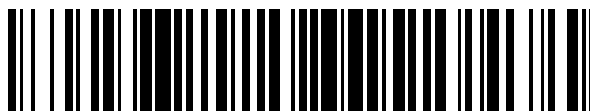


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 452 890**

51 Int. Cl.:

A61K 38/37 (2006.01)

C07K 14/755 (2006.01)

C12N 15/85 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.07.2010 E 10731905 (5)**

97 Fecha y número de publicación de la concesión europea: **18.12.2013 EP 2451474**

54 Título: **Variantes de factor VIII optimizadas mediante codones y promotor sintético específico para el hígado**

30 Prioridad:

08.07.2009 GB 0911870

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

03.04.2014

73 Titular/es:

**UCL BUSINESS PLC (33.3%)
The Network Building 97 Tottenham Court Road
London
Greater London W1T 4TP, GB;
THROMBOSIS RESEARCH INSTITUTE (33.3%) y
ST. JUDE CHILDREN'S RESEARCH HOSPITAL
(33.3%)**

72 Inventor/es:

**NATHWANI, AMIT;
WARD, NATALIE;
THRASHER, ADRIAN;
TUDDENHAM, EDWARD;
MCVEY, JOHN;
GRAY, JOHN y
DAVIDOFF, ANDREW**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 452 890 T3

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

DESCRIPCIÓN

Variantes de factor VIII optimizadas mediante codones y promotor sintético específico para el hígado

5 **Campo de la invención**

La invención se refiere a una secuencia codificadora optimizada de factor ocho (VIII) de coagulación sanguínea humana. Además, la invención se refiere a vectores como rAAV para la introducción de la secuencia de factor VIII optimizado en células para llevar a cabo su expresión. La invención se refiere también a células transformadas con estos vectores, animales transgénicamente modificados que contienen células modificadas usando los vectores y usos terapéuticos que utilizan los vectores en una propuesta de sustitución génica.

Antecedentes de la invención

15 Los inventores están interesados en el desarrollo de una estrategia de transferencia génica segura y eficaz para el tratamiento de la hemofilia A (HA), el trastorno sanguíneo heredado más común. Esto representaría un avance clínico considerable con implicaciones significativas para otros trastornos congénitos que carecen de tratamiento eficaz. Los inventores ya han desarrollado una propuesta alentadora de terapia génica para la hemofilia B usando vectores virales recombinantes adeno-asociados (rAAV). La hemofilia A plantea varios retos nuevos debido a las propiedades moleculares y bioquímicas distintas del factor VIII humano (hFVIII), una molécula que es mutada en esta enfermedad. Estas incluyen el tamaño relativamente grande del hFVIII cDNA y el hecho de que la expresión de proteínas de hFVIII es altamente ineficaz. Los inventores han comenzado a abordar algunas de estas limitaciones a través de avances de la tecnología de vectores y el desarrollo de una variante de hFVIII nueva más potente (codop-hFVIII) que puede ser eficazmente envasada en rAAV.

25 La hemofilia A (HA) es un trastorno sanguíneo conectado a x que afecta aproximadamente a uno de cada uno 5.000 varones, que está provocada por mutaciones en el gen de factor VIII (FVIII), que codifica un cofactor esencial en la cascada de coagulación. Los pacientes de HA grave (>50% de los pacientes) tienen menos de 1% de la actividad normal de FVIII y padecen de hemorragias espontáneas en articulaciones que soportan peso y tejidos blandos, que provocan una incapacidad permanente y ocasionalmente la muerte. El patrón habitual para el cuidado de la HA consiste en concentrados de proteína de hFVIII recombinante, que puede detener la hemorragia pero que no evita el deterioro crónico que surge después de una hemorragia. La administración profiláctica de concentrados de factor para mantener los niveles de FVIII en plasma por encima de 1% (>2 ng/ml) conduce a una reducción considerable de la hemorragia espontánea y la artropatía crónica. Sin embargo, la semivida del FVIII es corta (8-12 horas) necesitando tres administraciones intravenosas de concentrados por semana. Esto es prohibitivamente caro (>100 mil libras esterlinas/año/paciente), es altamente invasiva y requiere tiempo. Debido al coste y el suministro limitado, por encima de un 75% de los pacientes de HA grave no reciben ningún tratamiento o solamente esporádicos con concentrados de FVIII. Estos individuos se enfrentan a una vida enormemente acortada con dolor e incapacidad.

40 Por lo tanto, se ha prestado atención a una terapia génica somática para HA debido a su capacidad potencial para una curación a través de la producción endógena continua de FVIII a continuación de una administración única de vector. La hemofilia A está de hecho bien adecuada para una propuesta de sustitución génica porque sus manifestaciones clínicas son completamente atribuibles a la falta de un producto de gen único (FVIII) y circula en cantidades diminutas (200 ng/ml) en el plasma. No es esencial un control estrictamente regulada de la expresión génica y un aumento modesto en el nivel de FVIII (>1% del normal) puede mejorar el fenotipo grave. La disponibilidad de modelos animales que incluyen ratones desprovistos de FVIII y perros con hemofilia A puede facilitar una evaluación preclínica extensiva de estrategias de terapias génicas. Finalmente, las consecuencias de la transferencia de genes puede ser valorada usando una cuantificación simple en lugar de puntos finales cualitativos, que puede ser fácilmente ensayada en la mayoría de los laboratorios clínicos, lo que contrasta con otros objetivos de terapias génicas en los que la expresión es difícil de valorar o está influenciada por factores adicionales como el flujo de sustrato.

55 Han sido realizado tres ensayos de fase I de transferencia génica hasta ahora para la HA usando un suministro de genes *in vivo* directo de vectores onco-retro-o adeno-virales así como un trasplante autólogo de fibroblastos autólogos modificados mediante plásmidos. No se consiguió la expresión estable de hFVIII por encima de 1%. Estos estudios preclínicos y los posteriores resaltaron los obstáculos biológicos críticos graves para una terapia génica satisfactoria de la HA.

60 El tratamiento celular de la molécula de FVIII de longitud completa de tipo salvaje es altamente complejo y la expresión se confunde con la inestabilidad de mRNA, la interacción con acompañantes del retículo endoplásmico (ER), y una necesidad de un transporte facilitado desde el ER hasta el aparato de Golgi a través de una interacción con lectina LMAN1 de unión a manosa. Sin embargo, han sido desarrolladas variantes de FVIII más potentes y nuevas a través de avances progresivos en la comprensión de la biología de la expresión de FVIII. Por ejemplo, los estudios bioquímicos demostraron que el dominio B de FVIII era suministrable para la actividad del cofactor VIII. La supresión del dominio B dio lugar a un aumento de 17 veces en los niveles de mRNA sobre el FVIII de tipo salvaje y de longitud completa y un aumento de 30% en la proteína secretada. Esto condujo al desarrollo de concentrado de

proteína FVIII desprovista de dominio B (BDD), que es ampliamente usada en la actualidad en el campo clínico. Sin embargo, unos estudios recientes indican que el hFVIII de longitud completa y BDD se pliegan mal en el Lumen del ER, dando lugar a la activación de la respuesta de proteínas no plegadas (UPR) y la apoptosis de hepatocitos de mórvidos. Sin embargo, la adición de un separador de dominios B corto, con elevado contenido de oligosacáridos conectados a asparagina, a BDD-FVIII (=N6-FVIII) resuelve este problema en parte a través de un transporte mejorado desde el ER hasta el aparato de Golgi. El N6-FVIII es secretado a niveles 100 veces superiores al FVIII de tipo salvaje y de longitud completa, pero los inventores creen que la secreción de FVIII puede ser mejorada adicionalmente a través de la modificación del genoma de FVIII.

El rAAV muestra habitualmente las mejores previsiones para los trastornos crónicos como la HA debido a su excelente perfil de seguridad. Además, los inventores y otros han mostrado que una administración única de vector rAAV es suficiente para dirigir la expresión transgénica a largo plazo sin una toxicidad significativa en una diversidad de modelos animales que incluyen primates no humanos. La integración del provirus rAAV ha sido descrita, pero a una frecuencia que es excesivamente baja y comparable a la del DNA de plásmido. Por lo tanto, la expresión transgénica estable está mediada principalmente por genomas de rAAV episomalmente retenidos en los tejidos post-mitóticos, reduciendo así el riesgo de una oncogénesis insercional. Esto contrasta con los vectores integrantes que se ha mostrado que provocan un trastorno linfoproliferativo en niños con SCID-XI. Aunque recientemente se ha informado de resultados alentadores en pacientes que sufren enfermedad de Parkinson y amaurosis congénita de Lever a continuación de una transferencia génica mediada por rAAV, hasta recientemente el tamaño grande del hFVIII c de NA (-7 Kb), que sobrepasa la capacidad de envasado relativamente pequeña de rAAV de ~ 4,7 a 4,9 Kb, a limitado el uso de este vector para la HA. Un informe reciente de Pierce y colaboradores demostró la expresión a largo plazo (>4 años) de FVIII canino con supresión de dominio B (BDD) a un 2,5-5% de lo normal a continuación de una administración única de rAAV en perros con hemofilia A. Sin embargo, la expresión mediada por rAAV de FVIII humano no ha sido establecida hasta el mismo grado.

Actualmente, la complicación más grave y comprometedora del tratamiento con concentrados de FVIII es el desarrollo de anticuerpos neutralizantes para FVIII (inhibidores de FVIII) que se produce en hasta un 30% de pacientes con HA. Estos inhibidores anulan los efectos biológicos de los concentrados de FVIII y hacen que sea difícil tratar episodios de hemorragias, excepto con agentes de derivación como factor VII activado recombinante (rFVIIa). El coste significativo del rFVIIa (~ 500.000 libras esterlinas por episodio de cirugía ortopédica) y la toxicidad (por ejemplo, trombosis) Excluyen el uso profiláctico. La inducción de la tolerancia inmune (ITI) es un objetivo alternativo, pero es menos eficaz en pacientes con inhibidores de titulación elevada de permanencia prolongada. De hecho, se ha conseguido una tolerancia periférica en algunos pacientes con inhibidores de FVIII irresolubles a continuación de un trasplante de hígado, sugiriendo que la expresión endógena a largo plazo y estable de hFVIII puede ser importante para conseguir la tolerancia. Los datos de los inventores en ratones y primates no humanos y los de otros muestran que la transferencia génica dirigida al hígado con rAAV favorecen un estado de tolerancia permanente hacia el transgen a través de la expresión de células T reguladoras específicas de transgenes (Tregs).

Por lo tanto, la transferencia génica puede proporcionar un medio alternativo para la prevención y erradicación de inhibidores irresolubles.

Una lección clave de los ensayos clínicos previos con rAAV es que los estudios clínicos previos necesitan ser evaluados en un contexto relevante para seres humanos. Por lo tanto, han sido enfocados en primates no humanos para la evaluación de vectores rAAV porque, como en seres humanos, los macacos son hospedantes naturales para una infección de AAV. Esto proporciona una oportunidad importante para evaluar la eficacia de la transferencia génica con vectores de rAAV en animales de raza no pura previamente sensibilizados con AAV de tipo salvaje, lo que no es posible con modelos de mórvidos o caninos. Finalmente, las autoridades reguladoras en Europa y los estados Unidos están requiriendo ahora estudios preclínicos de seguridad y eficacia en primates no humanos como una condición para la autorización de un ensayo clínico.

Para superar las desventajas anteriormente mencionadas, los inventores han creado una secuencia mejorada de nucleótidos aislados que codifican el factor VIII junto con un nuevo promotor.

Sumario de la invención

En un primer aspecto, la presente invención proporciona una molécula aislada de ácido nucleico que comprende una secuencia de nucleótidos que tiene una identidad de al menos 90% respecto a la secuencia de nucleótidos de la SEC ID N° 1 y que codifica factor VIII funcional (fVIII o FVIII).

La presente invención proporciona también un vector que comprende una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene una identidad de al menos 90% respecto a la secuencia de nucleótidos de la SEC ID N° 1 y que codifica factor VIII funcional.

Además, la presente invención proporciona un célula hospedante que comprende una molécula de ácido nucleico como se describió anteriormente o un vector como se describió anteriormente.

Adicionalmente, la presente invención proporciona un animal transgénico que comprende células que comprenden una molécula de ácido nucleico o un vector como se describió anteriormente.

5 Además, la presente invención proporciona una molécula de ácido nucleico como se describió anteriormente o un vector como se describió anteriormente para un uso en terapia.

Adicionalmente, la presente invención proporciona una molécula de ácido nucleico como se describió anteriormente o un vector como se describió anteriormente para un uso en el tratamiento de hemofilia.

10 La molécula de ácido nucleico anteriormente descrita puede comprender adicionalmente una secuencia de nucleótidos que tiene una homología de al menos 85% respecto a la secuencia de nucleótidos de la SEC ID N° 3. Esta secuencia de nucleótidos es un promotor.

15 La presente invención proporciona también un método para la preparación de un vector de suministro de genes parvovirales, comprendiendo el método las etapas de:

(a) proporcionar una célula de insecto que comprende uno o más constructos de ácidos nucleicos que comprenden:

20 (i) una molécula de ácido nucleico como se describió anteriormente que está flanqueada por al menos una secuencia de nucleótidos repetida, terminal, invertida y parvoviral;

(ii) un primer cassette de expresión que comprende una secuencia de nucleótidos que codifica una o más proteínas Rep parvovirales que está funcionalmente conectado a un promotor que es capaz de conducir la expresión de la(s) proteína(s) Rep en la célula de insecto;

25 (iii) un segundo cassette de expresión que comprende una secuencia de nucleótidos que codifica una o más proteínas de la capsid que está funcionalmente conectado a un promotor que es capaz de conducir la expresión de la(s) proteína(s) de la cápside en la célula de insecto;

30 (b) cultivar la célula de insecto definida en el apartado (a) bajo condiciones que conduzcan a la expresión de la Rep y las proteínas de la cápside; y, opcionalmente,

(c) recuperar el vector de suministro de gen parvoviral.

35 **Descripción detallada de la invención**

Según un aspecto de la invención, se proporciona una molécula aislada de ácido nucleico que comprende una secuencia de nucleótidos que tiene una homología sustancial respecto a la secuencia de nucleótidos de la SEC ID N° 1. La expresión "homología sustancial" puede ser adicionalmente definida con respecto a una homología porcentual. Esto se expone más en detalle con posterioridad pero, de forma breve, la secuencia de nucleótidos tiene una homología de al menos 90% (también definida como identidad) respecto a la secuencia de nucleótidos de la SEC ID N° 1.

45 El término "aislada", cuando se usa en relación con una molécula de ácidos nucleicos de la invención, se refiere normalmente a una secuencia de ácidos nucleicos que está identificada y separa de al menos un ácido nucleico contaminante que está ordinariamente asociado en su fuente natural. El ácido nucleico aislado puede estar presente en una forma o ajuste que es diferente del que se encuentra en la naturaleza. Por el contrario, los ácidos nucleicos no aislados son ácidos nucleicos como DNA y RNA que se encuentran en el estado que existe en la naturaleza. Por ejemplo, una secuencia de DNA dada (por ejemplo, un gen) se encuentra en el cromosoma de células hospedantes en las proximidades de genes vecinos; las secuencias de RNA, como una secuencia de mRNA específica que codifica una proteína específica, se encuentran en la célula en forma de una mezcla con otros numerosos mRNAs que codifican una multitud de proteínas. La molécula de ácidos nucleicos aislada de la invención puede estar presente en una forma de cadena única o de cadena doble. Cuando una molécula de ácidos nucleicos aislada va a ser utilizada para expresar una proteína, normalmente contendrá como mínimo la cadena de sentido o codificadora (es decir, la molécula de ácidos nucleicos puede ser de cadena única), pero puede contener las cadenas tanto de sentido como de anti-sentido (es decir, la molécula de ácidos nucleicos puede ser de cadena doble).

60 La molécula de ácidos nucleicos de la invención tiene una homología de al menos 90% y, más preferentemente, al menos 95%, por ejemplo una homología de al menos 98% respecto a la secuencia de nucleótidos de la SEC ID N° 1. Preferentemente tiene también una homología de al menos 70%, más preferentemente al menos 75% y más preferentemente al menos 80% respecto al factor VIII de tipo salvaje. Además, la molécula de ácidos nucleicos codifica una proteína de factor VIII funcional, es decir, codifica el cofactor VIII que cuando es expresado tiene la funcionalidad del factor VIII de tipo salvaje. La Molécula de ácidos nucleicos, cuando es expresada en un sistema adecuado (por ejemplo, una célula hospedante), produce una proteína de factor VIII funcional a un nivel relativamente elevado. Como el factor VIII que es producido es funcional, tendrá una conformación que es la misma que al menos una parte del factor VIII de tipo salvaje. En una realización, el factor VIII producido mediante el ácido

- nucleico tendrá la misma conformación que el factor VIII N6 que ha sido previamente descrito. Una proteína de factor VIII funcional producida mediante la invención permite que tenga lugar al menos algo de coagulación de sangre en un sujeto. Esto provoca una disminución en el tiempo que tarda la sangre en coagular en un sujeto que padezca de hemofilia. El factor VIII normal participa en la coagulación sanguínea a través de la cascada de coagulación. El factor VIII normal es un cofactor para el factor IXa que, en presencia de Ca^{+2} y fosfolípidos forma un complejo que convierte el factor X en la forma Xa activada. Por lo tanto, una proteína de factor VIII funcional según la invención puede formar un complejo funcional con el factor IXa que puede convertir el factor X en forma Xa activada.
- 5
- 10 Las secuencias de nucleótidos de factor XVIII previamente usadas han tenido problemas con la expresión de la proteína funcional. Esto se cree que es debido a un expresión ineficaz de mRNA, plegado erróneo de la proteína con posterior degradación intracelular y transporte ineficaz del producto de traducción primaria desde el retículo endoplásmico hasta el aparato de Golgi. Los inventores han encontrado que la molécula de ácidos nucleicos proporcionada por la invención provoca niveles sorprendentemente elevados de expresión de una proteína de factor
- 15 VIII tanto *in vitro* como *in vivo*. Esto significa que esta molécula de ácidos nucleicos podría ser usada en terapia génica para tratar hemofilia como la hemofilia A.
- La secuencia de nucleótidos de la SEC ID N° 1 es una secuencia de ácidos nucleicos de factor VIII humano optimizado mediante codones que está basada en la secuencia correspondiente a la secuencia de nucleótidos de factor VIII N6. La secuencia de nucleótidos de factor VIII N6 es una secuencia de factor VIII de la que el dominio B ha sido agotado y sustituido con un separador corto de dominio B, con elevado contenido de oligosacáridos conectados a asparagina, que mejora el transporte de N6-FVIII desde el ER hasta el aparato de Golgi.
- 20
- Los inventores han mostrado que la SEC ID N° 1 y las secuencias que son similares a la misma, es decir, las secuencias que tienen un nivel relativamente elevado de homología, muestran todas sorprendentemente unos niveles elevados de expresión de proteína funcional. A este respecto, las SEC ID N° 4, 5, 6 y 7 son también secuencias de ácidos nucleicos de factor VIII optimizadas mediante codones, cuyo % de homología es 88%, 77%, 82% y 97%, respectivamente, en comparación con la SEC ID N° 1.
- 25
- 30 Una secuencia de nucleótidos de la invención puede tener al menos aproximadamente 400, al menos aproximadamente 650, al menos aproximadamente 890, al menos aproximadamente 1140, al menos aproximadamente 1380, al menos aproximadamente 1530 de todos los codones que codifican el factor VIII funcional que es idéntico a los codones (en las posiciones correspondientes) en la SEC ID N° 1.
- 35 La secuencia de nucleótidos de la SEC ID N° 4 es una secuencia de ácidos nucleicos de factor VIII optimizada mediante codones que está basada en la secuencia de una secuencia de nucleótidos de factor VIII SQ N6. La secuencia de nucleótidos de factor VIII SQ N6 es una secuencia de factor VIII de la que el dominio B ha sido agotado y suprimido con una conexión SQ de 14 aminoácidos entre los dominios a2 y a3. En la conexión SQ, ha sido insertado un dominio N6 B.
- 40
- La secuencia de nucleótidos de la SEC ID N° 5 es una secuencia de ácidos nucleicos de factor VIII optimizada mediante codones que está basada en la secuencia de una secuencia de nucleótidos de factor VIII SQ. La secuencia de nucleótidos de factor VIII SQ es una secuencia de factor VIII de la que el dominio B ha sido agotado y sustituido con una conexión SQ de 14 aminoácidos entre los dominios a2 y a3. La presencia de la conexión SQ en el complejo favorece una escisión intracelular eficaz del producto de traducción de cadena única primario de 170 kDa debido a los residuos básicos de arginina que forman un resto de reconocimiento para la escisión proteolítica mediante la proteasa furina de tipo subtilisina unida a membrana.
- 45
- La secuencia de nucleótidos de la SEC ID N° 6 es una secuencia de ácidos nucleicos de factor VIII optimizado mediante codones que está basada en una secuencia de la secuencia de nucleótidos de factor VIII de factor B SQ. PEZ teleosteo *Fugu rubripes*. El dominio de Fugu B tiene una concentración de sitios de glicosilación conectados a N que mejora grandemente el tráfico intracelular y la expresión de la secuencia.
- 50
- La secuencia de nucleótidos de la SEC ID N° 7 es una secuencia de ácidos nucleicos de factor VIII optimizado mediante codones que está basada en la secuencia de la secuencia de nucleótidos de factor VIII N6.
- 55
- En una realización, cualquiera de las moléculas de ácidos nucleicos de la invención puede comprender una secuencia de nucleótidos que codifica una conexión SQ en la proteína de factor VIII. La secuencia de aminoácidos del conector SQ es preferentemente la secuencia de la SEC ID N° 18.
- 60
- En otra realización, cualquiera de las moléculas de ácidos nucleicos de la invención puede comprender una secuencia de nucleótidos que codifica un dominio b N6 en la proteína de factor VIII.
- 65
- En una realización adicional, cualquiera de las moléculas de ácidos nucleicos de la invención puede comprender una secuencia de nucleótidos que codifica un dominio Fugu B en la proteína de factor VIII.

El ácido nucleico de la invención puede comprender una conexión SQ y un dominio B N6 en la proteína de factor VIII, o una conexión SQ y un dominio Fugu B en la proteína de Factor VIII.

5 Generalmente, la optimización mediante codones no cambia el aminoácido para el que codifica cada codón. Simplemente cambia la secuencia de nucleótidos de forma que es más probable que sea expresada a un nivel relativamente elevado en comparación con la secuencia no optimizada con codones. Por lo tanto, en una realización, la molécula de ácidos nucleicos de la invención codifica una proteína que tiene entre 0 y 350, entre 0 y 300, entre 0 y 250, entre 0 y 200, entre 0 y 150, entre 0 y 100, entre 0 y 50, entre 0 y 30, entre 0 y 20, entre 0 y 15, entre 0 y 10, entre 0 y 5 cambios de aminoácidos para la proteína codificada mediante la secuencia de nucleótidos de la SEC ID N° 1 o la SEC ID N° 7. Esto significa que la secuencia de nucleótidos del ácido nucleico de la invención puede ser diferente pero cuando son traducidas a la secuencia de aminoácidos de la proteína que es producida difiere solamente en entre 0 y 10 aminoácidos. Preferentemente, cualesquiera cambio de aminoácidos codificados para el ácido nucleico de la invención, en comparación con la SEC ID N° 1 o la SEC ID N° 7 están en la parte de la secuencia que sustituyó el dominio B de la proteína de factor VIII, es decir, los cambios no se producen en otros dominios de la proteína como los dominios A1, a1, A2, a2, a3, A3, C1 o C2. Los cambios de aminoácidos en los otros dominios de la proteína de factor VIII afectan a la actividad biológica de la proteína de factor VIII.

Además, la molécula de ácidos nucleicos de la invención puede codificar una proteína que es codificada por la secuencia de nucleótidos de la SEC ID N° 1 o la SEC ID N° 7. Esto significa que las secuencias de nucleótidos del ácido nucleico de la invención pueden ser diferentes, pero cuando son traducidas, la secuencia de aminoácidos de la proteína que es producida es la misma.

En una realización preferida de la invención, la secuencia de nucleótidos que codifica un factor VIII funcional tiene una desviación mejorada del uso de codones para la célula humana en comparación con la secuencia de nucleótidos que se produce de forma natural que codifica la correspondiente secuencia no optimizada mediante codones. La adaptabilidad de una secuencia de nucleótidos que codifica un factor VIII funcional para el uso de codones de células humanas puede ser expresada en forma de un índice de adaptación de codones (CAI). Un índice de adaptación de codones se define en la presente memoria descriptiva como una medición de la adaptabilidad relativa del uso de codones de un gen hacia el uso de codones de genes humanos altamente expresados. La adaptabilidad relativa (w) de cada codón es la relación del uso de cada codón, respecto a la del codón más abundante para el mismo aminoácido. El CAI se define como la media geométrica de estos valores de la adaptabilidad relativa. Los codones no sinónimos y los codones de terminación (dependientes del código genético) son excluidos. Los valores del CAI varían en el intervalo de 0 a 1, en que los valores suprimidos indican una proporción superior de los codones más abundantes (véase Sharp and Li, 1987, Nucleic Acids Research 15: 1281-1295; véase también: Kim et al., Gene. 1997, 199:293-301; zur Megede et al., Journal of Virology, 2000, 74: 2628-2635). Preferentemente, una molécula de ácidos nucleicos que codifica un factor VIII tiene un CAI de al menos 0,8, 0,85, 0,90, 0,92, 0,94, 0,95, 0,96 o 0,97.

En una realización particular, la molécula de ácidos nucleicos codifica una proteína que comprende la secuencia de la SEC ID N° 2 o la SEC ID N° 21 que tiene entre 0 y 250, entre 0 y 200, entre 0 y 150, entre 0 y 100, entre 0 y 50, entre 0 y 30, entre 0 y 20, entre 0 y 15, entre 0 y 10 o entre 0 y 5 cambios de aminoácidos en la misma. Si la molécula de ácidos nucleicos codifica una proteína que comprende la secuencia de la SEC ID N° 2 o la SEC ID N° 21 que tiene cambios para cualquiera de los aminoácidos, la proteína debe ser todavía una proteína funcional. Un experto en la técnica apreciará que se pueden hacer cambios menores para algunos de los aminoácidos de la proteína sin afectar a la función de la proteína. Preferentemente, los cambios de aminoácidos están en la parte de la secuencia que sustituye al dominio B de la proteína del factor VIII, es decir, los cambios no se producen en otros dominios de la proteína como los dominios A1, a1, A2, a2, a3, A3, C1 o C2. En otras realizaciones, la molécula de ácidos nucleicos puede codificar una proteína que comprende o consiste en la secuencia de la SEC ID N° 2 o la SEC ID N° 21.

50 Estaría dentro de las capacidades de un experto en la técnica producir una molécula de ácidos nucleicos según la invención. Esto se podría hacer, por ejemplo, usando una síntesis química de una secuencia dada.

Además, un experto en la técnica sería capaz fácilmente de determinar si un ácido nucleico según la invención expresa una proteína funcional. Los métodos adecuados serían evidentes para los expertos en la técnica. Por ejemplo, un método *in vitro* adecuado incluye insertar el ácido nucleico en un vector, como un vector lentiviral o AAV transducir células hospedantes, como células 293T o HeLa con el vector y valorar la actividad de factor VIII. Alternativamente, un método *in vivo* adecuado incluye transducir un vector que contiene el ácido nucleico en ratones hemofílicos y valorar el factor VIII funcional en el plasma de los ratones. Se describen métodos adecuados más en detalle con posterioridad.

El ácido nucleico puede ser cualquier tipo de ácido nucleico compuesto por nucleótidos. El ácido nucleico debe ser capaz también de ser expresado de forma que se produzca una proteína. Preferentemente, el ácido nucleico es DNA o RNA.

65 La molécula de ácidos nucleicos comprende preferentemente una secuencia de nucleótidos seleccionada entre la

secuencia de la SEC ID N° 1 y la SEC ID N° 7. La molécula de ácidos nucleicos puede consistir en una secuencia de nucleótidos seleccionada entre la secuencia de la SEC ID N° 1 y la SEC ID N° 7. En una realización, la molécula de ácidos nucleicos consiste en la secuencia de nucleótidos de la SEC ID N° 1.

5 También se proporciona un vector que comprende la molécula de ácidos nucleicos de la invención. El vector puede ser cualquier vector apropiado, incluidos vectores virales y no virales. Los vectores virales incluyen vectores lenti-
adeno- y herpes-virales. Preferentemente es un vector viral recombinante adeno-asociado (rAAV) o un vector
lentiviral. Alternativamente, pueden ser usados sistemas no virales que incluyen DNA purificado (con o sin regiones
10 de unión a cromatina) o DNA conjugado que es introducido en la células mediante diversos métodos de transfección
como lípidos o electroporación.

El vector comprende también preferentemente otros componentes necesarios para la expresión de la molécula de
ácidos nucleicos, como promotores. Pueden ser usados cualesquiera promotores, como LP1, HCR-hAAT, ApoE-
hAAT y LSP. Estos promotores se describen más en detalle en las siguientes referencias: LP1: Nathwani A. et al.
15 Blood. 2006 April 1; 107(7): 2653-2661; HCR-hAAT: Miao CH et al. Mol Ther. 2000;1: 522-532; ApoE-hAAT:
Okuyama T et al. Human Gene Therapy, 7, 637-645 (1996); and LSP: Wang L et al. Proc Natl Acad Sci U S A. 1999
March 30; 96(7): 3906-3910..

Se proporciona un promotor particular preferido particular. Este es un promotor que comprende una secuencia de
nucleótidos que tiene una homología sustancial a la secuencia de nucleótidos de la SEC ID N° 3. El promotor es
20 específico para el hígado. En una realización, la molécula de ácidos nucleicos anteriormente descrita comprende
además una secuencia de nucleótidos que tiene una homología sustancial a la secuencia de nucleótidos de la SEC
ID N° 3. La expresión "homología sustancial" puede ser definida adicionalmente con referencia a un porcentaje de
homología. Esto se expone más en detalle con posterioridad.

25 Un vector según la invención puede ser un vector de suministro de genes. Este vector de suministro de genes puede
ser un vector de suministro de genes viral o un vector de suministro de genes no viral.

El suministro de genes no viral se puede llevar a cabo usando DNA purificado que es el método más simple de
30 transfección no viral. Puede ser posible, por ejemplo, administrar un ácido nucleico de la invención usando un DNA
de plásmido purificado. Alternativamente, pueden ser usados métodos como electroporación, sonoporación o
mediante el uso de una "pistola de genes", que dispara partículas de oro revestidas con DNA en la célula usando,
por ejemplo, un gas a presión elevada o una pistola de calibre 22 invertida.

35 Para mejorar el suministro de un ácido nucleico a una célula, puede ser necesario protegerlo del deterioro y puede
ser facilitada su entrada en la célula. Con esta finalidad, pueden ser usados lipoplejos y poliplexos que tienen la
capacidad de proteger un ácido nucleico de una degradación no deseable durante el procedimiento de transfección.

El DNA de plásmido puede ser revestido con lípidos en una estructura organizada como una micela o un liposoma.
40 Cuando la estructura organizada es complejada con DNA se denomina lipoplejo. Pueden ser usados lípidos
aniónicos y neutros para la construcción de lipoplejos para vectores sintéticos. Sin embargo, preferentemente,
pueden ser usados lípidos catiónicos, debido a su carga positiva, para condensar negativamente moléculas de DNA
con carga negativa con el fin de facilitar la encapsulación de DNA en forma de liposomas. Puede ser necesario
añadir lípidos Helper (habitualmente lípidos eléctricamente neutros como DOPE) a lípidos catiónicos con el fin de
45 formar lipoplejos.

Los complejos de polímeros con DNA, denominado poliplexos, pueden ser usados para suministrar un ácido nucleico
de la invención. La mayoría de los poliplexos consisten en polímeros catiónicos y su producción está regulada
50 mediante interacciones iónicas. Los poliplexos normalmente no pueden liberar su contenido de DNA en el
citoplasma. Por tanto, puede ser necesaria una co-transfección con agentes endosoma-líticos (para lisar el
endosoma que se prepara durante la endocitosis, el procedimiento mediante el cual el poliplexo entra en la célula),
como adenovirus inactivados.

Pueden ser usados métodos híbridos para suministrar un ácido nucleico de la invención que combine dos o más
55 técnicas. Los virosomas son un ejemplo, combinan liposomas con un HIV inactivado o virus gripal. Otros métodos
incluyen la mezcla de otros vectores virales con lípidos catiónicos o virus hibridantes y pueden ser usados para
suministrar un ácido nucleico de la invención.

Puede ser usado un dendrímero para suministrar un ácido nucleico de la invención, en particular, un dendrímero
60 catiónico, es decir, uno con una carga superficial positiva. Cuando esté en presencia de un material genético como
DNA o RNA, la complementaridad de la carga conduce a una asociación temporal del ácido nucleico con el
dendrímero catiónico. Al alcanzar su destino, el complejo de dendrímero-ácido nucleico es seguidamente importado
a la célula a través de endocitosis.

65 Más normalmente, puede ser usado un vector de suministro de genes virales adecuado para suministrar un ácido
nucleico de la invención. Los vectores virales adecuados para ser usados en la invención pueden ser un parvovirus,

un adenovirus, un retrovirus, un lentivirus o un virus simple de herpes. El parvovirus puede ser un virus asociado a adenovirus (AAV).

5 Como se usa en la presente invención, en el contexto del suministro de genes, el término "vector" o la expresión "vector de suministro de genes" se puede referir a una partícula que funciona como un vehículo de suministro de genes y que comprende ácido nucleico "es decir, el genoma del vector) envasado, por ejemplo, en una envoltura o cápside. Alternativamente, en algunos contextos, el término "vector" puede ser usado para hacer referencia solamente al genoma del vector.

