

PRODUCCIÓN DE ADN CON EXTREMOS CERRADOS CON SECUENCIAS DE REPETICIONES TERMINALES INVERTIDAS

REFERENCIA AL LISTADO DE SECUENCIAS PRESENTADO ELECTRÓNICAMENTE

[0001] El contenido del listado de secuencias presentado electrónicamente en formato XML (Nombre: SA9-481_SeqListing.xml; Tamaño: 115 kB; y Fecha de Creación: 19 de agosto de 2022) se incorpora en la presente memoria por referencia en su totalidad.

ANTECEDENTES

[0002] La terapia génica ofrece el potencial de un medio duradero para tratar diversas enfermedades. En el pasado, muchos tratamientos de terapia génica se basaban normalmente en el uso de vectores virales. Existen numerosos agentes virales que podrían seleccionarse para este fin, cada uno con propiedades distintas que los hacen más o menos adecuados para la terapia génica. Sin embargo, las propiedades no deseadas de algunos vectores virales han dado lugar a problemas de seguridad clínica y han limitado su uso terapéutico.

[0003] El virus adeno-asociado (AAV) es un vector común de terapia génica, pero no está exento de inconvenientes. Las secuencias codificantes del genoma del AAV están flanqueadas por repeticiones terminales invertidas (ITR), necesarias para la replicación y el empaquetamiento del virus, así como para la expresión del transgén. Las estructuras de horquilla en forma de T de las ITR del AAV son susceptibles de ser unidas por proteínas de la célula huésped que inhiben la expresión del transgén en los vectores del AAV. Existe la necesidad de proporcionar una expresión eficiente y persistente de las secuencias diana, evitando al mismo tiempo las limitaciones de la tecnología actual de vectores AAV.

RESUMEN DE LA DIVULGACIÓN

[0004] Se divulgan en la presente memoria moléculas de ácido nucleico y usos de las mismas que comprenden una primera repetición terminal invertida (ITR) y/o una segunda ITR que flanquea un casete genético que comprende una secuencia de polinucleótidos heteróloga.

[0005] En un aspecto, se proporciona en la presente memoria una molécula de ácido nucleico que comprende una primera ITR y una segunda ITR que flanquean un casete genético que comprende una secuencia de polinucleótidos heteróloga, en donde la primera ITR y la segunda ITR son ITR de bocavirus o fragmentos/derivados de las mismas (p. ej., ITR de bocavirus humano 1). En otro aspecto, se proporciona en la presente memoria una molécula de ácido nucleico que comprende una primera ITR y una segunda ITR, en donde la primera ITR comprende una secuencia de polinucleótidos al menos aproximadamente un 75 % idéntica a la SEQ ID NO:1, y la segunda ITR comprende una secuencia de polinucleótidos al menos un 75 % idéntica a la SEQ ID NO:2.

[0006] En algunas realizaciones, la primera ITR comprende una secuencia de polinucleótidos al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente

un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 %, al menos aproximadamente un 99 % idéntica a la SEQ ID NO:1. En algunas realizaciones, la primera ITR comprende la secuencia de polinucleótidos establecida en la SEQ ID NO: 1. En algunas realizaciones, la primera ITR comprende una secuencia de polinucleótidos al menos aproximadamente un 50 % idéntica a la SEQ ID NO: 1.

[0007] En algunas realizaciones, la segunda ITR comprende una secuencia de polinucleótidos al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 %, al menos aproximadamente un 99 % idéntica a la SEQ ID NO:2. En algunas realizaciones, la segunda ITR comprende la secuencia de polinucleótidos establecida en la SEQ ID NO: 2. En algunas realizaciones, la primera ITR comprende una secuencia de polinucleótidos que al menos aproximadamente un 50 % idéntica a la SEQ ID NO: 2.

[0008] En algunas realizaciones, la primera ITR comprende la secuencia de polinucleótidos establecida en la SEQ ID NO:1, y la segunda ITR comprende la secuencia de polinucleótidos establecida en la SEQ ID NO:2.

[0009] En algunas realizaciones, la molécula de ácido nucleico comprende además un casete genético que comprende una secuencia de polinucleótidos heteróloga y al menos una secuencia de control de la expresión, como un promotor, un potenciador, un intrón, una señal de terminación de la transcripción o un elemento regulador postranscripcional.

[0010] En algunas realizaciones, el casete genético comprende además un promotor. En algunas realizaciones, el promotor es un promotor específico de tejido. En algunas realizaciones, el promotor dirige la expresión de la secuencia de polinucleótidos heteróloga en un órgano, en donde el órgano comprende músculo, sistema nervioso central (SNC), ocular, hígado, corazón, riñón, páncreas, pulmones, piel, vejiga, tracto urinario, bazo, linajes celulares mieloides y linfoides, o cualquier combinación de los mismos. En algunas realizaciones, el promotor dirige la expresión de la secuencia de polinucleótidos heteróloga en hepatocitos, células epiteliales, células endoteliales, células musculares cardíacas, células musculares esqueléticas, células sinusoidales, neuronas aferentes, neuronas eferentes, interneuronas, células gliales, astrocitos, oligodendrocitos, microglía, células ependimarias, células epiteliales pulmonares, células de Schwann, células satélite, células fotorreceptoras, células ganglionares de la retina, células T, células B, células NK, macrófagos, células dendríticas, o cualquier combinación de las mismas. En algunas realizaciones, el promotor se sitúa en 5' de la secuencia de polinucleótidos heteróloga. En algunas realizaciones, el promotor es un promotor de la transtiretina de ratón (mTTR), un promotor nativo del factor VIII humano, un promotor de la antitripsina alfa-1 humana (hAAT), un promotor mínimo de la albúmina humana, un promotor de la albúmina de ratón, un promotor de la tristetraprolina (TTP) un promotor CASI, un promotor CAG, un promotor de citomegalovirus (CMV), antitripsina $\alpha 1$ (AAT), creatina quinasa muscular (MCK), cadena pesada

de miosina alfa (α MHC), mioglobina (MB), desmina (DES), SPc5-12, 2R5Sc5-12, dMCK, tMCK, o un promotor de fosfoglicerato quinasa (PGK).

[0011] En algunas realizaciones, el casete genético comprende además una secuencia intrónica. En algunas realizaciones, la secuencia intrónica se sitúa en 5' de la secuencia de polinucleótidos heteróloga. En algunas realizaciones, la secuencia intrónica se sitúa en 3' del promotor. En algunas realizaciones, la secuencia intrónica comprende una secuencia intrónica sintética.

[0012] En algunas realizaciones, el casete genético comprende además un elemento regulador postranscripcional. En algunas realizaciones, el elemento regulador se sitúa en 3' de la secuencia de polinucleótidos heteróloga. En algunas realizaciones, el elemento regulador comprende un elemento regulador postranscripcional mutado del virus de la hepatitis de la marmota (WPRE), un sitio de unión de microARN, una secuencia de direccionamiento nuclear del ADN, o cualquier combinación de los mismos.

[0013] En algunas realizaciones, el casete genético comprende además una secuencia de cola poli(A) 3'UTR. En algunas realizaciones, la secuencia de cola de poli(A) 3'UTR se selecciona del grupo que consiste en poli(A) de bGH, poli(A) de actina, poli(A) de hemoglobina y cualquier combinación de las mismas.

[0014] En algunas realizaciones, el casete genético comprende además una secuencia potenciadora. En algunas realizaciones, la secuencia potenciadora se sitúa entre la primera ITR y la segunda ITR.

[0015] En algunas realizaciones, la molécula de ácido nucleico comprende de 5' a 3': la primera ITR, el casete genético y la segunda ITR, en donde el casete genético comprende una secuencia promotora específica de tejido, una secuencia intrónica, la secuencia de polinucleótidos heteróloga, un elemento regulador postranscripcional y una secuencia de cola de poli(A) 3'UTR.

[0016] En algunas realizaciones, el casete genético comprende de 5' a 3': una secuencia promotora específica de tejido, una secuencia intrónica, la secuencia de polinucleótidos heteróloga, un elemento regulador postranscripcional y una secuencia de cola de poli(A) 3'UTR.

[0017] En algunas realizaciones, el casete genético es un ácido nucleico monocatenario. En algunas realizaciones, el casete genético es un ácido nucleico bicatenario.

[0018] En algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica una proteína terapéutica.

[0019] En algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica un factor de coagulación, un factor de crecimiento, una hormona, una citoquina, un anticuerpo, un fragmento del mismo, o cualquier combinación de los mismos. En algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica un factor de coagulación. En algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica un factor de crecimiento. En

algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica una hormona. En algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica una citoquina.

[0020] En algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica una proteína FVIII.

[0021] En algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica distrofina ligada a X, MTM1 (miotubularina), tirosina hidroxilasa, AADC, ciclohidrolasa, SMN1, FXN (frataxina), GUCY2D, RS1, CFH, HTRA, ARMS, CFB/CC2, CNGA/CNGB, Prf65, ARSA, PSAP, IDUA (MPS I), IDS (MPS II), PAH, GAA (alfa-glucosidasa ácida), GALT, OTC, CMD1A, LAMA2, o cualquier combinación de los mismos.

[0022] En algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica un microARN (miARN). En algunas realizaciones, el miARN regula a la baja la expresión de un gen diana que comprende SOD1, HTT, RHO, CD38, o cualquier combinación de los mismos.

[0023] En algunas realizaciones, la secuencia de polinucleótidos heteróloga codifica un factor de coagulación, en donde el factor de coagulación comprende el factor I (FI), el factor II (FII), el factor III (FIII), el factor IV (FIV) el factor V (FV), el factor VI (FVI), el factor VII (FVII), el factor VIII (FVIII), el factor IX (FIX), el factor X (FX), el factor XI (FXI), el factor XII (FXII), el factor XIII (FXIII), el factor de Von Willebrand (VWF), precalicreína, cininógeno de alto peso molecular, fibronectina, antitrombina III, cofactor II de la heparina, proteína C, proteína S, proteína Z, inhibidor de la proteasa relacionada con la proteína Z (ZPI), plasminógeno, antiplasmina alfa-2, activador del plasminógeno tisular (tPA), uroquinasa, inhibidor del activador del plasminógeno 1 (PAI-1), inhibidor del activador del plasminógeno 2 (PAI2), o cualquier combinación de los mismos.

[0024] En algunas realizaciones, la secuencia de polinucleótidos heteróloga está optimizada en codones. En algunas realizaciones, la secuencia de polinucleótidos heteróloga está optimizada en codones para su expresión en un ser humano.

[0025] En algunas realizaciones, la molécula de ácido nucleico se formula con un agente de administración. En algunas realizaciones, el agente de administración comprende una nanopartícula lipídica. En algunas realizaciones, la nanopartícula lipídica es ionizable. En algunas realizaciones, el agente de administración comprende liposomas, moléculas poliméricas no lipídicas, endosomas o cualquier combinación de los mismos.

[0026] En algunas realizaciones, la molécula de ácido nucleico se formula para la administración intravenosa, transdérmica, intradérmica, subcutánea, pulmonar, intraneural, intraocular, intratecal, oral, o cualquier combinación de las mismas. En algunas realizaciones, la molécula de ácido nucleico se formula para la administración intravenosa. En algunas realizaciones, la molécula de ácido nucleico se formula para la administración por inyección in situ. En algunas realizaciones, la molécula de ácido nucleico se formula para la administración por inhalación.

[0027] En otro aspecto, se proporciona en la presente memoria un vector que comprende una molécula de ácido nucleico descrita en la presente memoria.

[0028] En otro aspecto, se proporciona en la presente memoria una célula huésped que comprende una molécula de ácido nucleico descrita en la presente memoria, o un vector descrito en la presente memoria. En algunas realizaciones, la célula huésped es una célula de insecto.

[0029] En otro aspecto, se proporciona en la presente memoria una composición farmacéutica que comprende una molécula de ácido nucleico descrita en la presente memoria.

[0030] En otro aspecto, se proporciona en la presente memoria una composición farmacéutica que comprende un vector descrito en la presente memoria y un excipiente farmacéuticamente aceptable.

[0031] En otro aspecto, se proporciona en la presente memoria una composición farmacéutica que comprende una célula huésped descrita en la presente memoria y un excipiente farmacéuticamente aceptable.

[0032] En otro aspecto, se proporciona en la presente memoria un kit que comprende una molécula de ácido nucleico descrita en la presente memoria e instrucciones para administrar la molécula de ácido nucleico a un sujeto que lo necesite.

[0033] En otro aspecto, se proporciona en la presente memoria un sistema de baculovirus para la producción de una molécula de ácido nucleico descrita en la presente memoria.

[0034] En algunas realizaciones, la molécula de ácido nucleico se produce en células de insecto.

[0035] En otro aspecto, se proporciona en la presente memoria un sistema de administración de nanopartículas que comprende una molécula de ácido nucleico descrita en la presente memoria.

[0036] En otro aspecto, se proporciona en la presente memoria un método para expresar una secuencia de polinucleótidos heteróloga en un sujeto que lo necesita, que comprende administrar al sujeto una molécula de ácido nucleico descrita en la presente memoria, un vector descrito en la presente memoria o una composición farmacéutica descrita en la presente memoria.

[0037] En otro aspecto, se proporciona en la presente memoria un método para tratar una enfermedad o trastorno en un sujeto que lo necesita, que comprende administrar al sujeto una molécula de ácido nucleico descrita en la presente memoria, un vector descrito en la presente memoria o una composición farmacéutica descrita en la presente memoria.

[0038] En algunas realizaciones, la molécula de ácido nucleico se administra por vía intravenosa, transdérmica, intradérmica, subcutánea, oral, pulmonar, intraneural, intraocular, intratecal o cualquier combinación de las mismas. En algunas realizaciones, la molécula de ácido nucleico se administra por vía intravenosa. En algunas realizaciones, la molécula de ácido nucleico se administra por inyección in situ. En algunas realizaciones, la molécula de ácido nucleico se administra por inhalación.

[0039] En algunas realizaciones, el sujeto es un mamífero. En algunas realizaciones, el sujeto es un ser humano.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0040] Las **FIGs. 1A-1C** son una representación esquemática de los enfoques utilizados para la producción de ADNce en el sistema de baculovirus según una realización de la invención. La **FIG. 1A** muestra un diagrama esquemático del enfoque de un BAC, en donde se utilizó un único BEV recombinante que codifica los genes FVIIIXTEN y Rep en diferentes loci para la infección en células Sf9 para la producción de ADNce. La **FIG. 1B** muestra un diagrama esquemático del enfoque de dos BAC, en donde las células Sf9 fueron coinfectadas con BEV recombinantes que codifican los genes FVIIIXTEN y/o Rep para la producción de ADNce. La **FIG. 1C** muestra un diagrama esquemático de un enfoque de línea celular estable, en donde el casete de expresión de FVIIIXTEN se integró de forma estable en el genoma de las células Sf9 y se rescató infectando el BEV recombinante que codifica el gen Rep para la producción de ADNce.

[0041] Las **FIGs. 2A-2B** son una representación esquemática de la construcción de expresión del FVIIIXTEN humano. La **FIGs. 2A** muestra el mapa lineal esquemático de la construcción de expresión según una realización de la invención que comprende el Factor VIII humano optimizado en codones (coFVIII) con el dominio B delecionado (BDD) fusionado con el péptido XTEN 144 (FVIIIXTEN) bajo la regulación del promotor de la transtiretina de ratón modificado (mTTR) específico del hígado (mTTR482) con el elemento potenciador (A1MB2), intrón sintético híbrido (Intrón Quimérico), el Elemento Regulador Postranscripcional de Marmota (WPRE), y la señal de Poliadenilación de la Hormona de Crecimiento Bovina (bGHpA). El casete de expresión del FVIIIXTEN está flanqueado por las ITR del Bocavirus humano tipo 1 (HBoV1) de tipo salvaje (WT). (SEQ ID NO: 1 y SEQ ID NO: 2). La **FIG. 2B** muestra un mapa esquemático del vector de transferencia Tn7 según una realización de la invención realizado mediante la inserción del casete de expresión de FVIIIXTEN (SEQ ID NO: 3) en el vector pFastBac1 (Invitrogen).

[0042] Las **FIGs. 3A-3C** son una representación esquemática de construcciones de expresión del gen de Replicación (Rep) según las realizaciones de la invención. La **FIG. 3A** muestra un mapa lineal esquemático de un ADN sintético que codifica el gen NS1 de HBoV1 optimizado en codones para Sf bajo el promotor de *polihedrina* AcMNPV seguido de la señal de poliadenilación del SV40 (SV40 PAS). La **FIG. 3B** muestra un mapa esquemático de un vector de transferencia Tn7 según una realización de la invención preparado mediante la inserción del ADN sintético de NS1 de HBoV1 (SEQ ID NO: 4) en el vector pFastBac1 (Invitrogen). La **FIG. 3C** muestra un mapa esquemático de un vector donante Cre-LoxP según una realización de la invención preparado mediante la inserción del ADN sintético de NS1 de HBoV1 (SEQ ID NO: 4) en el vector donante Cre-LoxP creado como se describe en el Sistema de Expresión de Baculovirus, Solicitud de Patente de EE. UU. No. 63/069.073, incorporada en la presente memoria por referencia en su totalidad.

[0043] Las **FIGs. 4A-4C** son una representación esquemática de construcciones de expresión del gen de Replicación (Rep) según las realizaciones de la invención. La **FIG. 4A** muestra un mapa lineal esquemático de un ADN sintético que codifica el gen NS1 de HBoV1 optimizado en codones para Sf bajo el promotor *inmediato-temprano1* de AcMNPV (pIE1) precedido por el elemento potenciador transcripcional *hr5* del AcMNPV seguido por la señal de poliadenilación del SV40 (SV40 PAS). La **FIG. 4B** muestra un mapa esquemático de un vector de transferencia Tn7 según una realización de la invención preparado mediante la inserción del ADN sintético de NS1 de HBoV1 (SEQ ID NO: 4) en el vector pFastBac1 (Invitrogen). La **FIG. 4C** muestra un mapa esquemático de un vector donante Cre-LoxP según una realización de la invención preparado mediante la inserción del ADN sintético de NS1 de HBoV1 (SEQ ID NO: 4) en el vector donante Cre-LoxP.

[0044] Las **FIGs. 5A-5D** son representaciones esquemáticas de construcciones de expresión del gen de Replicación (Rep) según las realizaciones de la invención. La **FIG. 5A** muestra un mapa lineal esquemático de un ADN sintético que codifica el gen NS1 de HBoV1 optimizado en codones para Sf bajo el promotor *inmediato-temprano2* de OpMNPV (OpIE2) seguido de la señal de poliadenilación SV40 (SV40 PAS). La **FIG. 5B** muestra un mapa lineal esquemático de un ADN sintético que codifica el gen NS1 de HBoV1 optimizado en codones para Sf bajo el promotor *inmediato-temprano1* de AcMNPV (pIE1) seguido de la señal de poliadenilación de SV40 (SV40 PAS). La **FIG. 5C** muestra un mapa esquemático de un vector de transferencia Tn7 según una realización de la invención preparado mediante la inserción del ADN sintético de NS1 de HBoV1 (SEQ ID NO: 4) en el vector pFastBac1 (Invitrogen). La **FIG. 5D** muestra un mapa esquemático de un vector donante Cre-LoxP según una realización de la invención preparado mediante la inserción del ADN sintético de NS1 de HBoV1 (SEQ ID NO: 4) bajo el promotor *inmediato-temprano1* de AcMNPV (pIE1) en el vector donante Cre-LoxP creado como se describe en la Solicitud de Patente de EE. UU. No. 63/069.073.

[0045] Las **FIGs. 6A-6B** muestran la generación del vector de expresión de baculovirus recombinante (BEV) que codifica FVIIIIXTEN humano con ITR de HBoV1. La **FIG. 6A** es una imagen de electroforesis en gel de agarosa del mapeo de enzimas de restricción de los clones de bácmido BIVVBac recombinantes que codifican FVIIIIXTEN humano con ITR de HBoV1 (BIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7}). Las **FIG. 6B** es una representación esquemática de BEV recombinante que codifica el casete de expresión de FVIIIIXTEN flanqueado por las ITR de HBoV1 (SEQ ID NO: 3) como se indica (AcBIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7}).

[0046] Las **FIGs. 7A-7F** son una representación esquemática de un BAC que comprende casetes de expresión de los genes de FVIIIIXTEN humano y Rep y estudios de confirmación de los mismos. Las **FIGs. 7A-C** son imágenes de electroforesis en gel de agarosa de clones de bácmido recombinantes de BIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP} (**FIG. 7A**), BIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)IE1.HBoV1.NS1^{LoxP} (**FIG. 7B**), y BIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)HR5.IE1.HBoV1.NS1^{LoxP} (**FIG. 7C**) cribados por los

cebadores de PCR exterior/interior (SEQ ID NO: 5 y SEQ ID NO: 6) como se indica por las flechas rojas en las **FIG. 7D-7F**. La **FIG. 7D** muestra un mapa esquemático de los vectores de expresión de baculovirus (BEV) recombinantes que codifican NS1 de HBoV1 bajo el promotor de *Polihedrina* de AcMNPV (pPolh) y el casete de expresión FVIIIIXTEN flanqueado por las ITR de HBoV1 como se indica (AcBIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP}). La **FIG. 7E** muestra un mapa esquemático de BEV recombinante que codifica NS1 de HBoV1 bajo el promotor *inmediato-temprano1* de AcMNPV (pIE1) y el casete de expresión FVIIIIXTEN flanqueado por las ITR de HBoV1 como se indica (AcBIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)IE1.HBoV1.NS1^{LoxP}). La **FIG. 7F** muestra un mapa esquemático de BEV recombinante que codifica NS1 de HBoV1 bajo el promotor *inmediato-temprano1* de AcMNPV precedido por el elemento potenciador transcripcional *hr5* de AcMNPV (pHR5.IE1) y el casete de expresión FVIIIIXTEN flanqueado por las ITR de HBoV1 como se indica (AcBIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)HR5.IE1.HBoV1.NS1^{LoxP}).

[0047] Las **FIGs. 8A-8C** muestran la producción del vector de ADNce de FVIIIIXTEN humano utilizando el enfoque Un BAC según una realización de la invención. La **FIG. 8A** es un diagrama esquemático del enfoque Un BAC de la producción del vector de ADNce de FVIIIIXTEN en células Sf9 utilizando BEV recombinante que codifica el gen NS1 de HBoV1 bajo el promotor de *polihedrina* de AcMNPV y el casete de expresión de FVIIIIXTEN humano flanqueado por las ITR de HBoV1 (AcBIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP}). La **FIG. 8B** muestra el mapa esquemático de AcBIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP} BEV. La **FIG. 8C** es una imagen de electroforesis en gel de agarosa del vector de ADNce aislado de células Sf9 infectadas con una preparación madre de virus titulada (P2) de (AcBIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP}) BEV. Las bandas de ADN correspondientes al tamaño del ADNce de FVIIIIXTEN (ADNce), el ADN baculoviral (ADNv) y el ADN genómico de las células Sf9 (ADNg) se indican con flechas.

[0048] Las **FIGs. 9A-9D** muestran esquemas de vectores de expresión de baculovirus (BEV) recombinantes que comprenden la secuencia que codifica NS1 de HBoV1 y estudios de confirmación de los mismos. La **FIG. 9A** es una imagen de electroforesis en gel de agarosa del mapeo con enzimas de restricción de clones de bácmido recombinantes de HBoV1.NS1 bajo el promotor de *polihedrina* de AcMNPV, *inmediato-temprano1* precedido por el elemento potenciador transcripcional *hr5* de AcMNPV o el promotor *inmediato-temprano2* de OpMNPV (BIVVBac.Polh.HBoV1.NS1^{Tn7}, BIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}, y BIVVBac.OpIE2.HBoV1.NS1^{Tn7}, respectivamente). La **FIG. 9B** muestra un mapa esquemático de AcBIVVBac.Polh.HBoV1.NS1^{Tn7}. La **FIG. 9C** muestra un mapa esquemático de AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}. La **FIG. 9D** muestra un mapa esquemático de AcBIVVBac.OpIE2.HBoV1.NS1^{Tn7}.

[0049] Las **FIGs. 10A-10C** muestran la producción del vector de ADNce de FVIIIIXTEN humano utilizando el enfoque de Dos BAC según una realización de la invención. La **FIG. 10A** es un diagrama esquemático del enfoque de Dos BAC para la producción del vector de ADNce de FVIIIIXTEN, en donde las células Sf9 se coinfectan con un BEV recombinante que codifica el casete de expresión de FVIIIIXTEN flanqueado por las ITR de HBoV1 (AcBIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7}) y/o que codifica el gen NS1 de HBoV1 bajo el promotor de *polihedrina* de AcMNPV (AcBIVVBac.Polh.HBoV1.NS1^{Tn7}). La **FIG. 10B** muestra los mapas esquemáticos de BEV AcBIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7} y AcBIVVBac.Polh.HBoV1.NS1^{Tn7}. La **FIG. 10C** es una imagen de electroforesis en gel de agarosa del vector de ADNce aislado de células Sf9 coinfectadas a diferentes MOI de proporción constante o diferentes proporciones de MOI constantes de BEV AcBIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7} y AcBIVVBac.Polh.HBoV1.NS1^{Tn7} como se indica. Las bandas de ADN correspondientes al tamaño del vector de ADNce de FVIIIIXTEN (ADNce), el ADN baculoviral (ADNv) y el ADN genómico de las células Sf9 (ADNg) se indican con flechas.

[0050] Las **FIGs. 11A-11C** muestran la producción del vector de ADNce de FVIIIIXTEN humano utilizando el enfoque de Dos BAC según una realización de la invención. La **FIG. 11A** es un diagrama esquemático del enfoque de Dos BAC para la producción del vector de ADNce de FVIIIIXTEN, en donde las células Sf9 se coinfectan con un BEV recombinante que codifica el casete de expresión de FVIIIIXTEN flanqueado por las ITR de HBoV1 (AcBIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7}) y/o que codifica el gen NS1 de HBoV1 bajo el promotor de *polihedrina* de AcMNPV (AcBIVVBac.Polh.HBoV1.NS1^{Tn7}) o el promotor *inmediato-temprano1* precedido por el elemento potenciador transcripcional *hr5* de AcMNPV (AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}). La **FIG. 11B** muestra los mapas esquemáticos de BEV AcBIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7} y AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}. La **FIG. 11C** es una imagen de electroforesis en gel de agarosa del vector de ADNce aislado de células Sf9 coinfectadas a diferentes MOI de BEV AcBIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7} y AcBIVVBac.Polh.HBoV1.NS1^{Tn7} (imagen izquierda) o AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7} (imagen derecha). Las bandas de ADN correspondientes al tamaño del vector de ADNce de FVIIIIXTEN (ADNce), el ADN baculoviral (ADNv) y el ADN genómico de las células Sf9 (ADNg) se indican con flechas.

[0051] Las **FIGs. 12A-12C** muestran los materiales utilizados en la generación de una línea celular estable que codifica el casete de expresión del FVIIIIXTEN flanqueado por las ITR del HBoV1. La **FIG. 12A** muestra un mapa esquemático de un plásmido que codifica el marcador de resistencia a la neomicina bajo el promotor *inmediato-temprano1* de AcMNPV (IE1) precedido por el elemento potenciador transcripcional *hr5* de AcMNPV y seguido por la señal de poliadenilación p10 del AcMNPV (P10 PAS). La **FIG. 12B** muestra un mapa esquemático de un plásmido que codifica el marcador de la proteína fluorescente verde mejorada (eGFP) bajo el

promotor *inmediato-temprano1* de AcMNPV (IE1) precedido por el elemento potenciador transcripcional *hr5* de AcMNPV y seguido por la señal de poliadenilación p10 del AcMNPV (P10 PAS). La **FIG. 12C** muestra un mapa esquemático del casete de expresión de FVIIIIXTEN flanqueado por las ITR del HBoV1 (SEQ ID NO: 3), que se integró de forma estable en el genoma de las células Sf9 para generar una línea celular estable.

[0052] La **FIGs. 13A-13E** muestra el flujo de trabajo de la producción y purificación del vector de ADNce de FVIIIIXTEN utilizando el enfoque de Dos BAC según una realización de la invención. La **FIG. 13A** muestra un esquema de la expansión y duración de las células Sf9 (Día 0-2), donde las células se someten a aumento de escala secuencialmente desde frascos de cultivo a pequeña escala (0,5L) hasta gran escala (1,5L) para alcanzar la densidad celular de 2,5 a 3,0 x 10⁶/mL en medio ESF921 libre de suero. La **FIG. 13B** muestra un esquema de la infección de matraces de cultivo a gran escala (1,5L) de Sf9 y la duración de la incubación (Día 2-6), donde las células se coinfectan con un BEV recombinante que codifica un BEV recombinante que codifica un casete de expresión de FVIIIIXTEN flanqueado por las ITR de HBoV1 (AcBIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7}) y/o que codifica el gen NS1 de HBoV1 bajo el promotor de *polihedrina* de AcMNPV (AcBIVVBac.Polh.HBoV1.NS1^{Tn7}) a una MOI de 0,1 y 0,01 unidades formadoras de placas (pfu)/célula, respectivamente. La **FIG. 13C** muestra una imagen del kit de purificación del Plásmido Giga Prep y de la electroforesis en gel de agarosa con la duración del procesamiento (Día 6-7), donde se miden diariamente la densidad y la viabilidad celular de las células infectadas, y las células se sedimentaron por centrifugación a baja velocidad una vez que la viabilidad celular alcanzó el 70-80 %. El vector de ADNce de FVIIIIXTEN se purificó a partir de los sedimentos de las células infectadas mediante el kit PureLinkTM HiPure Expi Plasmid Gigaprep (Invitrogen) y se pasó una alícuota por electroforesis en gel de agarosa para determinar la productividad del ADNce de FVIIIIXTEN (ADNce), el ADN baculoviral (ADNv) y/o el ADN genómico de células Sf9 (ADNg). La **FIG. 13D** muestra una imagen de la Bio-Rad Modelo 491 Prep Cell y la electroforesis en gel de agarosa con la duración del procesamiento (Día 7-12), donde los ADN purificados por Giga-prep se cargaron en el gel de agarosa preparativo en Prep Cell para separar el ADNce de FVIIIIXTEN (fragmento de ~8,5kb) de los ADN de alto peso molecular. Las fracciones de elución recogidas a intervalos de 70-80 minutos de la electroforesis en gel de agarosa preparativa se analizaron en un gel de agarosa al 0,8 a 1,2 % para determinar la pureza del ADNce de FVIIIIXTEN. La **FIG. 13E** muestra una imagen de la electroforesis en gel de agarosa, donde las fracciones recogidas de la Prep Cell se combinaron y precipitaron con 1/10^o vol de NaOAc 3M (pH 5,5) y 3 vol de etanol al 100 % para obtener el ADNce de FVIIIIXTEN purificado. La imagen del gel muestra la pureza del ADNce de FVIIIIXTEN en comparación con el material de partida, con flechas que indican las bandas de ADN correspondientes al tamaño del vector de ADNce de FVIIIIXTEN (ADNce), el ADN baculoviral (ADNv) y el ADN genómico de las células Sf9 (ADNg).

[0053] Las **FIGs. 14A-14B** muestran la representación gráfica de los niveles de actividad plasmática del FVIII medidos mediante los ensayos cromogénicos del Factor VIII Chromogenix Coatest® SP. La **FIG. 14A** muestra la representación gráfica de los niveles de actividad plasmática del FVIII medidos en muestras de sangre recogidas a diferentes intervalos de ratones hFVIII^{R593C^{+/+}}/HemA a los que se inyectaron sistémicamente a *mediante una* inyección hidrodinámica en la vena de la cola 1.600 o 400 µg/kg de ADN de ADN monocatenario (ADNss) de FVIII^{XTEN} ITR de HBoV1. La **FIG. 14B** muestra la representación gráfica de los niveles plasmáticos de actividad del FVIII medidos en muestras de sangre recogidas a diferentes intervalos de ratones hFVIII^{R593C^{+/+}}/HemA a los que se inyectaron sistémicamente *mediante* inyección hidrodinámica en la vena de la cola 80, 40 o 12 µg/kg de ADN^{ce} (ADN^{ce}) de FVIII^{XTEN} ITR de HBoV1. Las barras de error representan la desviación estándar.

[0054] Las **FIGs. 15A-15C** muestran las imágenes microscópicas de fluorescencia roja (panel superior) o de campo brillante (panel inferior) de células Sf9 cotransfectadas con ADN de bácmido AcBIVVBac.Polh.HBoV1.NS1^{Tn7} y ARNsg de VP80 según una realización de la invención. La **FIG. 15A** muestra las imágenes microscópicas de células cotransfectadas con ADN de bácmido AcBIVVBac.Polh.HBoV1.NS1^{Tn7} y Cas9 solo. La **FIG. 15B** muestra las imágenes microscópicas de células cotransfectadas con ADN de bácmido AcBIVVBac.Polh.HBoV1.NS1^{Tn7} y sgRNA.VP80.T1. La **FIG. 15C** muestra las imágenes microscópicas de células cotransfectadas con ADN de bácmido AcBIVVBac.Polh.HBoV1.NS1^{Tn7} y sgRNA.VP80.T2.

[0055] Las **FIGs. 16A-16C** muestran la generación de BEV de VP80KO. La **FIG. 16A** y la **FIG. 16B** ilustran los análisis TIDE de los BEV clonales AcBIVVBac.Polh.HBoV1.NS1ΔVP80^{Tn7} y AcBIVVBac.OpIE2.HBoV1.NS1ΔVP80^{Tn7} para determinar los índices inducidos por CRISPR/Cas9. La **FIG. 16C** es una imagen de electroforesis en gel de agarosa del vector de ADN^{ce} de FVIII^{XTEN} aislado de células Sf9 coinfectadas a diferentes MOI de BEV AcBIVVBac.mTTR.FVIII^{XTEN}.HBoV1.ITRs^{Tn7} y AcBIVVBac.Polh.HBoV1.NS1ΔVP80^{Tn7} o AcBIVVBac.OpIE2.HBoV1.NS1ΔVP80^{Tn7} como se indica. Las bandas de ADN correspondientes al tamaño del vector de ADN^{ce} FVIII^{XTEN} (ADN^{ce}), el ADN baculoviral (ADN^v) y el ADN genómico de las células Sf9 (ADN^g) se indican con flechas.

[0056] Las **FIGs. 17A-17C** muestran la producción de ADN^{ce} FVIII^{XTEN} HBoV1 humano utilizando el enfoque DosBAC. La **FIG. 17A** es un diagrama esquemático del enfoque DosBAC de la producción del vector de ADN^{ce} de FVIII^{XTEN}. La **FIG. 17B** muestra los mapas esquemáticos de BEV AcBIVVBac.mTTR.FVIII^{XTEN}.HBoV1.ITRs^{Tn7} y AcBIVVBac.Polh.HBoV1.NS1^{Tn7}. La **FIG. 17C** es una imagen de electroforesis en gel de agarosa del vector de ADN^{ce} de FVIII^{XTEN} aislado de células Sf9 coinfectadas con BEV AcBIVVBac.mTTR.FVIII^{XTEN}.HBoV1.ITRs^{Tn7} y AcBIVVBac.Polh.HBoV1.NS1^{Tn7} a 1,0, 2,0, 3,0, 4,0, o 5,0 MOI. Las bandas de ADN correspondientes al tamaño del vector de ADN^{ce} de

FVIIIIXTEN (ADNce), el ADN baculoviral (ADNv) y el ADN genómico de las células Sf9 (ADNg) se indican con flechas.

[0057] Las **FIGs. 18A-18D** muestran la producción del vector de ADNce FVIIIIXTEN HBoV1 humano utilizando el enfoque UnBAC. La **FIG. 18A** es un diagrama esquemático del enfoque UnBAC de la producción del vector de ADNce de FVIIIIXTEN en células Sf9. La **FIG. 18B** muestra el mapa esquemático de BEV AcBIVVBac(mTTR.FVIIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP}. La **FIG. 18C** es una imagen de electroforesis en gel de agarosa del ADNce aislado de los BEV clonales HBoV1 UnBAC amplificados a P2 en células Sf9. La **FIG. 18D** es una imagen de electroforesis en gel de agarosa de ADNce aislado de células Sf9 infectadas con BEV HBoV1 UnBAC a 0,1, 0,2, 0,3, 0,4, o 0,5 MOI. Las bandas de ADN correspondientes al tamaño del ADNce de FVIIIIXTEN (ADNce), el ADN baculoviral (ADNv) y el ADN genómico de las células Sf9 (ADNg) se indican con flechas.

[0058] Las **FIGs. 19A-19C** muestran la generación y el ensayo de ADNss y ADNce de HBoV1 in vivo. La **FIG. 19A** es una imagen de electroforesis en gel de agarosa del ADN monocatenario (ADNss) FVIIIIXTEN HBoV1. La **FIG. 19B** es una imagen de electroforesis en gel de agarosa del ADNce FVIIIIXTEN HBoV1. La **FIG. 19C** muestra los niveles de expresión del FVIII normalizados respecto al porcentaje de lo normal para FVIIIIXTENss y FVIIIIXTENce. Las barras de error representan la desviación estándar.

[0059] Las **FIGs. 20A-20B** muestran el ensayo de las formas monoméricas y multiméricas del ADNce FVIIIIXTEN HBoV1. La **FIG. 20A** es una imagen de electroforesis en gel de agarosa de las formas monoméricas y multiméricas de ADNce FVIIIIXTEN HBoV1. La **FIG. 20B** muestra los niveles de expresión de FVIII normalizados respecto al porcentaje de lo normal en ratones a los que se inyectaron formas monoméricas y multiméricas de ADNce FVIIIIXTEN HBoV1. Las barras de error representan la desviación estándar.

[0060] Las **FIGs. 21A-21C** muestran el ensayo del promotor mTTR específico del hígado y A1AT humano que dirige la expresión de FVIIIIXTEN en construcciones de ITR de HBoV1. La **FIG. 21A** es un diagrama esquemático de los casetes de expresión de FVIIIIXTEN con el promotor mTTR específico del hígado o A1AT flanqueado por ITR de HBoV1 WT. La **FIG. 21B** es una imagen de electroforesis en gel de agarosa del ADN monocatenario (ADNss) FVIIIIXTEN HBoV1 generado por digestión con enzimas de restricción, tal como se describe. La **FIG. 21C** muestra los niveles de expresión del FVIII normalizados respecto al porcentaje de lo normal en ratones a los que se inyectaron las construcciones del promotor mTTR o A1AT representadas en la FIG. 21A. Las barras de error representan la desviación estándar.

DESCRIPCIÓN DETALLADA

[0061] Se divulgan en la presente memoria moléculas de ácido nucleico y usos de las mismas que comprenden una primera repetición terminal invertida (ITR) modificada y/o una segunda ITR modificada que flanquea un casete genético que comprende una secuencia de polinucleótidos

heterólogo. En algunas realizaciones, la primera y/o segunda ITR se derivan del bocavirus humano 1 (HBoV1).

[0062] Las construcciones ejemplares de la divulgación se ilustran en las figuras y el listado de secuencias adjuntos. Con el fin de proporcionar una clara comprensión de la memoria descriptiva y de las reivindicaciones, las siguientes definiciones se proporcionan a continuación.

Definiciones

[0063] Cabe señalar que el término "una" o "un" entidad se refiere a una o más de esa entidad: por ejemplo, "una secuencia de nucleótidos" se entiende que representa una o más secuencias de nucleótidos. Del mismo modo, "una proteína terapéutica" y "un miARN" se entiende que representan una o más proteínas terapéuticas y uno o más miARN, respectivamente. Por lo tanto, los términos "un" (o "una"), "uno o más" y "al menos uno" pueden utilizarse indistintamente en la presente memoria.

[0064] El término "aproximadamente" se utiliza en la presente memoria para significar aproximadamente, más o menos, alrededor de, o en las regiones de. Cuando el término "aproximadamente" se utiliza junto con un rango numérico, modifica dicho rango ampliando los límites por encima y por debajo de los valores numéricos establecidos. En general, el término "aproximadamente" se utiliza en la presente memoria para modificar un valor numérico por encima y por debajo del valor establecido con una variación del 10 por ciento, hacia arriba o hacia abajo (mayor o menor).

[0065] También y como se utiliza en la presente memoria, "y/o" se refiere y abarca todas y cada una de las posibles combinaciones de uno o más de los elementos asociados enumerados, así como la falta de combinaciones cuando se interpreta en la alternativa ("o").

[0066] "Ácidos nucleicos", "moléculas de ácido nucleico", "nucleótidos", "secuencia de nucleótidos" y "polinucleótido" se utilizan indistintamente y se refieren a la forma polimérica de éster de fosfato de los ribonucleósidos (adenosina, guanosina, uridina o citidina; "moléculas de ARN") o desoxirribonucleósidos (desoxiadenosina, desoxiguanosina, desoxitiimidina o desocitidina; "moléculas de ADN"), o cualquier análogo de fosfoéster de los mismos, como los fósforotioatos y los tioésteres, ya sea en forma monocatenaria o en forma de hélice bicatenaria. Las secuencias de ácido nucleico monocatenario se refieren al ADN monocatenario (ADNss) o al ARN monocatenario (ARNss). Son posibles las hélices bicatenarias ADN-ADN, ADN-ARN y ARN-ARN. El término molécula de ácido nucleico, y en particular molécula de ADN o ARN, se refiere únicamente a la estructura primaria y secundaria de la molécula, y no se limita a ninguna forma terciaria particular. Así, este término incluye el ADN bicatenario que se encuentra, *inter alia*, en las moléculas de ADN lineal o circular (p. ej., fragmentos de restricción), plásmidos, ADN superenrollado y cromosomas. Al discutir la estructura de moléculas particulares de ADN bicatenario, las secuencias pueden describirse en la presente memoria según la convención normal de dar sólo la secuencia en la dirección 5' a 3' a lo largo de la cadena no transcrita del

ADN (es decir, la cadena que tiene una secuencia homóloga al ARNm). Una "molécula de ADN recombinante" es una molécula de ADN que ha sido sometida a una manipulación de biología molecular. El ADN incluye, pero no se limita a, ADNc, ADN genómico, ADN plasmídico, ADN sintético y ADN semisintético. Una "composición de ácido nucleico" de la divulgación comprende uno o más ácidos nucleicos como se describe en la presente memoria.

[0067] Tal y como se utiliza en la presente memoria, una "repetición terminal invertida" (o "ITR") se refiere a una subsecuencia de ácido nucleico situada en el extremo 5' o 3' de una secuencia de ácido nucleico monocatenaria, que comprende un conjunto de nucleótidos (secuencia inicial) seguido aguas abajo por su complemento inverso, es decir, una secuencia palindrómica. La secuencia intermedia de nucleótidos entre la secuencia inicial y el complemento inverso puede tener cualquier longitud, incluso cero. En una realización, la ITR útil para la presente divulgación comprende una o más "secuencias palindrómicas". Una ITR puede tener cualquier número de funciones. En algunas realizaciones, una ITR descrita en la presente memoria forma una estructura de horquilla. En algunas realizaciones, la ITR forma una estructura de horquilla en forma de T. En algunas realizaciones, la ITR forma una estructura de horquilla no en forma de T, p. ej., una estructura de horquilla en forma de U. En algunas realizaciones, la ITR promueve la supervivencia a largo plazo de la molécula de ácido nucleico en el núcleo de una célula. En algunas realizaciones, la ITR promueve la supervivencia permanente de la molécula de ácido nucleico en el núcleo de una célula (p. ej., durante toda la vida de la célula). En algunas realizaciones, la ITR promueve la estabilidad de la molécula de ácido nucleico en el núcleo de una célula. En algunas realizaciones, la ITR promueve la retención de la molécula de ácido nucleico en el núcleo de una célula. En algunas realizaciones, la ITR promueve la persistencia de la molécula de ácido nucleico en el núcleo de una célula. En algunas realizaciones, la ITR inhibe o evita la degradación de la molécula de ácido nucleico en el núcleo de una célula.

[0068] En una realización, la secuencia inicial de la ITR y/o el complemento inverso comprenden aproximadamente 2-600 nucleótidos, aproximadamente 2-550 nucleótidos, aproximadamente 2-500 nucleótidos, aproximadamente 2-450 nucleótidos, aproximadamente 2-400 nucleótidos, aproximadamente 2-350 nucleótidos, aproximadamente 2-300 nucleótidos o aproximadamente 2-250 nucleótidos. En algunas realizaciones, la secuencia inicial y/o el complemento inverso comprenden aproximadamente 5-600 nucleótidos, aproximadamente 10-600 nucleótidos, aproximadamente 15-600 nucleótidos, aproximadamente 20-600 nucleótidos, aproximadamente 25-600 nucleótidos, aproximadamente 30-600 nucleótidos, aproximadamente 35-600 nucleótidos, aproximadamente 40-600 nucleótidos, aproximadamente 45-600 nucleótidos, aproximadamente 50-600 nucleótidos, aproximadamente 60-600 nucleótidos, aproximadamente 70-600 nucleótidos, aproximadamente 80-600 nucleótidos, aproximadamente 90-600 nucleótidos, aproximadamente 100-600 nucleótidos, aproximadamente 150-600 nucleótidos, aproximadamente 200-600 nucleótidos, aproximadamente 300-600 nucleótidos,

aproximadamente 350-600 nucleótidos, aproximadamente 400-600 nucleótidos, aproximadamente 450-600 nucleótidos, aproximadamente 500-600 nucleótidos, o aproximadamente 550-600 nucleótidos. En algunas realizaciones, la secuencia inicial y/o el complemento inverso comprenden aproximadamente 5-550 nucleótidos, aproximadamente 5-500 nucleótidos, aproximadamente 5-450 nucleótidos, aproximadamente 5-400 nucleótidos, aproximadamente 5-350 nucleótidos, aproximadamente 5-300 nucleótidos o aproximadamente 5-250 nucleótidos. En algunas realizaciones, la secuencia inicial y/o el complemento inverso comprenden aproximadamente 10-550 nucleótidos, aproximadamente 15-500 nucleótidos, aproximadamente 20-450 nucleótidos, aproximadamente 25-400 nucleótidos, aproximadamente 30-350 nucleótidos, aproximadamente 35-300 nucleótidos, o aproximadamente 40-250 nucleótidos. En ciertas realizaciones, la secuencia inicial y/o el complemento inverso comprenden aproximadamente 225 nucleótidos, aproximadamente 250 nucleótidos, aproximadamente 275 nucleótidos, aproximadamente 300 nucleótidos, aproximadamente 325 nucleótidos, aproximadamente 350 nucleótidos, aproximadamente 375 nucleótidos, aproximadamente 400 nucleótidos, aproximadamente 425 nucleótidos, aproximadamente 450 nucleótidos, aproximadamente 475 nucleótidos, aproximadamente 500 nucleótidos, aproximadamente 525 nucleótidos, aproximadamente 550 nucleótidos, aproximadamente 575 nucleótidos o aproximadamente 600 nucleótidos. En realizaciones particulares, la secuencia inicial y/o el complemento inverso comprenden aproximadamente 400 nucleótidos.

[0069] En otras realizaciones, la secuencia inicial de la ITR y/o el complemento inverso comprenden aproximadamente 2-200 nucleótidos, aproximadamente 5-200 nucleótidos, aproximadamente 10-200 nucleótidos, aproximadamente 20-200 nucleótidos, aproximadamente 30-200 nucleótidos, aproximadamente 40-200 nucleótidos, aproximadamente 50-200 nucleótidos, aproximadamente 60-200 nucleótidos, aproximadamente 70-200 nucleótidos, aproximadamente 80-200 nucleótidos, aproximadamente 90-200 nucleótidos, aproximadamente 100-200 nucleótidos, aproximadamente 125-200 nucleótidos, aproximadamente 150-200 nucleótidos, o aproximadamente 175-200 nucleótidos. En otras realizaciones, la secuencia inicial y/o el complemento inverso comprenden aproximadamente 2-150 nucleótidos, aproximadamente 5-150 nucleótidos, aproximadamente 10-150 nucleótidos, aproximadamente 20-150 nucleótidos, aproximadamente 30-150 nucleótidos, aproximadamente 40-150 nucleótidos, aproximadamente 50-150 nucleótidos, aproximadamente 75-150 nucleótidos, aproximadamente 100-150 nucleótidos o aproximadamente 125-150 nucleótidos. En otras realizaciones, la secuencia inicial y/o el complemento inverso comprenden aproximadamente 2-100 nucleótidos, aproximadamente 5-100 nucleótidos, aproximadamente 10-100 nucleótidos, aproximadamente 20-100 nucleótidos, aproximadamente 30-100 nucleótidos, aproximadamente 40-100 nucleótidos, aproximadamente 50-100 nucleótidos o aproximadamente 75-100 nucleótidos. En otras realizaciones, la secuencia inicial y/o el complemento inverso comprenden aproximadamente 2-50 nucleótidos, aproximadamente 10-50 nucleótidos, aproximadamente 20-

50 nucleótidos, aproximadamente 30-50 nucleótidos, aproximadamente 40-50 nucleótidos, aproximadamente 3-30 nucleótidos, aproximadamente 4-20 nucleótidos, o aproximadamente 5-10 nucleótidos. En otra realización, la secuencia inicial y/o el complemento inverso consisten en dos nucleótidos, tres nucleótidos, cuatro nucleótidos, cinco nucleótidos, seis nucleótidos, siete nucleótidos, ocho nucleótidos, nueve nucleótidos, diez nucleótidos, 11 nucleótidos, 12 nucleótidos, 13 nucleótidos, 14 nucleótidos, 15 nucleótidos, 16 nucleótidos, 17 nucleótidos, 18 nucleótidos, 19 nucleótidos o 20 nucleótidos. En otras realizaciones, un nucleótido intermedio entre la secuencia inicial y el complemento inverso es (p. ej., consiste en) 0 nucleótidos, 1 nucleótido, dos nucleótidos, tres nucleótidos, cuatro nucleótidos, cinco nucleótidos, seis nucleótidos, siete nucleótidos, ocho nucleótidos, nueve nucleótidos, 10 nucleótidos, 11 nucleótidos, 12 nucleótidos, 13 nucleótidos, 14 nucleótidos, 15 nucleótidos, 16 nucleótidos, 17 nucleótidos, 18 nucleótidos, 19 nucleótidos o 20 nucleótidos.

