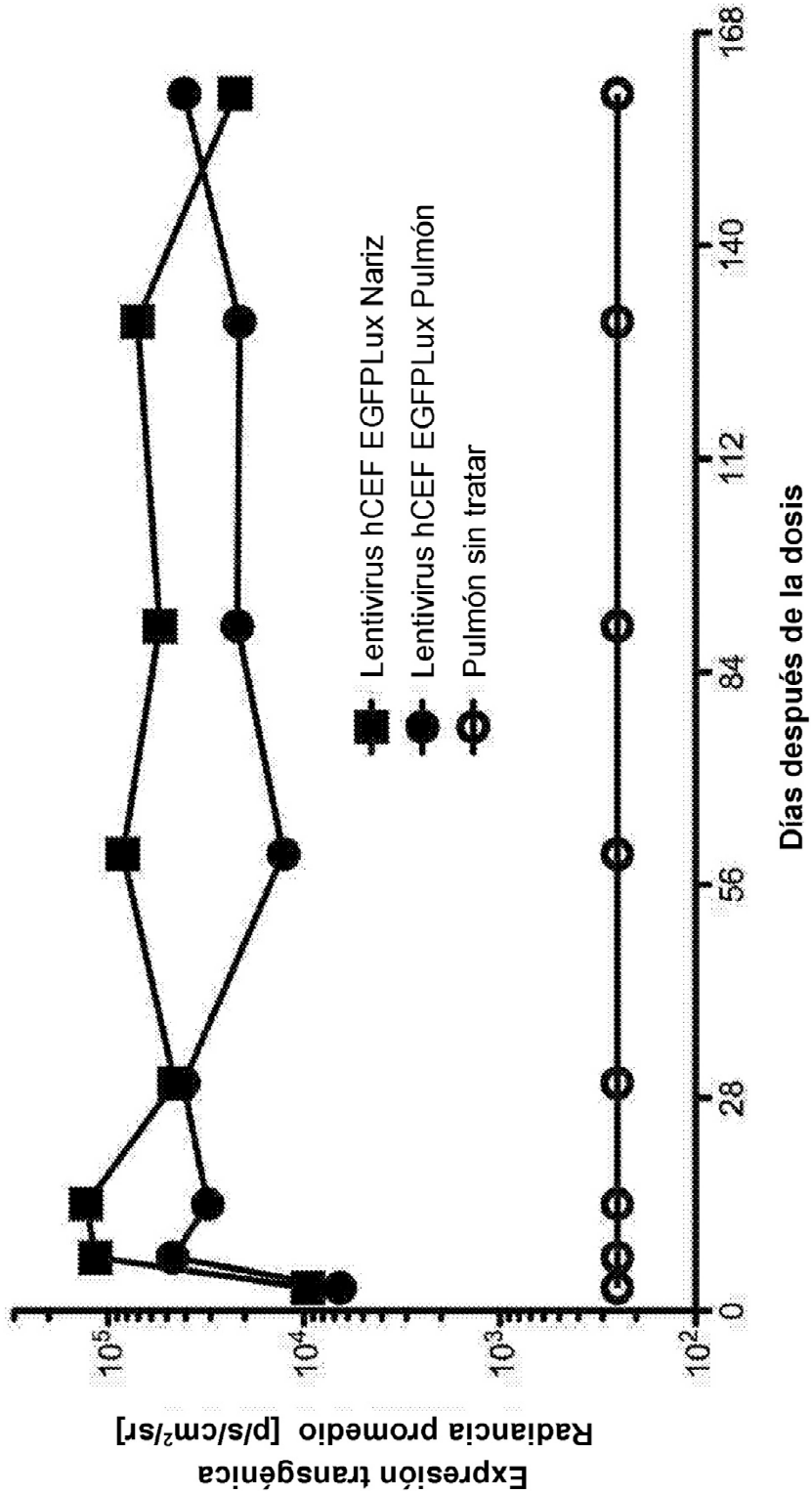
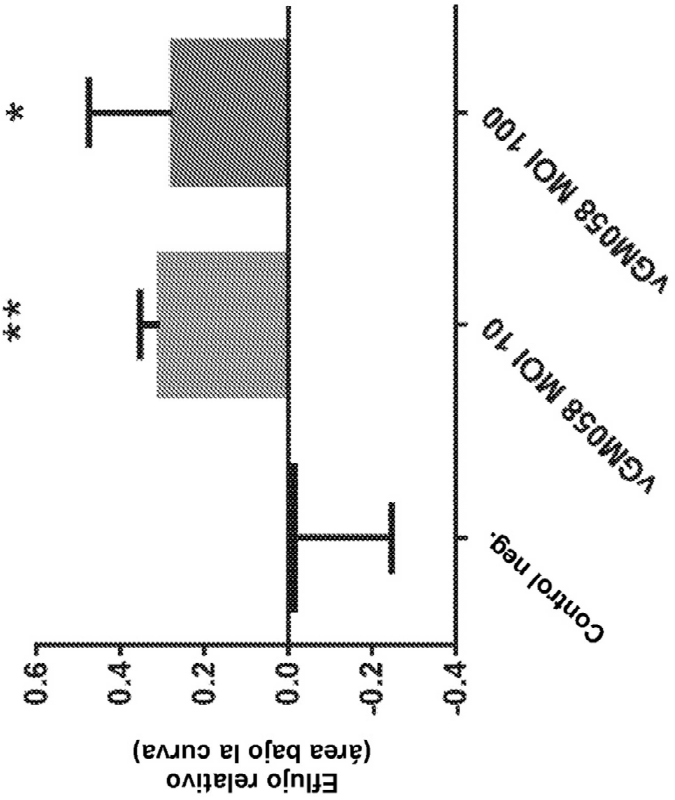

Expresión transgénica
Radiancia promedio [p/s/cm²/sr]
10⁵
10⁴
10³
10²
0
28
56
84
112
140
168
Días después de la dosis
Lentivirus hCEF EGFPLux Nariz
Lentivirus hCEF EGFPLux Pulmón
Pulmón sin tratar

Figura 13


Eflujo relativo
(área bajo la curva)
0.6
0.4
0.2
0.0
-0.2
-0.4
Control neg.
VGM058 MOI 10
VGM058 MOI 100
**
*