10 Consecuentemente, la presente invención proporciona vectores de suministro de genes (que comprenden un ácido nucleico de la invención) basados en parvovirus de animales, en particular dependovirus como AAV humano infeccioso o de simio y sus componentes (por ejemplo, un genoma de parvovirus animal) para ser usado como vectores para la introducción y/o expresión de un polipéptido de factor VIII en una célula de mamífero. El término parvoviral", como se usa en la presente memoria descriptiva, abarca por tanto a dependovirus como cualquier tipo
15 de AAV.

Los virus de la familia de parvoviridos son virus de animales de DNA pequeños. La familia de los parvoviridos puede ser dividida en dos subfamilias: los parvoviridos, que infectan a vertebrados y los densoviridos que infectan a insectos. Los miembros de la subfamilia de los parvoviridos se denominan en la presente memoria descriptiva parvovirus e incluyen el género dependovirus. Como se puede deducir a partir del nombre de su género, los miembros de los dependovirus son únicos en cuanto que habitualmente requieren una co-infección con un virus
20 Helper como un adenovirus o un virus del herpes para una infección productiva en el cultivo celular. El género dependovirus incluye AAV, que normalmente infecta seres humanos (por ejemplo, serotipos, 2, 3a, 3b, 4, 5 y 6) o de primates (por ejemplo, serotipos 1 y 4) y virus relacionados que infectan otros animales de sangre caliente (por ejemplo, virus de adenovirus bovinos, caninos, equinos, y ovinos. Otra información sobre parvovirus y otros miembros de los parvoviridos se describe por Kenneth I. Berns, "Parvoviridae: The Viruses and Their Replication," capítulo 69 en Fields Virology (3d Ed. 1996). Por motivos de conveniencia, la presente invención se ilustra adicionalmente como se describe en la presente memoria descriptiva en relación a los AAV. Sin embargo, debe entenderse que la invención no está limitada a los AAV sino que puede ser igualmente aplicada a otros parvovirus.
30

La organización genómica de todos los serotipos de AAV conocidos es muy similar. El genoma de un AAV es una molécula de DNA de cadena única lineal, que tiene menos de aproximadamente 5 mil nucleótidos (nt) de longitud. Repeticiones terminales invertidas (ITRs) flanquean las secuencias de nucleótidos de codificación única para las proteínas de replicación no estructurales (Reb) y las proteínas estructurales (VP). Las proteínas VP (VB1, -2 y -3) forman la cápside IDE. Los 145 nt son auto-complementarios y están organizados de forma que se puede formar un dúplex intramolecular energéticamente estable que forma una horquilla en forma de T. Estas estructuras de horquillas funcionan como n origen para la replicación de DNA viral, sirviendo como cebadores para el complejo celular de DNA-polimerasa. A continuación de una infección de AAV de tipo salvaje (WT) en células de mamíferos, los genes de Rep (es decir, que codifican proteínas Rep78 y Rep52) son expresados a partir del promotor P5 y el promotor P19, respectivamente, y ambas proteínas de Rep tienen una función en la replicación del genoma viral. Un acontecimiento de escisión en el Rep o RF da lugar a la expresión de realmente cuatro proteínas de Rep (es decir, Rep78, Rep68, Rep52 y Rep 40). Sin embargo, se ha mostrado que el eRNA no escindido, que codifica proteínas de Rep78 y Rep52, en células de mamíferos, es suficiente para la producción del vector de AAV. También, en células de insectos, las proteínas de Reo 78 y Rep52 son suficientes para la producción de vector de AAV.
40

En un AAV adecuado para ser usado como un vector de terapia génica el genoma del vector comprende normalmente un ácido nucleico de la invención (como se describe en la presente memoria descriptiva) que va a ser envasado para un suministro a una célula diana. Según esta realización particular, la secuencia de nucleótidos heterólogos está ubicada entre las ITRs virales en cualquier extremo del genoma del vector. En realizaciones preferidas adicionales, los genes CAP de parvovirus (por ejemplo, AAV) y los genes Rep de parvovirus (por ejemplo AAV) son suprimidos del genoma de la plantilla (y, por tanto del DNA de virión producido a partir del mismo). Esta configuración maximiza el tamaño de la(s) secuencia(s) de ácidos nucleicos que se puede(n) llevar a cabo por la cápside del parvovirus.
50

Según esta realización particular, el ácido nucleico de la invención está ubicado entre las ITRs virales en cualquier extremo del sustrato. Es posible que un genoma parvoviral funcione como solamente una ITR. Por tanto, en un vector de terapia génica de la invención basado en un parvovirus, el genoma del vector esta flanqueado por al menos una ITR pero, más normalmente, por dos ITRs de AAV (generalmente con un lado del genoma del vector, es decir, uno en el extremo 5' y uno en el extremo 3'). Puede haber secuencias que intervienen entre el ácido nucleico de la invención en el genoma del vector y una o más de las ITRs.
60

Preferentemente, el ácido nucleico que codifica un polipéptido de factor VII funcional (para la expresión en la célula de mamífero) será incorporado en un genoma parvoviral ubicado entre dos ITRs regulares o ubicado en cualquier lado de ITR tratada por ingeniería genética con dos regiones D.
65

Las secuencias de AAV que pueden ser usadas en la presente invención para la producción de vectores de terapia

génica de AAV pueden ser derivadas del genoma de cualquier serotipo de AAV. Generalmente, los serotipos de AAV tienen secuencias genómicas de una homología significativa en los niveles de aminoácidos y ácidos nucleicos, proporcionan un conjunto idéntico de funciones genéticas, producen viriones que son de forma esencial física y funcionalmente equivalentes y se replican y ensamblan mediante mecanismos prácticamente idénticos. Para la secuencia genómica de los diversos serotipos de AAV y para una visión global de las analogías genómicas, véase, por ejemplo, GenBank Accession number U89790; GenBank Accession number J01901; GenBank Accession number AF043303; GenBank Accession number AF085716; Chiorini et al. (1997, J. Vir. 71: 6823-33); Srivastava et al. (1983, J. Vir. 45:555-64); Chiorini et al. (1999, J. Vir. 73:1309-1319); Rutledge et al. (1998, J. Vir. 72:309-319); and Wu et al. (2000, J. Vir. 74: 8635-47). Pueden ser usados serotipos de AAV 1, 2, 3, 4, 5, 6, 7, 8 ó 9 en la presente invención. Sin embargo, los serotipos de AAV 1, 5 ó 8 son fuentes preferidas de secuencias de AAV para ser usadas en el contexto de la presente invención.

Preferentemente, las secuencias de ITR de AAV para ser usadas en el contexto de la presente invención son derivadas de AA V1, AAV2, AAV4 y/o AAV6. Análogamente, las secuencias codificadoras de Rep (Rep78 y Rep52) son preferentemente derivadas de AAV1, AAV2, AAV4 y/o AAV6. Las secuencias que codifican las proteínas de cápsides VP1, VP2 y VP3 para ser usadas en el contexto de la presente invención, sin embargo, pueden ser tomadas de cualquiera de los 42 serotipos conocidos, más preferentemente de AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 o AAV9 o partículas de tipo AAV nuevamente desarrolladas obtenidas, por ejemplo, mediante técnicas de recomposición de cápsides y bibliotecas de cápsides de AAV.

Las secuencias de Rep e ITR de AAV son particularmente conservadas entre las mayorías de los serotipos. Las proteínas de Rep78 de diversos serotipos de AAV son, por ejemplo, más de un 89% idénticas y la identidad de la secuencia de nucleótidos totales al nivel del genoma entre AAV2, AAV3a, AAV3b y AAV6 es de aproximadamente 82% (Bantel-Schaal et al., 1999, J. Virol., 73(2):939-947). Además, las secuencias de Rep y las ITRs de muchos serotipos de AAV se conoce que se cruzan-complementan eficazmente (es decir, sustituyen funcionalmente) a las secuencias correspondientes de otros serotipos en la producción de partículas de AAV en células de mamíferos. El documento US 2003148506 refiere que las secuencias de Rep y de ITR de AAV también se cruzan-complementan, eficazmente respecto a otras secuencias de Rep y de ITR de AAV en células de insectos.

Las proteínas VP de AAV se conoce que determinan la tropicidad células del virión de AAV. Las secuencias que codifican proteínas de VP son significativamente menos conservadas que las proteínas de Rep y los genes entre diferentes serotipos de AAV. La capacidad de secuencias de Rep y de ITR para conocer-complementar las correspondientes secuencias de otros serotipos permite la producción de partículas de AAV pseudo-tipadas que comprenden las proteínas de cápsides de un serotipo (por ejemplo AAV1, 5 o 8) y las secuencias de Rep y/o ITR de otro serotipo AAV (por ejemplo, AAV2). Estas partículas de rAAV pseudotipadas son una parte de la presente invención.

Las secuencias de "AAV" modificadas pueden ser también en el contexto de la presente invención, por ejemplo para la producción de vectores de terapia génica de AAV. Estas secuencias modificadas incluyen, por ejemplo, secuencias que tienen al menos aproximadamente 75%, al menos es aproximadamente 80%, al menos aproximadamente, 85%, al menos aproximadamente 90%, al menos aproximadamente 95% o más identidad de secuencias de nucleótidos y/o amino vacíos (por ejemplo, una secuencia que tiene una identidad de secuencias de nucleótidos de aproximadamente 75-99%) respecto a un AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 o AAV9 ITR, Rep o VP pueden ser usadas en lugar de las secuencias de AAV ITR, Rep o VP de tipo salvaje.

Aunque es similar a otros serotipos de AAV en muchos aspectos, el AAV5 difiere de otros serotipos de AAV humanos y de simios más que otros serotipos humanos y de simios conocidos. Considerando esto, la producción de AAV5 puede diferir de la producción de otros serotipos en células de insectos. Cuando los métodos de la invención son empleados para producir rAAV5, es preferido que uno o más constructos que comprendan, colectivamente en el caso de más de un constructo, una secuencia de nucleótidos que comprende una IPR de AAV5, una secuencia de nucleótidos comprende una secuencia codificadora Rep de AAV5 (es decir, una secuencia de nucleótidos comprende una Rep78 de AAV5). Estos ITR y secuencias de Rep pueden ser modificadas en la medida deseada para obtener una producción eficaz de AAV5 o vectores de AAV5 pseudotipados. Por ejemplo, el codón de comienzo de las secuencias de Rep puede ser modificado, los sitios de escisión de VP pueden ser modificados o eliminados y/o el codón de comienzo de VP1 y los nucleótidos cercanos pueden ser modificados para mejorar la producción de vectores de AAV5.

Por tanto, la cápside viral usado en la invención puede ser de cualquier parvovirus, ya sea un parvovirus autónomo o un dependovirus, como se describió anteriormente. Preferentemente, la cápside viral es un cápside de AAV (por ejemplo, cápside de AAV1, AAV2, AAV3, AAV4, AAV5 o AAV6). En general, se prefieren la cápside de AAV1 o cápside de AAV6. La elección de la cápside parvovirus puede estar basada en un cierto número de consideraciones como es conocido en la técnica, por ejemplo, el tipo de célula diana, el nivel de expresión deseado, la naturaleza de la secuencia de nucleótidos heteróloga que va a ser expresada, cuestiones relacionadas con la producción viral y similares. Por ejemplo, la cápside de AAV1 y AAV6 pueden ser ventajosamente empleados para músculos esqueléticos; AAV1, AAV5 y AAV8 para el hígado y células del sistema nervioso central (por ejemplo, cerebro); AAV5 para células en las vías respiratorias y el pulmón o cerebro; AAV3 para células de médula ósea; y AAV4 para

células particulares en el cerebro (por ejemplo, células adjuntables).

Está dentro de los conocimientos técnicos de un experto en la técnica, seleccionar el virus, subtipo de virus o serotipo de virus más apropiado. Algunos subtipos o serotipos pueden ser más apropiados que otros para un cierto tipo de tejido.

Por ejemplo, la expresión específica para el hígado de un ácido nucleico de la invención puede ser ventajosamente inducida mediante la transducción mediada por AAV de células de hígado. El hígado es susceptible a una transducción mediada por AAV y pueden ser usados muchos serotipos diferentes (por ejemplo, AAV1, AAV5 o AAV8). La transducción de músculos puede ser realizada mediante la administración de un AAV que codifique un ácido nucleico de la invención a través de la corriente sanguínea. Por tanto, es aplicable la administración intravenosa o intra-arterial.

Un vector de terapia génica de parvovirus preparado según la invención puede ser una partícula "híbrida" en la que las TRs virales y la cápside viral son de parvovirus diferentes. Generalmente, las TRs virales y la cápside son de serotipos diferentes de AAV. Análogamente, el parvovirus puede tener un cápside "quimérico" (por ejemplo, que contiene secuencia de parvovirus diferentes, preferentemente diferentes serotipos de AAV) o un cápside "hecho diana" (por ejemplo, un tropismo dirigido).

En el contexto de la invención, "al menos una secuencia de nucleótidos de IPR parvoviral" se entiende que significa una secuencia palindrómica, que comprende principalmente secuencias dispuestas de forma complementaria o simétrica también denominadas regiones "A", "B" y "C". La ITR funciona como un origen de replicación, un sitio que tiene una función "cis" en la replicación, es decir, que es un sitio de reconocimiento para proteínas de replicación de actuación en trans, como, por ejemplo, Rep78 (o Rep68) que reconoce el palíndromo y las secuencias específicas internas para el palíndromo. Una excepción a la simetría de la secuencia de ITR es la región "D" de la ITR. Es única (que no tiene un complemento en una ITR). El mellado de DNA de cadena única se produce en la unión entre las regiones A y D. Es la región en la que se inicia la síntesis de nuevo DNA. La región D normalmente se sitúa en un lado del palíndromo y proporciona direccionalmente la etapa de replicación del ácido nucleico. Un parvovirus replicante en una célula de mamífero tiene normalmente dos secuencias de ITR. Sin embargo, es posible tratar por ingeniería genética una ITR de forma que los sitios de unión estén en las dos cadenas de las regiones A y regiones D ubicadas simétricamente, una a cada lado del palíndromo. En una plantilla de DNA circular de cadena doble (por ejemplo, un plásmido), la replicación de ácido nucleico asistida por Rep78 o Rep68 tiene lugar entonces en ambas direcciones y es suficiente una ITR única para la replicación parvoviral de un vector circular. Por tanto, puede ser usada una secuencia de nucleótidos de ITR en el contexto de la presente invención. Sin embargo, preferentemente, se usan dos o incluso otro número de ITRs regulares. Lo más preferentemente, se usan dos secuencias de ITR. Una ITR parvoviral preferida es una ITR de AAV. Por razones de seguridad puede ser deseable construir un vector parvoviral (AAV) que es incapaz de propagarse adicionalmente después de la introducción inicial en una célula. Este mecanismo de seguridad para limitar la propagación no deseable del vector en un receptor puede ser proporcionado usando AAV con ITR quimérica como se describe en el documento US 2003148506.

Los expertos en la técnica apreciarán que la(s) proteína(s) de Rep usada(s) para producir un vector de AAV de la invención se pueden seleccionar teniendo en consideración la fuente de las ITRs virales. Por ejemplo, la ITR de AAV5 interacciona normalmente más eficazmente con la proteína Rep de AAV5, aunque no es necesario que el serotipo de la ITR y la proteína de Rep sea(n) coincidente(s).

La(s) ITR(s) usada(s) en la invención es (o son) normalmente funcional(es), es decir, pueden ser completamente resolubles y son preferentemente secuencias de AAV, siendo preferidos los serotipos 1, 2, 3, 4, 5 o 6. Las ITRs de AAV resolubles según la presente invención no necesitan una secuencia de ITR salvaje (es decir, una secuencia de tipo salvaje puede ser alterada mediante inserción, delección, truncación o mutaciones sin sentido), en la medida en que la ITR medie las funciones deseadas, por ejemplo, envasado de virus, integración y/o rescate de provirus, y similares.

Ventajosamente, usando el vector de terapia génica en comparación con propuestas previas, la restauración de la proteína de síntesis, es decir, la síntesis de factor VIII, es una característica que las células transduccidas adquieren permanentemente o durante un período de tiempo sostenido, evitando así la necesidad de una administración continua para conseguir un efecto terapéutico.

Consecuentemente, los vectores de la invención representan por lo tanto una herramienta para el desarrollo de estrategias para el suministro *in vivo* de un ácido nucleico de la invención, tratando por ingeniería genética el ácido nucleico con un vector de terapia génica que transduce eficazmente el tipo de célula apropiada, como una célula de hígado.

En un aspecto adicional de la invención, se proporciona un hospedante que comprende el vector anteriormente descrito. Preferentemente, el vector es capaz de expresar la molécula de ácido nucleico de la invención en el hospedante. El hospedante puede ser cualquier hospedante adecuado.

Como se usa en la presente memoria descriptiva, el término "hospedante" se refiere a organismos y/o células que albergan una molécula de ácido nucleico o un vector de la invención, así como organismos y/o células que son adecuados para ser usados en la expresión en un gen o proteína recombinante. No está previsto que la presente invención esté limitada a algún tipo particular de célula u organismo. De hecho, está contemplado que cualquier

5 organismos y/o célula adecuados encuentren uso en la presente invención como un hospedante. Una célula hospedante puede estar en la forma de una célula única, una población de células similares o diferentes, por ejemplo en la forma de un cultivo (como un cultivo líquido o un cultivo en un sustrato celular sólido), un organismo o parte de los mismos.

10 Una célula hospedante según la invención puede permitir la expresión de una molécula de ácido nucleico de la invención. Por tanto, la célula hospedante puede ser, por ejemplo, una bacteria, una levadura, un insecto o una célula de mamífero.

15 Cualquier célula de insecto que permita la replicación de un vector parvoviral recombinante (rAAV) y que pueda ser mantenida en un cultivo puede ser usada de acuerdo con la presente invención. Por ejemplo, la línea celular usada puede ser de *Spodoptera frugiperda*, líneas celulares de drosófilas, líneas celulares de mosquito, por ejemplo, líneas celulares derivadas de *Aedes albopictus*. Las células de insectos o líneas celulares preferidas son células de las especies de insectos que son susceptibles de infección con baculovirus que incluyen, por ejemplo, Se301, SelZD2109, SeUCR1, Sf9, Sf900+, Sf21, BTI-TN-5B1-4, MG-1, Tn368, HzAm1, Ha2302, Hz2E5, High Five

20 (Invitrogen, CA, USA) y *expresSF+*[®] (US 6.103.526; Protein Sciences Corp., CT, USA)..

Además, la invención proporciona un método para la preparación de un vector de suministro de genes parvovirales, comprendiendo el método las etapas de:

25 (a) proporcionar una célula de insecto que comprende uno o más constructos de ácidos nucleicos que comprenden:

(i) una molécula de ácido nucleico de la invención que está flanqueada por al menos una secuencia de nucleótidos repetida terminal invertida parvoviral;

30 (ii) un primer cassette de expresión que comprende una secuencia de nucleótidos que codifica una o más proteínas Rep parvovirales que está funcionalmente conectado a un promotor que es capaz de conducir la expresión de la(s) proteína(s) Rep en la célula de insecto;

35 (iii) un segundo cassette de expresión que comprende una secuencia de nucleótidos que codifica una o más proteínas de cápsides parvovirales que está funcionalmente conectada a un promotor que es capaz de conducir la expresión de la(s) proteína(s) de la cápside en la célula de insecto;

(b) cultivar la célula de insecto definida en el apartado (a) bajo condiciones que conduzcan a la expresión del Rep y las proteínas de cápside; y opcionalmente,

40 (c) recuperar el vector de suministro de genes parvovirales.

Por lo tanto, en general, el método de la invención permite la producción de un vector de suministro de genes parvovirales (que comprende un ácido nucleico de la invención) en una célula de insecto. Preferentemente, el

45 método comprende las etapas de: (a) cultivar una célula de insecto como se definió anteriormente bajo condiciones tales que se produzca el vector parvoviral (por ejemplo, AAV); y (b) recuperar el vector parvoviral recombinante (por ejemplo, AAV). Preferentemente, el vector de suministro de genes parvovirales es un vector de suministro de genes AAV.

50 Debe entenderse que el vector (AAV) producido en este método es preferentemente un virión parvoviral o AAV infeccioso que comprende un genoma parvoviral, que comprende en sí mismo un ácido nucleico de la invención. Las condiciones de crecimiento para células de insectos en un cultivo, y la producción de productos heterólogos en células de insectos en cultivo, son bien conocidas en la técnica y están descritas, por ejemplo, en las referencias anteriormente citadas sobre tratamiento por ingeniería molecular de células de insectos.

55 En un método de la invención, se proporciona un ácido nucleico de la invención que está flanqueado por al menos una secuencia de ITR parvoviral. Este tipo de secuencia se describe en detalle con anterioridad. Preferentemente, el ácido nucleico de la invención está ubicado entre dos secuencias ITR parvovirales.

60 El primer cassette de expresión comprende una secuencia de nucleótidos que codifica una o más proteínas Rep parvovirales que están funcionalmente conectadas a un primer promotor que es capaz de conducir la expresión de la(s) proteína(s) Rep en la célula de insecto.

65 Una secuencia de nucleótidos que codifica proteínas Rep de parvovirus de animales es entendida en la presente memoria descriptiva como una secuencia de nucleótidos que codifica las proteínas no estructurales que son necesarias y suficientes para la producción de vector parvoviral en células de insectos como las proteínas Rep78 y

Rep52 o las proteínas Rep68 y Rp 40 o las combinaciones de dos o más de las mismas.

La secuencia de nucleótidos de parvovirus animal es preferentemente de un dependovirus, más preferentemente de un virus adeno-asociado (AAV) humano o de simio y, lo más preferentemente, de un AAV que normalmente infecta seres humanos (por ejemplo, serotipos, 1, 2, 3A, 3B, 4, 5 y 6) o primates (por ejemplo serotipos 1 y 4). Las secuencias que codifican Rep son bien conocidas por los expertos en la técnica y se mencionan y se describen en detalle secuencias adecuadas en el documento WO 2007/148971 y también en el WO 2009/14445.

Preferentemente, la secuencia de nucleótidos codifica proteínas Rep de parvovirus de animales que son necesarias y suficientes para la producción de vectores parvovirales en células de insectos.

El segundo cassette de expresión comprende una secuencia de nucleótidos que codifica una o más proteínas de cápsides parvovirales que está funcionalmente conectada a un promotor que es capaz de conducir la expresión de la(s) proteína(s) de la cápside en la célula de insecto. La(s) proteína(s) de cápside expresada puede ser una o más de las anteriormente descritas.

Preferentemente, la secuencia de nucleótidos codifica proteínas cap de parvovirus de animales que son necesarias y suficientes para la producción del vector parvoviral en células de insectos.

Estas tres secuencias (genoma, codificación, rep y codificación cap) son proporcionadas en una célula de insecto por medio de uno o más constructos de ácidos nucleicos, por ejemplo, uno, dos o tres constructos de ácidos nucleicos. Así, preferentemente, el uno o más constructos de ácidos nucleicos para el genoma y expresión del vector del Rep parvoviral y proteínas cap en células de insecto es un vector compatible con células de insectos. Un "Vector compatible con células de insectos" o "vector" se entiende que es una molécula de ácido nucleico capaz de una transformación productora o transfección de un insecto o célula de insecto. Ejemplos de vectores biológicos incluyen plásmidos, moléculas de ácidos nucleicos lineales y virus recombinantes. Puede ser empleado cualquier vector en la medida en que sea compatible con células de insectos. El vector se puede integrar en el genoma de las células de insectos, pero la presencia del vector en la célula de insecto necesita ser permanente y son incluidos también vectores episomales transitorios. Los vectores pueden ser introducidos por cualquier medio conocido, por ejemplo, mediante tratamiento químico de las células, electroporación o infección. En una realización preferida, el vector es un vaculovirus, un vector viral o un plásmido. En una realización más preferida, el vector es un vaculovirus, es decir, el constructo es un vector vaculoviral. Los vectores vaculovirales y los métodos para su uso son bien conocidos por los expertos en la técnica.

Por tanto, normalmente un método de la invención para producir un vector de suministro de genes parvovirales comprende: proporcionar a una célula permisiva para la replicación de parvovirus (a) una secuencia de nucleótidos que codifica una plantilla para producir genoma de vector de la invención (como se describe en detalle en la presente memoria descriptiva); (b) secuencias de nucleótidos suficientes para la replicación de la plantilla para producir un genoma de vector (el primer cassette de expresión anteriormente definido); (c) secuencias de nucleótidos suficientes para envasar el genoma del vector en un cápside de parvovirus (el segundo cassette de expresión anteriormente definido), bajo condiciones suficientes para la replicación y envasado del genoma de vector en la cápside de parvovirus, con lo que se producen en la célula partículas de parvovirus que comprenden el genoma del vector encapsulado en la cápside de parvovirus. Preferentemente, la replicación de parvovirus y/o secuencias codificadoras de la cápside son secuencias de AAV.

Un método de la invención puede comprender preferentemente la etapa de purificación por afinidad del vector parvoviral recombinante (rAAV) (viriones que lo comprenden) usando un anticuerpo anti-AAV, preferentemente un anticuerpo inmovilizado. El anticuerpo anti-AAV es preferentemente un anticuerpo monoclonal. Un anticuerpo particularmente adecuado es un anticuerpo de camélida de cadena única o un fragmento del mismo que puede ser obtenido, por ejemplo a partir de camellos o llamas (véase, por ejemplo, Muyldermans, 2001, *Biotechnol.* 74: 277-302). El anticuerpo para la purificación por afinidad de rAAV es preferentemente un anticuerpo que se une a un epitopo en una proteína de cápside de AAV, con lo que el epitopo es preferentemente un epitopo que está presente en la proteína de cápside y más de un serotipo de AAV. Por ejemplo, el anticuerpo puede ser constituido o seleccionado sobre la base de unión específica a cápside AAV2 pero, al mismo tiempo, se puede unir también específicamente a cápsides AAV1, AAV3, AAV5, AAV6, AAV8 o AAV9.

La invención proporciona también un medio para suministrar un ácido nucleico de la invención en una amplia gama de células, incluidas células en división y que no están en división. La presente invención puede ser empleada para suministrar un ácido nucleico de la invención a una célula *in vitro*, por ejemplo, para producir un polipéptido codificado mediante esta molécula de ácido nucleico *in vitro* o para una terapia génica *ex vivo*.

La molécula de ácido nucleico, vector, células y usos de la presente invención son adicionalmente útiles en un método para suministrar un ácido nucleico de la invención a un hospedante que lo necesita, normalmente un hospedante que padece de hemofilia A.

La presente invención encuentra uso en aplicaciones veterinarias y médicas. Los sujetos adecuados para métodos

de suministro génico como se describen en la presente memoria descriptiva incluyen aves y mamíferos, siendo preferidos los mamíferos. El término "ave" como se usa en la presente memoria descriptiva incluye, pero sin limitación, pollos, patos, gansos, codornices, pavos y faisanes. El término mamífero, como se usa en la presente memoria descriptiva incluye pero sin limitación seres humanos, bovinos, caprinos, equinos, felinos, caninos, lagomorfos, etc. Los más preferidos son los sujetos humanos. Los sujetos humanos incluyen neonatos, niños, jóvenes y adultos.

Por tanto, la invención proporciona una composición farmacéutica que comprende un ácido nucleico o un vector de la invención y un vehículo o diluyente farmacéuticamente aceptable y/o otro agente medicinal, agente o adyuvante farmacéutico, etc.

Para una inyección, el vehículo será normalmente un líquido. Para otros métodos de administración el vehículo puede ser sólido o líquido. Para una administración por inhalación, el vehículo será respirable y, preferentemente, estará en forma de partículas sólidas o líquidas. Como un medio de inyección, es preferido usar agua que contenga los aditivos habituales para soluciones de inyección, como agentes estabilizantes, sales o solución salina y/o tampones.

En general, un vehículo "farmacéuticamente aceptable" es uno que es no tóxico ni indebidamente perjudicial para las células. Ejemplos de vehículos farmacéuticamente aceptables incluyen agua esterilizada exenta de pirógenos y solución salina tamponada de fosfato esterilizada y exenta de pirógenos. Los vehículos farmacéuticamente aceptables incluyen vehículos fisiológicamente aceptables. La expresión "vehículo farmacéuticamente aceptable" incluye cualquiera y la totalidad de disolventes, medios dispersantes, revestimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardantes de la absorción y similares, que sean fisiológicamente compatibles.

Mediante "farmacéuticamente aceptable" se quiere indicar que no es biológicamente o de alguna otra forma no deseable, es decir, el material puede ser administrado a un sujeto sin provocar ningún efecto biológico no deseable. Por tanto, esta composición farmacéutica puede ser usada, por ejemplo, en la transfección de una célula *ex vivo* o en la administración de una partícula viral o célula directamente a un sujeto.

Un vehículo puede ser adecuado para una administración parenteral, que incluye una administración intravenosa, intraperitoneal o intramuscular. Alternativamente, el vehículo puede ser adecuado para una administración sublingual u oral. Los vehículos farmacéuticamente aceptables incluyen soluciones o dispersiones acuosas esterilizadas y polvos esterilizados para una preparación extemporánea de soluciones o dispersiones inyectables esterilizadas. El uso de estos medios y agentes para sustancias farmacéuticamente activas es bien conocido en la técnica. Salvo que muchos medios o agentes convencionales sean incompatibles con el compuesto activo, está contemplado su uso en las composiciones farmacéuticas de la invención.

Las composiciones farmacéuticas están normalmente esterilizadas y son estables bajo las condiciones de elaboración y uso. Las composiciones farmacéuticas pueden ser formuladas como una solución, microemulsión, liposoma u otra estructura ordenada adecuada para acomodar una concentración de fármaco elevada. El vehículo puede ser un disolvente o medio de dispersión que contenga, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido, y similares) y sus mezclas adecuadas. La fluidez apropiada puede ser mantenida, por ejemplo, mediante el uso de un revestimiento como lecitina, mediante el mantenimiento del tamaño de partículas requerido en el caso de una dispersión y mediante el uso de tensioactivos. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares, polialcoholes como manitol, sorbitol o cloruro de sodio en la composición. La absorción prolongada de las composiciones inyectables se puede llevar a cabo incluyendo en la composición un agente que retrase la absorción, por ejemplo, sales de monoestearato y gelatina. Un ácido nucleico o vector de la invención puede ser administrado en un período de tiempo o en una formulación de liberación controlada, por ejemplo, en una composición que incluya un polímero de liberación lenta u otros vehículos que protejan el compuesto frente a una liberación rápida, que incluyen implantes y sistemas de suministro microencapsulados. Pueden ser usados, por ejemplo, polímeros biodegradables o biocompatibles como etilenoacetato de vinilo, polianhídridos, poli(ácido glicólico), colágeno, poliortoésteres, poli(ácido láctico) y copolímeros poliglicólicos (PLG).

El vector parvoviral, por ejemplo AAV de la invención, puede ser usado para transferir material genético a una célula. Esta transferencia puede tener lugar *in vitro*, *ex vivo* o *in vivo*.

Consecuentemente, se proporciona un método para suministrar una secuencia de nucleótidos a una célula, método que comprende poner en contacto un ácido nucleico, un vector o una composición farmacéutica como se describe en la presente memoria descriptiva bajo condiciones tales que el ácido nucleico o vector de la invención entre en la célula. La célula puede ser una célula *in vitro*, *ex vivo* o *in vivo*.

También se proporciona un método para tratar la hemofilia que comprende administrar una cantidad eficaz de un ácido nucleico, una proteína o un vector a un paciente que padece hemofilia. Preferentemente, el paciente está padeciendo hemofilia A. Preferentemente, el paciente es un ser humano.

Adicionalmente, se proporciona también un método para suministrar o administrar una secuencia de nucleótidos a un sujeto, método que comprende administrar a dicho sujeto un ácido nucleico, un vector o una composición farmacéutica como se describe en la presente memoria descriptiva. En particular, se proporciona un método para administrar una molécula de ácido nucleico de la invención a un sujeto, que comprende administrar al sujeto un vector de terapia génica parvoviral según la invención, opcionalmente junto con un vehículo farmacéuticamente aceptable. Preferentemente, el vector de terapia génica parvoviral es administrado en una cantidad terapéuticamente eficaz a un sujeto que lo necesita. Es decir, la administración según la invención se lleva a cabo normalmente bajo condiciones que dan lugar a la expresión de factor VIII funcional a un nivel que proporciona un efecto terapéutico en un sujeto que lo necesita.

El suministro de un ácido nucleico o vector de la invención a una célula hospedante *in vivo* puede dar lugar a un aumento de factor VIII funcional en el hospedante, por ejemplo, hasta un nivel que mejore uno o más síntomas de un trastorno de coagulación sanguínea como la hemofilia A.

El nivel de factor VIII que se produce de forma natural en un sujeto que padece hemofilia A varía dependiendo de la gravedad de la hemofilia. Los pacientes con una forma grave de la enfermedad tienen niveles de factor VIII de menos de aproximadamente 1% del nivel encontrado en un sujeto sano normal (denominado en la presente memoria descriptiva "un nivel normal", un nivel normal es de aproximadamente 50-150 IU/dl). Los pacientes con una forma moderada de la enfermedad tienen niveles de factor VIII entre aproximadamente 1% y aproximadamente 5% de un nivel normal. Los pacientes con una forma leve de la enfermedad tienen niveles de factor VIII de más de aproximadamente 5% de un nivel normal; normalmente entre aproximadamente 5% y aproximadamente 30% de un nivel normal.