[0070] Por lo tanto, una "ITR", tal y como se utiliza en la presente memoria, puede plegarse sobre sí misma y formar un segmento bicatenario. Por ejemplo, la secuencia GATCXXXXGATC comprende una secuencia inicial de GATC y su complemento (3'CTAG5') cuando se pliega para formar una doble hélice. En algunas realizaciones, la ITR comprende una secuencia palindrómica continua (p. ej., GATCGATC) entre la secuencia inicial y el complemento inverso. En algunas realizaciones, la ITR comprende una secuencia palindrómica interrumpida (p. ej., GATCXXXXGATC) entre la secuencia inicial y el complemento inverso. En algunas realizaciones, las secciones complementarias de la secuencia palindrómica continua o interrumpida interaccionan entre sí para formar una estructura de "bucle horquilla". Tal y como se utiliza en la presente memoria, una estructura de "bucle horquilla" resulta cuando al menos dos secuencias complementarias en una molécula de nucleótidos monocatenaria se emparejan por bases para formar una sección bicatenaria. En algunas realizaciones, sólo una parte de la ITR forma un bucle horquilla. En otras realizaciones, toda la ITR forma un bucle horquilla. En algunas realizaciones, la ITR conserva el Elemento de Unión a Rep (RBE) de la ITR de tipo salvaje de la que deriva. La conservación del RBE puede ser importante para la estabilidad de la ITR y para fines de fabricación.

[0071] El término "parvovirus", tal y como se utiliza en la presente memoria, abarca la familia *Parvoviridae*, incluyendo, pero sin limitarse a, los parvovirus de replicación autónoma y los Dependovirus. Los parvovirus autónomos incluyen, por ejemplo, miembros de los géneros *Bocavirus*, *Dependovirus*, *Erythrovirus*, *Amdovirus*, *Parvovirus*, *Densovirus*, *Iteravirus*, *Contravirus*, *Aveparvovirus*, *Copiparvovirus*, *Protoparvovirus*, *Tetraparvovirus*, *Ambidensovirus*, *Brevidensovirus*, *Hepandensovirus* y *Penstylidensovirus*. Los parvovirus autónomos ejemplares incluyen, pero no se limitan a, bocavirus humano 1 (HBoV1), parvovirus porcino, virus diminuto de los ratones, parvovirus canino, virus de enteritis de visón, parvovirus bovino, parvovirus del pollo, virus de la panleucopenia felina, parvovirus felino, parvovirus de oca (GPV), parvovirus H1, parvovirus del pato de muscovy, parvovirus de la serpiente y virus B19. Los expertos en la

técnica conocen otros parvovirus autónomos. Véase, p.ej., Fields et al. *Virology*, Vol. 2, Cap. 69 (4ª ed., Lippincott-Raven Publishers).

[0072] El término "no AAV", tal y como se utiliza en la presente memoria, abarca los ácidos nucleicos, las proteínas y los virus de la familia *Parvoviridae*, excluyendo cualquier virus adeno-asociado (AAV) de la familia *Parvoviridae*. Los "no AAV" incluyen, pero no se limitan a, los miembros de replicación autónoma de los géneros *Bocavirus*, *Dependovirus*, *Erythrovirus*, *Amdovirus*, *Parvovirus*, *Densovirus*, *Iteravirus*, *Contravirus*, *Aveparvovirus*, *Copiparvovirus*, *Protoparvovirus*, *Tetraparvovirus*, *Ambidensovirus*, *Brevidensovirus*, *Hepandensovirus* y *Penstyldensovirus*.

[0073] Tal y como se utiliza en la presente memoria, el término "virus adeno-asociado" (AAV), incluye pero no se limita a, AAV tipo 1, AAV tipo 2, AAV tipo 3 (incluyendo los tipos 3A y 3B), AAV tipo 4, AAV tipo 5, AAV tipo 6, AAV tipo 7, AAV tipo 8, AAV de tipo 9, AAV de tipo 10, AAV de tipo 11, AAV de tipo 12, AAV de tipo 13, AAV de serpiente, AAV aviar, AAV bovino, AAV canino, AAV equino, AAV ovino, AAV caprino, AAV de camarón, los serotipos y clados de AAV divulgados por Gao et al. (*J. Virol.* 78:6381 (2004)) y Moris et al. (*Virol.* 33:375 (2004)), y cualquier otro AAV conocido ahora o descubierto posteriormente. Véase, p. ej., FIELDS et al. *VIROLOGY*, volumen 2, capítulo 69 (4ª ed., Lippincott-Raven Publishers).

[0074] El término "derivado de", tal y como se utiliza en la presente memoria, se refiere a un componente que se aísla o se prepara usando una molécula u organismo especificado, o una información (p. ej., una secuencia de aminoácidos o de ácidos nucleicos) de la molécula u organismo especificado. Por ejemplo, una secuencia de ácido nucleico (p. ej., ITR) que se deriva de una segunda secuencia de ácido nucleico (p. ej., ITR) puede incluir una secuencia de nucleótidos que es idéntica o sustancialmente similar a la secuencia de nucleótidos de la segunda secuencia de ácido nucleico. En el caso de nucleótidos o polipéptidos, las especies derivadas pueden obtenerse, por ejemplo, mediante mutagénesis natural, mutagénesis artificial dirigida o mutagénesis artificial aleatoria. La mutagénesis utilizada para derivar nucleótidos o polipéptidos puede ser intencionalmente dirigida o intencionalmente aleatoria, o una mezcla de ambas. La mutagénesis de un nucleótido o polipéptido para crear un nucleótido o polipéptido diferente derivado del primero puede ser un evento aleatorio (p. ej., causado por la infidelidad de la polimerasa) y la identificación del nucleótido o polipéptido derivado puede hacerse mediante métodos de cribado apropiados, p. ej., como se discute en la presente memoria. La mutagénesis de un polipéptido implica típicamente la manipulación del polinucleótido que codifica el polipéptido.

[0075] Un vector o molécula de ácido nucleico "libre de cápside" o "sin cápside" se refiere a una construcción vectorial sin cápside.

[0076] Tal y como se utiliza en la presente memoria, una "región codificante" o "secuencia codificante" es una porción de polinucleótido que consiste en codones traducibles a aminoácidos. Aunque un "codón de parada" (TAG, TGA o TAA) no suele traducirse en un

aminoácido, puede considerarse parte de una región codificante, pero cualquier secuencia flanqueante, por ejemplo, promotores, sitios de unión a ribosomas, terminadores transcripcionales, intrones y similares, no forman parte de una región codificante. Los límites de una región codificante suelen estar determinados por un codón de inicio en el extremo 5', que codifica el extremo amino del polipéptido resultante, y un codón de parada de la traducción en el extremo 3', que codifica el extremo carboxilo del polipéptido resultante. Dos o más regiones codificantes pueden estar presentes en una única construcción polinucleotídica, p. ej., en un único vector, o en construcciones polinucleotídicas separadas, p. ej., en vectores separados (diferentes). De ello se deduce que un único vector puede contener una sola región codificante, o comprender dos o más regiones codificantes.

[0077] Ciertas proteínas secretadas por células de mamíferos están asociadas a un péptido señal secretor que se escinde de la proteína madura una vez que se ha iniciado la exportación de la cadena proteica en crecimiento a través del retículo endoplásmico rugoso. Los expertos en la técnica saben que los péptidos señal se fusionan generalmente con el extremo N del polipéptido y se escinden del polipéptido completo o "de longitud completa" para producir una forma secretada o "madura" del polipéptido. En ciertas realizaciones, un péptido señal nativo o un derivado funcional de esa secuencia que conserva la capacidad de dirigir la secreción del polipéptido que se asocia operativamente con él. Alternativamente, puede utilizarse un péptido señal heterólogo de mamífero, p. ej., un péptido señal del activador del plasminógeno tisular (TPA) humano o de la β -glucuronidasa de ratón, o un derivado funcional del mismo.

[0078] El término "aguas abajo" se refiere a una secuencia de nucleótidos que se encuentra en 3' de una secuencia de nucleótidos de referencia. En ciertas realizaciones, las secuencias de nucleótidos aguas abajo se refieren a las secuencias que siguen al punto de inicio de la transcripción. Por ejemplo, el codón de iniciación de la traducción de un gen está situado aguas abajo del punto de inicio de la transcripción.

[0079] El término "aguas arriba" se refiere a una secuencia de nucleótidos que se encuentra en 5' de una secuencia de nucleótidos de referencia. En ciertas realizaciones, las secuencias de nucleótidos aguas arriba se refieren a secuencias que están situadas en el lado 5' de una región codificante o punto de inicio de la transcripción. Por ejemplo, la mayoría de los promotores están situados aguas arriba del sitio de inicio de la transcripción.

[0080] Tal y como se utiliza en la presente memoria, el término "casete genético" significa una secuencia de ADN capaz de dirigir la expresión de una secuencia de polinucleótidos particular en una célula huésped apropiada, que comprende un promotor unido operativamente a una secuencia de polinucleótidos de interés. Un casete genético puede abarcar secuencias de nucleótidos situadas aguas arriba (secuencias no codificantes en 5'), dentro o aguas abajo (secuencias no codificantes en 3') de una región codificante, y que influyen en la transcripción, el procesamiento del ARN, la estabilidad o la traducción de la región codificante asociada. Si una región codificante está destinada a expresarse en una célula eucariota, una señal de

poliadenilación y una secuencia de terminación de la transcripción se situarán normalmente en 3' de la secuencia codificante. En algunas realizaciones, el casete genético comprende un polinucleótido que codifica un producto génico. En algunas realizaciones, el casete genético comprende un polinucleótido que codifica un miARN. En algunas realizaciones, el casete genético comprende una secuencia de polinucleótidos heteróloga.

[0081] Un polinucleótido que codifica un producto, p. ej., un miARN o un producto génico (p. ej., un polipéptido como una proteína terapéutica), puede incluir un promotor y/u otras secuencias de control de la expresión (p. ej., de la transcripción o la traducción) asociadas operativamente a una o más regiones codificantes. En una asociación operativa, una región codificante para un producto génico, p. ej., un polipéptido, se asocia con una o más regiones reguladoras de tal manera que la expresión del producto génico queda bajo la influencia o el control de la(s) región(es) reguladora(s). Por ejemplo, una región codificante y un promotor están "asociados operativamente" si la inducción de la función del promotor da lugar a la transcripción del ARNm que codifica el producto génico codificado por la región codificante, y si la naturaleza de la unión entre el promotor y la región codificante no interfiere con la capacidad del promotor para dirigir la expresión del producto génico ni interfiere con la capacidad del molde de ADN para ser transcrito. Otras secuencias de control de la expresión, además de un promotor, por ejemplo, potenciadores, operadores, represores y señales de terminación de la transcripción, también pueden asociarse operativamente con una región codificante para dirigir la expresión del producto génico.

[0082] Las "secuencias de control de la expresión" se refieren a las secuencias de nucleótidos reguladoras, como promotores, potenciadores, terminadores y similares, que permiten la expresión de una secuencia codificante en una célula huésped. Las secuencias de control de la expresión generalmente abarcan cualquier secuencia de nucleótidos reguladora que facilita la transcripción y traducción eficientes del ácido nucleico codificante al que unido operativamente. Los ejemplos no limitantes de secuencias de control de la expresión incluyen promotores, potenciadores, secuencias líderes de traducción, intrones, secuencias de reconocimiento de poliadenilación, sitios de procesamiento de ARN, sitios de unión de efectores o estructuras de bucle-tallo. Los expertos en la técnica conocen una variedad de secuencias de control de la expresión. Éstas incluyen, sin limitación, secuencias de control de la expresión que funcionan en células de vertebrados, tales como, pero no limitadas a, segmentos promotores y potenciadores de citomegalovirus (el promotor temprano inmediato, junto con el intrón-A), el virus del simio 40 (el promotor temprano), y retrovirus (como el virus del sarcoma de Rous). Otras secuencias de control de la expresión incluyen las derivadas de genes de vertebrados como la actina, la proteína de choque térmico, la hormona de crecimiento bovina y la β -globina de conejo, así como otras secuencias capaces de controlar la expresión de genes en células eucariotas. Otras secuencias de control de la expresión adecuadas incluyen los promotores y potenciadores específicos de tejido, así como los promotores inducibles por linfoquinas (p. ej.,

promotores inducibles por interferones o interleuquinas). Otras secuencias de control de la expresión incluyen secuencias intrónicas, elementos reguladores postranscripcionales y señales de poliadenilación. Otras secuencias ejemplares de control de la expresión se discuten en otra parte de la presente divulgación.

[0083] De forma similar, los expertos en la técnica conocen una variedad de elementos de control de la traducción. Estos incluyen, pero no se limitan a, los sitios de unión al ribosoma, los codones de inicio y terminación de la traducción, y los elementos derivados de los picornavirus (en particular, un sitio de entrada al ribosoma interno, o IRES).

[0084] El término "expresión", tal como se utiliza en la presente memoria, se refiere a un proceso por el cual un polinucleótido produce un producto génico, por ejemplo, un ARN o un polipéptido. Incluye, sin limitación, la transcripción del polinucleótido en ARN mensajero (ARNm), ARN de transferencia (ARNt), ARN de horquilla pequeño (ARNsh), ARN de interferencia pequeño (ARNsi) o cualquier otro producto de ARN, y la traducción de un ARNm en un polipéptido. La expresión produce un "producto génico". Tal y como se utiliza en la presente memoria, un producto génico puede ser un ácido nucleico, p. ej., un ARN mensajero producido por la transcripción de un gen, o un polipéptido que se traduce a partir de un transcrito. Los productos génicos descritos en la presente memoria incluyen además ácidos nucleicos con modificaciones postranscripcionales, p. ej., poliadenilación o corte y empalme, o polipéptidos con modificaciones posteriores a la traducción, p. ej., metilación, glicosilación, adición de lípidos, asociación con otras subunidades proteicas o escisión proteolítica. El término "rendimiento", tal y como se utiliza en la presente memoria, se refiere a la cantidad de un polipéptido producido por la expresión de un gen.

[0085] Un "vector" se refiere a cualquier vehículo para la clonación y/o transferencia de un ácido nucleico a una célula huésped. Un vector puede ser un replicón al que se puede unir otro segmento de ácido nucleico para provocar la replicación del segmento unido. Un "replicón" se refiere a cualquier elemento genético (p. ej., un plásmido, fago, cósmido, cromosoma, virus) que funciona como una unidad autónoma de replicación *in vivo*, es decir, capaz de replicarse bajo su propio control. El término "vector" incluye vehículos para introducir el ácido nucleico en una célula *in vitro*, *ex vivo* o *in vivo*. Se conoce y se utiliza un gran número de vectores en la técnica, incluyendo, por ejemplo, plásmidos, virus eucariotas modificados o virus bacterianos modificados. La inserción de un polinucleótido en un vector adecuado puede llevarse a cabo ligando los fragmentos de polinucleótidos apropiados en un vector elegido que tenga extremos cohesivos complementarios.

[0086] Los vectores pueden ser diseñados para codificar marcadores o informadores seleccionables que permiten la selección o identificación de las células que han incorporado el vector. La expresión de marcadores o informadores seleccionables permite la identificación y/o selección de las células huésped que incorporan y expresan otras regiones codificantes contenidas en el vector. Los ejemplos de genes marcadores seleccionables conocidos y

utilizados en la técnica incluyen: genes que proporcionan resistencia a la ampicilina, estreptomicina, gentamicina, kanamicina, higromicina, herbicida bialafos, sulfonamida, y similares; y genes que se utilizan como marcadores fenotípicos, es decir, genes reguladores de antocianina, gen de isopentanol transferasa, y similares. Los ejemplos de informadores conocidos y utilizados en la técnica incluyen: luciferasa (Luc), proteína verde fluorescente (GFP), cloranfenicol acetiltransferasa (CAT), β -galactosidasa (LacZ), β -glucuronidasa (Gus), y similares. Los marcadores seleccionables también pueden considerarse como informadores.

[0087] El término "célula huésped", tal y como se utiliza en la presente memoria, se refiere, por ejemplo, a microorganismos, células de levadura, células de insectos y células de mamíferos, que pueden ser, o han sido, utilizadas como receptores de ADNss o vectores. El término incluye la progenie de la célula original que ha sido transducida. Así, una "célula huésped", tal y como se utiliza en la presente memoria, se refiere generalmente a una célula que ha sido transducida con una secuencia de ADN exógena. Se entiende que la progenie de una sola célula parental puede no ser necesariamente completamente idéntica en morfología o en complemento de ADN genómico o total al progenitor original, debido a una mutación natural, accidental o deliberada. En algunas realizaciones, la célula huésped puede ser una célula huésped *in vitro*.

[0088] El término "marcador seleccionable" se refiere a un factor de identificación, normalmente un gen de resistencia a un antibiótico o a un producto químico, que puede seleccionarse en función del efecto del gen marcador, es decir, resistencia a un antibiótico, resistencia a un herbicida, marcadores colorimétricos, enzimas, marcadores fluorescentes y similares, en los que el efecto se utiliza para seguir la herencia de un ácido nucleico de interés y/o para identificar una célula u organismo que ha heredado el ácido nucleico de interés. Los ejemplos de genes marcadores seleccionables conocidos y utilizados en la técnica incluyen: genes que proporcionan resistencia a la ampicilina, estreptomicina, gentamicina, kanamicina, higromicina, herbicida bialafos, sulfonamida, y similares; y genes que se utilizan como marcadores fenotípicos, es decir, genes reguladores de antocianina, gen de isopentanol transferasa, y similares.

[0089] El término "gen informador" se refiere a un ácido nucleico que codifica un factor de identificación que puede ser identificado sobre la base del efecto del gen informador, en donde el efecto se utiliza para rastrear la herencia de un ácido nucleico de interés, para identificar una célula u organismo que ha heredado el ácido nucleico de interés, y/o para medir la inducción de la expresión génica o la transcripción. Los ejemplos de genes informadores conocidos y utilizados en la técnica incluyen: luciferasa (Luc), proteína verde fluorescente (GFP), cloranfenicol acetiltransferasa (CAT), β -galactosidasa (LacZ), β -glucuronidasa (Gus), y similares. Los genes marcadores seleccionables también pueden considerarse genes informadores.

[0090] "Promotor" y "secuencia promotora" se utilizan indistintamente y se refieren a una secuencia de ADN capaz de controlar la expresión de una secuencia codificante o un ARN funcional. En general, una secuencia codificante está situada en 3' de una secuencia promotora.

Los promotores pueden derivarse en su totalidad de un gen nativo, o estar compuestos por diferentes elementos derivados de diferentes promotores encontrados en la naturaleza, o incluso comprender segmentos de ADN sintético. Los expertos en la técnica entienden que diferentes promotores pueden dirigir la expresión de un gen en diferentes tejidos o tipos de células, o en diferentes etapas del desarrollo, o en respuesta a diferentes condiciones ambientales o fisiológicas. Los promotores que hacen que un gen se exprese en la mayoría de los tipos de células en la mayoría de los momentos se denominan comúnmente "promotores constitutivos". Los promotores que hacen que un gen se exprese en un tipo de célula específico suelen denominarse "promotores específicos de célula" o "promotores específicos de tejido". Los promotores que hacen que un gen se exprese en una etapa específica del desarrollo o de la diferenciación celular se denominan comúnmente "promotores específicos del desarrollo" o "promotores específicos de la diferenciación celular." Los promotores que se inducen y hacen que un gen se exprese tras la exposición o el tratamiento de la célula con un agente, molécula biológica, producto químico, ligando, luz o similar que induce el promotor se denominan comúnmente "promotores inducibles" o "promotores regulables". Se reconoce además que, dado que en la mayoría de los casos los límites exactos de las secuencias reguladoras no se han definido completamente, fragmentos de ADN de diferentes longitudes pueden tener una actividad promotora idéntica. En otra parte de la presente divulgación se discuten otros promotores ejemplares.

[0091] La secuencia promotora suele estar limitada en su extremo 3' por el sitio de inicio de la transcripción y se extiende aguas arriba (dirección 5') para incluir el número mínimo de bases o elementos necesarios para iniciar la transcripción a niveles detectables por encima del fondo. Dentro de la secuencia promotora se encontrará un sitio de inicio de la transcripción (convenientemente definido, por ejemplo, mediante el mapeo con la nucleasa S1), así como dominios de unión a proteínas (secuencias consenso) responsables de la unión de la ARN polimerasa.

[0092] En algunas realizaciones, la molécula de ácido nucleico comprende un promotor específico de tejido. En ciertas realizaciones, el promotor específico de tejido dirige la expresión de la proteína terapéutica en el hígado, en los hepatocitos, y/o en las células endoteliales. En una realización particular, el promotor comprende un promotor TTP. En una realización particular, el promotor comprende un promotor mTTR. En una realización particular, el promotor comprende un promotor A1AT.

[0093] El término "plásmido" se refiere a un elemento extracromosómico que a menudo porta un gen que no forma parte del metabolismo central de la célula, y que suele estar en forma de moléculas circulares de ADN bicatenario. Dichos elementos pueden ser secuencias de replicación autónoma, secuencias que se integran en el genoma, fagos o secuencias de nucleótidos, lineales, circulares o superenrolladas, de ADN o ARN mono o bicatenario, derivadas de cualquier fuente, en las que varias secuencias de nucleótidos se han unido o recombinado

en una construcción única, que es capaz de introducir en una célula un fragmento promotor y una secuencia de ADN para un producto génico seleccionado junto con la secuencia no traducida en 3' apropiada.

[0094] Los vectores virales eucariotas que pueden utilizarse incluyen, pero no están limitados a, vectores de adenovirus, vectores de retrovirus, vectores de virus adeno-asociados, poxvirus, p. ej., vectores de virus vaccinia, vectores de baculovirus o vectores de herpesvirus. Los vectores no virales incluyen plásmidos, liposomas, lípidos cargados eléctricamente (citofectinas), complejos ADN-proteína y biopolímeros.

[0095] Un "vector de clonación" se refiere a un "replicón", que es una unidad de longitud de un ácido nucleico que se replica secuencialmente y que comprende un origen de replicación, como un plásmido, fago o cósmido, al que se puede unir otro segmento de ácido nucleico para provocar la replicación del segmento unido. Algunos vectores de clonación son capaces de replicarse en un tipo de célula, p. ej., bacterias, y de expresarse en otra, p. ej., células eucariotas. Los vectores de clonación comprenden típicamente una o más secuencias que pueden utilizarse para la selección de células que comprenden el vector y/o uno o más sitios de clonación múltiple para la inserción de secuencias de ácido nucleico de interés.

[0096] El término "vector de expresión" se refiere a un vehículo diseñado para permitir la expresión de una secuencia de ácido nucleico insertada tras su inserción en una célula huésped. La secuencia de ácido nucleico insertada se coloca en asociación operativa con regiones reguladoras como se ha descrito anteriormente.

[0097] Los vectores se introducen en las células huésped por métodos bien conocidos en la técnica, p. ej., transfección, electroporación, microinyección, transducción, fusión celular, dextrano DEAE, precipitación con fosfato de calcio, lipofección (fusión de lisosomas), uso de una pistola de genes o un transportador de vectores de ADN. "Cultivo", "cultivar" y "cultivando", tal y como se utilizan en la presente memoria, significa incubar células en condiciones *in vitro* que permiten el crecimiento o la división celular o mantener las células en un estado vivo. "Células cultivadas", tal y como se utiliza en la presente memoria, significa células que se propagan *in vitro*.

[0098] Tal y como se utiliza en la presente memoria, el término "polipéptido" pretende abarcar un "polipéptido" en singular, así como "polipéptidos" en plural, y se refiere a una molécula compuesta de monómeros (aminoácidos) unidos linealmente por enlaces amida (también conocidos como enlaces peptídicos). El término "polipéptido" se refiere a cualquier cadena o cadenas de dos o más aminoácidos, y no se refiere a una longitud específica del producto. Por lo tanto, los péptidos, dipéptidos, tripéptidos, oligopéptidos, "proteína", "cadena de aminoácidos" o cualquier otro término utilizado para referirse a una cadena o cadenas de dos o más aminoácidos, están incluidos en la definición de "polipéptido", y el término "polipéptido" puede utilizarse en lugar de, o de forma intercambiable con, cualquiera de estos términos. El término "polipéptido" también pretende referirse a los productos de las modificaciones posteriores a la

expresión del polipéptido, incluyendo sin limitación glicosilación, acetilación, fosforilación, amidación, derivación por grupos protectores/bloqueantes conocidos, escisión proteolítica o modificación por aminoácidos no naturales. Un polipéptido puede derivarse de una fuente biológica natural o producirse mediante tecnología recombinante, pero no se traduce necesariamente a partir de una secuencia de ácido nucleico designada. Puede generarse de cualquier manera, incluso mediante síntesis química.

[0099] El término "aminoácido" incluye alanina (Ala o A); arginina (Arg o R); asparagina (Asn o N); ácido aspártico (Asp o D); cisteína (Cys o C); glutamina (Gln o Q); ácido glutámico (Glu o E); glicina (Gly o G); histidina (His o H); isoleucina (Ile o I); leucina (Leu o L); lisina (Lys o K); metionina (Met o M); fenilalanina (Phe o F); prolina (Pro o P); serina (Ser o S); treonina (Thr o T); triptófano (Trp o W); tirosina (Tyr o Y); y valina (Val o V). Los aminoácidos no tradicionales también están dentro del alcance de la divulgación e incluyen norleucina, omitina, norvalina, homoserina y otros análogos de residuos de aminoácidos como los descritos en Ellman *et al.* Meth. Enzym. 202:301-336 (1991). Para generar dichos residuos de aminoácidos no naturales, pueden utilizarse los procedimientos de Noren *et al.* Science 244:182 (1989) y Ellman *et al.*, supra. Brevemente, estos procedimientos implican la activación química de un ARNt supresor con un residuo de aminoácido no natural seguido de la transcripción y traducción in vitro del ARN. La introducción del aminoácido no tradicional también puede lograrse utilizando químicas peptídicas conocidas en la técnica. Tal y como se utiliza en la presente memoria, el término "aminoácido polar" incluye aminoácidos que tienen carga neta cero, pero que tienen cargas parciales no nulas en diferentes porciones de sus cadenas laterales (p. ej., M, F, W, S, Y, N, Q, C). Estos aminoácidos pueden participar en interacciones hidrófobas e interacciones electrostáticas. Tal y como se utiliza en la presente memoria, el término "aminoácido cargado" incluye aminoácidos que pueden tener una carga neta distinta de cero en sus cadenas laterales (p. ej., R, K, H, E, D). Estos aminoácidos pueden participar en interacciones hidrófobas e interacciones electrostáticas.

[0100] También se incluyen en la presente divulgación fragmentos o variantes de polipéptidos, y cualquier combinación de los mismos. El término "fragmento" o "variante", al referirse a los dominios de unión de polipéptidos o a las moléculas de unión de la presente divulgación, incluye cualquier polipéptido que conserve al menos algunas de las propiedades (p. ej., la afinidad de unión al FcRn para un dominio de unión al FcRn o una variante del Fc, la actividad de coagulación para una variante del FVIII, o la actividad de unión al FVIII para el fragmento del VWF) del polipéptido de referencia. Los fragmentos de polipéptidos incluyen fragmentos proteolíticos, así como fragmentos de delección, además de los fragmentos de anticuerpos específicos discutidos en otra parte de la presente memoria, pero no incluyen el polipéptido natural de longitud completa (o polipéptido maduro). Las variantes de los dominios de unión a polipéptidos o de las moléculas de unión de la presente divulgación incluyen fragmentos como los descritos anteriormente, y también polipéptidos con secuencias de aminoácidos alteradas

debido a sustituciones, deleciones o inserciones de aminoácidos. Las variantes pueden ser naturales o no naturales. Las variantes no naturales pueden producirse mediante técnicas de mutagénesis conocidas en la técnica. Los polipéptidos variantes pueden incluir sustituciones, deleciones o adiciones de aminoácidos conservativas o no conservativas.

[0101] Una "sustitución de aminoácidos conservativa" es aquella en la que el residuo de aminoácido se reemplaza por un residuo de aminoácido que tiene una cadena lateral similar. En la técnica, se han definido familias de residuos de aminoácidos que tienen cadenas laterales similares, incluyendo cadenas laterales básicas (p. ej., lisina, arginina, histidina), cadenas laterales ácidas (p. ej., ácido aspártico, ácido glutámico), cadenas laterales polares no cargadas (p. ej., glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (p. ej., alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales con ramificación beta (p. ej., treonina, valina, isoleucina) y cadenas laterales aromáticas (p. ej., tirosina, fenilalanina, triptófano, histidina). Así, si un aminoácido de un polipéptido se reemplaza por otro aminoácido de la misma familia de cadenas laterales, la sustitución se considera conservativa. En otra realización, una cadena de aminoácidos puede ser reemplazada de forma conservativa por una cadena estructuralmente similar que difiere en el orden y/o la composición de los miembros de la familia de cadenas laterales.

[0102] El término "porcentaje de identidad", tal y como se conoce en la técnica, es una relación entre dos o más secuencias polipeptídicas o dos o más secuencias polinucleotídicas, tal y como se determina al comparar las secuencias. En la técnica, "identidad" también significa el grado de relación de la secuencia entre las secuencias polipeptídicas o polinucleotídicas, según sea el caso, como se determina por la coincidencia entre las cadenas de dichas secuencias. La "identidad" puede calcularse fácilmente mediante métodos conocidos, incluidos, pero no limitados a, los descritos en: *Computational Molecular Biology* (Lesk, A. M., ed.) Oxford University Press, Nueva York (1988); *Biocomputing: Informatics and Genome Projects* (Smith, D. W., ed.) Academic Press, Nueva York (1993); *Computer Analysis of Sequence Data, Parte I* (Griffin, A. M., y Griffin, H. G., eds.) Humana Press, Nueva Jersey (1994); *Sequence Analysis in Molecular Biology* (von Heinje, G., ed.) Academic Press (1987); y *Sequence Analysis Primer* (Gribskov, M. y Devereux, J., eds.) Stockton Press, Nueva York (1991). Los métodos preferidos para determinar la identidad están diseñados para ofrecer la mejor coincidencia entre las secuencias ensayadas. Los métodos para determinar la identidad están codificados en programas informáticos de acceso público. Los alineamientos de secuencias y los cálculos del porcentaje de identidad pueden realizarse utilizando software de análisis de secuencias como el programa Megalign del paquete informático bioinformático LASERGENE (DNASTAR Inc., Madison, WI), el paquete programas GCG (Wisconsin Package Versión 9.0, Genetics Computer Group (GCG), Madison, WI), BLASTP, BLASTN, BLASTX (Altschul *et al.*, *J. Mol. Biol.* 215:403 (1990)), y DNASTAR (DNASTAR, Inc. 1228 S. Park St. Madison, WI 53715 EE. UU.). En el contexto de esta solicitud, se entenderá que cuando se utilice un software de análisis de

secuencias para el análisis, los resultados del mismo se basarán en los "valores por defecto" del programa al que se haga referencia, a menos que se especifique lo contrario. Tal y como se utiliza en la presente memoria, "valores por defecto" significará cualquier conjunto de valores o parámetros que se cargan originalmente con el software cuando se inicializa por primera vez. Para determinar el porcentaje de identidad entre una secuencia de consulta (p. ej., una secuencia de ácido nucleico) y una secuencia de referencia, sólo se utilizan los nucleótidos de la secuencia de consulta que coinciden con los nucleótidos de la secuencia de referencia para calcular el porcentaje de identidad. Así, al determinar el porcentaje de identidad entre una secuencia de consulta o una porción designada de la misma (p. ej., los nucleótidos 1-522) y una secuencia de referencia, el porcentaje de identidad se calculará dividiendo el número de nucleótidos coincidentes por el número total de nucleótidos de la secuencia de consulta completa.

[0103] Tal y como se utiliza en la presente memoria, los nucleótidos correspondientes a los nucleótidos de una secuencia particular de la divulgación se identifican mediante la alineación de la secuencia de la divulgación para maximizar la identidad con una secuencia de referencia. El número utilizado para identificar un aminoácido equivalente en una secuencia de referencia se basa en el número utilizado para identificar el aminoácido correspondiente en la secuencia de la divulgación.

[0104] "Tratar", "tratamiento", "tratando", tal y como se utiliza en la presente memoria, se refiere, p. ej., a la reducción de la gravedad de una enfermedad o afección; a la reducción de la duración del curso de una enfermedad; a la mejoría de uno o más síntomas asociados a una enfermedad o afección; a la provisión de efectos beneficiosos a un sujeto con una enfermedad o afección, sin necesariamente curar la enfermedad o afección, o a la profilaxis de uno o más síntomas asociados a una enfermedad o afección.

[0105] "Administrar", tal y como se utiliza en la presente memoria, significa proporcionar una molécula de ácido nucleico farmacéuticamente aceptable, un polipéptido expresado a partir de ella o un vector que comprende la molécula de ácido nucleico de la divulgación a un sujeto a través de una vía farmacéuticamente aceptable. Las vías de administración pueden ser intravenosa, p. ej., inyección intravenosa e infusión intravenosa. Otras vías de administración incluyen, p. ej., la administración subcutánea, intramuscular, oral, nasal y pulmonar. Las moléculas de ácido nucleico, los polipéptidos y los vectores pueden administrarse como parte de una composición farmacéutica que comprende al menos un excipiente.

[0106] "Nanopartícula lipídica", tal y como se utiliza en la presente memoria, se refiere a una partícula que tiene al menos una dimensión en la escala nanométrica (p. ej., de 1nm a 1.000nm) que comprende uno o más lípidos catiónicos. En algunas realizaciones, la nanopartícula lipídica se incluye en una formulación que puede utilizarse para suministrar un agente activo o terapéutico, como un ácido nucleico (p. ej., ARNm), a un sitio diana asociado (p. ej., célula, tejido, órgano, tumor, etc.). En algunas realizaciones, la nanopartícula lipídica divulgada en la presente

memoria comprende un ácido nucleico. Dichas nanopartículas lipídicas comprenden típicamente uno o más excipientes seleccionados entre lípidos neutros, lípidos cargados, esteroides y lípidos conjugados con polímeros. En algunas realizaciones, un agente activo o terapéutico, como un ácido nucleico, puede estar encapsulado en la porción lipídica de la nanopartícula lipídica, o en el espacio acuoso encapsulado por una parte o la totalidad de la porción lipídica de la nanopartícula lipídica, protegiéndolo así de la degradación enzimática, o de otros efectos indeseables desencadenados por los mecanismos del organismo o célula huésped, como una respuesta inmunitaria adversa.

[0107] El término "farmacéuticamente aceptable", tal y como se utiliza en la presente memoria, se refiere a entidades moleculares y composiciones que son fisiológicamente tolerables y que no suelen producir toxicidad o una reacción alérgica o inapropiada similar, como molestias gástricas, mareos y similares, cuando se administran a un ser humano. Opcionalmente, tal y como se utiliza en la presente memoria, el término "farmacéuticamente aceptable" significa que ha sido aprobado por una agencia reguladora del gobierno federal o estatal o que figura en la Farmacopea de los EE. UU. o en otra farmacopea generalmente reconocida para su uso en animales, y más concretamente en seres humanos.

[0108] Tal y como se utiliza en la presente memoria, la frase "sujeto que lo necesita" incluye a sujetos, como los mamíferos, que se beneficiarían de la administración de una molécula de ácido nucleico, polipéptido o vector de la divulgación. En algunas realizaciones, el sujeto es un sujeto humano. En algunas realizaciones, los sujetos son individuos con hemofilia. El sujeto puede ser un adulto o un menor (p. ej., menor de 12 años).

[0109] Tal y como se utiliza en la presente memoria, el término "proteína terapéutica" se refiere a cualquier polipéptido conocido en la técnica que pueda administrarse a un sujeto. En algunas realizaciones, la proteína terapéutica comprende una proteína seleccionada entre un factor de coagulación, un factor de crecimiento, un anticuerpo, un fragmento funcional del mismo, o una combinación de los mismos. Tal y como se utiliza en la presente memoria, el término "factor de coagulación" se refiere a moléculas, o análogos de las mismas, producidas de forma natural o recombinante, que evitan o disminuyen la duración de un episodio de hemorragia en un sujeto. En otras palabras, se trata de moléculas que tienen una actividad procoagulante, es decir, que son responsables de la conversión del fibrinógeno en una malla de fibrina insoluble que provoca la coagulación de la sangre. El "factor de coagulación", tal y como se utiliza en la presente memoria, incluye un factor de coagulación activado, su zimógeno o un factor de coagulación activable. Un "factor de coagulación activable" es un factor de coagulación en una forma inactiva (p. ej., en su forma zimógena) que es capaz de convertirse en una forma activa. El término "factor de coagulación" incluye, pero no se limita a, factor I (FI), factor II (FII), factor III (FIII), factor IV (FIV), factor V (FV), factor VI (FVI), factor VII (FVII), factor VIII (FVIII), factor IX (FIX), factor X (FX), factor XI (FXI), factor XII (FXII), factor XIII (FXIII), factor de Von Willebrand (VWF), precalicreína, cininógeno de alto peso molecular, fibronectina, antitrombina III, cofactor II de la

heparina, proteína C, proteína S, proteína Z, inhibidor de la proteasa relacionada con la proteína Z (ZPI), plasminógeno, alfa-2-antiplasmina, activador del plasminógeno tisular (tPA), uroquinasa, inhibidor del activador del plasminógeno 1 (PAI-1), inhibidor del activador del plasminógeno 2 (PAI2), zimógenos de los mismos, formas activadas de los mismos, o cualquier combinación de los mismos.

[0110] "Actividad de coagulación", tal y como se utiliza en la presente memoria, significa la capacidad de participar en una cascada de reacciones bioquímicas que culmina en la formación de un coágulo de fibrina y/o reduce la gravedad, la duración o la frecuencia de la hemorragia o del episodio de sangrado.

[0111] Un "factor de crecimiento", tal y como se utiliza en la presente memoria, incluye cualquier factor de crecimiento conocido en la técnica, incluidas las citoquinas y las hormonas.

[0112] Tal y como se utilizan en la presente memoria, los términos "heterólogo" o "exógeno" se refieren a moléculas que no se encuentran normalmente en un contexto determinado, p. ej., en una célula o en un polipéptido. Por ejemplo, una molécula exógena o heteróloga puede introducirse en una célula y estar presente sólo después de manipular la célula, p. ej., mediante transfección u otras formas de ingeniería genética, o una secuencia de aminoácidos heteróloga puede estar presente en una proteína en la que no se encuentra de forma natural.

[0113] Una "secuencia de nucleótidos de referencia", cuando se utiliza en la presente memoria como comparación con una secuencia de nucleótidos de la divulgación, es una secuencia de polinucleótidos esencialmente idéntica a la secuencia de nucleótidos de la divulgación, excepto que las porciones correspondientes a la secuencia del FVIII no están optimizadas. En algunas realizaciones, la secuencia de nucleótidos de referencia para una molécula de ácido nucleico divulgada en la presente memoria es la SEQ ID NO: 32.

[0114] Tal y como se utiliza en la presente memoria, el término "optimizado", con respecto a las secuencias de nucleótidos, se refiere a una secuencia de polinucleótidos que codifica un polipéptido, en donde la secuencia de polinucleótidos ha sido mutada para mejorar una propiedad de dicha secuencia de polinucleótidos. En algunas realizaciones, la optimización se hace para aumentar los niveles de transcripción, aumentar los niveles de traducción, aumentar los niveles de ARNm en estado estacionario, aumentar o disminuir la unión de proteínas reguladoras como los factores de transcripción generales, aumentar o disminuir el corte y empalme, o aumentar el rendimiento del polipéptido producido por la secuencia de polinucleótidos. Los ejemplos de cambios que pueden realizarse en una secuencia de polinucleótidos para optimizarla incluyen optimización de codones, optimización del contenido de G/C, eliminación de secuencias repetidas, eliminación de elementos ricos en AT, eliminación de sitios de corte y empalme crípticos, eliminación de elementos que actúan en cis que reprimen la transcripción o la traducción, adición o eliminación de secuencias poli T o poli A, adición de secuencias alrededor del sitio de inicio de la transcripción que mejoran la transcripción, como las secuencias de consenso de Kozak, eliminación de secuencias que podrían formar

estructuras de tallo bucle, eliminación de secuencias desestabilizadoras, eliminación de restos CpG, y dos o más combinaciones de las mismas.

Moléculas de Ácido nucleico

[0115] Ciertos aspectos de la presente divulgación se dirigen a superar las deficiencias de los vectores AAV para la terapia génica. En particular, ciertos aspectos de la presente divulgación se dirigen a una molécula de ácido nucleico que comprende una primera ITR, una segunda ITR y un casete genético. En algunas realizaciones, el casete genético codifica una proteína terapéutica y/o un miRNA. En algunas realizaciones, la primera ITR y la segunda ITR flanquean un casete genético que comprende una secuencia de polinucleótidos heteróloga. En algunas realizaciones, la molécula de ácido nucleico no comprende un gen que codifique una proteína de la cápside, una proteína de replicación y/o una proteína de ensamblaje. En algunas realizaciones, el casete genético codifica una proteína terapéutica. En algunas realizaciones, la proteína terapéutica comprende un factor de coagulación. En algunas realizaciones, el casete genético codifica un miRNA. En ciertas realizaciones, el casete genético se sitúa entre la primera ITR y la segunda ITR. En algunas realizaciones, la molécula de ácido nucleico comprende además una o más regiones no codificantes. En ciertas realizaciones, la una o más región no codificante comprende una secuencia promotora, un intrón, un elemento regulador postranscripcional, una secuencia poli(A) 3'UTR, o cualquier combinación de las mismas.

[0116] En una realización, el casete genético es un ácido nucleico monocatenario. En otra realización, el casete genético es un ácido nucleico bicatenario. En otra realización, el casete genético es un ácido nucleico bicatenario con extremos cerrados (ADNce).

[0117] En algunas realizaciones, la molécula de ácido nucleico comprende: (a) una primera ITR que es una ITR derivada de un miembro de la familia *Parvoviridae* que no es AAV (p. ej., una ITR de HBoV1); (b) una secuencia promotora específica de tejido, p. ej., el promotor TTP o TTR; (c) un intrón, p. ej., un intrón sintético; (d) un nucleótido que codifica un miRNA o una proteína terapéutica, p. ej., un factor de coagulación; (e) un elemento regulador postranscripcional, p. ej., WPRE; (f) una secuencia de cola de poli(A) 3' UTR, p. ej., bGHpA; (g) una segunda ITR que es una ITR derivada de un miembro de la familia *Parvoviridae* que no es AAV (p. ej., una ITR de HBoV1). En algunas realizaciones, la molécula de ácido nucleico comprende: (a) una primera ITR que es una ITR derivada de un miembro de la familia *Parvoviridae* que no es AAV (p. ej., una ITR de HBoV1); (b) una secuencia promotora específica de tejido, p. ej., el promotor mTTR; (c) un intrón, p. ej., un intrón sintético; (d) un nucleótido que codifica un miRNA o una proteína terapéutica, p. ej., un factor de coagulación; (e) un elemento regulador postranscripcional, p. ej., WPRE; (f) una secuencia de cola de poli(A) 3'UTR, p. ej., bGHpA; (g) una segunda ITR que es una ITR derivada de un miembro de la familia *Parvoviridae* que no es un AAV (p. ej., una ITR de HBoV1). En algunas realizaciones, el promotor específico

del tejido es el promotor de la antitripsina alfa-1 (A1AT) humana. En algunas realizaciones, el promotor específico de tejido comprende la secuencia de nucleótidos de la SEQ ID NO: 36.

[0118] En algunas realizaciones, se divulgan en la presente memoria moléculas de ácido nucleico aisladas que comprenden un casete genético que comprende una secuencia de nucleótidos al menos aproximadamente un 75 % idéntica a la SEQ ID NO: 9. En algunas realizaciones, se divulga en la presente memoria una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene al menos un 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 100 % de identidad de secuencia con la SEQ ID NO: 9.

[0119] En algunas realizaciones, se divulgan en la presente memoria moléculas de ácido nucleico aisladas que comprenden un casete genético que comprende una secuencia de nucleótidos al menos aproximadamente un 75 % idéntica a la SEQ ID NO: 33. En algunas realizaciones, se divulga en la presente memoria una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene al menos un 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 100 % de identidad de secuencia con la SEQ ID NO: 33.

[0120] En algunas realizaciones, se divulgan en la presente memoria moléculas de ácido nucleico aisladas que comprenden un casete genético que comprende una secuencia de nucleótidos al menos aproximadamente un 75 % idéntica a la SEQ ID NO: 14. En algunas realizaciones, se divulga en la presente memoria una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene al menos un 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 100 % de identidad de secuencia con la SEQ ID NO: 14.

[0121] En otro aspecto, se divulga en la presente memoria una molécula de ácido nucleico aislada que comprende un casete genético que expresa un polipéptido de factor VIII (FVIII), en donde el casete genético comprende una secuencia de nucleótidos al menos un 85 % idéntica a la SEQ ID NO: 35. En algunas realizaciones, el casete genético comprende una secuencia de nucleótidos que es al menos un 90 % idéntica a la SEQ ID NO: 35. En algunas realizaciones, el casete genético comprende una secuencia de nucleótidos que es al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % o un 100 % idéntica a la SEQ ID NO: 35. En algunas realizaciones, la secuencia de nucleótidos es al menos un 50 % idéntica a la SEQ ID NO: 35.

[0122] También se divulga en la presente memoria una molécula de ácido nucleico aislada que comprende un casete genético que expresa un polipéptido de factor VIII (FVIII), en donde el casete genético comprende la secuencia de nucleótidos de la SEQ ID NO: 35.

[0123] En ciertas realizaciones, la molécula de ácido nucleico divulgada en la presente memoria comprende secuencias de ITR del bocavirus humano 1 (HBoV1). En ciertas realizaciones, la molécula de ácido nucleico divulgada en la presente memoria comprende una primera ITR que es al menos aproximadamente un 75 % idéntica a la SEQ ID NO: 1 o SEQ ID NO: 2.

A. Repeticiones Terminales Invertidas (ITR)

[0124] Ciertos aspectos de la presente divulgación se dirigen a una molécula de ácido nucleico que comprende una primera ITR, p. ej., una ITR 5', y una segunda ITR, p. ej., una ITR 3'. Típicamente, las ITR están implicadas en la replicación del ADN de parvovirus (p. ej., AAV) y en el rescate, o escisión, de plásmidos procariotas (Samulski et al., 1983, 1987; Senapathy et al., 1984; Gottlieb y Muzyczka, 1988). Además, las ITR parecen ser las secuencias mínimas requeridas para la integración proviral del AAV y para el empaquetamiento del ADN del AAV en los viriones (McLaughlin et al., 1988; Samulski et al., 1989). Estos elementos son esenciales para la multiplicación eficiente del genoma de un parvovirus. Se plantea la hipótesis de que los elementos definitorios mínimos indispensables para la función de ITR son un sitio de unión a Rep y un sitio de resolución terminal, además de una secuencia palindrómica variable que permite la formación de horquillas. Las regiones palindrómicas de nucleótidos normalmente funcionan juntas en *cis* como orígenes de replicación del ADN y como señales de empaquetamiento para el virus. Las secuencias complementarias en las ITR se pliegan en una estructura de horquilla durante la replicación del ADN. En algunas realizaciones, las ITR se pliegan en una estructura de horquilla en forma de T. En otras realizaciones, las ITR se pliegan en estructuras de horquilla que no tienen forma de T, p. ej., en una estructura de horquilla en forma de U. Los datos sugieren que las estructuras de horquilla en forma de T de las ITR de AAV pueden inhibir la expresión de un transgén flanqueado por las ITR. Véase, p. ej., Zhou et al. (2017) *Scientific Reports* 7:5432. Al utilizar una ITR que no forma estructuras de horquilla en forma de T, puede evitarse esta forma de inhibición. Por lo tanto, en ciertos aspectos, un polinucleótido que comprende una ITR no AAV tiene una expresión de transgenes mejorada en comparación con un polinucleótido que comprende una ITR de AAV que forma una horquilla en forma de T.