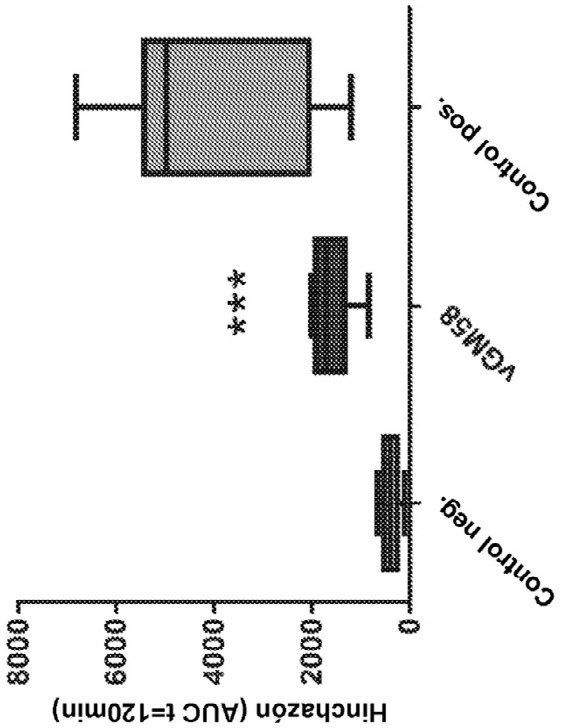
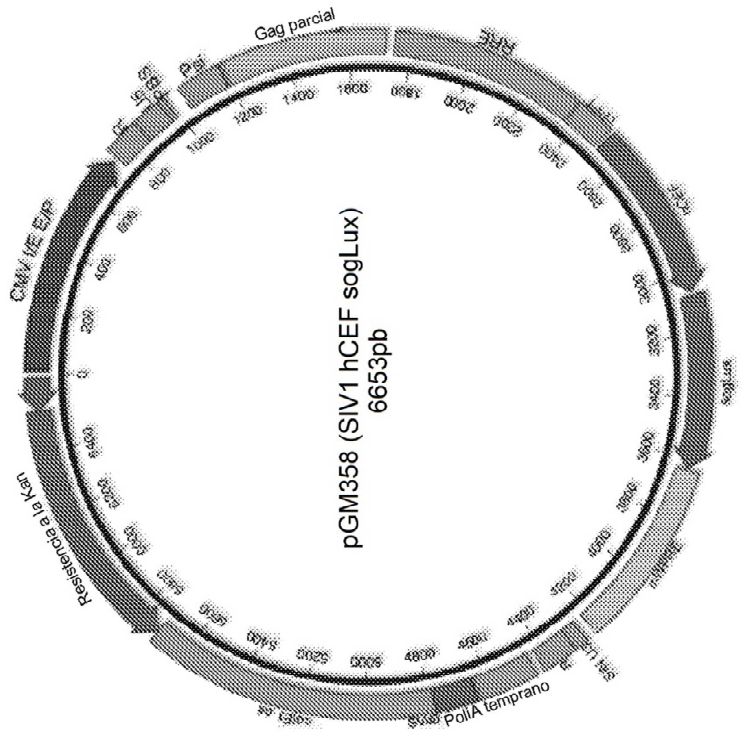
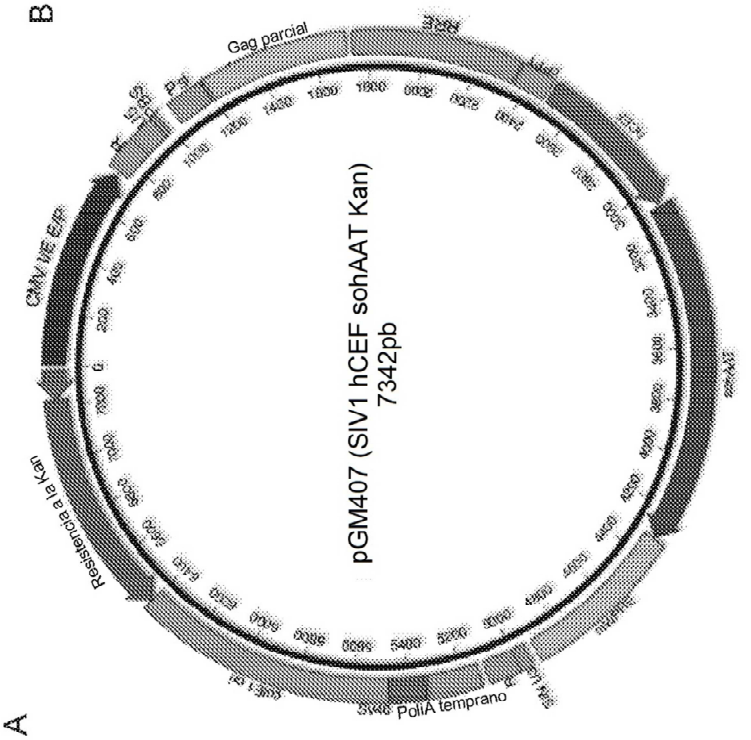

Hinchazón (AUC t=120min)
8000
6000
4000
2000
0
Control neg.
VGM58
Control pos.
***

Figura 14


pGM358 (SIV1 hCEF sogLux)
6653pb
Gag parcial
RRE
Rev
Tat
Vif
Vpr
Tat
Rev
Vpu
Env
Nef
PolIA temprano
Resistencia a la Kan
CMV IE E/P
Pst
Sal
Xho
Not


pGM407 (SIV1 hCEF sohAAT Kan)
7342pb
Gag parcial
RRE
Rev
Tat
Vif
Vpr
Tat
Rev
Vpu
Env
Nef
PolIA temprano
Resistencia a la Kan
CMV IE E/P
Pst
Sal
Xho
Not

B

A

Figure 15

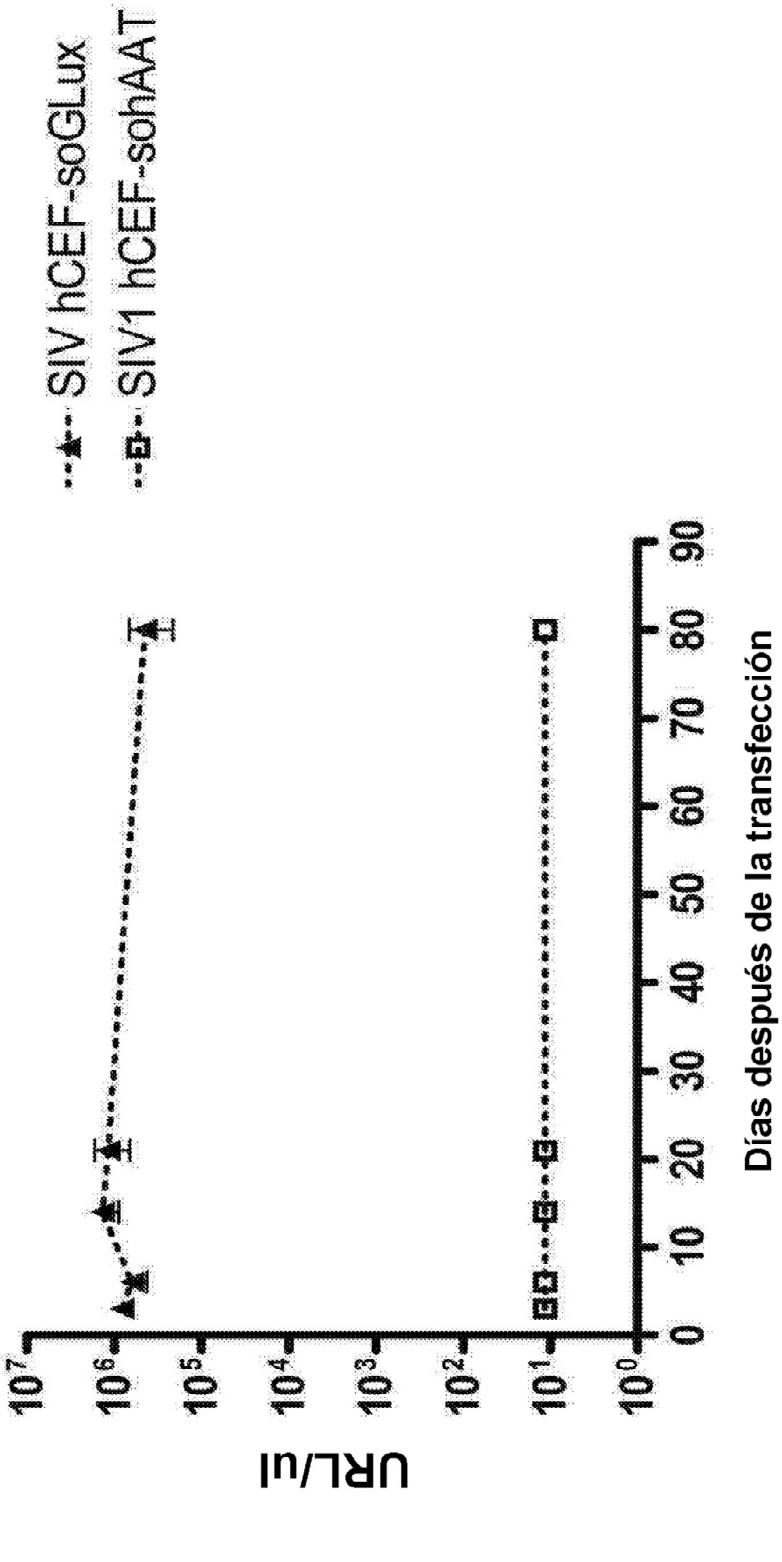
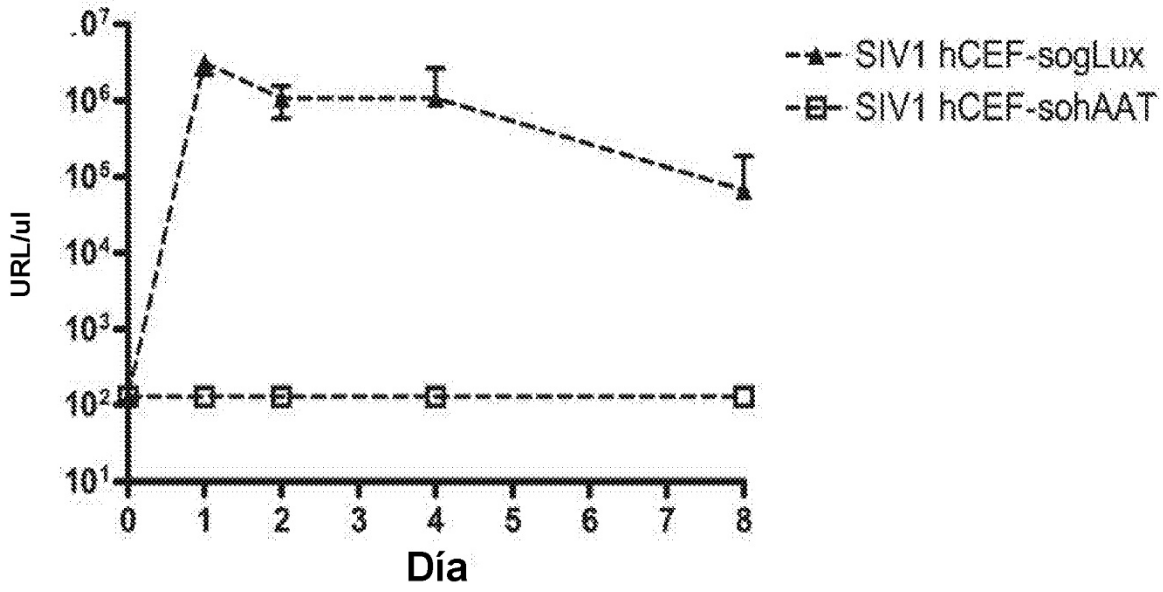