Se ha encontrado que cuando se usa el método de tratamiento anteriormente descrito, puede provocar un aumento en el nivel de factor VIII funcional de al menos 1% de los niveles normales, es decir, además del nivel de factor VIII presente en el sujeto antes del tratamiento. En un sujeto que padece hemofilia, este aumento puede provocar una mejora de un síntoma de la hemofilia. En particular, un aumento de al menos 1% puede reducir la frecuencia de la hemorragia que se produce en pacientes de hemofilia, especialmente aquellos con una forma grave de la enfermedad. En una realización, el método de tratamiento provoca un aumento en el nivel de factor VIII funcional de al menos aproximadamente 5% de los niveles normales. Esto podría cambiar el fenotipo de la enfermedad de grave hasta leve. Los pacientes con una forma leve de la enfermedad raramente tienen hemorragias espontáneas. En otras realizaciones, el método de tratamiento de la invención provoca un aumento en el nivel de factor VIII funcional de al menos aproximadamente 2%, al menos aproximadamente 3%, al menos aproximadamente 4%, al menos aproximadamente 10%, al menos aproximadamente 15%, al menos aproximadamente 20% o al menos aproximadamente 25% de los niveles normales. En una realización particular, el método de tratamiento de la invención provoca un aumento en el nivel de factor VIII funcional de al menos aproximadamente 30% de los niveles normales. Este nivel de aumento normalizaría virtualmente la coagulación de sangre en sujetos que padecen hemofilia. Estos sujetos es improbable que requieran concentrados de factor VIII a continuación de un trauma o durante una cirugía.

En otra realización, el método de tratamiento puede provocar un aumento en el nivel de factor VIII funcional hasta al menos aproximadamente un 1% de los niveles normales. El método de tratamiento puede provocar un aumento en el nivel de factor VIII funcional hasta al menos aproximadamente un 5% de los niveles normales. En otras realizaciones, el método de tratamiento puede provocar un aumento en el nivel de factor VIII funcional hasta al menos aproximadamente 2%, al menos aproximadamente 3% al menos aproximadamente 4%, al menos aproximadamente 10%, al menos aproximadamente 15%, al menos aproximadamente 20 % o al menos aproximadamente 25% de los niveles normales. En una realización particular, el método de tratamiento provoca un aumento en el nivel de factor VIII funcional hasta al menos aproximadamente 30% de los niveles normales. Un sujeto cuyo nivel de factor VIII funcional ha sido aumentado hasta 30% o más tendrá una coagulación de sangre virtualmente normal.

En una realización, el método de tratamiento provoca un aumento en el nivel de factor VIII funcional, como máximo, hasta los niveles normales.

El nivel de factor VIII funcional puede ser medido de forma relativamente fácil y son conocidos métodos para medir los niveles de factor VIII por los expertos en la técnica. Están disponibles muchos ensayos de coagulación, que incluyen ensayos cromogénicos y basados en la coagulación. Están también ampliamente disponibles ensayos ELISA. Un método particular es medir el nivel de factor VIII: C que es una medida de laboratorio de la actividad coagulante del factor VIII. Un nivel normal de factor VIII: C es 46,8 a 141,8 IU/dl o 0,468-1,4 IU/ml.

Un método adicional es medir el tiempo de tromboplastina parcial activado (aPTT) que es una medida de la capacidad de la sangre para coagular. Un aPTT normal es entre aproximadamente 24 y aproximadamente 34 segundos. Un sujeto que padece hemofilia tendrá un aPTT más largo. Este método puede ser usado en combinación con la medición de tiempo de protrombina.

También se proporciona una molécula de ácido nucleico o vector de la invención para uso en terapia, especialmente

en el tratamiento de hemofilia, particularmente hemofilia A.

Se proporciona también el uso de una molécula de ácido nucleico o vector de la invención en la fabricación de un medicamento para el tratamiento de hemofilia, particularmente hemofilia A.

5 La invención proporciona también un ácido nucleico o vector de la invención para ser usado en el tratamiento del cuerpo humano o animal mediante terapia. En particular, se proporciona un ácido nucleico o vector de la invención para ser usado en el tratamiento de un trastorno de coagulación de la sangre como hemofilia, por ejemplo, hemofilia A. Se proporciona un ácido nucleico o un vector de la invención para ser usado en la mejora de uno o más síntomas de un trastorno de coagulación de la sangre, por ejemplo, reduciendo la frecuencia y/o la gravedad de episodios de hemorragia.

Puede ser usado un método de tratamiento de un trastorno de coagulación de la sangre, método que comprende la etapa de administrar una cantidad eficaz de un ácido nucleico o vector de la invención a un sujeto que lo necesita.

15 Consecuentemente, la invención proporciona adicionalmente el uso de un ácido nucleico o vector como se describe en la presente memoria descriptiva en la fabricación de un medicamento para ser usado en la administración de un ácido nucleico a un sujeto. Además, la invención proporciona un ácido nucleico o vector como se describe en la presente memoria descriptiva en la fabricación de un medicamento para ser usado en el tratamiento de un trastorno de coagulación de la sangre.

Normalmente, un ácido nucleico o vector de la invención puede ser administrado a un sujeto mediante terapia génica, en particular, mediante el uso de un vector de terapia génica parvoviral como AAV. Los métodos generales para una terapia génica son conocidos en la técnica. El vector, composición o composición farmacéutica puede ser suministrado a una célula *in vitro* o *ex vivo* o a un sujeto *in vivo* mediante cualquier método adecuado conocido en la técnica. Alternativamente, el vector puede ser suministrado a una célula *ex vivo*, y la célula puede ser administrada a un sujeto, como es conocido en la técnica. En general, la presente invención puede ser empleada para suministrar cualquier ácido nucleico de la invención a una célula *in vitro*, *ex vivo* o *in vivo*.

30 También se proporciona un método para suministrar un ácido nucleico a una célula. Normalmente, para métodos *in vitro*, el virus puede ser introducido en la célula mediante métodos de transducción viral estándar, como son conocidos en la técnica.

Preferentemente, las partículas de virus son añadidas a las células a la multiplicidad apropiada de infección según métodos de transducción estándar apropiados para las células dianas particulares. Los títulos del virus que va a ser administrado pueden variar, dependiendo del tipo de célula diana y el vector del virus particular y pueden ser determinados por los expertos en la técnica sin experimentación excesiva.

40 Las células pueden ser retiradas de un sujeto, el vector de parvovirus es introducido en las mismas y las células se vuelven a colocar seguidamente en el sujeto. Los métodos para retirar células de un sujeto para un tratamiento *ex vivo*, seguido de nueva introducción en el sujeto, son conocidos en la técnica. Alternativamente, un vector AAV puede ser introducido en las células de otros sujetos, en células cultivadas o en células de cualquier otra fuente adecuada y las células son administradas a un sujeto que lo necesita.

45 Un aspecto adicional es un método para tratar sujetos *in vivo* con un ácido nucleico o vector de la invención. La administración de un ácido nucleico o vector de la presente invención a un sujeto humano o un animal que lo necesita pueden ser cualesquiera medios conocidos en la técnica para administrar vectores de virus.

50 Un ácido nucleico o vector de la invención será incluido normalmente en una composición farmacéutica como se expuso con anterioridad. Estas composiciones incluyen el ácido nucleico o vector en una cantidad eficaz, suficiente para proporcionar un efecto terapéutico o profiláctico deseado y un vehículo o excipiente farmacéuticamente aceptable. Una "cantidad eficaz" incluye una cantidad terapéuticamente eficaz o una cantidad profilácticamente eficaz.

55 Una "cantidad terapéuticamente eficaz" se refiere a una cantidad eficaz, a dosificaciones y durante períodos de tiempo necesarios, para conseguir el resultado terapéutico deseado, como elevar el nivel de factor VIII funcional en un sujeto (con el fin de conducir a una producción de factor VIII funcional hasta un nivel suficiente para mejorar los síntomas de la enfermedad asociada con una falta de esa proteína).

60 Una cantidad terapéuticamente eficaz de una molécula de ácido nucleico o vector de la invención puede variar según factores como el estado de la enfermedad, edad, sexo y peso del individuo y la capacidad de la molécula de ácido nucleico o vector para provocar una respuesta deseada en el individuo. Los regímenes de dosificación pueden ser ajustados para proporcionar la respuesta terapéutica óptima. Una cantidad terapéuticamente eficaz es también normalmente una en la que cualesquiera efectos tóxicos o perjudiciales de la molécula de ácido nucleico o vector sean sobrepasados por los efectos terapéuticos beneficiosos.

65

Los vectores de terapia génica viral pueden ser administrados a una célula u hospedante en una cantidad biológicamente eficaz. Una cantidad "biológicamente eficaz" del vector de virus es una cantidad que sea suficiente para dar lugar a la infección (o transducción) y la expresión de la secuencia de ácido nucleico heterólogo en la célula. Si el virus es administrado a una célula *in vivo* (por ejemplo, el virus es administrado a un sujeto) una cantidad biológicamente eficaz" del vector del virus es una cantidad que sea suficiente para dar lugar a la transducción y expresión del ácido nucleico según la invención en una célula diana.

Para una molécula de ácido nucleico o vector de la invención, como un vector de terapia génica, la dosificación que va a ser administrada puede depender en gran medida del estado y el tamaño del sujeto que esté siendo tratado, así como de la formulación terapéutica, frecuencia de tratamiento y la vía de administración. Los regímenes para continuar la terapia, que incluyen la dosis, formulación y frecuencia, pueden estar guiados por la respuesta inicial y el criterio clínico. La vía parenteral de inyección en el espacio intersticial del tejido puede ser preferida, aunque pueden ser necesarias otras vías parenterales, como la inhalación de una formulación de aerosol en una administración específica. En algunos protocolos, una formulación que comprende el gen y el sistema de suministro de genes en un vehículo acuoso es inyectada en el tejido en cantidades apropiadas.

Ejemplos de modos de administración incluyen la administración oral, rectal, transmucosa, tópica, transdermal, inhalación, parenteral (por ejemplo, intravenosa, subcutánea, intradérmica, intramuscular e intraarticular), y similares, así como una inyección directa en tejidos u órganos alternativamente, inyecciones intratecal, intramuscular directa, intraventricular, intravenosa, intraperitoneal, intranasal o intraocular. Los productos inyectables pueden ser preparados en formas convencionales, como soluciones o suspensiones líquidas, formas sólidas adecuadas para una solución o suspensión en un líquido antes de la inyección o como emulsiones. Alternativamente, se puede administrar el virus de una manera local en lugar de sistémica, por ejemplo, en un depósito o una formulación de liberación sostenida.

El tipo de tejido/célula al que va a ser administrado una molécula de ácido nucleico o vector de la invención puede ser de cualquier tipo pero, normalmente, será una célula hepática/de hígado. No está previsto que la presente invención esté limitada a ninguna vía particular de administración. Sin embargo, con el fin de que las células de hígado sean transducidas, puede ser satisfactoriamente administrada una molécula de ácido nucleico o vector de la presente invención a través de la vasculatura portal o arterial. Alternativamente, la célula puede ser cualquier célula progenitora. Como una alternativa adicional, la célula puede ser una célula madre (por ejemplo, una célula madre del hígado). El tejido diana puede ser específico o puede ser una combinación de diversos tejidos, por ejemplo tejidos de hígado y músculos.

En el caso de un vector de terapia génica, el intervalo de dosis eficaz varía para animales pequeños como ratones, a continuación de una inyección intramuscular, puede ser entre aproximadamente 1×10^{11} y aproximadamente 1×10^{12} copias de genoma (gc)/kg y para animales mayores (gatos) y posiblemente para seres humanos, entre aproximadamente 1×10^{10} y aproximadamente 1×10^{13} gc/kg. Las dosificaciones de vector de terapia génica de parvovirus de la invención dependerán del modo de administración, la enfermedad o estado que esté siendo tratado, el estado del sujeto individual y el vector de virus particular y pueden ser determinadas de una manera rutinaria. Normalmente, puede ser adecuada una cantidad de aproximadamente 10^3 a aproximadamente 10^{16} partículas de virus por dosis. Preferentemente, se usa una cantidad de aproximadamente 10^9 a aproximadamente 10^{14} partículas de virus por dosis. Cuando es tratado de esta forma, un sujeto puede recibir una dosis única de partículas de virus de forma que las partículas virales efectúen el tratamiento en una administración única.

La cantidad de compuesto activo en las composiciones de la invención puede variar según factores como el estado de enfermedad, edad, sexo y peso del individuo. Los regímenes de dosificación pueden ser ajustados para proporcionar la respuesta terapéutica óptima. Por ejemplo, puede ser administrado un bolo único, pueden ser administradas varias dosis divididas a lo largo del tiempo o la dosis puede ser proporcionalmente reducida o aumentada según esté indicado por las exigencias de la situación terapéutica.

Puede ser ventajoso formular composiciones parenterales en forma unitaria de dosificación para una mayor facilidad de la administración y uniformidad de la dosificación. "forma unitaria de dosificación" como se usa en la presente memoria descriptiva, se refiere a unidades físicamente discretas adecuadas como dosificaciones unitarias para los sujetos que van a ser tratados conteniendo cada unidad una cantidad predeterminada de compuesto activo calculada para producir el efecto terapéutico deseado en asociación con el vehículo farmacéutico requerido. La especificación para las formas unitarias de dosificación de la invención puede estar dictada por las características únicas del compuesto activo y el efecto terapéutico particular que se quiere conseguir y por las limitaciones inherentes en la técnica de componer este compuesto activo para el tratamiento de un estado en individuos.

Muchos métodos para la preparación de estas formulaciones están patentados o son ampliamente conocidos por los expertos en la técnica.

Se proporciona adicionalmente un animal transgénico que comprende células que comprenden un vector según la invención. Preferentemente, el animal es un mamífero no humano, especialmente un primate como un macaco. Alternativamente, el animal puede ser un roedor, especialmente un ratón o puede ser canino, felino, ovino o porcino.

En el aspecto de la invención en el que se proporciona un promotor que comprende una secuencia de nucleótidos que tiene una homología sustancial respecto a la secuencia de nucleótidos de la SEC ID N° 3, el promotor tiene preferentemente al menos un 85%, más preferentemente al menos un 90%, incluso más preferentemente al menos un 95% de homología respecto a la secuencia de nucleótidos de la SEC ID N° 3. El promotor tiene preferentemente menos de 400 bp, más preferentemente menos de 350 bp, incluso más preferentemente menos de bp de tamaño.

Se proporciona también una segunda molécula de ácidos nucleicos aislada. La molécula de ácidos nucleicos aislada comprende una primera secuencia de nucleótidos que tiene una identidad de al menos 90% respecto a la secuencia de nucleótidos de la SEC ID N° 1; y una secuencia de nucleótidos que tiene homología sustancial respecto a la secuencia de nucleótidos de la SEC ID N° 3. La expresión "homología sustancial" puede ser adicionalmente definida con referencia al porcentaje de homología. Esto se expone más en detalle en la presente memoria descriptiva.

En la segunda molécula de ácidos nucleicos (también denominadas anteriormente en el primer aspecto de la invención), las dos secuencias pueden ser contiguas o pueden estar separadas por un cierto número de nucleótidos. Por ejemplo, las dos secuencias pueden estar separadas por una secuencia kozak o uno o más intrones. Las secuencias preferentemente están funcionalmente conectadas, es decir, la segunda secuencia, que codifica un promotor, está conectada a la primera secuencia de forma que la primera secuencia puede ser expresada cuando es introducida en una célula usando un vector.

También se proporciona un vector que comprende la segunda molécula de ácidos nucleicos de la invención.

Se proporciona adicionalmente una célula hospedante que comprende un vector según la invención. La célula hospedante puede ser cualquier célula apropiada pero, preferentemente, es una célula de mamífero no humano, especialmente una célula de primate. Pueden ser usadas células para producir la proteína de forma recombinante, y puede ser usada cualquier célula apropiada, como una célula CHO.

Se proporciona adicionalmente un animal transgénico que comprende células que comprenden un vector según la invención. Preferentemente, el animal es un mamífero no humano, especialmente un primate como un macaco. Alternativamente, el animal puede ser un roedor, especialmente un ratón.; o puede ser canino, felino, ovino o porcino.

En la descripción que antecede, el término "homología se usa para hacer referencia a la analogía de dos secuencias. Esto puede ser descrito también usando el término "identidad". Los términos "homología" y "identidad" pueden ser usados de forma intercambiable en la presente memoria descriptiva. Para los fines de esta invención, se define en la presente memoria descriptiva que, con el fin de determinar el porcentaje de identidad de dos secuencias de ácidos nucleicos, las secuencias están alineadas para fines de comparación óptima (por ejemplo, las separaciones pueden ser introducidas en la secuencia de un primer ácido nucleico para una alineación óptima con una segunda secuencia amino o de ácidos nucleicos). Seguidamente se comparan los residuos de nucleótidos en las partes de nucleótidos. Cuando una posición en la primera en la primera secuencia está ocupada por el mismo residuo de aminoácido o nucleótido que la correspondiente posición en la segunda secuencia, entonces las moléculas son idénticas en esa posición. El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias (es decir, % identidad= número de posiciones idénticas/número total de posiciones (es decir, posiciones solapadas) x 100). Preferentemente, las dos secuencias tienen la misma longitud.

Se puede llevar a cabo una comparación de secuencias sobre las longitudes completas de las dos secuencias que están siendo comparadas o sobre un fragmento de las dos secuencias. Normalmente, la comparación se llevará a cabo sobre la longitud completa de las dos secuencias que están siendo comparadas. Sin embargo, la identidad de secuencias se puede llevar a cabo sobre una región, por ejemplo, de aproximadamente veinte, aproximadamente cincuenta, aproximadamente cien, aproximadamente doscientos, aproximadamente quinientos, aproximadamente 1.000, aproximadamente 2.000, aproximadamente 3.000, aproximadamente 4.000, aproximadamente 4.500, aproximadamente 5.000 o más residuos de ácidos nucleicos contiguos.

El experto en la técnica será consiente del hecho de que están disponibles varios programas de ordenador diferentes para determinar la homología entre dos secuencias. Por ejemplo, se puede realizar una comparación de secuencias y una determinación del porcentaje de identidad entre dos secuencias usando un algoritmo matemático. En una realización preferida, el porcentaje de identidad entre dos secuencias aminoácidos o ácidos nucleicos se determina usando el algoritmo de Needleman y Wunsch (J. Mol. Biol. (48): 444-453 (1970)), que ha sido incorporado en el programa GAP en un paquete de software Accelrys GCG (disponible en <http://www.accelrys.com/products/gcg/>), usando una matriz aproximadamente Blosum 62 o una matriz PAM250 y un peso de la separación de 16, 14, 12, 10, 8, 6, o 4 y un peso de la longitud de 1, 2, 3, 4, 5, o 6. El experto en la técnica apreciará que estos diferentes parámetros producirán resultados ligeramente diferentes pero que el porcentaje de identidad global de dos secuencias no está significativamente alterado cuando se usan algoritmos diferentes.

Las secuencias de ácidos nucleicos de la presente invención pueden ser adicionalmente usadas como una "secuencia de consulta" para realizar una búsqueda frente a bases de datos públicas, por ejemplo, para identificar otros miembros de la familia o secuencias relacionadas. Estas búsquedas se pueden realizar usando los programas BLASTN y BLASTP (versión 2.0) de Altschul, et al. Las búsquedas de proteínas de BLAST se pueden realizar también con el programa BLASTP, puntuación = 50, longitud de palabras=3 para obtener secuencias de aminoácidos homologas respecto a las moléculas de proteínas anteriormente descritas. Para obtener alineaciones separadas para fines de comparación se puede utilizar BLAST separada como se describe por Altschul et al., (1997) Nucleic Acids Res. 25(17): 3389-3402. Cuando se utilizan programas BLAST y BLAST separado, pueden ser usados los parámetros por defecto de los respectivos programas (por ejemplo, BLASTP y BLASTN). Véase la página principal de la entidad National Center for Biotechnology Information en <http://www.ncbi.nlm.nih.gov/>.

Salvo que se defina otra cosa, todos los términos técnicos y científicos usados en la presente memoria descriptiva tienen el mismo significado que se entiende comúnmente por un experto en la técnica a la que pertenece esta invención. La terminología usada en la descripción de la invención en la presente memoria descriptiva es para los fines de describir realizaciones particulares solamente y no está previsto que limite la invención.

En este documento y en las reivindicaciones, el verbo "comprender" y sus conjugaciones se usan en su sentido no limitativo para indicar que están incluidos los objetos que siguen a la palabra pero los objetos que no estén específicamente mencionados no están excluidos. Además, la referencia a un elemento mediante el artículo indefinido "uno" o "una" no excluye la posibilidad de que esté presente más de un elemento, salvo que el contexto requiera claramente que hay solamente uno de los elementos. El artículo indefinido "uno" o "una" indica por tanto habitualmente "al menos uno".

Un experto en la técnica apreciará que todos los aspectos de la invención, ya se refieran, por ejemplo al ácido nucleico, al vector a la célula hospedante o al uso, son igualmente aplicables a todos los otros aspectos de la invención. En particular, algunos aspectos del método de tratamiento, por ejemplo, la administración del ácido nucleico o vector, pueden haber sido descritos más en detalle en alguno de los otros aspectos de la invención, por ejemplo, relativos al uso del ácido nucleico o vector para tratar la hemofilia. Sin embargo, el experto en la técnica apreciará cuando que se haya proporcionado más información detallada para un aspecto de la invención, esta información probablemente es igualmente aplicable a otros aspectos de la invención. Por ejemplo, el experto en la técnica apreciará que la descripción relativa a los vectores y las células hospedantes para el primer aspecto de la invención es aplicable a todos los vectores y células hospedantes de la invención. Además, el experto en la técnica apreciará también que la descripción relativa al método de tratamiento es igualmente aplicable al uso del ácido nucleico o vector en el tratamiento de la hemofilia.

La invención se describirá seguidamente en detalle, a modo de ejemplo solamente, con referencia a los dibujos en los cuales:

Figura 1: Expresión de FVIII humano en ratones con hemofilia A. Zona superior: un esquema de vector rAAV que codifica el BDD hFVIII bajo el control del promotor específico del hígado LP1. Zona inferior: actividad de FVIII humano en plasma de ratón a las ocho semanas posteriores a la administración en vena de la cola de dos $\times 10^9$ vg/ratón (n = 4). Ratones inmunológicamente desprotegidos fueron inyectados con excipiente.

Figura 2: A. Actividad de FVIII humano en plasma de ratón a los 30 días posteriores a la administración en vena temporal de 1×10^8 TU de vectores lentivirales que codifican el BDD, N6 o el codop-FVIII bajo el control del promotor SFFV (N = 4). B. Actividad de hFVIII en materia sobrenadante recolectada a partir de células HepG2 transfectadas con plásmido rAAV que codifica variantes de FVIII bajo el control del promotor LP1 o el cassette de expresión más pequeño rAAV HLP-codop-FVIII (N = 3).

Figura 3: A. Rendimiento de rAAV-HLP-codop-hFVIII (n = 5) pseudotipado con cápside de serotipo 5 en comparación con el rendimiento de scAAV-FIX que contiene un cassette auto-complementario de 2,3 kb y vectores de rAAV de cadena única que contienen un cassette de expresión de 4,6 kb. B. gel nativo teñido con bromuro de etidio y C. Análisis Southern en gel alcalino del genoma viral rAAV-HLP-codop-hFVIII derivado de dos preparaciones separadas (1 y 2).

Figura 4: A. Niveles medios de FVIII \pm SEM en plasma de múridos después de una administración única en vena de cola de constructos de rAAV-hFVIII pseudotipados con cápside de serotipo 5 (dosis= 4×10^{11} uvg/ratón, N = 3). B. número de copias provirales medias (\pm SEM) en hígado de múridos transducido con variantes de rAAV5-hFVIII.

Figura 5: A. transferencia Southern: la zona izquierda muestra el digesto doble (Kpn-1) de DNA genómico de hígado derivado de ratones (M1 y M2) transducido con rAAV-HLP-codo-HFVIII. La zona derecha muestra DNA no cortado o el digerido con un cortador único (Not-1). HH y HT = con cantémeros de cabeza a cabeza y cabeza a cola. B. Transferencia Western que muestra una banda única de ~ 210 Kd en el plasma de ratón transducido con rAAV-HLP-codop-HFVIII que no está presente en plasma de ratón inmunológicamente desprotegido o testigo positivo que consiste en FVIII humano recombinante de longitud completa (rhFVIII) diluido con plasma de ratón.

Figura 6: A. relación entre dosis de rAAV5-HLP-codop-hFVIII y niveles de hFVIII en plasma de múridos y número de copias de transgenes a las 6 semanas posteriores a la transferencia génica. B. Características cinéticas de la expresión de hFVIII a continuación de una administración única en vena de cola de 4×10^{12} vg/ratón de rAAV-HLP-codop-hFVIII pseudotipado con cápside de serotipo 5 o 8. Se muestran niveles medios \pm SEM.

Figura 7: A. Actividad de FVIII y nivel de antígeno en ratones F8/- a continuación de una administración única en vena de cola de una dosis baja y elevada de rAAV HLP codop-hFVIII. B y C. tiempo de hemorragia y pérdida de sangre en ratones F8/- a continuación de transferencia génica de rAAV-HLP-codop-hFVIII en comparación con ratones F8/- no transducidos y animales normales y de tipo salvaje.

Figura 8: (A) Representación esquemática de variantes de FVIII humano diseñadas y clonadas en una cadena principal de vector lentiviral SIN. Se diseñaron nueve variantes de FVIII humano diferentes y se clonaron en plásmido de cadena principal de vector lentiviral: BDD FVIII; FVIII humano suprimido en dominio B (tamaño total 4,3 kb). Fugu B de FVIII; BDD FVIII que contiene el dominio B de Fugu (4,9 kb de tamaño total). FVIII N6; FVIII de BDD que contiene el dominio B de N6 humano (tamaño total de 5,0 kb). SQ FVIII; FVIII de BDD que contiene una versión modificada de la secuencia de aminoácidos SQ^m (tamaño total de 4,4 kb). SQ FVIII Fugu B; FVIII de SQ que contiene el dominio B de Fugu entre la secuencia SQ^m para crear la secuencias SQ^a de N-terminal y SQ^b C-terminal (tamaño total de 5,0 kb). SQ FVIII N6; FVIII de SQ que contiene el dominio B humano (tamaño total de 5,1 kb). Constructos SQ FVIII (ca) (tamaño total de 4,4 kb), SQ FVIII Fugu B (ca) (tamaño total de 5,0 kb) y SQ FVIII N6 (co) (tamaño total de 5,1 kb) son la misma estructura de aminoácidos que los constructos SQ FVIII, SQ FVIII Fugu B y SQ FVIII N6, respectivamente, pero son producidos a partir de una secuencia de cDNA optimizada mediante codones. El tamaño del dominio relativo no es exacto. Las rayas en los constructos marcan los sitios de glicosilación (N)-conectados a asparagina en el dominio B solamente. (B) representaciones esquemáticas de secuencias de SQ y SQ modificada; SQ^m, SQ^a y SQ^b. La secuencia de SQ es un puente de 14 aminoácidos entre los dominios a2 y a3 de FVIII creada mediante fusión de Ser 743 y Gln 1638 en el dominio B. La secuencia favorece una escisión intracelular eficaz conteniendo el sitio de reconocimiento de proteasa de 4 aminoácidos RHQR. Se creó una secuencia SQ modificada (SQ^m) que contenía una mutación sin sentido de Lys 1644 a Thr 1644 provocada por la creación de un sitio de enzima de restricción *MluI* en la secuencia de cDNA para la inserción de Fugu y dominios B de N6. SQ^a es la secuencia 11aa creada en el N terminal del dominio B después de la inserción de las secuencias de N6 o de dominio B Fugu en el constructo SQ FVIII. SQ^b es la secuencia de 5 aminoácidos creada en el C terminal del dominio B después de la inserción de las secuencias N6 o de dominio B Fugu en el constructo SQ FVIII, esta secuencia retiene el sitio de reconocimiento de proteasa de 4 aminoácidos. Los sitios de restricción *MluI* se muestran subrayados y la mutación sin sentido de K a T está en la posesión de los aminoácidos del lado izquierdo del sitio de restricción *MluI*.

Figura 9: Actividad de FVIII humano relativa de constructos de FVIII *in vitro* determinada mediante ensayo cromogénico. Se transdujeron 1×10^5 células 293T con diluciones en serie de BDD FVIII, FVIII Fugu B, FVIII N6, SQ FVIII, SQ FVIII Fugu B, SQ FVIII N6, SQ FVIII (co), SQ FVIII Fugu B (co) o SQ FVIII N6 (co). A las 48 horas se cambiaron 500 μ l de medios exentos de suero con medios celulares. Después de 24 horas adicionales se recogieron medios de incubación de todos los pocillos y se valoraron en cuanto a la expresión de factor VIII usando un ensayo basado en cromogénesis la actividad de cofactor de factor VIII. Los resultados fueron seguidamente normalizados en un número de copias por célula determinados mediante qPCR. Se muestran medias y SD para $m = 5$. Las barras anteriores de los valores representan el aumento en veces de la expresión de FVIII a partir de constructos optimizados mediante codones en comparación con secuencias equivalentes no optimizadas mediante codones. Se realizaron análisis estadísticos usando el modelo lineal general (GLM) basado en un análisis de dos vías de la varianza (ANOVA) con comparaciones por pares de individuos realizadas usando ensayos simultáneos de Bonferroni (Minitab software, Myerstown, PA). Los resultados muestran un aumento altamente significativo para SQ FVIII (co) SQ FVIII Fugu B (co) y SQ FVIII N6 (co) en comparación con sus equivalentes no optimizados mediante codones SQ FVIII, SQ FVIII Fugu B, SQ FVIII N6, respectivamente ($P < 0,001$). Además, los resultados para vectores optimizados mediante codones muestran también un aumento significativo para SQ FVIII N6 (co) en comparación con SQ FVIII (co) ($p < 0,0001$) y un aumento significativo para SQ FVIII Fugu B (co) en comparación con SQ FVIII (co) y SQ FVIII N6 (co) ($B < 0,0001$).

Figura 10: Expresión de actividad de FVIII humano *in vivo* en plasma sanguíneo en ratones hemofílicos después de una inyección intravenosa de vectores lentivirales SIN que expresan constructos de FVIII tratados por bioingeniería. Seis a diez ratones neonatos hemofílicos F8^{tm2Kaz} fueron inyectados por vía intravenosa a través de la vena temporal superficial con vectores lentivirales SIN que expresan constructos de FVIII humanos tratados por bioingeniería. Se tomó sangre de los ratones para diversos valores de tiempo durante aproximadamente 250 días y se usó un ensayo cromogénico para calcular la actividad de FVIII humano en plasma sanguíneo tomado de cada ratón como un porcentaje de los niveles humanos normales. (A) SQ FVIII (rombos blancos) frente SQ FVIII (co) (triángulos negros). SQ FVIII Fugu B (rombos blancos) frente a SQ FVIII Fugu B (co) (triángulos negros). SQ FVIII N6 (rombos blancos) frente a SQ FVIII N6 (co) (triángulos negros). Los puntos en las gráficas representan la media. Las barras de error representan la desviación típica. Se realizaron análisis estadísticos usando el modelo lineal general (GLM) basado en un análisis por dos vías de la varianza (ANOVA) con comparaciones emparejadas individuales realizadas usando ensayos simultáneos de Bonferroni (Minitab software, Myerstown, PA).

Figura 11: Niveles de actividad de FVIII *in vivo* en plasma tomado de ratones inyectados con vector que expresa FVIII a partir de secuencias de cDNA optimizadas mediante codones. La actividad de FVIII humano en plasma sanguíneo tomado de ratones inyectados con vector lentiviral que expresa SQ FVIII (co) (círculos grises), SQ FVIII Fugu B (Co) (rombos blancos) y SQ FVIII N6 (co) (triángulos negros) fue cuantificada. Los puntos en los gráficos representan la media, las barras de error representan la SD (desviación típica). No se apreció ninguna diferencia significativa en la expresión entre constructos que expresan dominios B diferentes ($P > 0,5$, ensayo simultáneo de Bonferroni).

Figura 12: Cuantificación de número de copias de vectores en tejidos de ratones hemofílicos después de una inyección intravenosa de vectores lentivirales SIN que expresan constructos de FVIII tratados por bioingeniería. Se tomó tejido de hígado, corazón, pulmón y riñón de ratones sacrificados ~ 250 días después de una inyección neonatal de vectores lentivirales que expresan SQ FVIII, SQ FVIII Fugu B, SQ FVIII N6, SQ FVIII (co), SQ FVIII Fugu B (co) y SQ FVIII N6 (co). Se extrajo DNA genómico y el número de copias virales se determinó usando qPCR. La línea representa la media de todos los puntos. No se observó ninguna diferencia significativa en el número de copias entre cualquier grupo de vectores ($P > 0,1$, ensayo simultáneo de Bonferroni).

Ejemplo 1

Invasado de cassette de expresión de hFVIII en rAAV

Los inventores han establecido que un cassette de expresión de 6,0 kb que contiene el BDD-FVIII cDNA bajo el control del promotor específico de hígado anteriormente descrito (LP1) puede ser eficazmente invasado en vectores rAAV pseudotipados con proteínas de cápsidas de serotipo 5 (rAAV5-LP1-BDD-h FVIII) usando el método de transfección transitoria de 3 plásmidos convencional. La administración en vena de solamente 2×10^9 partículas de rAAV-LP1-BDD-hFVIII en ratones FVIII KO machos adultos dio lugar a una actividad de coagulación de FVIII de $18 \pm 5,3\%$ usando un ensayo cromogénico (Figura 1), que está significativamente por encima del nivel requerido para una mejora de la diátesis hemorrágica en seres humanos ($> 1\%$ de lo normal).