[0125] Tal y como se utiliza en la presente memoria, una "repetición terminal invertida" (o "ITR") se refiere a una subsecuencia de ácido nucleico situada en el extremo 5' o 3' de una secuencia de ácido nucleico monocatenaria, que comprende un conjunto de nucleótidos (secuencia inicial) seguido aguas abajo por su complemento inverso, es decir, una secuencia palindrómica. La secuencia intermedia de nucleótidos entre la secuencia inicial y el complemento inverso puede tener cualquier longitud, incluso cero. En una realización, la ITR útil para la presente divulgación comprende una o más "secuencias palindrómicas". Una ITR puede tener cualquier número de funciones. En algunas realizaciones, una ITR descrita en la presente memoria forma una estructura de horquilla. En algunas realizaciones, la ITR forma una estructura de horquilla en forma de T. En algunas realizaciones, la ITR forma una estructura de horquilla no en forma de T, p. ej., una estructura de horquilla en forma de U. En algunas realizaciones, la ITR promueve la supervivencia a largo plazo de la molécula de ácido nucleico en el núcleo de una célula. En algunas realizaciones, la ITR promueve la supervivencia permanente de la molécula de ácido nucleico en el núcleo de una célula (p. ej., durante toda la vida de la célula).

En algunas realizaciones, la ITR promueve la estabilidad de la molécula de ácido nucleico en el núcleo de una célula. En algunas realizaciones, la ITR promueve la retención de la molécula de ácido nucleico en el núcleo de una célula. En algunas realizaciones, la ITR promueve la persistencia de la molécula de ácido nucleico en el núcleo de una célula. En algunas realizaciones, la ITR inhibe o evita la degradación de la molécula de ácido nucleico en el núcleo de una célula.

[0126] Por lo tanto, una "ITR", tal y como se utiliza en la presente memoria, puede plegarse sobre sí misma y formar un segmento bicatenario. Por ejemplo, la secuencia GATCXXXXGATC comprende una secuencia inicial de GATC y su complemento (3'CTAG5') cuando se pliega para formar una doble hélice. En algunas realizaciones, la ITR comprende una secuencia palindrómica continua (p. ej., GATCGATC) entre la secuencia inicial y el complemento inverso. En algunas realizaciones, la ITR comprende una secuencia palindrómica interrumpida (p. ej., GATCXXXXGATC) entre la secuencia inicial y el complemento inverso. En algunas realizaciones, las secciones complementarias de la secuencia palindrómica continua o interrumpida interaccionan entre sí para formar una estructura de "bucle horquilla". Tal y como se utiliza en la presente memoria, una estructura de "bucle horquilla" resulta cuando al menos dos secuencias complementarias en una molécula de nucleótidos monocatenaria se emparejan por bases para formar una sección bicatenaria. En algunas realizaciones, sólo una parte de la ITR forma un bucle horquilla. En otras realizaciones, toda la ITR forma un bucle horquilla.

[0127] En algunas realizaciones, la ITR comprende una ITR de origen natural, p. ej., la ITR comprende toda o una parte de una ITR derivada de un miembro de la familia *Parvoviridae*. En algunas realizaciones, la ITR comprende una secuencia sintética. En una realización, la primera ITR o la segunda ITR comprende una secuencia sintética. En otra realización, cada una de la primera ITR y la segunda ITR comprende una secuencia sintética. En algunas realizaciones, la primera ITR o la segunda ITR comprende una secuencia natural. En otra realización, cada una de la primera ITR y la segunda ITR comprende una secuencia natural.

[0128] En algunas realizaciones, la ITR comprende o consiste en una porción de una ITR natural, p. ej., una ITR truncada. En algunas realizaciones, la ITR comprende o consiste en un fragmento de una ITR natural, en donde el fragmento comprende al menos aproximadamente 5 nucleótidos, al menos aproximadamente 10 nucleótidos, al menos aproximadamente 15 nucleótidos, al menos aproximadamente 20 nucleótidos, al menos aproximadamente 25 nucleótidos, al menos aproximadamente 30 nucleótidos, al menos aproximadamente 35 nucleótidos, al menos aproximadamente 40 nucleótidos, al menos aproximadamente 45 nucleótidos, al menos aproximadamente 50 nucleótidos, al menos aproximadamente 55 nucleótidos, al menos aproximadamente 60 nucleótidos, al menos aproximadamente 65 nucleótidos, al menos aproximadamente 70 nucleótidos, al menos aproximadamente 75 nucleótidos, al menos aproximadamente 80 nucleótidos, al menos aproximadamente 85 nucleótidos, al menos aproximadamente 90 nucleótidos, al menos aproximadamente 95

nucleótidos, al menos aproximadamente 100 nucleótidos, al menos aproximadamente 125 nucleótidos, al menos aproximadamente 150 nucleótidos, al menos aproximadamente 175 nucleótidos, al menos aproximadamente 200 nucleótidos, al menos aproximadamente 225 nucleótidos, al menos aproximadamente 250 nucleótidos, al menos aproximadamente 275 nucleótidos, al menos aproximadamente 300 nucleótidos, al menos aproximadamente 325 nucleótidos, al menos aproximadamente 350 nucleótidos, al menos aproximadamente 375 nucleótidos, al menos aproximadamente 400 nucleótidos, al menos aproximadamente 425 nucleótidos, al menos aproximadamente 450 nucleótidos, al menos aproximadamente 475 nucleótidos, al menos aproximadamente 500 nucleótidos, al menos aproximadamente 525 nucleótidos, al menos aproximadamente 550 nucleótidos, al menos aproximadamente 575 nucleótidos, o al menos aproximadamente 600 nucleótidos; en donde la ITR conserva una propiedad funcional de la ITR natural. En ciertas realizaciones, la ITR comprende o consiste en un fragmento de una ITR natural, en donde el fragmento comprende al menos aproximadamente 129 nucleótidos; en donde la ITR conserva una propiedad funcional de la ITR natural. En ciertas realizaciones, la ITR comprende o consiste en un fragmento de una ITR de origen natural, en donde el fragmento comprende al menos aproximadamente 102 nucleótidos; en donde la ITR conserva una propiedad funcional de la ITR natural. En algunas realizaciones, la ITR conserva el Elemento de Unión a Rep (RBE) de la ITR de tipo salvaje de la que se deriva. En algunas realizaciones, la ITR conserva al menos uno de los RBE de la ITR de tipo salvaje de la que deriva. En algunas realizaciones, la ITR conserva al menos uno de los RBE o una porción funcional del mismo de la ITR de tipo salvaje de la que se deriva. La conservación del RBE puede ser importante para la estabilidad de la ITR y para fines de fabricación.

[0129] En algunas realizaciones, la ITR comprende o consiste en una porción de una ITR natural, en donde el fragmento comprende al menos aproximadamente un 5 %, al menos aproximadamente un 10 %, al menos aproximadamente un 15 %, al menos aproximadamente un 20 %, al menos aproximadamente un 25 %, al menos aproximadamente un 30 %, al menos aproximadamente un 35 %, al menos aproximadamente un 40 %, al menos aproximadamente un 45 %, al menos aproximadamente un 50 %, al menos aproximadamente un 55 %, al menos aproximadamente un 60 %, al menos aproximadamente un 65 %, al menos aproximadamente un 70 %, al menos aproximadamente un 75 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 %, o al menos aproximadamente un 99 % de la longitud de la ITR natural; en donde el fragmento conserva una propiedad funcional de la ITR natural. En algunas realizaciones, la primera ITR y/o la segunda ITR se derivan de una ITR de HBoV1 de tipo salvaje. En algunas realizaciones, la primera ITR y/o la segunda ITR se derivan de una ITR de B19 de tipo salvaje. En algunas realizaciones, la primera ITR y/o la segunda ITR se derivan de una ITR de GPV de tipo salvaje.

[0130] En ciertas realizaciones, la ITR comprende o consiste en una secuencia que tiene una identidad de secuencia de al menos un 50 %, al menos un 51 %, al menos un 52 %, al menos un 53 %, al menos un 54 %, al menos un 55 %, al menos un 56 %, al menos un 57 %, al menos un 58 %, al menos un 59 %, al menos un 60 %, al menos un 61 %, al menos un 62 %, al menos un 63 %, al menos un 64 %, al menos un 65 %, al menos un 66 %, al menos un 67 %, al menos un 68 %, al menos un 69 %, al menos un 70 %, al menos un 71 %, al menos un 72 %, al menos un 73 %, al menos un 74 %, al menos un 75 %, al menos un 76 %, al menos un 77 %, al menos un 78 %, al menos un 79 %, al menos un 80 %, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % con una porción homóloga de una ITR natural, cuando está correctamente alineada; en donde la ITR conserva una propiedad funcional de la ITR natural. En otras realizaciones, la ITR comprende o consiste en una secuencia que tiene una identidad de secuencia de al menos un 90 % con una porción homóloga de una ITR natural, cuando está correctamente alineada; en donde la ITR conserva una propiedad funcional de la ITR natural. En algunas realizaciones, la ITR comprende o consiste en una secuencia que tiene una identidad de secuencia de al menos un 80 % con una porción homóloga de una ITR natural, cuando está correctamente alineada; en donde la ITR conserva una propiedad funcional de la ITR natural. En algunas realizaciones, la ITR comprende o consiste en una secuencia que tiene una identidad de secuencia de al menos un 70 % con una porción homóloga de una ITR natural, cuando está correctamente alineada; en donde la ITR conserva una propiedad funcional de la ITR natural. En algunas realizaciones, la ITR comprende o consiste en una secuencia que tiene una identidad de secuencia de al menos un 60 % con una porción homóloga de una ITR natural, cuando está correctamente alineada; en donde la ITR conserva una propiedad funcional de la ITR natural. En algunas realizaciones, la ITR comprende o consiste en una secuencia que tiene una identidad de secuencia de al menos un 50 % con una porción homóloga de una ITR natural, cuando está correctamente alineada; en donde la ITR conserva una propiedad funcional de la ITR natural.

[0131] En algunas realizaciones, la ITR comprende una ITR de un genoma de AAV. En algunas realizaciones, la ITR es una ITR de un genoma de AAV seleccionado entre AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 y AAV11, y cualquier combinación de los mismos. En algunas realizaciones, la ITR es una ITR de cualquier genoma de AAV conocido por los expertos en la técnica, incluyendo un aislado natural, p. ej., un aislado humano natural. En una realización particular, la ITR es una ITR del genoma de AAV2. En otra realización, la ITR es una secuencia sintética diseñada genéticamente para incluir en sus extremos 5' y 3' ITR derivadas de uno o más genomas de AAV.

[0132] En algunas realizaciones, la ITR no se deriva de un genoma de AAV (es decir, la ITR se deriva de un virus que no es AAV). En algunas realizaciones, la ITR es una ITR de un virus que no es AAV. En algunas realizaciones, la ITR es una ITR de un genoma no AAV de la familia viral *Parvoviridae* seleccionada de, pero no limitada a, el grupo que consiste en *Bocavirus*, *Dependovirus*, *Erythrovirus*, *Amdovirus*, *Parvovirus*, *Densovirus*, *Iteravirus*, *Contravirus*, *Aveparvovirus*, *Copiparvovirus*, *Protoparvovirus*, *Tetraparvovirus*, *Ambidensovirus*, *Brevidensovirus*, *Hepandensovirus*, *Penstyldensovirus* y cualquier combinación de los mismos. En ciertas realizaciones, la ITR se deriva del bocavirus humano 1 (HBoV1). En otra realización, la ITR se deriva del parvovirus eritrovirus B19 (virus humano). En otra realización, la ITR se deriva de una cepa de parvovirus del pato Muscovy (MDPV). En ciertas realizaciones, la cepa de MDPV está atenuada, p. ej., la cepa de MDPV FZ91-30. En otras realizaciones, la cepa de MDPV es patógena, p. ej., la cepa de MDPV YY. En algunas realizaciones, la ITR se deriva de un parvovirus porcino, p. ej., el parvovirus porcino U44978. En algunas realizaciones, la ITR se deriva de un virus diminuto de los ratones, p. ej., el virus diminuto de los ratones U34256. En algunas realizaciones, la ITR se deriva de un parvovirus canino, p. ej., el parvovirus canino M19296. En algunas realizaciones, la ITR se deriva de un virus de la enteritis de visión, p. ej., el virus de la enteritis de visión D00765. En algunas realizaciones, la ITR se deriva de un *Dependoparvovirus*. En una realización, el *Dependoparvovirus* es una cepa de parvovirus de *Dependovirus* de Oca (GPV). En una realización específica, la cepa de GPV está atenuada, p. ej., la cepa de GPV 82-0321V. En otra realización específica, la cepa de GPV es patógena, p. ej., la cepa de GPV B.

[0133] La primera ITR y la segunda ITR de la molécula de ácido nucleico pueden derivarse del mismo genoma, p. ej., del genoma del mismo virus, o de genomas diferentes, p. ej., de los genomas de dos o más genomas de virus diferentes. En ciertas realizaciones, la primera ITR y la segunda ITR se derivan del mismo genoma de AAV. En una realización específica, las dos ITR presentes en la molécula de ácido nucleico de la invención son las mismas y pueden ser en particular ITR de AAV2. En otras realizaciones, la primera ITR se deriva de un genoma de AAV y la segunda ITR no se deriva de un genoma de AAV (p. ej., un genoma no AAV). En otras realizaciones, la primera ITR no se deriva de un genoma de AAV (p. ej., un genoma no AAV) y la segunda ITR se deriva de un genoma de AAV. En otras realizaciones más, tanto la primera ITR como la segunda ITR no se derivan de un genoma de AAV (p. ej., un genoma no AAV). En una realización particular, la primera ITR y la segunda ITR son idénticas.

[0134] En algunas realizaciones, la primera ITR se deriva de un genoma no AAV y la segunda ITR se deriva de un genoma no AAV, en donde la primera ITR y la segunda ITR se derivan del mismo genoma. Los ejemplos no limitantes de genomas virales no AAV son de *Bocavirus*, *Dependovirus*, *Erythrovirus*, *Amdovirus*, *Parvovirus*, *Densovirus*, *Iteravirus*, *Contravirus*, *Aveparvovirus*, *Copiparvovirus*, *Protoparvovirus*, *Tetraparvovirus*, *Ambidensovirus*, *Brevidensovirus*, *Hepandensovirus* y *Penstyldensovirus*. En algunas realizaciones, la primera

ITR se deriva de un genoma no AAV y la segunda ITR se deriva de un genoma no AAV, en donde la primera ITR y la segunda ITR se derivan de diferentes genomas virales.

[0135] En algunas realizaciones, la primera ITR se deriva de un genoma de AAV, y la segunda ITR se deriva del bocavirus humano 1 (HBoV1). En otras realizaciones, la segunda ITR se deriva de un genoma de AAV, y la primera ITR se deriva del bocavirus humano 1 (HBoV1).

[0136] En algunas realizaciones, la primera ITR comprende o consiste en toda o una porción de una ITR derivada de un genoma de AAV o de no AAV, y la segunda ITR comprende o consiste en toda o una porción de una ITR derivada de un genoma de AAV o de noAAV. En algunas realizaciones, una porción de una ITR derivada de un genoma de de AAV o de no AAV es una versión truncada de una ITR natural derivada de un genoma de AAV o de no AAV. En algunas realizaciones, una porción de una ITR derivada de un genoma de AAV o de no AAV comprende porciones de una ITR natural derivada de un genoma de AAV o de no AAV. Por ejemplo, una porción de una ITR derivada de un genoma de AAV o de no AAV comprende porciones de una ITR natural derivada de un genoma de AAV o de no AAV, en donde se conserva al menos un RBE o una porción funcional del mismo.

[0137] En ciertas realizaciones, la primera ITR y/o la segunda ITR comprende o consiste en toda o una porción de una ITR derivada del HBoV1. En ciertas realizaciones, la primera ITR y/o la segunda ITR comprende o consiste en toda o una porción de una ITR derivada del HBoV1. En algunas realizaciones, la segunda ITR es un complemento inverso de la primera ITR. En algunas realizaciones, la primera ITR es un complemento inverso de la segunda ITR. En algunas realizaciones, la primera ITR y/o la segunda ITR derivada de HBoV1 es capaz de formar una estructura de horquilla. En algunas realizaciones, la estructura de horquilla no comprende una horquilla en forma de T.

[0138] En algunas realizaciones, la primera ITR y/o la segunda ITR comprende o consiste en una secuencia de nucleótidos al menos aproximadamente un 50 %, al menos aproximadamente un 55 %, al menos aproximadamente un 60 %, al menos aproximadamente un 65 %, al menos aproximadamente un 70 %, al menos aproximadamente un 75 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 %, al menos aproximadamente un 99 % o un 100 % idéntica a una secuencia de nucleótidos establecida en las SEQ ID NO: 1 o 2, en donde la primera ITR y/o la segunda ITR conserva una propiedad funcional de la ITR de HBoV1 de la que se deriva. En algunas realizaciones, la primera ITR y/o la segunda ITR comprende o consiste en una secuencia de nucleótidos al menos aproximadamente un 50 %, al menos aproximadamente un 55 %, al menos aproximadamente un 60 %, al menos aproximadamente un 65 %, al menos aproximadamente un 70 %, al menos aproximadamente un 75 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente

un 97 %, al menos aproximadamente un 98 %, al menos aproximadamente un 99 % o un 100 % idéntica a una secuencia de nucleótidos seleccionada de las SEQ ID NO: 1 o 2, en donde la primera ITR y/o la segunda ITR es capaz de formar una estructura de horquilla. En ciertas realizaciones, la estructura de horquilla no comprende una horquilla en forma de T.

[0139] En algunas realizaciones, la primera ITR y/o la segunda ITR comprende o consiste en la secuencia de nucleótidos de la SEQ ID NO: 1. En algunas realizaciones, la primera ITR y/o la segunda ITR comprende o consiste en la secuencia de nucleótidos de la SEQ ID NO: 2. En algunas realizaciones, la primera ITR comprende o consiste en la secuencia de nucleótidos establecida en la SEQ ID NO: 1. En algunas realizaciones, la segunda ITR comprende o consiste en la secuencia de nucleótidos establecida en la SEQ ID NO: 2. En algunas realizaciones, la primera ITR comprende o consiste en la secuencia de nucleótidos establecida en la SEQ ID NO: 1 y la segunda ITR comprende o consiste en la secuencia de nucleótidos establecida en la SEQ ID NO: 2.

[0140] Los expertos en la técnica apreciarán que cualquiera de las secuencias de la primera ITR descritas en la presente memoria puede emparejarse con cualquiera de las secuencias de la segunda ITR descritas en la presente memoria. En algunas realizaciones, la secuencia de la primera ITR descrita en la presente memoria es una secuencia ITR 5'. En algunas realizaciones, la secuencia de la segunda ITR descrita en la presente memoria es una secuencia ITR 3'. En algunas realizaciones, la secuencia de la segunda ITR descrita en la presente memoria es una secuencia ITR 5'. En algunas realizaciones, la secuencia de la primera ITR descrita en la presente memoria es una secuencia ITR 3'. Los expertos en la técnica podrán determinar la orientación adecuada de la primera y la segunda ITR descritas en la presente memoria con respecto a la arquitectura de un casete genético.

[0141] En otra realización particular, la ITR es una secuencia sintética diseñada genéticamente para incluir en sus extremos 5' y 3' ITR no derivados de un genoma de AAV. En otra realización particular, la ITR es una secuencia sintética diseñada genéticamente para incluir en sus extremos 5' y 3' ITR derivadas de uno o más genomas de no AAV. Las dos ITR presentes en la molécula de ácido nucleico de la invención pueden ser del mismo o de diferentes genomas de no AAV. En particular, las ITR pueden derivarse del mismo genoma de no AAV. En una realización específica, las dos ITR presentes en la molécula de ácido nucleico de la invención son las mismas y pueden ser, en particular, ITR de AAV2.

[0142] En algunas realizaciones, la secuencia de ITR comprende una o más secuencias palindrómicas. Una secuencia palindrómica de una ITR divulgada en la presente memoria incluye, pero no se limita a, secuencias palindrómicas nativas (es decir, secuencias que se encuentran en la naturaleza), secuencias sintéticas (es decir, secuencias que no se encuentran en la naturaleza), tales como secuencias pseudopalindrómicas, y combinaciones o formas modificadas de las mismas. Una "secuencia pseudopalindrómica" es una secuencia palindrómica de ADN, incluyendo una secuencia palindrómica imperfecta, que comparte menos

del 80 %, incluyendo menos del 70 %, 60 %, 50 %, 40 %, 30 %, 20 %, 10 %, o 5 %, o ninguna, identidad de secuencia de ácido nucleico con secuencias en la secuencia palindrómica nativa de AAV o de no AAV que forman una estructura secundaria. Las secuencias palindrómicas nativas pueden obtenerse o derivarse de cualquier genoma divulgado en la presente memoria. La secuencia palindrómica sintética puede basarse en cualquier genoma divulgado en la presente memoria.

[0143] La secuencia palindrómica puede ser continua o interrumpida. En algunas realizaciones, la secuencia palindrómica está interrumpida, en donde la secuencia palindrómica comprende una inserción de una segunda secuencia. En algunas realizaciones, la segunda secuencia comprende un promotor, un potenciador, un sitio de integración para una integrasa (p. ej., sitios para la recombinasa Cre o Flp), un marco de lectura abierto para un producto génico, o una combinación de los mismos.

[0144] En algunas realizaciones, las ITR forman estructuras de bucle horquilla. En una realización, la primera ITR forma una estructura de horquilla. En otra realización, la segunda ITR forma una estructura de horquilla. En otra realización más, tanto la primera ITR como la segunda ITR forman estructuras de horquilla. En algunas realizaciones, la primera ITR y/o la segunda ITR no forman una estructura de horquilla en forma de T. En ciertas realizaciones, la primera ITR y/o la segunda ITR forman una estructura de horquilla no en forma de T. En algunas realizaciones, la estructura de horquilla no en forma de T comprende una estructura de horquilla en forma de U.

[0145] En algunas realizaciones, una ITR en una molécula de ácido nucleico descrita en la presente memoria puede ser una ITR activada transcripcionalmente. Una ITR activada transcripcionalmente puede comprender toda o una porción de una ITR de tipo salvaje que ha sido activada transcripcionalmente mediante la inclusión de al menos un elemento transcripcionalmente activo. Varios tipos de elementos transcripcionalmente activos son adecuados para su uso en este contexto. En algunas realizaciones, el elemento transcripcionalmente activo es un elemento transcripcionalmente activo constitutivo. Los elementos transcripcionalmente activos constitutivos proporcionan un nivel continuo de transcripción del gen y se prefieren cuando se desea que el transgén se exprese de forma continua. En otras realizaciones, el elemento transcripcionalmente activo es un elemento transcripcionalmente activo inducible. Los elementos transcripcionalmente activos inducibles generalmente exhiben una baja actividad en ausencia de un inductor (o condición inductora) y son regulados al alza en presencia del inductor (o al cambiar a una condición inductora). Los elementos transcripcionalmente activos inducibles pueden ser preferibles cuando se desea la expresión sólo en ciertos momentos o en ciertos lugares, o cuando es deseable titular el nivel de expresión utilizando un agente inductor. Los elementos transcripcionalmente activos también pueden ser específicos de un tejido; es decir, sólo muestran actividad en determinados tejidos o tipos de células.

[0146] Los elementos transcripcionalmente activos pueden incorporarse a una ITR de diversas maneras. En algunas realizaciones, un elemento transcripcionalmente activo se incorpora en 5' de cualquier porción de una ITR o en 3' de cualquier porción de una ITR. En otras realizaciones, un elemento transcripcionalmente activo de una ITR activada transcripcionalmente se encuentra entre dos secuencias ITR. Si el elemento transcripcionalmente activo comprende dos o más elementos que deben estar espaciados, dichos elementos pueden alternar con porciones de la ITR. En algunas realizaciones, la estructura de horquilla de una ITR se elimina y se reemplaza por repeticiones invertidas de un elemento transcripcional. Esta última disposición crearía una horquilla que imita la estructura de la porción delecionada. También pueden estar presentes múltiples elementos transcripcionales activos en tándem en una ITR activada transcripcionalmente, y estos pueden ser adyacentes o estar espaciados. Además, pueden introducirse sitios de unión a proteínas (p. ej., sitios de unión a Rep) en los elementos transcripcionalmente activos de las ITR activadas transcripcionalmente. Un elemento transcripcionalmente activo puede comprender cualquier secuencia que permita la transcripción controlada del ADN por la ARN polimerasa para formar ARN, y puede comprender, por ejemplo, un elemento transcripcionalmente activo, como se define a continuación.

[0147] Las ITR activadas transcripcionalmente proporcionan tanto la activación transcripcional como las funciones de ITR a la molécula de ácido nucleico en una longitud de secuencia de nucleótidos relativamente limitada que maximiza efectivamente la longitud de un transgén que puede ser portado por y expresado a partir de la molécula de ácido nucleico. La incorporación de un elemento transcripcionalmente activo en una ITR puede llevarse a cabo de varias maneras. Una comparación de la secuencia de la ITR y de los requisitos de la secuencia del elemento transcripcionalmente activo puede proporcionar información sobre las formas de codificar el elemento dentro de una ITR. Por ejemplo, se puede añadir actividad transcripcional a una ITR mediante la introducción de cambios específicos en la secuencia de la ITR que replica los elementos funcionales del elemento transcripcionalmente activo. Existen varias técnicas en la técnica para añadir, delecionar y/o cambiar eficientemente secuencias de nucleótidos particulares en sitios específicos (véase, por ejemplo, Deng y Nickoloff (1992) Anal. Biochem. 200:81-88). Otra forma de crear ITR activadas transcripcionalmente implica la introducción de un sitio de restricción en un lugar deseado de la ITR. Además, se pueden incorporar múltiples elementos activadores de la transcripción en una ITR transcripcionalmente activada, utilizando métodos conocidos en la técnica.

[0148] A modo de ilustración, las ITR activadas transcripcionalmente pueden generarse mediante la inclusión de uno o más elementos transcripcionalmente activos como: caja TATA, caja GC, caja CCAAT, sitio Sp1, región Inr, sitio CRE (elemento regulador de AMPc), sitio ATF-1/CRE, caja APB β , caja APB α , caja CArG, caja CCAC, o cualquier otro elemento implicado en la transcripción como se conoce en la técnica.

B. Proteínas terapéuticas

[0149] Ciertos aspectos de la presente divulgación se dirigen a una molécula de ácido nucleico que comprende una primera ITR, una segunda ITR y un casete genético que codifica una secuencia diana, en donde la secuencia diana codifica una proteína terapéutica. En algunas realizaciones, el casete genético codifica una proteína terapéutica. En algunas realizaciones, el casete genético codifica más de una proteína terapéutica. En algunas realizaciones, el casete genético codifica dos o más copias de la misma proteína terapéutica. En algunas realizaciones, el casete genético codifica dos o más variantes de la misma proteína terapéutica. En algunas realizaciones, el casete genético codifica dos o más proteínas terapéuticas diferentes.

[0150] Ciertas realizaciones de la presente divulgación se dirigen a una molécula de ácido nucleico que comprende una primera ITR, una segunda ITR y un casete genético que codifica una proteína terapéutica, en donde la proteína terapéutica comprende un factor de coagulación. En algunas realizaciones, el factor de coagulación se selecciona del grupo que consiste en FI, FII, FIII, FIV, FV, FVI, FVII, FVIII, FIX, FX, FXI, FXII, FXIII, VWF, precalicreína, cininógeno de alto peso molecular, fibronectina, antitrombina III, cofactor de heparina II, proteína C, proteína S, proteína Z, inhibidor de la proteasa relacionada con la Proteína Z (ZPI), plasminógeno, antiplasmina alfa 2, activador del plasminógeno tisular (tPA), uroquinasa, inhibidor del activador del plasminógeno 1 (PAI-1), inhibidor del activador del plasminógeno 2 (PAI2), cualquier zimógeno de los mismos, cualquier forma activa de los mismos y cualquier combinación de los mismos. En una realización, el factor de coagulación comprende FVIII o una variante o fragmento del mismo. En otra realización, el factor de coagulación comprende FIX o una variante o fragmento del mismo. En otra realización, el factor de coagulación comprende FVII o una variante o fragmento del mismo. En otra realización, el factor de coagulación comprende VWF o una variante o fragmento del mismo.

[0151] En algunas realizaciones, la molécula de ácido nucleico comprende una primera ITR, una segunda ITR y un casete genético que codifica una secuencia diana, en donde la secuencia diana codifica una proteína terapéutica, en donde la proteína terapéutica comprende un polipéptido del factor VIII. "Factor VIII", abreviado a lo largo de la presente solicitud como "FVIII", tal y como se utiliza en la presente memoria, significa el polipéptido FVIII funcional en su papel normal en la coagulación, a menos que se especifique lo contrario. Así, el término FVIII incluye polipéptidos variantes que son funcionales. "Una proteína FVIII" se utiliza indistintamente con el polipéptido (o proteína) FVIII o FVIII. Los ejemplos de las funciones del FVIII incluyen, pero no se limitan a, la capacidad de activar la coagulación, la capacidad de actuar como cofactor del factor IX, o la capacidad de formar un complejo de tenasa con el factor IX en presencia de Ca^{2+} y fosfolípidos, que luego convierte el Factor X en la forma activada Xa.

[0152] La porción de FVIII en la proteína terapéutica utilizada en la presente memoria tiene actividad de FVIII. La actividad de FVIII puede medirse por cualquier método conocido en la técnica. Se dispone de una serie de ensayos para evaluar la función del sistema de coagulación: ensayo de tiempo de tromboplastina parcial activada (aPTT), ensayo cromogénico, ensayo

ROTEM, ensayo de tiempo de protrombina (PT) (también utilizado para determinar el INR), ensayo de fibrinógeno (a menudo por el método de Clauss), recuento de plaquetas, ensayos de la función plaquetaria (a menudo mediante PFA-100), TCT, tiempo de sangrado, ensayo de mezcla (si una anomalía se corrige si el plasma del paciente se mezcla con plasma normal), ensayos de factores de coagulación, anticuerpos antifosfolípidos, dímero D, ensayos genéticos (p. ej., factor V de Leiden, mutación de protrombina G20210A), tiempo de veneno de víbora de Russell diluido (dRVVT), ensayos misceláneos de la función plaquetaria, tromboelastografía (TEG o Sonoclot), tromboelastometría (TEM®, p. ej., ROTEM®), o el tiempo de lisis de la euglobulina (ELT).

[0153] El ensayo aPTT es un indicador de rendimiento que mide la eficacia tanto de la vía "intrínseca" (también denominada vía de activación por contacto) como de las vías comunes de coagulación. Este ensayo se utiliza habitualmente para medir la actividad de coagulación de los factores de coagulación recombinantes disponibles comercialmente, p. ej., FVIII. Se utiliza junto con el tiempo de protrombina (TP), que mide la vía extrínseca.

[0154] El análisis ROTEM proporciona información sobre toda la cinética de la hemostasia: tiempo de coagulación, formación del coágulo, estabilidad del coágulo y lisis. Los diferentes parámetros de la tromboelastometría dependen de la actividad del sistema de coagulación plasmática, de la función plaquetaria, de la fibrinólisis o de muchos factores que influyen en estas interacciones. Este ensayo puede proporcionar una visión completa de la hemostasia secundaria.

[0155] El mecanismo del ensayo cromogénico se basa en los principios de la cascada de la coagulación sanguínea, donde el FVIII activado acelera la conversión del Factor X en Factor Xa en presencia del Factor IX activado, fosfolípidos y iones de calcio. La actividad del Factor Xa se evalúa mediante la hidrólisis de un sustrato de p-nitroanilida (pNA) específico del Factor Xa. La tasa inicial de liberación de p-nitroanilina medida a 405 nm es directamente proporcional a la actividad del Factor Xa y, por tanto, a la actividad del FVIII en la muestra. El ensayo cromogénico está recomendado por el Subcomité del Comité Científico y de Estandarización (SSC) de FVIII y Factor IX de la Sociedad Internacional de Trombosis y Hemostasia (ISTH). Desde 1994, el ensayo cromogénico ha sido también el método de referencia de la Farmacopea Europea para la asignación de la potencia del concentrado de FVIII.

[0156] En algunas realizaciones, el casete genético comprende una secuencia de nucleótidos que codifica un polipéptido FVIII, en donde la secuencia de nucleótidos está optimizada en codones. En algunas realizaciones, el casete genético comprende una secuencia de nucleótidos que codifica un FVIII optimizado en codones, dirigido por un promotor mTTR y un intrón sintético. En algunas realizaciones, el casete genético comprende una secuencia de nucleótidos que se divulga en la Solicitud Internacional No. PCT/US2017/015879, que se incorpora por referencia en su totalidad. En algunas realizaciones, el casete genético es un casete genético

"hFVIIIco6XTEN" como se describe en PCT/US2017/015879. En algunas realizaciones, el casete genético comprende la SEQ ID NO: 32.

[0157] En algunas realizaciones, el casete genético comprende ADNc optimizado en codones que codifica el Factor VIII humano optimizado en codones con delección del dominio B (BDD) (BDDcoFVIII) fusionado con el péptido XTEN 144. En algunas realizaciones, el casete genético comprende la secuencia de nucleótidos establecida como SEQ ID NO: 9. En algunas realizaciones, el casete genético comprende la secuencia de nucleótidos establecida como SEQ ID NO: 33. En algunas realizaciones, el casete genético comprende la secuencia de nucleótidos establecida como SEQ ID NO: 14. En algunas realizaciones, el casete genético tiene la secuencia de nucleótidos de la SEQ ID NO: 14. En algunas realizaciones, la molécula de ácido nucleico comprende la secuencia de nucleótidos de la SEQ ID NO: 35.

[0158] En algunas realizaciones, el casete genético comprende una secuencia de nucleótidos que codifica un FVIII optimizado en codones dirigido por un promotor mTTR. En algunas realizaciones, el casete genético comprende además un elemento potenciador A1MB2. En algunas realizaciones, el casete genético comprende además un intrón quimérico o sintético. En algunas realizaciones, el casete genético comprende además un Elemento Regulador Postranscripcional de Marmota (WPRe). En algunas realizaciones, el casete genético comprende además una señal de Poliadenilación de la Hormona de Crecimiento Bovina (bGHpA).

[0159] En algunas realizaciones, la presente divulgación se dirige a moléculas de ácido nucleico optimizadas en codones que codifican un polipéptido con actividad de FVIII. En algunas realizaciones, el polinucleótido codifica un polipéptido FVIII de longitud completa. En otras realizaciones, la molécula de ácido nucleico codifica un polipéptido FVIII con delección del dominio B (BDD), en donde se deleciona todo o una parte del dominio B del FVIII.

[0160] En otras realizaciones, las moléculas de ácido nucleico divulgadas en la presente memoria se optimizan aún más mediante la eliminación de uno o más restos CpG y/o la metilación de al menos un resto CpG. Tal y como se utiliza en la presente memoria, "resto CpG" se refiere a una secuencia de dinucleótidos que contiene una citosina no metilada unida por un enlace fosfato a una guanosina. El término "resto CpG" abarca tanto dinucleótidos CpG metilados como no metilados. Los restos CpG no metilados son comunes en el ácido nucleico de origen bacteriano y viral (p. ej., ADN de los plásmidos), pero están suprimidos y en gran medida metilados en el ADN de los vertebrados. Así, los restos CpG no metilados estimulan al huésped mamífero a montar una rápida respuesta inflamatoria. Klinman, et al. (1996). PNAS 93:2879-2883. Se describen métodos ejemplares de eliminación de CpG en Yew, N. S., et al. (2002). Mol Ther. 5(6):731-738 y la Solicitud Internacional No. PCT/US2001/010309. En algunas realizaciones, las moléculas de ácido nucleico divulgadas en la presente memoria han sido modificadas para contener menos restos CpG (es decir, CpG reducidos o CpG deplecionados). En una realización, los restos CpG situados dentro de un triplete de codones para un aminoácido

seleccionado se cambian por un triplete de codones para el mismo aminoácido que carece de un resto CpG. En algunas realizaciones, las moléculas de ácido nucleico divulgadas en la presente memoria se han optimizado para reducir la respuesta inmunitaria innata.

[0161] En una realización particular, la molécula de ácido nucleico codifica un polipéptido que comprende una secuencia de aminoácidos que tiene al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 86 %, al menos aproximadamente un 87 %, al menos aproximadamente un 88 %, al menos aproximadamente un 89 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 % o al menos aproximadamente un 99 % de identidad de secuencia con la SEQ ID NO: 10 o un fragmento de la misma. En algunas realizaciones, la molécula de ácido nucleico de la divulgación codifica un polipéptido FVIII que comprende un péptido señal o un fragmento del mismo. En otras realizaciones, la molécula de ácido nucleico codifica un polipéptido FVIII que carece de péptido señal. En algunas realizaciones, el péptido señal comprende la secuencia de aminoácidos de la SEQ ID NO: 11. En algunas realizaciones, el péptido señal comprende los aminoácidos 1-19 de la SEQ ID NO: 10.

[0162] En algunas realizaciones, la molécula de ácido nucleico comprende una primera ITR, una segunda ITR y un casete genético que codifica una secuencia diana, en donde la secuencia diana codifica una proteína terapéutica, y en donde la proteína terapéutica comprende un factor de crecimiento. El factor de crecimiento puede seleccionarse entre cualquier factor de crecimiento conocido en la técnica. En algunas realizaciones, el factor de crecimiento es una hormona. En otras realizaciones, el factor de crecimiento es una citoquina. En algunas realizaciones, el factor de crecimiento es una quimioquina.

[0163] En algunas realizaciones, el factor de crecimiento es la adrenomedulina (AM). En algunas realizaciones, el factor de crecimiento es la angiopoyetina (Ang). En algunas realizaciones, el factor de crecimiento es un factor de motilidad autocrino. En algunas realizaciones, el factor de crecimiento es una proteína morfogénica ósea (BMP). En algunas realizaciones, la BMP se selecciona entre BMP2, BMP4, BMP5 y BMP7. En algunas realizaciones, el factor de crecimiento es un miembro de la familia del factor neurotrófico ciliar. En algunas realizaciones, el miembro de la familia del factor neurotrófico ciliar se selecciona entre el factor neurotrófico ciliar (CNTF), el factor inhibidor de la leucemia (LIF) y la interleuquina-6 (IL-6). En algunas realizaciones, el factor de crecimiento es un factor estimulante de colonias. En algunas realizaciones, el factor estimulante de colonias se selecciona entre el factor estimulante de colonias de macrófagos (m-CSF), el factor estimulante de colonias de granulocitos (G-CSF) y el factor estimulante de colonias de granulocitos y macrófagos (GM-CSF). En algunas realizaciones, el factor de crecimiento es un factor de crecimiento epidérmico (EGF). En algunas realizaciones, el factor de crecimiento es una efrina. En algunas realizaciones, la efrina se selecciona entre la efrina A1, la efrina A2, la efrina A3, la efrina A4, la

efrina A5, la efrina B1, la efrina B2 y la efrina B3. En algunas realizaciones, el factor de crecimiento es la eritropoyetina (EPO). En algunas realizaciones, el factor de crecimiento es un factor de crecimiento de fibroblastos (FGF). En algunas realizaciones, el FGF se selecciona entre FGF1, FGF2, FGF3, FGF4, FGF5, FGF6, FGF7, FGF8, FGF9, FGF10, FGF11, FGF12, FGF13, FGF14, FGF15, FGF16, FGF17, FGF19, FGF20, FGF21, FGF22 y FGF23. En algunas realizaciones, el factor de crecimiento es somatotropina bovina fetal (FBS). En algunas realizaciones, el factor de crecimiento es un miembro de la familia GDNF. En algunas realizaciones, el miembro de la familia GDNF se selecciona entre el factor neurotrófico derivado de la línea celular glial (GDNF), la neurturina, la persefina y la artemina. En algunas realizaciones, el factor de crecimiento es el factor de diferenciación de crecimiento-9 (GDF9). En algunas realizaciones, el factor de crecimiento es el factor de crecimiento de hepatocitos (HGF). En algunas realizaciones, el factor de crecimiento es el factor de crecimiento derivado del hepatoma (HDGF). En algunas realizaciones, el factor de crecimiento es la insulina. En algunas realizaciones, el factor de crecimiento es un factor de crecimiento similar a la insulina. En algunas realizaciones, el factor de crecimiento similar a la insulina es el factor de crecimiento similar a la insulina-1 (IGF-1) o el IGF-2. En algunas realizaciones, el factor de crecimiento es una interleuquina (IL). En algunas realizaciones, la IL se selecciona entre IL-1, IL-2, IL-3, IL-4, IL-5, IL-6 e IL-7. En algunas realizaciones, el factor de crecimiento es el factor de crecimiento de queratinocitos (KGF). En algunas realizaciones, el factor de crecimiento es el factor estimulante de la migración (FEM). En algunas realizaciones, el factor de crecimiento es la proteína estimulante de los macrófagos (MSP) o la proteína similar al factor de crecimiento de los hepatocitos (HGFLP). En algunas realizaciones, el factor de crecimiento es la miostatina (GDF-8). En algunas realizaciones, el factor de crecimiento es una neuregulina. En algunas realizaciones, la neuregulina se selecciona entre la neuregulina 1 (NRG1), NRG2, NRG3 y NRG4. En algunas realizaciones, el factor de crecimiento es una neurotrofina. En algunas realizaciones, el factor de crecimiento es el factor neurotrófico derivado del cerebro (BDNF). En algunas realizaciones, el factor de crecimiento es el factor de crecimiento nervioso (NGF). En algunas realizaciones, el NGF es la neurotrofina-3 (NT-3) o NT-4. En algunas realizaciones, el factor de crecimiento es el factor de crecimiento placentario (PGF). En algunas realizaciones, el factor de crecimiento es el factor de crecimiento derivado de las plaquetas (PDGF). En algunas realizaciones, el factor de crecimiento es la renalasa (RNLS). En algunas realizaciones, el factor de crecimiento es el factor de crecimiento de células T (TCGF). En algunas realizaciones, el factor de crecimiento es la trombopoyetina (TPO). En algunas realizaciones, el factor de crecimiento es un factor de crecimiento transformante. En algunas realizaciones, el factor de crecimiento transformante es el factor de crecimiento transformante alfa (TGF- α) o TGF- β . En algunas realizaciones, el factor de crecimiento es el factor de necrosis tumoral alfa (TNF- α). En algunas realizaciones, el factor de crecimiento es el factor de crecimiento del endotelio vascular (VEGF).

C. Secuencias de Control de la Expresión

[0164] En algunas realizaciones, la molécula de ácido nucleico o el vector de la divulgación comprende además al menos una secuencia de control de la expresión. Por ejemplo, la molécula de ácido nucleico aislada de la divulgación puede estar unida operativamente a al menos una secuencia de control de la expresión. La secuencia de control de la expresión puede ser, por ejemplo, una secuencia promotora o una combinación promotor-potenciador.

[0165] Los promotores constitutivos de mamíferos incluyen, pero no se limitan a, los promotores de los siguientes genes: hipoxantina fosforibosil transferasa (HPRT), adenosina desaminasa, piruvato quinasa, promotor de la actina beta y otros promotores constitutivos. Los promotores virales ejemplares que funcionan de forma constitutiva en las células eucariotas incluyen, por ejemplo, promotores del citomegalovirus (CMV), virus simio (p. ej., SV40), virus del papiloma, adenovirus, virus de la inmunodeficiencia humana (VIH), virus del sarcoma de Rous, citomegalovirus, repeticiones terminales largas (LTR) del virus de la leucemia de Moloney y otros retrovirus, y el promotor de la timidina quinasa del virus del herpes simple. Otros promotores constitutivos son conocidos por los expertos en la técnica. Los promotores útiles como secuencias de expresión génica de la divulgación también incluyen promotores inducibles. Los promotores inducibles se expresan en presencia de un agente inductor. Por ejemplo, el promotor de la metalotioneína se induce para promover la transcripción y la traducción en presencia de ciertos iones metálicos. Otros promotores inducibles son conocidos por los expertos en la técnica.

[0166] En una realización, la divulgación incluye la expresión de un transgén bajo el control de un promotor y/o potenciador específico de tejido. En otra realización, el promotor u otra secuencia de control de la expresión potencia selectivamente la expresión del transgén en células hepáticas. En ciertas realizaciones, el promotor u otra secuencia de control de la expresión potencia selectivamente la expresión del transgén en hepatocitos, células sinusoidales y/o células endoteliales. En una realización particular, el promotor u otra secuencia de control de la expresión potencia selectivamente la expresión del transgén en las células endoteliales. En ciertas realizaciones, el promotor u otra secuencia de control de la expresión potencia selectivamente la expresión del transgén en las células musculares, el sistema nervioso central, el ojo, el hígado, el corazón, o cualquier combinación de los mismos. Los ejemplos de promotores específicos de hígado incluyen, pero no se limitan a, un promotor de transtiretina de ratón (mTTR), un promotor de factor VIII humano nativo, un promotor de antitripsina alfa-1 humana (hAAT), un promotor mínimo de albúmina humana y un promotor de albúmina de ratón. En algunas realizaciones, las moléculas de ácido nucleico divulgadas en la presente memoria comprenden un promotor mTTR. El promotor mTTR se describe en Costa et al. (1986) Mol. Cell. Biol. 6:4697. El promotor de FVIII se describe en Figueiredo y Brownlee, 1995, J. Biol. Chem. 270:11828-11838. En algunas realizaciones, el promotor se selecciona entre un promotor específico del hígado (p. ej., antitripsina $\alpha 1$ (AAT)), un promotor específico del músculo (p. ej.,

creatina quinasa muscular (MCK), cadena pesada de miosina alfa (α MHC), mioglobina (MB) y desmina (DES)), un promotor sintético (p. ej., SPc5-12, 2R5Sc5-12, dMCK y tMCK), o cualquier combinación de los mismos.

[0167] En algunas realizaciones, la expresión del transgén se dirige al hígado. En ciertas realizaciones, la expresión del transgén se dirige a los hepatocitos. En otra realización, la expresión del transgén se dirige a las células endoteliales. En una realización particular, la expresión del transgén se dirige a cualquier tejido que exprese naturalmente el FVIII endógeno. En algunas realizaciones, la expresión del transgén se dirige al sistema nervioso central. En ciertas realizaciones, la expresión del transgén se dirige a las neuronas. En algunas realizaciones, la expresión del transgén se dirige a las neuronas aferentes. En algunas realizaciones, la expresión del transgén se dirige a las neuronas eferentes. En algunas realizaciones, la expresión del transgén se dirige a las interneuronas. En algunas realizaciones, la expresión del transgén se dirige a las células gliales. En algunas realizaciones, la expresión del transgén se dirige a los astrocitos. En algunas realizaciones, la expresión del transgén se dirige a los oligodendrocitos. En algunas realizaciones, la expresión del transgén se dirige a la microglía. En algunas realizaciones, la expresión del transgén se dirige a las células ependimarias. En algunas realizaciones, la expresión del transgén se dirige a las células de Schwann. En algunas realizaciones, la expresión del transgén se dirige a las células satélite. En algunas realizaciones, la expresión del transgén se dirige al tejido muscular. En algunas realizaciones, la expresión del transgén se dirige al músculo liso. En algunas realizaciones, la expresión del transgén se dirige al músculo cardíaco. En algunas realizaciones, la expresión del transgén se dirige al músculo esquelético. En algunas realizaciones, la expresión del transgén se dirige al ojo. En algunas realizaciones, la expresión del transgén se dirige a una célula fotorreceptora. En algunas realizaciones, la expresión del transgén se dirige a la célula ganglionar de la retina.

[0168] Otros promotores útiles en las moléculas de ácido nucleico divulgadas en la presente memoria incluyen un promotor de la transtiretina de ratón (mTTR), un promotor del FVIII humano nativo, un promotor de la antitripsina alfa-1 humana (hAAT), un promotor mínimo de la albúmina humana, un promotor de la albúmina de ratón, un promotor de la tristetraprolina (TTP), un promotor del CASI, un promotor del CAG, un promotor del citomegalovirus (CMV), un promotor de la antitripsina α 1 (AAT), un promotor de la creatina quinasa muscular (MCK), un promotor de la cadena pesada de miosina alfa (α MHC), un promotor de la mioglobina (MB), un promotor de la desmina (DES), un promotor SPc5-12, un promotor 2R5Sc5-12, un promotor de la dMCK, un promotor de la tMCK, un promotor de la fosfoglicerato quinasa (PGK) o un promotor de la antitripsina alfa-1 humana (A1AT), o cualquier combinación de los mismos.

[0169] En algunas realizaciones, las moléculas de ácido nucleico divulgadas en la presente memoria comprenden un promotor de transtiretina (TTR). En algunas realizaciones, el promotor es un promotor de transtiretina de ratón (mTTR). Los ejemplos no limitantes de

promotores de mTTR incluyen el promotor mTTR202, el promotor mTTR202opt, y el promotor mTTR482, como se divulga en la Publicación de EE. UU. No. US2019/0048362, que se incorpora por referencia en la presente memoria en su totalidad. En algunas realizaciones, el promotor es un promotor de transtiretina de ratón (mTTR) modificado específico de hígado. En algunas realizaciones, el promotor es el promotor de transtiretina de ratón (mTTR) modificado específico de hígado mTTR482. Se describen ejemplos de promotores mTTR482 en Kyostio-Moore et al. (2016) *Mol Ther Methods Clin Dev.* 3:16006, y Nambiar B. et al. (2017) *Hum Gene Ther Methods*, 28(1):23-28. En algunas realizaciones, el promotor es un promotor de transtiretina de ratón (mTTR) modificado específico de hígado que comprende la secuencia de ácido nucleico de la SEQ ID NO: 16. En algunas realizaciones, el promotor específico de tejido es el promotor de antitripsina alfa-1 (A1AT) humana. En algunas realizaciones, el promotor específico de tejido comprende la secuencia de nucleótidos de la SEQ ID NO: 36.