SIV hCEF-soGLux
SIV1 hCEF-sohAAT
URL/ul
10^7
10^6
10^5
10^4
10^3
10^2
10^1
10^0
0
10
20
30
40
50
60
70
80
90
Días después de la transfección

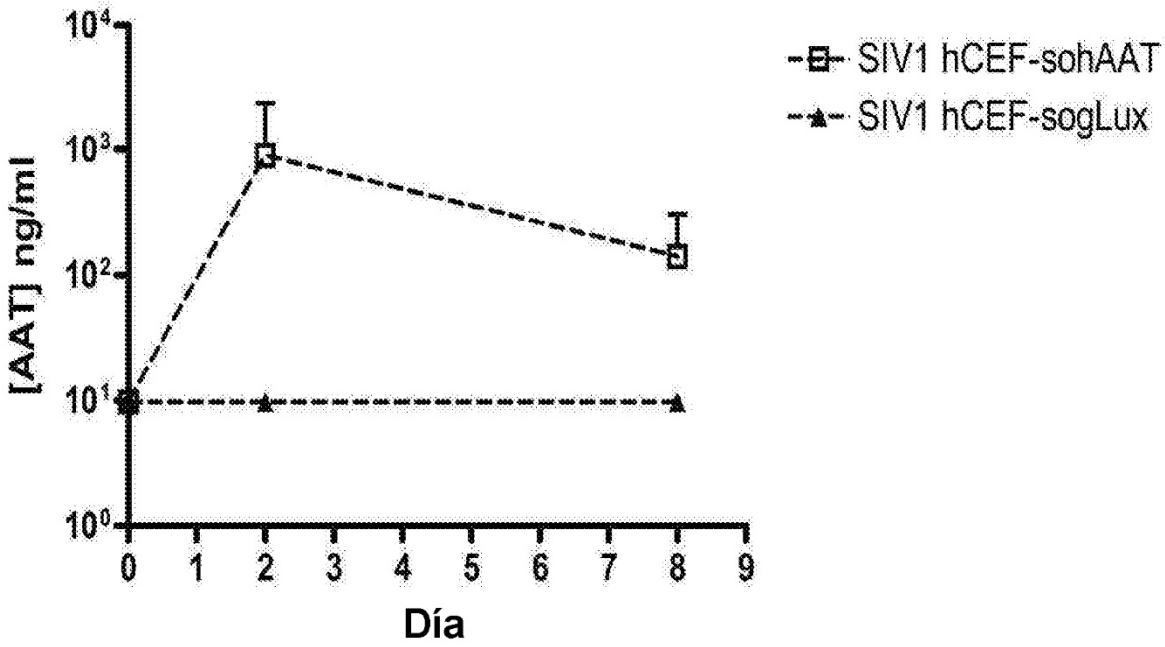
Figura 16

Figura 17

A


SIV1 hCEF-sogLux
SIV1 hCEF-sohAAT
URL/ul
10^7
10^6
10^5
10^4
10^3
10^2
10^1
0 1 2 3 4 5 6 7 8
Día

B


SIV1 hCEF-sohAAT
SIV1 hCEF-sogLux
[AAT] ng/ml
10^4
10^3
10^2
10^1
10^0
0 1 2 3 4 5 6 7 8 9
Día