Adaptación escala de producción de vector de rAAV-hFVIII

Los inventores han establecido un método de producción de rAAV adaptable a escala, sencillo y compatible con GMP usando el vector de expresión de vaculovirus y células de insectos. Una ventaja clave del sistema de vaculovirus es la facilidad con que la producción se puede llevar a escala. Ha sido posible generar 1×10^{14} genomas de vectores (vg) a partir de un único ensayo de producción usando un bio-reactor. Esta cantidad sería suficiente para un ensayo clínico de HA de fase I/II. Los rendimientos iniciales en el vector de FVIII de primera generación (rAAV5-LP1-BDD-hFVIII) son del orden de 5×10^{11} uvg a partir de 1 litro de cultivo celular.

Expresión a partir de hFVIII optimizado mediante codones

Los inventores han diseñado un constructo de hFVIII alternativo (codop-FVIII) para ensayar la hipótesis de que sustituyendo codones infrecuentemente usados en el cDNA con los más comúnmente encontrados en genes de mamíferos ("optimización mediante codones") generará una expresión aumentada de hFVIII a continuación de la transferencia génica. Un ejercicio similar para los factores de coagulación IX y VII mejoró la expresión hasta 10 veces en comparación con los cognados de tipo salvaje. La estrategia para el diseño del codop-h FVIII implicó volver a traducir la secuencia de aminoácidos de hFVIII con un conjunto de codones más frecuentemente encontrados en genes de mamíferos altamente expresados. Esta secuencia modificada seguidamente fue cuidadosamente explorada y los codones fueron adicionalmente modificados para mejorar la estabilidad de mRNA y suprimir secuencias no deseables, como nucleótidos de CpG en exceso y sitios de empalme crípticos. La secuencia de codop-h FVIII diseñada final contiene 1076 cambios bp únicos a partir de la secuencia de N6-FVIII de tipo salvaje y es un 42% a+t con relación a un 56% de contenido de A + T de la secuencia de tipo salvaje. La secuencia de FVIII codop es la secuencia de la SEC ID N° 1. Inicialmente, esta variante de codop- FVIII fue clonada en un vector lentiviral en dirección descendente del promotor del virus formador de focos de bazo constitutivo (SFFV) y su potencia es valorada en ratones FVIII KO recién nacidos inyectando 1×10^8 TQ en la vena temporal. Para hacer una comparación, dos grupos separados de ratones FVIII KO recién nacido fueron transducidos con un título equivalente de vector idéntico que codifica las variantes de FVIII BDD o la más potente N6. Como se muestra en la Figura 2^a, y de forma congruente con informes previos, los vectores lentivirales que codifican el N6-FVIII ($15 \pm 0,8\%$ de lo normal) mediaron niveles cinco veces superiores de expresión transgénica en comparación con la variante BDDD ($3 \pm 0,6\%$ de lo normal). En comparación, la expresión de hFVIII en el plasma del grupo de codop-FVIII de ratones ($283 \pm 0,21\%$ de lo normal) era al menos 18 veces mayor que el conseguido con el N6-FVIII. Los inventores han clonado las variantes de FVIII BDD, N6 y codop en su vector rAAV estándar (Nathwani A. et al. Blood. 15 de Febrero de 2007; 109(4): 1414-1421) bajo el promotor LP1. Además, el codop-FVIII ha sido clonado también en dirección descendente de un nuevo promotor específico para el hígado híbrido y más pequeño (HLP). El promotor HLP tiene la secuencia de la SEC ID N° 3. La evaluación de estos plásmidos de vectores rAAV en un ensayo de transfección transitoria en la línea celular de hígado HepG2 (Figura 2b) mostró que los vectores LP1 rAAV que codificaban la expresión de FVIII mediada por codop-FVIII ($0,38 \pm 0,06$ IU/ml) a niveles que estaban entre $4 (0,09 \pm 0,02$ IU/ml) y 8

(0,05 ± 0,02 IU/ml) veces mayor que la conseguida con rAAV-LP1-N6-FVIII y rAAV-LP1-BDD-FVIII, respectivamente. Por lo tanto, colectivamente, estos datos sugieren que la molécula de codop-FVIII de los inventores es más potente que la variante N6-FVIII. Apreciablemente, el plásmido de vector rAAV-HLP-codop-FVIII ligeramente más pequeño genera consistentemente rendimientos entre 30-50% superiores de vector que el rAAV-LP1-codop-FVIII.

5

El cassette de expresión de HLP-codop- h FVIII puede ser envasado en viriones AAV

Los cassettes de expresión de rAAV-HLP-codop-hFVIII de 5,6 kb sobrepasaron el límite de envasado de 4,6 kb de vectores AAV pero fueron satisfactoriamente envasados en viriones AAV con la misma eficacia que el vector scAAV-FIX que está siendo usado en ensayos clínicos continuados (Figura 3A) Usando el método de transfección transitoria HEK293T convencional. Otros han mostrado que una secuencia de vector de 6,6kb puede ser envasado en viriones AAV. Adicionalmente, el grupo del High de la Universidad de Pensilvania verificó independientemente que hasta 6 c 10¹³ partículas pseudotipadas de rAAV 8 de rAAV-HEDP-codop-hFVIII podían ser derivadas a continuación de una transfección transitoria a partir de justo 20 botellas en rodillos de células HEK293 (rendimiento = 6 x 10⁴ uvg/célula 293t). Para demostrar que el genoma del vector rAAV-HLP-codop-hFVIII era envasado en su totalidad, se extrajo DNA de viriones derivados a partir de dos tandas separadas, tratamiento con DNaseI y separadas en agarosa nativa y alcalina y seguidamente valoradas siguiendo un teñido con bromuro de etidio o análisis de transferencia Southern, respectivamente. Se apreció una banda prominente de aproximadamente 5,7 kb con ambos métodos de valoración (Figura 3B y C).

10

El codop-hFVIII es más potente pero igual de seguro que las variantes de hFVIII N6 o BDD

Vectores rAAV pseudotipados con cápside de serotipo 5 que codifican la variante de hFVIII codop N6 y BDD bajo el control de los promotores Lp1 o HLP fueron inyectados a través de la vena de la cola (4 x 10¹¹ uvg/ratón, N = 3/grupo) de ratones C57B1/6 de 4-semanas machos. Como se muestra en la figura 4, una administración única en la vena de cola de rAAV-LP1-codop-hFVIII dio lugar 0,20 ± 0,03 IU/ml (= 20% de los niveles normales) de hFVIII en plasma de muridos sin toxicidad alguna- las expresión de hFVIII fue 10 veces inferior en ratones transducidos con 4 x 10¹¹ uvg/ratón de rAAV-LP1-N6-hFVIII (0,02 ± 0,0003 IU/ml = 2% de lo normal) que codifica la secuencia de DNA de hFVIII de tipo salvaje en lugar de la secuencia de nucleótidos de FVIII optimizada con codones en codop hFVIII. Esta diferencia de expresión entre estos dos vectores que por lo demás son idénticos es altamente significativa (p = 0,0003, ANOVA de una vía). La sustitución del promotor LP1 con el promotor HLP específico de hígado más pequeño dio lugar a niveles marginalmente superiores (0,22 ± 0,04 IU/ml) de hFVIII en el plasma de ratones transducidos con 4 x 10¹¹ uvg/kg de rAAV-HLP-codop-hFVIII en comparación con el grupo de rAAV-LP1-codop-hFVIII pero esta diferencia no fue significativa (p = 0,6). El nivel más bajo de expresión de hFVIII fue observado en el plasma de ratones que recibieron 4 x 10¹¹ uvg/ratón de rAAV LP1-BDD-hFVIII (0,01 ± 0,001 IU/ml), que se aproxima a un 1% de los niveles normales. De forma importante, estas diferencias en el nivel de expresión de hFVIII no estaban relacionadas con el número de copias de vectores, ya que el análisis PCR muestra un número de copias de vectores similar en el DNA genómico extraído del hígado de animales en cada grupo, que varían en el intervalo de 0,9-1 copias provirales/célula. El análisis por transferencia Southern del DNA genómico de los hígados de ratones transducidos con LP1-codop-hFVIII a las seis semanas después de la transferencia génica digerida con Kpn-1, que corta dos veces en el cassette de expresión codop-hFVIII, liberó una banda del tamaño esperado de aproximadamente 1,9 kb (Figura 5 A). La digestión con Not-I, que es un cortador único, generó dos bandas de ~ 5 kb y de ~ 10 kb correspondientes a fragmentos de concatémeros de cabeza a cola y de cabeza a cabeza en una relación de 3:1, respectivamente. Los análisis de transferencia Western mostraron que el codop-hFVIII es secretado en forma de una proteína de 210 kd de cadena única, que como era de esperar tiene un tamaño más pequeño en comparación con el FVIII recombinante de longitud completa (Helixate, Figura 5B, columna izquierda) ya que dos tercios del dominio B han sido suprimidos de codop-hFVIII.

30

35

40

45

50

55

60

65

Seguidamente, se administraron dosis diferentes de rAAV5-HLP-codop-hFVIII a través de la vena de la cola a ratones C57B1/6 machos y se valoraron los niveles de hFVIII en plasma a las 6 semanas. Como se muestra en la Figura 6^a, se observó una relación relativamente lineal entre la dosis de vector, niveles en plasma de hFVIII y número de copias de transgenes sin ninguna evidencia de cinéticas de saturación incluso a los niveles de dosis más elevados. La administración de 4 x 10¹⁰ uvg/ratón de rAAV5-HLP-codop-hFVIII dio lugar a una expresión baja pero detectable de hFVIII a un 0,5% de lo normal. El número de copias de transgenes de rAAV-HLP-codop-hFVIII en el hígado de estos animales fue también 7 veces inferior (0,12 ± 0,06 copias/células) al de 4 x 10¹¹ uvg/ grupo de dosis de ratones. Un aumento en la dosis de vector de 4 x 10¹² uvg/ratón dio lugar a niveles de hFVIII en plasma de aproximadamente un 190% de los niveles fisiológicos (1,9 ± 0,3 IU/ml). El número de copias de transgenes de rAAV-HPL-codop-hFVIII en el hígado de estos ratones rea por encima de 330 veces superior (43,5 ± 2,5 copias provirales/célula) a los niveles observados en animales transducidos con 4 x 10¹¹ uvg/ratón y aproximadamente en el hígado. No se observó ninguna toxicidad a cualquiera de los niveles de dosis y un examen histológico del órgano después de necropsia a las 6 semanas no mostró ninguna patología significativa. El perfil de expresión de transgenes se valoró seguidamente en dos grupos de ratones (n = 3) a continuación de una administración en la vena de la cola de 4 x 10¹² uvg/ratón de rAAV-HLP-codop-hFVIII pseudotipado con proteínas de cápsides de serotipo 5 y 8. Como para los informes previos por los inventores con otros vectores rAAV de cadena única, el hFVIII fue detectable en dos semanas de transferencia génica antes de alcanzar niveles de estado estacionario de 23 + 6

IU/ml y 54 ± 12 IU/ml en 10 semanas en ratones transducidos con rAAV-HLP-codop-hFVIII pseudotipado con cápside de serotipo 5 y 8, respectivamente (Figura 6B). En todos los valores de tiempo, el nivel de hFVIII en el grupo de rAAV8-HLP-codop-hFVIII era entre 2-10 veces superior en comparación con los niveles conseguidos en ratones que recibieron vector pseudotipado de cápside de serotipo 5. Esta diferencia es altamente significativa ($p < 0,001$) y es congruente con diferencias específicas de serotipos similares en transducción mediada por rAAV previamente descritas. Los complejos en plasma de trombina-antitrombina ($2,2 \pm 0,2$ $\mu\text{g/l}$) no eran elevados, indicando que los niveles suprahistológicos de hFVIII no inducen un estado hipercoagulable apreciable en ratones. Finalmente, los anticuerpos anti-FVIII no fueron detectados en los ratones de rAAV-HLP-codop-hFVIII en ninguna fase después de la transferencia génica.

El rAAV-HLP-codon-hFVIII corrige la diátesis de hemorragia en ratones con hemofilia A

Para confirmar la corrección del fenotipo de hemorragia, los inventores inyectaron 4×10^{11} (grupo de dosis baja, $n = 3$) o 5×10^{12} (grupo de dosis elevada, $n = 3$) de genomas de vector rAAV5-HLP-codop-hFVIII en la vena de la cola de ratones con desprotección inmunológica con hemofilia A, que son de fondo C57B16/J-129 Sv y contienen una delección en el exón 16 de FVIII de múridos. Los niveles picos de hFVIII, según se determinaron mediante un ensayo de coagulación en una fase, fueron $137 \pm 27\%$ y $374 \pm 18\%$ de los niveles normales en los grupos de dosis baja y elevada de ratones, respectivamente (Figura 7A). Estos niveles estaban significativamente por encima del fondo (ratones hemofílicos HA sin tratar (FVIIIKO) nivel de hFVIII: $C = < 2\%$ de la normal) y eran significativamente superiores al terapéutico de $>5\%$ de lo normal. Hubo una concordancia muy estrecha entre la actividad de hFVIII y los niveles de antígeno en todos los valores de tiempo examinados con una relación media de 1,16. El tiempo de hemorragia en los ratones F8-/- tratados con AAV y sin tratar, así como los ratones testigos de tipo salvaje, se valoró usando un ensayo de pinza en la cola. El tiempo para la primera detención de la hemorragia en el rAAV5-HLP-codop-hFVIII fue significativamente más corto ($p = 0,003$) a 114 ± 3 y 74 ± 14 segundos en los grupos de dosis baja y elevada, respectivamente, en comparación con los ratones F8-/- sin tratar (311 ± 3 segundos) y comparable al de los animales de tipo salvaje testigos (74 ± 20 segundos). Análogamente, la cantidad de pérdida de sangre valorada mediante un análisis espectrofotométrico del contenido de hemoglobina en solución salina en la que es sumergida la cola del ratón con pinzas fue significativamente inferior ($p = 0,002$) en los ratones rAAV5-HLP-codop-hFVIII en comparación con los animales F8-/- sin tratar. Los anticuerpos anti-h-FVIII no fueron detectados en ratones HA tratados con rAAV en ninguna fase después de la transferencia génica.

Por lo tanto, colectivamente, estos datos sugieren que la molécula de codop-hFVIII es más potente que la variante N6-hFVIII. Adicionalmente, el cassette de expresión de codop-hFVIII parece que está bien envasado en los viriones rAAV a pesar de un tamaño relativamente grande en comparación con el genoma de AAV de tipo salvaje. El hFVIII es expresado como una proteína biológicamente activa de cadena única a continuación de la transferencia génica de rAAV que es capaz de corregir el fenotipo de la hemorragia en animales con hemofilia A inmunológicamente desprotegidos.

Ejemplo 2

Introducción

La hemofilia A es un trastorno de hemorragia grave provocado por una deficiencia o ausencia completa del factor VIII de coagulación sanguínea (FVIII). Es el trastorno de coagulación hereditario más común con una incidencia que se acerca a aproximadamente uno de cada 5.000 varones¹. El trastorno es un candidato atractivo para una terapia génica, porque solamente es necesario un aumento modesto de la concentración en plasma de FVIII para un beneficio terapéutico, y los niveles de $<1\%$ son capaces de conseguir tasas considerablemente reducidas de hemorragia espontánea y artropatía² a largo plazo. Sin embargo, aunque los resultados preclínicos usando una terapia génica en modelos de animales de hemofilia A han sido alentadores, ninguna propuesta se ha traducido todavía en un éxito clínico cuando se han observado niveles insuficientes de expresión de FVIII³.

La baja expresión de FVIII está causada principalmente por una expresión ineficaz del mRNA⁴⁻⁶, una proporción significativa del mal plegado de proteínas con degradación intracelular consiguiente y transporte ineficaz del producto de traducción primario desde el retículo endoplásmico (ER) hasta el aparato de Golgi^{7,8}. Esto da lugar a niveles de expresión de FVIII de aproximadamente 2 a 3 de magnitud inferior a los de otras proteínas secretadas de tamaño comparable⁴. Los estudios durante las dos últimas décadas sobre la trayectoria de secreción, estructura y función de proteína de FVIII y mecanismos del desarrollo inhibitor han conducido a la incorporación de formas tratadas por bioingeniería de FVIII en los sistemas de transferencia génica. La bioingeniería está dirigida a mejorar propiedades como la biosíntesis, eficacia de secreción, actividad funcional, semivida en plasma y a reducir la antigenicidad/inmunogenicidad⁹. El FVIII es producido en forma de una glicoproteína larga de 330 kDa con la estructura de dominios A1-A2-B-A3-C1-C2^{10,11}, en la que los dominios A y C tienen una homología de secuencias internas y una identidad de secuencias de aproximadamente 40% para los dominios A y C de factor V (FV), que comparte la misma estructura de dominios^{12,13}. El dominio B, que constituye un 38% de la secuencia total, no comparte identidad de secuencia de aminoácidos con otras proteínas conocidas, incluido el dominio B de FV. Sin embargo, está extensivamente glicosilado y contiene 19 de los 26 sitios de glicosilación de asparagina unidos a (N)

en la molécula de FVIII completa. El dominio B de FVIII es suministrable para la actividad pro-coagulante. El FVIII en el que es suprimido el dominio B (BDD) y sustituido con un conector corto de 11 aminoácidos (FVIII SQ; Figura 8b) es de uso clínico como un producto de FVIII recombinante de sustitución (Refacto, Wyeth Pharma)¹⁵.

5 Se ha mostrado que la delección del dominio B completo conduce a un aumento de 17 veces en el mRNA y producto de traducción primario, sin embargo, solamente un aumento de 30% en los niveles de proteína secretada, sugiriendo que la velocidad de transporte de PR-Golgi es realmente reducida¹⁶. La secreción eficaz de FVIII requiere un transporte facilitado por hidratos de carbono mediante LMAN1 (enmascaramiento-1 de lectina-manosa) mediado por residuos de manosa de oligosacáridos unidos en N post-translacionalmente unidos al dominio B. para construir sobre las ventajas de BDD-FVIII mientras se ayuda al transporte mediado por LMAN1, Miao *et al.* (2004)¹⁷ volvieron a añadir una secuencia corta de dominio B al BDD-FVIII, óptimamente 226 aminoácidos y reteniendo 6 sitios para la glicosilación unida a N (226/N6). Esto dio lugar a un aumento de 10 veces en la secreción *in vitro* a partir de células COS-1 transfectadas y un aumento de 5 veces *in vivo* a continuación de un suministro de genes hepáticos hidrodinámicos¹⁷.

15 El teleosteo pez globo *Fugu rubripes* es un organismo comúnmente usado para la investigación de características genéticas. El Fugu tiene un genoma de vertebrado básico y contiene un repertorio similar de genes respecto a seres humanos, sin embargo, en 1993 se mostró que el genoma del Fugu es de solamente 390 mb, aproximadamente una octava parte del tamaño del genoma humano¹⁸. Esto hace que el Fugu sea un modelo extremadamente útil para anotar el genoma humano y un genoma "de referencia" valioso para identificar genes y otros elementos funcionales. El análisis de secuencia de genes en el sistema de coagulación sanguínea mostró que las secuencias de aminoácidos del Fugu están altamente conservadas con relación a sus ortólogos humanos. Para secuencias de FVIII cDNA, Los dominios de Fugu A1, A2, A3, C1 y C2 muestran una identidad de secuencias de 46, 43, 47, 52 y 50% respecto a los ortólogos humanos, respectivamente. Inversamente, el dominio B de factor VIII del Fugu comparte solamente un 6% de identidad de secuencias respecto a su correspondiente humano. Sin embargo, aunque no hay una conservación de secuencias evidente entre los dominios B, el dominio B del Fugu está también altamente glicosilado con 11 sitios de unión de glicosilación unidos a asparagina (N) a través de su longitud de 224 aminoácidos¹⁹.

20 En este estudio los inventores examinaron la expresión de constructos BDD FVIII humano que contienen el elemento de dominio B "SQ" anteriormente descrito, el fragmento del dominio B 226/N6 y el dominio B de Fugu. Los constructos fueron ensayados bajo el control de promotor del virus formador de foco de vaso (SFFV en el contexto de un vector lentiviral (LV) HIV-1 auto-inactivante (SIN). Además de ello, los constructos fueron expresados a partir de su secuencia de cDNA optimizada mediante codones o no optimizada mediante codones. Múltiples silenciadores transcripcionales y restos inhibidores están ampliamente distribuidos a través del FVIII cDNA^{4,6;20-22}, y estas secuencias actúan como potentes inhibidores de la producción de RNA y la formación de proteínas que pueden obstaculizar la expresión *in vivo*. La expresión de FVIII en todos los constructos se comparó *in vitro* mediante transducción de células 293T y *in vivo* mediante inyección intravenosa de vector en ratones neonatos con hemofilia A. La variación del dominio B constituyó una diferencia significativa para la expresión de factor VIII a partir de secuencias de cDNA optimizadas mediante codones *in vitro*, sin embargo, no se observaron diferencias *in vivo*. La comparación directa de constructos de FVIII tratados por bioingeniería mostró que se detectaron niveles significativamente mayores (un aumento hasta 44 veces y por encima de 200% de los niveles humanos normales) de proteína de FVIII activa en el plasma de ratones transducidos con vector que expresa FVIII a partir de una secuencia de cDNA optimizada mediante codones, corrigiendo satisfactoriamente el modelo de enfermedad. Hasta la fecha, este es el aumento relativo más elevado en la expresión de FVIII a continuación de un tratamiento por bioingeniería de BDDD FVIII que da lugar a una expresión de FVIII estable sin precedentes *in vivo* usando una propuesta basada en lentivirales.

Métodos

50 Construcción de transgen de FVIII y vector lentiviral

El plásmido de expresión pMT2-FVIII fue obtenido por amable donación del Dr. Steven W. Pipe (University of Michigan). El plásmido contiene el gen FVIII humano con un dominio B de Fugu. El gen HFVIII tenía una delección de dominio B de los aminoácidos 740-1649 y un sitio de restricción *MluI* (ACCGT) tratado por ingeniería genética mediante mutagénesis dirigida al sitio en las posiciones de los aminoácidos 739-740 provocando la mutación sin sentido de Pro739 AThr739 en el dominio a2. El dominio B de Fugu había sido clonado usando sitios de restricción *MluI* en 5' y 3' creando un gen de Fugu B HFVIII de 4935 bp. El gen de Fugu B de FVIII fue suprimido en tres partes usando un digesto con *XhoI* y *KpnI* para suprimir un fragmento de 1,83 kb, un digesto parcial con *KpnI* y *MluI* para suprimir un fragmento de 1,06 kb y una amplificación por PCR de la sección de al menos 2066 kb usando cebadores que crearon sitios *MluI* and *SbfI* en los extremos 5' y 3' respectivamente. Cada sección fue secuencialmente clonada en pLNT/SFFV-MCS usando las mismas enzimas para crear pLNT/SFFV-FVIII Fugu B. El constructo fue completamente secuenciado tras la compleción. Se produjo pLNT/SFFV-BDD FVIII mediante digesto pLNT/SFFV-FVIII Fugu B con *MluI* para separar el dominio B de Fugu y la religadura. La secuencia del dominio B de 226/N6 fue fabricada por la entidad GeneArt (Regensburg, Alemania) para producir un plásmido de GeneArt estándar que contenía 226/N6; pGA_N6_nopt, la secuencia fue obtenida tomando los primeros 688 bp del dominio B de FVIII

humano (cDNA encontrado en la entidad Genbank: A05328), seguidamente fueron añadidos sitios *MluI* que flanquean en 5' y 3'. Seguidamente fue retirado N6 de pGA_N6_nonopt y ligado en pLNT/SFFV-BDD FVIII usando *MluI* para crear pLNT/SFFV N6. La secuencia de cDNA SQ fue obtenida a partir de ²³ y fue modificada para que contuviera un sitio *MluI* (subrayado) para proporcionar la secuencia de cDNA SQ^m: 5'-AGC'TTC'AGC'CAG'AAC'CCC'CCC'GTG'CTG'ACG'CGT'CAC'CAG'CGG-3' (SEC ID N° 8) (Figura 8b). El Fugu B de LNT/SFFV-SQ FVIII fue producido mediante mutagénesis dirigida al sitio realizada por la entidad Eurofins MWG Operon (Ebersberg, Alemania) para añadir secuencias SQ^a y SQ^b flanqueantes (Figura 8b) en el plásmido pLNT/SFFV-FVIII Fugu B para producir pLNT/SFFV Fugu B. El pLNT/SFFV FVIII fue seguidamente producido mediante supresión del dominio B de Fugu de pLNT/SFFV Fugu B mediante digesto con *MluI* y religadura. El pLNT/SFFV FVIII N6 fue producido mediante supresión del dominio B de 226/N6 a partir de from pGA_N6_nonopt mediante digestión con *MluI* y ligadura en pLNT/SFFV FVIII. En este constructo hay una repetición de la secuencia SQ^a 11 ab provocada por la inserción del dominio B de N6 en la secuencia SQ^m. Se crearon secuencias optimizadas mediante codones mediante análisis de cDNA de SQ FVIII Fugu B y adaptación del uso de codones a la desviación de *Homo sapiens* Usando un índice de adaptación de codones (CAI) realizado por la entidad GeneArt (Regensburg, Alemania) usando su Software propiedad de la entidad GeneOptimizer[®]. La optimización suprimió también restos de secuencias de actuación en cis que incluyen cajas de DATA, sitios chi y sitios de entradas ribosomales, extensiones de secuencias con elevado contenido de AT o GC, elementos con elevado contenido de AU, elementos de secuencias inhibitoras y represoras de actuación en cis, secuencias repetidas, estructuras secundarias de RNA y sitios de escisiones crípticas. La optimización de Fugu B de SQ FVIII incluyó la supresión de 14 sitios de escisión, un aumento en el contenido de GC de ~ 45% a ~ 60% y un aumento en CAI desde 0,74 hasta 0,97. Se introdujo una secuencia KOZAK para aumentar el inicio de la traducción, y se añadieron dos codones de parada para asegurar una terminación eficaz. El gen optimizado retuvo en los sitios de restricción *MluI* que flanquean el dominio B en el dominio de Fugu b y tiene una analogía de secuencias de 75,8% respecto a la secuencia original no optimizada. El gen optimizado fue clonado en pLNT/SFFV-MCS para proporcionar el plásmido pLNT/SFFV-SQ Fugu B (co). El plásmido pLNT/SFFV FVIII (co) fue creado mediante digestión de Fugu B de TPP con *MluI* y religadura. La secuencia del dominio B 226/N6 de pGA_N6_nonopt fue optimizada mediante codones y fabricada por la entidad GeneArt. Fue recibida en el plásmido pGA_N6_opt y como los sitios de restricción *MluI* fueron mantenidos clonados directamente en el plásmido pLNT/SFFV FVIII (co) para obtener el constructo pLNT/SFFV N6(co), nuevamente, este constructo contendrá una secuencia repetida SQ_a 11 aa causada por la inserción del dominio b en la secuencia SQ^m. Cada constructo fue completamente secuenciado antes del ensayo. La secuencia SQ FVIII N6 optimizada mediante codones es la secuencia de la SEC ID N° 4. La secuencia SQ FVIII optimizada mediante codones es la secuencia de la SEC ID N° 5. La secuencia de SQ FVIII Fugu B mediante codones es la secuencia de la SEC ID N° 6.

Producción y titulación de vectores lentivirales.

Se produjeron vectores lentivirales con transfección transitoria de células HEK293T (293T) con 3 plásmidos (el vector lentiviral), pMD.G2 [plásmido de envoltura de glicoproteína del virus de estomatosis vesicular (VSV-G)] y pCMVΔ8.91 [plásmido de envoltura, ambos producidos por la entidad Plasmid Factory, Bielefeld, Alemania] empleando polietilenimina (Sigma-Aldrich, Poole, Reino Unido). La materia sobrenadante viral fue recolectada y concentrada usando ultracentrifugación (25.000 x g durante 2 h a 4°C. Se almacenaron partes alícuotas de los virus a -80°C. Los títulos de todos los vectores lentivirales se determinaron usando un estuche de ensayo para ensayo inmunoabsorbente asociado a enzima (ELISA) de transcriptasa inversa (TR) colorimétrico (Roche, West Sussex, Reino Unido) según las instrucciones del fabricante y qPCR para determinar un título aproximado en genomas del vector por ml (vg/ml).

Medición de la actividad de FVIII

La actividad del cofactor de muestras de plasma sanguíneo y medios de cultivos celulares *in vitro* se valoró usando el ensayo cromogénico de factor VIII:C Biophen (Biophen, Quadrtech Diagnostics, Epsom, Reino Unido) como se dice en las instrucciones del fabricante. Las muestras fueron diluidas de 1:20 a 1:40 en el diluyente de muestras proporcionado y se analizaron por duplicado. Se construyó una curva estándar en % de actividad de cofactor VIII diluyendo plasma testigo I (Biophen, Quadrtech Diagnostics) 1:20, llevando a cabo cuatro diluciones en serie 1:2 y realizando un duplicado. Se usó también plasma testigo anormal (Biophen, Quadrtech Diagnostics) como un testigo de calidad adicional para el ensayo.

Transducción lentiviral

Se mantuvieron 293T en medio Eagle modificado de Dulbecco (DMEM) (Life Technologies Ltd, Paisley, Reino Unido) y se complementaron 50 IU/ml de penicilina, 50 µg/ml de estreptomina y 10% de suero de ternera fetal inactivado con calor (FCS; Gibco). Para la transducción lentiviral se transdujeron cinco pocillos de 1 x 10⁵ células 293T con diluciones en serie de vector en un volumen total de 300 µl de DMEM + 10% de FCS. 48 horas después de la transducción, los medios celulares sustituyeron a 500 µl de OptiMEM (Gibco). Después de 24 horas adicionales los medios de incubación fueron recogidos de todos los pocillos y ensayados en cuanto a actividad de factor VIII usando un ensayo cromogénico de FVIII. Seguidamente se extrajo DNA de las fórmulas y el número de copias virales se cuantificó PCR cuantitativo en tiempo real (qPCR).

Métodos *in vivo*

5 Todos los ratones fueron manejados según los procedimientos aprobados por la entidad UK Home Office and the Imperial College London Research Ethics Committee. Ratones con hemofilia A (F8^{tm2Kaz}) generados mediante
 10 deleción de exón 17²⁴ fueron mantenidos en un fonde de 129SV. Ratones neonatales de 0-1 día de edad fueron sometidos a anestesia hipotérmica breve (<5 minutos) y se inyectaron 40 µm de vector lentiviral concentrado (equivalente a 4×10^7 - 1×10^8 unidades de transducción por ratón) en la vena temporal superficial. Para los ensayos del factor de coagulación se recogieron 100 µl de sangre periférica de ratones anestesiados mediante hemorragia en la vena de la cola. La sangre se mezcló inmediatamente en una relación 1:9 con citrato de sodio, se centrifugó a
 15 13.000 fpm en una microcentrifugadora durante cinco minutos y el plasma se transfirió a tubos de microcentrifugadora nuevos y se almacenó a -20°C antes de ensayar.

Determinación del número de copias de vectores mediante PCR cuantitativo en tiempo real

15 Se extrajo DNA genómico de células usando un método estándar de desalación²⁵. La qPCR en tiempo real se llevó a cabo por triplicado para cada muestra para determinar el número de copias virales. Se realizó una qPCR usando un dispositivo (ABI 7000 Sequence Detection System (ABI, Applied Biosystems, Warrington, Reino Unido)). El DNA viral total se cuantificó usando los cebadores -TGTGTGCCCGTCTGTTGTGT-3' (SEC ID N° 9) y 5'-GAGTCTGCGTCGAGAGAGC-3' (SEC ID N° 10) y una sonda Taqman (FAM) 5'-CGCCCGAACAGGACTTGAA-3' (TAMRA) (SEC ID N° 11). El gen titin de ratón (Ttn) fue usado como un testigo de gen de 2 copias endógenas para células de ratón y fue cuantificado usando los cebadores 5'-AAAACGAGCAGTGACCTGAGG-3' (SEC ID N° 12) y 5'-TTCAGTCATGCTGCTAGCGC-3' (SEC ID N° 13) y sonda Taqman (FAM) 5'-TGCACGGAATCTCGTCTCAGTC-3' (TAMRA) (SEC ID N° 14). Se usó el gen de beta-actina humano (ACTV) como testigo génico de dos copias endógenas para células HEK-293T y se cuantificó usando los cebadores 5'-TCACCCACAAGTTGCCCATCTACGA - 3' (SEC ID N° 15) y 5' - CAGCGGAACCGCTCATTGCCAATGG - 3' (SEC ID N° 16) y sonda Taqman (FAM) 5'-ATGCCCTCCCCATGCCATCTGCGT - 3' (TAMRA) (SEC ID N° 17).

Análisis estadístico

30 Los datos se expresan como valores medios más o menos SD (desviación típica). Se realizaron análisis estadísticos usando un modelo lineal general (GLM) basado en un análisis de una vía de la varianza (ANOVA) con comparaciones emparejadas individuales realizadas usando ensayos simultáneos de Bonferroni (Minitab software, Myerstown, PA).