[0170] Los niveles de expresión pueden mejorarse aún más para lograr la eficacia terapéutica utilizando uno o más elementos potenciadores. Uno o más potenciadores pueden proporcionarse solos o junto con uno o más elementos promotores. Típicamente, la secuencia de control de la expresión comprende una pluralidad de elementos potenciadores y un promotor específico de tejido. En una realización, un potenciador comprende una o más copias del potenciador de microglobulina α -1/bicunina (Rouet et al. (1992) *J. Biol. Chem.* 267:20765-20773; Rouet et al. (1995), *Nucleic Acids Res.* 23:395-404; Rouet et al (1998) *Biochem. J.* 334:577-584; Ill et al. (1997) *Blood Coagulation Fibrinolysis* 8:S23-S30). En algunas realizaciones, el potenciador se deriva de sitios de unión de factores de transcripción específicos del hígado, como EBP, DBP, HNF1, HNF3, HNF4, HNF6, con Enh1, que comprende HNF1, (sentido)-HNF3, (sentido)-HNF4, (antisentido)-HNF1, (antisentido)-HNF6, (sentido)-EBP, (antisentido)-HNF4-(antisentido).

[0171] En algunas realizaciones, el elemento potenciador comprende uno o dos potenciadores de protrombina modificados (pPrT2), uno o dos potenciadores de microbicunina alfa 1 (A1MB2), un potenciador de albúmina de ratón modificado (mEalb), un potenciador del virus de la hepatitis B II (HE11) o un potenciador CRM8. En algunas realizaciones, el potenciador A1MB2 es el potenciador divulgado en la Solicitud Internacional No. PCT/US2019/055917. En algunas realizaciones, el elemento potenciador es A1MB2. En algunas realizaciones, el elemento potenciador incluye múltiples copias de la secuencia potenciadora A1MB2. En algunas realizaciones, el potenciador A1MB2 se sitúa en 5' de la secuencia de ácido nucleico que codifica el polipéptido FVIII. En algunas realizaciones, el potenciador A1MB2 se sitúa en 5' de la secuencia promotora, como el promotor mTTR. En algunas realizaciones, el elemento potenciador es el potenciador A1MB2 que comprende la secuencia de ácido nucleico de la SEQ ID NO: 15.

[0172] En algunas realizaciones, las moléculas de ácido nucleico divulgadas en la presente memoria comprenden un intrón o una secuencia intrónica. En algunas realizaciones, la

secuencia intrónica es una secuencia intrónica natural. En algunas realizaciones, la secuencia intrónica es una secuencia sintética. En algunas realizaciones, la secuencia intrónica se deriva de una secuencia intrónica natural. En algunas realizaciones, la secuencia intrónica es un intrón sintético híbrido o un intrón quimérico. En algunas realizaciones, la secuencia intrónica es un intrón quimérico que consiste en el intrón de actina beta de pollo/ globina beta de conejo y ha sido modificado para eliminar cinco secuencias ATG existentes para reducir los falsos inicios de traducción. En algunas realizaciones, el intrón quimérico comprende la secuencia de ácido nucleico de la SEQ ID NO: 17. En algunas realizaciones, la secuencia intrónica está posicionada en 5' de la secuencia de ácido nucleico que codifica el polipéptido FVIII. En algunas realizaciones, el intrón quimérico está posicionado en 5' de una secuencia promotora, como el promotor de mTTR.

[0173] En algunas realizaciones, las moléculas de ácido nucleico divulgadas en la presente memoria comprenden un elemento regulador postranscripcional. En ciertas realizaciones, el elemento regulador postranscripcional comprende un elemento regulador del virus de la hepatitis de la marmota mutado (WPRE). Se cree que el WPRE mejora la expresión de los transgenes administrados por el vector viral. Se describen ejemplos de WPRE en Zufferey et al. (1999) *J Virol.*, 73(4):2886-2892; Loeb et al. (1999) *Hum Gene Ther.* 10(14):2295-2305. En algunas realizaciones, el WPRE está situado en 3' de la secuencia de ácido nucleico que codifica el polipéptido FVIII. En algunas realizaciones, el WPRE comprende la secuencia de ácido nucleico de la SEQ ID NO: 18.

[0174] En algunas realizaciones, las moléculas de ácido nucleico divulgadas en la presente memoria comprenden un terminador de la transcripción. En algunas realizaciones, el terminador de la transcripción es una secuencia de poliadenilación (poli(A)). Los ejemplos no limitantes de terminadores de la transcripción incluyen aquellos derivados de la señal de poliadenilación de la hormona de crecimiento bovina (BGHpA), la señal de poliadenilación del virus de simio 40 (SV40pA), o una señal de poliadenilación sintética. En una realización, la cola de poli(A) 3'UTR comprende un sitio de poli(A) de actina. En una realización, la cola de poli(A) 3'UTR comprende un sitio de poli(A) de hemoglobina. En algunas realizaciones, el terminador transcripcional es BGHpA. Los ejemplos de terminadores transcripcionales BGHpA se describen en Woychik et al. (1984) *PNAS* 81:3944-3948. En algunas realizaciones, el terminador transcripcional se sitúa en el extremo 3' del casete genético que codifica la secuencia de ácido nucleico que codifica el polipéptido FVIII. En algunas realizaciones, el terminador transcripcional es un BGHpA que comprende la secuencia de ácido nucleico de la SEQ ID NO: 19.

[0175] En algunas realizaciones, la molécula de ácido nucleico divulgada en la presente memoria comprende una o más secuencias de direccionamiento nuclear del ADN (DTS). Una DTS promueve la translocación de moléculas de ADN que contienen dichas secuencias al núcleo. En ciertas realizaciones, la DTS comprende una secuencia potenciadora de SV40. En ciertas realizaciones, la DTS comprende una secuencia potenciadora de c-Myc. En algunas

realizaciones, la molécula de ácido nucleico comprende DTS que se localizan entre la primera ITR y la segunda ITR. En algunas realizaciones, la molécula de ácido nucleico comprende una DTS situada en 3' de la primera ITR y en 5' del transgén (p. ej., la proteína FVIII). En algunas realizaciones, la molécula de ácido nucleico comprende una DTS situada en 3' del transgén y en 5' de la segunda ITR en la molécula de ácido nucleico.

[0176] En algunas realizaciones, la molécula de ácido nucleico divulgada en la presente memoria comprende una secuencia de inhibición del receptor tipo Toll 9 (TLR9). Se describen secuencias ejemplares de inhibición de TLR9, por ejemplo, en Trieu et al. (2006) Crit Rev Immunol. 26(6):527-44; Ashman et al. Int'l Immunology 23(3): 203-14.

[0177] En algunas realizaciones, la molécula de ácido nucleico divulgada en la presente memoria comprende una secuencia de ácido nucleico que codifica una proteína no estructural del HBoV1. "Proteínas no estructurales" se refiere a cualquiera de las seis proteínas, a saber, NS1, NS1-70, NS2, NS3, NS4 y NP1, que son expresadas por el HBoV1. Las proteínas no estructurales se expresan mediante transcritos de ARNm generados a través del corte y empalme alternativo y la poliadenilación de un único pre-ARNm viral. Las proteínas NS1 a NS4 están codificadas en diferentes regiones del mismo marco de lectura abierto (ORF). La NS1 se une al origen de replicación del HBoV1 y presumiblemente mella el ADN monocatenario (ADNss) del origen durante la replicación de la horquilla rodante. La NS1 desempeña un papel importante en la expresión de la producción del vector mediada por la ITR de HBoV1 en células eucariotas. En una realización, se generaron construcciones de expresión que expresan NS1 de HBoV1. En algunas realizaciones, la molécula de ácido nucleico divulgada en la presente memoria codifica una proteína no estructural descrita en Shen et al. (2015) J Virology 89(19): 10097-10109.

[0178] En algunas realizaciones, la molécula de ácido nucleico comprende un sitio de unión de microARN (miARN). En una realización, el sitio de unión de miARN es un sitio de unión de miARN para miR-142-3p. En otras realizaciones, el sitio de unión de miARN es un sitio de unión de miARN descrito por Rennie et al. (2016) RNA Biol. 13(6):554-560.

Producción de ADNce en baculovirus

[0179] Los baculovirus son los virus más destacados que infectan a los insectos. Se han identificado más de 500 aislados de baculovirus, la mayoría de los cuales se originan en insectos del orden *Lepidoptera*. Los dos aislados más comunes son el nucleopoliedrovirus múltiple de *Autographa californica* (AcMNPV) y el nucleopoliedrovirus de *Bombyx mori* (BmNPV). Entre los vectores de expresión, los baculovirus destacan por su enorme capacidad de carga genética - hasta varias decenas de kb, con algunos informes de hasta 100 kb. Esta capacidad transgénica se ha utilizado para la producción de vectores AAV recombinantes (casetes de expresión de hasta 38 kb). Sin embargo, cuando se producen vectores virales o no virales para la terapia génica, a menudo se requieren varios vectores de expresión de baculovirus para ser infectados en células huésped de insectos. La generación de cada uno de los vectores de expresión de

baculovirus requiere mucho tiempo y eleva el coste de producción, lo que representa una desventaja significativa de la mayoría de los sistemas de vectores de expresión de baculovirus. Sin embargo, se generó un nuevo y versátil vector lanzadera de baculovirus (bácmido) diseñado específicamente para dar cabida a múltiples transgenes que podrían lograrse con las herramientas de bácmidos existentes. Este bácmido versátil (llamado "BIVVBac") también podría utilizarse para la producción de vectores rAAV para la terapia génica in vivo, así como para la producción de cualquier proteína deseada, p. ej., una proteína recombinante. Este sistema de expresión de bácmidos se describe con más detalle en la Solicitud de Patente de EE. UU. No. 63/069.073, que se incorpora por el presente documento por referencia en su totalidad.

[0180] En ciertas realizaciones, las moléculas de ácido nucleico divulgadas se producen utilizando un sistema de vector de expresión de baculovirus que comprende el bácmido recombinante "BIVVBac". En ciertas realizaciones, el BIVVBac es un AcMNPV genéticamente modificado que comprende al menos dos sitios de inserción de secuencias extrañas. Un sistema de vectores de expresión de baculovirus que comprende un bácmido que comprende al menos dos sitios de inserción de secuencias extrañas permite reducir el número total de vectores de expresión de baculovirus que deben generarse.

[0181] En ciertas realizaciones, el BIVVBac comprende un primer y un segundo sitio de inserción de secuencias extrañas. El primer y el segundo sitio de inserción de secuencias extrañas pueden ser diferentes, utilizando una maquinaria diferente para dirigir la inserción de una secuencia extraña (p. ej., una secuencia heteróloga, un gen heterólogo). La inserción de una secuencia extraña puede ser dirigida por cualquier método conocido en la técnica. Por ejemplo, una secuencia extraña puede insertarse por transposición o recombinación específica de sitio. Un sitio de inserción de la secuencia extraña puede diseñarse para que esté comprendido en un gen informador, de manera que, al insertar la secuencia extraña, el gen informador se interrumpa. La interrupción del gen informador puede ayudar a identificar los clones de bácmidos que tienen una secuencia extraña insertada en ellos. En tales realizaciones, el sitio de inserción de la secuencia extraña se fusiona en marco con el gen informador, o el gen informador se fusiona en marco con el sitio de inserción de la secuencia extraña.

[0182] En ciertas realizaciones, el primer sitio de inserción de secuencias extrañas permite la inserción de una secuencia extraña por transposición. En ciertas realizaciones, el primer sitio de inserción de secuencias extrañas comprende un sitio diana preferente para la inserción de un transposón. En ciertas realizaciones, el primer sitio de inserción de secuencias extrañas es un sitio diana preferente para la inserción de un transposón. En ciertas realizaciones, el primer sitio de inserción de secuencias extrañas es un sitio diana preferente que es un sitio de unión para un transposón bacteriano. Los transposones bacterianos adecuados y sus correspondientes sitios de unión son conocidos por los expertos en la técnica. Por ejemplo, el transposón Tn7 es conocido por su capacidad de transposición a un sitio específico de un cromosoma bacteriano (attTn7) con una alta frecuencia. En consecuencia, en ciertas realizaciones, el primer sitio de

inserción de secuencias extrañas es un sitio diana preferente que es un sitio de unión para un transposón Tn7 (p. ej., attTn7). En algunas realizaciones, el primer sitio de inserción de secuencias extrañas es un sitio diana preferente que es un sitio de unión para un transposón mini-Tn7 (p. ej., mini-attTn7, la secuencia de ADN mínima requerida para el reconocimiento por los factores de transposición Tn7 y la inserción de un transposón Tn7).

[0183] En ciertas realizaciones, el segundo sitio de inserción de secuencias extrañas permite la inserción de una secuencia extraña a través de la recombinación específica de sitio. En ciertas realizaciones, el segundo sitio de inserción de secuencias extrañas comprende un sitio diana preferente capaz de mediar un evento de recombinación específica de sitio. Los expertos en la técnica conocen varias tecnologías de recombinasa específica de sitio. Por ejemplo, el sistema Cre-loxP media la recombinación específica de sitio a través de la recombinasa Cre, que es capaz de reconocer secuencias de ADN de 34 pares de bases llamadas sitios loxP. En consecuencia, el segundo sitio de inserción de secuencias extrañas es un sitio diana preferente para la recombinación mediada por Cre. En ciertas realizaciones, el segundo sitio de inserción de secuencias extrañas es un sitio diana preferente que comprende un sitio loxP o una variante del mismo capaz de ser reconocido por la recombinasa Cre.

[0184] En algunas realizaciones, el bácmido recombinante comprende una variante del gen VP80, de manera que el bácmido exhibe una expresión reducida de su proteína codificada. Por ejemplo, se divulga en la presente memoria un esqueleto de ADN de baculovirus que comprende un gen VP80 inactivado debido a una inserción y/o delección en el locus del gen VP80. En algunas realizaciones, el bácmido recombinante comprende un bácmido divulgado en la Solicitud de Patente de EE. UU. No. US 63/069.115.

[0185] En ciertas realizaciones, el ADNce se produce utilizando un único vector de expresión de baculovirus. En este enfoque "UnBAC", un único vector de expresión de baculovirus (p. ej., BIVVBac) codifica todos los elementos esenciales requeridos para la producción de ADNce en el sistema de baculovirus y podría utilizarse potencialmente en cualquier línea celular permisiva para baculovirus para la producción de ADNce. Este enfoque se representa en la **FIG. 1A**.

[0186] En ciertas realizaciones, el ADNce se produce utilizando múltiples vectores de expresión de baculovirus. En este enfoque "DosBAC", los elementos esenciales requeridos para la producción de ADNce se insertan en dos baculovirus diferentes (p. ej., dos bácmidos BIVVBac) y podrían utilizarse potencialmente para la coinfección en cualquier línea celular permisiva para la infección por baculovirus. Este enfoque se representa en la **FIG. 1B**.

[0187] En ciertas realizaciones, el ADNce es producido por una línea celular estable. En este enfoque, los elementos esenciales requeridos para la producción de ADNce se insertan en ambos componentes del sistema de baculovirus. Este enfoque se representa en la **FIG. 1C**. Puede generarse una línea celular estable integrando de forma estable una secuencia codificante de proteína bajo el control de un promotor génico de baculovirus (p. ej., un promotor

génico constitutivo de baculovirus). En ciertas realizaciones, la línea celular estable es una línea celular de insectos estable.

[0188] Los métodos para la integración estable de ácidos nucleicos en una variedad de líneas celulares huésped son conocidos en la técnica. Por ejemplo, la selección repetida (p. ej., mediante el uso de un marcador seleccionable) puede utilizarse para seleccionar las células que han integrado un ácido nucleico que contiene un marcador seleccionable (y los genes de AAV cap y rep y/o un genoma de rAAV). En otras realizaciones, los ácidos nucleicos pueden integrarse de forma específica de sitio en una línea celular para generar una línea celular productora. Se conocen en la técnica varios sistemas de recombinación específica de sitio, como FLP/FRT (véase, p. ej., O'Gorman, S. et al. (1991) *Science* 251:1351-1355), Cre/loxP (véase, p. ej., Sauer, B. y Henderson, N. (1988) *Proc. Acad. Sci.* 85:5166-5170), y phi C31-att (véase, p. ej., Groth, A. C. et al. (2000) *Proc. Acad. Sci.* 97:5995-6000).

[0189] La divulgación también proporciona un polipéptido codificado por una molécula de ácido nucleico de la divulgación. En algunas realizaciones, el polipéptido de la divulgación está codificado por un vector que comprende las moléculas nucleicas aisladas divulgadas en la presente memoria. En otras realizaciones más, el polipéptido de la divulgación es producido por una célula huésped que comprende las moléculas nucleicas aisladas divulgadas en la presente memoria.

Células Huésped

[0190] La divulgación también proporciona una célula huésped que comprende una molécula de ácido nucleico o vector de la divulgación. Tal y como se utiliza en la presente memoria, el término "transformación" se utilizará en un sentido amplio para referirse a la introducción de ADN en una célula huésped receptora que cambia el genotipo y, en consecuencia, da lugar a un cambio en la célula receptora.

[0191] "Células huésped" se refiere a células que han sido transformadas con vectores contruidos mediante técnicas de ADN recombinante y que codifican al menos un gen heterólogo. Las células huésped de la presente divulgación son preferentemente de origen mamífero; más preferentemente de origen humano o de ratón. Los expertos en la técnica tienen la capacidad de determinar preferentemente las líneas celulares huésped particulares que mejor se adaptan a su propósito. Las líneas celulares huésped ejemplares incluyen, pero no se limitan a, CHO, DG44 y DUXB11 (líneas de ovario de hámster chino, DHFR menos), HELA (carcinoma de cuello uterino humano), CVI (línea de riñón de mono), COS (un derivado de CVI con antígeno T de SV40), R1610 (fibroblasto de hámster chino) BALBC/3T3 (fibroblasto de ratón), HAK (línea de riñón de hámster), SP2/O (mieloma de ratón), P3x63-Ag8.653 (mieloma de ratón), BFA-1c1BPT (células endoteliales bovinas), RAJI (linfocitos humanos), PER.C6®, NS0, CAP, BHK21 y HEK 293 (riñón humano). En una realización particular, la célula huésped se selecciona del grupo que consiste en: una célula CHO, una célula HEK293, una célula BHK21, una célula

PER.C6®, una célula NS0, una célula CAP y cualquier combinación de las mismas. En algunas realizaciones, las células huésped de la presente divulgación son de origen insecto. En una realización particular, las células huésped son células SF9. Las líneas de células huésped suelen estar disponibles en servicios comerciales, en la American Tissue Culture Collection o en la bibliografía publicada.

[0192] La introducción de las moléculas de ácido nucleico o los vectores de la divulgación en la célula huésped puede llevarse a cabo mediante diversas técnicas bien conocidas por los expertos en la técnica. Estas incluyen, pero no se limitan a, transfección (incluyendo la electroforesis y la electroporación), fusión de protoplastos, precipitación con fosfato de calcio, fusión celular con ADN envuelto, microinyección e infección con virus intactos. Véase, Ridgway, A. A. G. "*Mammalian Expression Vectors*" Capítulo 24.2, p. 470-472 Vectors, Rodriguez y Denhardt, Eds. (Butterworths, Boston, Mass. 1988). Lo más preferible es que la introducción del plásmido en el huésped se haga por electroporación. Las células transformadas se cultivan en condiciones apropiadas para la producción de cadenas ligeras y cadenas pesadas y se ensaya la síntesis de proteínas de cadenas pesadas y/o ligeras. Las técnicas de ensayo ejemplares incluyen el ensayo inmunoabsorbente ligado a enzimas (ELISA), el radioinmunoensayo (RIA), o el análisis de clasificación de células activada por fluorescencia (FACS), la inmunohistoquímica y similares.

[0193] Las células huésped que comprenden las moléculas de ácido nucleico aisladas o los vectores de la divulgación se cultivan en un medio de crecimiento apropiado. Tal y como se utiliza en la presente memoria, el término "medio de crecimiento apropiado" significa un medio que contiene los nutrientes necesarios para el crecimiento de las células. Los nutrientes necesarios para el crecimiento de las células pueden incluir una fuente de carbono, una fuente de nitrógeno, aminoácidos esenciales, vitaminas, minerales y factores de crecimiento. Opcionalmente, el medio puede contener uno o más factores de selección. Opcionalmente, el medio puede contener suero de ternera bovina o suero de ternera fetal (FCS). En una realización, el medio no contiene sustancialmente ninguna IgG. El medio de crecimiento generalmente seleccionará las células que contengan la construcción de ADN mediante, por ejemplo, la selección con fármacos o la deficiencia de un nutriente esencial que se complementa con el marcador seleccionable de la construcción de ADN o la cotransfección con la construcción de ADN. Las células de mamífero cultivadas generalmente se cultivan en medios disponibles comercialmente que contienen suero o desprovistos de suero (p. ej., MEM, DMEM, DMEM/F12). En una realización, el medio es CDoptiCHO (Invitrogen, Carlsbad, CA.). En otra realización, el medio es CD17 (Invitrogen, Carlsbad, CA.). La selección de un medio apropiado para la línea celular particular utilizada está dentro del nivel de los expertos en la técnica.

[0194] Los aspectos de la presente divulgación proporcionan un método de clonación de una molécula de ácido nucleico descrita en la presente memoria, que comprende la inserción de una molécula de ácido nucleico capaz de estructuras secundarias complejas en un vector adecuado,

y la introducción del vector resultante en una cepa huésped bacteriana adecuada. Como se conoce en la técnica, las estructuras secundarias complejas (p. ej., las regiones palindrómicas largas) de los ácidos nucleicos pueden ser inestables y difíciles de clonar en las cepas huésped bacterianas. Por ejemplo, las moléculas de ácido nucleico que comprenden una primera ITR y una segunda ITR (p. ej., ITR parvovirales no AAV, p. ej., ITR de HBoV1) de la presente divulgación pueden ser difíciles de clonar utilizando metodologías convencionales. Los palíndromos de ADN largos inhiben la replicación del ADN y son inestables en los genomas de *E. coli*, *Bacillus*, *Streptococcus*, *Streptomyces*, *S. cerevisiae*, ratones y seres humanos. Estos efectos son el resultado de la formación de estructuras de horquilla o cruciformes por el emparejamiento de bases dentro de la cadena. En *E. coli*, la inhibición de la replicación del ADN puede superarse significativamente en los mutantes SbcC o SbcD. SbcD es la subunidad de la nucleasa y SbcC es la subunidad de la ATPasa del complejo SbcCD. El complejo SbcCD de *E. coli* es un complejo de exonucleasas responsable de evitar la replicación de palíndromos largos. El complejo SbcCD es nuclear con actividad exonucleasa de ADN bicatenario dependiente de ATP y actividad endonucleasa de ADN monocatenario independiente de ATP. SbcCD puede reconocer los palíndromos de ADN y colapsar las horquillas de replicación atacando las estructuras de horquilla que surgen.

[0195] En ciertas realizaciones, una cepa huésped bacteriana adecuada es incapaz de resolver estructuras de ADN cruciformes. En ciertas realizaciones, una cepa huésped bacteriana adecuada comprende una disrupción en el complejo SbcCD. En algunas realizaciones, la disrupción en el complejo SbcCD comprende una disrupción genética en el gen SbcC y/o en el gen SbcD. En ciertas realizaciones, la disrupción en el complejo SbcCD comprende una disrupción genética en el gen SbcC. Se conocen en la técnica diversas cepas de huéspedes bacterianos que comprenden una disrupción genética en el gen SbcC. Por ejemplo, sin limitación, la cepa huésped bacteriana PMC103 comprende el genotipo *sbcC*, *recD*, *mcrA*, $\Delta mcrBCF$; la cepa huésped bacteriana PMC107 comprende el genotipo *recBC*, *recJ*, *sbcBC*, *mcrA*, $\Delta mcrBCF$; y la cepa huésped bacteriana SURE comprende el genotipo *recB*, *recJ*, *sbcC*, *mcrA*, $\Delta mcrBCF$, *umuC*, *uvrC*. En consecuencia, en algunas realizaciones, un método de clonación de una molécula de ácido nucleico descrita en la presente memoria comprende insertar una molécula de ácido nucleico capaz de estructuras secundarias complejas en un vector adecuado, e introducir el vector resultante en la cepa huésped PMC103, PMC107 o SURE. En ciertas realizaciones, el método de clonación de una molécula de ácido nucleico descrita en la presente memoria comprende insertar una molécula de ácido nucleico capaz de estructuras secundarias complejas en un vector adecuado e introducir el vector resultante en la cepa huésped PMC103.

[0196] Los vectores adecuados son conocidos en la técnica. En ciertas realizaciones, un vector adecuado para su uso en una metodología de clonación de la presente divulgación es un

vector con un número bajo de copias. En ciertas realizaciones, un vector adecuado para su uso en una metodología de clonación de la presente divulgación es pBR322.

[0197] En consecuencia, la presente divulgación proporciona un método de clonación de una molécula de ácido nucleico, que comprende insertar una molécula de ácido nucleico capaz de estructuras secundarias complejas en un vector adecuado, e introducir el vector resultante en una cepa huésped bacteriana que comprende una disrupción en el complejo SbcCD, en donde la molécula de ácido nucleico comprende una primera repetición terminal invertida (ITR) y una segunda ITR, en donde la primera ITR y/o la segunda ITR comprende una secuencia de nucleótidos al menos aproximadamente un 75 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 %, al menos aproximadamente un 99 %, o un 100 % idéntica a una secuencia de nucleótidos establecida en las SEQ ID NO. 1 o 2 o un derivado funcional de la misma.

Producción de Polipéptidos FVIII

[0198] La divulgación también proporciona un polipéptido codificado por una molécula de ácido nucleico de la divulgación. En algunas realizaciones, el polipéptido de la divulgación está codificado por un vector que comprende las moléculas nucleicas aisladas divulgadas en la presente memoria. En otras realizaciones más, el polipéptido de la divulgación es producido por una célula huésped que comprende las moléculas nucleicas aisladas de la divulgación.

[0199] Están disponibles diversos métodos para producir recombinantemente una proteína FVIII a partir de la molécula de ácido nucleico optimizada de la divulgación. Puede producirse un polinucleótido de la secuencia deseada mediante síntesis de ADN en fase sólida *de novo* o mediante mutagénesis por PCR de un polinucleótido preparado anteriormente. La mutagénesis mediada por oligonucleótidos es un método para preparar una sustitución, inserción, delección o alteración (p. ej., un codón alterado) en una secuencia de nucleótidos. Por ejemplo, el ADN de partida se altera mediante la hibridación de un oligonucleótido que codifica la mutación deseada con un molde de ADN monocatenario. Tras la hibridación, se utiliza una ADN polimerasa para sintetizar una segunda cadena complementaria completa del molde que incorpora el cebador oligonucleotídico. En una realización, la ingeniería genética, p. ej., la mutagénesis por PCR basada en el cebador, es suficiente para incorporar una alteración, como se define en la presente memoria, para producir un polinucleótido de la divulgación.

[0200] Para la producción de la proteína recombinante, se inserta una secuencia de polinucleótidos optimizada de la divulgación que codifica la proteína FVIII en un vehículo de expresión apropiado, es decir, un vector que contiene los elementos necesarios para la transcripción y traducción de la secuencia codificante insertada, o en el caso de un vector viral de ARN, los elementos necesarios para la replicación y traducción.

[0201] La secuencia de polinucleótidos de la divulgación se inserta en el vector en el marco de lectura adecuado. A continuación, el vector de expresión se transfecta en una célula diana adecuada que expresará el polipéptido. Las técnicas de transfección conocidas en la técnica incluyen, pero no se limitan a, la precipitación con fosfato de calcio (Wigler et al. 1978, *Cell* 14: 725) y la electroporación (Neumann et al. 1982, *EMBO, J.* 1: 841). Se puede utilizar una variedad de sistemas de huésped-vectores de expresión para expresar las proteínas FVIII descritas en la presente memoria en células eucariotas. En una realización, la célula eucariota es una célula animal, incluyendo células de mamífero (p. ej., células HEK293, células PER.C6®, CHO, BHK, Cos, HeLa). Una secuencia de polinucleótidos de la divulgación también puede codificar una secuencia señal que permitirá la secreción de la proteína FVIII. Un experto en la técnica entenderá que mientras la proteína FVIII se traduce, la secuencia señal es escindida por la célula para formar la proteína madura. Se conocen varias secuencias señal en la técnica, p. ej., la secuencia señal nativa del factor VII, la secuencia señal nativa del factor IX y la secuencia señal de la cadena ligera de la IgK de ratón. Alternativamente, cuando no se incluye una secuencia señal, la proteína FVIII puede recuperarse lisando las células.

[0202] La proteína FVIII de la divulgación puede sintetizarse en un animal transgénico, como un roedor, una cabra, una oveja, un cerdo o una vaca. El término "animales transgénicos" se refiere a animales no humanos que han incorporado un gen extraño en su genoma. Como este gen está presente en los tejidos de la línea germinal, se transmite de padres a hijos. Los genes exógenos se introducen en embriones unicelulares (Brinster et al. 1985, *Proc. Natl. Acad. Sci. USA* 82:4438). Los métodos de producción de animales transgénicos son conocidos en la técnica, incluidos los transgénicos que producen moléculas de inmunoglobulina (Wagner et al. 1981, *Proc. Natl. Acad. Sci. USA* 78: 6376; McKnight et al. 1983, *Cell* 34: 335; Brinster et al. 1983, *Nature* 306: 332; Ritchie et al. 1984, *Nature* 312: 517; Baldassarre et al. 2003, *Theriogenology* 59 : 831; Robl et al. 2003, *Theriogenology* 59: 107; Malassagne et al. 2003, *Xenotransplantation* 10 (3): 267).

[0203] Los vectores de expresión pueden codificar etiquetas que permitan una fácil purificación o identificación de la proteína producida recombinantemente. Los ejemplos incluyen, pero no se limitan a, el vector pUR278 (Ruther et al. 1983, *EMBO J.* 2: 1791) en el que la secuencia codificante de la proteína FVIII descrita en la presente memoria puede ligarse al vector en marco de la región codificante lac Z de modo que se produzca una proteína híbrida; los vectores pGEX pueden utilizarse para expresar proteínas con una etiqueta de glutatión S-transferasa (GST). Estas proteínas suelen ser solubles y pueden purificarse fácilmente a partir de las células mediante la adsorción en perlas de agarosa con glutatión, seguida de una elución en presencia de glutatión libre. Los vectores incluyen sitios de escisión (p. ej., PreCission Proteasa (Pharmacia, Peapack, N. J.)) para facilitar la eliminación de la etiqueta después de la purificación.

[0204] Para los fines de esta divulgación, pueden emplearse numerosos sistemas de vectores de expresión. Estos vectores de expresión son típicamente replicables en los organismos huésped, ya sea como episomas o como parte integral del ADN cromosómico del huésped. Los vectores de expresión pueden incluir secuencias de control de la expresión que incluyen, pero no se limitan a, promotores (p. ej., promotores asociados naturalmente o heterólogos), potenciadores, secuencias señal, señales de corte y empalme, elementos potenciadores y secuencias de terminación de la transcripción. Preferiblemente, las secuencias de control de la expresión son sistemas promotores eucariotas en vectores capaces de transformar o transfectar células huésped eucariotas. Los vectores de expresión también pueden utilizar elementos de ADN derivados de virus animales como virus del papiloma bovino, virus del poliovirus, adenovirus, virus vaccinia, baculovirus, retrovirus (RSV, MMTV o MOMLV), citomegalovirus (CMV) o virus SV40. Otros implican el uso de sistemas policistronicos con sitios de unión a ribosomas internos.

[0205] Comúnmente, los vectores de expresión contienen marcadores de selección (p. ej., resistencia a la ampicilina, resistencia a la higromicina, resistencia a la tetraciclina o resistencia a la neomicina) para permitir la detección de las células transformadas con las secuencias de ADN deseadas (véase, p. ej., Itakura *et al.*, Patente de EE. UU. 4,704,362). Las células que han integrado el ADN en sus cromosomas pueden seleccionarse introduciendo uno o más marcadores que permitan la selección de las células huésped transfectadas. El marcador puede proporcionar prototrofia a un huésped auxotrófico, resistencia a biocidas (p. ej., antibióticos) o resistencia a metales pesados como el cobre. El gen marcador seleccionable puede estar directamente ligado a las secuencias de ADN que se van a expresar o introducirse en la misma célula por cotransformación.

[0206] Un ejemplo de un vector útil para expresar una secuencia optimizada de FVIII es NEOSPLA (Patente de EE. UU. No. 6,159,730). Este vector contiene el promotor/potenciador del citomegalovirus, el promotor principal de la globina beta del ratón, el origen de replicación del SV40, la secuencia de poliadenilación de la hormona de crecimiento bovina, el exón 1 y el exón 2 de la neomicina fosfotransferasa, el gen de la dihidrofolato reductasa y la secuencia líder. Se ha descubierto que este vector da lugar a una expresión de muy alto nivel de anticuerpos tras la incorporación de los genes de la región variable y constante, la transfección en células, seguida de la selección en medio que contiene G418 y la amplificación con metotrexato. Los sistemas de vectores también se enseñan en las Pat. de EE. UU. No. 5,736,137 y 5,658,570, cada una de las cuales se incorpora por referencia en su totalidad en la presente memoria. Este sistema proporciona altos niveles de expresión, p. ej., > 30 pg/célula/día. Otros sistemas de vectores ejemplares se divulgan, p. ej., en la Patente de EE. UU. No. 6,413,777.

[0207] En otras realizaciones, los polipéptidos de la divulgación de la presente divulgación pueden expresarse utilizando construcciones policistronicas. En estos sistemas de expresión, pueden producirse múltiples productos génicos de interés, tales como múltiples polipéptidos de proteína de unión multimérica, a partir de una única construcción policistronica. Estos sistemas

utilizan ventajosamente un sitio de entrada del ribosoma interno (IRES) para proporcionar niveles relativamente altos de polipéptidos en las células huésped eucariotas. Las secuencias de IRES compatibles se divulgan en la Pat. de EE. UU. No. 6,193,980 que también se incorpora en la presente memoria.

[0208] De forma más general, una vez que se ha preparado el vector o la secuencia de ADN que codifica un polipéptido, el vector de expresión puede introducirse en una célula huésped adecuada. Es decir, las células huésped pueden ser transformadas. La introducción del plásmido en la célula huésped puede llevarse a cabo mediante diversas técnicas bien conocidas por los expertos en la técnica, como se ha comentado anteriormente. Las células transformadas se cultivan en condiciones apropiadas para la producción del polipéptido FVIII y se ensayan para determinar la síntesis del polipéptido FVIII. Las técnicas de ensayo ejemplares incluyen el ensayo inmunoabsorbente ligado a enzimas (ELISA), el radioinmunoensayo (RIA), o el análisis de clasificación de células activada por fluorescencia (FACS), la inmunohistoquímica y similares.

[0209] En las descripciones de los procesos para el aislamiento de polipéptidos a partir de huéspedes recombinantes, los términos "célula" y "cultivo celular" se utilizan indistintamente para denotar la fuente de polipéptidos, a menos que se especifique claramente lo contrario. En otras palabras, la recuperación del polipéptido a partir de las "células" puede significar tanto a partir de células enteras sedimentadas, como del cultivo celular que contiene tanto el medio como las células en suspensión.

[0210] La línea celular huésped utilizada para la expresión de la proteína es preferiblemente de origen mamífero; lo más preferiblemente de origen humano o de ratón, ya que los ácidos nucleicos aislados de la divulgación han sido optimizados para la expresión en células humanas. Se han descrito anteriormente líneas celulares huésped ejemplares. En una realización del método para producir un polipéptido con actividad de FVIII, la célula huésped es una célula HEK293. En otra realización del método para producir un polipéptido con actividad de FVIII, la célula huésped es una célula CHO.

[0211] Los genes que codifican los polipéptidos de la divulgación también pueden expresarse en células no de mamíferos como bacterias o levaduras o células vegetales. A este respecto, se apreciará que varios microorganismos unicelulares no mamíferos, como las bacterias, también pueden transformarse, es decir, aquellos capaces de crecer en cultivos o fermentación. Las bacterias, que son susceptibles de ser transformadas, incluyen miembros de las enterobacterias, como cepas de *Escherichia coli* o *Salmonella*; *Bacillaceae*, como *Bacillus subtilis*; *Pneumococcus*; *Streptococcus*, y *Haemophilus influenzae*. Se apreciará además que, cuando se expresan en bacterias, los polipéptidos suelen formar parte de cuerpos de inclusión. Los polipéptidos deben ser aislados, purificados y luego ensamblados en moléculas funcionales.

[0212] Alternativamente, las secuencias de nucleótidos optimizadas de la divulgación pueden incorporarse en transgenes para su introducción en el genoma de un animal transgénico y su posterior expresión en la leche del animal transgénico (véanse, p. ej., Deboer *et al.*, documento

US 5,741,957, Rosen, documento US 5,304,489, y Meade *et al.*, documento US 5,849,992). Los transgenes adecuados incluyen secuencias codificantes para polipéptidos en unión operativa con un promotor y potenciador de un gen específico de la glándula mamaria, como la caseína o la beta-lactoglobulina.

[0213] La producción *in vitro* permite el aumento de escala para obtener grandes cantidades de los polipéptidos deseados. Las técnicas para el cultivo de células de mamíferos en condiciones de cultivo de tejidos son conocidas en la técnica e incluyen el cultivo en suspensión homogénea, p. ej., en un reactor de aire o en un reactor de agitación continua, o el cultivo de células inmovilizadas o atrapadas, p. ej., en fibras huecas, microcápsulas, en microperlas de agarosa o en cartuchos de cerámica. Si es necesario y/o se desea, las soluciones de polipéptidos pueden purificarse mediante los métodos de cromatografía habituales, por ejemplo, filtración en gel, cromatografía de intercambio iónico, cromatografía sobre DEAE-celulosa o cromatografía de (inmuno)afinidad, p. ej., después de la biosíntesis preferente de un polipéptido sintético de la región bisagra o antes o después de la etapa de cromatografía HIC descrita en la presente memoria. Una secuencia de etiqueta de afinidad (p. ej., una etiqueta His(6)) puede unirse o incluirse opcionalmente dentro de la secuencia del polipéptido para facilitar la purificación aguas abajo.

[0214] Una vez expresada, la proteína FVIII puede purificarse de acuerdo con los procedimientos estándar de la técnica, incluyendo precipitación con sulfato de amonio, cromatografía en columna de afinidad, purificación por HPLC, electroforesis en gel y similares (véase, en general, Scopes, Protein Purification (Springer-Verlag, N.Y., (1982)). Para usos farmacéuticos se prefieren proteínas sustancialmente puras con una homogeneidad de al menos aproximadamente el 90 al 95 %, siendo lo más preferible una homogeneidad del 98 al 99 % o más.

Composiciones Farmacéuticas

[0215] Las composiciones que contienen una molécula de ácido nucleico aislada, un polipéptido que tiene actividad de FVIII codificado por la molécula de ácido nucleico, un vector o una célula huésped de la presente divulgación pueden contener un vehículo adecuado farmacéuticamente aceptable. Por ejemplo, pueden contener excipientes y/o auxiliares que faciliten el procesamiento de los compuestos activos en preparaciones diseñadas para su administración en el sitio de acción.

[0216] La composición farmacéutica puede formularse para administración parenteral (es decir, intravenosa, subcutánea o intramuscular) mediante inyección en bolo. Las formulaciones para inyección pueden presentarse en forma de dosificación unitaria, p. ej., en ampollas o en envases multidosis con un conservante añadido. Las composiciones pueden adoptar formas tales como suspensiones, soluciones o emulsiones en vehículos oleosos o acuosos, y contener agentes formuladores tales como agentes de suspensión, estabilización y/o dispersión.

Alternativamente, el ingrediente activo puede estar en forma de polvo para su constitución con un vehículo adecuado, p. ej., agua libre de pirógenos.

[0217] Las formulaciones adecuadas para la administración parenteral también incluyen soluciones acuosas de los compuestos activos en forma soluble en agua, por ejemplo, sales solubles en agua. Además, pueden administrarse suspensiones de los compuestos activos en forma de suspensiones oleosas de inyección adecuadas. Los disolventes o vehículos lipofílicos adecuados incluyen aceites grasos, por ejemplo, aceite de sésamo, o ésteres sintéticos de ácidos grasos, por ejemplo, oleato de etilo o triglicéridos. Las suspensiones acuosas inyectables pueden contener sustancias que aumentan la viscosidad de la suspensión, incluyendo, por ejemplo, carboximetilcelulosa sódica, sorbitol y dextrano. Opcionalmente, la suspensión puede contener también estabilizadores. Los liposomas también pueden utilizarse para encapsular las moléculas de la divulgación para su administración en células o espacios intersticiales. Los vehículos farmacéuticamente aceptables ejemplares son disolventes fisiológicamente compatibles, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardadores de la absorción, agua, solución salina, solución salina tamponada con fosfato, dextrosa, glicerol, etanol y similares. En algunas realizaciones, la composición comprende agentes isotónicos, por ejemplo, azúcares, polialcoholes como manitol, sorbitol o cloruro de sodio. En otras realizaciones, las composiciones comprenden sustancias farmacéuticamente aceptables como agentes humectantes o cantidades menores de sustancias auxiliares como agentes humectantes o emulsionantes, conservantes o tampones, que mejoran la vida útil o la eficacia de los ingredientes activos.

[0218] Las composiciones de la divulgación pueden estar en una variedad de formas, incluyendo, por ejemplo, líquido (p. ej., soluciones inyectables e infundibles), dispersiones, suspensiones, formas de dosificación semisólidas y sólidas. La forma preferida depende del modo de administración y de la aplicación terapéutica.

[0219] La composición puede formularse como solución, microemulsión, dispersión, liposoma u otra estructura ordenada adecuada a la alta concentración del fármaco. Las soluciones inyectables estériles pueden prepararse incorporando el ingrediente activo en la cantidad requerida en un disolvente apropiado con uno o una combinación de los ingredientes enumerados anteriormente, según sea necesario, seguido de esterilización por filtración. Generalmente, las dispersiones se preparan incorporando el ingrediente activo en un vehículo estéril que contiene un medio de dispersión básico y los demás ingredientes necesarios de los enumerados anteriormente. En el caso de los polvos estériles para la preparación de soluciones inyectables estériles, los métodos preferidos de preparación son el secado al vacío y la liofilización que produce un polvo del ingrediente activo más cualquier ingrediente adicional deseado a partir de una solución previamente esterilizada por filtración. La fluidez adecuada de una solución puede mantenerse, por ejemplo, mediante el uso de un recubrimiento como la lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de la dispersión

y mediante el uso de tensioactivos. La absorción prolongada de las composiciones inyectables puede conseguirse incluyendo en la composición un agente que retrase la absorción, por ejemplo, sales de monoestearato y gelatina.

[0220] El ingrediente activo puede formularse con una fórmula o dispositivo de liberación controlada. Los ejemplos de tales formulaciones y dispositivos incluyen implantes, parches transdérmicos y sistemas de administración microencapsulados. Pueden utilizarse polímeros biodegradables y biocompatibles, por ejemplo, acetato de vinilo y etileno, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres y ácido poliláctico. Los métodos para la preparación de tales formulaciones y dispositivos son conocidos en la técnica. Véase, p. ej., *Sustained and Controlled Release Drug Delivery Systems*, J. R. Robinson, ed., Marcel Dekker, Inc., Nueva York, 1978.

[0221] Las formulaciones inyectables de liberación lenta pueden hacerse formando matrices microencapsuladas del fármaco en polímeros biodegradables como la polilactida-poliglicolida. Dependiendo de la proporción entre el fármaco y el polímero, y de la naturaleza del polímero empleado, se puede controlar la tasa de liberación del fármaco. Otros polímeros biodegradables ejemplares son los poliortoésteres y los polianhídridos. Las formulaciones inyectables de liberación lenta también pueden prepararse atrapando el fármaco en liposomas o microemulsiones.

[0222] Pueden incorporarse a las composiciones compuestos activos suplementarios. En una realización, la proteína quimérica de la divulgación se formula con otro factor de coagulación, o una variante, fragmento, análogo o derivado del mismo. Por ejemplo, el factor de coagulación incluye, pero no se limita a, factor V, factor VII, factor VIII, factor IX, factor X, factor XI, factor XII, factor XIII, protrombina, fibrinógeno, factor de von Willebrand o factor tisular soluble recombinante (rsTF) o las formas activadas de cualquiera de los anteriores. El factor de coagulación del agente hemostático también puede incluir fármacos antifibrinolíticos, p. ej., ácido épsilon-amino-caproico, ácido tranexámico.

[0223] Los regímenes de dosificación pueden ajustarse para proporcionar la respuesta óptima deseada. Por ejemplo, puede administrarse un único bolo, varias dosis divididas a lo largo del tiempo, o la dosis puede reducirse o aumentarse proporcionalmente según las exigencias de la situación terapéutica. Es ventajoso formular las composiciones parenterales en forma de unidades de dosificación para facilitar la administración y la uniformidad de la dosificación. Véase, p. ej., *Remington's Pharmaceutical Sciences* (Mack Pub. Co., Easton, Pa. 1980).

[0224] Además del compuesto activo, la forma de dosificación líquida puede contener ingredientes inertes como agua, alcohol etílico, carbonato de etilo, acetato de etilo, alcohol bencílico, benzoato de bencilo, propilenglicol, 1,3-butilenglicol, dimetilformamida, aceites, glicerol, alcohol tetrahidrofurfurílico, polietilenglicoles y ésteres de ácidos grasos de sorbitán.

[0225] En Remington's Pharmaceutical Sciences, de E. W. Martin, se describen también ejemplos no limitativos de vehículos farmacéuticos adecuados. Algunos ejemplos de excipientes incluyen almidón, glucosa, lactosa, sacarosa, gelatina, malta, arroz, harina, tiza, gel de sílice, estearato de sodio, monoestearato de glicerol, talco, cloruro de sodio, leche desnatada en polvo, glicerol, propileno, glicol, agua, etanol y similares. La composición también puede contener reactivos tamponadores del pH y agentes humectantes o emulsionantes.

[0226] Para la administración oral, la composición farmacéutica puede adoptar la forma de comprimidos o cápsulas preparadas por medios convencionales. La composición también puede prepararse como un líquido, por ejemplo, un jarabe o una suspensión. El líquido puede incluir agentes de suspensión (p. ej., jarabe de sorbitol, derivados de la celulosa o grasas comestibles hidrogenadas), agentes emulsionantes (lecitina o goma arábica), vehículos no acuosos (p. ej., aceite de almendras, ésteres oleosos, alcohol etílico o aceites vegetales fraccionados) y conservantes (p. ej., metil o propil-p-hidroxibenzoatos o ácido sórbico). Las preparaciones también pueden incluir agentes aromatizantes, colorantes y edulcorantes. Alternativamente, la composición puede presentarse como un producto seco para su constitución con agua u otro vehículo adecuado.

[0227] Para la administración bucal, la composición puede adoptar la forma de comprimidos o pastillas para chupar los protocolos convencionales.

[0228] Para la administración por inhalación, los compuestos para su uso según la presente divulgación se administran convenientemente en forma de aerosol nebulizado con o sin excipientes o en forma de pulverización en aerosol de un envase presurizado o nebulizador, con opcionalmente un propulsor, p. ej., diclorodifluorometano, triclorofluorometano, diclorotetrafluorometano, dióxido de carbono u otro gas adecuado. En el caso de un aerosol presurizado, la unidad de dosificación puede determinarse proporcionando una válvula para suministrar una cantidad medida. Las cápsulas y cartuchos de, p. ej., gelatina para su uso en un inhalador o insuflador pueden formularse conteniendo una mezcla de polvo del compuesto y una base de polvo adecuada, como lactosa o almidón.

[0229] La composición farmacéutica también puede formularse para administración rectal como supositorio o enema de retención, p. ej., conteniendo bases convencionales para supositorios como la manteca de cacao u otros glicéridos.

[0230] En una realización, una composición farmacéutica comprende un polipéptido que tiene actividad de Factor VIII, una molécula de ácido nucleico optimizada que codifica el polipéptido que tiene actividad de Factor VIII, el vector que comprende la molécula de ácido nucleico, o la célula huésped que comprende el vector, y un vehículo farmacéuticamente aceptable. En algunas realizaciones, la composición se administra por una vía seleccionada del grupo que consiste en administración tópica, administración intraocular, administración parenteral, administración intratecal, administración subdural y administración oral. La administración parenteral puede ser administración intravenosa o subcutánea.

Métodos de Tratamiento

[0231] En algunos aspectos, la presente divulgación se dirige a métodos para tratar una enfermedad o afección en un sujeto que lo necesita, que comprenden administrar una molécula de ácido nucleico, un vector, un polipéptido o una composición farmacéutica divulgada en la presente memoria.