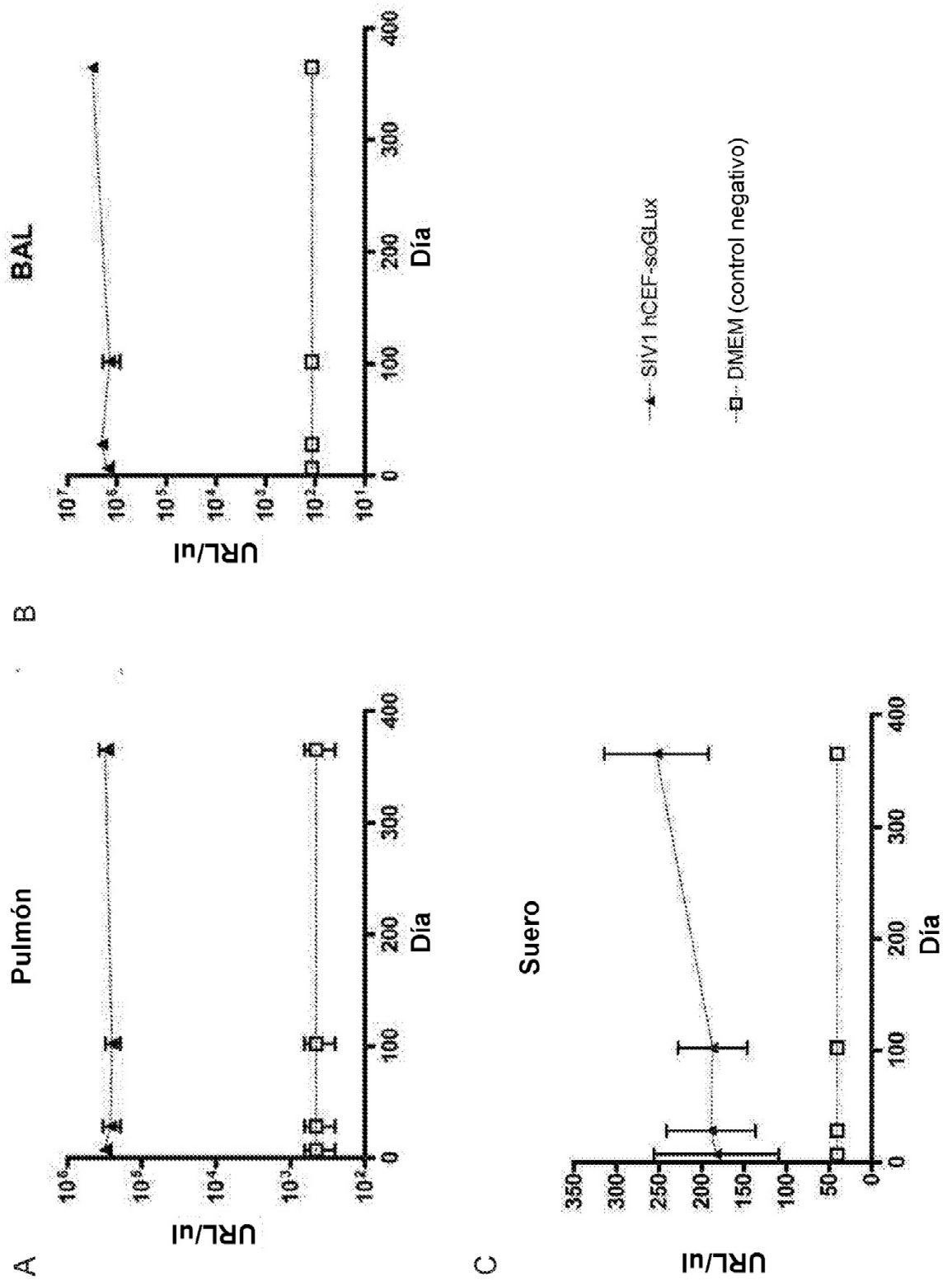
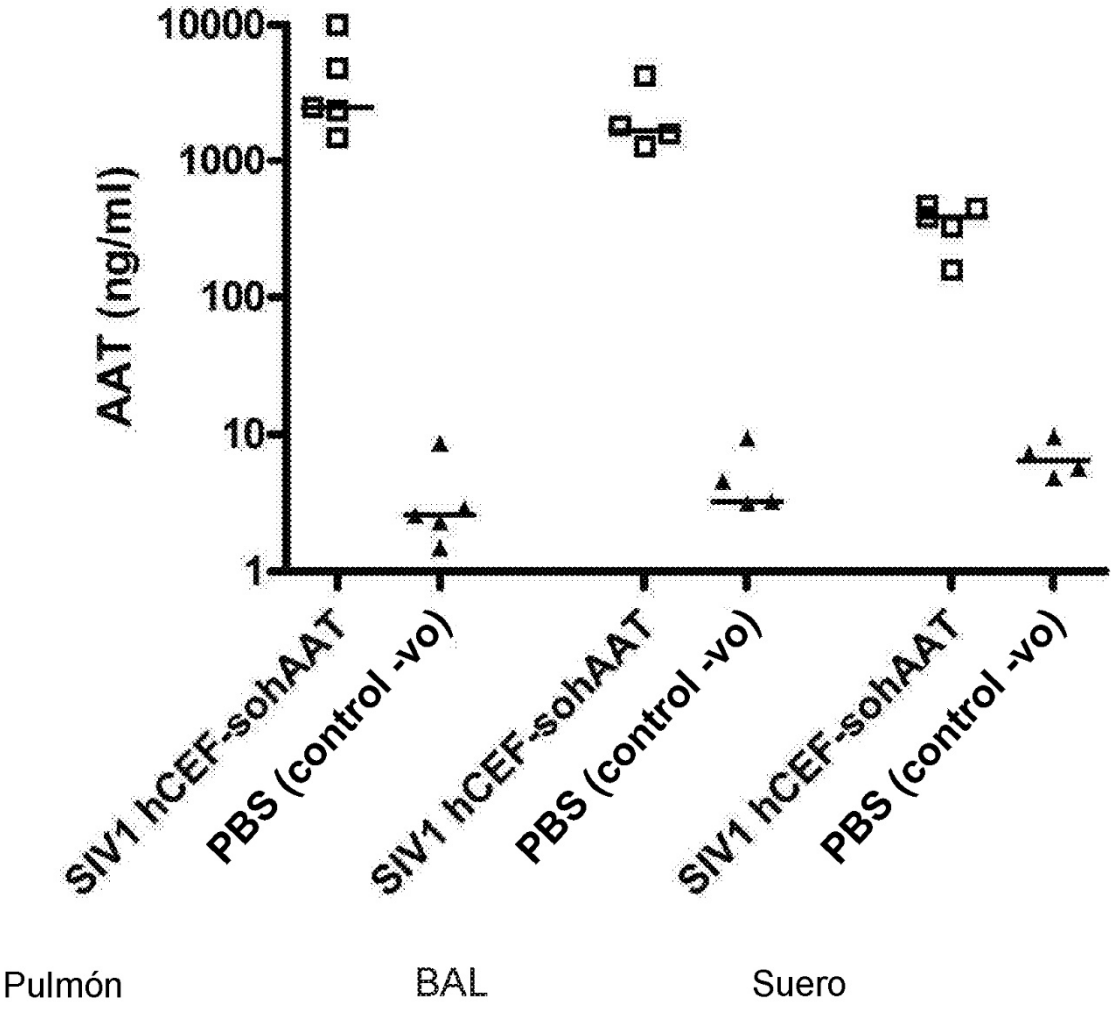

A
Pulmón
URL/µl
Día
B
BAL
URL/µl
Día
C
Suero
URL/µl
Día
SIV1 hCEF-soGLux
DMEM (control negativo)

Figura 18

Figura 19


10000
1000
100
10
1
AAT (ng/ml)
SIV1 hCEF-sohAAT
PBS (control -vo)
SIV1 hCEF-sohAAT
PBS (control -vo)
SIV1 hCEF-sohAAT
PBS (control -vo)
Pulmón
BAL
Suero

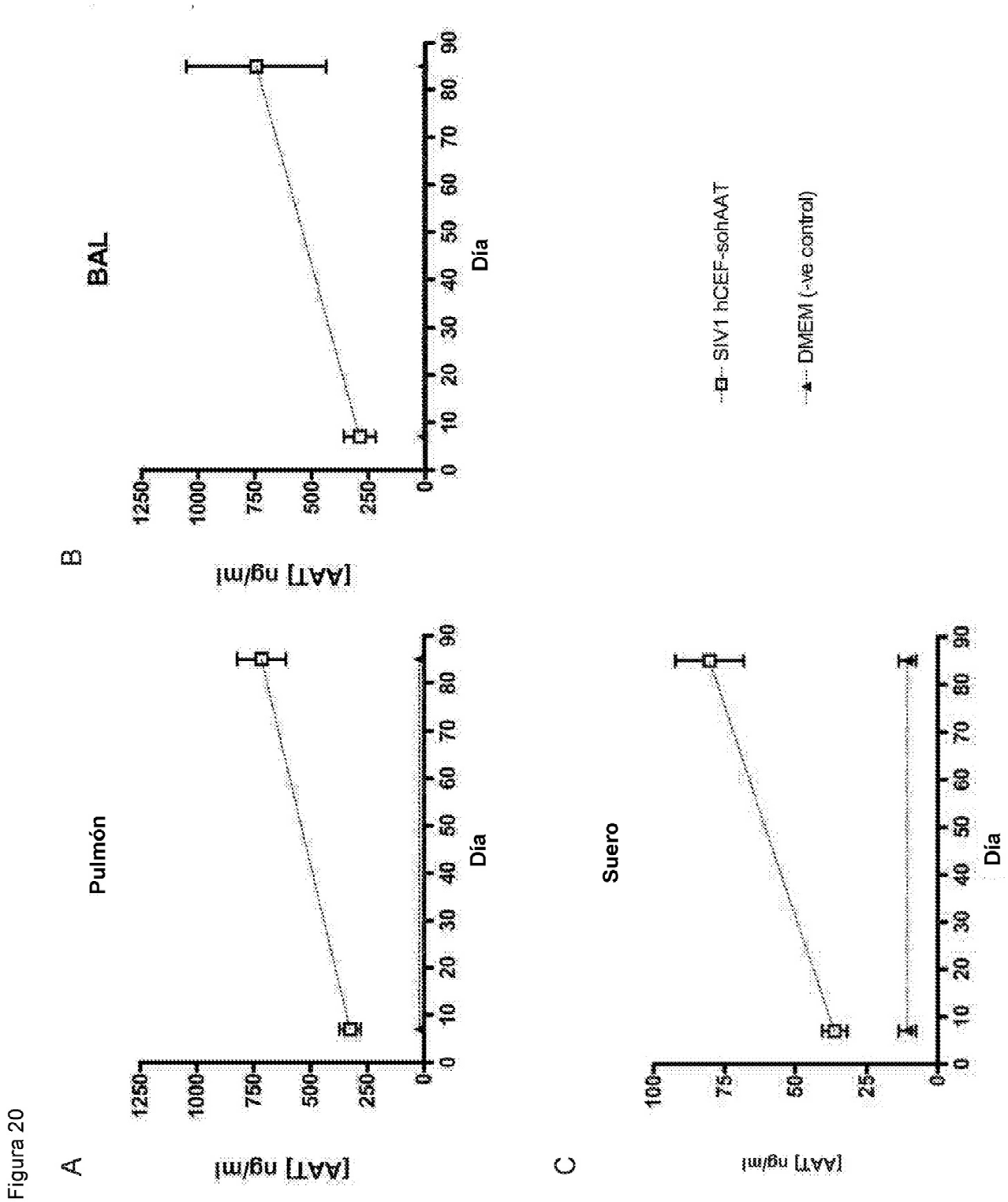

A
Pulmón
[AAT] ng/ml
Día
B
BAL
[AAT] ng/ml
Día
C
Suero
[AAT] ng/ml
Día
SIV1 hCEF-sohAAT
DMEM (-ve control)