ResultadosGeneración de variantes de FVIII tratadas por bioingeniería y producción de vectores lentivirales SIN que expresan FVIII

40 Para superar la baja expresión de proteínas asociada con las aplicaciones de transferencias génicas de hemofilia A, los inventores investigaron la expresión de transgenes de FVIII tratados por bioingeniería que contenían varios elementos de dominio B de secuencias de cDNA optimizadas mediante codones o no optimizadas mediante codones. Se generaron las siguientes variantes de FVIII (figura 8^a): FVIII humano BDD que contiene una deleción de dominio B entre los aminoácidos 740-1649 con una mutación sin sentido Pro 739 a Thr 739 en el dominio a2
 45 previamente descrito por Miao *et al.* (2004)¹⁷ (denominada en lo sucesivo BDD rFVIII); BDD FVIII que contiene el dominio B de Fugu 201 aa que contiene 11 sitios de glicosilación unidos en N entre los aa (aminoácidos) 740 y 1469(denominado en la presente memoria descriptiva FVIII Fugu B); BDD FVIII que contiene el fragmento de dominio B humano 226 aa N6 que contiene 6 sitios de glicosilación unidos a N los aa 740 y 1649 previamente descritos por Miao *et al.* (2004)¹⁷ denominados en la presente memoria descriptiva FVIII N6); BDD FVIII que contiene un péptido de activación SQ de 14 aminoácidos SQ^{nm} entre los aa 740 y 1649 (SFSQNPPVLTRHQR) (SEC ID N° 18) (mutación sin sentido Lys a Thr subrayada), contiene la secuencia de reconocimiento de furina RHQR para aumentar la escisión intracelular, la secuencia del péptido de activación original descrita por Sandberg *et al.* (2001)²³ (denominada en la presente memoria descriptiva SQ FVIII); SQ FVIII que contiene el dominio B de Fugu insertado en la secuencia SQ^m. Esto provoca que la secuencia SQ^m se escinda en cada lado del inserto del dominio B con la
 55 secuencia N-terminal (SFSQNPPVLTR) (SEC ID N° 19) que se denomina SQ^a y la secuencia C-terminal que contiene el sitio de reconocimiento de furina RHQR como SQ^b (TRHQR) (SEC ID N° 20) (denominada en la presente memoria descriptiva SQ FVIII Fugu B); SQ FVIII que contiene el dominio B de 226 aa/N6 insertado en la secuencia SQ^m creando las secuencias SQ^a y SQ^b en los lados N- y C-terminal del dominio B, respectivamente. En este constructo hay una repetición de la secuencia SQ^a de 11 aa (aminoácidos) provocada por la inserción del dominio B N6 en la secuencia SQ^m. Es desconocido el efecto que esta repetición tendrá sobre la secreción y función de FVIII (denominada en la presente memoria descriptiva SQ FVIII N6). Los constructos "SQ FVIII (ca)", "SQ FVIII Fugu B (cb)" y "SQ FVIII N6 (Cb)" son iguales en la estructura de aminoácidos que los constructos "SQ FVIII", "SQ FVIII Fugu B" y "SQ FVIII N6", respectivamente, pero están traducidos a partir de una secuencia de cDNA optimizada mediante codones (Figura 8a). La representación de SQ, SQ^m, SQ^a y SQ^b se muestra en la Figura 8b. Todos los
 60 constructos fueron clonados en una cadena principal lentiviral SIN bajo el control del promotor SFFV y las secuencias de transgenes fueron confirmadas mediante secuenciado de DNA automatizado.

Los vectores se produjeron para la totalidad de los nueve constructos de factor VIII y se ensayaron en cuanto al título físico usando el ensayo de proteínas de transcriptasa inversa. Seguidamente fueron ensayados usando qPCR para determinar un título aproximado en genomas de vector por ml (uvg/ml) (tabla 1). No hubo diferencia sustancial en el título entre los constructos.

Tabla 1. Título físico de vectores de FVIII determinados mediante ensayo de transcriptasa inversa y qPCR. Cuantificación de concentración de proteína de transcriptasa inversa (RP) en poblaciones virales, medida realizando un ensayo colorimétrico RT, cuantificada en ng/μl y título estimado calculado a partir de esto. Media mostrada de m=3. Cuantificación de título en genomas de vectores por ml se determinó usando qPCR. Se transdujeron 1×10^5 células 93T con un dilución en células de vector, posteriormente se extrajo DNA genómico después de 72 horas a partir de las células y se llevó a cabo una qPCR para WPRE y el gen de doméstico humano β-actina. Media mostrada de n = 5.

Virus	Transcriptasa inversa media (ng/μl)	Título estimado (TU/ml)	Título (vg/ml)
BDD FVIII	10,9	$3,71 \times 10^9$	$1,14 \times 10^8$
FVIII Fugu B	46,5	$1,58 \times 10^{10}$	$1,58 \times 10^9$
FVIII N6	30,7	$1,04 \times 10^{10}$	$1,07 \times 10^9$
SQ FVIII	68,3	$2,32 \times 10^{10}$	$2,91 \times 10^9$
SQ FVIII Fugu B	44,8	$1,52 \times 10^{10}$	$1,18 \times 10^9$
SQ FVIII N6	78,0	$2,65 \times 10^{10}$	$2,0 \times 10^9$
SQ FVIII (co)	69,6	$2,37 \times 10^{10}$	$4,45 \times 10^9$
SQ FVIII Fugu B (co)	71,8	$2,40 \times 10^{10}$	$2,65 \times 10^9$
SQ FVIII N6 (co)	87,9	$2,99 \times 10^{10}$	$3,39 \times 10^9$

Expresión de FVIII *in vitro*

La expresión de proteínas de FVIII relativa se midió para cada constructo en la línea celular de riñón embrionario humano 293T. Las células fueron transducidas con una dilución en serie de vector y cultivadas durante 48 horas, después de lo cual las células fueron lavadas, se añadieron medios libres de suero de nueva aportación y se realizaron ensayos cromogénicos después de 24 horas adicionales para determinar la actividad de FVIII. Se extrajo también DNA genómico de las células para determinar el número de copias virales mediante qPCR. Los valores de la expresión fueron seguidamente normalizados frente al número de copias, permitiendo que se determinaran valores exactos de expresión de proteínas de FVIII por copia de gen (figura 9). Todos los constructos produjeron actividad de FVIII detectable usando un ensayo cromogénico (Figura 9) y antígeno de FVIII mediante ELISA (datos no mostrados).

Las células transducidas con constructos expresados a partir de secuencias cDNA no optimizadas mediante codones produjeron como promedio de 1,40 a 2,89% de actividad de FVIII/ml 24 h/número de copias de vector. No hubo diferencia significativa en la expresión de FVIII entre constructos equivalentes en los que estaban presentes las secuencias de péptidos de activación SQ^m, SQ^a y SQ^b ($T > 0,05$). Además, no hubo ningún aumento significativo en la expresión cuando estaban presentes dominios B de Fugu B o 226/N6 en comparación con constructos de SQ FVIII o BDD FVIII ($p > 0,05$).

Sin embargo, se observó un aumento altamente significativo en la expresión con constructos expresados a partir de secuencias de cDNA optimizadas mediante codones. Las células que expresan SQ FVIII (co), SQ FVIII Fugu B (co) y SQ FVIII N6 (co) produjeron $22,89 \pm 3,68$, $47,20 \pm 2,71$ y $35,8 \pm 2,39\%$ de actividad de FVIII/ml/24 h/ número de copias de vector, respectivamente, un aumento de 13 a 16 veces en comparación con la expresión a partir de secuencias de cDNA equivalentes no optimizadas mediante codones ($P < 0,001$). Se observó también un aumento significativo en la expresión a partir de constructos que contenían los dominio B de Fugu y 226/N6 en comparación con SQ FVIII (co) ($P < 0,0001$), además, la SQ FVIII Fugu B (co) tenía una expresión significativamente mayor que SQ FVIII (co) y SQ FVIII N6 (co) ($P < 0,0001$).

Comparación de la expresión de FVIII *in vivo* después de suministro intravenoso de vector en ratones con hemofilia A neonatos

Se ensayaron *in vivo* cassettes de expresión de FVIII que contienen SQ. Se ensayaron 6 constructos; SQ FVIII, SQ FVIII Fugu B, SQ FVIII N6, SQ FVIII (co), SQ FVIII Fugu B (co) y SQ FVIII N6 (co) mediante inyección intravenosa directa de vector lentiviral en ratones hemofílicos (FVIII KO) neonatos (0-1 día de edad). Todos los ratones

recibieron entre $4,72 \times 10^7$ y $1,78 \times 10^8$ genomas de vectores (vg) con 6 a 10 ratones inyectados con productos de vectores. Se recogieron muestras de plasma sanguíneo a través de hemorragia en la vena de la cola aproximadamente cada 30 días durante un total ~ 250 días. La actividad de FVIII se valoró usando un ensayo cromogénico funcional.

5 El FVIII funcional se detectó en el plasma de todos los ratones transducidos para todos los valores de tiempo (Figura 10). El plasma de ratones transducidos con vector que contenía secuencias de FVIII no optimizadas mediante codones; SQ FVIII, SQ FVIII Fugu B o SQ FVIII N6 contenían una media de $5,72 \pm 2,31\%$, $7,79 \pm 3,66\%$ y $9,53 \pm 2,24\%$ de actividad de FVIII humano normal, respectivamente, para la duración del experimento. La capacidad para coagular rápidamente a continuación de hemorragias de vena de cola indicó que los ratones tratados con secuencias SQ FVIII Fugu B o SQ FVIII N6 eran capaces de conseguir una hemostásis adecuada, sin embargo, cuatro de cada seis ratones inyectados en el grupo de vector SQ FVIII no sobrevivieron, indicando que los niveles de FVIII fueron insuficientes para corregir el fenotipo de hemofilia A de múridos. Ninguno de los grupos de vectores mostró una morbilidad asociada con una baja expresión de FVIII. Para ratones transducidos con vector que contiene secuencias de FVIII cDNA optimizadas mediante codones; SQ FVIII (co), SQ FVIII Fugu B (co) o SQ FVIII N6 (co), se detectaron niveles medios de FVIII a $256,1\% \pm 63,4\%$, $230,2\% \pm 74,1\%$ y $283,7\% \pm 56,2\%$ de actividad normal de FVIII, respectivamente, durante el período del experimento. Es decir, un aumento de 44, 29 y 29 veces en la expresión de SQ FVIII (c), SQ FVIII Fugu B (c) y SQ FVIII N6 (c), respectivamente, en comparación con la expresión a partir de secuencias equivalentes no optimizadas mediante codones ($P < 0,0001$, ensayo simultáneo de Bonferroni). Además, no se observó ninguna pérdida sustancial en la expresión de FVIII en ninguno de los grupos de vectores. De forma importante, no se observó ninguna diferencia significativa en la expresión para constructos que contienen diferentes elementos de dominio B para vectores que contienen secuencias de cDNA optimizadas con codones o no optimizadas con codones (Figura 11).

25 Análisis del número de copias virales en órganos de ratones transducidos

A partir de los 187 y 246 días posteriores a la inyección, los ratones fueron sacrificados para determinar el número de copias de vectores en tejido de hígado, bazo, corazón, pulmón y riñón mediante qPCR en tiempo real (Figura 12). Se detectaron genomas de vectores predominantemente en el tejido del hígado y bazo con copias insignificantes en los tejidos de corazón, pulmón y riñón para todos los ratones en todos los grupos de vectores. El tejido de hígado tomado de ratones transducidos con vector que contiene secuencias de cDNA no optimizadas mediante codones contenía una media de 5,75, 6,97 y 5,25 copias de vectores por célula para SQ FVIII, SQ FVIII Fugu B y SQ FVIII N6, respectivamente. En el tejido de bazo, el número de copias media fue de 1,50, 3,13 y 2,75 copias por célula para SQ FVIII, SQ FVIII Fugu B y SQ FVIII N6, respectivamente. No hubo ninguna diferencia significativa en el número de copias de vectores detectado en tejidos de hígado de animales inyectados con vector que contenía secuencias optimizadas mediante codones ($P > 0,1$, ensayo simultáneo de Bonferroni). El número medio de copias en el tejido de hígado fue detectado a 5,04, 9,17 y 8,80 copias por células y la copia en el tejido de bazo fue de 2,28, 2,57 y 2,60 copias por célula, para SQ FVIII (co), SQ FVIII Fugu B (co) y SQ FVIII N6 (co), respectivamente. En todos los casos, se encontró un número de copias similar en todos los tejidos para los animales independientemente del grupo de vectores.

Explicación

45 La inestabilidad de mRNA, las interacciones con proteínas, las interacciones con proteínas de chaperona residentes en el ER y el requisito de transporte facilitado de hidratos de carbono desde ER hasta el aparato de Golgi significa que el FVIII es expresado a niveles mucho más bajos a partir de células de mamíferos que otras proteínas de tamaño y complejidad similar^{7,26}. Esto ha sido un factor limitante para la producción comercial de FVIII recombinante para una terapia de sustitución y en el éxito de una terapia génica para la hemofilia A. Ha sido incorporado un cierto número de formas tratadas por bioingeniería de FVIII humana en sistemas de transferencia génica y se ha mostrado que tienen una expresión mejorada tanto *in vitro* como *in vivo*. Los constructos de FVIII con supresión de dominio B (BDD) son usados ampliamente en experimentos de transferencia génica ya que no hay pérdida de la función procoagulante del FVIII y su tamaño más pequeño es más fácilmente incorporado en vectores. Una variación de este constructo es un BDD FVIII que contiene la conexión SQ de 14 aminoácidos entre los dominios A2 y A3, actualmente producido como un producto recombinante y comercializado por la entidad Refacto™ (Wyeth)²³. La conexión SQ se ha mostrado previamente que favorece una escisión intracelular eficaz del producto de traducción de cadena única primario de FVIII ya que contiene el sitio de reconocimiento y escisión de furina intracelular^{23,27}. Este constructo ha sido incorporado en vectores de plásmidos en los que ha conferido niveles terapéuticos de expresión²⁸⁻³⁰. Miao *et al.*, in 2004¹⁷ han mostrado también que después de la transfección de plásmidos de células COS-1 de un constructo de BDD FVIII humano que contiene los primeros 226 aminoácidos de los 6 sitios de glicosilación de asparaginas unidos en N fue secretado de forma cuatro veces más eficaz en comparación con BDD FVIII y cinco veces más eficaz *in vivo* a continuación del suministro de genes hepáticos hidrodinámicos¹⁷. Este constructo ha sido actualmente incorporado en muchos vectores de transferencia génica que incluyen plásmidos³¹, vectores lentivirales³² y vectores gamma – retrovirales³³ y es más eficazmente secretado tanto *in vitro*¹⁷; 33-35 como *in vivo*^{17,35}.

65 Una de las limitaciones significativas de la generación de sistemas de suministro de genes virales eficaces para el

tratamiento de la hemofilia A mediante terapia génica es el gran tamaño del FVIII cDNA. El objetivo de este estudio es investigar el efecto de los cassettes de expresión FVIII con diversos constructos de dominio B.

5 Estos consisten en constructo de SQ FVIII, FVIII N6 y BDD FVIII que contienen el dominio B completo del pez globo *Fugu rubripes* que contiene 11 sitios de glicosilación de asparagina unidos en N que potencialmente podrían favorecer un tráfico más eficaz desde el ER hasta el aparato de Golgi y, por lo tanto, que sean más eficazmente secretados. Se investigó también la expresión de estos constructos a partir de secuencias de cDNA que han sido optimizadas mediante codones para la expresión *Homo Sapiens*. Todos los constructos fueron ensayados usando un vector lentiviral SIN, sin embargo, los resultados son aplicables a cualquier sistema de suministro génico. El estudio de los solicitantes encontró que *in vitro* no se encontró ninguna diferencia en la expresión de FVIII entre constructos con o sin la secuencia SQ modificada. La incorporación de regiones de dominio B en constructos no provocó tampoco ningún aumento significativo en la expresión para constructos no optimizados mediante codones en comparación con sus equivalentes con supresión de dominio B. Sin embargo, para secuencias optimizadas mediante codones, se observó una expresión significativamente de SQ FVIII Fugu B (co), así como de SQ FVIII N6 (co) en comparación con SQ FVIII (co). Se observó también un aumento de 13 a 16 veces en la expresión de factor VIII funcional por copia de gen integrado a partir de secuencias optimizadas mediante codones.

10 *In vivo*, después de una inyección neonatal de un número similar de genomas de vectores lentivirales, la presencia de dominio B no afectó significativamente a los niveles de estado estacionario de la actividad de FVIII en circulación para cualquiera de los constructos optimizados mediante codones o no optimizados mediante codones. Sin embargo, se observó un aumento de 29 a 44 veces en los niveles de plasma en estado estacionario de FVIII funcional en ratones con hemofilia A a niveles por encima de 200% de la expresión FVIII humano normal a partir de constructos optimizados mediante codones en comparación con los equivalentes no optimizados mediante codones. De forma importante, estos niveles de FVIII en circulación estaban asociados con una correlación de las diátesis hemorrágicas. Por el contrario, los niveles de actividad de FVIII observados en ratones tratados con cassettes de expresión de FVIII no optimizados mediante codones están asociados con una hemorragia mortal a continuación de hemorragias en la cola.

15 Múltiples silenciadores transcripcionales, y secuencias inhibitoras están ampliamente distribuidos en el FVIII cDNA ^{4,6;21,22} y la expresión aumentada que sigue la optimización mediante codones puede ser debida en parte a la eliminación de estas secuencias. Sin embargo, la delección del dominio B completo que condujo a un aumento de 17 veces en nRNA y producto de traducción primaria solo dio lugar a un aumento de 30% en los niveles de proteína secretada, sugiriendo que la velocidad de transporte de ER-aparato de Golgi se redujo y que los niveles de FVIII nRNA no estaban limitando la expresión. La introducción de múltiples sitios de glicosilación unidos en N que son importantes en el transporte de ER-aparato de Golgi de niveles aumentados de FVIII de FVIII secretado, sugiriendo que la velocidad de transporte de ER-aparato de Golgi puede ser una etapa limitante de la velocidad ¹⁷. Sin embargo, una cantidad significativa de FVIII en el ER nunca transita hasta el compartimento de Golgi debido a un fallo para la acumulación de FVIII correctamente plegado y mal plegado en el ER que puede dar lugar a un deterioro oxidativo y apoptosis sugiriendo quizás que el plegado de FVIII es la etapa limitante de la velocidad en la expresión de FVIII ³⁴.

20 Aunque la estructura secundaria de las proteínas está determinada principalmente por la secuencia de aminoácidos, el plegado de proteínas en la célula está afectado por una gama de factores: estos incluyen la interacción con otras proteínas (chaperonas) y ligandos, translocación a través de la membrana del ER y condiciones Redox. La velocidad de traducción puede afectar también al plegado de proteínas y se ha sugerido que el uso de codones puede ser un mecanismo para regular la velocidad de traducción y permitir así un plegado por etapas de dominios de proteínas individuales ^{36;37}. El FVIII es una proteína de dominios múltiples compleja en la que los segmentos no secuenciales de cadena de polipéptidos naciente pueden interaccionar en el plegado tridimensional. El ribosoma inactivado en codones "raros", por lo tanto, puede conducir a trayectorias de plegado alternativas, generando conformaciones alteradas y una proteína potencialmente mal plegada. Una explicación potencial para el efecto observado de las secuencias optimizadas mediante codones utilizadas en este estudio puede ser que permiten una traducción y transporte eficaz a través de la membrana del E permitiendo que la cadena de polipéptido de FVIII naciente se pliegue correctamente, conduciendo a los niveles aumentados de FVIII secretado observados *in vitro* e *in vivo*.

25 La expresión de >200 no es requerida en pacientes de hemofilia y la producción de estos niveles elevados de FVIII puede ser perjudicial para las células productoras ^{4,34}. Sin embargo, una ventaja principal de la secuencia optimizada es la capacidad de minimizar el número de células genéticamente modificadas necesarias para producir niveles terapéuticos, reduciendo así el riesgo de mutagénesis insercional y los efectos posicionales dependientes del sitio de inserción. También, el uso de elementos promotores fuertes o ubicuos como SFFV que eran previamente necesarios para conducir la expresión elevada de constructos de FVIII podrían ser sustituidos por promotores específicos de tejidos más débiles, que son menos propensos a un silenciado transcripcional ³¹.

30 Los estudios *in vivo* previos han demostrado la expresión de niveles terapéuticos de FVIII *in vivo* en ratones adultos con hemofilia A después de una inyección sistémica de vector ^{32;38-40}, trasplante de células de médula ósea transducidas ^{31;33}, células de médula ósea trasplantadas o transducidas con expresión específica para plaqueta dirigida a diana ^{41;42} y trasplante de células endoteliales resultante de la sangre transducida ⁴³. Sin embargo, los

niveles de expresión de FVIII mediados a partir de estas propuestas han sido bajos (1-5% del humano normal) y la expresión transitoria debida a la formación de anticuerpos neutralizantes. En este estudio se usó un sistema de suministro de genes lentivirales para investigar la expresión de FVIII a partir de constructos de FVIII que contienen diversos dominios B de secuencias de cDNA no optimizadas mediante codones y optimizadas mediante codones. Se observó un aumento enorme en el nivel de FVIII secretado a partir de un gDNA optimizado mediante codones usando este sistema, sin embargo, como este cassette de expresión tiene un tamaño de solamente ~ 5 kb, es aplicable para cualquier sistema de suministro génico viral (incluido AAV) o no viral y permitirá el desarrollo de vectores más seguros y más eficaces para una terapia génica de hemofilia A.

10 **Referencias**

1. Hoyer LW. Hemophilia A. *N.Engl.J.Med.* 1994;330:38-47.
2. High KA. Gene transfer as an approach to treating hemophilia. *Semin.Thromb.Hemost.* 2003;29:107-120.
3. Viiala NO, Larsen SR, Rasko JE. Gene therapy for hemophilia: clinical trials and technical tribulations. *Semin.Thromb.Hemost.* 2009;35:81-92.
4. Lynch CM, Israel DI, Kaufman RJ, Miller AD. Sequences in the coding region of clotting factor VIII act as dominant inhibitors of RNA accumulation and protein production. *Hum.Gene Ther.* 1993;4:259-272.
5. Kaufman RJ, Wasley LC, Davies MV et al. Effect of von Willebrand factor coexpression on the synthesis and secretion of factor VIII in Chinese hamster ovary cells. *Mol.Cell Biol.* 1989;9:1233-1242.
6. Hoeben RC, Fallaux FJ, Cramer SJ et al. Expression of the blood-clotting factor-VIII cDNA is repressed by a transcriptional silencer located in its coding region. *Blood* 1995;85:2447-2454.
7. Domer AJ, Bole DG, Kaufman RJ. The relationship of N-linked glycosylation and heavy chain-binding protein association with the secretion of glycoproteins. *J.Cell Biol.* 1987;105:2665-2674.
8. PIPE SW, Kaufinan RJ. Factor VIII C2 domain missense mutations exhibit defective trafficking of biologically functional proteins. *J.Biol.Chem.* 1996;271:25671-25676.
9. PIPE SW. The promise and challenges of bioengineered recombinant clotting factors. *J.Thromb.Haemost.* 2005;3:1692-1701.
10. Fang H, Wang L, Wang H. The protein structure and effect of factor VIII. *Thrombosis Research* 2007;119:1-13.
11. Lenting PJ, van Mourik JA, Mertens K. The life cycle of coagulation factor VIII in view of its structure and function. *Blood.* 1998;92:3983-3996.
12. Kane WH, Davie EW. Cloning of a cDNA coding for human factor V, a blood coagulation factor homologous to factor VIII and ceruloplasmin. *Proc.Natl.Acad.Sci.U.S.A* 1986;83:6800-6804.
13. Jenny RJ, Pittman DD, Toole JJ et al. Complete cDNA and derived amino acid sequence of human factor V. *Proc.Natl.Acad.Sci.U.S.A* 1987;84:4846-4850.
14. PIPE SW. Functional roles of the factor VIII B domain. *Haemophilia.* 2009
15. Toole JJ, Pittman DD, Orr EC et al. A large region (approximately equal to 95 kDa) of human factor VIII is dispensable for in vitro procoagulant activity. *Proc.Natl.Acad.Sci.U.S.A.* 1986;83:5939-5942.
16. Pittman DD, Marquette KA, Kaufman RJ. Role of the B domain for factor VIII and factor V expression and function. *Blood.* 1994;84:4214-4225.
17. Miao HZ, Sirachainan N, Palmer L et al. Bioengineering of coagulation factor VIII for improved secretion. *Blood.* 2004;103:3412-3419.
18. Brenner S, Elgar G, Sandford R et al. Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature* 1993;366:265-268.
19. Davidson CJ, Hirt RP, Lal K et al. Molecular evolution of the vertebrate blood coagulation network. *Thromb.Haemost.* 2003;89:420-428.
20. Chuah MK, VANDENDRIESSCHE T, Morgan RA. Development and analysis of retroviral vectors expressing human factor VIII as a potential gene therapy for hemophilia A. *Hum.Gene Ther.* 1995;6:1363-1377.

21. Koeberl DD, Halbert CL, Krumm A, Miller AD. Sequences within the coding regions of clotting factor VIII and CFTR block transcriptional elongation. *Hum.Gene Ther.* 1995;6:469-479.
- 5 22. Fallaux FJ, Hoeben RC, Cramer SJ et al. The human clotting factor VIII cDNA contains an autonomously replicating sequence consensus- and matrix attachment region-like sequence that binds a nuclear factor, represses heterologous gene expression, and mediates the transcriptional effects of sodium butyrate. *Mol.Cell Biol.* 1996;16:4264-4272.
- 10 23. Sandberg H, Almstedt A, Brandt J et al. Structural and functional characteristics of the B-domain-deleted recombinant factor VIII protein, r-VIII SQ. *Thromb.Haemost.* 2001;85:93-100.
- 15 24. Chuah MK, Schiedner G, THORREZ L et al. Therapeutic factor VIII levels and negligible toxicity in mouse and dog models of hemophilia A following gene therapy with high-capacity adenoviral vectors. *Blood* 2003;101:1734-1743.
- 25 25. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988;16:1215.
- 20 26. Marquette KA, Pittman DD, Kaufman RJ. A 110-amino acid region within the A1-domain of coagulation factor VIII inhibits secretion from mammalian cells. *J.Biol.Chem.* 1995;270:10297-10303.
27. Lind P, Larsson K, Spira J et al. Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction and biochemical characterization. *Eur.J.Biochem.* 1995;232:19-27.
- 25 28. Doering CB, Denning G, Dooriss K et al. Directed engineering of a high-expression chimeric transgene as a strategy for gene therapy of hemophilia A. *Mol.Ther.* 2009;17:1145-1154.
- 30 29. Doering CB, Healey JF, Parker ET, Barrow RT, Lollar P. High level expression of recombinant porcine coagulation factor VIII. *J.Biol.Chem.* 2002;277:38345-38349.
- 30 30. Ye P, Thompson AR, Sarkar R et al. Naked DNA transfer of Factor VIII induced transgene-specific, species-independent immune response in hemophilia A mice. *Mol.Ther.* 2004;10:117-126.
- 35 31. Dooriss KL, Denning G, Gangadharan B et al. Comparison of factor VIII transgenes bioengineered for improved expression in gene therapy of hemophilia A. *Hum.Gene Ther.* 2009;20:465-478.
- 40 32. Sinn PL, Goreham-Voss JD, Arias AC et al. Enhanced gene expression conferred by stepwise modification of a nonprimate lentiviral vector. *Hum.Gene Ther.* 2007;18:1244-1252.
- 40 33. Ramezani A, Hawley RG. Correction of murine hemophilia A following nonmyeloablative transplantation of hematopoietic stem cells engineered to encode an enhanced human factor VIII variant using a safety-augmented retroviral vector. *Blood* 2009; 114:526-534.
- 45 34. Malhotra JD, Miao H, Zhang K et al. Antioxidants reduce endoplasmic reticulum stress and improve protein secretion. *Proc.Natl.Acad.Sci.U.S.A* 2008;105:18525-18330.
- 50 35. Cerullo V, Seiler MP, Mane V et al. Correction of Murine Hemophilia A and Immunological Differences of Factor VIII Variants Delivered by Helper-dependent Adenoviral Vectors. *Mol.Ther.* 2007;.:
- 50 36. Marin M. Folding at the rhythm of the rare codon beat. *Biotechnol.J.* 2008;3:1047-1057.
- 55 37. Tsai CJ, Sauna ZE, Kimchi-Sarfaty C et al. Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. *J.Mol.Biol.* 2008;383:281-291.
- 55 38. Kootstra NA, Matsumura R, Verma IM. Efficient production of human FVIII in hemophilic mice using lentiviral vectors. *Mol.Ther.* 2003;7:623-631.
- 60 39. Park F, Ohashi K, Kay MA. Therapeutic levels of human factor VIII and IX using HIV-1-based lentiviral vectors in mouse liver. *Blood* 2000;96:1173-1176.
- 60 40. Kang Y, Xie L, Tran DT et al. Persistent expression of factor VIII in vivo following nonprimate lentiviral gene transfer. *Blood.* 2005;106:1552-1558.
- 65 41. Shi Q, Wilcox DA, Fahs SA et al. Factor VIII ectopically targeted to platelets is therapeutic in hemophilia A with high-titer inhibitory antibodies. *J.Clin.Invest* 2006;116:1974-1982.

42. Ohmori T, Mimuro J, Takano K et al. Efficient expression of a transgene in platelets using simian immunodeficiency virus-based vector harboring glycoprotein I α promoter: in vivo model for platelet-targeting gene therapy. *FASEB J.* 2006;20:1522-1524.

5
43. Matsui H, Shibata M, Brown B et al. Ex Vivo Gene Therapy for Hemophilia A That Enhances Safe Delivery and Sustained In Vivo FVIII Expression From Lentivirally-engineered Endothelial Progenitors. *Stem Cells.* 2007.