[0232] En algunas realizaciones, la divulgación se dirige a métodos para tratar un trastorno hemorrágico. En algunas realizaciones, la divulgación se dirige a métodos para tratar la hemofilia A.

[0233] La molécula de ácido nucleico aislada, el vector o el polipéptido pueden administrarse por vía intravenosa, subcutánea, intramuscular o a través de cualquier superficie mucosal, p. ej., por vía oral, bucal, sublingual, nasal, rectal, vaginal o pulmonar. La proteína del factor de coagulación puede implantarse en o unirse a un soporte sólido de biopolímero, lo que permite la liberación lenta de la proteína quimérica en el sitio deseado.

[0234] En una realización, la vía de administración de la molécula de ácido nucleico aislada, el vector o el polipéptido es parenteral. El término parenteral, tal y como se utiliza en la presente memoria, incluye la administración intravenosa, intraarterial, intraperitoneal, intramuscular, subcutánea, rectal o vaginal. En algunas realizaciones, la molécula de ácido nucleico aislada, el vector o el polipéptido se administran por vía intravenosa. Aunque todas estas formas de administración se contemplan claramente dentro del ámbito de la divulgación, una forma de administración sería una solución para inyección, en particular para inyección o goteo intravenoso o intraarterial.

[0235] Las dosis eficaces de las composiciones de la presente divulgación, para el tratamiento de afecciones, varían en función de muchos factores diferentes, incluidos los medios de administración, el sitio diana, el estado fisiológico del paciente, si el paciente es humano o animal, otros medicamentos administrados y si el tratamiento es profiláctico o terapéutico. Por lo general, el paciente es un ser humano, pero también pueden tratarse mamíferos no humanos, incluidos los mamíferos transgénicos. Las dosificaciones de tratamiento se pueden titular utilizando métodos rutinarios conocidos por los expertos en la técnica para optimizar la seguridad y la eficacia.

[0236] La molécula de ácido nucleico, el vector o los polipéptidos de la divulgación pueden administrarse opcionalmente en combinación con otros agentes que sean eficaces para tratar el trastorno o la afección que necesita tratamiento (p. ej., profiláctico o terapéutico).

[0237] Tal y como se utiliza en la presente memoria, la administración de moléculas de ácido nucleico aisladas, vectores o polipéptidos de la divulgación en conjunción o combinación con una terapia adyuvante significa la administración o aplicación secuencial, simultánea, coextensiva, concurrente, concomitante o contemporánea de la terapia y los polipéptidos divulgados. Los expertos en la técnica apreciarán que la administración o aplicación de los diversos componentes del régimen terapéutico combinado puede programarse en el tiempo para

mejorar la eficacia global del tratamiento. Un experto en la técnica (p. ej., un médico) sería capaz de discernir fácilmente regímenes terapéuticos combinados eficaces sin una experimentación indebida sobre la base de la terapia adyuvante seleccionada y las enseñanzas de la presente memoria descriptiva.

[0238] Se apreciará además que la molécula de ácido nucleico aislada, el vector o el polipéptido de la presente divulgación pueden usarse en conjunción o en combinación con un agente o agentes (p. ej., para proporcionar un régimen terapéutico combinado). Los agentes ejemplares con los que se puede combinar un polipéptido o polinucleótido de la divulgación incluyen agentes que representan el estándar actual de atención para un trastorno particular que se está tratando. Tales agentes pueden ser de naturaleza química o biológica. El término "biológico" o "agente biológico" se refiere a cualquier agente farmacéuticamente activo preparado a partir de organismos vivos y/o sus productos que está destinado a ser utilizado como agente terapéutico.

[0239] La cantidad de agente que debe utilizarse en combinación con los polinucleótidos o polipéptidos de la presente divulgación puede variar según el sujeto o puede administrarse según lo que se conoce en la técnica. Véase, p. ej., Bruce A Chabner *et al.*, *Antineoplastic Agents*, en GOODMAN & GILMAN'S THE PHARMACOLOGICAL BASIS OF THERAPEUTICS 1233-1287 (Joel G. Hardman *et al.*, eds., 9ª ed. 1996). En otra realización, se administra una cantidad de dicho agente consistente con el estándar de cuidado.

[0240] En una realización, también se divulga en la presente memoria un kit, que comprende la molécula de ácido nucleico divulgada en la presente memoria e instrucciones para administrar la molécula de ácido nucleico a un sujeto que la necesita. En otra realización, se divulga en la presente memoria un sistema de baculovirus para la producción de la molécula de ácido nucleico proporcionada en la presente memoria. La molécula de ácido nucleico se produce en células de insecto. En otra realización, se proporciona un sistema de administración de nanopartículas para construcciones de expresión. La construcción de expresión comprende la molécula de ácido nucleico descrita en la presente memoria.

Terapia Génica

[0241] En algunas realizaciones, la molécula de ácido nucleico divulgada en la presente memoria se utiliza en terapia génica. Las moléculas de ácido nucleico de FVIII optimizadas divulgadas en la presente memoria pueden utilizarse en cualquier contexto en el que se requiera la expresión de FVIII. En algunas realizaciones, las moléculas de ácido nucleico comprenden la secuencia de nucleótidos de la SEQ ID NO: 2. En algunas realizaciones, las moléculas de ácido nucleico comprenden la secuencia de nucleótidos de la SEQ ID NO: 1.

[0242] Por ejemplo, se ha explorado la terapia génica somática como un posible tratamiento para la hemofilia A. La terapia génica es un tratamiento particularmente atractivo para la hemofilia debido a su potencial para curar la enfermedad a través de la producción endógena

continua de FVIII tras una única administración de vector. La hemofilia A es muy adecuada para un enfoque de reemplazo génico porque sus manifestaciones clínicas son totalmente atribuibles a la falta de un único producto génico (FVIII) que circula en cantidades mínimas (200ng/ml) en el plasma.

[0243] En un aspecto, la molécula de ácido nucleico descrita en la presente memoria puede utilizarse en la terapia génica con AAV. El AAV es capaz de infectar una serie de células de mamífero. Véase, p. ej., Tratschin et al. (1985) Mol. Cell Biol. 5:3251-3260 y Grimm et al. (1999) Hum. Gene Ther. 10:2445-2450. Un vector rAAV lleva una secuencia de ácido nucleico que codifica un gen de interés, o un fragmento del mismo, bajo el control de secuencias reguladoras que dirigen la expresión del producto del gen en las células. En algunas realizaciones, el rAAV se formula con un vehículo y componentes adicionales adecuados para su administración.

[0244] En otro aspecto, la molécula de ácido nucleico descrita en la presente memoria puede utilizarse en la terapia génica lentiviral. Los lentivirus son virus de ARN en los que el genoma viral es ARN. Cuando una célula huésped se infecta con un lentivirus, el ARN genómico se transcribe de forma inversa en un ADN intermedio que se integra de forma muy eficiente en el ADN cromosómico de las células infectadas. En algunas realizaciones, el lentivirus se formula con un vehículo y componentes adicionales adecuados para su administración. En otro aspecto, la molécula de ácido nucleico descrita en la presente memoria puede utilizarse en la terapia adenoviral. Una revisión del uso de adenovirus para la terapia génica puede encontrarse, p. ej., en Wold et al. (2013) Curr Gene Ther. 13(6):421-33). En otro aspecto, la molécula de ácido nucleico descrita en la presente memoria puede utilizarse en terapia génica no viral. Una proteína FVIII optimizada de la divulgación puede producirse *in vivo* en un mamífero, p. ej., un paciente humano, utilizando un enfoque de terapia génica para el tratamiento de una enfermedad o trastorno hemorrágico seleccionado del grupo que consiste en un trastorno de coagulación hemorrágica, hemartrosis, hemorragia muscular, hemorragia oral, hemorragia, hemorragia en los músculos, hemorragia oral, traumatismo, traumatismo capitis, hemorragia gastrointestinal, hemorragia intracraneal, hemorragia intraabdominal, hemorragia intratorácica, fractura ósea, hemorragia del sistema nervioso central, hemorragia en el espacio retrofaríngeo, hemorragia en el espacio retroperitoneal y hemorragia en la vaina del iliopsoas serían terapéuticamente beneficiosas. En una realización, la enfermedad o trastorno hemorrágico es la hemofilia. En otra realización, la enfermedad o trastorno hemorrágico es la hemofilia A. Esto implica la administración de un ácido nucleico optimizado que codifica el FVIII unido operativamente a secuencias de control de la expresión adecuadas. En cierta realización, estas secuencias se incorporan a un vector viral. Los vectores virales adecuados para dicha terapia génica incluyen vectores adenovirales, vectores lentivirales, vectores baculovirales, vectores virales de Epstein Barr, vectores papovirales, vectores virales de vaccinia, vectores virales de herpes simple y vectores de virus adeno asociados (AAV). El vector viral puede ser un vector viral de replicación defectuosa. En otras realizaciones, un vector adenoviral tiene una delección en su gen E1 o en

su gen E3. En otras realizaciones, las secuencias se incorporan a un vector no viral conocido por los expertos en la técnica.

[0245] En otro aspecto, las moléculas de ácido nucleico divulgadas en la presente memoria se utilizan para la alteración específica de la información genética (p. ej., el genoma) de los organismos vivos. Tal y como se utiliza en la presente memoria, el término "alteración" o "alteración de la información genética" se refiere a cualquier cambio en el genoma de una célula. En el contexto del tratamiento de los trastornos genéticos, las alteraciones pueden incluir, pero no se limitan a, inserción, delección y/o corrección.

[0246] En algunos aspectos, las alteraciones también pueden incluir un *knock-in*, *knock-out* o *knock down* de un gen. Tal y como se utiliza en la presente memoria, el término "*knock-in*" se refiere a la adición de una secuencia de ADN, o de un fragmento de la misma en un genoma. Dichas secuencias de ADN que se van a someter a *knock-in* pueden incluir un gen o genes completos, pueden incluir secuencias reguladoras asociadas a un gen o cualquier porción o fragmento de lo anterior. Por ejemplo, puede insertarse un ADNc que codifique la proteína de tipo salvaje en el genoma de una célula portadora de un gen mutante. Las estrategias de *knock-in* no tienen por qué reemplazar el gen defectuoso, en su totalidad o en parte. En algunos casos, una estrategia de *knock-in* puede implicar además la sustitución de una secuencia existente por la secuencia proporcionada, p. ej., la sustitución de un alelo mutante por una copia de tipo salvaje. El término "*knock-out*" se refiere a la eliminación de un gen o de la expresión de un gen. Por ejemplo, un gen puede ser sometido a *knock-out* mediante una delección o una adición de una secuencia de nucleótidos que conduce a una interrupción del marco de lectura. Como otro ejemplo, un gen puede ser sometido a *knock-out* mediante el reemplazo de una parte del gen por una secuencia irrelevante. El término "*knock-down*", tal y como se utiliza en la presente memoria, se refiere a la reducción de la expresión de un gen o de su(s) producto(s) génico(s). Como resultado del *knock-down* de un gen, la actividad o función de la proteína puede ser atenuada o los niveles de la proteína pueden ser reducidos o eliminados.

[0247] En algunas realizaciones, las secuencias de ácido nucleico divulgadas en la presente memoria se utilizan para la edición del genoma. La edición del genoma se refiere generalmente al proceso de modificación de la secuencia de nucleótidos de un genoma, preferentemente de una manera precisa o predeterminada. Los ejemplos de métodos de edición del genoma descritos en la presente memoria incluyen métodos de uso de nucleasas dirigidas a sitio para cortar el ácido desoxirribonucleico (ADN) en ubicaciones diana precisas en el genoma, creando así roturas de ADN monocatenarias o bicatenarias en ubicaciones particulares dentro del genoma. Estas roturas pueden ser reparadas, y de hecho lo son, por procesos celulares naturales y endógenos, como la reparación dirigida por homología (HDR) y la unión de extremos no homólogos (NHEJ), como se ha revisado recientemente en Cox et al. (2015). Nature Medicine 21(2): 121-31. Estos dos procesos principales de reparación del ADN consisten en una familia de vías alternativas. La NHEJ une directamente los extremos del ADN resultantes de una rotura

bicatenaria, a veces con la pérdida o adición de una secuencia de nucleótidos, lo que puede interrumpir o mejorar la expresión de los genes. La HDR utiliza una secuencia homóloga, o secuencia donante, como molde para insertar una secuencia de ADN definida en el punto de rotura. La secuencia homóloga puede estar en el genoma endógeno, como una cromátida hermana. Alternativamente, el donante puede ser un ácido nucleico exógeno, como un plásmido, un oligonucleótido monocatenario, un oligonucleótido bicatenario, un oligonucleótido dúplex o un virus, que tiene regiones de alta homología con el locus escindido por la nucleasa, pero que también puede contener una secuencia adicional o cambios en la secuencia, incluyendo deleciones que pueden incorporarse al locus diana escindido. Un tercer mecanismo de reparación puede ser la unión de extremos mediada por microhomología (MMEJ), también denominada "NHEJ alternativa", en donde el resultado genético es similar a la NHEJ en el sentido de que pueden producirse pequeñas deleciones e inserciones en el sitio de escisión. La MMEJ puede hacer uso de secuencias homólogas de unos pocos pares de bases que flanquean el sitio de rotura del ADN para dirigir un resultado de reparación de la unión de extremos del ADN más favorecido, e informes recientes han dilucidado aún más el mecanismo molecular de este proceso, véase, p. ej., Cho y Greenberg (2015). Nature 518, 174-76. En algunos casos, puede ser posible predecir los resultados probables de la reparación basándose en el análisis de microhomologías potenciales en el sitio de la rotura del ADN.

[0248] Cada uno de estos mecanismos de edición del genoma puede utilizarse para crear las alteraciones genómicas deseadas. Un paso en el proceso de edición del genoma puede ser la creación de una o dos roturas de ADN, estas últimas como roturas bicatenarias o como dos roturas monocatenarias, en el locus diana tan cerca del sitio de la mutación deseada. Esto puede lograrse mediante el uso de polipéptidos dirigidos a sitio, como el sistema de endonucleasas CRISPR y otros.

[0249] En otro aspecto, la molécula de ácido nucleico descrita en la presente memoria puede utilizarse en la administración mediada por nanopartículas lipídicas (LNP) de ADNce de FVIII. Las nanopartículas lipídicas formadas a partir de lípidos catiónicos con otros componentes lipídicos, como lípidos neutros, colesterol, PEG, lípidos PEGilados y oligonucleótidos, se han utilizado para bloquear la degradación de los ácidos nucleicos en el plasma y facilitar la captación celular de los oligonucleótidos. Dichas nanopartículas lipídicas pueden utilizarse para administrar a los sujetos la molécula de ácido nucleico descrita en la presente memoria.

[0250] La divulgación proporciona un método para aumentar la expresión de un polipéptido con actividad de FVIII en un sujeto que comprende administrar la molécula de ácido nucleico aislada de la divulgación a un sujeto que la necesita, en donde la expresión del polipéptido se aumenta en relación con una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 6. La divulgación también proporciona un método para aumentar la expresión de un polipéptido con actividad de FVIII en un sujeto que comprende administrar un vector de la

divulgación a un sujeto que lo necesita, en donde la expresión del polipéptido se aumenta en relación con un vector que comprende una molécula de ácido nucleico de referencia.

[0251] Todos los diversos aspectos, realizaciones y opciones descritas en la presente memoria pueden combinarse en todas y cada una de las variaciones.

[0252] Todas las publicaciones, patentes y solicitudes de patentes mencionadas en esta memoria descriptiva se incorporan en la presente memoria por referencia en la misma medida como si se indicara específica e individualmente que cada publicación, patente o solicitud de patente individual se incorpora por referencia.

EJEMPLOS

[0253] Una vez proporcionada la divulgación anterior, se puede obtener una mayor comprensión mediante la referencia a los ejemplos que se proporcionan en la presente memoria. Estos ejemplos son solo para propósitos de ilustración y no pretenden ser limitantes.

Ejemplo 1: Enfoques para la producción de ADNce

[0254] En el sistema baculovirus-célula de insecto, el BEV recombinante administra el gen de interés bajo un promotor fuerte y proporciona el complejo transcripcional esencial para la replicación del virus en las células de insecto. Este sistema ofrece la flexibilidad de insertar el transgén de interés bien en el genoma del baculovirus y/o en el genoma de la célula de insecto en forma de línea celular estable. Aprovechando estas ventajas del sistema baculovirus-célula de insecto, se diseñaron tres enfoques diferentes de producción de ADNce para proporcionar una amplia selección según la facilidad de escalabilidad.

1. UnBAC

[0255] Para investigar el uso de un enfoque UnBac para la expresión de transgenes, el casete de expresión de FVIII^{WT} optimizado se insertó con ITR parvovirales en el sitio *mini-attTn7* en el locus de *Polihedrina* en BIVVBac a través de la transposición de Tn7 y en el mismo esqueleto, el casete de expresión del gen de Replicación (Rep) específico de ITR se insertó en el sitio LoxP en el locus de *EGT* a través de la recombinación Cre-LoxP. A continuación, se generó el BEV recombinante y se utilizó para la infección en células Sf9 con el fin de producir el ADNce de FVIII^{WT}, como se muestra en la **FIG. 1A**. Se utilizaron diferentes promotores para controlar los niveles de expresión de Rep para probar el concepto del enfoque UnBac para la producción de ADNce, como se describe a continuación.

2. DosBAC:

[0256] Para investigar el uso de un enfoque DosBac para la expresión de transgenes, el casete de expresión de FVIII^{WT} optimizado se insertó con ITR parvovirales y el casete de expresión del gen de replicación (Rep) específico de ITR en el sitio *mini-attTn7* del locus de la *polihedrina* en dos bácmidos BIVVBac diferentes mediante transposición de Tn7. A continuación, se generaron los BEV recombinantes y se utilizaron para la coinfección en células Sf9 con el fin de producir ADNce de FVIII^{WT}, como se muestra en la **FIG. 1B**. Se investigaron los desafíos

asociados con el enfoque DosBAC utilizando diferentes proporciones de multiplicidad de infección (MOI) de dos baculovirus y el ajuste de los niveles de expresión de Rep para obtener una productividad de ADNce reproducible, como se describe en los siguientes experimentos.

3. Línea Celular Estable:

[0257] Para investigar el uso de un enfoque de una línea celular estable para la expresión de transgenes, se generó una línea celular estable con el casete de expresión FVIIIIXTEN optimizado con ITR parvovirales. También se generó un bácmido recombinante insertando el casete de expresión del gen de la replicación (Rep) específico de ITR en el sitio *mini-attTn7* del locus de *la polihedrina* en el bácmido BIVVBac mediante la transposición de Tn7. A continuación, se generó el Rep.BEV recombinante y se utilizó para la infección en la línea celular estable de FVIIIIXTEN para producir ADNce de FVIIIIXTEN, como se muestra en la **FIG. 1C**. Los desafíos asociados con el enfoque de la Línea Celular Estable se investigaron enriqueciendo los transformantes FVIIIIXTEN a través de la clasificación celular FACS utilizando GFP como aproximación para acelerar el proceso de generación de la línea celular estable, como se describe en los siguientes experimentos.

Ejemplo 2: Construcción de expresión de FVIIIIXTEN ITR de HBoV1

[0258] El Bocavirus humano 1 (HBoV1), un parvovirus autónomo, es un virus auxiliar que apoya la replicación del virus adeno-asociado 2 (AAV2) de tipo salvaje. Se ha demostrado el uso de ITR parvovirales de AAV y de no AAV para la producción de ADNce de FVIIIIXTEN en el sistema de baculovirus (véase, p. ej., la Solicitud de Patente de EE. UU. No. 63/069.073). Las ITR de HBoV1 tienen un tamaño y una forma únicos en comparación con otras ITR parvovirales. La ITR 5' (REH) del HBoV1 tiene una longitud de 140 pb (SEQ ID NO: 1) y forma una horquilla en forma de "U" con un emparejamiento de bases perfecto, mientras que la ITR 3' (LEH) tiene una longitud de 200 pb (SEQ ID NO: 2) y forma un bucle con una unión de tres vías, lo que hace que tenga una naturaleza asimétrica distinta de las regiones terminales de otras ITR parvovirales (**FIG. 2A**).

[0259] Se investigaron las ITR de HboV1 para su uso en la producción de ADNce de FVIIIIXTEN. Se planteó la hipótesis de que las ITR asimétricas podrían mejorar la expresión persistente a largo plazo mediante la estabilización del transgén. Para ensayar esta hipótesis, se sintetizó una construcción de ADN que comprendía un Factor VIII humano optimizado en codones con dominio B deletado (BDD) (BDDcoFVIII) que comprendía un péptido XTEN 144 (FVIIIIXTEN) bajo la regulación del promotor de la transtiretina de ratón (mTTR) modificado específico del hígado (mTTR482) con un elemento potenciador (A1MB2), intrón sintético híbrido (Intrón Quimérico), el Elemento Regulador Postranscripcional de Marmota (WPRE), la señal de Poliadenilación de la Hormona de Crecimiento Bovina (bGHpA) y las ITR 5' y 3' flanqueantes del HBoV1 humano a través de GenScript® (Piscataway, NJ) ® para generar la secuencia de ácido nucleico establecida como SEQ ID NO: 3 (**FIG. 2A**). Este ADN sintético se clonó en el vector

pFastBac1 (Invitrogen) para generar el vector de transferencia pFastBac.mTTR.FVIIIXTEN.HBoV1.ITRs (**FIG. 2B**). A continuación, este vector se transformó en *E. coli* BIVVBac^{DH10B} para producir un BEV recombinante, AcBIVVBac.Polh.GPV.Rep^{Tn7}, como se describe a continuación.

Ejemplo 3: Construcciones de Expresión de NS1 de HBoV1 (no estructural)

Vectores de transferencia HBoV1 NS1 Tn7

[0260] Se sabe que HBoV1 expresa cinco proteínas no estructurales, a saber, NS1, NS2, NS3, NS4 y NP1, mediante transcritos de ARNm generados a través del corte y empalme alternativo y la poliadenilación de un único pre-ARNm viral. Las proteínas NS1 a NS4 están codificadas en diferentes regiones del mismo marco de lectura abierto (ORF). La NS1 consiste en un dominio de unión al origen/endonucleasa (OBD), un dominio de helicasa y un dominio putativo de transactivación (TAD) en el extremo N, región media y extremo C, respectivamente. La NS1 se une al origen de replicación del HBoV1 y presumiblemente mella el ADN monocatenario (ADNss) del origen durante la replicación de la horquilla rodante.

[0261] Para investigar el papel de NS1 en la producción de vectores mediada por ITR en células eucariotas y para "rescatar" un genoma de vector de ADNce de FVIIIXTEN flanqueado por ITR de HBoV1 en células Sf9, se generó una construcción de expresión de NS1 de HBoV1 y se insertó en el BIVVBac para producir BEV recombinante que expresara NS1 de HBoV1 en células Sf9.

[0262] Para generar el vector de expresión, la secuencia codificante de NS1 de HBoV1 se obtuvo del genoma del HBoV1 (no. de acceso de GenBank: JQ923422) y se optimizó en codones para el genoma de las células Sf antes de sintetizarlo mediante GenScript® para generar la secuencia de ácido nucleico establecida como SEQ ID NO: 4. El ADN sintético de NS1 de HBoV1 se clonó entonces en el vector pFastBac1 (Invitrogen) bajo el control del promotor de *Polihedrina* de AcMNPV (**FIG. 3A**) para generar el vector de transferencia pFastBac.Polh.HBoV1.NS1 (**FIG. 3B**). El ADN sintético de NS1 de HBoV1 también se clonó en el vector pFastBac1 (Invitrogen) bajo el control del promotor *inmediato-temprano1* (IE1) precedido por el elemento *hr5* del factor transcripcional de AcMNPV (**FIG. 4A**) para generar el vector de transferencia pFastBac.HR5.IE1.HBoV1.NS1 (**FIG. 4B**). El ADN sintético de NS1 de HBoV1 también se clonó en el vector pFastBac1 (Invitrogen) bajo el promotor *inmediato-temprano2* (IE2) de OpMNPV (**FIG. 5A**) para generar el vector de transferencia pFastBac.OpIE2.HBoV1.NS1 (**FIG. 5C**). A continuación, estos vectores se transformaron en *E. coli* BIVVBac^{DH10B} para producir BEV recombinantes: AcBIVVBac.Polh.HBoV1.NS1^{Tn7}, AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}, o AcBIVVBac.OpIE2.HBoV1.NS1^{Tn7}, respectivamente.

[0263] Estos BEV recombinantes se utilizaron entonces para la coinfección (DosBAC) junto con BEV FVIIIXTEN en células Sf9 para generar vectores de ADNce de FVIIIXTEN.

Vectores donantes Cre-LoxP de NS1 de HBoV1

[0264] Para investigar el uso de un enfoque UnBac para la producción de ADNce, se insertó el gen de NS1 de HBoV1 en el sitio LoxP en el casete de expresión recombinante BIVVBac que codifica FVIIIIXTEN en el sitio Tn7 en el locus de la *polihedrina*. La razón de insertar estos dos genes en estos sitios fue evitar la interferencia de la secuencia de repetición terminal invertida (ITR) que flanquea al FVIIIIXTEN con la secuencia LoxP, que también es una repetición palindrómica.

[0265] Además, para hacer frente a los desafíos asociados con el sistema UnBAC como se ha descrito anteriormente, se ensayaron diferentes promotores de genes de baculovirus que se expresan en diferentes momentos y niveles durante el ciclo de infección para controlar los niveles de expresión de NS1 de HBoV1 en UnBAC que codifican tanto las ITR de HBoV1 como la NS1.

[0266] El ADN sintético de NS1 de HBoV1 optimizado en codones para Sf se clonó en el vector donante Cre-LoxP (como se describe en la Solicitud de Patente de EE. UU. No. 63/069.073) bajo el control del promotor de la *polihedrina* de AcMNPV (**FIG. 3A**) o del promotor *inmediato-temprano1* precedido con (**FIG. 4A**) y sin (**FIG. 5B**) el elemento potenciador transcripcional *hr5* de AcMNPV para generar los vectores donantes Cre-LoxP pCL.Polh.HBoV1.NS1 (**FIG. 3C**), pCL.HR5.IE1.HBoV1.NS1 (**FIG. 4C**) y pCL.IE1.HBoV1.NS1 (**FIG. 5D**), respectivamente. Estas construcciones se designaron con el prefijo "pCL" para "Plásmido Cre-LoxP". (véanse las **FIGs. 3C, 4C, 5D**). Los vectores donantes Cre-LoxP resultantes se insertaron entonces en el sitio LoxP en el bácmido BIVVBac que codifica FVIIIIXTEN ITR de HBoV1 en el sitio Tn7 (**FIG. 6B**), como se describe a continuación.

Ejemplo 4: Vectores de Expresión de Baculovirus (BEV) FVIIIIXTEN ITR de HBoV1

[0267] Para generar BEV recombinante que codifica el casete de expresión de FVIIIIXTEN (**FIG. 2A**) con ITR de HBoV1, primero, *E. coli* BIVVBac^{DH10B} (descrita en la Solicitud de Patente de EE. UU. No. 63/069.073) se supertransformó con el vector de transferencia Tn7, pFastBac.mTTR.FVIIIIXTEN.HBoV1.ITRs (**FIG. 2B**). Los transformantes se seleccionaron con kanamicina, gentamicina, X-Gal e IPTG. La transposición específica de sitio del casete de expresión de FVIIIIXTEN y el gen de resistencia a la gentamicina en el sitio de inserción *del mini-attTn7* en el BIVVBac interrumpió el LacZα (fusionado en marco con *mini-attTn7*) y dio lugar a colonias blancas de *E. coli* en la selección dual de antibióticos mediada por X-Gal. Los ADN de bácmido recombinantes se aislaron de las colonias blancas de *E. coli* mediante miniprep por lisis alcalina y se digirieron con enzimas de restricción para determinar las estructuras genéticas correctas. Los resultados del mapeo con enzimas de restricción mostraron los fragmentos esperados para cada bácmido recombinante, sugiriendo la transposición específica de sitio del casete de expresión de FVIIIIXTEN con ITR de HBoV1 en el locus de la *Polihedrina* de BIVVBac (**FIG. 6A**). Se obtuvo una confirmación adicional mediante la amplificación por PCR de las regiones que abarcan el sitio de inserción esperado utilizando cebadores internos y externos al plásmido de transferencia y secuenciando los amplímeros resultantes (datos no mostrados).

[0268] El bácmido recombinante correcto que codifica el casete de expresión de FVIIIIXTEN con ITR de HBoV1 se purificó por maxi prep y se transfectó en células Sf9 utilizando el reactivo de transfección Cellfectin® (Invitrogen), según las instrucciones del fabricante. A los 4-5 días después de la transfección, se recogió la progenie de baculovirus y se purificó en placa en células Sf9, como se describió anteriormente. Jarvis *et al.* (2014), *Methods Enzymol*, 536: 149-163. Seis clones RFP+ purificados en placa del BEV recombinante, AcBIVVBac.mTTR.FVIIIIXTEN. HBoV1.ITRs^{Tn7} (**FIG. 6B**) se amplificaron hasta P1 (Paso 1) en células Sf9 sembradas a $0,5 \times 10^6$ por mL en frascos T25 en medio ESF-921 suplementado con Suero Bovino Fetal (FBS) inactivado por calor al 10 %. A los 4-5 días después de la infección, todos los clones mostraron una progresión de la infección, determinada por el número de células RFP+, para cada BEV clonal, lo que sugiere que el virus era capaz de replicarse normalmente y que la inserción del transgén FVIIIIXTEN con ITR de HBoV1 en el genoma del baculovirus no tenía ningún efecto adverso en la producción de la progenie de los virus. El clon más alto de RFP+ se seleccionó para su posterior amplificación en células Sf9 para producir preparaciones madre de trabajo de BEV (P2) y luego se utilizó para la coinfección con BEV de NS1 de HBoV1 en el sistema DosBAC para la producción de vectores de ADNce.

Ejemplo 5: Vectores de Expresión de Baculovirus (BEV) FVIIIIXTEN ITR de HBoV1 + NS1 de HBoV1

[0269] Para ensayar si BIVVBac podía utilizarse para albergar múltiples transgenes, se generó una familia de vectores derivados que codificaban dos casetes de expresión de transgenes: 1) FVIIIIXTEN ITR de HBoV1 y 2) NS1 de HBoV1 bajo el control de diferentes promotores. Estos BEV se produjeron en dos pasos. En primer lugar, el casete de expresión FVIIIIXTEN con ITR de HBoV1 se insertó en el sitio *mini-attTn7* en el locus de la *Polihedrina* en BIVVBac mediante de la transposición de Tn7, como se ha descrito anteriormente. A continuación, el bácmido resultante, BIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs (**FIG. 6B**) se utilizó para insertar el casete de expresión de NS1 de HBoV1 en el sitio LoxP del locus *EGT* mediante recombinación Cre-LoxP *in vitro* utilizando la recombinasa Cre (New England Biolabs).

[0270] En el proceso, los vectores donantes Cre-LoxP que codifican NS1 de HBoV1 bajo el promotor de la *polihedrina* de AcMNPV (**FIG. 3C**) o el promotor *inmediato-temprano1* (IE1) precedido con (**FIG. 4C**) y sin (**FIG. 5D**) el elemento potenciador transcripcional *hr5* de AcMNPV se insertaron en el bácmido BIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs (**FIG. 6B**). Las reacciones de recombinación se transformaron en *E. coli* DH10B y los transformantes se seleccionaron con kanamicina, gentamicina y ampicilina. Las colonias con triple resistencia a los antibióticos se cribaron mediante el mapeo con enzimas de restricción y/o PCR (**FIGs. 7A, 7B y 7C**) amplificando las regiones que abarcan el sitio de inserción esperado utilizando cebadores internos y externos al plásmido de transferencia y secuenciando los amplímeros resultantes.

[0271] El bácmido recombinante correcto que codificaba ambos casetes transgénicos se purificó con maxi prep y se transfectó en células Sf9 utilizando el reactivo de transfección Cellfectin® (Invitrogen). A los 4-5 días después de la transfección, se recogió la progenie de baculovirus y se purificó en placa en células Sf9. Seis clones RFP+ y GFP+ purificados en placa de cada BEV recombinante (AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1. NS1^{LoxP}: **FIG. 7D**; AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)IE1.HBoV1. NS1^{LoxP}: **FIG. 7E**; y AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)HR5.IE1.HBoV1.NS1^{LoxP}: **FIG. 7F**) se amplificaron hasta P1 (Paso 1) en células Sf9 sembradas a $0,5 \times 10^6$ por mL en frascos T25 en medio ESF921 suplementado con suero fetal bovino (FBS) inactivado por calor al 10 %. A los 4-5 días después de la infección, todos los clones mostraron una progresión de la infección, determinada por el número de células GFP+ y RFP+, para cada BEV recombinante, lo que sugiere que el virus fue capaz de replicarse normalmente y que la inserción de múltiples transgenes en el mismo genoma de baculovirus no tuvo ningún efecto adverso en la producción de la progenie de los virus. El virus P1 se recogió por centrifugación a baja velocidad y los sedimentos de las células infectadas se procesaron para la detección de NS1 de HBoV1 por inmunotransferencia. Finalmente, el clon con mayor expresión de NS1 de HBoV1 de cada BEV se amplificó aún más para producir preparaciones madre de trabajo de BEV (P2), seguido de la titulación en células Sf9. Los BEV titulados se utilizaron para la infección en células Sf9 con el fin de producir vectores de ADNce de FVIIIXTEN, como se describe a continuación.

Ejemplo 6: Producción de Vectores de ADNce de FVIIIXTEN ITR de HBoV1 a partir de UnBAC

[0272] Los BEV UnBAC que codifican los genes tanto FVIIIXTEN ITR de HBoV1 como NS1 de HBoV1 (**FIG. 7D-7F**) se ensayaron para determinar la producción de ADNce de FVIIIXTEN en células Sf9. Se infectaron aproximadamente $2,5 \times 10^6$ /mL células con la preparación madre de trabajo titulada (P2) de cada BEV con una multiplicidad de infecciones (MOI) de 0,1, 0,5, 1,0, 2,0 o 3,0 unidades formadoras de placas (pfu)/célula (**FIG. 8A**). Las células se suspendieron en 50mL de medio ESF-921 libre de suero y luego se incubaron durante 72-96h o hasta que la viabilidad celular alcanzó el 60-70 % en un incubador con agitación a 28° C. A las ~96h después de la infección, se recogieron las células infectadas y los sedimentos se procesaron para el aislamiento del vector de ADNce de FVIIIXTEN mediante el kit de aislamiento de ADN PureLink Maxi Prep (Invitrogen), de acuerdo con las instrucciones del fabricante. Las fracciones de elución finales se analizaron en electroforesis en gel de agarosa del 0,8 al 1,2 % para determinar la productividad del vector de ADNce de FVIIIXTEN.

[0273] El análisis en gel de agarosa del BEV AcBIVVBac (mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP} (**FIG. 8B**) que codifica FVIIIXTEN con ITR de HBoV1 y NS1 de HBoV1 dirigido por *polihedrina* se muestra en la **FIG. 8C**. Los resultados mostraron que la banda de ADN corresponde al tamaño del ADNce de FVIIIXTEN ITR de HBoV1

(~8,5 kb) en todas las dosis ensayadas con una productividad creciente a medida que aumenta la MOI.

[0274] Este resultado fue contrario a la productividad de ADNce obtenida con ITR de AAV2 UnBAC, donde se observó previamente una reducción de la productividad con el aumento de la carga viral. Sin estar limitado por la teoría, la proteína NS1 de HBoV1 puede tener un mecanismo único de unión y actividad endonucleasa en el sitio de resolución terminal de las ITR de HBoV1 para la replicación del ADN, que podría deberse a las distintas estructuras de las ITR REH y LEH (**FIG. 2A**).

[0275] En conclusión, estos experimentos han mostrado que el enfoque UnBAC sí demuestra el concepto de producción de ADNce a partir de un único BEV recombinante que codifica FVIIIIXTEN con los transgenes de ITR y NS1 de HBoV1. También muestra la viabilidad y funcionalidad de múltiples transgenes insertados en diferentes loci en un vector lanzadera de baculovirus (BIVVBac), y su uso potencial para la producción de vectores AAV recombinantes en el sistema de baculovirus células de insectos.

Ejemplo 7: Vectores de Expresión de Baculovirus (BEV) de NS1 de HBoV1 (no estructural)

[0276] El único dominio de nucleasa N-terminal de NS1 de parvovirus caracterizado estructuralmente es de Rep de AAV2, que se une a las repeticiones de tetranucleótidos consecutivas en el origen de replicación (Ori). Sin embargo, estas repeticiones de tetranucleótidos son específicas de AAV y no están presentes en HBoV1. De hecho, la LEH (3' ITR) del genoma de HBoV1 forma un bucle con una unión de tres vías, mientras que la REH (5' ITR) es una horquilla con un emparejamiento perfecto de bases (**FIG. 2A**), que se conservan en los bocavirus pero son distintas de las regiones terminales de los genomas de AAV y parovirus B19 (B19V). Estos hallazgos sugieren que el modo de reconocimiento de NS1 del Ori en HBoV1 puede ser distinto al de AAV. Además, no se sabe que AAV cause enfermedades humanas y es un dependovirus porque la replicación del virus requiere un virus auxiliar como el herpesvirus o el adenovirus. La NS1 de HBoV1 comparte tan poco como un 14 % de identidad de secuencia con Rep de AAV. Se ha demostrado que NS1 de HBoV1 contiene una superficie cargada positivamente que es el sitio de unión putativo para el Ori y apoya directamente la replicación del ADN del HBoV, como en el mecanismo común de la horquilla rodante propuesto para los parvovirus.

[0277] La NS1 de HBoV1 parece ser esencial para la producción de vectores mediada por ITR en células eucariotas. Para investigar el "rescate" potencial de un genoma de vector FVIIIIXTEN flanqueado por ITR a partir de células Sf9 o BEV FVIIIIXTEN, se generaron BEV recombinantes que codificaban NS1 de HBoV1 bajo diferentes promotores baculovirales para optimizar los niveles de expresión de NS1 en células Sf9.

[0278] Para generar estos BEV, se supertransformaron *E. coli* BIVVBac^{DH10B} (véase la Solicitud de Patente de EE. UU. No. 63/069.073) con vectores de transferencia Tn7,

pFastBac.Polh.HBoV1-NS1 (**FIG. 3B**), pFastBac.HR5.IE1.HBoV1-NS1 (**FIG. 4B**), o pFastBac.OpIE2.HBoV1-NS1 (**FIG. 5C**). Los transformantes se seleccionaron con kanamicina, gentamicina, X-Gal e IPTG. La transposición específica de sitio del casete de expresión HBoV1-NS1 y del gen de resistencia a la gentamicina en el sitio de inserción *mini-attTn7* en BIVVBac interrumpió LacZα (fusionado en marco con *mini-attTn7*) y dio lugar a colonias blancas de *E. coli* en la selección dual de antibióticos mediada por X-Gal. Así, se aislaron los ADN de báculo recombinantes de las colonias blancas de *E. coli* mediante miniprep por lisis alcalina y se digirieron con enzimas de restricción para determinar las estructuras genéticas correctas. Los resultados del mapeo con enzimas de restricción mostraron fragmentos esperados para cada báculo recombinante, lo que sugiere la transposición específica de sitio de NS1 de HBoV1 en el locus de la *Polihedrina* de BIVVBac (**FIG. 9A**). Se obtuvo una confirmación adicional mediante la amplificación por PCR de las regiones que abarcan el sitio de inserción esperado, utilizando cebadores internos y externos al plásmido de transferencia y secuenciando los amplímeros resultantes (datos no mostrados).

[0279] Los báculos recombinantes confirmados como correctos que codifican Polh.HBoV1-NS1, HR5.IE1.HBoV1-NS1 u OpIE2-HBoV1-NS1 se transfectaron en células Sf9 utilizando el reactivo de transfección Cellfectin® (Invitrogen), según las instrucciones del fabricante. A los 4-5 días después de la transfección, se recogió la progenie de baculovirus y se purificó en placa en células Sf9, como se describió previamente. Jarvis *et al.* (2014), *Methods Enzymol*, 536: 149-163. Seis clones RFP+ purificados en placa de cada BEV recombinante, AcBIVVBac.Polh.HBoV1-NS1^{Tn7} (**FIG. 9B**), AcBIVVBac.HR5.IE1.HBoV1-NS1^{Tn7} (**FIG. 9C**), y AcBIVVBac.OpIE2.HBoV1-NS1^{Tn7} (**FIG. 9D**) se amplificaron hasta P1 (Paso 1) en células Sf9 sembradas a $0,5 \times 10^6$ por mL en frascos T25 en medio ESF-921 suplementado con suero bovino fetal (FBS) inactivado por calor al 10%. A los 4-5 días después de la infección, todos los clones mostraron una progresión de la infección, determinada por el número de células RFP+, para cada BEV clonal, lo que sugiere que el virus fue capaz de replicarse normalmente y que la inserción de NS1 de HBoV1 en el genoma del baculovirus no tuvo ningún efecto adverso en la producción de la progenie de los virus.

[0280] El clon RFP+ más alto se seleccionó para su posterior amplificación en células Sf9 para producir preparaciones madre de trabajo de BEV (P2). La preparación madre de virus titulada se utilizó entonces para la coinfección con BEV FVIIIXTEN en el sistema DosBAC o en la línea celular estable FVIIIXTEN ITR de HBoV1 para la producción de vectores de ADNce.

Ejemplo 8: Producción del Vector de ADNce FVIIIXTEN ITR de HBoV1 a partir de DosBAC

[0281] Para investigar el enfoque de DosBAC para la expresión de transgenes, se ensayaron las BEV recombinantes clonales que codifican FVIIIXTEN ITR de HBoV1 con BEV de HBoV1-NS1 dirigidos por *polihedrina* para coinfecciones en diferentes proporciones de MOI de proporciones 1:10 y 1:5 o en diferentes proporciones de una MOI de 0,3, 1,0, 3,0 y 5,0 pfu/célula

para la producción de vectores de ADNce de FVIIIXTEN en células Sf9 (**FIG. 10A**). Específicamente, se sembraron $\sim 2,0 \times 10^6$ /mL células en 50mL de medio ESF-921 libre de suero y se coinfectaron con preparaciones madre de trabajo tituladas (P2) de BEV AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7} a una MOI de 0,1, 0,3, 0,5, 1,0, 3,0, 5,0 pfu/célula con BEV AcBIVVBac.Polh.HBoV1-NS1^{Tn7} a una MOI de 0,01, 0,03, 0,05, 0,1, 0,3, 0,5 pfu/célula para mantener la proporción constante 1:10 o a una MOI de 0,02, 0,06, 0,1, 0,2, 0,6, 1,0 pfu/célula para mantener la proporción constante 1:5, respectivamente. Del mismo modo, las células también se coinfectaron en una proporción 1:1, 1:2, 1:5 o 1:10 con una MOI constante de 0,3, 1,0, 3,0 o 5,0 pfu/célula (**FIG. 10B**). En cada caso, no se eliminó el inóculo del virus y las células se incubaron hasta que la viabilidad alcanzó el 60-70 % en un incubador con agitación a 28° C. A las ~96h después de la infección, se recogieron las células infectadas y los sedimentos se procesaron para el aislamiento del vector de ADNce de FVIIIXTEN mediante el kit de aislamiento de ADN PureLink Maxi Prep (Invitrogen), de acuerdo con las instrucciones del fabricante. Las fracciones de elución final se analizaron en electroforesis en gel de agarosa del 0,8 al 1,2 % para determinar la productividad del ADNce.

[0282] Como era de esperar, los análisis en gel de agarosa mostraron un grado variable de productividad de ADNce de FVIIIXTEN con diferentes condiciones. Sin embargo, DosBAC coinfectado a una MOI de 3,0 pfu/célula mostró niveles crecientes de productividad de ADNce de FVIIIXTEN con aumentos en las proporciones de carga de virus, siendo 1:10 la más alta en comparación con las otras condiciones ensayadas (**FIG. 10C**). La mayor carga de virus parece mejorar la productividad del ADNce de FVIIIXTEN ITR de HBoV1, lo que es consistente con la observación en los BEV UnBAC (véase el Ejemplo 6). Esto sugiere además el requisito de un mayor nivel de HBoV1-NS1 para la replicación de ADNce de FVIIIXTEN dependiente de ITR de HBoV1 en las células Sf9.

[0283] Los resultados con UnBAC o DosBAC indican que el nivel de replicación de HBoV1-NS1 tiene un impacto significativo en la productividad del ADNce de FVIIIXTEN en el sistema de baculovirus.

[0284] Como alternativa al ensayo de varias condiciones diferentes de coinfección, como se ha comentado anteriormente, se exploraron otras formas de mejorar la productividad del ADNce de FVIIIXTEN aprovechando diferentes promotores del genoma del baculovirus. Los promotores de genes de baculovirus se dividen en promotores inmediatos tempranos, tempranos, tardíos y muy tardíos, según su inicio de transcripción en el ciclo de infección. Entre ellos, como su nombre indica, los promotores génicos *inmediatos-tempranos (ie)* se activan inmediatamente después de la infección viral y permanecen activos durante todo el ciclo de infección. Sin embargo, los promotores génicos tardíos o muy tardíos, como *la polihedrina*, permanecen en silencio hasta que el virus alcanza la fase tardía de la infección.

[0285] Para aprovechar esta amplia gama de selección de promotores del genoma de baculovirus, se ensayó el promotor *inmediato-temprano1* (IE1) para el HBoV1-NS1. El elemento

potenciador de la transcripción *hr5*, que ha mostrado aumentar los niveles de expresión en las células Sf9, se incluyó antes del promotor IE1. Esto generó un BEV recombinante que codifica HBoV1-NS1 bajo el control del promotor *inmediato-temprano1* (IE1) del AcMNPV precedido por el elemento potenciador transcripcional *hr5* del AcMNPV, como se muestra en la **FIG. 9C**.

[0286] Las células Sf9 fueron coinfectadas con BEV que codificaban FVIIIIXTEN ITR de HBoV1 y HBoV1-NS1 dirigido por *hr5.IE1* a diferentes MOI manteniendo la proporción constante 1:10, basándose en los resultados obtenidos en la **FIG. 10C**. Como control positivo, se incluyó el BEV HBoV1-NS1 dirigido por *polihedrina* y se ensayó de nuevo en el mismo conjunto de experimentos. Más concretamente, se coinfectaron $\sim 2,0 \times 10^6$ /mL células Sf9 con preparaciones madre tituladas (P2) de BEV AcBIVVBac.mTTR.FVIIIIXTEN.HBoV1.ITRs^{Tn7} (**FIG. 10B**) a una MOI de 0,1, 0,3, 0,5, 1,0, 3,0, 5,0 pfu/célula con BEV AcBIVVBac.Polh.HBoV1-NS1^{Tn7} O AcBIVVBac.hr5.IE1.HBoV1-NS1^{Tn7} a una MOI de 0,01, 0,03, 0,05, 0,1, 0,3, 0,5 pfu/célula para mantener la proporción constante de 1:10 (**FIG. 11B**). El procedimiento restante se siguió como se ha descrito anteriormente (véase el Ejemplo 6).

[0287] Las fracciones de elución final se analizaron en electroforesis en gel de agarosa del 0,8 al 1,2 % para determinar la productividad del ADNce. La coinfección de HBoV1-NS1 dirigida por *polihedrina* mostró niveles crecientes de productividad de ADNce de FVIIIIXTEN con incrementos en las MOI, lo que confirma aún más la reproducibilidad del enfoque DosBAC para la producción de ADNce. Sin embargo, sorprendentemente, la coinfección de HBoV1-NS1 dirigida por *hr5.IE1* mostró niveles apenas detectables de ADNce de FVIIIIXTEN sin aumento aparente de la productividad con aumentos de MOI, como se observa en la **FIG. 10C**.

[0288] Estos datos sugieren que el inicio temprano de HBoV1-NS1 puede no ser crítico para rescatar el ADNce de FVIIIIXTEN con ITR de HBoV1. En cambio, es posible que se requieran niveles más altos de expresión más tarde en la infección para el rescate eficiente y la productividad de ADNce de FVIIIIXTEN con ITR de HBoV1. Estos resultados confirman además el requisito de niveles más altos de HBoV1-NS1 para la replicación de ADNce de FVIIIIXTEN dependiente de HBoV1-ITR en células Sf9.

[0289] En conclusión, estos experimentos han mostrado que el enfoque DosBAC sí demuestra el concepto de producción de ADNce a partir de dos BEV recombinantes que codifican FVIIIIXTEN con ITR de HBoV1 y/o transgén NS1. Estos experimentos también demuestran la importancia de la proporción óptima de MOI y/o del promotor para lograr una mayor productividad de ADNce de FVIIIIXTEN en células Sf9.