Figura 20

Figura 21

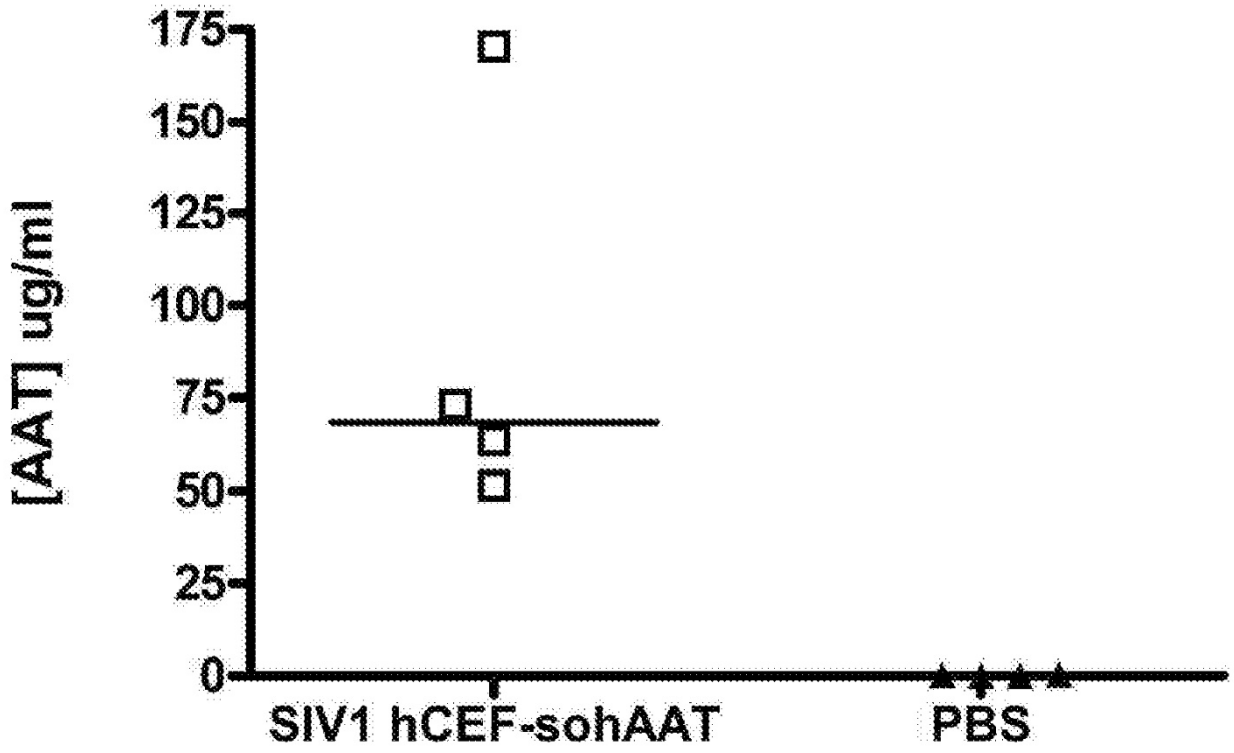
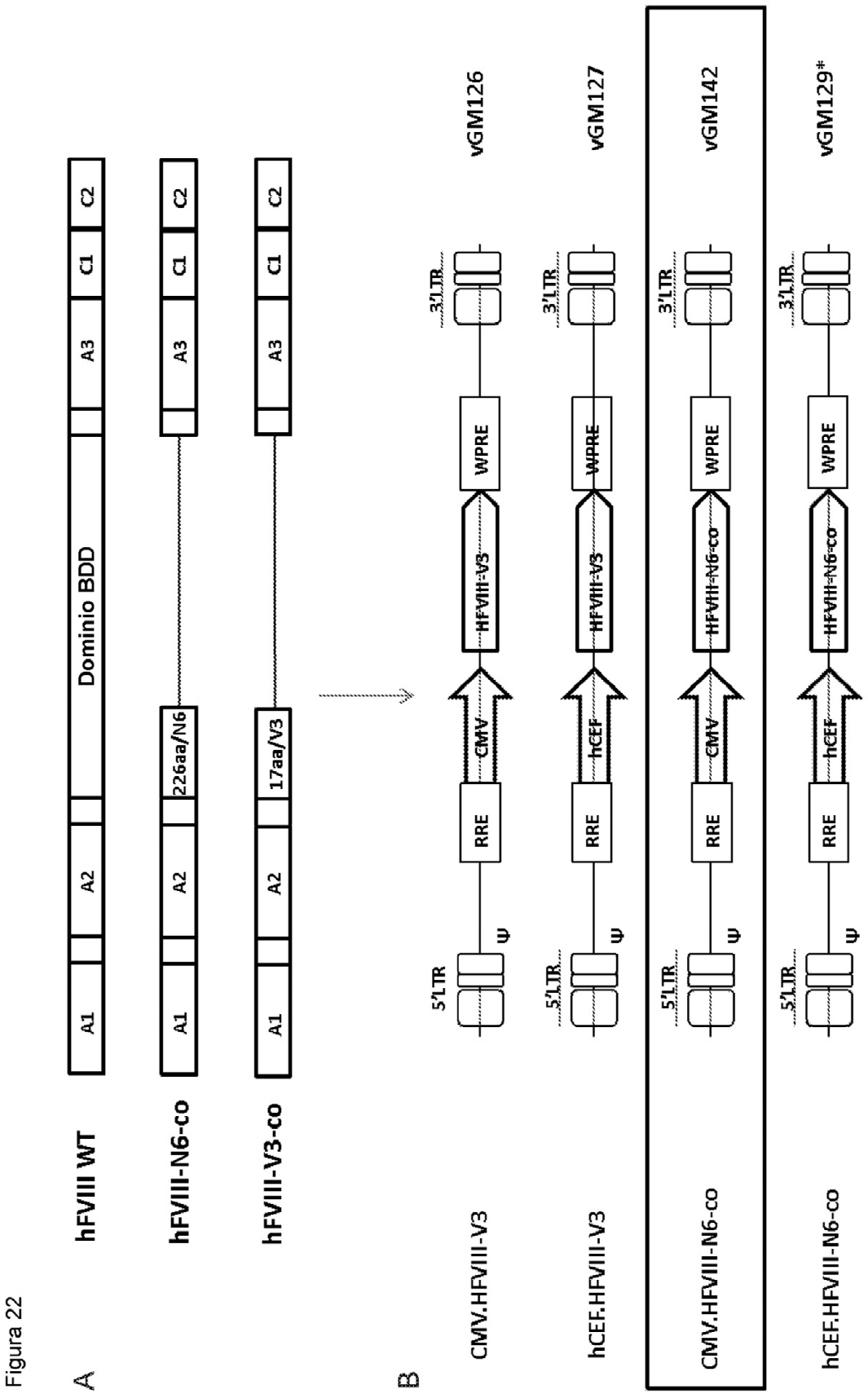