Listado de secuencias

	<110> UCL Business plc	
	<120> Secuencia codificadora optimizada y promotor Optimised	
5	<130> P524906PCT	
	<150> GB 0911870.4	
	<151> 2009-07-08	
	<160> 21	
	<170> PatentIn version 3.5	
10	<210> 1	
	<211> 4890	
	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
15	<223> Secuencia de nucleótidos de factor VIII modificada optimizada mediante codones	
	<220>	
	<221> CDS	
	<222> (1)..(4890)	
	<400> 1	
	atg cag att gag ctg agc acc tgc ttc ttc ctg tgc ctg ctg agg ttc	48
	Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe	
	1 5 10 15	
	tgc ttc tct gcc acc agg aga tac tac ctg ggg gct gtg gag ctg agc	96
	Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser	
	20 25 30	
	tgg gac tac atg cag tct gac ctg ggg gag ctg cct gtg gat gcc agg	144
	Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg	
	35 40 45	
	ttc ccc ccc aga gtg ccc aag agc ttc ccc ttc aac acc tct gtg gtg	192
	Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val	
	50 55 60	
	tac aag aag acc ctg ttt gtg gag ttc act gac cac ctg ttc aac att	240
	Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile	
	65 70 75 80	
	gcc aag ccc agg ccc ccc tgg atg ggc ctg ctg ggc ccc acc atc cag	288
	Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln	
	85 90 95	
	gct gag gtg tat gac act gtg gtg atc acc ctg aag aac atg gcc agc	336
	Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser	
	100 105 110	
	cac cct gtg agc ctg cat gct gtg ggg gtg agc tac tgg aag gcc tct	384
	His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser	
	115 120 125	
	gag ggg gct gag tat gat gac cag acc agc cag agg gag aag gag gat	432
	Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp	
20	130 135 140	

ES 2 452 890 T3

gac aag gtg ttc cct ggg ggc agc cac acc tat gtg tgg cag gtg ctg Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu 145 150 155 160	480
aag gag aat ggc ccc atg gcc tct gac ccc ctg tgc ctg acc tac agc Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser 165 170 175	528
tac ctg agc cat gtg gac ctg gtg aag gac ctg aac tct ggc ctg att Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile 180 185 190	576
ggg gcc ctg ctg gtg tgc agg gag ggc agc ctg gcc aag gag aag acc Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr 195 200 205	624
cag acc ctg cac aag ttc atc ctg ctg ttt gct gtg ttt gat gag ggc Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly 210 215 220	672
aag agc tgg cac tct gaa acc aag aac agc ctg atg cag gac agg gat Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp 225 230 235 240	720
gct gcc tct gcc agg gcc tgg ccc aag atg cac act gtg aat ggc tat Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr 245 250 255	768
gtg aac agg agc ctg cct ggc ctg att ggc tgc cac agg aag tct gtg Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val 260 265 270	816
tac tgg cat gtg att ggc atg ggc acc acc cct gag gtg cac agc atc Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile 275 280 285	864
ttc ctg gag ggc cac acc ttc ctg gtc agg aac cac agg cag gcc agc Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser 290 295 300	912
ctg gag atc agc ccc atc acc ttc ctg act gcc cag acc ctg ctg atg Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met 305 310 315 320	960
gac ctg ggc cag ttc ctg ctg ttc tgc cac atc agc agc cac cag cat Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His 325 330 335	1008
gat ggc atg gag gcc tat gtg aag gtg gac agc tgc cct gag gag ccc Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro 340 345 350	1056
cag ctg agg atg aag aac aat gag gag gct gag gac tat gat gat gac Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp 355 360 365	1104
ctg act gac tct gag atg gat gtg gtg agg ttt gat gat gac aac agc Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser 370 375 380	1152
ccc agc ttc atc cag atc agg tct gtg gcc aag aag cac ccc aag acc Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr 385 390 395 400	1200
tgg gtg cac tac att gct gct gag gag gag gac tgg gac tat gcc ccc	1248

ES 2 452 890 T3

Trp	Val	His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro		
				405						410				415			
ctg	gtg	ctg	gcc	cct	gat	gac	agg	agc	tac	aag	agc	cag	tac	ctg	aac		1296
Leu	Val	Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn		
			420					425				430					
aat	ggc	ccc	cag	agg	att	ggc	agg	aag	tac	aag	aag	gtc	agg	ttc	atg		1344
Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met		
		435					440					445					
gcc	tac	act	gat	gaa	acc	ttc	aag	acc	agg	gag	gcc	atc	cag	cat	gag		1392
Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu		
	450					455					460						
tct	ggc	atc	ctg	ggc	ccc	ctg	ctg	tat	ggg	gag	gtg	ggg	gac	acc	ctg		1440
Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu		
465					470					475					480		
ctg	atc	atc	ttc	aag	aac	cag	gcc	agc	agg	ccc	tac	aac	atc	tac	ccc		1488
Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro		
				485					490					495			
cat	ggc	atc	act	gat	gtg	agg	ccc	ctg	tac	agc	agg	agg	ctg	ccc	aag		1536
His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys		
			500					505					510				
ggg	gtg	aag	cac	ctg	aag	gac	ttc	ccc	atc	ctg	cct	ggg	gag	atc	ttc		1584
Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe		
		515					520					525					
aag	tac	aag	tgg	act	gtg	act	gtg	gag	gat	ggc	ccc	acc	aag	tct	gac		1632
Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp		
	530					535					540						
ccc	agg	tgc	ctg	acc	aga	tac	tac	agc	agc	ttt	gtg	aac	atg	gag	agg		1680
Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg		
545					550					555				560			
gac	ctg	gcc	tct	ggc	ctg	att	ggc	ccc	ctg	ctg	atc	tgc	tac	aag	gag		1728
Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu		
				565					570					575			
tct	gtg	gac	cag	agg	ggc	aac	cag	atc	atg	tct	gac	aag	agg	aat	gtg		1776
Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val		
			580					585					590				
atc	ctg	ttc	tct	gtg	ttt	gat	gag	aac	agg	agc	tgg	tac	ctg	act	gag		1824
Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu		
		595				600						605					
aac	atc	cag	agg	ttc	ctg	ccc	aac	cct	gct	ggg	gtg	cag	ctg	gag	gac		1872
Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp		
	610					615					620						
cct	gag	ttc	cag	gcc	agc	aac	atc	atg	cac	agc	atc	aat	ggc	tat	gtg		1920
Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val		
625					630					635				640			
ttt	gac	agc	ctg	cag	ctg	tct	gtg	tgc	ctg	cat	gag	gtg	gcc	tac	tgg		1968
Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp		
				645				650					655				
tac	atc	ctg	agc	att	ggg	gcc	cag	act	gac	ttc	ctg	tct	gtg	ttc	ttc		2016
Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe		

ES 2 452 890 T3

660					665					670					
tct ggc tac acc ttc aag cac aag atg gtg tat gag gac acc ctg acc															2064
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr															
675															
680															
685															
ctg ttc ccc ttc tct ggg gag act gtg ttc atg agc atg gag aac cct															2112
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro															
690															
695															
700															
ggc ctg tgg att ctg ggc tgc cac aac tct gac ttc agg aac agg ggc															2160
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly															
705															
710															
715															
720															
atg act gcc ctg ctg aaa gtc tcc agc tgt gac aag aac act ggg gac															2208
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp															
725															
730															
735															
tac tat gag gac agc tat gag gac atc tct gcc tac ctg ctg agc aag															2256
Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys															
740															
745															
750															
aac aat gcc att gag ccc agg agc ttc agc cag aac agc agg cac ccc															2304
Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro															
755															
760															
765															
agc acc agg cag aag cag ttc aat gcc acc acc atc cct gag aat gac															2352
Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp															
770															
775															
780															
ata gag aag aca gac cca tgg ttt gcc cac cgg acc ccc atg ccc aag															2400
Ile Glu Lys Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys															
785															
790															
795															
800															
atc cag aat gtg agc agc tct gac ctg ctg atg ctg ctg agg cag agc															2448
Ile Gln Asn Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser															
805															
810															
815															
ccc acc ccc cat ggc ctg agc ctg tct gac ctg cag gag gcc aag tat															2496
Pro Thr Pro His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr															
820															
825															
830															
gaa acc ttc tct gat gac ccc agc cct ggg gcc att gac agc aac aac															2544
Glu Thr Phe Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn															
835															
840															
845															
agc ctg tct gag atg acc cac ttc agg ccc cag ctg cac cac tct ggg															2592
Ser Leu Ser Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly															
850															
855															
860															
gac atg gtg ttc acc cct gag tct ggc ctg cag ctg agg ctg aat gag															2640
Asp Met Val Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu															
865															
870															
875															
880															
aag ctg ggc acc act gct gcc act gag ctg aag aag ctg gac ttc aaa															2688
Lys Leu Gly Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys															
885															
890															
895															
gtc tcc agc acc agc aac aac ctg atc agc acc atc ccc tct gac aac															2736
Val Ser Ser Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn															
900															
905															
910															
ctg gct gct ggc act gac aac acc agc agc ctg ggc ccc ccc agc atg															2784
Leu Ala Ala Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met															
915															
920															
925															

ES 2 452 890 T3

cct gtg cac tat gac agc cag ctg gac acc acc ctg ttt ggc aag aag Pro Val His Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys 930 935 940	2832
agc agc ccc ctg act gag tct ggg ggc ccc ctg agc ctg tct gag gag Ser Ser Pro Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu 945 950 955 960	2880
aac aat gac agc aag ctg ctg gag tct ggc ctg atg aac agc cag gag Asn Asn Asp Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu 965 970 975	2928
agc agc tgg ggc aag aat gtg agc acc agg agc ttc cag aag aag acc Ser Ser Trp Gly Lys Asn Val Ser Thr Arg Ser Phe Gln Lys Lys Thr 980 985 990	2976
agg cac tac ttc att gct gct gtg gag agg ctg tgg gac tat ggc atg Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly Met 995 1000 1005	3024
agc agc agc ccc cat gtg ctg agg aac agg gcc cag tct ggc tct Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser 1010 1015 1020	3069
gtg ccc cag ttc aag aag gtg gtg ttc cag gag ttc act gat ggc Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly 1025 1030 1035	3114
agc ttc acc cag ccc ctg tac aga ggg gag ctg aat gag cac ctg Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu 1040 1045 1050	3159
ggc ctg ctg ggc ccc tac atc agg gct gag gtg gag gac aac atc Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile 1055 1060 1065	3204
atg gtg acc ttc agg aac cag gcc agc agg ccc tac agc ttc tac Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr 1070 1075 1080	3249
agc agc ctg atc agc tat gag gag gac cag agg cag ggg gct gag Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu 1085 1090 1095	3294
ccc agg aag aac ttt gtg aag ccc aat gaa acc aag acc tac ttc Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe 1100 1105 1110	3339
tgg aag gtg cag cac cac atg gcc ccc acc aag gat gag ttt gac Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp 1115 1120 1125	3384
tgc aag gcc tgg gcc tac ttc tct gat gtg gac ctg gag aag gat Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp 1130 1135 1140	3429
gtg cac tct ggc ctg att ggc ccc ctg ctg gtg tgc cac acc aac Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His Thr Asn 1145 1150 1155	3474
acc ctg aac cct gcc cat ggc agg cag gtg act gtg cag gag ttt Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln Glu Phe 1160 1165 1170	3519

ES 2 452 890 T3

gcc ctg ttc ttc acc atc ttt gat gaa acc aag agc tgg tac ttc	3564
Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe	
1175 1180 1185	
act gag aac atg gag agg aac tgc agg gcc ccc tgc aac atc cag	3609
Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn Ile Gln	
1190 1195 1200	
atg gag gac ccc acc ttc aag gag aac tac agg ttc cat gcc atc	3654
Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His Ala Ile	
1205 1210 1215	
aat gcc tac atc atg gac acc ctg cct gcc ctg gtg atg gcc cag	3699
Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met Ala Gln	
1220 1225 1230	
gac cag agg atc agg tgg tac ctg ctg agc atg gcc agc aat gag	3744
Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu	
1235 1240 1245	
aac atc cac agc atc cac ttc tct gcc cat gtg ttc act gtg agg	3789
Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr Val Arg	
1250 1255 1260	
aag aag gag gag tac aag atg gcc ctg tac aac ctg tac cct ggg	3834
Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly	
1265 1270 1275	
gtg ttt gag act gtg gag atg ctg ccc agc aag gct gcc atc tgg	3879
Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp	
1280 1285 1290	
agg gtg gag tgc ctg att ggg gag cac ctg cat gct gcc atg agc	3924
Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser	
1295 1300 1305	
acc ctg ttc ctg gtg tac agc aac aag tgc cag acc ccc ctg gcc	3969
Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly	
1310 1315 1320	
atg gcc tct gcc cac atc agg gac ttc cag atc act gcc tct gcc	4014
Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly	
1325 1330 1335	
cag tat gcc cag tgg gcc ccc aag ctg gcc agg ctg cac tac tct	4059
Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser	
1340 1345 1350	
ggc agc atc aat gcc tgg agc acc aag gag ccc ttc agc tgg atc	4104
Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile	
1355 1360 1365	
aag gtg gac ctg ctg gcc ccc atg atc atc cat gcc atc aag acc	4149
Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile Lys Thr	
1370 1375 1380	
cag ggg gcc agg cag aag ttc agc agc ctg tac atc agc cag ttc	4194
Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe	
1385 1390 1395	
atc atc atg tac agc ctg gat ggc aag aag tgg cag acc tac agg	4239
Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr Tyr Arg	
1400 1405 1410	
ggc aac agc act gcc acc ctg atg gtg ttc ttt gcc aat gtg gac	4284

ES 2 452 890 T3

Gly	Asn	Ser	Thr	Gly	Thr	Leu	Met	Val	Phe	Phe	Gly	Asn	Val	Asp	
	1415					1420					1425				
agc	tct	ggc	atc	aag	cac	aac	atc	ttc	aac	ccc	ccc	atc	att	gcc	4329
Ser	Ser	Gly	Ile	Lys	His	Asn	Ile	Phe	Asn	Pro	Pro	Ile	Ile	Ala	
	1430					1435					1440				
aga	tac	atc	agg	ctg	cac	ccc	acc	cac	tac	agc	atc	agg	agc	acc	4374
Arg	Tyr	Ile	Arg	Leu	His	Pro	Thr	His	Tyr	Ser	Ile	Arg	Ser	Thr	
	1445					1450					1455				
ctg	agg	atg	gag	ctg	atg	ggc	tgt	gac	ctg	aac	agc	tgc	agc	atg	4419
Leu	Arg	Met	Glu	Leu	Met	Gly	Cys	Asp	Leu	Asn	Ser	Cys	Ser	Met	
	1460					1465					1470				
ccc	ctg	ggc	atg	gag	agc	aag	gcc	atc	tct	gat	gcc	cag	atc	act	4464
Pro	Leu	Gly	Met	Glu	Ser	Lys	Ala	Ile	Ser	Asp	Ala	Gln	Ile	Thr	
	1475					1480					1485				
gcc	agc	agc	tac	ttc	acc	aac	atg	ttt	gcc	acc	tgg	agc	ccc	agc	4509
Ala	Ser	Ser	Tyr	Phe	Thr	Asn	Met	Phe	Ala	Thr	Trp	Ser	Pro	Ser	
	1490					1495					1500				
aag	gcc	agg	ctg	cac	ctg	cag	ggc	agg	agc	aat	gcc	tgg	agg	ccc	4554
Lys	Ala	Arg	Leu	His	Leu	Gln	Gly	Arg	Ser	Asn	Ala	Trp	Arg	Pro	
	1505					1510					1515				
cag	gtc	aac	aac	ccc	aag	gag	tgg	ctg	cag	gtg	gac	ttc	cag	aag	4599
Gln	Val	Asn	Asn	Pro	Lys	Glu	Trp	Leu	Gln	Val	Asp	Phe	Gln	Lys	
	1520					1525					1530				
acc	atg	aag	gtg	act	ggg	gtg	acc	acc	cag	ggg	gtg	aag	agc	ctg	4644
Thr	Met	Lys	Val	Thr	Gly	Val	Thr	Thr	Gln	Gly	Val	Lys	Ser	Leu	
	1535					1540					1545				
ctg	acc	agc	atg	tat	gtg	aag	gag	ttc	ctg	atc	agc	agc	agc	cag	4689
Leu	Thr	Ser	Met	Tyr	Val	Lys	Glu	Phe	Leu	Ile	Ser	Ser	Ser	Gln	
	1550					1555					1560				
gat	ggc	cac	cag	tgg	acc	ctg	ttc	ttc	cag	aat	ggc	aag	gtg	aag	4734
Asp	Gly	His	Gln	Trp	Thr	Leu	Phe	Phe	Gln	Asn	Gly	Lys	Val	Lys	
	1565					1570					1575				
gtg	ttc	cag	ggc	aac	cag	gac	agc	ttc	acc	cct	gtg	gtg	aac	agc	4779
Val	Phe	Gln	Gly	Asn	Gln	Asp	Ser	Phe	Thr	Pro	Val	Val	Asn	Ser	
	1580					1585					1590				
ctg	gac	ccc	ccc	ctg	ctg	acc	aga	tac	ctg	agg	att	cac	ccc	cag	4824
Leu	Asp	Pro	Pro	Leu	Leu	Thr	Arg	Tyr	Leu	Arg	Ile	His	Pro	Gln	
	1595					1600					1605				
agc	tgg	gtg	cac	cag	att	gcc	ctg	agg	atg	gag	gtg	ctg	ggc	tgt	4869
Ser	Trp	Val	His	Gln	Ile	Ala	Leu	Arg	Met	Glu	Val	Leu	Gly	Cys	
	1610					1615					1620				
gag	gcc	cag	gac	ctg	tac	tga									4890
Glu	Ala	Gln	Asp	Leu	Tyr										
	1625														

<210> 2

<211> 1629

<212> PRT

5 <213> Secuencia Artificial

<220>

<223> Proteína de factor VIII modificado

<400> 2

ES 2 452 890 T3

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

Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser
20 25 30

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
85 90 95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser
100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
165 170 175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
180 185 190

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

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
225 230 235 240

ES 2 452 890 T3

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
 245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
 260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
 275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
 290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
 305 310 315 320

Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
 325 330 335

Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
 340 345 350

Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365

Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380

Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400

Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415

Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
 420 425 430

Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
 435 440 445

Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu
 450 455 460

Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
 465 470 475 480

Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
 485 490 495

His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys

ES 2 452 890 T3

Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp
770 775 780

Ile Glu Lys Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys
785 790 795 800

Ile Gln Asn Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser
805 810 815

Pro Thr Pro His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr
820 825 830

Glu Thr Phe Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn
835 840 845

Ser Leu Ser Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly
850 855 860

Asp Met Val Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu
865 870 875 880

Lys Leu Gly Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys
885 890 895

Val Ser Ser Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn
900 905 910

Leu Ala Ala Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met
915 920 925

Pro Val His Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys
930 935 940

Ser Ser Pro Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu
945 950 955 960

Asn Asn Asp Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu
965 970 975

Ser Ser Trp Gly Lys Asn Val Ser Thr Arg Ser Phe Gln Lys Lys Thr
980 985 990

Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly Met
995 1000 1005

Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser
1010 1015 1020

ES 2 452 890 T3

Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly
 1025 1030 1035

Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu
 1040 1045 1050

Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile
 1055 1060 1065

Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr
 1070 1075 1080

Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu
 1085 1090 1095

Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe
 1100 1105 1110

Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp
 1115 1120 1125

Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp
 1130 1135 1140

Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His Thr Asn
 1145 1150 1155

Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln Glu Phe
 1160 1165 1170

Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe
 1175 1180 1185

Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn Ile Gln
 1190 1195 1200

Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His Ala Ile
 1205 1210 1215

Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met Ala Gln
 1220 1225 1230

Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu
 1235 1240 1245

Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr Val Arg
 1250 1255 1260

ES 2 452 890 T3

Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly
 1265 1270 1275
 Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp
 1280 1285 1290
 Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser
 1295 1300 1305
 Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly
 1310 1315 1320
 Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly
 1325 1330 1335
 Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser
 1340 1345 1350
 Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile
 1355 1360 1365
 Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile Lys Thr
 1370 1375 1380
 Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe
 1385 1390 1395
 Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr Tyr Arg
 1400 1405 1410
 Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn Val Asp
 1415 1420 1425
 Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile Ile Ala
 1430 1435 1440
 Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg Ser Thr
 1445 1450 1455
 Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys Ser Met
 1460 1465 1470
 Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln Ile Thr
 1475 1480 1485
 Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser Pro Ser
 1490 1495 1500
 Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp Arg Pro

ES 2 452 890 T3

acccggegggt actacctggg cgccgtggag ctgtcctggg actacatgca gagcgacctg 120
 ggcgagetgc ccgtggacgc ccggttcccc cccagagtgc ccaagagctt ccccttcaac 180
 accagcgtgg tgtacaagaa aaccctgttc gtggagtcca ccgaccacct gttcaatatac 240
 gccaaagccca ggccccctg gatgggcctg ctgggccccca ccatccaggc cgaggtgtac 300
 gacaccgtgg tgatcacctt gaagaacatg gccagccacc ccgtgagcct gcacgccgtg 360
 ggcgtgagct actggaaggc cagcaggggc gccgagtacg acgaccagac cagccagcgg 420
 gagaaaagag atgacaaggt gttccctggc ggcagccaca cctacgtgtg gcaggtgctg 480
 aaagaaaacg gccccatggc ctccgacccc ctgtgcctga cctacageta cctgagccac 540
 gtggacctgg tgaaggacct gaacagcggc ctgatcggcg ctctgctcgt ctgccgggag 600
 ggcagcctgg ccaaagagaa aaccagacc ctgcacaagt tcatcctgct gttcgccgtg 660
 ttcgacgagg gcaagagctg gcacagcggc acaaagaaca gcctgatgca ggaccgggac 720
 gccgcctctg ccagagcctg gcccaagatg cacaccgtga acggctacgt gaacagaagc 780
 ctgcccgccc tgattggctg ccaccggaag agcgtgtact ggcacgtgat cggcatgggc 840
 accacacccg aggtgcacag catctttctg gaagggcaca cctttctggt ccggaaccac 900
 cggcagggcca gcctggaaat cagccctatc accttctga ccgcccagac actgctgatg 960
 gacctgggcc agttcctgct gttttgccac atcagctctc accagcacga cggcatggaa 1020
 gcctacgtga aggtggactc ttgccccgag gaaccccagc tggggatgaa gaacaacgag 1080
 gaagccgagg actacgacga cgacctgacc gacagcgaga tggacgtggt gcggttcgac 1140
 gacgacaaca gccccagett catecagatc agaagcgtgg ccaagaagca ccccaagacc 1200
 tgggtgcact atatcgccgc cgaggaagag gactgggact acgccccctt ggtgctggcc 1260
 cccgacgaca gaagctacaa gagccagtac ctgaacaatg gccccagcg gatcggccgg 1320
 aagtacaaga aagtgcggtt catggcctac accgacgaga cattcaagac ccgggagggc 1380
 atccagcacg agagcggcat cctgggcccc ctgctgtacg gcgaagtggg cgacacactg 1440
 ctgatcatct tcaagaacca ggetagccgg ccctacaaca tctacccccca cggcatcacc 1500
 gacgtcgggc ccctgtacag caggcggctg cccaagggcg tgaagcacct gaaggacttc 1560
 cccatcctgc ccggcgagat ctcaagtac aagtggaccg tgaccgtgga ggacggcccc 1620
 accaagagcg accccagatg cctgacccgg tactacagca gcttcgtgaa catggaacgg 1680
 gacctggcct ccgggctgat cggacctctg ctgatctgct acaaagaaag cgtggaccag 1740
 cggggcaacc agatcatgag cgacaagcgg aacgtgatcc tgttcagcgt gttcgatgag 1800
 aaccggtcct ggtatctgac cgagaacatc cagcggtttc tgcccaacc tgccggcgtg 1860
 cagctggaag atcccagtt ccaggccagc aacatcatgc actccatcaa tggctacgtg 1920
 ttcgaactctc tgcagctctc cgtgtgtctg cacgaggtgg cctactggta catcctgagc 1980

ES 2 452 890 T3

atcggcgccc agaccgactt cctgagcgtg ttcttcagcg gctacacctt caagcacaag 2040
 atggtgtacg aggacaccct gaccctgttc cctttcagcg gcgagacagt gttcatgagc 2100
 atggaaaacc cgggcctgtg gattctgggc tgccacaaca gcgacttccg gaaccggggc 2160
 atgaccgccc tgctgaaggt gtccagctgc gacaagaaca cgggcgacta ctacgaggac 2220
 agctacgagg atatcagcgc ctacctgctg tccaagaaca acgccatcga accccggagc 2280
 ttcagccaga accccccctg gctgacgctg agcttcagcc agaacagccg gcaccccagc 2340
 acccggcaga agcagttcaa cgccaccacc atccccgaga acgacatcga gaaaaccgac 2400
 ccttggtttg cccaccggac ccccatgccc aagatccaga acgtgtccag cagcgacctg 2460
 ctgatcgtgc tgcggcagag ccccaccctt cagggcctga gcctgagcga cctgcaggaa 2520
 gccaaagtacg agacattcag cgacgacccc agccctggcg ccatcgacag caacaacagc 2580
 ctgtccgaga tgaccactt ccggccccag ctgcaccaca gggcgacat ggtgttcacc 2640
 cccgagagcg gcctgcagct gcggtgaac gagaagctgg gcaccaccgc cgccaccgag 2700
 ctgaagaagc tggacttcaa ggtctccagc accagcaaca acctgatcag caccatcccc 2760
 agcgacaacc tggccctgtg caccgacaac accagcagcc tgggccctcc cagcatgccc 2820
 gtgcactacg acagccagct ggacaccacc ctgttcggca agaagtccag ccccctgacc 2880
 gagtccggcg gaccctgtc cctgagcggg gaaaacaacg acagcaagct gctggaagc 2940
 ggctgatga acagccagga aagcagctgg ggcaagaatg tgtccagcac gcgtcaccag 3000
 cgggagatca cccggacaac cctgcagtcc gaccaggaag agatcgatta cgacgacacc 3060
 atcagcgtgg agatgaagaa agaggatttc gatatctacg acgaggacga gaaccagagc 3120
 cccagaagct tccagaagaa aaccggcac tacttcattg ccgccgtgga gaggtgtgg 3180
 gactacggca tgagttctag cccccacgtg ctgcggaacc gggcccagag cggcagcgtg 3240
 ccccagttca agaaagtggg gttccaggaa ttcacagacg gcagcttcac ccagcctctg 3300
 tatagaggcg agctgaacga gcacctgggg ctgctggggc cctacatcag ggccgaagtg 3360
 gaggacaaca tcatggtgac cttccggaat caggccagca gaccctactc cttctacagc 3420
 agcctgatca gctacgaaga ggaccagcgg cagggcgccg aaccgccgaa gaacttcgtg 3480
 aagcccaacg aaaccaagac ctacttctgg aaagtgcagc accacatggc ccccaccaag 3540
 gacgagttcg actgcaaggc ctgggcctac ttcagcagc tggatctgga aaaggacgtg 3600
 cactctggac tgattggccc actcctggtc tgccacacta acaccctcaa ccccggccac 3660
 ggccgccagc tgaccgtgca ggaattcgcc ctgttcttca ccatcttcga cgagacaaag 3720
 tcctgggtact tcaccgagaa tatggaacgg aactgcagag ccccctgcaa catccagatg 3780
 gaagatccta ccttcaaaga gaactaccgg ttccacgcca tcaacggcta catcatggac 3840
 acctgcctg gcctggtgat ggcccaggac cagagaatcc ggtggtatct gctgtccatg 3900
 ggcagcaacg agaatatcca cagcatccac ttcagcggcc acgtgttcac cgtgcggaag 3960

ES 2 452 890 T3

aaagaagagt acaagatggc cctgtacaac ctgtaccccg gcgtgttcga gacagtggag 4020
atgctgcca gcaaggccgg catctggcgg gtggagtgtc tgatcggcga gcacctgcac 4080
gctggcatga gcacctgtt tctggtgtac agcaacaagt gccagacccc actgggcatg 4140
gcctctggcc acatccggga cttccagatc accgcctccg gccagtacgg ccagtgggcc 4200
cccaagctgg ccagactgca ctacagcggc agcatcaacg cctgggtccac caaagagccc 4260
ttcagctgga tcaaggtgga cctgctggcc cctatgatca tccacggcat taagaccag 4320
ggcgccaggc agaagttcag cagcctgtac atcagccagt tcatcatcat gtacagcctg 4380
gacggcaaga agtggcagac ctaccggggc aacagcaccg gcacctgat ggtgttcttc 4440
ggcaatgtgg acagcagcgg catcaagcac aacatcttca accccccat cattgcccgg 4500
tacatccggc tgcaccccac ccaactacagc attagatcca cactgagaat ggaactgatg 4560
ggctgcgacc tgaactcctg cagcatgcct ctgggcatgg aaagcaaggc catcagcgcac 4620
gcccagatca cagccagcag ctacttcacc aacatgttcg ccacctggtc cccctccaag 4680
gccaggctgc acctgcaggg ccggtccaac gcctggcggc ctcaggtcaa caaccccaaa 4740
gaatggctgc agtgaggact tcagaaaacc atgaaggtga ccggcgtgac caccagggc 4800
gtgaaaagcc tgetgaccag catgtacgtg aaagagtttc tgatcagcag ctctcaggat 4860
ggccaccagt ggacctgtt ctttcagaac ggcaaggtga aagtgttcca gggcaaccag 4920
gactcctca cccccgtggt gaactccctg gacccccccc tgctgaccgg ctacctgaga 4980
atccaccccc agtcttgggt gcaccagatc gccctcagga tggaaagtccct gggatgtgag 5040
gcccaggatc tgtactgatg a 5061

<210> 5

<211> 4377

<212> DNA

5 <213> Secuencia Artificial

<220>

<223> Secuencia de nucleótidos de factor VIII modificada optimizada mediante codones

<400> 5

atgcagatcg agctgtccac ctgctttttt ctgtgcctgc tgcggttctg cttcagcgcc	60
acccggcggg actacctggg cgccgtggag ctgtcctggg actacatgca gagcgacctg	120
ggcgagctgc ccgtggacgc ccggttcccc cccagagtgc ccaagagctt ccccttcaac	180
accagcgtgg tgtacaagaa aaccctgttc gtggagttca ccgaccacct gttcaatate	240
gccaagccca ggccccctg gatgggcctg ctgggcccga ccatccaggc cgagggtgtac	300
gacaccgtgg tgatcacctt gaagaacatg gccagccacc ccgtgagcct gcacgccgtg	360
ggcgtgagct actggaaggc cagcgagggc gccgagtacg acgaccagac cagccagcgg	420
gagaaagaag atgacaaggt gttccctggc ggcagccaca cctacgtgtg gcagggtgctg	480

ES 2 452 890 T3

aaagaaaacg gccccatggc ctccgacccc ctgtgcctga cctacagcta cctgagccac 540
gtggacctgg tgaaggacct gaacagcggc ctgatcggcg ctctgctcgt ctgccgggag 600
ggcagcctgg ccaaagagaa aaccagacc ctgcacaagt tcatcctgct gttcgcctg 660
ttcgacgagg gcaagagctg gcacagcggg acaaagaaca gcctgatgca ggaccgggac 720
gccgcctctg ccagagcctg gcccaagatg cacaccgtga acggctacgt gaacagaagc 780
ctgcccggcc tgattggctg ccaccggaag agcgtgtact ggcacgtgat cggcatgggc 840
accacacccc aggtgcacag catctttctg gaagggcaca cctttctggt cgggaaccac 900
cggcaggcca gcctggaaat cagccctatc accttcctga ccgccagac actgctgatg 960
gacctgggcc agttcctgct gttttgccac atcagctctc accagcacga cggcatggaa 1020
gcctacgtga aggtggactc ttgccccgag gaaccccagc tgcggatgaa gaacaacgag 1080
gaagccgagg actacgacga cgacctgacc gacagcgaga tggacgtggt gcggttcgac 1140
gacgacaaca gccccagctt catccagatc agaagcgtgg ccaagaagca cccaagacc 1200
tgggtgcact atatcgccgc cgaggaagag gactgggact acgccccctt ggtgctggcc 1260
cccagcgaca gaagctacaa gagccagtac ctgaacaatg gccccagcg gatcggccgg 1320
aagtacaaga aagtgcggtt catggcctac accgacgaga cattcaagac ccgggaggcc 1380
atccagcacg agagcggcat cctgggcccc ctgctgtacg gcgaagtggg cgacacactg 1440
ctgatcatct tcaagaacca ggctagccgg ccctacaaca tctaccccca cggcatcacc 1500
gacgtgcggc ccctgtacag caggcggctg cccaagggcg tgaagcacct gaaggacttc 1560
cccacctctg ccggcgagat cttcaagtac aagtggaccg tgaccgtgga ggacggcccc 1620
accaagagcg accccagatg cctgaccggg tactacagca gcttcgtgaa catggaacgg 1680
gacctggcct ccgggctgat cggacctctg ctgatctgct acaaagaaa gctggaccag 1740
cggggcaacc agatcatgag cgacaagcgg aacgtgatcc tgttcagcgt gttcgatgag 1800
aaccggctct ggtatctgac cgagaacatc cagcggtttc tgcccaacct tgccggcgtg 1860
cagctggaag atcccagatt ccaggccagc aacatcatgc actccatcaa tggctacgtg 1920
ttcgactctc tgcagctctc cgtgtgtctg cagcaggtgg cctactggta catcctgagc 1980
atcggcgcce agaccgactt cctgagcgtg ttcttcagcg gctacacctt caagcacaag 2040
atggtgtacg aggacacct gacctgttc cctttcagcg gcgagacagt gttcatgagc 2100
atgaaaaacc ccggcctgtg gattctgggc tgccacaaca gcgacttccg gaaccggggc 2160
atgaccgccc tgctgaaggt gtccagctgc gacaagaaca ccggcgacta ctacgaggac 2220
agctacgagg atatcagcgc ctacctgctg tccaagaaca acgcatcga accccggagc 2280
ttcagccaga accccccctg gctgacgcgt caccagcggg agatcaccoc gacaacctg 2340
cagtcggacc aggaagagat cgattacgac gacacatca gcgtggagat gaagaaagag 2400
gatttcgata tctacgacga ggacgagaac cagagcccca gaagcttcca gaagaaaacc 2460

ES 2 452 890 T3

cggcactact tcattgccgc cgtggagagg ctgtgggact acggcatgag ttctagcccc 2520
 cacgtgctgc ggaaccgggc ccagagcggc agcgtgcccc agttcaagaa agtgggtgtc 2580
 caggaattca cagacggcag cttcaccag cctctgtata gaggcgagct gaacgagcac 2640
 ctggggctgc tggggcccta catcagggcc gaagtggagg acaacatcat ggtgaccttc 2700
 cggaatcagg ccagcagacc ctactccttc tacagcagcc tgatcagcta cgaagaggac 2760
 cagcggcagg gcgccgaacc ccggaagaac ttcgtgaagc ccaacgaaac caagacctac 2820
 ttctggaaaag tgcagcacca catggccccc accaaggacg agttcgactg caaggcctgg 2880
 gcctacttca gcgacgtgga tctggaaaag gacgtgcaact ctggactgat tggcccactc 2940
 ctggtctgcc acactaacac cctcaacccc gccacggcc gccaggtgac cgtgcaggaa 3000
 ttgccectgt tcttcacat cttcgacgag acaaagtctt ggtacttca cagagaatag 3060
 gaacggaact gcagagcccc ctgcaacatc cagatggaag atcctacctt caaagagaac 3120
 taccggttcc acgccatcaa cggctacatc atggacacc tgcttggcct ggtgatggcc 3180
 caggaccaga gaatccggtg gtatctgctg tccatgggca gcaacgagaa tatccacagc 3240
 atccacttca gcggccacgt gttcaccgtg cgggaagaaag aagagtacaa gatggccctg 3300
 tacaacctgt accccggcgt gttcgagaca gtggagatgc tgcccagcaa ggccggcatc 3360
 tggcgggtgg agtgtctgat cggcgagcac ctgcacgctg gcatgagcac cctgtttctg 3420
 gtgtacagca acaagtgcc aacccccactg ggcatggcct ctggccacat ccgggaette 3480
 cagatcaccg cctccggcca gtacggccag tgggccccca agctggccag actgcaactac 3540
 agcggcagca tcaacgcctg gtccacaaaa gagcccttca gctggatcaa ggtggacctg 3600
 ctggccccta tgatcatcca cggcattaag acccagggcg ccaggcagaa gttcagcagc 3660
 ctgtacatca gccagttcat catcatgtac agcctggacg gcaagaagtg gcagacctac 3720
 cggggcaaca gcaccggcac cctgatggtg ttcttcggca atgtggacag cagcggcatc 3780
 aagcacaaca tcttcaaccc ccccatcatt gcccggtaca tccggctgca ccccaccac 3840
 tacagcatta gatccacact gagaatggaa ctgatgggct gcgacctgaa ctctgcagc 3900
 atgcctctgg gcatggaaaag caaggccatc agcagcggcc agatcacagc cagcagctac 3960
 ttcaccaaca tgttcgccac ctggtccccc tccaaggcca ggctgcacct gcagggccgg 4020
 tccaacgcct ggcggcctca ggtcaacaac cccaaagaat ggctgcaggt ggactttcag 4080
 aaaaccatga aggtgaccgg cgtgaccacc cagggcgtga aaagcctgct gaccagcatg 4140
 tacgtgaaaag agtttctgat cagcagctct caggatggcc accagtggac cctgttcttt 4200
 cagaacggca aggtgaaaag gttecagggc aaccaggact ccttcacccc cgtgggtgac 4260
 tccttgacc cccccctgct gacccgctac ctgagaatcc acccccagtc ttgggtgcac 4320
 cagatcgccc tcaggatgga agtccctggga tgtgaggccc aggatctgta ctgatga 4377

<210> 6

<211> 4980

<212> DNA

5 <213> Secuencia Artificial

<220>

<223> Secuencia de nucleótidos de factor VIII modificada optimizada mediante codones