Ejemplo 9: Línea Celular Estable FVIIIIXTEN ITR de HBoV1

[0290] Se planteó la hipótesis de que el genoma de la célula de insecto podría modificarse potencialmente para producir ADNce para aplicaciones terapéuticas tras la infección por baculovirus. Para ensayar esta hipótesis, se utilizaron plásmidos que codifican el marcador de resistencia a la neomicina (pUC57.HR5.IE1.NeoR.P10PAS: SEQ ID NO: 7) (**FIG. 12A**) o la

proteína verde fluorescente mejorada (eGFP) (pUC57.HR5.IE1.eGFP.P10PAS: SEQ ID NO: 8) (**FIG. 12B**) bajo el control del promotor inmediato temprano (*ie1*) del AcMNPV precedido por el elemento potenciador transcripcional *hr5* y seguido por la señal de poliadenilación *p10* de AcMNPV fue sintetizado de GenScript® (Piscataway, NJ)

[0291] Estos plásmidos fueron cotransfectados con el plásmido que codifica FVIIIXTEN con ITR de HBoV1 (Sf.mTTR.FVIIIXTEN.HBoV1.ITRs) (**FIG. 12C**) en células Sf9 utilizando un método de transfección con fosfato de calcio modificado. A las 24 horas después de la transfección, las células se visualizaron bajo el microscopio de fluorescencia para determinar la eficiencia de transfección y los resultados mostraron >80 % de células GFP+, lo que sugiere una mayor eficiencia de transfección. A las 72h después de la transfección, las células se seleccionaron con el antibiótico G418 (Sigma Aldrich) suspendido en medio completo TNMFH (Medio de Insecto de Grace suplementado con FBS al 10 % + Plurónico F68 al 0,1 %) a una concentración final de 1,0 mg/mL. Después de aproximadamente una semana de selección, se recuperó un ~50 % de células transformadas, lo que sugiere que el marcador resistente a la neomicina se integró de forma estable en esta población celular. Las células supervivientes se retiraron del medio de selección y se alimentaron con un medio TNMFH completo fresco hasta el crecimiento de confluencia. Las células confluentes se expandieron progresivamente como un cultivo adherente en recipientes de cultivo más grandes a medida que continuaban dividiéndose. Posteriormente, la población de células policlonales se adaptó al cultivo en suspensión mediante el crecimiento en frascos agitados durante un subcultivo en medio TNMFH completo y un subcultivo en medio ESF-921 suplementado con FBS al 10 %. Finalmente, las células se adaptaron a ESF-921 sin suero en matraces de agitación como cultivos en suspensión. Estos cultivos en matraces agitados se mantuvieron de forma rutinaria en medio ESF-921 sin suero con subcultivos cada cuatro días y se monitorizó el crecimiento celular.

Ejemplo 10: Purificación de ADNce de FVIIIXTEN

[0292] En el sistema baculovirus-célula de insecto, el BEV recombinante entrega el gen de interés bajo un fuerte promotor y proporciona el complejo transcripcional esencial para la replicación del virus en las células de insecto. Normalmente, el genoma de ADN de baculovirus se replica en el núcleo y produce varias decenas de millones de partículas virales de progenie, cada una de las cuales contiene un genoma de ADN de longitud completa. Se ha demostrado que el ADN genómico baculoviral se copurifica con el ADNce mientras se aísla el ADN de las células de los insectos utilizando un método de purificación basado en el ADN plasmídico, como las columnas de gel de sílice. Las columnas comerciales del kit de ADN plasmídico no suelen estar diseñadas para separar el ADN en función de su peso molecular y, por tanto, normalmente todas las formas de ADN presentes en la muestra pueden unirse a estas columnas. Además, la capacidad de unión del ADN de gran peso molecular podría ser diferente a la del ADN de bajo

peso molecular y las columnas del kit basadas en el intercambio aniónico no están optimizadas en función de la eficiencia de unión de los diferentes tamaños de ADN.

[0293] Se planteó la hipótesis de que los ADN de alto peso molecular (>20 kb) observados en las preparaciones de ADNce eran probablemente ADN genómico baculoviral y/o de células Sf9 que se copurificaron con el ADNce de FVIIIIXTEN de bajo peso molecular (~8,5 kb) (véanse, p. ej., las **FIGs. 8C, 10C y 11C**).

[0294] Anteriormente, se empleó un enfoque indirecto para reducir el ADN baculoviral sometiendo a *knock-out* un gen de la cápside baculoviral, como el VP80, que es necesario para la producción de la progenie infecciosa del virus. Este enfoque mostró una reducción significativa del ADN baculoviral en la preparación de ADNce obtenida del BEV con *knock-out* (véase la Solicitud de Patente de EE. UU. No. 63/069.115). Aunque este enfoque fue eficiente en la reducción de la contaminación por ADN baculoviral, no pudo reducir los ADN genómicos celulares, que estaban presentes en una cantidad significativa (~60 %) de los ADN totales obtenidos de los sedimentos de células infectadas.

[0295] Por lo tanto, se empleó un enfoque directo para separar el ADNce de FVIIIIXTEN del resto de los ADN no deseados y se demostró que se podía obtener eficientemente el FVIIIIXTEN purificado (>95 % de pureza) a partir de la preparación de ADN total del sedimento de células infectadas. Este novedoso enfoque aprovecha la electroforesis preparativa, que se utiliza ampliamente para separar diferentes moléculas de proteínas según su tamaño y carga. Véase, p. ej., Michov, B. (2020) Electrophoresis. Berlín, Boston: De Gruyter, p. 405-424. Por ejemplo, la Bio-Rad Modelo 491 prep cell u otras unidades similares pueden utilizarse para separar moléculas complejas en función de su tamaño.

[0296] El flujo de trabajo completo de la purificación de ADNce se muestra en la **FIG. 13**, donde el proceso comienza con el aumento de escala del cultivo de células Sf9 de 0,5L a 1,5L o un volumen mayor en un medio de cultivo de células de insectos sin suero (**FIG. 13A**). Una vez alcanzada la densidad celular deseada de $\sim 2,5 \times 10^6$ /mL, normalmente después de 2 días de incubación con una densidad de siembra de $\sim 1,3 \times 10^6$ /mL, las células se infectan con BEV UnBAC o DosBAC (dependiendo del enfoque utilizado para la producción de ADNce) a una MOI optimizada y se dejan incubar las células en un incubador con agitación a 28 °C hasta que la viabilidad alcance el ~60-70 %, lo que suele tardar aproximadamente 4 días (**FIG. 13B**). Una vez que la viabilidad alcanza el ~70 %, las células se recogen y se procesan para la purificación del ADN total mediante columnas de kits cromatográficos de intercambio aniónico, como el kit de purificación PureLink HiPure Expi Plasmid Gigaprep (Invitrogen), de acuerdo con las instrucciones del fabricante. Una alícuota del material de ADN purificado se comprueba en una electroforesis en gel de agarosa del 0,8 al 1,2 % para determinar la productividad e integridad del ADN (**FIG. 13C**).

[0297] A continuación, el material purificado se carga en una Unidad de Electroforesis en Gel de Agarosa Preparativo, que contiene un gel de agarosa preparativo al 0,5 % y un gel de agarosa

concentrador al 0,25 %, montados según las instrucciones del fabricante. Las muestras se ejecutan a bajo voltaje (~40 voltios constantes) a 4 °C durante 6-7 días con una tasa de flujo de recirculación de tampón de ~50 mL/min y la tasa de tampón de elución de 50 µL/min para recoger cada fracción a 70-80 min en la cámara de recogida de fracciones. Después de la electroforesis de elución continua, se comprueban 20 µl de cada fracción en una electroforesis en gel de agarosa del 0,8 al 1,2 % para determinar la pureza del ADNce de FVIIIXTEN (**FIG. 13D**). Las fracciones deseadas se combinan para precipitar con NaOAc 3M, pH 5,5 y EtOH al 100 % a -20C durante 1-2h. Finalmente, el ADNce de FVIIIXTEN precipitado se sedimenta a alta velocidad y se lava una vez con EtOH al 70 % antes de resuspenderlo en el tampón TE, pH 8,0. El ADNce de FVIIIXTEN purificado se comprueba de nuevo en una electroforesis en gel de agarosa del 0,8 al 1,2 % para confirmar la pureza y la integridad antes de inyectarlo en los animales para los estudios de eficacia *in vivo* (**FIG. 13E**).

Ejemplo 11: Eficacia *in vivo* de FVIIIXTEN ITR de HBoV1

FVIIIXTENss ITR de HBoV1 (ADN monocatenario)

[0298] Se planteó la hipótesis de que la horquilla formada dentro de la región ITR permite mayores niveles de expresión transgénica persistente a largo plazo. Para investigar la funcionalidad de las ITR de HBoV1 *in vivo*, se ensayó el ADN monocatenario (ADNss) que comprendía FVIIIXTEN humano optimizado en codones con ITR de HBoV1 preformadas en ratones hFVIII^{R593C^{+/+}}/HemA. Estos ratones contienen un transgén *FVIII-R593C* humano, diseñado con el promotor de la albúmina murina (*Alb*) que dirige la expresión de un ADNc del factor VIII de coagulación humano (FVIII) alterado que alberga una mutación que se observa con frecuencia en pacientes con hemofilia A leve. Estos ratones también portan un *knock-out* del gen *FVIII* y son defectuosos para la proteína FVIII endógena. Estos ratones dobles mutantes toleran la inyección de FVIII humano y no tienen actividad de FVIII. Producen muy pocos anticuerpos inhibidores y carecen de células T o B que respondan al FVIII tras el tratamiento con FVIII humano. El ratón hFVIII^{R593C^{+/+}}/HemA se describe con más detalle en Brill, et al. (2006) *Thromb. Haemost.* 95(2): 341-7.

[0299] El FVIIIXTEN monocatenario (FVIIIXTENss) con ITR de HBoV1 preformadas se generó desnaturalizando los productos de fragmentos de ADN bicatenarios (casete de expresión del FVIII y esqueleto del plásmido) de la digestión con PvuII a 95 °C y enfriando después a 4 °C para permitir que las secuencias palindrómicas de ITR se plegaran. A continuación, el FVIIIXTENss se inyectó sistémicamente *mediante* inyecciones hidrodinámicas en la vena de la cola a 10 µg o 40 µg/ratón, lo que equivale a 400 µg o 1.600 µg/kg, respectivamente. Se recogieron muestras de plasma de los ratones inyectados con un intervalo de 7 días durante 5,5 meses. La actividad plasmática de FVIII se midió mediante el ensayo cromogénico Chromogenix Coatest® SP Factor VIII, según las instrucciones del fabricante.

[0300] La actividad plasmática del FVIII normalizada respecto al porcentaje de lo normal para los animales a los que se inyectó FVIIIXTENss se muestra en la **FIG. 14A**. Los resultados mostraron una respuesta dependiente de la dosis en los ratones HemA en el transcurso de 5,5 meses con niveles suprafisiológicos (>1.000 % de lo normal) de la expresión del FVIII en ambas dosis ensayadas. Sin embargo, hubo una caída inicial en la expresión del FVIII observada hasta el día 56, pero luego los niveles se estabilizaron hasta el día 168, lo que sugiere la expresión persistente de FVIIIXTENss flanqueado por las ITR de HBoV1 del hígado de los animales inyectados. Así, estos resultados validan la funcionalidad de las ITR de HBoV1 para la expresión persistente a largo plazo de FVIIIXTEN *in vivo*.

FVIIIXTENce IYT de HBoV1 (ADN con extremos cerrados)

[0301] Para validar la funcionalidad de las ITR de HBoV1 en el ADNce, el FVIIIXTENce purificado a partir de los sedimentos de células Sf9 infectadas, como se ha descrito anteriormente, se inyectó sistémicamente *mediante* inyecciones hidrodinámicas en la vena de la cola en ratones hFVIIR593C^{+/+}/HemA a 0,3 µg, 1,0 µg o 2,0 µg/ratón, lo que equivale a 12 µg, 40 µg y 80 µg/kg, respectivamente. Se recogieron muestras de plasma de los ratones inyectados con un intervalo de 7 días y se midió la actividad de FVIII mediante el ensayo cromogénico, como se ha descrito anteriormente.

[0302] La actividad plasmática del FVIII normalizada respecto al porcentaje de lo normal para los animales a los que se inyectó FVIIIXTENce se muestra en la **FIG. 14B**. Los resultados de este estudio muestran una respuesta dependiente de la dosis en los ratones HemA con niveles suprafisiológicos (>500 % de lo normal) de expresión del FVIII observados en la dosis más alta ensayada hasta el día 56 después de la inyección. Curiosamente, se alcanzó un nivel similar de expresión cuando se inyectó a los ratones FVIIIXTENss a 1.600 µg/kg, que es al menos 20 veces mayor que la dosis de FVIIIXTENce (80 µg/kg) (**FIG. 14A-14B**). Estos datos sugieren que el ADNce proporciona niveles más altos de expresión del FVIII en comparación con la forma ADNss.

[0303] En conclusión, estos estudios *in vivo* validan la funcionalidad de las ITR de HBoV1 tanto en forma de ADNss como de ADNce y demuestran que las ITR de HBoV1 pueden utilizarse para producir ADNce funcional que codifica un transgén de interés en el sistema de baculovirus células de insecto.

Ejemplo 12: Mejora de la pureza del vector de ADNce utilizando CRISPR Cas *knock out* de VP80 en BEV NS1 de HBoV1

[0304] La NS1 de HBoV1 expresada bajo el promotor de *polihedrina* del AcMNPV fue efectivamente capaz de rescatar el FVIIIXTEN flanqueado por ITR de HBoV1 y demuestra el concepto de producción de ADNce con ITR de HBoV1 en el sistema de baculovirus. Sin embargo, se observaron niveles significativos de contaminación por ADN baculoviral (ADNv) en las preparaciones de ADNce, presumiblemente debido a la mayor carga de virus requerida en comparación con el BEV Rep de AAV2 para lograr una mayor productividad de ADNce. El ADN

de alto peso molecular (>20kb) observado en estas preparaciones de ADNce (**FIG. 8C, FIG. 10C**) era probablemente el ADN genómico baculoviral que se copurificó con el ADNce de bajo peso molecular (~8 kb).

[0305] Para reducir la contaminación del ADN baculoviral en las preparaciones de ADNce, se implementó un enfoque indirecto de un proceso de *knock-out* de VP80, un gen esencial del genoma del baculovirus que se requiere para producir partículas de virus infecciosas en células de insectos (Sf9). Se sometió VP80 a *knock-out* en los tres BEV de NS1 (**FIGs. 9B, 9C y 9D**) utilizando el sistema Alt-R CRISPR-Cas9 (véase la Solicitud de Patente de EE. UU. No. 63/069.115). Este enfoque reduce potencialmente el número de partículas virales de progenie y, en última instancia, la contaminación de ADN baculoviral en las preparaciones de ADNce.

Knock-out por CRISPR-Cas9 del gen VP80 del AcMNPV:

[0306] Los BEV recombinantes que codifican NS1 de HBoV1 bajo el promotor de la polihedrina del AcMNPV (**FIG. 9B**) o OpIE2 de OpMNPV (**FIG. 9C**) fueron seleccionados para someter el gen vp80 a *knock-out* mediante el sistema CRISPR-Cas9, como se describió anteriormente (véase, p. ej., la Solicitud Internacional No. PCT/US2021/047202).

[0307] Brevemente, se diseñaron dos ARNcr dirigidos a la secuencia codificante y se utilizaron para generar ARNsg funcionales utilizando el sistema Alt-R CRISPR-Cas9 (Integrated DNA Technology™), de acuerdo con las instrucciones del fabricante. Cada ARNsg se cotransfectó con la nucleasa SpCas9 y ADN de bácmido AcBIVVBac.Polh.HBoV1.NS1^{Tn7} o AcBIVVBac.OpIE2.HBoV1.NS1^{Tn7} en células Sf9, sembradas a 0,5 x 10⁶ por mL en frascos T25 en medio ESF-921 libre de suero, utilizando el reactivo de transfección Cellfectin® (Invitrogen™). A los 4-5 días después de la transfección, las células se visualizaron bajo el microscopio de fluorescencia y los resultados mostraron un ~10 % de células RFP+ para ambas dianas de ARNsg. En la **FIG. 15** se muestran ejemplos de imágenes de microscopía de fluorescencia de células infectadas. Las células infectadas con BEV AcBIVVBac.Polh.HBoV1.NS1^{Tn7} en Cas9 solo mostraron una progresión de la infección como se esperaba (**FIG. 15A**), sin embargo, en contraste, las células tratadas con ARNsg.VP80.T1 (**FIG. 15B**) o ARNsg.VP80.T2 (**FIG. 15C**) mostraron una infección restringida a células individuales, lo que probablemente se deba al *knockout* de VP80.

[0308] Para determinar las indels inducidas por cada ARNsg, se recogió la progenie de baculovirus y se purificó en placa en una línea celular de complemento Sf.39K.VP80, como se ha descrito anteriormente. Jarvis *et al.* (2014), Methods Enzymol., 536: 149-163. A los 5-6 días después de la infección, doce clones RFP+ purificados en placa se amplificaron a P1 en células Sf.39K.VP80 sembradas a 0,5 x 10⁶ por mL en matraces T25 en medio ESF-921 suplementado con FBS al 10 %. La observación al microscopio de fluorescencia de los clones amplificados mostró un ~80 % de células RFP+, lo que sugiere que la línea celular Sf.39K.VP80 fue capaz de complementar la función de VP80 en trans para la producción de virus de progenie. Cada BEV clonal se recogió mediante centrifugación a baja velocidad y el sedimento celular se utilizó

entonces para el aislamiento del ADN total mediante el kit de aislamiento de ADN genómico DNeasy Blood and Tissue de Qiagen (No. de catálogo 69506), de acuerdo con las instrucciones del fabricante. El ADN resultante se utilizó como molde para la amplificación por PCR de cada secuencia diana con cebadores específicos para la secuencia codificante de vp80 de AcMNPV. Los amplímeros de la PCR se purificaron en gel y se secuenciaron directamente a través de la instalación de secuenciación Genewiz. Las secuencias resultantes fueron analizadas por el programa TIDE (Tracking of Indels by DEcomposition) (tide.deskgen.com) utilizando los ajustes por defecto para determinar las indels inducidas por cada ARNsg. Los análisis TIDE mostraron mutaciones de cambio de marco en el Clon no. 4 de BEV AcBIVVBac.Polh.HBoV1.NS1^{Tn7} tratado con ARNsg.T1 con las mayores deleciones (97,1 %) -15 pb (**FIG. 16A**) y el Clon no. 4 BEV AcBIVVBac.OpIE2.HBoV1.NS1^{Tn7} con las mayores deleciones (37,4 %) -4 pb y (26,9 %) -3 pb (**FIG. 16B**) en la secuencia codificante de vp80 sin inserciones detectables. Cada clon se amplificó a P2 para generar una preparación madre de trabajo de BEV seguido de la titulación en células Sf.39K.VP80, como se ha descrito anteriormente. Jarvis *et al.* (2014), *Methods Enzymol*, 536: 149-163. La preparación madre de trabajo titulada de BEV vp80KO se utilizó entonces para la coinfección en el sistema DosBAC para la producción del vector de ADNce de FVIIIIXTEN HBoV1.

Producción de ADNce de FVIIIIXTEN humano utilizando BEV vp80KO:

[0309] Se sembraron aproximadamente $2,0 \times 10^6$ células en 100 mL de medio ESF-921 sin suero y se coinfectaron con preparaciones madre de trabajo tituladas PP1P2 de BEV AcBIVVBac.FVIIIIXTEN.HBoV1.ITRs^{Tn7} y AcBIVVBac.Polh.HBoV1.NS1ΔVP80^{Tn7} o AcMNPV.OpIE2.HBoV1.NS1Δ VP80^{Tn7} a una MOI de 1,0, 2,0, 3,0, 4,0 y 5,0 pfu/célula. En cada caso, no se eliminó el inóculo de virus y las células se incubaron hasta que la viabilidad alcanzó el 60-70 % en un incubador con agitación a 28 °C. A las ~96 horas después de la infección, se recogieron las células infectadas y los sedimentos se procesaron para el aislamiento de ADNce de FVIIIIXTEN HBoV1 mediante el kit de aislamiento de ADN PureLink Maxi Prep (Invitrogen), de acuerdo con las instrucciones del fabricante.

[0310] Las fracciones de elución final se analizaron mediante electroforesis en gel de agarosa del 0,8 al 1,2 % para determinar la productividad y la pureza del ADNce. Los análisis en gel de agarosa mostraron una contaminación de ADN baculoviral (ADNv) de alto peso molecular (>20kb) muy baja a no detectable en los BEV vp80KO que expresaban NS1 de HBoV1 bajo el promotor de polihedrina del AcMNPV o OpIE2 de OpMNPV (**FIG. 16C**). Esto sugiere que el enfoque vp80KO fue capaz de reducir el ADN baculoviral contaminante y simultáneamente mejorar el rendimiento de ADNce de FVIIIIXTEN HboV1 cuando las células fueron coinfectadas a unas MOI de 2,0, 3,0 o 4,0 pfu/célula (FIG. 11C, FIG. 16C).

Ejemplo 13: Producción del vector de ADNce de FVIIIIXTEN HBoV1 a partir del sistema DosBAC

[0311] La inestabilidad genética es una de las principales preocupaciones en el campo de la baculovirología y especialmente a lo largo de varios subcultivos de baculovirus recombinantes en células Sf9. Además, los genomas de los baculovirus contienen varias regiones homólogas (hr) que son propensas a recombinarse a lo largo de subcultivos en células Sf9 y pueden potencialmente perder el transgén en los BEV recombinantes. Las repeticiones terminales invertidas (ITR) son también secuencias de repetición palindrómica y pueden potencialmente recombinarse en diferentes loci del genoma del baculovirus, considerando el gran tamaño del ADN baculoviral. Por lo tanto, para determinar la estabilidad genética del BEV recombinante que codifica un gen FVIIIIXTEN bajo el promotor mTTR específico del hígado con ITR de HBoV1 WT, el BEV se amplificó secuencialmente infectando células Sf9 a una MOI de 0,1 pfu/célula, como se describió previamente. Jarvis *et al.* (2014), *Methods Enzymol*, 536: 149-163. Las BEV recombinantes resultantes se ensayaron para la producción de ADNce de FVIIIIXTEN HBoV1 utilizando el sistema DosBAC (véase la **FIG. 17A** y la construcción representada en la **FIG. 17B**).

[0312] Se sembraron aproximadamente $2,0 \times 10^6$ células en 100 mL de medio ESF-921 libre de suero y se coinfectaron con preparaciones madre de trabajo tituladas (P3 o P4) de BEV AcBIVVBac.mTTR.FVIIIIXTEN. HBoV1.ITRs^{Tn7} y AcBIVVBac.Polh.HBoV1.NS1^{Tn7} a una MOI de 1,0, 2,0, 3,0, 4,0 y 5,0 pfu/célula. En cada caso, no se eliminó el inóculo de virus, y las células se incubaron hasta que la viabilidad alcanzó el 60-70 % en un incubador con agitación a 28 °C. A las 96 horas después de la infección, se recogieron las células infectadas y los sedimentos se procesaron para el aislamiento del ADNce de FVIIIIXTEN HBoV1 mediante el kit de aislamiento de ADN PureLink Maxi Prep (Invitrogen), de acuerdo con las instrucciones del fabricante.

[0313] Las fracciones de elución final se analizaron mediante electroforesis en gel de agarosa del 0,8 al 1,2 % para determinar la productividad del ADNce. El análisis en gel de agarosa, como se muestra en la **FIG. 17C**, mostró niveles casi equivalentes de productividad de ADNce de FVIIIIXTEN HBoV1 con los BEV P3 o P4 (y P5, datos no mostrados), lo que sugiere que los BEV recombinantes que codifican FVIIIIXTEN ITR de HBoV1 son genéticamente estables en subcultivos tardíos en células Sf9.

Ejemplo 14: Producción del vector de ADNce de FVIIIIXTEN HBoV1 a partir del sistema UnBAC

[0314] Se ha mostrado que el sistema HBoV1 UnBAC produce el vector de ADNce FVIIIIXTEN HBoV1 en células Sf9 (véase, p. ej., la **FIG. 8C**). Sin embargo, la prueba de concepto se logró utilizando BEV recombinantes policlonales. Para apoyar producciones a mayor escala, es necesario generar BEV clonales. En consecuencia, en este estudio, los BEV policlonales HBoV1 UnBAC se purificaron en placa y se amplificaron en células Sf9 (véase la **FIG. 18A**). Estos BEV clonales UnBAC fueron entonces cribados para determinar la producción del vector de ADNce FVIIIIXTEN HBoV1 en células Sf9.

[0315] La purificación en placa y la amplificación de los BEV recombinantes de HBoV1 UnBAC se realizó como se ha descrito anteriormente. Jarvis *et al.* (2014), *Methods Enzymol*, 536: 149-163. Seis clones purificados en placa se amplificaron a P2 infectando $\sim 1,0 \times 10^6$ /mL de células Sf9 en 100mL de medio ESF-921 suplementado con suero fetal-bovino al 10 % y se incubaron durante 4-5 días o hasta que la viabilidad celular alcanzó el 60-70 % en un incubador con agitación a 28 °C. A los 4-5 días después de la infección, se recogió el sobrenadante libre de células y se almacenó como preparaciones madre de trabajo P2 y los sedimentos celulares se procesaron para el aislamiento de ADNce de FVIIIIXTEN HBoV1 mediante el kit de aislamiento de ADN PureLink Maxi Prep (Invitrogen), de acuerdo con las instrucciones del fabricante.

[0316] Las fracciones de elución final se analizaron mediante electroforesis en gel de agarosa del 0,8 al 1,2 % para determinar la productividad del vector de ADNce FVIIIIXTEN HBoV1. La **FIG. 18C** muestra el análisis en gel de agarosa de HBoV1 UnBAC que codifica FVIIIIXTEN con ITR de HBoV1 y HBoV1-NS1 dirigido por polihedrina (construcción representada en la **FIG. 18B**). Los resultados mostraron un grado variable de productividad de ADNce de HBoV1 para diferentes clones, siendo el clon no. 2 y el clon no. 4 los mayores productores de ADNce de HBoV1 en comparación con otros clones ensayados (**FIG. 18C**). Este resultado muestra la variabilidad de los diferentes clones de baculovirus obtenidos de la misma preparación madre y destaca la importancia de utilizar BEV recombinantes clonales para la fabricación de ADNce a gran escala.

[0317] Para determinar la productividad óptima del clon de BEV HBoV1 UnBAC, se infectaron aproximadamente $2,0 \times 10^6$ células con una preparación madre de trabajo titulada (P2) del clon no. 5 de BEV HBoV1 UnBAC a una MOI de 0,1, 0,2, 0,3, 0,4, 0,5, 1,0, 2,0, 3,0, 4,0 o 5,0 pfu/célula. En cada caso, no se eliminó el inóculo del virus y las células se incubaron hasta que la viabilidad celular alcanzó el 60-70 % en un incubador con agitación a 28 °C. A las ~ 96 h después de la infección, se recogieron las células infectadas y los sedimentos se procesaron para el aislamiento del ADNce de FVIIIIXTEN HBoV1 mediante el kit de aislamiento de ADN PureLink Maxi Prep (Invitrogen), de acuerdo con las instrucciones del fabricante. Las fracciones de elución final se analizaron mediante electroforesis en gel de agarosa del 0,8 al 1,2 % para determinar la productividad del ADNce de FVIIIIXTEN HBoV1.

[0318] El análisis en gel de agarosa mostró una banda de ADN correspondiente al tamaño del ADNce de FVIIIIXTEN HBoV1 ($\sim 8,5$ kb) en todas las dosis ensayadas, con una productividad creciente a medida que aumenta la MOI. Este resultado fue contrario a la productividad de ADNce obtenida con ITR de AAV2 UnBAC, donde se observó una reducción de la productividad con el aumento de la carga viral. Este enfoque HBoV1 UnBAC demuestra el concepto de producción de ADNce a partir de un único BEV recombinante que codifica FVIIIIXTEN con ITR de HBoV1 y transgenes NS1. También muestra la viabilidad y funcionalidad de múltiples transgenes insertados en diferentes loci en un vector lanzadera de baculovirus (BIVVBac).

Ejemplo 15: Eficacia *in vivo* de ADNss frente a ADNce de FVIII^{IX}TEN HBoV1

FVIII^{IX}TENss ITR de HBoV1 (ADN monocatenario)

[0319] Se planteó la hipótesis de que la horquilla formada dentro de la región ITR de HBoV1 dirige la expresión persistente a largo plazo del transgén a niveles más altos. Para validar la funcionalidad de las ITR de HBoV1 *in vivo*, se ensayó en ratones hFVIII^{R593C}+/+/HemA el ADN monocatenario (ADNss) que comprende FVIII^{IX}TEN humano optimizado en codones (FVIII^{IX}TENss) con ITR de HBoV1 preformadas.

[0320] El FVIII^{IX}TENss con ITR de HBoV1 preformadas se generó desnaturalizando los productos de fragmentos de ADN bicatenario (ADNds) (casete de expresión del FVIII y esqueleto del plásmido) de la digestión con PmlI a 95 °C y enfriando después a 4 °C para permitir que las secuencias palindrómicas de ITR se plegaran. El FVIII^{IX}TENss resultante se confirmó mediante electroforesis en gel de agarosa del 0,8 al 1,2 %. El análisis del gel mostró la mitad del tamaño del ADNds para el FVIII^{IX}TENss, lo que sugiere una formación eficiente de la horquilla (**FIG. 19A**). El FVIII^{IX}TENss se inyectó sistémicamente mediante inyecciones hidrodinámicas en la vena de la cola a 10 µg o 40 µg/ratón, lo que equivale a 400 µg o 1.600 µg/kg, respectivamente. Se recogieron muestras de plasma de los ratones inyectados a intervalos de 7 días durante 5,5 meses. La actividad plasmática de FVIII se midió mediante el ensayo cromogénico Chromogenix Coatest® SP Factor VIII, según las instrucciones del fabricante.

[0321] La actividad plasmática del FVIII normalizada respecto al porcentaje de lo normal para los animales a los que se inyectó FVIII^{IX}TENss se muestra en la **FIG. 19C**. Los resultados mostraron una respuesta dependiente de la dosis en los ratones HemA en el transcurso de 5,5 meses con niveles suprafisiológicos (>1.000 % de lo normal) de la expresión del FVIII en cohortes de dosis más altas. Sin embargo, hubo una caída inicial en la expresión del FVIII observada hasta el día 56 y los niveles se estabilizaron hasta el día 140, lo que sugiere una expresión persistente de FVIII^{IX}TENss flanqueado por ITR de HBoV1 del hígado. Estos resultados validan la funcionalidad de las ITR de HBoV1 para la expresión persistente a largo plazo de FVIII^{IX}TEN *in vivo*.

FVIII^{IX}TENce ITR de HBoV1 (ADN con extremos cerrados)

[0322] Existe una importante diferencia estructural entre el ADN con extremos cerrados (ADNce) y el ADN monocatenario (ADNss), donde el primero es bicatenario y el segundo es monocatenario, respectivamente. Esta diferencia puede influir en los niveles de expresión, así como en la estabilidad de la molécula de ácido nucleico. Este estudio muestra la funcionalidad de las ITR de HBoV1 en la forma de ADNce *in vivo*. Para ensayar esto, se obtuvo FVIII^{IX}TENce mediante el enfoque DosBac, como se describe en el Ejemplo 8, se purificó a partir de sedimentos de células Sf9 infectadas y se determinó la calidad mediante electroforesis en gel de agarosa al 0,8 al 1,2 %. El análisis del gel de agarosa mostró una pureza >90 % de FVIII^{IX}TENce sin ADN contaminantes detectables (**FIG. 19B**).

[0323] El FVIIIXTEN_{ce} resultante se inyectó sistémicamente mediante inyecciones hidrodinámicas en la vena de la cola en ratones hFVIII^{R593C}+/-HemA a 0,3 µg, 1,0 µg o 2,0 µg/ratón, lo que equivale a 12 µg, 40 µg y 80 µg/kg, respectivamente. Se recogieron muestras de plasma de los ratones inyectados a intervalos de 7 días y se midió la actividad del FVIII mediante el ensayo cromogénico, como se ha descrito anteriormente. La actividad plasmática del FVIII normalizada respecto al porcentaje de lo normal para los animales a los que se inyectó FVIIIXTEN_{ce} se muestra en la **FIG. 19C**.

[0324] Los resultados mostraron una respuesta dependiente de la dosis en ratones HemA con niveles suprafisiológicos (>500 % de lo normal) de expresión de FVIII a la dosis más alta (80 µg/kg) ensayada para FVIIIXTEN_{ce}. Los niveles más altos de expresión del FVIII para el ADN_{ce} fueron 2 veces menores que los niveles más altos alcanzados por el FVIIIXTEN_{ss} a 1.600 µg/kg. Sin embargo, el ADN_{ss} fue dosificado en cantidades mucho más altas para lograr estos altos niveles de expresión del FVIII. El ADN_{ce} parece proporcionar niveles más altos de expresión del FVIII por cantidad de dosificación. Por ejemplo, los niveles de expresión del FVIII para el ADN_{ss} a 400 µg/kg y el ADN_{ce} a 40 µg/kg fueron comparables (**FIG. 19C**). Estos estudios *in vivo* validan la funcionalidad de las ITR de HBoV1, ya sea en forma de ADN_{ss} o ADN_{ce}, y muestran que las ITR de HBoV1 pueden utilizarse para producir ADN_{ce} funcional que codifique el transgén de interés en el sistema de baculovirus células de insecto.

Ejemplo 16: Eficacia in vivo de ADN_{ce} de FVIIIXTEN HBoV1 monomérico frente a multimérico

[0325] Se ha mostrado que los genomas recombinantes de AAV persisten de forma episomal y se cree que su existencia episomal está correlacionada con la expresión de transgenes a largo plazo. Al parecer, estos genomas se originan mediante un proceso de circularización monomérica, que da lugar a un genoma circular de AAV de cabeza a cola. Sin embargo, con el tiempo, se produce una disminución de los intermedios circulares monoméricos en favor de concatámeros circulares de alto peso molecular. En Duan *et al.* (1998), J Virol. 72(11);8568-8577 se divulgan detalles adicionales. Actualmente, se sabe poco sobre la existencia episomal de ADN de extremos cerrados (ADN_{ce}) y las ventajas de las formas monoméricas sobre las concataméricas de ADN_{ce} *in vivo*.

[0326] Este estudio se realizó para determinar el impacto de las formas monoméricas frente a las multiméricas de ADN_{ce} *in vivo*, ensayando ambas formas de ADN_{ce} de FVIIIXTEN HBoV1 (FVIIIXTEN_{ce}) en ratones hFVIII^{R593C}+/-HemA mediante inyecciones hidrodinámicas en la vena de la cola.

[0327] Las formas monoméricas y multiméricas de FVIIIXTEN_{ce} se generaron por purificación PAGE, como se describió anteriormente (véase la Solicitud Internacional No. PCT/US2021/047218). La calidad de las formas concataméricas de FVIIIXTEN_{ce} se determinó mediante electroforesis en gel de agarosa del 0,8 al 1,2 % y los resultados mostraron que la

mayoría de las especies eran la forma monomérica o multimérica de FVIII_{XTEN}ce (**FIG. 20A**). El FVIII_{XTEN}ce monomérico o multimérico purificado se inyectó sistemáticamente en ratones hFVIII_{R593C}+/-HemA mediante inyecciones hidrodinámicas en la vena de la cola a 40 µg/kg. Se recogieron muestras de plasma de los ratones inyectados a intervalos de 7 días durante aproximadamente 3 meses. La actividad plasmática del FVIII se midió mediante el ensayo cromogénico Chromogenix Coatest® SP Factor VIII, según las instrucciones del fabricante.

[0328] La actividad plasmática del FVIII normalizada con respecto al porcentaje de lo normal para los animales a los que se inyectó FVIII_{XTEN}ce se muestra en la **FIG. 20B**. Los resultados no mostraron diferencias significativas en los niveles de expresión del FVIII entre las formas monoméricas o multiméricas de FVIII_{XTEN}ce en el transcurso de 3 meses. Estos datos sugieren que tanto las formas monoméricas como las multiméricas de FVIII_{XTEN}ce tienen una potencia y estabilidad *in vivo* comparables.

Ejemplo 17: Eficacia *in vivo* de ADNss de FVIII_{XTEN} HBoV1 mTTR frente a A1AT

[0329] El casete de expresión del FVIII_{XTEN} utilizado en los experimentos divulgados anteriormente contiene un promotor de mTTR y un elemento potenciador (V2.0, **FIG. 1**). Este promotor es específico de hígado de ratón, pero su expresión específica en el hígado no se ha estudiado en modelos de animales grandes ni en sujetos humanos. Por lo tanto, en este estudio, el casete de expresión FVIII_{XTEN} V3.0 (SEQ ID NO: 35) se generó reemplazando el promotor de mTTR y el elemento potenciador por el promotor de antitripsina alfa-1 (A1AT) específico del hígado humano (SEQ ID NO: 36) en el casete de expresión V2.0 (**FIG. 1**).

[0330] Para validar la funcionalidad del promotor mTTR frente al A1AT *in vivo*, se ensayó en ratones hFVIII_{R593C}+/-HemA el ADN monocatenario (ADNss) que comprende FVIII_{XTEN} humano optimizado en codones (FVIII_{XTEN}ss) con ITR de HBoV1 preformadas (**FIG. 21A**). El FVIII_{XTEN}ss con ITR de HBoV1 preformadas se generó desnaturalizando los productos de fragmentos de ADN bicatenario (ADNds) (casete de expresión de FVIII mTTR o A1AT y esqueleto del plásmido) de la digestión con PmlI a 95 °C y luego enfriando a 4 °C para permitir que las secuencias palindrómicas de ITR se plieguen. El FVIII_{XTEN}ss resultante se comprobó mediante electroforesis en gel de agarosa del 0,8 al 1,2 %. El análisis del gel mostró la mitad del tamaño del ADNds para el FVIII_{XTEN}ss, lo que sugiere una formación eficiente de la horquilla (**FIG. 21B**). El FVIII_{XTEN}ss se inyectó sistémicamente en ratones hFVIII_{R593C}+/-HemA mediante inyecciones hidrodinámicas en la vena de la cola a 10 µg/ratón. Se recogieron muestras de plasma de los ratones inyectados a intervalos de 7 días durante 5,5 meses. La actividad plasmática del FVIII se midió mediante el ensayo cromogénico Chromogenix Coatest® SP Factor VIII, según las instrucciones del fabricante.

[0331] La actividad plasmática del FVIII normalizada respecto al porcentaje de lo normal para los animales a los que se inyectó FVIII_{XTEN}ss se muestra en la **FIG. 21C**. Estos resultados mostraron niveles equivalentes de expresión de FVIII hasta el día 21 después de la inyección,

sugiriendo que no hay diferencia significativa en los niveles de FVIII_{XTEN} expresado por el promotor mTTR o A1AT en el modelo animal de ratones hFVIII^{R593C}^{+/+}/HemA.

SECUENCIAS

Tabla 2: Secuencias Adicionales de Nucleótidos y Aminoácidos

SEQ ID NO / Descripción	Secuencia de nucleótidos o aminoácidos
SEQ ID NO. 1: HBoV1 5' ITR	GTGGTTGTACAGACGCCATCTTGAATCCAATATGTCTGCCGGCTCAGTCATGCCTGCGCTGCGCGCAGCGCGCTGC GCGCGCGCATGATCTAATCGCCGGCAGACATATTGGATTCCAAGATGGCGTCTGTACAACCAC
SEQ ID NO. 2: HBoV1 3' ITR	TTGCTTATGCAATCGCGAACTCTATATCTTTTAAATGTGTTGTTGTTGTACATGCGCCATCTTAGTTTTATATCAGC TGGCGCCTTAGTTATATAACATGCATGTTATATAACTAAGGCGCCAGCTGATATAAACTAAGATGGCGCATGTACA ACAACAACACATTAAGATATAGAGTTTCGCGATTGCATAAGCAA
SEQ ID NO. 3: HBoV1-5'ITR- mTTR482- Intrón- coBDDFVIII TEN (V2.0)- WPRE- bGHPoliA- HBoV1-3'ITR	GTATACCTGCAGGCTAGCCACGTGTTGTTGTTGTACATGCGCCATCTTAGTTTTATATCAGCTGGCGCCTTAGTTAT ATAACATGCATGTTATATAACTAAGGCGCCAGCTGATATAAACTAAGATGGCGCATGTACAACAACAACACATTAA AAGATATAGAGTTTCGCGATTGCAAGCTTGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGGTCAAAGTGGCCCTT GGCAGCATTTACTCTCTCTATTGACTTTGGTTAATAATCTCAGGAGCACAAACATTCCTGGAGGCAGGAGAAGAAAT CAACATCTGGACTTATCTCTGGGCCTCTCCACCTTCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGG TCAAAGTGGCCCTTGGCAGCATTTACTCTCTCTATTGACTTTGGTTAATAATCTCAGGAGCACAAACATTCCTGGAG GCAGGAGAAGAAATCAACATCTGGACTTATCTCTGGGCCTCTCCACCGATATCTACCTGCTGATCGCCCCGCC CCTGTTCAAACATGTCCTAATACTCTGTCGGGGCAAAGGTCGGCAGTAGTTTTCCATCTTACTCAACATCCTCCAG TGTACGTAGGATCCTGTCTGTCTGCACATTTCTGAGAGCGAGTGTTCCGATACTCTAATCTCCCGGGCAAAGGTCG TATTGACTTAGGTTACTTATTCTCCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGC AGGGATCAGCAGCCTGGGTTGGAAGGAGGGGGTATAAAGCCCTTACCAGGAGAAGCCGTCACACAGATCCACAA GCTCCTGCTAGGAATTCAGGAGCACAAACATTCCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTATCCTCT GGGCCTCTCCACCGATATCTACCTGCTGATCGCCCGGCCCTGTTCAAACATGTCCTAATACTCTGTCTGGGGCAA AGGTCTGGCAGTAGTTTTCCATCTTACTCAACATCCTCCAGTGACGTAGGATCCTGTCTGTCTGCACATTTCTGAG AGCGAGTGTTCCGATACTCTAATCTCCCGGGCAAAGGTCGATTGACTTAGGTTACTTATTCTCCTTTTGTGACT AAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGGTATA AAAGCCCTTACCAGGAGAAGCCGTACACAGATCCACAAGCTCCTGCTAGAGTCGCTGCGCGCTGCCCTTCGCCCC GTGCCCCGCTCCGCGCCGCCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCACAGGTGAGCGGGCG GGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTATTGACGGCTTGTTCCTTTCTGTGGCTGCGTGAAA GCCTTGAGGGGCTCCGGGAAGGCCCTTGTGCGGGGGAGCGGCTCGGGGGTGCGTGCGTGTGTGTGCGTGGG AGCGCCGCTGCGGCTCCGCGCTGCCGCGGCTGTGAGCGCTGCGGGCGCGCGCGGGGCTTTGTGCGCTCCGAG TGTGCGCGAGGGGAGCGCGGCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGC GGGGTGTTGCGTGGGGGGTGAGCAGGGGGTGTTGGGCGCGTGGTGGGCTGCAACCCCCCTGCACCCCCCTCCC CGAGTTGCTGAGCACGGCCCGGCTTCCGGTGCGGGGCTCCGTACGGGGCTGGCGCGGGGCTCGCCGTGCCGGCGG GGGTGCGCGCAGGTGGGGGTGCGGGCGGGGCGGGGCGGCTCGGGCGGGGAGGGCTCGGGGAGGGGCGCGGCG GCCCCGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGAGCCATTGCCTTTATGGTAATCGTGCGAGAGGGCG CAGGGACTTCTTTGTCCAAATCTGTGCGGAGCCGAAATCGGGAGGCGCGCGCACCCCCCTAGCGGGCGCGG GGCGAAGCGGTGCGGCGCGGAGGAAATGGCGGGGAGGGCTTCGTGCGTGCCTGCGCGCGCGCTCCCTTC TCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGAGCGGCTGCCCTCGGGGGGAGCGGGGAGGGCGGGGTTCCGCTTC TGGCGTGTGACCGGCGGCTAGAGCCTCTGCTAACCTTGTCTTGCCTTCTTTTCTACAGCTCCTGGGCAA CGTGCTGGTTATTGTGCTGTCTCATATTTGGCAAAGAATTACTCGAGGCCACCATGCAGATTGAACTGTCCACTT GCTTCTTCTGTGCTCCTGCGGTTTTGCTTCTCGGCCACCGCCGGTATTACTTAGGTGCTGTGGAACAGAGCTGG GACTACATGCAGTCCGACCTGGGAGAAGTCCGCTGGACGCGAGATTCCACCTAGAGTCCCGAAGTCCCTCCCAT CAACACCTCCGTGGTCTACAAAAAGACCTGTTCTGTGGAGTTCACTGACCACCTTTTCAATATTGCCAAGCCGCGCC CCCCCTGGATGGGCTGCTTGGTCTACGATCCAAGCAGAGGTCTACGACACCGTGGTCATCACACTGAAGAACATG GCCTCACACCCGTTGCTGCTGATGCTGTGGAGTGCTCTACTGGAAGGCTCAGAGGGTGCCGAATATGATGACCA GACCAGCCAGAGGGAAAAGGAGGATGACAAAGTGTTCGCGGTGGCAGCCACACTTACGTGTGGCAAGTGCTGAAGG AAAACGGGCTATGGCGTCGGACCCCCATGCTGACCTACTCTACCTGTCCTATGTGGACCTTGTGAAGGATCTC AACTCGGGACTGATCGGCGCCCTTGGTGTCAGAGAAGGCAGCTGGCGAAGGAAAAGACTCAGACCCTGCACAA GTTCACTTGTGTTTGTGTTGCTGATGAAGGAAAGTCTGGCACTCAGAAACCAAGAACTCGCTGATGCAGGATA GAGATGCGGCCTCGGCCAGAGCCTGGCTAAATGCACACCGTCAACGGATATGTGAACAGGTGCTCCCTGGCCTC ATCGGCTGCCACAGAAAGTCCGTGTATTGGCATGTGATCGGCATGGGTACTACTCCGGAAGTGATAGTATCTTCT GGAGGGCCATACCTTCTTGGTGCGCAACCACAGACAGGCCTCGCTGGAATCTCGCTATCACTTTCTTGAAGTGC AGACCTCCTTATGGACCTTGGACAGTTCTGCTGTCTGTACATCAGCTCCCATCAGCATGATGGGATGGAGGCC