Figura 22


A
hFVIII WT
A1
A2
Dominio BDD
A3
C1
C2
hFVIII-N6-co
226aa/N6
hFVIII-V3-co
17aa/V3
B
CMV.HFVIII-V3
hCEF.HFVIII-V3
CMV.HFVIII-N6-co
hCEF.HFVIII-N6-co
5'LTR
Ψ
RRE
CMV
hCEF
HFVIII-V3
HFVIII-N6-co
WPRE
3'LTR
vGM126
vGM127
vGM142
vGM129*

C

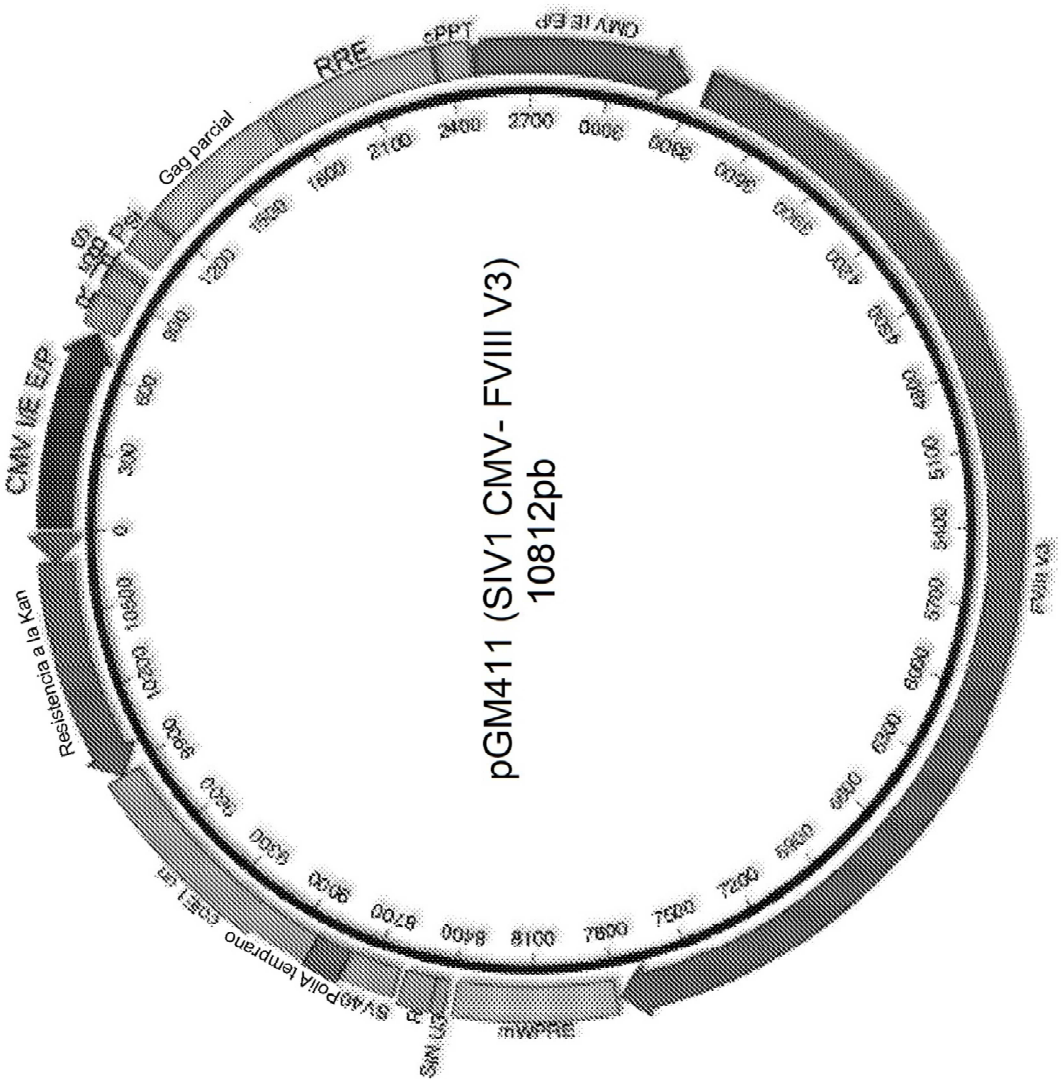


Figura 22 (continuación)