<400> 6

ES 2 452 890 T3

atgcagatcg agctgtccac ctgctttttt ctgtgcctgc tgcggttctg cttcagcgcc	60
accggcggt actacctggg cgccgtggag ctgtcctggg actacatgca gagcgacctg	120
ggcgagctgc cctgggacgc ccggttcccc cccagagtgc ccaagagctt ccccttcaac	180
accagcgtgg tgtacaagaa aaccctgttc gtggagttca ccgaccacct gttcaatatac	240
gccaagccca ggccccctg gatgggcctg ctgggccccca ccatccaggc cgaggtgtac	300
gacaccgtgg tgatcacctt gaagaacatg gccagccacc ccgtgagcct gcacgccgtg	360
ggcgtgagct actggaaggc cagcgagggc gccgagtacg acgaccagac cagccagcgg	420
gagaaagaag atgacaaggt gttccctggc ggcagccaca cctacgtgtg gcaggtgctg	480
aaagaaaacg gccccatggc ctccgacccc ctgtgcctga cctacagcta cctgagccac	540
gtggacctgg tgaaggacct gaacagcggc ctgatcggcg ctctgctcgt ctgccgggag	600
ggcagcctgg ccaaagagaa aaccagacc ctgcacaagt tcatcctgct gttcgcctg	660
ttcgacgagg gcaagagctg gcacagcgag acaaagaaca gcctgatgca ggaccgggac	720
gccgcctctg ccagagcctg gcccaagatg cacaccgtga acggctacgt gaacagaagc	780
ctgcccgccc tgattggctg ccaccggaag agcgtgtact ggcacgtgat cggcatgggc	840
accacacccg aggtgcacag catctttctg gaagggcaca cctttctggt ccggaaccac	900
cggcaggcca gcctggaaat cagccctatc accttctga ccgccagac actgetgatg	960
gacctgggcc agttcctgct gttttgccac atcagctctc accagcacga cggcatggaa	1020
gcctacgtga aggtggactc ttgccccgag gaacccagc tgcggatgaa gaacaacgag	1080
gaagccgagg actacgacga cgacctgacc gacagcgaga tggacgtggt gcggttcgac	1140
gacgacaaca gccccagctt catccagatc agaagcgtgg ccaagaagca cccaagacc	1200
tgggtgcact atatcgccgc cgaggaagag gactgggact acgccccctt ggtgctggcc	1260
cccgacgaca gaagctacaa gagccagtac ctgaacaatg gccccagcg gatcggccgg	1320
aagtacaaga aagtgcggtt catggcctac accgacgaga cattcaagac ccgggaggcc	1380
atccagcacg agagcggcat cctgggcccc ctgctgtacg gcgaagtggg cgacacactg	1440
ctgatcatct tcaagaacca ggctagccgg ccctacaaca tctaccccca cggcatcacc	1500
gacgtgcggc ccctgtacag caggcggctg cccaagggcg tgaagcacct gaaggacttc	1560
cccatcctgc ccggcgagat cttcaagtac aagtggaccg tgaccgtgga ggacggcccc	1620
accaagagcg accccagatg cctgacccgg tactacagca gcttcgtgaa catggaacgg	1680

ES 2 452 890 T3

gacctggcct	ccgggctgat	eggacctctg	ctgatctgct	acaagaaaag	cgtggaccag	1740
cggggcaacc	agatcatgag	cgacaagcgg	aacgtgatcc	tgttcagcgt	gttcgatgag	1800
aaccggtcct	ggtatctgac	cgagaacatc	cagcggtttc	tgcccaacce	tgccggcgtg	1860
cagctggaag	atcccagatt	ccaggccagc	aacatcatgc	actccatcaa	tggetacgtg	1920
ttcgactctc	tgcagctctc	cgtgtgtctg	cacgaggtgg	cctactggta	catcctgagc	1980
atcggcgccc	agaccgactt	cctgagcgtg	ttcttcagcg	gctacacctt	caagcacaag	2040
atgggtgtacg	aggacaccct	gacctgttc	cctttcagcg	gcgagacagt	gttcatgagc	2100
atggaaaacc	ccggcctgtg	gattctgggc	tgccacaaca	gcgacttccg	gaaccggggc	2160
atgaccgccc	tgctgaaggt	gtccagctgc	gacaagaaca	ccggcgacta	ctacgaggac	2220
agctacgagg	atatcagcgc	ctacctgctg	tccaagaaca	acgccatcga	accccgagc	2280
ttcagccaga	acccccctg	gctgacgcgt	ggcatgcggc	ccagaaacca	caccgtgctg	2340
ctgaaagtgt	gcaacaaggt	gcccataaac	aacagcacc	acagcgtgaa	caccagcacc	2400
ctgaacgcca	ccgacggcag	gcggtgctac	ctgaagaagg	tgacctgac	ccccctggc	2460
ggcgaggacg	ccaacatcac	ctctggcgag	cgcatccccg	aggacatcct	ggaagagctg	2520
cagctgggatg	gcgaagtgac	agccgccct	tccggcatcc	tgagagtggg	cgagcaggcc	2580
ctgcggcgga	gaaggcaggc	cgaggacaac	tccaccgact	tcgacatcta	cagcggcgac	2640
gagggcgggg	aagagcggag	cgagatccag	gaacgccccag	ccgagaaccg	gaccaccgcc	2700
aacgccaccg	gcatcaaccg	gaccaaagag	ctggaagagg	gcaacgagat	cctgctgacc	2760
ggcaaccccc	gcaacagcaa	gcggagcttc	accgtggagc	ccgtgaccag	cgaggaaatg	2820
gtgcagaact	acatcatgtc	tgaggaagag	aagctgaagc	cccaccatca	gcagcccggc	2880
aacgtggcgc	tgaacgagct	gtctgtgacg	cgtcaccagc	gggagatcac	ccggacaacc	2940
ctgcagtccg	accaggaaga	gatcgattac	gacgacacca	tcagcgtgga	gatgaagaaa	3000
gaggatttcg	atatctacga	cgaggacgag	aaccagagcc	ccagaagctt	ccagaagaaa	3060
acccggcact	acttcattgc	cgccgtggag	aggctgtggg	actacggcat	gagttctagc	3120
ccccacgtgc	tgcggaaccg	ggcccagagc	ggcagcgtgc	cccagttcaa	gaaagtgggtg	3180
ttccaggaat	tcacagacgg	cagettcacc	cagcctctgt	atagaggcga	gctgaacgag	3240
cacctggggc	tgctggggcc	ctacatcagg	gccgaagtgg	aggacaacat	catggtgacc	3300
ttccggaatc	aggccagcag	accetactcc	ttctacagca	gcctgatcag	ctacgaagag	3360
gaccagcggc	agggcgccga	accccggaag	aacttcgtga	agcccaacga	aaccaagacc	3420
tacttctgga	aagtgcagca	ccacatggcc	cccaccaagg	acgagttcga	ctgcaaggcc	3480
tgggcctact	tcagcgacgt	ggatctggaa	aaggacgtgc	actctggact	gattggccca	3540
ctcctggtct	gccacactaa	cacctcaac	cccggccacg	gccgccaggt	gaccgtgcag	3600

ES 2 452 890 T3

gaattcgccc tgttcttcac catcttcgac gagacaaagt cctggtactt caccgagaat 3660
atggaacgga actgcagagc cccctgcaac atccagatgg aagatcctac cttcaaagag 3720
aactaccggt tccacgccat caacggctac atcatggaca ccctgcctgg cctggtgatg 3780
gcccaggacc agagaatccg gtggtatctg ctgtccatgg gcagcaacga gaatatccac 3840
agcatccact tcagcggcca cgtgttcacc gtgcggaaga aagaagagta caagatggcc 3900
ctgtacaacc tgtaccccg cgtgttcgag acagtggaga tgctgcccag caaggccggc 3960
atctggcggg tggagtgtct gatcggcgag cacctgcacg ctggcatgag caccctgttt 4020
ctggtgtaca gcaacaagt cagacccca ctgggatgg cctctggcca catccgggac 4080
ttccagatca ccgcctccgg ccagtacggc cagtgggcc ccaagctggc cagactgcac 4140
tacagcggca gcatcaacgc ctggtccacc aaagagccct tcagctggat caaggtggac 4200
ctgctggccc ctatgatcat ccacggcatt aagaccagc gcgccaggca gaagttcagc 4260
agcctgtaca tcagccagtt catcatcatg tacagcctgg acggcaagaa gtggcagacc 4320
taccggggca acagcaccgg caccctgatg gtgttcttcg gcaatgtgga cagcagcggc 4380
atcaagcaca acatcttcaa ccccccatc attgcccggg acatccggct gcaccccacc 4440
cactacagca ttagatccac actgagaatg gaactgatgg gctgcgacct gaactcctgc 4500
agcatgcctc tgggcatgga aagcaaggcc atcagcgacg cccagatcac agccagcagc 4560
tacttcacca acatgttcgc cacctggtcc ccctccaagg ccaggctgca cctgcagggc 4620
cggccaacg cctggcggcc tcaggtaac aaccccaaag aatggctgca ggtggacttt 4680
cagaaaacca tgaaggtgac cggcgtgacc acccagggcg tgaaaagcct gctgaccagc 4740
atgtacgtga aagagtttct gatcagcagc tctcaggatg gccaccagtg gaccctgttc 4800
tttcagaacg gcaaggtgaa agtgttccag ggcaaccagg actccttcac ccccggtg 4860
aactccctgg accccccct gctgaccgc tacctgagaa tccacccca gtcttgggtg 4920
caccagatcg ccctcaggat ggaagtcctg ggatgtgagg cccaggatct gtactgatga 4980

<210> 7

<211> 5336

<212> DNA

5 <213> Secuencia Artificial

<220>

<223> Secuencia de nucleótidos de factor VIII modificada optimizada mediante codones

<400> 7

tgtttgctgc ttgcaatggt tgcccatttt agggtggaca caggacgctg tggtttctga 60
gccagggggc gactcagatc ccagccagtg gacttagccc ctgtttgctc ctccgataac 120
tggggtgacc ttggttaata ttcaccagca gcctccccg tgcccctct ggatccactg 180
cttaaatagc gacgaggaca gggccctgct tcctcagctt caggcaccac cactgacctg 240
ggacagtgaa tcgcgggccg caccatgcag attgagctga gcacctgctt cttcctgtgc 300

ES 2 452 890 T3

ctgctgaggt tctgcttctc tgccaccagg agatactacc tgggggctgt ggagctgagc	360
tgggactaca tgcagtctga cctgggggag ctgcctgtgg atgccagggt cccccccaga	420
gtgcccaaga gcttcccctt caacacctct gtgggtgata agaagaccct gtttgtggag	480
ttcactgacc acctgttcaa cattgccaaag cccaggcccc cctggatggg cctgctgggc	540
cccaccatcc aggctgaggt gtatgacact gtgggtgatca ccctgaagaa catggccagc	600
caccctgtga gcctgcatgc tgtgggggtg agctactgga aggcctctga gggggctgag	660
tatgatgacc agaccagcca gagggagaag gaggatgaca aggtgttccc tgggggcagc	720
cacacctatg tgtggcaggt gctgaaggag aatggcccc tggcctctga cccctgtgc	780
ctgacctaca gctacctgag ccatgtggac ctgggtgaagg acctgaactc tggcctgatt	840
ggggccctgc tgggtgtcag ggagggcagc ctggccaagg agaagacca gacctgcac	900
aagttcatcc tctgttttc tgtgtttgat gagggcaaga gctggcactc tgaaccaag	960
aacagcctga tgcaggacag ggatgctgcc tctgccaggg cctggcccaa gatgcacact	1020
gtgaatggct atgtgaacag gacctgcct ggcctgattg gctgccacag gaagtctgtg	1080
tactggcatg tgattggcat gggcaccacc cctgaggtgc acagcatctt cctggagggc	1140
cacaccttc tggtcaggaa ccacaggcag gccagcctgg agatcagccc catcacctc	1200
ctgactgccc agaccctgct gatggacctg ggccagtcc tctgtttctg ccacatcagc	1260
agccaccagc atgatggcat ggaggcctat gtgaagggtg acagctgccc tgaggagccc	1320
cagctgagga tgaagaacaa tgaggaggct gaggactatg atgatgacct gactgactct	1380
gagatggatg tggtgagggt tgatgatgac aacagcccca gttcatcca gatcaggctc	1440
gtggccaaga agcaccceaa gacctgggtg cactacattg ctgctgagga ggaggactgg	1500
gactatgccc ccctgggtgct ggcccctgat gacaggagct acaagagcca gtacctgaac	1560
aatggcccc agaggattgg caggaagtac aagaaggca ggttcatggc ctacactgat	1620
gaaaccttca agaccagga ggccatccag catgagtctg gcatcctggg ccccctgctg	1680
tatggggagg tgggggacac cctgctgate atcttcaaga accaggccag caggccctac	1740
aacatctacc cccatggcat cactgatgtg aggccctgt acagcaggag gctgcccaag	1800
gggggtgaagc acctgaagga ctccccatc ctgcctgggg agatcttcaa gtacaagtgg	1860
actgtgactg tggaggatgg ccccaccaag tctgacccca ggtgcctgac cagatactac	1920
agcagctttg tgaacatgga gagggacctg gcctctggcc tgattggccc cctgctgate	1980
tgctacaagg agtctgtgga ccagaggggc aaccagatca tgtctgacaa gaggaatgtg	2040
atcctgttct ctgtgtttga tgagaacagg agctggtacc tgactgagaa catccagagg	2100
ttcctgcccc acctgctgg ggtgcagctg gaggaccctg agttccaggc cagcaacatc	2160
atgcacagca tcaatggeta tgtgtttgac agcctgcagc tgtctgtgtg cctgcatgag	2220

ES 2 452 890 T3

gtggcctact ggtacatcct gagcattggg gccagactg acttcctgtc tgtgttcttc	2280
tctggctaca ccttcaagca caagatggtg tatgaggaca ccctgaccct gttccccttc	2340
tctggggaga ctgtgttcat gagcatggag aaccctggcc tgtggattct gggctgccac	2400
aactctgact tcaggaacag gggcatgact gccctgctga aagtctccag ctgtgacaag	2460
aacctgggg actactatga ggacagctat gaggacatct ctgcctacct gctgagcaag	2520
aacaatgcca ttgagcccag gagcttcagc cagaacagca ggcaccccag caccaggcag	2580
aagcagtta atgccaccac catccctgag aatgacatag agaagacaga cccatggttt	2640
gccaccgga ccccatgcc caagatccag aatgtgagca gctctgacct gctgatgctg	2700
ctgaggcaga gcccacccc ccatggcctg agcctgtctg acctgcagga ggccaagtat	2760
gaaaccttct ctgatgacct cagccctggg gccattgaca gcaacaacag cctgtctgag	2820
atgaccctact tcaggcccca gctgcaccac tctggggaca tgggtgtcac ccctgagtct	2880
ggcctgcagc tgaggetgaa tgagaagctg ggcaccactg ctgccactga gctgaagaag	2940
ctggacttca aagtctccag caccagcaac aacctgatca gcaccatccc ctctgacaac	3000
ctggtgctg gcactgacaa caccagcagc ctggggcccc ccagcatgcc tgtgcactat	3060
gacagccagc tggacaccac cctgtttggc aagaagagca gccccctgac tgagtctggg	3120
ggccccctga gcctgtctga ggagaacaat gacagcaagc tgetggagtc tggcctgatg	3180
aacagccagg agagcagctg gggcaagaat gtgagcagca gggagatcac caggaccacc	3240
ctgcagtctg accaggagga gattgactat gatgacacca tctctgtgga gatgaagaag	3300
gaggactttg acatctacga cgagagcag aaccagagcc ccaggagctt ccagaagaag	3360
accaggcact acttcattgc tgetgtggag aggctgtggg actatggcat gagcagcagc	3420
ccccatgtc tgaggaacag ggcccagtct ggctctgtgc ccagttcaa gaaggtggtg	3480
ttccaggagt tcactgatgg cagcttcacc cagcccctgt acagagggga gctgaatgag	3540
cacctgggcc tgctgggccc ctacatcagg gctgaggtgg aggacaacat catggtgacc	3600
ttcaggaacc aggccagcag gccctacagc ttctacagca gcctgatcag ctatgaggag	3660
gaccagagge agggggetga gcccaggaag aactttgtga agcccaatga aaccaagacc	3720
tacttctgga aggtgcagca ccacatggcc cccaccaagg atgagtttga ctgcaaggcc	3780
tgggcctact tctctgatgt ggacctggag aaggatgtgc actctggcct gattggcccc	3840
ctgctggtgt gccacaccaa caccctgaac cctgcccctg gcaggcaggt gactgtgcag	3900
gagtttgccc tgttcttcac catctttgat gaaaccaaga gctggtactt cactgagaac	3960
atggagagga actgcagggc cccctgcaac atccagatgg aggaccccac cttcaaggag	4020
aactacaggt tccatgccat caatggctac atcatggaca ccctgcctgg cctggtgatg	4080
gcccaggacc agaggatcag gtggtacctg ctgagcatgg gcagcaatga gaacatccac	4140
agcatccaet tctctggcca tgtgttcaet gtgaggaaga aggaggagta caagatggcc	4200

ES 2 452 890 T3

```

ctgtacaacc tgtaccctgg ggtgtttgag actgtggaga tgctgcccag caaggctggc      4260
atctggaggg tggagtgcct gattggggag cacctgcatg ctggcatgag caccctgttc      4320
ctgggtgtaca gcaacaagtg ccagaccccc ctgggcatgg cctctggcca catcagggac      4380
ttccagatca ctgcctctgg ccagtatggc cagtgggccc ccaagctggc caggctgcac      4440
tactctggca gcatcaatgc ctggagcacc aaggagccct tcagctggat caaggtggac      4500
ctgctggccc ccatgatcat ccatggcatc aagaccagg gggccaggca gaagttcagc      4560
agcctgtaca tcagccagtt catcatcatg tacagcctgg atggcaagaa gtggcagacc      4620
tacaggggca acagcactgg caccctgatg gtgttctttg gcaatgtgga cagctctggc      4680
atcaagcaca acatcttcaa ccccccatc attgccagat acatcagget gcaccccacc      4740
cactacagca tcaggagcac cctgaggatg gagctgatgg gctgtgacct gaacagctgc      4800
agcatgcccc tgggcatgga gagcaaggcc atctctgatg cccagatcac tgccagcagc      4860
tacttcacca acatgtttgc cacctggagc cccagcaagg ccaggctgca cctgcagggc      4920
aggagcaatg cctggaggcc ccaggccaac aacccaagg agtggctgca ggtggacttc      4980
cagaagacca tgaaggtgac tggggtgacc acccaggggg tgaagagcct gctgaccagc      5040
atgtatgtga aggagtccct gatcagcagc agccaggatg gccaccagtg gaccctgttc      5100
ttccagaatg gcaaggtgaa ggtgttccag ggcaaccagg acagcttcac cctgtggtg      5160
aacagcctgg accccccct gctgaccaga tacctgagga ttcacccca gagctgggtg      5220
caccagattg ccctgaggat ggaggtgctg ggctgtgagg cccaggacct gtactgatcg      5280
cgaataaaaag atctttattt tcattagatc tgtgtgttgg ttttttgtgt gatgca      5336
<210> 8
<211> 42
<212> DNA
5 <213> Secuencia Artificial
<220>
<223> Constructo sintético
<400> 8
agcttcagcc agaaccccc cgtgctgacg cgtcaccagc gg      42
10 <210> 9
<211> 20
<212> DNA
<213> Secuencia Artificial
<220>
15 <223> Cebador
<400> 9
tgtgtgcccg tctgtgtgt      20
<210> 10
<211> 20
20 <212> DNA
<213> Secuencia Artificial
<220>
<223> Cebador
<400> 10
25 gagtctctgcg tcgagagagc      20
<210> 11
<211> 20
<212> DNA
<213> Secuencia Artificial
30 <220>
<223> Cebador
<400> 11
cgcccgaaca gggacttgaa      20
<210> 12

```


<211> 21
 <212> DNA
 <213> Secuencia Artificial
 <220>
 5 <223> Cebador
 <400> 12
 aaaacgagca gtagcctgag g 21
 <210> 13
 <211> 20
 10 <212> DNA
 <213> Secuencia Artificial
 <220>
 <223> Cebador
 <400> 13
 15 ttcagtcatg ctgctagcgc 20
 <210> 14
 <211> 22
 <212> DNA
 <213> Secuencia Artificial
 <220>
 20 <223> Cebador
 <400> 14
 tgcacggaat ctgctctcag tc 22
 <210> 15
 25 <211> 25
 <212> DNA
 <213> Secuencia Artificial
 <220>
 <223> Cebador
 <400> 15
 30 tcaccacaaa gttgccatc tacga 25
 <210> 16
 <211> 25
 <212> DNA
 35 <213> Secuencia Artificial
 <220>
 <223> Cebador
 <400> 16
 40 cagcgggaacc gctcattgcc aatgg 25
 <210> 17
 <211> 26
 <212> DNA
 <213> Secuencia Artificial
 <220>
 45 <223> Cebador
 <400> 17
 atgccctccc ccatgccatc ctgcgt 26
 <210> 18
 <211> 14
 50 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Constructo sintético
 <400> 18
Ser Phe Ser Gln Asn Pro Pro Val Leu Thr Arg His Gln Arg
 55 **1 5 10**
 <210> 19
 <211> 11
 <212> PRT
 <213> Secuencia Artificial
 <220>
 60 <223> Constructo sintético
 <400> 19
Ser Phe Ser Gln Asn Pro Pro Val Leu Thr Arg
1 5 10
 <210> 20

<211> 5
 <212> PRT
 <213> Secuencia Artificial
 <220>
 5 <223> Constructo sintético
 <400> 20
Thr Arg His Gln Arg
1 5
 <210> 21
 <211> 1670
 10 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Modified Factor VIII protein
 <400> 21
Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
1 5 10 15

Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser
20 25 30

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
85 90 95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser
100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
165 170 175

 15 **Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile**
180 185 190

ES 2 452 890 T3

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

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
 210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
 225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
 245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
 260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
 275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
 290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
 305 310 315 320

Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
 325 330 335

Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
 340 345 350

Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365

Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380

Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400

Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415

Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
 420 425 430

Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
 435 440 445

Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu

ES 2 452 890 T3

450	455	460														
Ser 465	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu	480
				470					475							
Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro	495
				485					490							
His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys	510
			500					505								
Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe	525
			515					520								
Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp	540
				530					535							
Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg	560
				545					550					555		
Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu	575
				565					570							
Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val	590
			580					585								
Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu	605
			595					600					605			
Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp	620
				610					615							
Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val	640
				625					630					635		
Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp	655
			645					650								
Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe	670
			660					665								
Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr	685
				675					680					685		
Leu	Phe	Pro	Phe	Ser	Gly	Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro	700
				690					695							
Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly	720
				705					710					715		

ES 2 452 890 T3

Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
725 730 735

Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys
740 745 750

Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro
755 760 765

Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp
770 775 780

Ile Glu Lys Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys
785 790 795 800

Ile Gln Asn Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser
805 810 815

Pro Thr Pro His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr
820 825 830

Glu Thr Phe Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn
835 840 845

Ser Leu Ser Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly
850 855 860

Asp Met Val Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu
865 870 875 880

Lys Leu Gly Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys
885 890 895

Val Ser Ser Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn
900 905 910

Leu Ala Ala Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met
915 920 925

Pro Val His Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys
930 935 940

Ser Ser Pro Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu
945 950 955 960

Asn Asn Asp Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu
965 970 975

ES 2 452 890 T3

Ser Ser Trp Gly Lys Asn Val Ser Ser Arg Glu Ile Thr Arg Thr Thr
 980 985 990

Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val
 995 1000 1005

Glu Met Lys Lys Glu Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn
 1010 1015 1020

Gln Ser Pro Arg Ser Phe Gln Lys Lys Thr Arg His Tyr Phe Ile
 1025 1030 1035

Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser Ser Pro
 1040 1045 1050

His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser Val Pro Gln Phe
 1055 1060 1065

Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly Ser Phe Thr Gln
 1070 1075 1080

Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly Leu Leu Gly
 1085 1090 1095

Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe
 1100 1105 1110

Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile
 1115 1120 1125

Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn
 1130 1135 1140

Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln
 1145 1150 1155

His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp
 1160 1165 1170

Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly
 1175 1180 1185

Leu Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro
 1190 1195 1200

Ala His Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe
 1205 1210 1215

ES 2 452 890 T3

Thr Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met
 1220 1225 1230

Glu Arg Asn Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro
 1235 1240 1245

Thr Phe Lys Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile
 1250 1255 1260

Met Asp Thr Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile
 1265 1270 1275

Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser
 1280 1285 1290

Ile His Phe Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu
 1295 1300 1305

Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr
 1310 1315 1320

Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys
 1325 1330 1335

Leu Ile Gly Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu
 1340 1345 1350

Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly
 1355 1360 1365

His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln
 1370 1375 1380

Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn
 1385 1390 1395

Ala Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu
 1400 1405 1410

Leu Ala Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg
 1415 1420 1425

Gln Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr
 1430 1435 1440

Ser Leu Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr
 1445 1450 1455

Gly Thr Leu Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile

ES 2 452 890 T3

1460		1465		1470
Lys His Asn Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg 1475		1480		1485
Leu His Pro Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu 1490		1495		1500
Leu Met Gly Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met 1505		1510		1515
Glu Ser Lys Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr 1520		1525		1530
Phe Thr Asn Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu 1535		1540		1545
His Leu Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn 1550		1555		1560
Pro Lys Glu Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val 1565		1570		1575
Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met 1580		1585		1590
Tyr Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln 1595		1600		1605
Trp Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly 1610		1615		1620
Asn Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro 1625		1630		1635
Leu Leu Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His 1640		1645		1650
Gln Ile Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp 1655		1660		1665
Leu Tyr 1670				

REIVINDICACIONES

1. Una molécula de ácidos nucleicos aislada, que comprende una secuencia de nucleótidos que tiene una identidad de al menos 90% respecto a la secuencia de nucleótidos de la SEC ID N° 1 y que codifica factor VIII funcional.
- 5 2. La molécula de ácidos nucleicos de la reivindicación 1, que tiene una identidad de al menos 95% respecto a la secuencia de nucleótidos de la SEC ID N° 1.
- 10 3. La molécula de ácidos nucleicos de la reivindicación 1 o la reivindicación 2, que codifica una proteína que comprende la secuencia de la SEC ID N° 2 o la SEC ID N° 21 que tiene entre 0 y 10 cambios de aminoácidos para la misma.
- 15 4. La molécula de ácidos nucleicos de la reivindicación 3, que codifica una proteína que comprende la secuencia de la SEC ID N° 2 o la SEC ID N° 21.
- 20 5. La molécula de ácidos nucleicos de la reivindicación 1, que comprende la secuencia de nucleótidos seleccionada entre la secuencia de la SEC ID N° 1 y la SEC ID N° 7.
6. La molécula de ácidos nucleicos de una cualquiera de las reivindicaciones 1 a 5, que comprende adicionalmente una secuencia de nucleótidos que tiene una identidad de al menos 85% respecto a la secuencia de nucleótidos de la SEC ID N° 3.
7. Un vector, que comprende una molécula de ácidos nucleicos según una cualquiera de las reivindicaciones 1 a 6.
- 25 8. Una célula hospedante que comprende la molécula de ácidos nucleicos de una cualquiera de las reivindicaciones 1 a 6 o el vector de la reivindicación 7.
- 30 9. Un animal transgénico que comprenden células que comprenden la molécula de ácidos nucleicos de una cualquiera de las reivindicaciones 1 a 6 o el vector de la reivindicación 7.
- 35 10. La molécula de ácidos nucleicos de una cualquiera de las reivindicaciones 1 a 6, o un vector según la reivindicación 7, para ser usada en terapia.
11. Una molécula de ácidos nucleicos de una cualquiera de las reivindicaciones 1 a 6, o un vector según la reivindicación 7, para ser usados en el tratamiento de hemofilia.
12. Un método para la preparación de un vector de suministro de genes parvovirales, comprendiendo el método las etapas de
- 40 (a) proporcionar una célula de insecto que comprende uno o más constructos de ácidos nucleicos que comprenden:
- (i) una molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 6 que está flanqueada por al menos una secuencia de nucleótidos repetida, terminal, invertida y parvoviral;
- 45 (ii) un primer cassette de expresión que comprende una secuencia de nucleótidos que codifica una o más proteínas Rep parvovirales que está funcionalmente conectado a un promotor que es capaz de conducir la expresión de la(s) proteína(s) Rep en la célula de insecto;
- 50 (iii) un segundo cassette de expresión que comprende una secuencia de nucleótidos que codifica una o más proteínas de cápsides parvovirales que está funcionalmente conectado a un promotor que es capaz de conducir la expresión de la(s) proteína(s) de la cápside en la célula de insecto;
- (b) cultivar la célula de insecto definida en el apartado (a) bajo condiciones que conduzcan a la expresión de las proteínas Rep y de la cápside; y, opcionalmente,
- 55 (c) recuperar el vector de suministro de genes parvovirales.

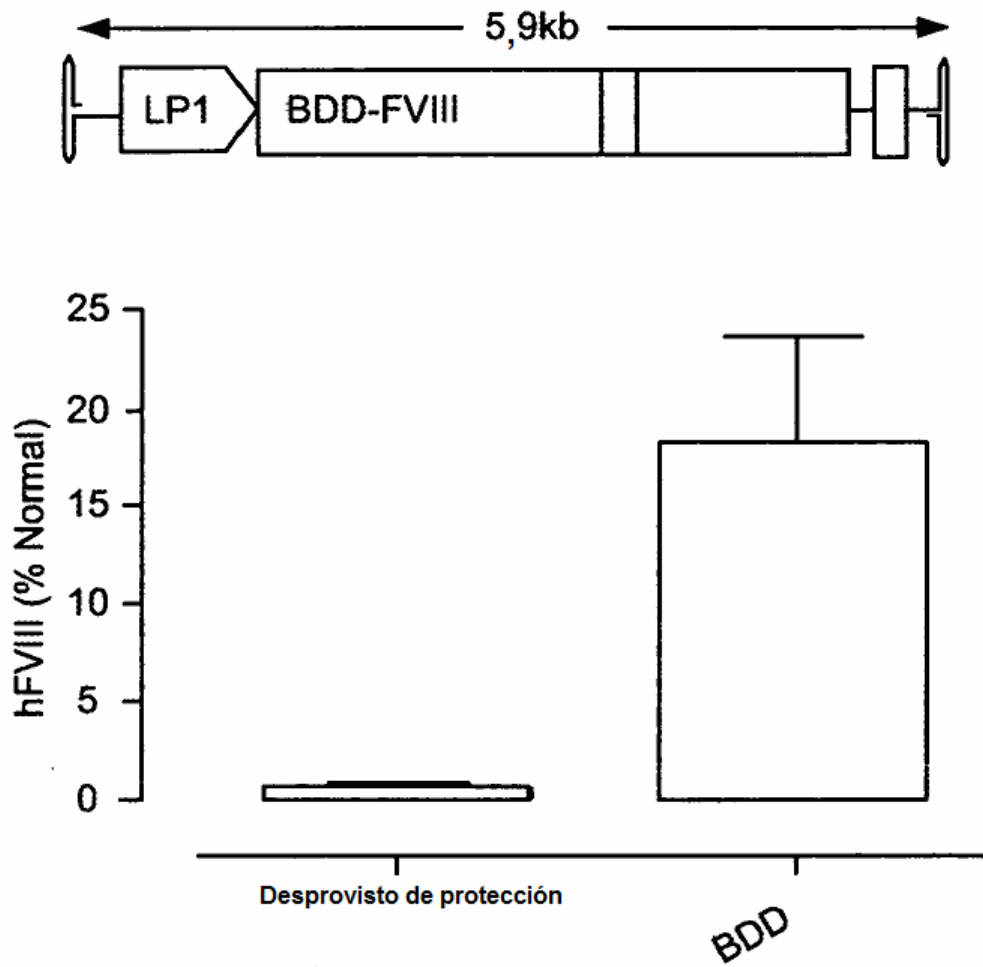


Figura 1

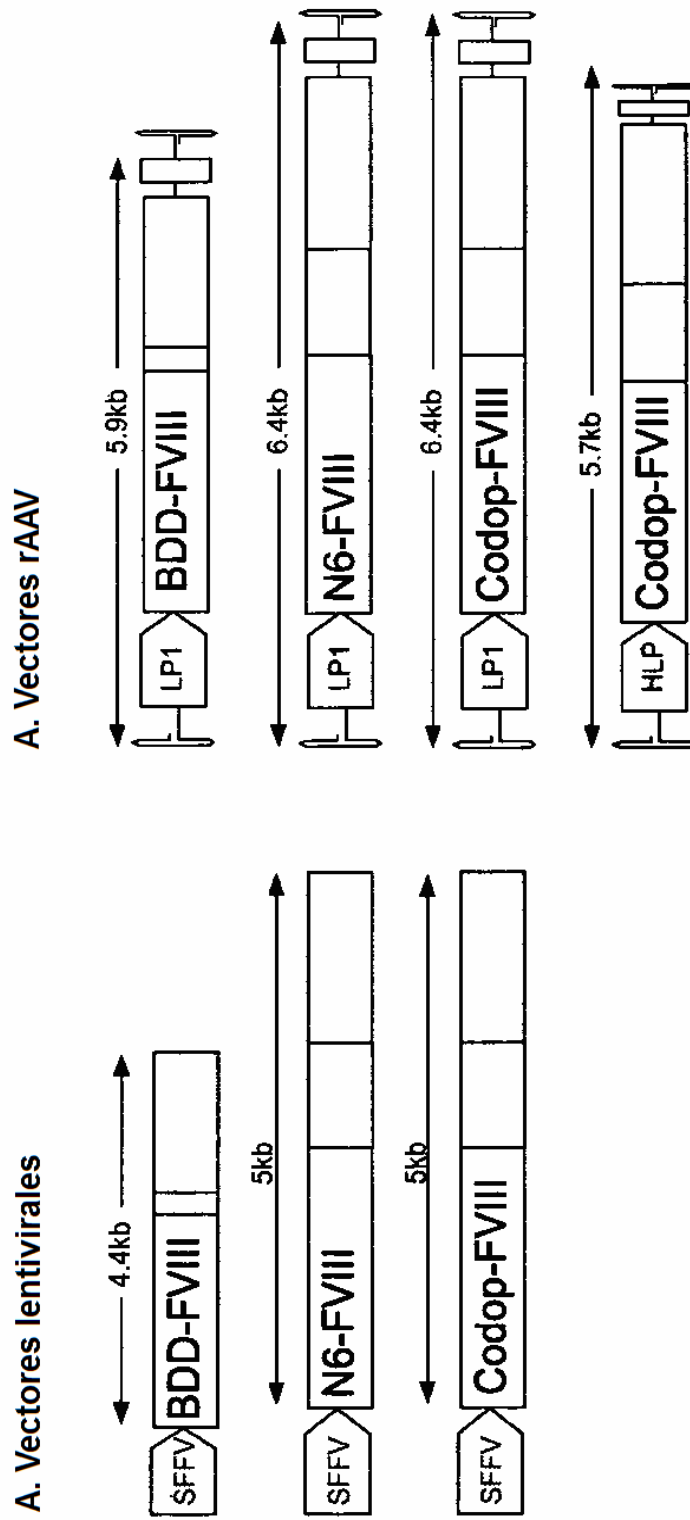


Figura 2(a)