	TATGTCAAAGTGGACTCCTGCCCTGAGGAGCCACAGCTCCGGATGAAGAACAATGAGGAAGCGGAGGATTACGACGACGACCTGACTGACAGCGAAATGGACGTCGTGCGATTTCGATGACGACAACAGCCCCTCTTCATCCAAATTAGATCAGTGGCGAAGAAGCACCCCAAGACCTGGGTGCACTACATTGCCGCGGAGGAAGAGGACTGGGACTACGCGCCGCTGGTGCTGGCGCCAGACGACAGGAGCTACAAGTCCCAGTACCTCAACAACGGGCCGAGCGCATTGGCAGGAAGTACAAGAAAGTCCGCTTCATGGCCTACACTGATGAAACCTTCAAGACGAGGGAAGCCATCCAGCAGGAGTCAGGCATCCTGGGACCGTCTCTTTACGGCGAAGTCGGGGATACCCTGCTCATCATTTTTCAAGAACCAGGCATCGCGGCCCTACAACATCTACCCTACGCGGATCACAGACGTGCGCCCCGCTCTACTCCCGCCGGCTGCCAAGGGAGTGAAGCACCTGAAGGATTTTCCCATCCTGCCGGGAGAAATCTTCAAGTACAAGTGGACCGTGACTGTGGAAGATGGCCCTACCAAGTCGGACCCTCGCTGTCTGACCCGGTACTATTCTCGTTTGTGAACATGGAGCGCGACCTGGCCTCGGGGTGATTGGTCCGCTGCTGATCTGCTACAAGGAGTCCGTGGACCAGCGCGGGAACCAGATCATGTCCGACAAGCGCAACGTGATCCTGTTCTCTGCTTTGATGAAAACAGATCGTGGTACTTGACTGAGAATATCCAGCGGTTCTGCCAACCAGCGGGAGTGCAACTGGAGGACCCGGAGTTCAGGCCTCAAACATTATGCACTCTATCAACGGCTATGTGTTTCGACTCGCTCCAATGAGCGTGTGCTGCATGAAGTGGCATACTGGTACATTCTGTCCATCGGAGCCAGACCGACTTCTGTCCGTGTTCTTCTCCGGATAACCTTCAAGCATAAGATGGTGTACGAGGACACTCTGACCTCTTCCCATTTTCCGGAGAACTGTGTTTCATGTCAATGGAAAACCCGGGCTTGTGGATTCTGGGTTGCCATAACTCGGACTTCCGGAATAGAGGGATGACCCGCTGCTGAAAGTGTCCAGCTGTGACAAGAATACCGGCGATTACTACGAGGACAGCTATGAGGACATCTCCGCTTATCTGCTGTCCAAGAACAACGCCATTGAACCCAGGTCCTTCTCCAAAACGGTGCACCGACCTCCGAAAGCGCCACCCAGAGTCAGGACCTGGCTCGGAACCGGCTACCTCGGGCTCAGAGACACCGGGGACTTCCGAGTCCGCAACCCCGAGAGTGGACCCGGAATCCGAACAGCAACCTCAGGATCAGAAACCCCGGGAACCTCGGAATCCGCCACTCCCGAGTCGGGACCAGGCACCTCACTGAGCCTTCCGAGGGAAGCGCCCCGGATCCCTGCTGGATCCCTACCAGCACTGAAGAAGGCACCTCAGAATCCGCGACCCCTGAGTCCGGCCCTGGAAGCGAACC CGCCACCTCCGTTCCGAAACCCCTGGGACTAGCGAGAGCGCCACTCCGGAATCGGGCCCAGGAAGCCCTGCCGATCCCGACAGCAGGAGGGAAGCCCGCCGGGTACCCGACTTCCACTGAGGAGGGAGCCTCATCCCCCGCTGCTGAAGCGGCATCAAAGAGAGATCACCAGGACCACTCTCCAGTCCGATCAGGAAGAAATTGACTACGACGATACTATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGATTGAGGATGAGAACCAGTCCCCTCGGAGCTTTCAGAAGAAAACCCGCCACTACTTCATCGCTGCCGTGGAGCGGCTGTGGGATTACGGGATGTCCAGCTACCCGATGTGCTGCGGAATAGAGCGCAGTCAGGATCGGTGCCCCAGTTCAAGAAGGTCTGTTCGAAGAGTTACCCGACGGGTCTTCACTCAACCCCTGTACCGGGGCGAACTCAACGAACACCTGGGACTGCTTGGGCGTATATCAGGGCAGAAGTGGAAAGATAACATCATGGTCACTTCCGCAACCAGGCCTCCCGCCGCTACAGCTTCTACTCTTCACTGATCTCTACGAGGAAGATCAGCGGCAGGGAGCCGAGCCCCGGAAGAAGTTCGTCAAGCCTAACGAACTAAGACCTACTTTTGGAAAGTCCAGCATCATATGGCCCCGACCAAAGACAGTTTCGACTGTAAGGCC TGGGCC TACTTCTCCGATGTGGACCTGGAGAAGGACGTGCACCTCGGACTCATTGGCCCCGCTTGTGTGCCATAATACTCCTGAACCCCTGCTCACGGTCGCAAGTACAGTGCAGGAGTTGCGCCTCTTCTTCAACATCTTCGATGAAACAAAGTCTTGGTACTTTACTGAGAACATGGAACGCAATTGCAAGGACCCCTGCAACATCCAGATGGAAGATCCCACCTTCAAGGAAAATACCGGTTTCATGCCATTAACGGCTACATAATGGACACGTTGCCAGGACTGGTCATGGCCAGGACCAGAGAATCCGGTGGTATCTGCTCTCCATGGGCTCCAACGAAAACATTACAGCATTCATTTTTCCGGCCATGTGTTACCGTCCGGAAGAAGGAAGATACAAGATGGCTCTGTACAACCTCTACCCTGGAGTGTTCGAGACTGTGGAAATGCTGCCTAGCAAGGCCGGCATTGGAGAGTGGAAATGCCTGATCGGAGAGCATTGACACGCCGGAATGTCCACCCTGTTTCTTGTGTACTCCAACAAGTGCCAGACCCCGCTGGGAATGGCCTCAGGTCAATTAGGGATTTCAGATCACTGCTTCGGGCGAGTACGGGACCTAAGTTGGCCCCGCTGCACCTGCTGCTCCATCAATGGCTGGTCCACCAAGAAACCTTCTCTGGATTAAAGGTGGACCTCTGGCCCCAATGATTATTCACGGTATTAAGACCCAGGGTGGCCGACAGAAAGTCTCCTCACTCTACATCTCGCAATTATCATATAATGTACAGCCTGGATGGGAAGAAGTGGCAGACCTACCGGGGAACTCCACTGGAACGCTCATGGTGTTTTTCGGCAACGTGGACTCCTCCGGCATTAAAGCACAAACATCTTCAACCTCCGATCATTTGCTCGGTACATCCGGCTGCACCCAACTCACTACAGCATCCGGTCCACCCTGCGGATGGAACGATGGGTTGTGACCTGAACTCCTGCTCCATGCCCCCTGGGATGGAATCCAAGGCCATTAGCGATGCACAGATCACCGCCTCTTCACTTCAACCAACATGTTGCGGACCTGGTCCCCGTCGAAGGCCCGCTGCACCTCCAAGGTGCTCCAATGCGTGGCGGCTCAAGTGAACAACCCCAAGGAGTGGCTCCAGGTCGACTTCAAAAAGACCATGAAGGTACCCGAGTGACCACCCAGGGCGTGAAGTCCCTGCTGACCTCTATGTACGTTAAGGAGTTCTCATCTCCTCAAGCCAAGACGACATCAGTGGACCCCTGTCTTCCAAAACGGAAGTCAAAGTATTCCAGGGCAACCAAGGACTCTTCAACCCTGTGGTCAACAGCCTGGACCCCCATTGCTGACCCGCTACCTCCGCATCAACCCCAAAGCTGGGTCCACCAGATCGCACTGCGCATGGAGTCTTGGGATGCGAAGCCCAAGATCTGTACTAAGCGGCCGCTCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACATATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCATTTTCTCCTCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTGAGGCAACGTGGCGTGGTGTGCACTGTGTTTGTGACGCAACCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACATCATCGCCGCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATCCGTGGTGTGTGCGGGAAATCATGCTCTTTCCTTGGCTGCTCGCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCTTTCGCCCCAATCCAGCGGACCTTCTTCCCGCGGCTGCTGCGGCTCTGCGGCTCTTCCGCTCTCGCCTCAGACGAGTCGGATCTCCTTTGGGCCGCTCCCCGCTGCTAGGCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGCTGCCTTCTTGACCCTGGAAGGTGCCACTCCCCTGCTCTTCTAATAAAATGAGGAAATGTCATGCTGCTGAGTAGGTGTCTATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAAGACCATGGGCGCGCAGGCTGTGACGTGGTTGTACAGACGCCATCTTGGAAATCCAATATGCTGCCGGCGATTAGATCATGCGCGCGCAGCGCGCTGCGCGCAGCGCAGGCATGACTGAGCCGGCAGACATATTGGATTCCAAGATGGCGTCTGTACAACCACGTGCTTAAGCTGCAGACTAGTGAGCTCGTTAAC
SEQ ID NO. 4:	GCGGCCGCGGATCCGCCACCATGGCATTCAATCCGCCCGTAATACGCGCATTTTCAACCCGCTTTACGTATGTCTTTAAGTTTCCGTACCCTCAATGGAAGAGAAAGAGTGGCTACTGCACGCGTTGCTTGGCCACGGCACCGAGCAGTC

hr5.ie1.eGFP. p10PAS	TTTGT TTTGCAATATGATATCATACAATATGACTCATTTGTTTTTCAAACCGAACTTGATTTACGGGTAGAATTCT ACTTGTAAGCACAATCAAAAGATGATGTCATTTGTTTTTCAAACCTGAACCTCGCTTACGAGTAGAATTCTACGT GTAAACACAATCAAGAAATGATGTCATTTGTTATAAAATAAAAGCTGATGTCATGTTTTGCACATGGCTCATAAC TAAACTCGCTTACGGGTAGAATTCTACGCGCGTCGATGTCCTTTGTGATGCGCGCGACATTTTTGTAGGTTATTGAT AAAATGAACGGATACGTTGCCCCGACATTATCATTAATCCTTGGCGTAGAATTTGTGCGGTCCATTGTCCGTGTGCG CTAGCATGCCCGTAACGGACCTCGTACTTTTGGCTTCAAAGGTTTTGCGCACAGACAAAATGTGCCACACTTGCAGC TCTGCATGTGTGCGCGTTACCAACAATCCCAACGCGCAGTGACTTGTGTTGATGCAAATAAATCTCGATAAAGGCG CGGCGCGCAATGCAGCTGATCACGTACGCTCCTCGTGTTCCGTTCAAGGACGGTGTTATCGACCTCAGATTAATGT TTATCGGCCGACTGTTTTCGTATCCGCTACCAACGCGTTTTTGCATTAACATTGTATGTGCGCGGATGTTCTATA TCTAATTTGAATAAATAAACGATAACCGCGTTGGTTTTAGAGGGCATAATAAAAGAAATATTGTTATCGTGTTCCG ATTAGGGCAGTATAAATTGACGTTTCATGTTGGATATTGTTTCAGTTGCAAGTTGACACTGGCGGCGACAAGATCGTG AAACAACCAAGTGACGCGGCCGCAATTTGTAAAAAATAAATAAATAAATGGTGTCGAAGGGCGAGGAACCTGTTACC GGTGTCGTGCCCATCCTGGTCGAACCTGGACGCGGACGTGAACGGTCACAAGTTCTCCGTGTCTGGCGAAGGCGAGGG CGACGCTACCTACGGAAGCTGACCTGAAGTTTCATCTGCACACCGGCAAGCTGCCGTGCTTGGCCTACCTGG TCACCACTCTGACCTACGGTGTCAGTGCTTCTCCGTTACCCGACCATGAAGCAGACGATTTCTTCAAGTCC GCTATGCCCCGAGGGTTACGTGCAAGAGCGTACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGTGCTGAAGT GAAGTTCGAAGGCGACACCTCGTGAACCGTATCGAGCTGAAGGGTATCGACTTCAAGGAAGATGGAACATCCTGG GCCACAAGCTCGAGTACAACATACTCCACAACGTGTACATCATGGCCGACAAGCAAAAGAACGGCATCAAAGTG AACTTCAAGATCCGCCACAACATCGAGGACGGTCCGTGACGTGGCTGACCACTACAGCAGAACACCCCATCGG CGACGGTCTGTGCTGCTGCTGACCACTACCTGTCCACCCAGTCCGCTCTGTCCAAGGACCCCAACGAGAAGC GTGACCACATGGTGCTGCTGAGTTCGTGACCGTGCTGGTATCACCTGGGCGATGGACGAGCTGTACAAGTAAGCC CCTTGTAACGCCACAATTGTGTTTGTGCAATAAACCCATGATTATTTGATTAAATTTGTTGTTTCTTTGTTCA TAGACAATAGTGTTTTCCTAAACGGGTACC
SEQ ID NO:9 Secuencia de nucleótidos que codifica coBDDFVIII TEN (V2.0)	ATGCAGATTGAACTGTCCACTTGCTTCTCTGTGCTCCTGCGGTTTTGCTTCTCGGCCACCCGCGGTATTACTT AGGTGCTGTGGAAGTGAAGTGGGACTACATGCAGTCCGACCTGGGAGAACTGCGGTTGGACGCGAGATTCCCACCTA GAGTCCCGAAGTCCTTCCATTCAACACCTCCGTGGTCTACAAAAGACCTGTTCTGTGGAGTTCACTGACACCTT TTCAATATTGCCAAGCCGCGCCCCCTGGATGGGCTGCTTGGTCTACGATCAAGCAGAGGTTCTACGACACCGT GGTCATCACACTGAAGAATGGCCTACACCCGCTGTCGCTGCATGCTGTGGAGTCTCTACTGGAAGGCTCAG AGGGTGCCGAATATGATGACCAGACCAAGAGAGGAAAAGGAGGATGACAAAGTGTTCGCGGTGGCAGCCACT TACGTGTGGCAAGTGTGAAGGAAAACGGGCTATGGCGTGGACCCCTATGCCTGACCTACTCTACCTGTCCCA TGTGGACCTTGTGAAGGATCTCAACTCGGACTGATCGGCGCCCTCTTGGTGTGAGAGAAGGCAGCCTGGCGAAGG AAAAGACTCAGACCTGCACAAGTTTCTGTTGTTGCTGTGTTGATGAAGGAAAGTCTGGCACTCAGAAACC AAGAACTCGCTGATGCAGGATAGAGATGCGGCCCTCGGCCAGAGCCTGGCCTAAAATGCACACCGTCAACGGATATGT GAACAGGTGCTCCCTGGCCTCATCGGCTGCCACAGAAAGTCCGTGTATTGGCATGTGATCGGCATGGGTACTACTC CGGAAGTGCATAGTATCTTTCTGGAGGGCCATACCTTCTTGGTGGCAACACAGACAGGCTCGCTGGAAATCTCG CCTATCACTTTCTTGTGCTGCGCAGACCTCTTATGGACCTGGACAGTTCCTGCTGTCTGTACATCAGTCCCCA TCAGCATGATGGGATGGAGGCTATGTCAAAGTGGACTCCTGCCCTGAGGAGCCACAGCTCCGGATGAAGAACAATG AGGAAGCGGAGGATTACGACGACGACCTGACTGACAGCGAAATGGACGTCGTGCGATTGATGACGACAACAGCCCG TCCTTCATCCAAATTAGATCAGTGGCGAAGAAGCACCCCAAGACCTGGGTGCACTACATTGCCGCCGAGGAAGAGGA CTGGGACTACGCGCCGCTGGTGCTGGCGCCAGACGACAGGAGCTACAAGTCCAGTACCTCAACAACGGGCGCAGC GCATTGGCAGGAAGTACAAGAAAGTCCGCTTCATGGCCTACACTGATGAAACCTTCAAGACGAGGGAAGCCATCCAG CACGAGTCAGGCATCCTGGGACCGCTCCTTACGGCGAAGTCCGGGATACCTGCTCATCTTTTCAAGAACCAGGC ATCGCGGCCCTACAACATCTACCTCACGGGATCACAGACGTGCGCCGCTCTACTCCCGCCGGCTGCCAAGGGAG TGAAGCACTGAAGGATTTTCCATCCTGCCGGGAGAAATCTCAAGTACAAGTGACCGTCACTGTGGAAGATGGC CCTACCAAGTCGGACCTCGCTGTCTGACCGGTACTATCTCGTTTGTGAACATGGAGCGGACCTGGCCTCGGG GCTGATTGGTCCGCTGCTGATCTGCTACAAGGAGTCCGTGGACGAGCGCGGGAACAGATCATGTCCGACAAGCGCA ACGTGATCCTGTTCTGTCTTTGATGAAAACAGATCGTGGTACTTGACTGAGAATATCCAGCGGTTCTTGCCCAAC CCAGCGGGAGTGCAACTGGAGGACCCGGAGTTCCAGGCCTCAAACATTATGACTCTATCAACGGCTATGTGTTCTGA CTGCTCCAACGAGCGTGTGCTGCTGATGAAGTGGCATACTGGTACATTTGTCCATCGAGGCCAGACCGACTTCC TGTCGTGTTCTTCTCCGGATACACCTTCAAGCATAAGATGGTGTAAGGACACTCTGACCTCTTCCATTTTCC GGAGAACTGTGTTTCATGTCAATGGAAACCCGGGCTTGTGGATTCTGGGTTGCCATAACTCGGACTTCCGGAATAG AGGGATGACCGCCCTGCTGAAAGTGTCCAGTGTGACAAGAATACCGGCGATTACTACGAGGACAGCTATGAGGACA TCTCCGCTTATCTGCTGTGTTCAAGAACAACGCCATTGAACCCAGGTCTTCTCCCAAACGGTGCACCGACC TCCGAA AGCGCCACCCAGAGTCAGGACCTGGCTCGGAACCGGTACCTCGGGCTCAGAGACACCGGGGACTTCCGAGTCCGC AACCCCGAGAGTGGACCCGGATCCGAACAGCAACCTCAGGATCAGAAACCCGGGAACCTCGGAATCCGCCACTC CCGAGTCGGGACAGGCACCTCCACTGAGCCTTCCGAGGGAAGCGCCCCGGATCCCTGCTGGATCCCTTACCAGC ACTGAAGAAGGCACCTCAGAATCCGCGACCCCTGAGTCCGGCCCTGGAAGCGAACCCGCCACCTCCGGTCCGAAAC CCCTGGGACTAGCGAGAGCGCCACTCCGAATCGGGCCAGGAAGCCCTGCCGGATCCCGACAGCAGCGAGGAGG GAAGCCCGCCGGGTACCGACTTCCACTGAGGAGGGAGCCTCATCCCCCGGTGCTGAAGCGGCATCAAAGAGAG ATCACCAGGACCACTCCAGTCCGATCAGGAAGAAATGACTACGACGATACTATCAGCGTGGAGATGAAGAAGGA GGACTTCGACATCTACGATGAGGATGAGAACCAGTCCCTCGGAGCTTTCAAGAAGAAACCCGCACTACTCTCATCG CTGCCGTGGAGCGGCTGTGGGATTACGGGATGTCCAGCTACCGCATGTGCTGCGGAATAGAGCGCAGTCAGGATCG GTGCCCCAGTTCAAGAAGTCTGTGTTCCAAGAGTTCACCGACGGTCTTCACTCAACCCCTGTACCGGGGCGAACT CAACGAACACCTGGGACTGCTTGGGCCGTATATCAGGGCAGAAAGTGAAGATAACATCATGGTCACCTTCCGCAACC AGGCCTCCCGCCGTACAGCTTCTACTCTTCACTGATCTCCTACGAGGAAGATCAGCGGCAGGGAGCCGAGCCCGG AAGAATTCGTCAAGCCTAACGAACTAAGACCTACTTTTGAAGGTCCAGCATCACATGGCCCGACCAAGACGA

	GTTCGACTGTAAAGCCTGGGCTACTTCTCCGATGTGGACCTGGAGAAGGACGTGCACTCGGGACTCATTGGCCCCGCTCCTTGTGTGCCATACTAATACCCTGAACCCTGCTACGGTCGCCAAGTCACAGTGCAGGAGTTGCGCCTCTTCTTCCACATCTTCGATGAAACAAAGTCTGGTACTTTACTGAGAACATGGAACGCAATTGCAGGGCACCCTGCAACATCCA GATGGAAGATCCCACCTTCAAGGAAAACTACCGGTTTCATGCCATTAACGGCTACATAATGGACACGTTGCCAGGAC TGGTCATGGCCAGGACCAGAGAATCCGGTGGTATCTGCTCTCCATGGGCTCCAACGAAAAACATTCACAGCATTAT TTTTCCGGCCATGTGTTACCGTCCGGAAGAAGGAAGAGTACAAGATGGCTCTGTACAACTCTACCCTGGAGTGT CGAGACTGTGGAATGCTGCTAGCAAGGCCGGCATTTGGAGAGTGAATGCCTGATCGGAGAGCATTTGCACGCCG GAATGTCCACCCTGTTTCTGTGTACTCCAACAAGTGCCAGACCCCGCTGGGAATGGCTCAGGTCATATTAGGGAT TTCCAGATCACTGCTTCGGGGCAGTACGGGCAGTGGGCACCTAAGTTGGCCCGGCTGCACTACTCTGGCTCCATCAA TGCTGGTCCACCAAGGAACCTTCTCCTGGATTAAGGTGGACCTCCTGGCCCCAATGATTATTCACGGTATTAAGA CCCAGGGTGCCCGACAGAAGTTCTCCTCACTCTACATCTCGCAATTCATATAATGTACAGCCTGGATGGGAAGAAG TGGCAGACCTACCGGGGAACTCCACTGGAACGCTCATGGTGTTCGCGCAACGTGGACTCCTCCGGCATTAAAGCA CAACATCTTCAACCTCCGATCATTGCTCGGTACATCCGGTCGACCCAACTCACTACAGCATCCGGTCCACCCTGC GGATGGAACGATGGGTTGTGACCTGAACTCTGCTCCATGCCCTTGGGATGGAATCCAAGGCCATTAGCGATGCA CAGATCACCGCCTCTTCATACTTACCAACATGTTGCGGACCTGGTCCCGTCGAAGGCCCGCTGCACCTCCAAGG TCGCTCCAATGCGTGGCGGCTCAAGTGAACAACCCCAAGGAGTGGCTCCAGGTCGACTTCCAAAAGACCATGAAGG TCACCGGAGTGACCAACCGGCGTGAAGTCCCTGCTGACCTCTATGTACGTTAAGGAGTTCCTCATCTCCTCAAGC CAAGACGGACATCAGTGGACCTGTTCTTCCAAAACGGAAGTCAAAGTATTCCAGGGCAACCAGGACTCCTTCAC CCCTGTGGTCAACAGCCTGGACCCCATGCTGACCCGCTACCTCCGCATCCACCCCAAGCTGGGTCCACCAGA TCGCACTGCGCATGGAGGTCCTTGATGCGAAGCCCAAGATCTGTACTAA
SEQ ID NO: 10 Secuencia de aminoácidos de coBDDFVIII X TEN (V2.0)	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRPVKSFPFNTSVVYKKTLFVEFTDHLFNIAPRPPWMGLLGPTIQ AEVYDVTVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCL TYSYLSHVDLVKDLNLSLIGALLVCREGLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKM HTVNGYVNRSLPGLIGCHRKSVMYWHVIGMGTTPVHSIFLEGHTFLVRNHRQASLEISPIITFLTAQTLLMDLGQFLL FCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDEMDVVRFDNNSPSFIQIRSVAKKHPKTWVHY IAAEEDWDYAPLV LAPDDRSYKSYQLNNGPQRIGRKYKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGD TLL IIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYSSFVNM ERDLASGLIGPLLYCYKESVDQRGNQIMSDKRNVLFSVF DENRSWYLTENIQRFLPNPAGVQLEDPEFQASIMHS INGYVFDLSQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMYEDTLTLFPFSGETVFMENPGLWLIGCH NSDFRNRGMTALLKVSSCDKNTGDYEDSYEDISAYLLSKNNAIEPRSFQNGAPTSESATPESGPGSEPATSGSET PGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGSPAGSPTSTEESGSPAGSPTSTEEGASSPVLKRHRQREITRTTLQSDQEEIDYDDTI SVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLNRNAQSGSVQPKVVFQEF TDG SFTQ PLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPKRNFKVKNETKTYFWKVQHH MAPTKDEFDCAWAYFSDVDLEKDVHSLIGPLLVCHTNTLNPAHGRQVTVEFALFFTIFDETKSWYFTENMERN RAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHSGHVFTVRKKEEYKMA LYNLYPGVFETVEMLPKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTP LGMASGHIRD FQITASGQYQGWAPKLARL HYSGSINAWSTKEPFSWIKVDLLAPMI IHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNV DSSGIKHNIFNPPIIARYIRLHPHTYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNM FATWSPSK ARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKV KVFQ GNQDSFTPVVNSLDPPLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY
SEQ ID NO: 11 Péptido señal de coBDDFVIII X TEN (V2.0)	MQIELSTCFFLCLLRFCFS
SEQ ID NO: 12 Secuencia de aminoácidos del FVIII humano maduro BDD	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRPVKSFPFNTSVVYKKTLFVEFTDHLFNIAPRPPWMGLLGPTIQ AEVYDVTVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCL TYSYLSHVDLVKDLNLSLIGALLVCREGLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKM HTVNGYVNRSLPGLIGCHRKSVMYWHVIGMGTTPVHSIFLEGHTFLVRNHRQASLEISPIITFLTAQTLLMDLGQFLL FCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDEMDVVRFDNNSPSFIQIRSVAKKHPKTWVHY IAAEEDWDYAPLV LAPDDRSYKSYQLNNGPQRIGRKYKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGD TLL IIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYSSFVNM ERDLASGLIGPLLYCYKESVDQRGNQIMSDKRNVLFSVF DENRSWYLTENIQRFLPNPAGVQLEDPEFQASIMHS INGYVFDLSQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMYEDTLTLFPFSGETVFMENPGLWLIGCH NSDFRNRGMTALLKVSSCDKNTGDYEDSYEDISAYLLSKNNAIEPRSFQNGPPVLKRHRQREITRTTLQSDQEEIDY DDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLNRNAQSGSVQPKVVFQEF TDG SFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPKRNFKVKNETKTYFWK VQHMAPTKDEFDCAWAYFSDVDLEKDVHSLIGPLLVCHTNTLNPAHGRQVTVEFALFFTIFDETKSWYFTENM ERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHSGHVFTVRKKEEYK MALYNLYPGVFETVEMLPKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTP LGMASGHIRD FQITASGQYQGWAPK LARLHYSGSINAWSTKEPFSWIKVDLLAPMI IHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFF

	FGNVDSSGIKHNI FNPP IIRYIRLHPHYSIRSLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATW SPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKV KVFQGNQDSF TPVVNSLD PPLLTRYLR IHPQSWVHQIALRMEVLGCEAQDLY
SEQ ID NO: 13 Secuencia de nucleótidos que codifica el FVIII humano maduro BDD	ATGCAAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCATTCTGCTTTAGTGCCACCAGAAGATACTACCT GGGTGCAGTGGAAGTGTCTATGGGACTATATGCAAAGTGATCTCGGTGAGCTGCCTGTGGACGCAAGATTTCTCTCTA GAGTGCCAAAAATCTTTTCCATTCAACACCTCAGTCGTGTACAAAAAGACTCTGTTTGTAGAATTCACGGATCACCTT TTCAACATCGTAAGCCAAGGCCACCTGGATGGGTCTGCTAGGTCTACCATCCAGGCTGAGGTTTATGATACAGT GGTCATTACACTTAAGAACATGGCTTCCCATCTGT CAGTCTTCATGCTGTTGGTGTATCCTACTGGAAAGCTTCTG AGGGAGCTGAATATGATGATCAGACCAGTCAAAGGGAGAAAGAAGATGATAAAGTCTTCCCTGGTGAAGCCATACA TATGTCTGGCAGGTCTGAAAGAGAATGGTCCAATGGCCTCTGACCCACTGTGCCTTACCTACTCATATCTTTCTCA TGTGGACCTGGTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGTATGTAGAGAAGGGAGTCTGGCCAAGG AAAAGACACAGACCTTGCACAAATTTATACTACTTTTGTCTGATTGTATGATGAAGGGAAGGTTGGCACTCAGAAACA AAGAAGTCTTGTATGCAGGATAGGGATGCTGCATCTGCTCGGGCTGGCTAAAATGCACACAGTCAATGGTTATGT AAACAGGTCTCTGCCAGGTCTGATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAATGGGCACCACTC CTGAAGTGCATCAATATTCCTCGAAGGTCACACATTTCTGTGAGGAACCATCGCCAGGCGTCTTGGAAATCTCG CCAATAACTTTCTTACTGCTCAAACACTCTTGATGGACCTTGGACAGTTTCTACTGTTTTGTATATCTCTTCCCA CCAACATGATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCAACTACGAATGAAAAAATATG AAGAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGGATGTGGTCAGGTTTGTATGATGACAACCTCCT TCCTTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCTAAACTTGGGTACATTACATTGCTGCTGAAGAGGAGGA CTGGGACTATGCTCCCTTAGTCTCTCGCCCCGATGACAGAAGTTATAAAGTCAATATTGAACAAATGGCCCTCAGC GGATTGGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAGACTCGTGAAGCTATTTCAG CATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACACTGTTGATTATATTTAAGAATCAAGC AAGCAGACCATATAACATCTACCTCACGGAATCACTGATGTCGCTCTTGTATTCAAGGAGATTACCAAAAGGTG TAAAACATTTGAAGGATTTTCCAATTCGTCAGGAGAAATATTCAAATATAAATGGACAGTGACTGTAGAAGATGGG CCAACTAAATCAGATCCTCGGTGCCTGACCCGCTATTACTCTAGTTTCGTTAATATGGAGAGAGATCTAGCTTCAGG ACTCATTGGCCCTCTCCTCATCTGCTACAAAGAATCTGTAGATCAAAGAGGAAACAGATAATGTGACACAAGAGGA ATGTCATCCTGTTTTCTGATTTGATGAGAACCAGGCTGGTACCTCAGAGAGAATATACAACGCTTTCTCCCAAT CCAGCTGGAGTGCAGCTTGAGGATCCAGAGTTCCAAGCCTCCAACATCATGCACAGCATCAATGGCTATGTTTTTGA TAGTTTGCAGTTGT CAGTTGTTTGCATGAGGTGGCATACTGGTACATTCTAAGCATTGGAGCACAGACTGACTTCC TTTCTGTCTTCTCTCTGGATATACCTTCAAACACAAAATGGTCTATGAAGACACACTCACCTATTCCCATTCTCA GGAGAACTGTCTTCATGTGATGGAAAAACCCAGGTCTATGGATTCTGGGGTGCCACAACCTCAGACTTTGGAACAG AGGCATGACCGCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGGTGATTATTACGAGGACAGTTATGAAGATA TTTCAGCATACTTGCTGAGTAAAAACAAATGCCATTGAACCAAGAAGCTTCTCTCAAAACCCACCAGTCTTGAAACGC CATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAATTGACTATGATGATACCATATCAGTTGA AATGAAGAAGGAAGATTTTGACATTTATGATGAGGATGAAAATCAGAGCCCCGACGCTTTCAAAGAAAAACACGAC ACTATTTTATTGCTGCAGTGGAGAGGCTCTGGGATTTATGGGATGAGTAGCTCCCCCATGTTCTTAAGAAACAGGGCT CAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTTCCAGGAATTTACTGATGGCTCCTTCTCAGCCCTTATA CCGTGGAGAACTAAATGAACATTTGGGACTCTGGGGCCATATATAAGAGCAGAAGTTGAAGATAATATCATGGTAA CTTTCAGAAATCAGGCCTCTCGTCCCTATTCTTCTATTCTAGCCTTATTTCTTATGAGGAAGATCAGAGGCAAGGA GCAGAACCTAGAAAAAATTTGTCAAGCCTAATGAAACCAAACCTTACTTTTGGAAAGTGCAACATCATATGGCACC CACTAAAGATGAGTTTGACTGCAAAGCCTGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACTCAGGCC TGATTGGACCCCTTCTGGTCTGCCACACTAACACACTGAACCTGCTCATGGGAGACAAGTGACAGTACAGGAATTT GCTCTGTTTTTACCATCTTTGATGAGACCAAAGCTGGTACTTCACTGAAAAATGGAAAGAACTGCAGGGCTCC CTGCAATATCCAGATGGAAGATCCCCTTTTAAAGAGAATTTATCGTTCATGCAATCAATGGCTACATAATGGATA CACTACCTGGCTTAGTAATGGCTCAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAAACATC CATTCTATTCTTTTCACTGGACATGTGTTCACTGTACGAAAAAAGAGGAGTATAAAATGGCACTGTACAATCTCTA TCCAGGTGTTTTTGGACAGTGGAAATGTTACCATCAAAGCTGGAATTTGGCGGGTGGAAATGCCTTATTGGCGAGC ATCTACATGCTGGGATGAGCACACTTTTTCTGGTGTACAGCAATAAGTGTGAGACTCCCCTGGGAATGGCTTCTGGA CACATTAGAGATTTTTCAGATTACAGCTTCAGGACAAATATGGACAGTGGGCCCCAAAGCTGGCCAGACTTCATTATTC CGGATCAATCAATGCCTGGAGCACCAAGGAGCCCTTTTCTGGATCAAGGTGGATCTGTTGGCACCATGATTATTC ACGGCATCAAGACCCAGGGTGCCCGTCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTT GATGGGAAGAAGTGGCAGACTTATCGAGGAAATTCAGTGGAACTTAAATGGTCTTCTTTGGCAATGTGGATTTCATC TGGGATAAAAAACAATATTTTAAACCTCCAATTTATTTGCTGATACATCCGTTTGACCCCAACTCATATAGCATTTC GCAGCACTCTTCGCATGGAGTTGATGGGCTGTGATTTAAATAGTTGCAGCATGCCATTGGGAATGGAGAGTAAAGCA ATATCAGATGCACAGATTACTGCTTCATCTTACTTTACCAATATGTTTGCCACCTGGTCTCCTTCAAAGCTCGACT TCACCTCCAAGGGAGGAGTAATGCCTGGAGACCTCAGGTGAATAATCCAAAGAGTGGCTGCAAGTGGACTTCCAGA AGACAATGAAAGTCACAGGAGTAACCTACTCAGGGAGTAAAATCTCTGCTTACCAGCATGTATGTGAAGGAGTTCTC ATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTTTTTTTCAGAATGGCAAAGTAAAGGTTTTTTCAGGGAAATCA AGACTCCTTCACACCTGTGGTGAACCTCTAGACCCACCGTTACTGACTCGCTACCTTCGAATTCACCCCAAGAGTT GGGTGCACCAGATTGCCCTGAGGATGGAGTTCTGGGCTGCGAGGCACAGGACCTCTAC
SEQ ID NO: 14 V2.0	GGCCCCAGGTTAATTTTTTAAAAAGCAGTCAAAGGTCAAAGTGGCCCTTGGCAGCATTTACTCTCTCTATTGACTTTG GTTAATAATCTCAGGAGCACAAACATTCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTATCCTCTGGGCCCTC TCCCCACCTTCGATGGCCCCAGGTTAATTTTTTAAAAAGCAGTCAAAGGTCAAAGTGGCCCTTGGCAGCATTTACTCT CTCTATTGACTTTGGTTAATAATCTCAGGAGCACAAACATTCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTT ATCCTCTGGGCCCTCTCCCCACCGATATCTACCTGCTGATCGCCCGGCCCTGTTCAAACATGTCCTAATACTCTGTCT

Casete de expresión mTTR482-Intrón-coBDDFVIII-X TEN (V2.0)-WPRE-bGHPoliA	GGGGCAAAGGTCGGCAGTAGTTTTCCATCTTACTCAACATCTCCAGTGTACGTAGGATCCTGTCTGTCTGCACAT TTCGTAGAGCGAGTGTTCCGATACTCTAATCTCCCGGGGCAAAGGTCGTATTGACTTAGGTTACTTATTCTCCTTTT GTTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGG GGGTATAAAAGCCCCCTTACCAGGAGAAGCCGTACACAGATCCACAAGCTCCTGCTAGGAATTCTCAGGAGCACAA ACATTCTTGGAGGCGAGGAGAAGAAATCAACATCTGGACTTATCCTCTGGGCCTCTCCCCACCGATATCTACCTGCT GATCGCCCCGGCCCTGTTCAAACATGTCTAATACTCTGTCTGGGGCAAAGGTCGGCAGTAGTTTTCCATCTTACTCA ACATCCTCCAGTGTACGTAGGATCCTGTCTGTCTGCACATTTTCGTAGAGCGAGTGTTCCGATACTCTAATCTCCCG GGGCAAAGGTCGTATTGACTTAGGTTACTTATTCTCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGTTTG AGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGGTATAAAAGCCCCCTTACCAGGAGAAGCCGTAC ACAGATCCACAAGCTCCTGCTAGAGTCGTGCGCGCTGCCCTCGCCCCGTGCCCGCTCCGCCCGCCTCGCGC CG CCCCCCCCGGCTCTGACTGACCGCGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAAAT AGCGCTTGTTTATTGACGGCTTGTCTTTCTGTGGCTGCGTGAAAGCCTTGAGGGGCTCCGGGAAGGCCCTTTG TGCGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGCGTGGGGAGCGCGCTGCGGCTCCGCGCTGCCCG CGGCTGTGAGCGCTGCGGGCGGGCGGGGCTTGTGCGCTCCGAGTGTGCGCGAGGGGAGCGCGGCGGGGGCG GTGCCCGCGGTGCGGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGG GGTGTGGGCGCGTGGTGGGCTGCAACCCCCCTGCACCCCCCTCCCGAGTTGCTGAGCACGGCCCCGGCTTCGGG TGCGGGGCTCCGTACGGGGCGTGCGCGGGGCTGCGCGTGCGGGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCG GGGCGGGGCCCTCGGGCCGGGAGGGCTCGGGGAGGGGCGCGCGGGCCCCGGAGCGCGGGCGCTGTGAGGC GCGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCTTTGTCCAAATCTGTGC GGAGCCGAAATCTGGGAGCGCGCGCCGACCCCCCTTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCGGCAGGAAGG AAATGGGCGGGGAGGGCTTCGTGCGTGC CGCGCGCGCGTCCCTTCTCCTCTCCAGCCTCGGGGCTGTCCGCGG GGGAGCGGCTGCCCTCGGGGGGAGCGGGCAGGGCGGGGTCGGCTTCTGGCGTGTGACCGGGCTCTAGAGCCTC TGCTAACCTTGTCTTGCTTCTTCTTTTCTTCTACAGCTCCTGGGCAACGTGTGGTTATTGTGCTGTCTCATCATT TTGGCAAAGAATTACTCGAGGCCACCATGCAGATTGAACTGTCCACTTGCTTCTTCTGTGCTCTGCGTTTTGTC TTCTCGGCCACCCGCGGTATTACTTAGGTGCTGTGGAAGTGAAGTGGGACTACATGCAGTCCGACCTGGGAGAACT GCCGGTGGACGCGAGATTCCACCTAGAGTCCCGAAGTCTTCCCATTCACACCTCCGTGGTCTACAAAAAGACCC TGTTCTGTGGAGTTCAGTACACCTTTTCAATATTGCCAAGCCGCGCCCCCTGGATGGGCTGTGGTCTACG ATCCAAGCAGAGGTCTACGACACCGTGGTCAACACTGAAGAACATGGCCTCACACCCCGTGTGCTGCATGCTGT GGGAGTGTCTACTGGAAGGCTCAGAGGGTGCCGAATATGATGACCAGACCAGCCAGAGGGAAAGGAGGATGACA AAGTGTTCGGGGTGGCAGCCACACTACGTGTGGCAAGTGTGAAGGAAACGGGCTATGGCGTCGGACCCCTA TGCTAGCTACTCTACCTGTCCCATGTGGACCTTGTGAAGGATCTCAACTCGGGACTGATCGGCGCCCTCTTGGT GTGCAGAGAAGGCAGCCTGGCGAAGGAAAAGACTCAGACCTGCACAAGTTTATTCTGTTGTTTGTGTGTTGATG AAGGAAAGTCTGGCACTCAGAAACCAAGAACTCGTGTATGCAGGATAGAGATGCGGCTCGGCCAGAGCCTGGCCT AAAATGCACACCGTCAACGGATATGTGAACAGGTGCTCCCTGGCCTCATCGGCTGCCACAGAAAGTCCGTGTATTG GCATGTGATCGGCATGGGTACTACTCCGGAAGTGCATAGTATCTTTCTGGAGGGCCATACCTTCTTGGTGCACAACC ACAGACAGGCCCTCGCTGGAATCTCGCCTATCACTTTCTTGACTGCGCAGACCCTCCTTATGGACCTTGGACAGTTT CTGCTGTTCTGTACATCAGCTCCCATCAGCATGATGGGATGGAGGCCTATGTCAAAGTGGACTCCTGCCCTGAGGA GCCACAGCTCCGGATGAAGAACAATGAGGAAGCGGAGGATTACGACGACGACCTGAGTACAGCGAAATGGACGTG TGCATTGATGACGACAACAGCCGCTCTTCTATCCAAATAGATCAGTGCGGAAGCAAGCACCACCAAGACCTGGGTG CACTACATTGCGCGGAGGAAAGAGGACTGGGACTACGCGCCGCTGGTGTGTCGCCAGACGACAGGAGC TACAAGTC CCAGTACCTCAACAACGGGCCGAGCGCATTTGGCAGGAAGTACAAGAAAGTCCGCTTCATGGCCTACACTGATGAAA CCTTCAAGACGAGGGAAGCCATCCAGCAGAGTCAGGCATCCTGGGACCGCTCCTTACGGCGAAGTGGGGATACC CTGCTCATCATTTTCAAGAACCAGGCATCGCGGCCCTACAACATCTACCCTCACGGGATCACAGACGTGCGCCCGCT CTACTCCCGCCGGCTGCCAAGGGAGTGAAGCACCTGAAGGATTTTCCATCCTGCGGGAGAAATCTTCAAGTACA AGTGGACCGTGACTGTGGAAGATGGCCCTACCAAGTCCGACCTCGTGTCTGACCCGGTACTATTCTCGTTTGTG AACATGGAGCGCGACCTGGCCTCGGGGCTGATTGGTCCGCTGCTGATCTGCTACAAGGAGTCCGTGGACAGCGCGG GAACCAGATCATGTCCGACAAGCGCAACGTGATCTGTTCTGTCTTTGATGAAAACAGATCGTGTTGACTG AGAATATCCAGCGGTTCTTGCCCAACCCAGCGGAGTGCAACTGGAGGACCCGGATTCAGGCTCAAACTATATG CACTCTATCAACGGCTATGTGTTGACTCGTCCAAGTGAAGCGTGTGCTGCATGAAGTGGCATACTGGTACATTCT GTCCATCGGAGCCAGACCGACTTCTGTCCGTGTTCTTCTCCGATACACCTTCAAGCATAAGATGGTGTACGAGG ACACTCTGACCTCTTCCCATTTTCCGGAGAACTGTGTTTATGTCAATGAAAACCCGGGCTTGTGGATTCTGGGT TGCCATAACTCGGACTTCCGGAATAGAGGGATGACCGCCCTGCTGAAAGTGTCCAGCTGTGACAAGAAATACCGGCGA TTACTACGAGGACAGCTATGAGGACATCTCCGCTTATCTGTGTCCAAGAACAACGCCATTGAACCCAGGTCTTCT CCCAAAACGGTGCACCGACCTCCGAAAGCGCCACCCAGAGTCAGGACCTGGCTCGGAACCGGCTACCTCGGGCTCA GAGACACCGGGGACTTCCGAGTCCGCAACCCCGAGAGTGGACCCGGATCCGAACAGCAACCTCAGGATCAGAAAC CCCGGGAACCTCGGAATCCGCCACTCCCGAGTCGGGACCAAGGCACCTTCACTGAGCCTTCCGAGGGAAGCGCCCCG GATCCCCTGTGGATCCCCTACCAGCACTGAAGAAGGCACCTCAGAATCCGCGACCCCTGAGTCCGGCCCTGGAAGC GAACCCGCCACCTCCGGTTCCGAAACCCCTGGGACTAGCGAGAGCGCCACTCCGGAATCGGGCCAGGAAGCCCTGC CGGATCCCCGACCAGCACCGAGGAGGGAAGCCCCCGGGTACCGGACTTCCACTGAGGAGGGAGCCTCATCCCCC CCGTGCTGAAGCGGCATCAAAGAGAGATCACCAGGACCACTTCCAGTCCGATCAGGAAGAAATTGACTACGACGAT ACTATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGATGAGGATGAGAACCAGTCCCCTCGGAGCTTCA GAAGAAAACCCGCCACTACTTATCGCTGCCGTGGAGCGGCTGTGGGATTACGGGATGTCCAGCTACCGCATGTGC TGCGGAATAGAGCGCAGTCAGGATCGGTGCCCAAGTTCAGAAGGTCGTGTTTCAAGAGTTACCGACGGGCTCTT ACTCAACCCCTGTACCGGGGCGAACTCAACGAACACCTGGGACTGCTTGGGCCGTATATCAGGGCAGAAGTGAAGA TAACATCATGGTCACCTTCCGCAACCAAGGCTCCCGGCCGTACAGCTTCTACTTCTACTGATCTCCTACGAGGAAG ATCAGCGGCAGGGAGCCGAGCCCCGGAAGAACTTCGTCAAGCCTAACGAACTAAGACCTACTTTTGAAGGTCCAG CATCATATGGCCCCGACCAAGACGAGTTCGACTGTAAAGCCTGGGCCTACTTCTCCGATGTGGACCTGGAGAAGGA
---	--

	CGTGCACTCGGGACTCATTGGCCCCGCTCCTTGTGTGCCATACTAATACCCTGAACCCTGCTCACGGTCGCCAAGTCA CAGTGCAGGAGTTCGCCCTCTTCTTACCACCTCTCGATGAAACAAAGTCTGGTACTTTACTGAGAACATGGAACGC AATTGCAGGGCACCCTGCAACATCCAGATGGAAGATCCCACCTTCAAGGAAAACCTACCGGTTTCATGCCATTAACGG CTACATAATGGACACGTTGCCAGGACTGGTTCATGGCCCAGGACCAGAGAATCCGGTGGTATCTGCTCTCCATGGGCT CCAACGAAAACATTACAGCATTATTTTTCCGGCCATGTGTTACCGTCCGGAAGAAGGAAGAGTACAAGATGGCT CTGTACAACCTCTACCCTGGAGTGTTCGAGACTGTGGAAATGCTGCCTAGCAAGGCCGGCATTGGAGAGTGGAAATG CCTGATCGGAGAGCATTTCACGCGCGAATGTCCACCCTGTTTCTTGTGTACTCCAACAAGTGCCAGACCCCGCTGG GAATGGCCTCAGGTTCATATTAGGGATTTCAGATCACTGCTTCGGGGCAGTACGGGCAGTGGGCACCTAAGTTGGCC CGGCTGCACTACTCTGGCTCCATCAATGCCTGGTCCACCAAGGAACCTTCTCCTGGATTAAAGTGGACCTCTGGC CCCAATGATTATTCACGGTATTAAGACCCAGGGTGCCCGACAGAAGTTCTCCTCACTCTACATCTCGCAATTCATCA TAATGTACAGCTGGATGGGAAGAAGTGGCAGACCTACCGGGGAACTCCACTGGAACGCTCATGGTGTTCCTGGC AACGTGGACTCCTCCGGCATTAAGCACAACATCTTCAACCTCCGATCATTGCTCGGTACATCCGGCTGCACCCAAC TCACTACAGCATCCGGTCCACCTGCGGATGGAAGTGTGGTGTGACCTGAACCTCCTGCTCCATGCCCTTGGGA TGAATCCAAGGCCATTAGCGATGCACAGATCACGCTCTTCATACTTCAACCAATGTTTCGCGACCTGGTCCCCG TCGAAGGCCCGCTGCACCTCCAAGGTGCTCCAATGCGTGGCGGCTCAAGTGAACAACCCCAAGGAGTGGCTCCA GGTCGACTTCCAAAAGACCATGAAGGTACCGGAGTGACCACCCAGGGCGTGAAGTCCCTGCTGACCTCTATGTACG TTAAGGAGTTCCTCATCTCTCAAGCCAAGACGGACATCAGTGGACCTGTTCTTCCAAAACGAAAAGTCAAAGTA TTCCAGGGCAACCAGGACTCCTTACCCTGTGGTCAACAGCTGGACCCCCATTGCTGACCCGCTACCTCCGCAT CCACCCCCAAAGCTGGGTCCACCAGATCGCACTGCGCATGGAGGTCTTGGATGCGAAGCCCAAGATCTGTACTAAG CGGCCGCTCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAATCTATGTTGCTCCTTT ACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCTCTCTCTT GTATAAATCCTGGTGTGCTCTCTTTATGAGGAGTTGTGGCCGTTGTGAGGCAACGTGGCGTGGTGTGCACTGTGT TTGCTGACGCAACCCCACTGGTTGGGCAATGGCACCACCTGTGAGTCTCTTCCGGGACTTTCGCTTTCCCTC CCTATTGCCACGGCGAATCATCGCCGCTGCCTTGCCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAA TTCCGTGGTGTGTGCGGGAAATCATCGTCTTCTTGGCTGCTGCGCTGTGTTGCCACCTGGATTCTGCGCGGGA CGTCTTCTGCTACGTCCCTTGGCCCTCAATCCAGCGGACCTTCTTCCCGCGGCTGCTGCCGCTCTGCGGCT CTTCGCGTCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT CCTTCTAGTTGCCAGCATCTGTTGTTTGCCTTCCCCGCTGCTTCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT CCTTCTAATAAAATGAGGAAATGTCATCGCATTGTCTGAGTAGGTGTCTATTCTTGGGGGTGGGGTGGGGC AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAAGACCATGGGCGCGCAGGCCTGTCGAC GCCCCGGCGGTACCGCATCGCTCGCGACGCATAAAG
SEQ ID NO: 15 Potenciador A1MB2	GGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGGTCAAAGTGCCCTTGGCAGCATTTACTCTCTCTATTGACTTTG GTTAATAATCTCAGGAGCACAAACATTCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTATCTCTGGGCTC TCCCCACCTTCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGGTCAAAGTGCCCTTGGCAGCATTTACTCT CTCTATTGACTTTGGTTAATAATCTCAGGAGCACAAACATTCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTT ATCTCTGGGCTCTCCCCACC
SEQ ID NO: 16 Promotor de mTTR	GATATCTACCTGCTGATCGCCCGGCCCTGTTCAAACATGTCCTAATACTCTGTGCGGGCAAAGGTGGCAGTAGTT TTCCATCTTACTCAACATCTCCAGTGTACGTAGGATCCTGTCTGTCTGCACATTCGTAGAGCGAGTGTCCGAT ACTCTAATCTCCCGGGCAAAGGTGCTATTGACTTAGGTTACTTATTCTCTTTTGTGACTAAGTCAATAATCAGA ATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCTGGGTGGAAGGAGGGGGTATAAAAGCCCCCTCACCA GGAGAAGCCGTACACAGATCCACAAGCTCTGCTAG
SEQ ID NO: 17 Intrón quimérico	TCAGGAGCACAAACATTCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTATCCTCTGGGCTCTCCCCACCGA TATCTACCTGCTGATCGCCCGGCCCTGTTCAAACATGTCCTAATACTCTGTGCGGGCAAAGGTGGCAGTAGTTTT CCATCTTACTCAACATCCTCCAGTGTACGTAGGATCCTGTCTGTCTGCACATTCGTAGAGCGAGTGTTCGATAC TCTAATCTCCCGGGCAAAGGTGCTATTGACTTAGGTTACTTATTCTCTTTTGTGACTAAGTCAATAATCAGAAT CAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCTGGGTGGAAGGAGGGGGTATAAAAGCCCCCTTACCAGG AGAAGCCGTACACAGATCCACAAGCTCTGCTAGAGTCGCTGCGCGCTGCCTTCCGCCGCTGCCCGCTCCGCCG CGCTCGCGCGCCCCGCCCGGCTCTGACTGACCGCTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCTC CGGGCTGTAATTAGCGCTTGGTTTATTGACGGCTTGTCTTTTCTGTGGCTGCGTGAAAGCCTTGAGGGGCTCCGG GAAGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCCTGTGTGTGCGTGGGGAGCGCCGCTGCGGCTC CGCGTGCCTCGGCGGCTGTGAGCGCTGCGGGCGCGCGGGGCTTTGTGCGCTCCGAGTGTGCGCAGGGGAGCG CGGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGGGGGTGTGCGTGGGG GGGTGAGCAGGGGGTGTGGGCGCTGCGTGGGCTGCAACCCCCCTGCACCCCCCTCCCCAGTTGCTGAGCACGG CCCGGCTTCCGGTGGGGGCTCCGTACGGGCGTGGCGGGGCTCGCGTGGCGGGGGGGTGGCGGCAGGTGG GGGTGCCGGGGGGGGGGGGCGCCTCGGGCGGGGAGGGCTCGGGGAGGGGCGCGCGGCCCGGAGCGCCGGC GGCTGTGAGGCGCGGCGAGCCGAGCATTGCCTTTATGGTAATCGTGCAGAGGGGCGAGGGACTTCTTTGTCT CCAAATCTGTGCGGAGCCGAAATCTGGGAGGCGCGCCGACCCCCCTTAGCGGGGCGGGGGCGAAGCGGTGCGGCG CCGGCAGGAAGGAAATGGGCGGGGAGGGCTTCTGCGTGCCTGCGCGCGCGCTCCCTTCTCCCTCTCCAGCTCGG GGCTGTGCGGGGGGACGGCTGCTTGGGGGGGACGGGGCAGGGCGGGGTTGCGCTTCTGGCGTGTGACCGGCGG CTCTAGAGCCTTGTAACTTGTCTTGCCTTCTTTTTCTACAGCTCCTGGGAACGTGCTGGTTATTGTGCTGCTCATCATTTTGGCAAAGAATTA