D

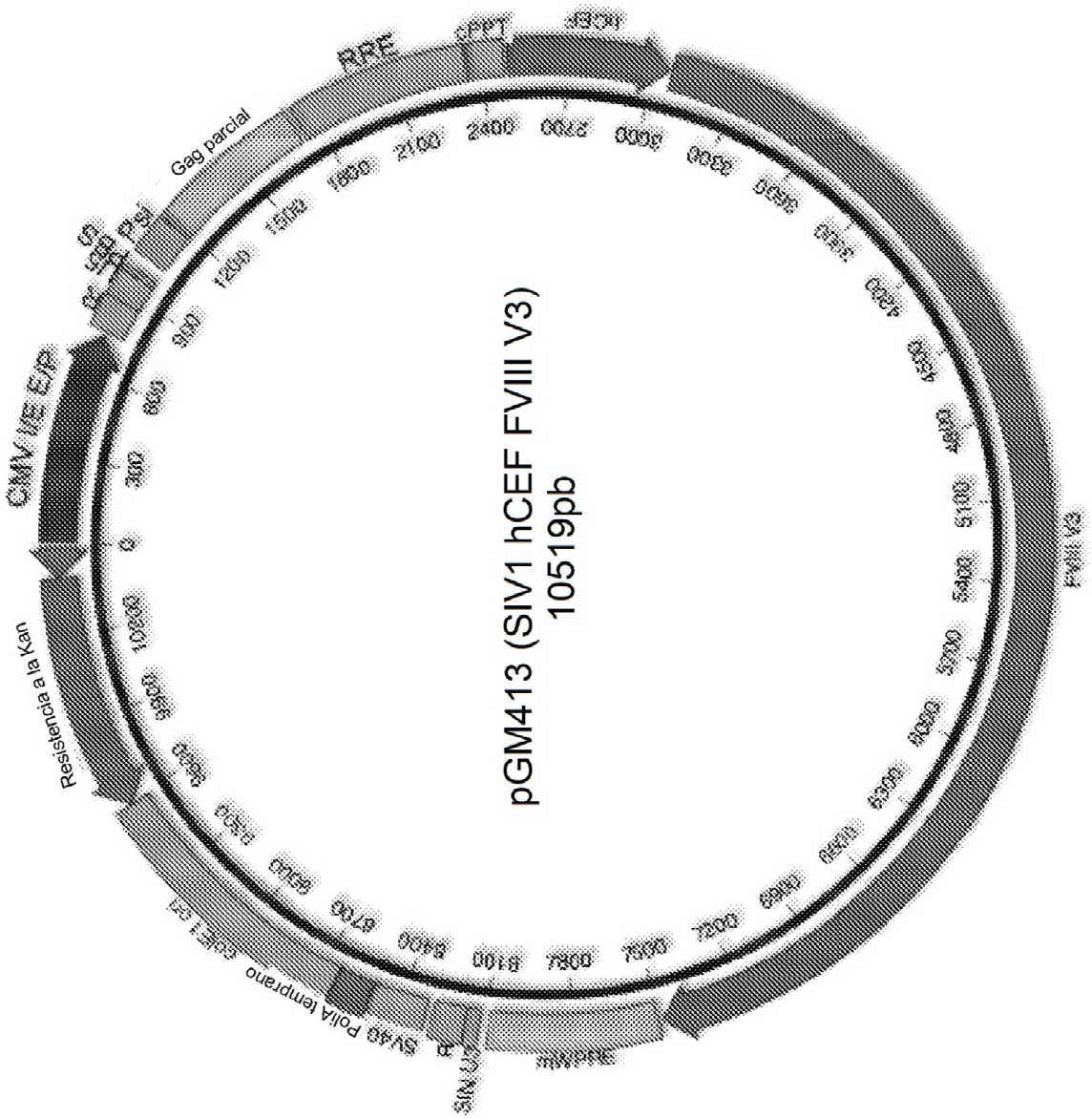
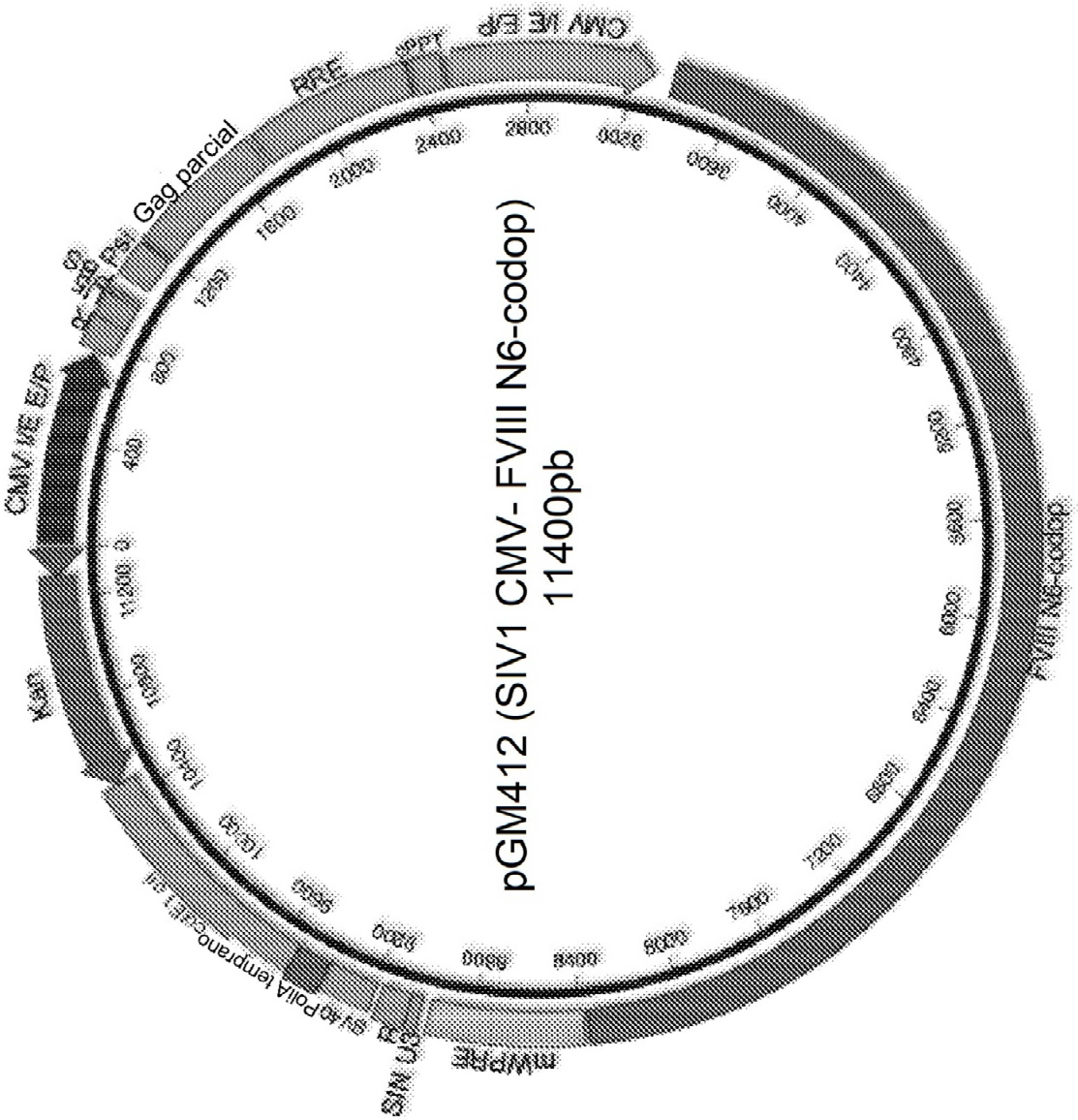
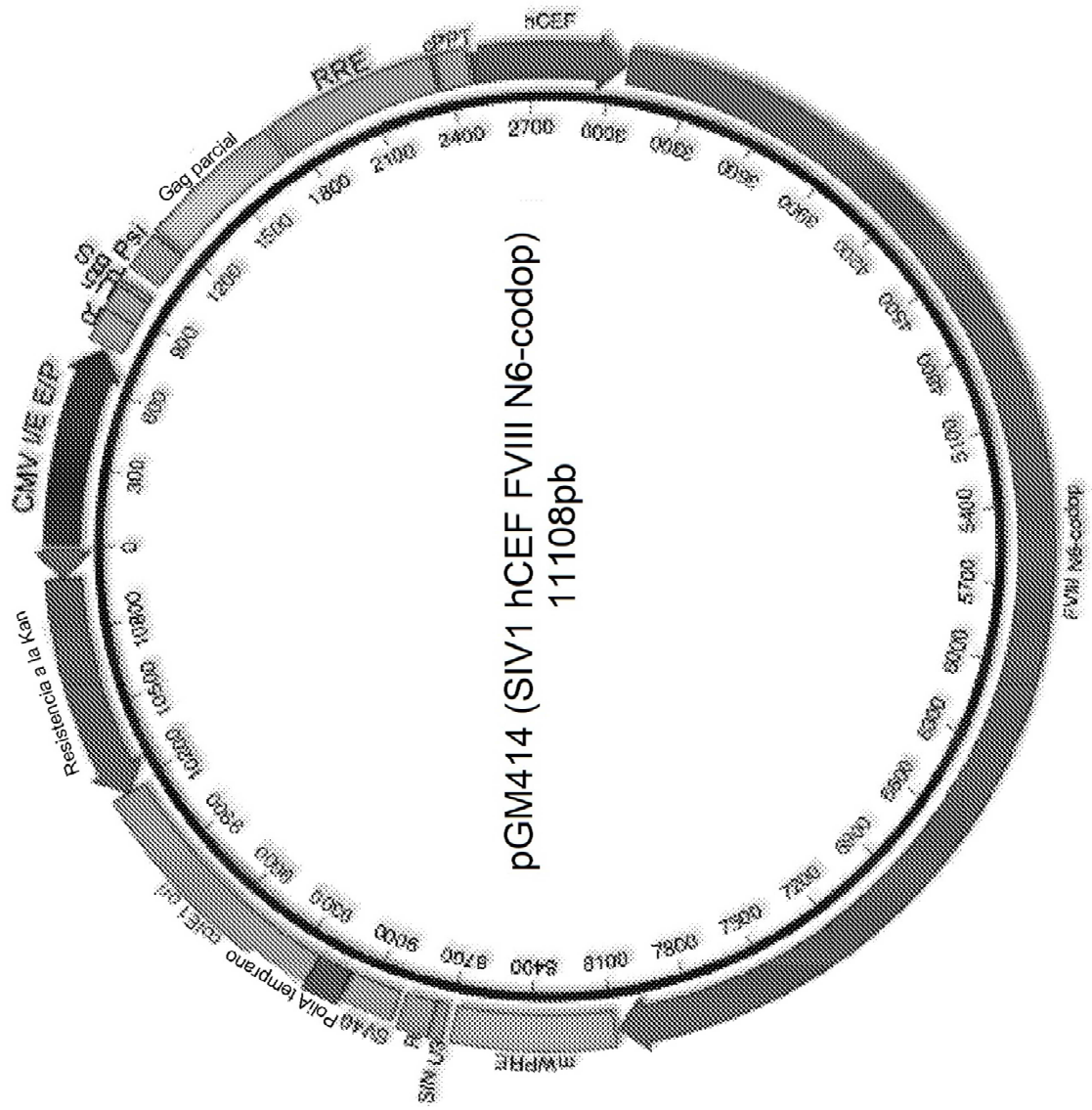


Figura 22 (continuación)