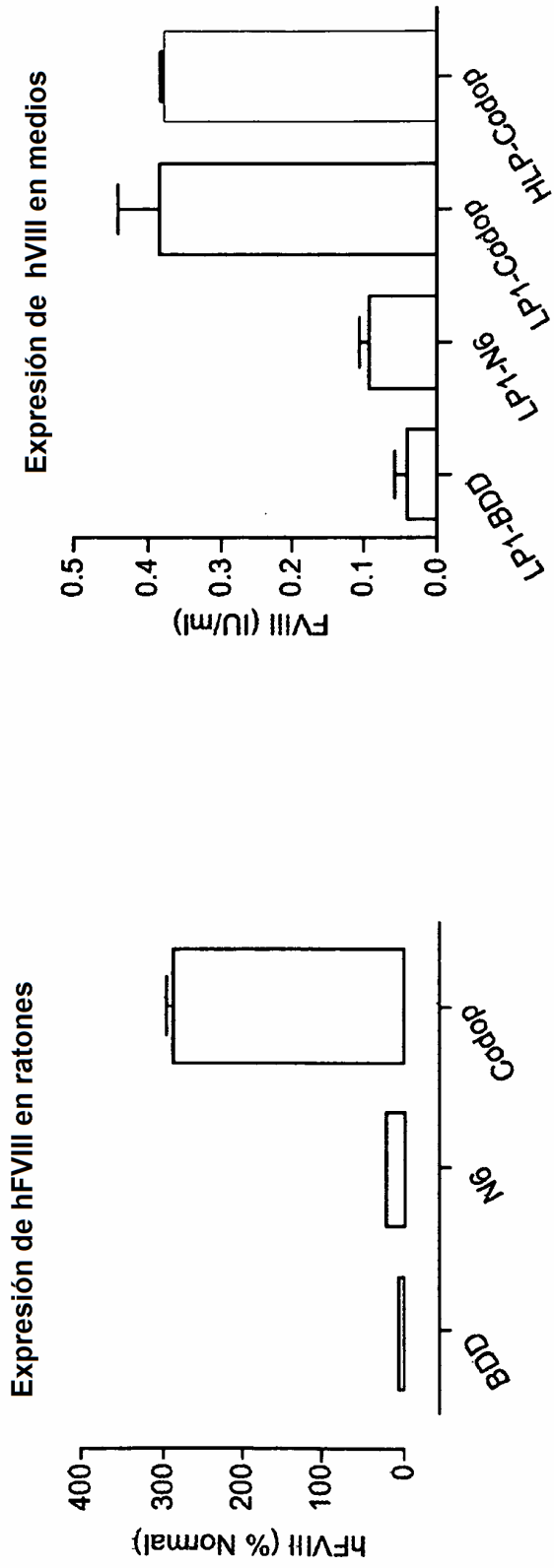
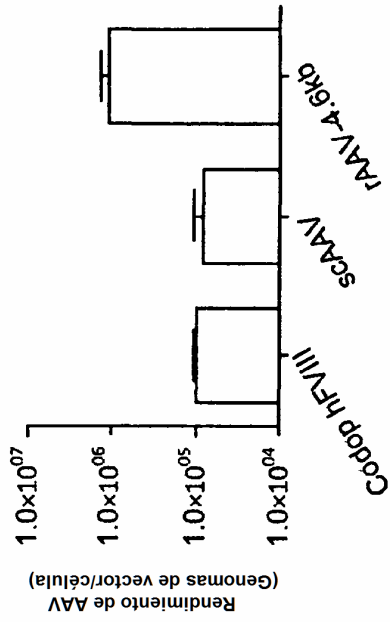
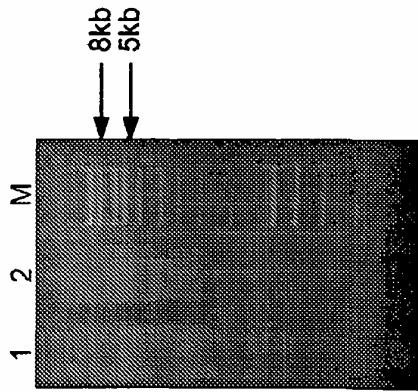


Figura 2(b)

A. Rendimiento de vector AAV a partir de células 293T



B. gel de agarosa nativo



C. Gel de agarosa - transferencia Southern

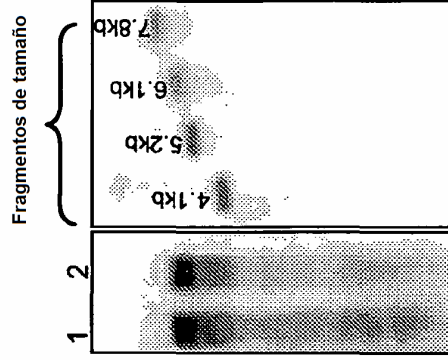


Figura 3

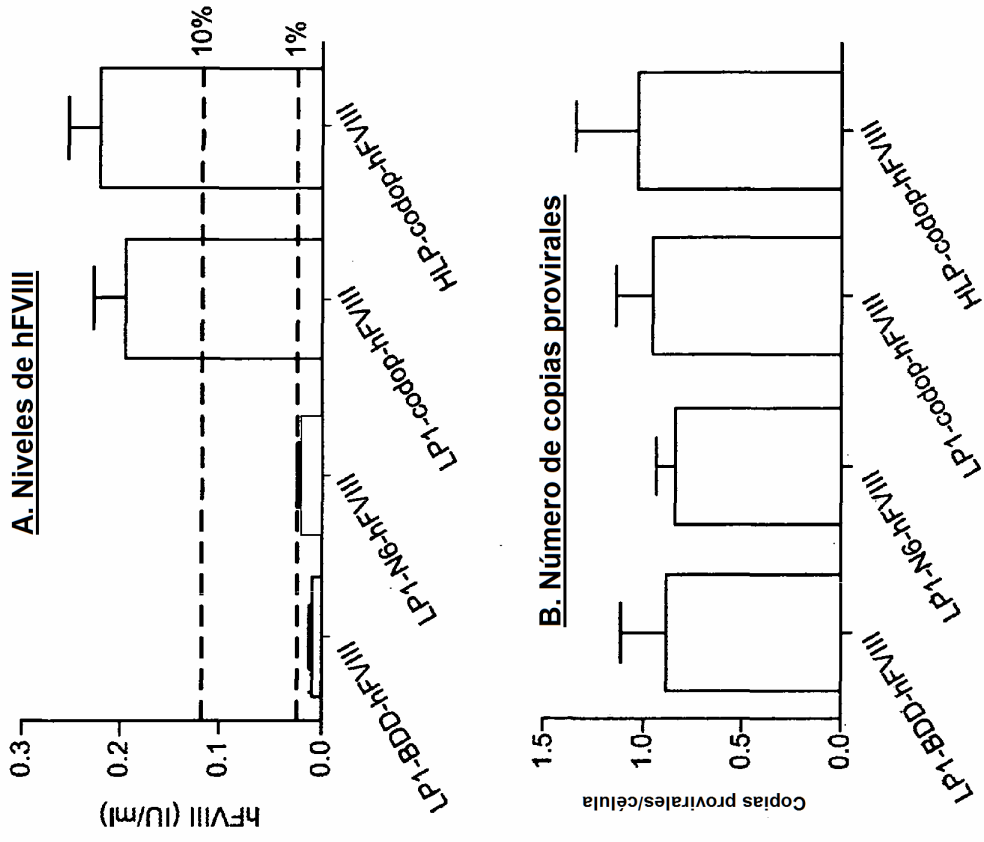


Figura 4

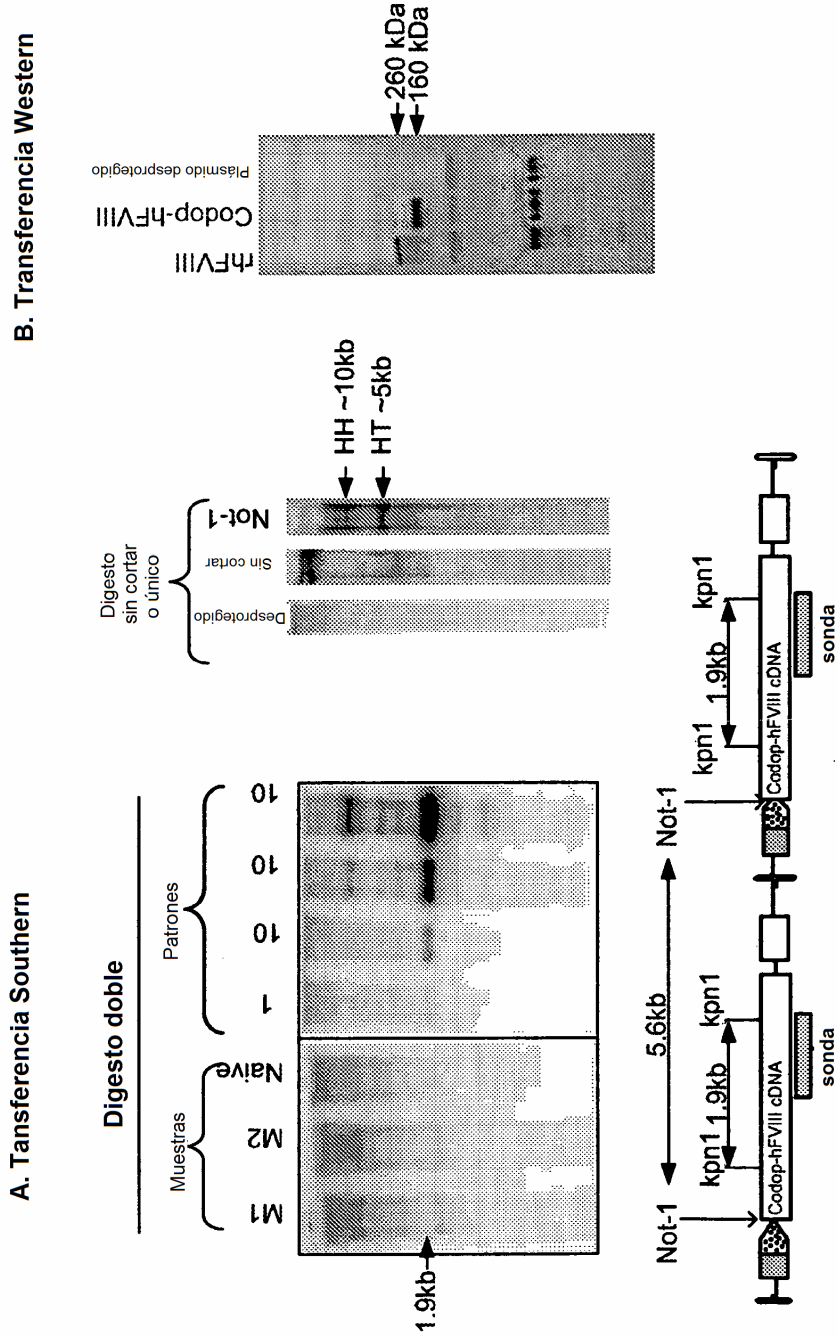


Figura 5

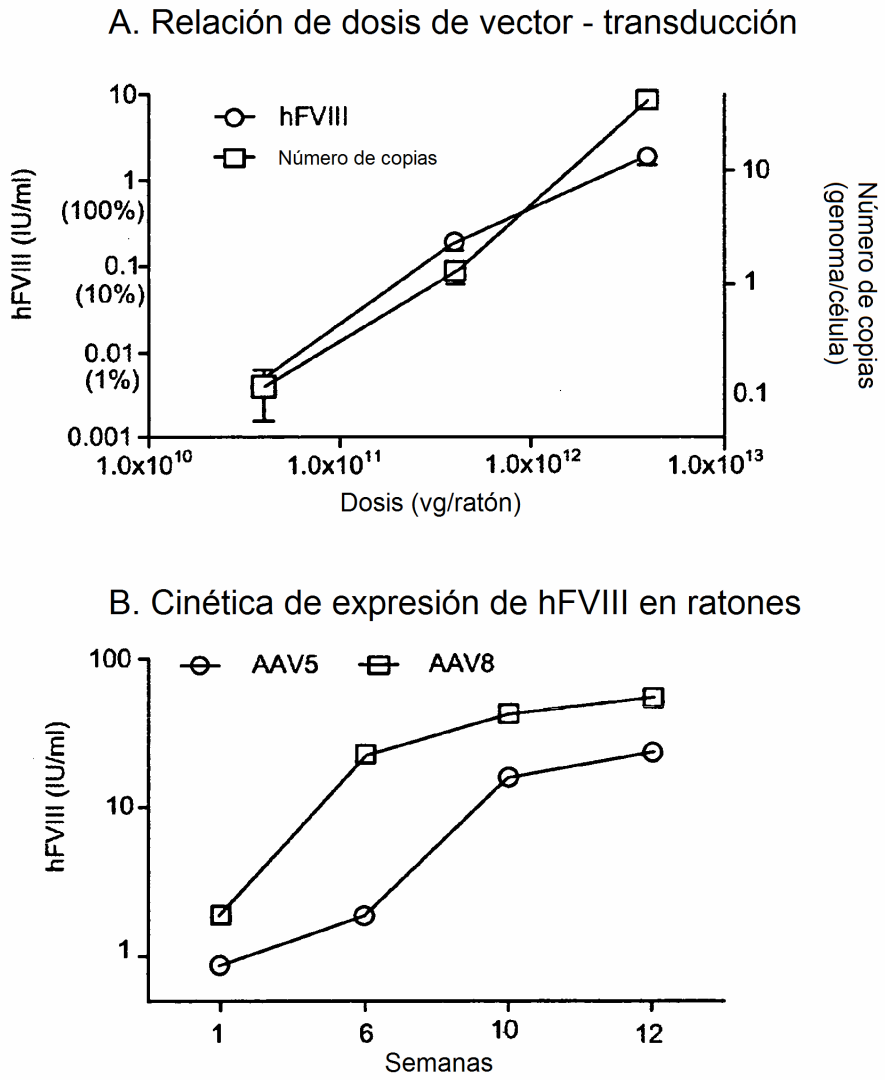


Figura 6

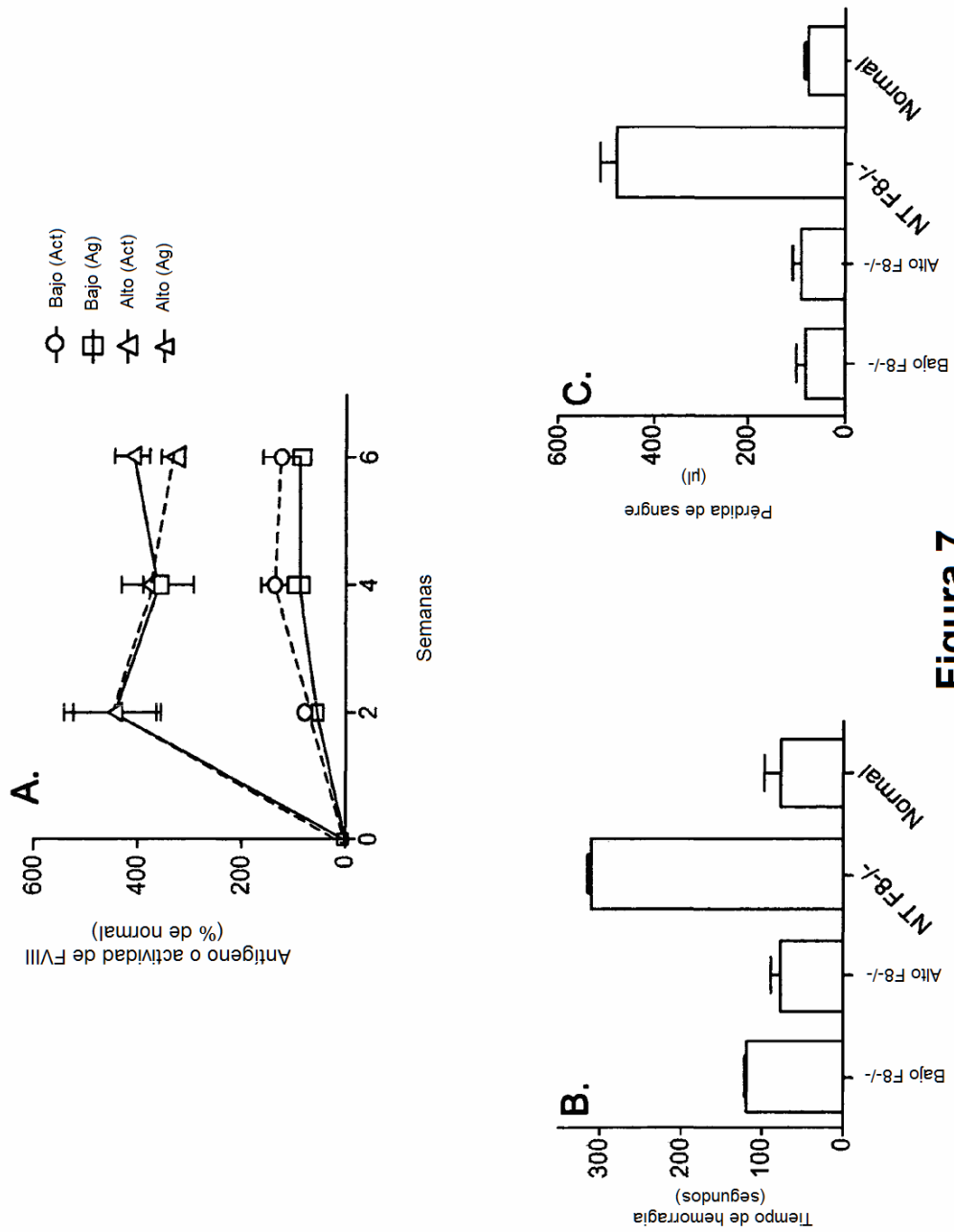
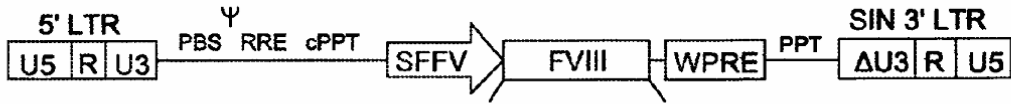


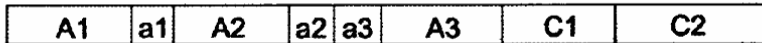
Figura 7

A

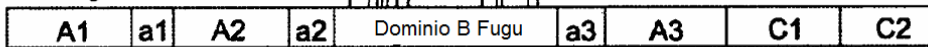


Delección de dominio B de hFVIII cDNA de longitud completa
(Bengbank: A05328) entre los aa 740-1649
sitio MluI (ACG'CGT) tratado por ingeniería en posición 740

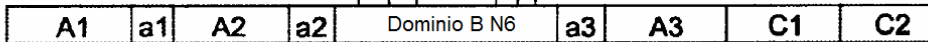
BDD FVIII



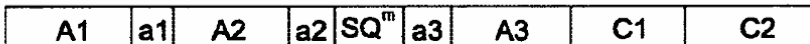
FVIII Fugu B



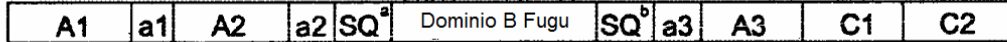
FVIII N6



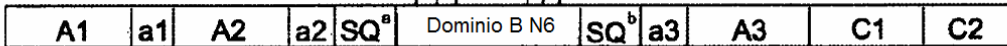
SQ FVIII



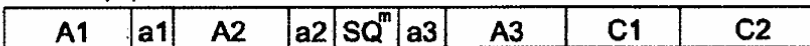
SQ FVIII Fugu B



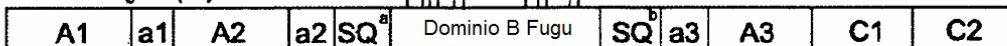
SQ FVIII N6



SQ FVIII (co)



SQ FVIII Fugu B (co)



SQ FVIII N6 (co)

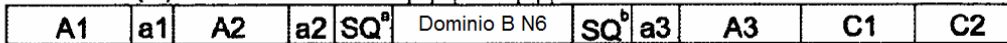


Figura 8(a)

B

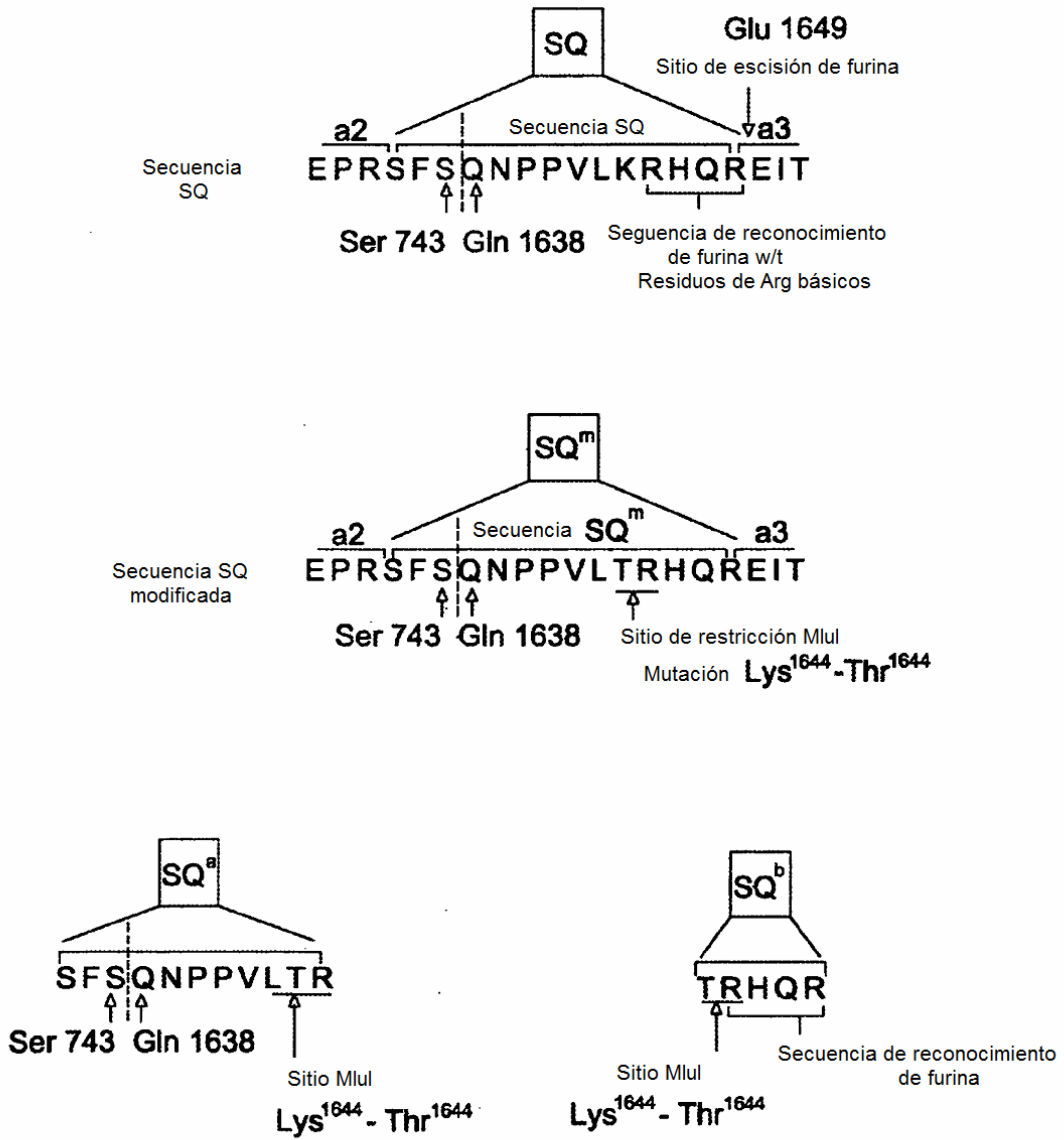


Figura 8(b)

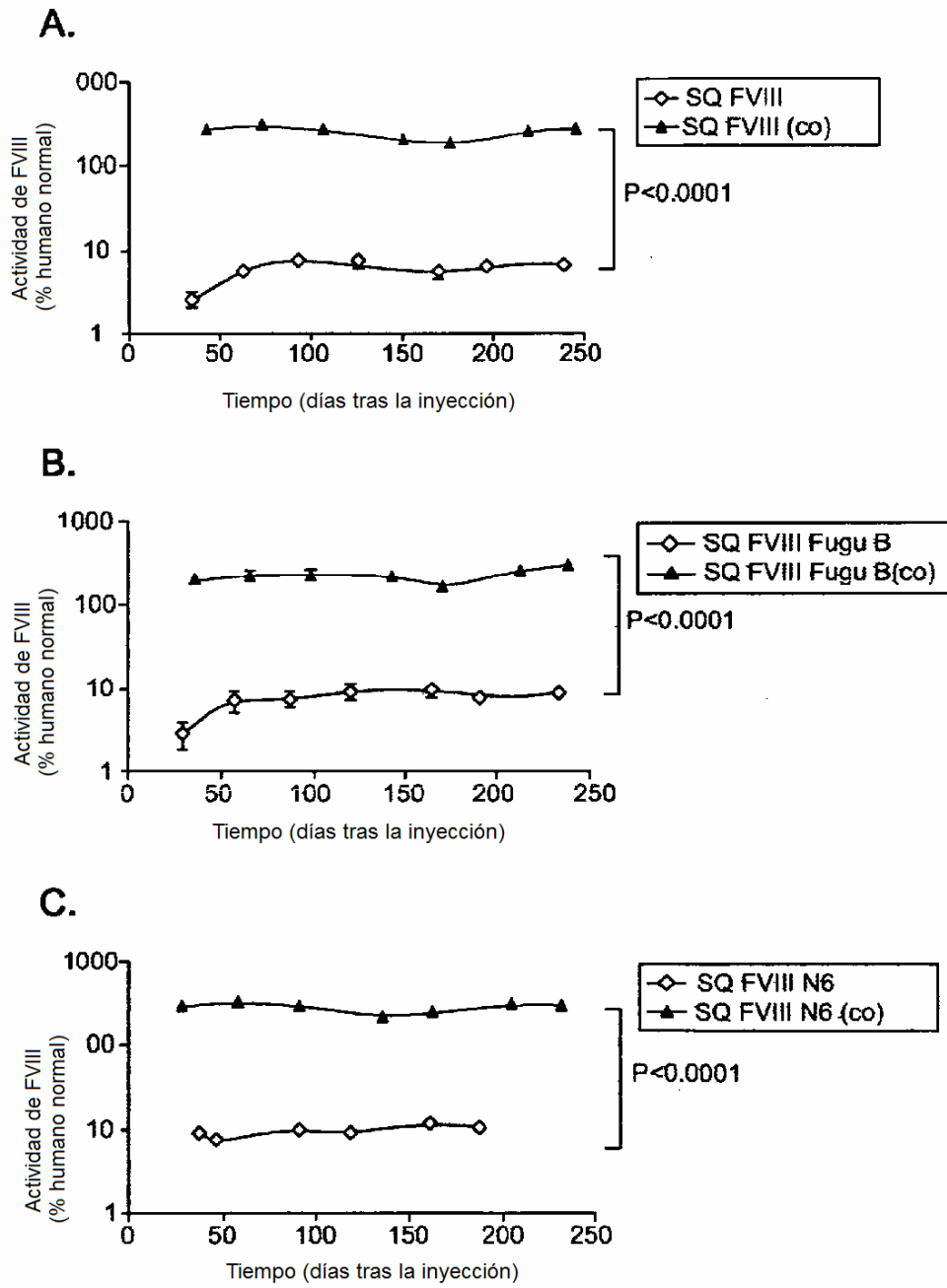


Figura 10

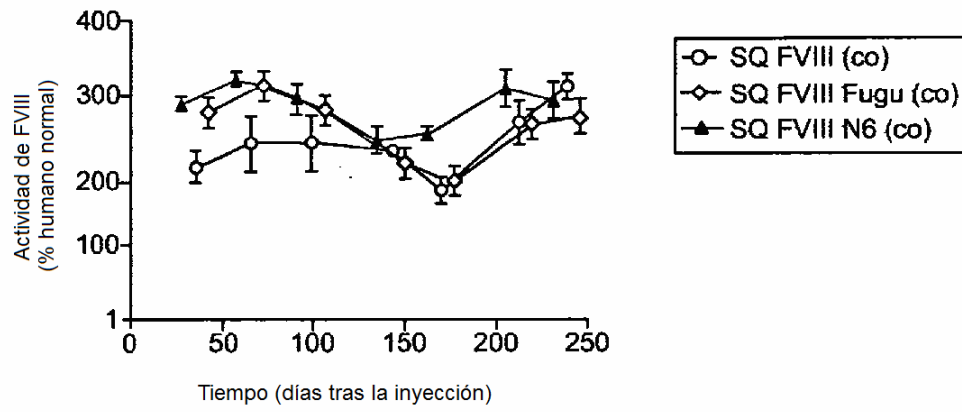


Figura 11

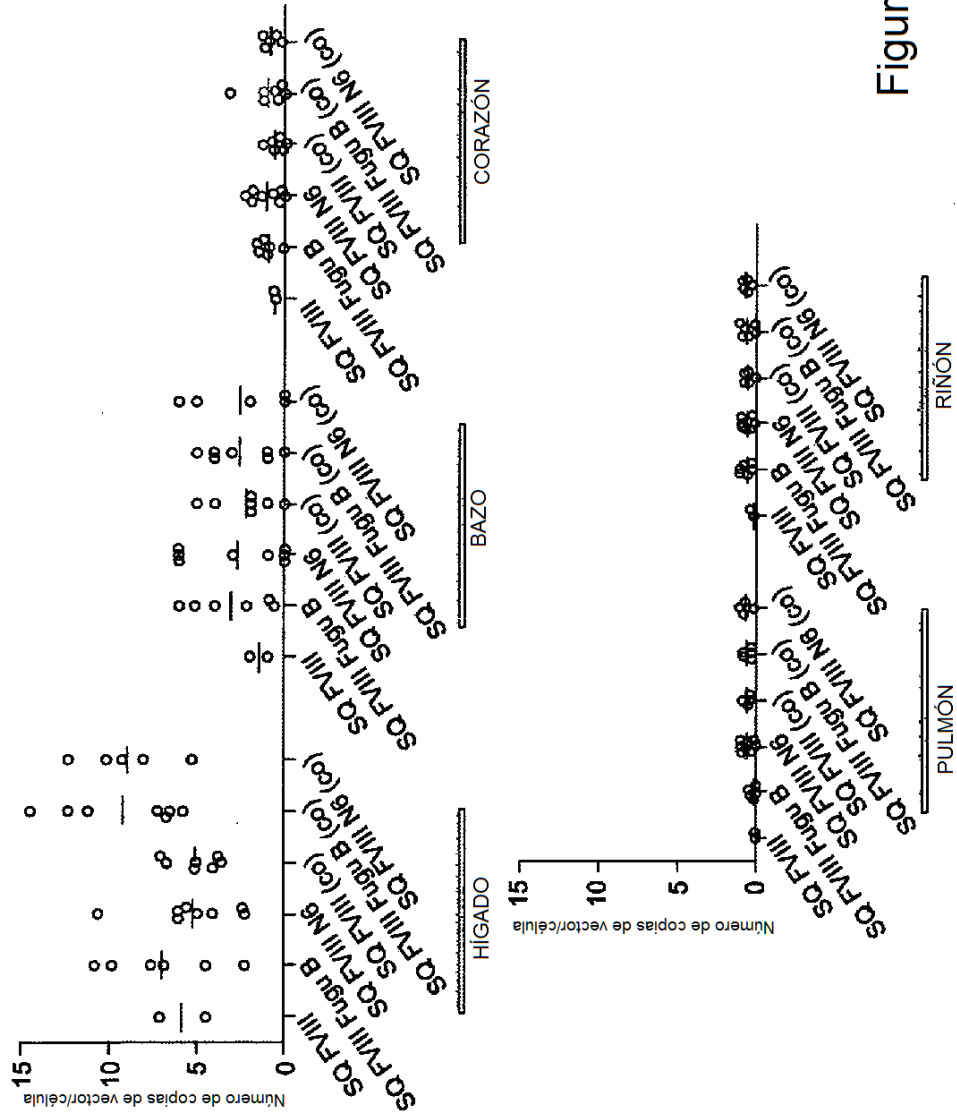


Figura 12