E

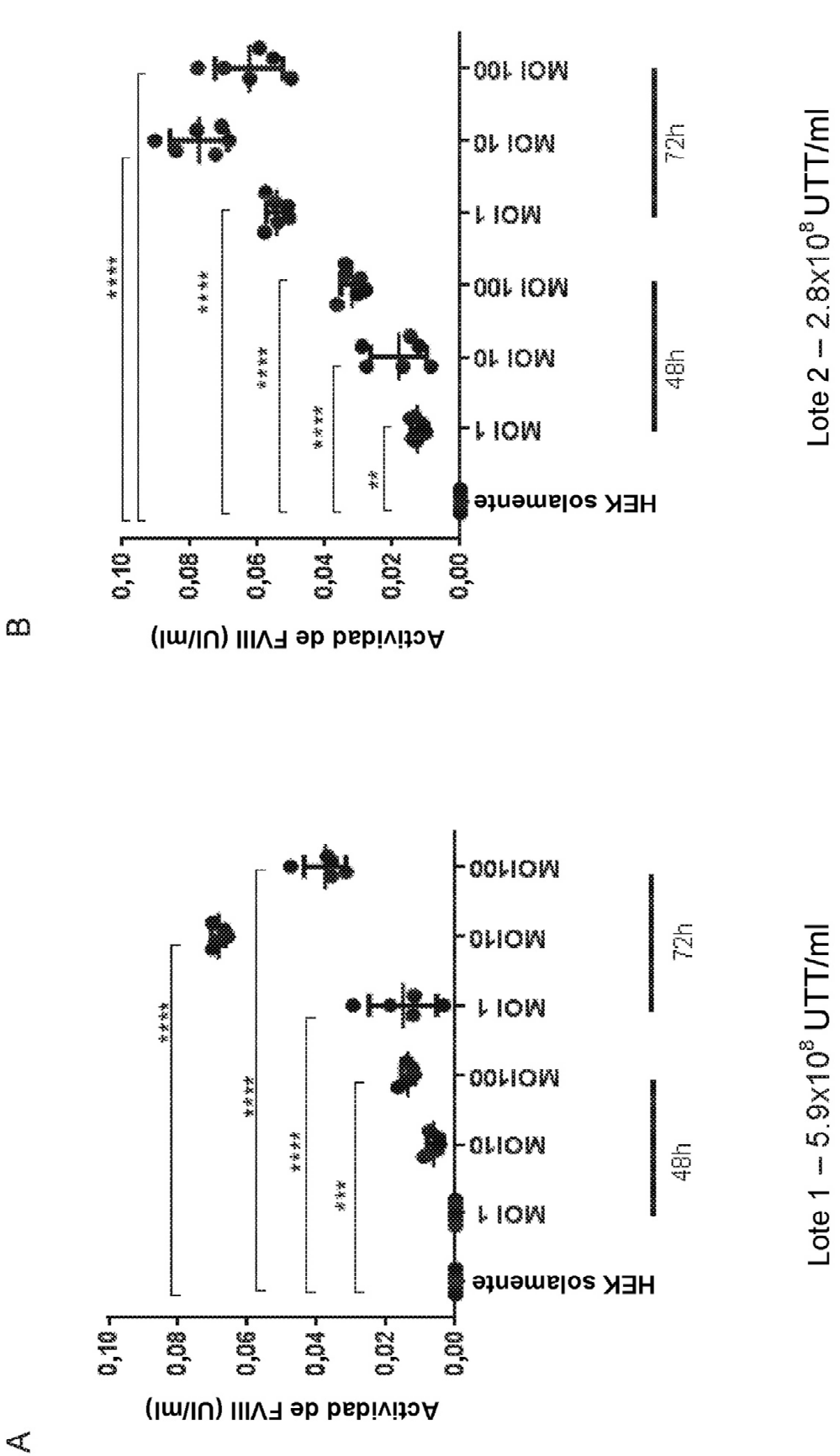
Figura 22 (continuación)



F

Figura 22 (continuación)

Figura 23


A
Actividad de FVIII (UI/ml)
0,10
0,08
0,06
0,04
0,02
0,00
HEK solamente
MOI 1
MOI10
MOI100
MOI 1
MOI10
MOI100
48h
72h
***
****
****
****
Lote 1 – 5.9x10^8 UTT/ml
B
Actividad de FVIII (UI/ml)
0,10
0,08
0,06
0,04
0,02
0,00
HEK solamente
MOI 1
MOI 10
MOI 100
MOI 1
MOI 10
MOI 100
48h
72h
**
****
****
****
Lote 2 – 2.8x10^8 UTT/ml

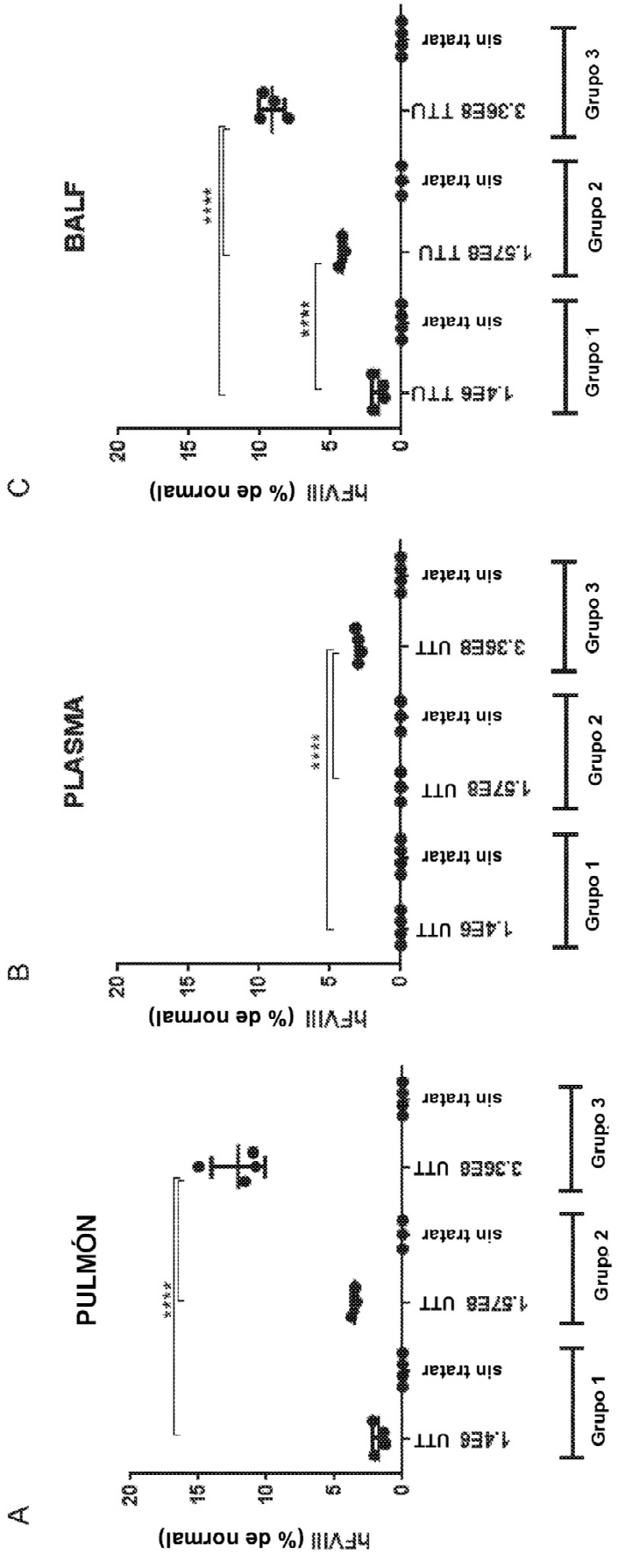

A
PULMÓN
hFVIII (% de normal)
20
15
10
5
0
****
1.4E6 UTT
sin tratar
1.57E8 UTT
sin tratar
3.36E8 UTT
sin tratar
Grupo 1
Grupo 2
Grupo 3
B
PLASMA
hFVIII (% de normal)
20
15
10
5
0
****
1.4E6 UTT
sin tratar
1.57E8 UTT
sin tratar
3.36E8 UTT
sin tratar
Grupo 1
Grupo 2
Grupo 3
C
BALF
hFVIII (% de normal)
20
15
10
5
0
****
****
1.4E6 TTU
sin tratar
1.57E8 TTU
sin tratar
3.36E8 TTU
sin tratar
Grupo 1
Grupo 2
Grupo 3

Figura 24