SEQ ID NO: 18 WPRE	TCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACATGTTGCTCCTTTTACGCTAT GTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAA TCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTGACGGCAACGTGGCGTGGTGTGCACTGTGTTTGTGA CGCAACCCCACTGGTTGGGGCATTGCCACCACCTGTGAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTG CCACGGCGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATCCGTG GTGTTGTGCGGGAAATCATCGTCTTTTCTTGGCTGCTCGCTGTGTTGCCACCTGGATTCTGCGCGGACGCTCCTT CTGTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCTTCCCGCGGCTGCTGCCGGCTCTGCGGCTCTTCCGC GTCTTCGCTTCGCCCTCAGACGAGTCGGATCTCCTTTGGGCCGCTCCCCGCTG
SEQ ID NO: 19 bGHPA	CGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGTCCTTCTTGACCCTGGAAGGTGCCACT CCCACTGTCTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTATTCTATTCTGGGGGTGG GGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA
SEQ ID NO: 20 Secuencia de aminoácidos de la proteína FVIII humana madura de tipo salvaje	ATRRYYLGAVELSDYMQSDLGELPVDARFPPRVPKSPFNSTSVVYKTLFVEFTDHLFNIAPRPPWMLLGPTIQ AEVYDVTVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCL TYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKM HTVNGYVNRSLPGLIGCHRSVYWHVIGMGTTPVEHVSIFLEGHTFLVRNHRQASLEISPIITFLTAQTLLMDLGQFLL FCHISSHQHDGMEAYVKVDSCEEPQLRMKNNEEAEDYDDDLTDEMDVVRFDNNSPSFIQIRSVAKKHPKTWVHY IAAEEDWDYAPLV LAPDDRSYKSYQLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGD TLL IIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFILPGEIFKYKWTVTVEDGPTKSDPRCLTRYSSFNVM ERDLASGLIGPLLICYSVESVDQRGNQIMSDKRNVLFSVF DENRSWYLTENIQRF LPNPAGVQLEDP EFQASINMHS INGYVFDLSQLSVCLHEVAYWYILSIGAQDFLSVFFSGYTFKHKMVYEDTLTFPFSGETVFMSENPGLWLILGCH NSDFRNRGMTALLKVSSCDKNTGDYEDSYEDISAYLLSKNNAIEPRSFQNSRHPSTRQKQFNATTIPENDIEKTD PWFAHRTMPMKIQNVSSD LMLLRQSPTPHGLSLSDLQEAKEYTFSDDPSPGAIDSNNSLSEMTFRPQLHHS GDM VFTPESGLQLRLNEKLGTTAATELKKLDFKVSSTSNLSTIPSDNLAAGTDNTSSLGPPSMPVHYDSQLD TTLFGK KSSPLTESGGPLSLSEENND SKLLESGLMNSQESSWGKNVSTESGRLFKGKRAHGPA L LTKDNALFKVSI SL LKTN KTSNNSATNRKTHIDGPSLLIENSPSVWQNI LESDTEFKKVTP LIHDRMLMDKNATALRLNHMSNKT TSSKNMEMVQ QKKEGPIPPDAQNPDM SFFKMLFLPESARWIQRTHGKNSLNSGQGPSKQLVSLGPEKSVEGQNF LSEKNKVVGKG EFTKDVGLKEMVFPSSRN LFLTNLDNLHENNTHNQEKKIQEEIEKKETLIQENVVL PQIHTVTGTKNFMKNL FLLST RQNVESGYDGAYAPVLQDFRSLNDSTNRKTKHTAHFSKKGEEENLEGLGNQTKQIVEKYACTTRISPNTSQQNFVTQ RSKRALKQFR LPLEETELEKRIIVDDTSTQWSKNMKHLTPSTLTQIDYNEKEKGAITQSP LSDCLTRSHSIPQANRS PLPIAKVSSFP SIRPIYLTRVLFQDNSSHLPAASYRKKDSGVQESSHFLQGA KKNLSAILTLEMTGDQREVGS LG TSATNSVTYKKVENTVLPKPDLPKTSKGVELLPKVHIYQKDLFPTETSNQSPGHDLVEGSL LQGT EGAIKWNEANR PGKVPFLRVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWKSQEKSPKTAFAKKKDTILSLNACESNHAIAAINEGQ NKPEIEVTWAKQGRTERLCSQNPV LKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKK TRHYFIAAVERLWDYGMSSSPHVLNRNQAQSGSV PQFKVVVFQFTDGSFTQPL YRGELNEHLGLLPYIRA EVEDNI MVTFRNQASRPYSFYSS LISYEEDQRQGAEPKRFVKNETKTYFWKVKQHMAPTKDEF DCKAWAYFSDVDLEKDVH SGLIGPLLVCHTNTLNPAHGRQVTQVEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYI MDTLPLGLVMAQDQIRIWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLP SKAGIWRVECLI GEHLHAGMSTLFLVYSNKCQTPLGMASGHIRD FQITASGQYQGWAPKLARLHYSGSINAWSTKEPFSWIKVDL LAMP I IHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSG IKHNI FNPPIIARYIRLHPHY SIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVD FQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSF TPVNSLDPL LTRYLRIHP QSWVHQIALRMEVLGCEAQDLY
SEQ ID NO: 21 B19 WT 5'	CCAAATCAGATGCCGCCGGTCGCCCGGTAGGCGGGACTTCCGGTACAAGATGGCGGACAATTACGTCATTTCTG TGACGTCATTTCTGTGACGTCACTTCGGTGGGCGGGACTTCCGGAATTAGGGTTGGCTCTGGGCCAGCTTGCTTG GGGTTGCTTGACACTAAGACAAGCGCGCGCGCTTGATCTTAGTGGCAGCTCAACCCCAAGCGCTGGCCAGAGC CAACCCTAATTCGGAAGTCCCGCCACCGGAAGTGACGTACAGGAAATGACGTACAGGAAATGACGTAATTGTC CGCCATCTTGACCGGAAGTCCCGCTACCGGCGGCGACCGGCGGCATCTGATTGGTGTCTTCTTTAAATTTT
SEQ ID NO: 22 B19 WT 3'	AAAATTTAAAAGAAGACACCAAATCAGATGCCGCCGGTCGCCCGGTAGGCGGGACTTCCGGTACAAGATGGCGGA CAATTACGTCATTTCTGTGACGTCACTTCTGTGACGTCACTTCCGGTGGGCGGGACTTCCGGAATTAGGGTTGGC TCTGGGCCAGCGCTTGGGGTTGACGTGCCACTAAGTCAAGCGGCGCGCGCTTGCTTAGTGTCAAGGCAACCCCA AGCAAGCTGGCCAGAGCCAACCTAATTCGGAAGTCCCGCCACCGGAAGTGACGTACAGGAAATGACGTACACA GGAAATGACGTAATTGTCCGCCATCTGTACCGGAAGTCCCGCTACCGGCGGCGACCGGCGGCATCTGATTGG
SEQ ID NO: 23 5'_B19_míni mo	CCAAATCAGATGCCGCCGGTCGCCCGGTAGGCGGGACTTCCGGTACAGCGCGCGCTGTACCGGAAGTCCCGCCT ACCGGCGGCGACCGGCGGCATCTGATTGGTGTCTTCTTTAAATTTT
SEQ ID NO: 24	AAAATTTAAAAGAAGACACCAAATCAGATGCCGCCGGTCGCCCGGTAGGCGGGACTTCCGGTACAGCGGCGCGCT GTACCGGAAGTCCCGCTACCGGCGGCGACCGGCGGCATCTGATTGG

3'_B19_míni mo	
SEQ ID NO: 25 5'_GPV_míni mo	CTCATTGGAGGGTTCGTTTCGTTTGAACGTTTCGTTTCGCATGCGAACGAACGTTTGAACGAACGAACCTCCAATGAGA CTCAAGGACAAGAGGATATTTTGC GCGCCAGGAAGTG
SEQ ID NO: 26 3'_GPV_míni mo	CACTTCCTGGCGCGCAAAATATCCTCTTGTCCTTGAGTCTCATTGGAGGGTTCGTTTCGTTTGAACGTTTCGTTTCGCAT GCGAACGAACGTTTGAACGAACGAACCTCCAATGAG
SEQ ID NO: 27 5'_GPV_Δ186	CTCATTGGAGGGTTCGTTTCGTTTGAACCAGCCAATCAGGGGAGGGGGAAGTGACGCAAGTTCCGGTCACATGCTTCC GGTGACGCACATCCGGTGACGTAGTTTCGCATGCCTGTCTATCGCCTACCCATCCCTGTCTGAGATCAAGGGCGTGAT CGTGACAGACTGGAGAGCGTGCTTATAATATCGGCTCTCAGGAGTGGAGCACCACAGTGCCAGATACGTGGCCA CCCAGGGCTATCTGATCTCCAACCTTCGACGCATGCGAACTACGTACCCGGATGTGCGTACCCGGAAGCATGTGACCG GAACCTTGCCTCACTTCCCCCTCCCCTGATTGGCTGGTTCGAACGAACGAACCTCCAATGAGACTCAAGGACAAGAG GATATTTTGC GCGCCAGGAAGTG
SEQ ID NO: 28 3'_GPV_Δ186	CACTTCCTGGCGCGCAAAATATCCTCTTGTCCTTGAGTCTCATTGGAGGGTTCGTTTCGTTTGAACCAGCCAATCAGG GGAGGGGGAAGTGACGCAAGTTCCGGTCACATGCTTCCGGTGACGCACATCCGGTGACGTAGTTTCGCATGCCTGTCT ATCGCCTACCCATCCCTGTCTGAGATCAAGGGCGTGATCGTGACAGACTGGAGAGCGTGCTTATAATATCGGCTCTCAGG TCAGGAGTGGAGCACCACAGTGCCAGATACGTGGCCACCCAGGGCTATCTGATCTCCAACCTTCGACGCATGCGAAC TACGTACCCGGATGTGCGTACCCGGAAGCATGTGACCGGAACCTGCGTCACTTCCCCCTCCCCTGATTGGCTGGTTC GAACGAACGAACCTCCAATGAG
SEQ ID NO: 29 5'_GPV_Δ120	CTCATTGGAGGGTTCGTTTCGTTTGAACCAGCCAATCAGGGGAGGGGGAAGTGACGCAAGTTCCGGTCACATGCTTCC GGTGACGCACATCCGGTGACGTAGTTCCGGTCACGTGCTTCTGTACGTGTTTCCGGTCGCATGCCTGTCTATCGC CTACCCATCCCTGTCTGAGATCAAGGGCGTGATCGTGACAGACTGGAGAGCGTGCTTATAATATCGGCTCTCAGG AGTGGAGCACCACAGTGCCAGATACGTGGCCACCCAGGGCTATCTGATCTCCAACCTTCGACGCATGCTCACGTGAC CGGAAACACGTGACAGGAAGCACGTGACCGGAACCTACGTACCCGGATGTGCGTACCCGGAAGCATGTGACCGGAAC TGCGTCACTTCCCCCTCCCCTGATTGGCTGGTTCGAACGAACGAACCTCCAATGAGACTCAAGGACAAGAGGATAT TTTGC GCGCCAGGAAGTG
SEQ ID NO: 30 3'_GPV_Δ120	CACTTCCTGGCGCGCAAAATATCCTCTTGTCCTTGAGTCTCATTGGAGGGTTCGTTTCGTTTGAACCAGCCAATCAGG GGAGGGGGAAGTGACGCAAGTTCCGGTCACATGCTTCCGGTGACGCACATCCGGTGACGTAGTTCCGGTCACGTGCT TCCTGTACGTGTTTCCGGTCACGTGAGCATGCCTGTCTATCGCCTACCCATCCCTGTCTGAGATCAAGGGCGTGAT CGTGACAGACTGGAGAGCGTGCTTATAATATCGGCTCTCAGGAGTGGAGCACCACAGTGCCAGATACGTGGCCA CCCAGGGCTATCTGATCTCCAACCTTCGACGGCATGCGACCGGAAACACGTGACAGGAAGCACGTGACCGGAAC TACCGGATGTGCGTACCCGGAAGCATGTGACCGGAACCTGCGTCACTTCCCCCTCCCCTGATTGGCTGGTTCGAAC GAACGAACCTCCAATGAG
SEQ ID NO: 31 Secuencia de ácido nucleico del FVIII humano de tipo salvaje	ATGCAAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGC GATTCTGCTTTAGTGCCACCAGAAGATACTACCT GGGTGCAGTGGAAGTGTCTGGGACTATATGCAAGTGATCTCGGTGAGTGCCTGTGGACGCAAGATTTCTCTCTA GAGTGCCAAAATCTTTTCCATTCAACACCTCAGTCGTGTACAAAAAGACTCTGTTTGTAGAATTACGGATCACCTT TTCAACATCGCTAAGCCAAGGCCACCCTGGATGGGTCTGCTAGGTCTTACCATCCAGGCTGAGGTTTATGATACAGT GGTCATTACACTTAAGAACATGGCTTCCCATCTGTCTGAGTCTTATGCTGTTGGTGTATCTTACTGGAAAGCTTCTG AGGGAGCTGAATATGATGATCAGACAGTCAAAGGGAGAAAGAAGATGATAAAGTCTTCCCTGGTGGAAAGCCATACA TATGTCTGGCAGGTCCTGAAAGAGAATGGTCCAATGGCCTCTGACCCACTGTGCCTTACCTACTCATATCTTTCTCA TGTGGACCTGGTAAAAGACTTGAATTAGGCCTCATTGGAGCCCTACTAGTATGTAGAGAAGGGAGTCTGGCCAAGG AAAAGACACAGACCTTGACAAAATTTATACTACTTTTTGCTGTATTTGATGAAGGGAAGGTTGGCACTCAGAAACA AAGAACTCCTTGATGCAGGATAGGGATGCTGCATCTGCTCGGGCTGGCTAAAATGCACACAGTCAATGGTTATGT AAACAGGTCTCTGCCAGGTCTGATTGGATGCCACAGGAAATCAGTCTATTGGCATGTATTGGAATGGGCACCACTC CTGAAGTGCACTCAATATTCCTCGAAGGTACACATTTCTTGTGAGGAACCATCGCCAGGCGTCTTGGAAATCTCG CCAATAACTTTCTTACTGCTCAAACACTCTTGATGGACCTTGGACAGTTTCTACTGTTTTGTCATATCTCTTCCCA CCAACATGATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAAAAATAATG AAGAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGGATGTGGTCAGGTTTGTATGATGACAACCTCTCT TCCTTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCTAAAACCTTGGGTACATTACATTGCTGCTGAAGAGGAGGA

	CTGGGACTATGCTCCCTTAGTCCTCGCCCCGATGACAGAAGTTATAAAAGTCAATATTTGAACAATGGCCCTCAGC GGATTGGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAGACTCGTGAAGCTATTACG CATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACACTGTTGATTATATTTAAGAATCAAGC AAGCAGACCATATAACATCTACCCTCACGGAATCACTGATGTCCTGCTTTGTATTCAAGGAGATTACCAAAAGGTG TAAACATTTGAAGGATTTTCAATTCTGCCAGGAGAAATATTCAAATATAAATGGACAGTGACTGTAGAAGATGGG CCAACTAAATCAGATCCTCGGTGCCTGACCCGCTATTACTCTAGTTTCGTTAATATGGAGAGAGATCTAGCTTCAGG ACTCATTGGCCCTCTCCTCATCTGCTACAAAGAATCTGTAGATCAAAGAGGAAACCAGATAATGTCTAGACAAGAGGA ATGTCATCCTGTTTTCTGTATTTGATGAGAACCAGGCTGGTACCTCACAGAGAATATACAACGCTTTCTCCCAAT CCAGCTGGAGTGCAGCTTGAGGATCCAGAGTTCCAAGCCTCCAACATCATGCACAGCATCAATGGCTATGTTTTGA TAGTTTGCAGTTGTCAGTTTGTGTCATGAGGTGGCATACTGGTACATTCTAAGCATTGGAGCACAGACTGACTTCC TTTCTGTCTTCTCTCTGGATATACCTTCAAACACAAAATGGTCTATGAAGACACACTACCCTATTCCCATTCTCA GGAGAACTGTCTTCATGTCGATGGAAACCCAGGTCTATGGATTCTGGGGTGCCACAACCTCAGACTTTCCGAACAG AGGCATGACCGCCTTACTGAAGGTTTTCTAGTTGTGACAAGAACACTGGTGATTATTACGAGGACAGTTATGAAGATA TTTCAGCATACTTGCTGAGTAAAAACAATGCCATTGAACCAAGAAGCTTCTCCAGAATTCAAGACACCCTAGCACT AGGCCAAAAGCAATTTAATGCCACCACAATTCAGAAAAATGACATAGAGAAGACTGACCCTTGTTTGCACACAGAAC ACCTATGCCTAAAATACAAAATGTCTCCTCTAGTGATTTGTTGATGCTCTTGCGACAGAGTCTACTCCACATGGGC TATCCTTATCTGATCTCCAAGAAGCCAAATATGAGACTTTTTCTGATGATCCATCACCTGGAGCAATAGACAGTAAT AACAGCCTGTCTGAAATGACACACTTCAGGCCACAGCTCCATCACAGTGGGGACATGGTATTTACCCTGAGTCAGG CCTCCAATTAAGATTAAATGAGAACTGGGGACAACCTGCAGCAACAGAGTTGAAGAACTTGATTTCAAAGTTTCTA GTACATCAAATAATCTGATTTCAACAATTCATCAGACAATTTGGCAGCAGGTACTGATAATACAAGTTCTTAGGA CCCCAAGTATGCCAGTTCATTATGATAGTCAATTAGATACCACTCTATTTGGCAAAAAGTCACTCCCCTTACTGA GTCTGGTGGACCTCTGAGCTTGAGTGAAGAAAATATGATTTCAAAGTTGTTAGAATCAGGTTTAAATGAATAGCCAAG AAAGTTTCATGGGGAAAAAATGTATCGTCAACAGAGAGTGGTAGGTTATTTAAAGGGAAAAAGAGCTCATGGACCTGCT TTGTTGACTAAAGATAATGCCTTATTCAAAGTTAGCATCTCTTTGTTAAAGACAAAACAAAACCTTCCAATAATTCAGC AACTAATAGAAAGACTCACATTGATGGCCCATCATTATTAATTGAGAATAGTCCATCAGTCTGGCAAAAATATATTAG AAAGTGACACTGAGTTTAAAAAAGTGACACCTTTGATTCATGACAGAATGCTTATGGACAAAAATGCTACAGCTTTG AGGCTAAATCATATGTCAAATAAACTACTTCATCAAAAAACATGGAAATGGTCCAACAGAAAAAAGAGGGCCCCAT TCCACCAGATGCACAAAATCCAGATATGTCGTTCTTTAAGATGCTATTCTTGCCAGAATCAGCAAGGTGGATACAAA GGACTCATGGAAAGAACTCTCTGAACTCTGGGCAAGGCCCCAGTCCAAAGCAATTAGTATCCTTAGGACCAGAAAAA TCTGTGGAAGGTGAGAAATTTCTGTCTGAGAAAAACAAAGTGGTAGTAGGAAAGGGTGAATTTACAAAGGAAGCTAGG ACTCAAAGAGATGGTTTTTCCAAGCAGCAGAAACCTATTTCTTACTAATTTGATATTTACATGAAAAATAATACAC ACAATCAAGAAAAAAAATTCAGGAAGAAATAGAAAAAGAGGAAACATTAATCCAAGAGAATGTAGTTTTGCTCTCAG ATACATACAGTGACTGGCACTAAGAATTTCTGAGAAACCTTTTCTTACTGAGCACTAGGCAAAATGTAGAAGGTTT ATATGACGGGCATATGCTCCAGTACTTCAAGATTTTAGGTCAATTAATGATTCAACAAATAGAACAAGAAACACA CAGCTCATTTCTCAAAAAAAGGGGAGGAAGAAAACCTTGAAGGCTTGGGAAATCAAACCAAGCAAAATGTAGAGAAA TATGATGCACCAAGGATATCTCCTAATACAAGCCAGCAGAAATTTGTGACGCAACGTAGTAAGAGAGCTTTGAA ACAATTCAGACTCCCCTAGAGAAACAGAACTTGAAGAAAGGATAAATGTTGGATGACACCTCAACCCAGTGGTCCA AAAAATGAAACATTTGACCCCGAGCACCTTCACACAGATAGACTACAATGAGAAGGAGAAAGGGGCCATTACACAG TCTCCCTTATCAGATTGCCCTTACGAGGAGTCAATGATCCTTCAAGCAAAATAGATCTCCATTACCAAGGTGAAAGG ATCATCATTTTCCATCTATTAGACCTATATCTGACAGGGTCTTATCCAGACAACCTTCTCATCTTCCAGCAG CATCTTATAGAAAGAAAGATTCTGGGGTCCAAGAAAGCAGTCAATTTCTTACAAGGAGCCAAAAAAAATAACCTTTCT TTAGCCATTCTAACCTTGGAGATGACTGGTGATCAAAGAGAGGTTGGCTCCCTGGGGACAAGTGCCACAAATTCAGT CACATACAAGAAAGTTGAGAACACTGTTCTCCCGAAACCAGACTTGCCCAAAACATCTGGCAAAGTTGAATTGCTTC CAAAAGTTCACATTTATCAGAAGGACCTATTCCCTACGGAACATAGCAATGGGTCTCCTGGCCATCTGGATCTCGTG GAAGGGAGCCTTCTTCAGGGAACAGAGGGAGCGATTAAAGTGAATGAAGCAACAGACCTGGAAAAAGTTCCCTTTCT GAGAGTAGCAACAGAAAGCTCTGCAAGACTCCCTCCAAGCTATTGGATCCTCTTGCTTGGGATAACCACTATGGTA CTCAGATACCAAAAGAGAGTGGAATCCCAAGAGAAGTCAACAGAAAAAACAGCTTTTAAAGAAAAAGGATACCAAT TTGTCCTGAAACGCTTGTGAAAGCAATCATGCAATTAGCAGCAATAAATGAGGACATTAAGCCGAAATAGAAAGT CACCTGGGCAAGCAAGGTAGGACTGAAAGGCTGTGCTCTCAAAACCCACCAGTCTTGAACGCCATCAACGGGAAA TAACCTGTAATACTCTTCTCAGTCAGATCAAGAGGAAATGACTATGATGATACCATATCAGTTGAAATGAAGAAGGAA GATTTTGACATTTATGATGAGGATGAAAATCAGAGCCCCGAGCTTTCAAAGAAAAACACGACACTATTTTATTGC TGCAGTGGAGAGGCTCTGGGATTATGGGATGAGTAGCTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTG TCCCTCAGTTCAAGAAAGTTGTTTTCCAGGAATTTACTGATGGCTCCTTTACTCAGCCCTTATACCGTGGAGAACTA AATGAACATTTGGGACTCCTGGGGCCATATATAAGAGCAGAAGTTGAAGATAATATCATGGTAACCTTCAGAAATCA GGCCTCTCGTCCCTATTCTTCTATTCTAGCCTTATTTCTATGAGGAAGATCAGAGGCAAGGAGCAGAACCTAGAA AAAACTTTGTCAAGCCTAATGAAACCAAACTTACTTTTGAAAGTGCAACATCATATGGCACCCCACTAAAGATGAG TTTGACTGCAAGCCTGGGCTTATTTCTGATGTTGACC TGAAAAAGATGTGCACTCAGGCCTGATTGGACCCCT TCTGGTCTGCCACACTAACACACTGAACCTGCTCATGGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTTCA CCATCTTTGATGAGACCAAAAGCTGGTACTTCACTGAAAATATGGAAAGAACTGCAGGGCTCCCTGCAATATCCAG ATGGAAGATCCCACCTTTTAAAGAGAATTATCGCTTCCATGCAATCAATGGCTACATAATGGATACACTACCTGGCTT AGTAATGGCTCAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAAACATCCATTCTATTCAAT TCAGTGGACATGTGTTCACTGTACGAAAAAAGAGGAGTATAAAATGGCACTGTACAATCTCTATCCAGGTGTTTTT GAGACAGTGGAAATGTTACCATCAAAGCTGGAATTTGGCGGGTGGAAATGCCTTATTGGCGAGCATCTACATGCTGG GATGAGCACACTTTTTCTGGTGTACAGCAATAAGGTGTCAGACTCCCCTGGGAATGGCTTCTGGACACATTAGAGATT TTCAGATTACAGCTTCAGGACAATATGGACAGTGGGCCCAAGCTGGCCAGACTTCATTATTCCGGATCAATCAAT GCCTGGAGCACCAAGGAGCCCTTTTCTGGATCAAGGTGGATCTGTTGGCACCAATGATTATTACGGCATCAAGAC CCAGGGTGCCCGTCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAAGT
--	---

	GGCAGACTTATCGAGGAAATTCCTGGAACCTTAATGGTCTTCTTTGGCAATGTGGATTATCTGGGATAAAACAC AATATTTTTAACCTCCAATTATTGCTCGATACATCCGTTTGACCCAACTCATTATAGCATTGCGAGCACTCTTCG CATGGAGTTGATGGGCTGTGATTTAAATAGTTGACGATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGCAC AGATTACTGCTTCATCTACTTTACCAATATGTTTGCCACCTGGTCTCTTCAAAAGCTCGACTTCACCTCCAAGGG AGGAGTAATGCCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTTCAGAAGACAATGAAAGT CACAGGAGTAATACTCAGGGAGTAAATCTCTGCTTACCAGCATGTATGTGAAGGAGTTCTCATCTCCAGCAGTC AAGATGGCCATCAGTGGACTCTCTTTTTTCAGAATGGCAAAGTAAAGGTTTTTCAGGGAAATCAAGACTCCTTCACA CCTGTGGTGAACCTCTAGACCCACCCTTACTGACTCGCTACCTTCGAATTCACCCCAAGATTGGGTGCACAGAT TGCCCTGAGGATGGAGGTTCTGGGCTGCGAGGCACAGGACCTCTAC
SEQ ID NO: 32 Secuencia de nucleótidos que codifica BDD-co6FVIII (V1.0) (no XTEN)	GCCACTCGCCGTAATACTTGGAGCCGTGGAGCTTTCATGGGACTACATGCAGAGCGACCTGGGCGAACTCCCCGT GGATGCCAGATTCCCCCCCCGCGTGCCAAAGTCCTTCCCCTTTAACACCTCCGTGGTGTACAAGAAAACCTCTTTG TCGAGTTCACTGACCACCTGTTCAACATCGCCAAGCCGCGCCACCTTGGATGGGCTCTCTGGACCGACCATCAA GCTGAAGTGTACGACACCTGGTGTACCCCTGAAGAACATGGCGTCCCACCCCGTGTCCCTGCATGCGGTGCGAGT GTCCTACTGGAAGGCTCCGAAGGAGTGAAGTACGACGACAGACTAGCCAGCGGAAAAGGAGGACGATAAAGTGT TCCCGGGCGGCTCGCATACTTACGTGTGGCAAGTCTGAAGGAAAACGGACCTATGGCATCCGATCTCTGTGCCTG ACTTACTCCTACCTTTCCATGTGGACCTCGTGAAGGACCTGAACAGCGGGCTGATTGGTGCATTTCTGTGTGCCG CGAAGGTTCTGCTCGTAAGGAAAAGACCCAGACCTCCATAAGTTTCATCTTTTGTTCGCTGTGTTTCGATGAAGGAA AGTCATGGCATTCCGAACTAAGAACTCGTGTATGCAGGACCGGGATGCCGCTCAGCCCGCGCTGGCCTAAATG CATACAGTCAACGGATACGTGAATCGGTCACTGCCCGGCTCATCGTTGTACAGAAAGTCCGTGTACTGGCAGT CATCGGCATGGGCACTACGCTGAAGTGCATCTCTTCTGGAAGGGCACCTTCTCTGTGCGCAACCCAGCC AGGCTCTCTGGAATCTCCCCGATTACCTTTCTGACCGCCAGACTCTGCTCATGGACCTGGGCGAGTTCCTTCTC TTCTGCCACATCTCCAGCCATCAGCACGACGGAATGGAGGCTACGTGAAGGTGGACTCATGCCGGAAGAACCTCA GTTGCGGATGAAGAACAACGAGGAGGCGGAGGACTATGACGACGATTTGACTGACTCCGAGATGGACGTCGTGCGGT TCGATGACGACAACAGCCCCAGCTTATCCAGATTGCGAGCGTGGCCAAGAGCACCCCAAAACCTGGGTGCATAC ATCGCGGCCGAGGAAGAAGATTGGGACTACGCCCGTGGTGTGGCACCCGATGACCGGTCTGTAAGTCCCAGTA TCTGAACAATGGTCCGACGCGATTGGCAGAAAGTACAAGAAAGTGGGTTTACATGGCGTACATGACGAAAGTTTA AGACCGGGAGGCCATTCAACATGAGAGCGGCTTCTGGACCACTGCTGTACGGAGAGTCCGCGATACCCGTCTC ATCATCTTCAAAAACAGGCTCTCCGCGCTTACAACATCTACCTCACGGAATCACGACGTGCGGCCACTCTACTC GCGGCGCTGCGGAAGGGCGTCAAGCACCTGAAAGACTTCCCTATCTGCGGGCGAAATCTTCAAGTATAAGTGA CCGTACCGTGGAGGACGGGCCCACCAAGAGCGATCTAGGTGTCTGACTCGGTACTACTCCAGCTCTGTGAACATG GAACGGGACCTGGCATCGGACTCATTGGACCGTCTGTATCTGCTACAAAGAGTGGTGGATCAACGCGGCAACCA GATCATGTCCGACAAGCGCAACGTGATCTGTTCTCCGTGTTGATGAAAACAGATCTGGTACCTCACTGAAAACA TCCAGAGGTTCTCCAAACCCCGAGGAGTGAACCTGGAGGACCTGAGTTTACGGCTCGAATATCATGCACTCG ATTAACGGTTACGTGTTGACTCGCTGCAGCTGAGCGTGTGCTCCATGAAGTGCCTTACTGGTACATTTCTGTCCAT CGGCGCCAGACTGACTTCTGAGCGTGTCTTTCCGGTTACACCTTTAAGCACAAGATGGTGTACGAAGATACCC TGACCTGTTCCCTTTCTCCGGCGAAACGGTGTTCATGTGATGGAGAACCAGGCTGTGGATTCTGGGATGCCAC AACAGCGACTTTCGGAACCGCGGAATGACTGCCCTGCTGAAGGTGCTCTCATGCGACAAGAACACCGGAGACTACTA CGAGGACTCTACGAGGATATCTCAGCCTACCTCTGTCCAAGAACAACGCGATCGAGCCGCGAGCTTACGCCAGA ACCCGCTGTGCTGAAGAGGCACCGAGGAAATACCCGGACACCTCCAATCGGATCAGGAGGAAATCGACTAC GACGACACCATCTCGGTGGAATGAAGAAGGAAGATTTGATATCTACGACGAGGACGAAAATCAGTCCCCTCGCTC ATTCAAAAAGAAAAGTAGACACTACTTTATCGCCGCGGTGGAAGACTGTGGGACTATGGAATGTCATCCAGCCCTC ACGTCTTCGGAACCGGGCCAGAGCGGATCGGTGCCTCAGTTCAAGAAAGTGGTGTTCAGGAGTTACCCGACGGC AGCTTACCCAGCGCTGTACCGGGGAGAACTGAACGAACCTGGGCTGCTCGGTCTATCTCCCTGACCTCGCGGAAAGT GGAGGATAACATCATGGTGACCTTCCGTAACCAAGCATCCAGACCTTACTCTTATCTCTCTCCCTGATCTCATACG AGGAGGACCGAGCGCAAGGCGCGAGCCCCGAAGAACTTCTGCAAGCCCAACGAGACTAAGACCTACTTCTGGAAG GTCCAACACCATATGGCCCCGACCAAGGATGAGTTTACTGCAAGGCTGGGCTACTTCTCCGACGTGGACCTTGA GAAGGATGTCCATTCCGGCTGATCGGGCCGCTGCTCGTGTGTACACCAACACCTGAACCCAGCGCATGGACGCC AGGTACCGTCCAGGAGTTTGTCTGTTCTTACCATTGTTGACGAAACTAAGTCTGGTACTTACCCGAGAATATG GAGCGAAACTGTAGAGCGCCTGCAATATCCAGATGGAAGATCCGACTTTCAAGGAGAACTATAGATTCCACGCCAT CAACGGGTACATCATGGATACTCTGCCGGGGCTGGTTCATGGCCAGGATCAGAGGATTCGGTGGTACTTGTGTCAA TGGGATCGAACGAAAACATTCACTCCATTCACTTCTCCGGTACGTGTTCACTGTGCGCAAGAAGGAGGATACAAG ATGGCGCTGTACAATCTGTACCCCGGGGTGTTTCAAGACTGTGGAGATGCTGCCGTCCAAGGCCGGCATCTGGAGAGT GGAGTGCCTGATCGGAGAGCACCTCCACGCGGGGATGTCCACCCTCTTCTGGTGTACTCGAATAAGTGCCAGACCC CGCTGGGATGGCTCGGGCCACATCAGAGACTTCCAGATCACAGCAAGCGGACAATACGGCAATGGGCGCCGAAG CTGGCCCGCTTGCACTACTCCGGATCGATCAACGCATGGTCCACCAAGGAACCGTTCTCGTGGATTAAGGTGGACCT CCTGGCCCTATGATTATCCACGGAATTAAGACCCAGGGCGCCAGGCAGAAGTCTCTCTCCCTGTACATCTCGCAAT TCATCATCATGTACAGCTGGACGGGAAGAAGTGGCAGACTTACAGGGGAACTCCACCGGCACCTGATGGTCTTT TTCGGCAACGTGGATTCTCCGGCATTAAGCACAAACATCTTCAACCCACCGATCATAGCCAGATATATTAGGTCCA CCCCACTCACTACTCAATCCGCTCAACTCTTCGGATGGAATCATGGGTGCGACCTGAACCTCTGCTCCATCCGCT TGGGGATGGAATCAAAGGCTATTAGCGACGCCAGATCACCGGAGCTCTACTTCACTAATGTTTCGCCACCTGG AGCCCTCAAAGGCCAGGCTGCACTTGCAAGGACGGTCAAATGCCTGGCGCCGCAAGTGAACAATCCGAAGGAATG GCTTCAAGTGGATTTCAAAAAGACCATGAAAGTGACCGGAGTCAACCCAGGGAGTGAAGTCCCTTCTGACCTCGA TGTATGTGAAGGAGTTCTGATTAGCAGCAGCCAGGACGGGACCAAGTGGACCTGTTCTTCAAAAACGGAAGGTC AAGTGTTTCCAGGGGAACAGGACTCGTTACACCCGTGGTGAACCTCTGGACCCCCACTGCTGACGCGGTACTT

	GAGGATTTCATCCTCAGTCCTGGGTCCATCAGATTGCATTGCGAATGGAAGTCTGGGCTGCGAGGCCAGGACCTGTACTGA
SEQ ID NO: 33 Secuencia de nucleótidos que codifica coBDDFVIII (V2.0) (no XTEN)	GCCACCCGCCGTATTACTTAGGTGCTGTGGAAGTGAAGTGGGACTACATGCAGTCCGACCTGGGAGAACTGCCGGT GGAGCGGAGATTCACCTAGAGTCCGGAAGTCTTCCATTCAACACCTCCGTGGTCTACAAAAAGACCCTGTTTCG TGGAGTTCACTGACCACCTTTTCAATATTGCCAAGCCGCGCCCCCTGGATGGGCTGCTTGGTCTACGATCCAA GCAGAGGTCTACGACACCTGGTCACTACACTGAAGAACATGGCCTCACACCCCGTGTCTGCTGATGCTGTGGGAGT GTCCTACTGGAAGGCTCAGAGGGTCCGAATATGATGACCAGACCAGCCAGAGGGAAAAGGAGGATGACAAAGTGT TCCCGGGTGGCAGCCACACTTACGTGTGGCAAGTGTGAAGGAAAACGGGCTATGGCGTCGGACCCCTATGCCTG ACCTACTCCTACCTGTCCCATGTGGACCTTGTGAAGGATCTCAACTCGGGACTGATCGGCGCCCTCTTGGTGTGCAG AGAAGGCAGCCTGGCGAAGGAAAAGACTCAGACCTGCACAAGTTCACTTGTGTTTGTCTGTGTTTCGATGAAGGAA AGTCCTGGCACTCAGAAACCAAGAACTCGCTGATGCAGGATAGAGATGCGGCTCGGCCAGAGCCTGGCCTAAATG CACACCGTCAACGGATATGTGAACAGGTGCTCCTGGCTCATCGGCTGCCACAGAAAGTCCGTGATTGGCATGT GATCGGCATGGGTACTACTCCGGAAGTGCATAGTATCTTTCTGGAGGGCCATACCTTCTTGGTGCACAACCACAGAC AGGCCTCGCTGGAATCTCGCTATCACTTTCTTGACTGCGCAGACCCTCTTATGGACCTGGACAGTTCCTGCTG TTCTGTACATCAGCTCCCATCAGCATGATGGGATGGAGGCCATGTCAAAGTGGACTCCTGCCCTGAGGAGCCACA GCTCCGGATGAAGAACAATGAGGAAGCGGAGGATTACGACGACGACCTGACTGACAGCGAAATGGACGTCTGCGCAT TCGATGACGACAACAGCCCGTCTTCAATCCAAATTAGATCAGTGGCGAAGAGCACCCCAAGACCTGGGTGCATAC ATTGCCGCGGAGGAAGAGGACTGGGACTACGCGCCGCTGGTGTGGCGCCAGACGACAGGAGCTACAAGTCCCAGTA CCTCAACAACGGGCGCAGCGCATTGGCAGGAAGTACAAGAAAGTCCGCTTCATGGCCTACACTGATGAAACCTTCA AGACGAGGAAGCCATCCAGCACGAGTCAGGCATCTGGGACCGCTCTTTACGGCGAAGTCGGGATACCTTGCTC ATCATTTTCAAGAACCAGGCATCGCGCCCTACAACATCTACCTCACGGGATCACAGACGTGCGCCCGCTCTACTC CCGCCGGCTGCCAAGGGAGTGAAGCACCTGAAGGATTTCCCATCTGCCGGGAGAAATCTTCAAGTACAAGTGA CCGTGAAGTGTGAAGATGGCCCTACCAAGTGGACCTCGCTGTCTGACCCGGTACTATTCTCGTTTGTGAACATG GAGCGGACCTGGCCTCGGGGCTGATTGGTCCGCTGCTGATCTGCTACAAGGAGTCCGTGGACAGCGCGGGAACCA GATCATGTCCGACAAGCGCAACGTGATCCTGTTCTGTCTTTGATGAAAACAGATCGTGGTACTTGACTGAGAATA TCCAGCGGTTCTGCCAACCCAGCGGGAGTGAAGTGGAGGACCCGGAGTTCCAGGCCTCAAACATTATGCACTCT ATCAACGGCTATGTGTTGCACTCGCTCCAAGTGAAGTGTGCTGATGAAGTGGCATACTGGTACATTCTGTCCAT CGGAGCCAGACCGACTTCTGTCCGTGTTCTTCTCCGGATACACCTTCAAGATAAGATGGTGTACGAGGACACTC TGACCTCTTCCCATTTTCCGGAGAACTGTGTTTATGTCAATGGAAAACCCGGGCTTGTGGATTCTGGGTTGCCAT AACTCGGACTTCCGGAATAGAGGGATGACCGCCCTGCTGAAAGTGTCCAGCTGTGACAAGAATACCGGCGATTACTA CGAGGACAGCTATGAGGACATCTCCGTTATCTGCTGTCCAAGAACAACGCCATTGAACCCAGGTCTTCTCCCAAA ACGGTGCACCGGCTCATCCCCCGCTGCTGAAGCGGCATCAAAGAGAGATCACCAGGACCACTCTCCAGTCCGAT CAGGAAGAAATTGACTACGACGATACTATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGATGAGGATGA GAACAGTCCCTCGGAGCTTTAGAAGAAAACCCGCCACTACTTCATCGCTGCCGTGGAGCGGCTGTGGGATTACG GGATGTCCAGCTCACCAGTGTGCTGCGGAATAGAGCGCAGTCAGGATCGGTGCCCGAGTTCAAGAAGGTCTGTGTT CAAGAGTTCACCGACGGGCTCTTCACTCAACCCCTGACCGGGCGAAGTCAACGAAACCTGGGACTGCTTGGGCC GTATATCAGGGCAGAAGTGAAGATAAATCATGTGTCACTTCCGCAACCGGCTCCGCGGTACAGCTTCTACT CTTCACTGATCTCTACGAGGAAGATCAGCGGAGGAGCGGAGCCCGGAAGAACTTCGTCAAGCCTAACGAACT AAGACCTACTTTTGAAGGTCCAGCATCACATGGCCCCGACCAAGACGAGTTGAGTGTAAAGCCTGGGCTACTT CTCCGATGTGGACCTGGAGAAGGACGTGCACTCGGACTCATTGGCCCGCTCTTGTGTGCCATACTAATACCCTGA ACCCTGCTCACGGTCGCCAAGTCAAGTGCAGGAGTTCGCCCTCTTCTTACCATCTTCGATGAAACAAAGTCTGG TACTTTACTGAGAACATGGAACGCAATTGCAGGGCACCTGCAACATCCAGATGGAAGATCCACCTTCAAGGAAAA CTACCGGTTTTCATGCCATTAACGGCTACATAATGGACAGTTGCCAGGACTGGTCATGGCCAGGACAGAGAATCC GGTGGTATCTGCTCTCCATGGGCTCCAACGAAACATTACAGCATTCTTTTCCGGCCATGTTTCCAGGCTCCG AAGAAGGAAGAGTACAAGATGGCTCTGTACAACCTTACCCTGGAGTGTTCGAGATGTGGAATGCTGCTACGCAA GGCCGGCATTTGGAGAGTGAATGCCTGATCGGAGAGCATTTGCACGCCGAATGTCCACCCTGTTTCTGTGTACT CCAACAAGTGCCAGACCCGCTGGGAATGGCCTCAGGTCAATTTAGGGATTTCCAGATCACTGCTTCCGGGACGATC GGGCAGTGGGCACCTAAGTTGGCCCGCTGCACTACTCTGGTCCATCAATGCCTGGTCCACCAAGGAACCTTCTC CTGGATTAAGGTGGACCTCTGGCCCCAATGATTATTCACGGTATTAAGACCCAGGGTGCCTGACAGAAGTCTCCT CACTCTACATCTCGCAATTCATCATAATGTACAGCTGGATGGGAAGAAGTGGCAGACCTACCGGGGAACTCCACT GGAACGCTCATGGTGTGTTTTCGGCAACGTGGACTCCTCCGGCATTAAGCACAACTCTTCAACCTCCGATCATTGC TCGGTACATCCGGCTGCACCAACTCACTACAGCATCCGGTCCACCTGCGGATGGAAGTGTGGTGTGACCTGA ACTCCTGCTCCATGCCCTTGGGATGGAATCCAAGGCCATTAGCGATGCACAGATCACCGCCTCTTCATACTTCAAC AACATGTTTCGCGACCTGGTCCCCGTCGAAGGCCCGCTGCACCTCCAAGGTGCTCCAATGCGTGGCGGCTCAAGT GAACAACCCCAAGGAGTGGTCCAGGTGCACTTCCAAAAGACCATGAAGTCAACGGAGTGACCAACAGGGCGTGA AGTCCCTGCTGACCTCTATGTACGTTAAGGAGTTCCTCATCTCTCAAGCAAGACGGACATCAGTGGACCTGTTC TTCCAAAACGGAAAAGTCAAAGTATTCCAGGGCAACAGGACTCCTTCAACCTGTGGTCAACAGCTGGACCCCC ATTGCTGACCCGCTACCTCCGCATCCACCCCAAAGCTGGGTCCACCAGATCGCACTGCGCATGGAGGTCTTGGAT GCGAAGCCCAAGATCTGTACTAA
SEQ ID NO: 34 V1.0 Casete de expresión	ATGCAGATTGAGCTGTCCACTTGTGTTCTTCTGTGCTCCTGCGCTTCTGTTTCTCCGCCACTCGCCGGTACTACCT TGGAGCCGTGGAGCTTTTCATGGGACTACATGCAGAGCGACCTGGGCGAATCCCCGTGGATGCCAGATTCCCCCCC GCGTGCCAAAGTCTTCCCTTTAACACCTCCGTGGTGTACAAGAAAACCTCTTTGTGAGTTCACTGACCACCTG TTCAACATCGCCAAGCCGCGCCACCTTGGATGGGCTCCTGGGACCGACCATTAAGTGAAGTGTACGACACCGT GGTGATCACCTGAAGAACATGGCGTCCACCCCGTGTCCCTGCATGCGGTGCGAGTGTCTACTGGAAGGCCCTCG AAGGAGCTGAGTACGACGACAGACTAGCCAGCGGGAAAAGGAGGACGATAAAGTGTTCGGGGCGGCTCGCATACT

TTP-Intrón-BDDFVIIIco6 XTEN (V1.0)- WPRE-bGHPolA	TACGTGTGGCAAGTCCTGAAGGAAAACGGACCTATGGCATCCGATCCTCTGTGCCTGACTTACTCCTACCTTTCCCA TGTGGACCTCGTGAAGGACCTGAACAGCGGGCTGATTGGTGCACCTTCTCGTGTGCCGGAAGGTTTCGCTCGCTAAGG AAAAGACCCAGACCTCCATAAGTTTATCCTTTTGTTCGCTGTGTTTCGATGAAGGAAAAGTCATGGCATTCCGAAACT AAGAAGTCGCTGATGCAGGACCGGGATGCCGCTCAGCCCCGCGCTGGCCTAAAATGCATACAGTCAACGGATACGT GAATCGGTCACTGCCCCGGGCTCATCGTTGTACAGAAAGTCCGTGTACTGGCACGTTCATCGGCATGGGCACTACGC CTGAAGTGCATCCATCTTCTGGAAGGGCACACCTTCTCGTGCACACCCAGGCCCTCTCTGGAAATCTCC CCGATTACCTTTCTGACCGCCAGACTCTGCTCATGGACCTGGGGCAGTTCTTCTCTTCTGCCACATCTCCAGCCA TCAGCACGACGGAATGGAGGCTACGTGAAGGTGGACTCATGCCCGGAAGAACCTCAGTTGCGGATGAAGAACAACG AGGAGGCCGAGGACTATGACGACGATTTGACTGACTCCGAGATGGACGTCGTGCGGTTTCGATGACGACAACAGCCCC AGCTTCATCCAGATTGCGAGCGTGGCCAAGAAGCACCCAAAACCTGGGTGCACTACATCGCGGCCGAGGAAGAAAG TTGGGACTACGCCCCGTTGGTGCTGGCACCCGATGACCGGTCTGACAAGTCCAGTATCTGAACAATGGTCCGACG GGATTGGCAGAAAGTACAAGAAAGTCCGTTTCATGGCGTACACTGACGAAACGTTTAAAGACCCGGGAGGCCATTCAA CATGAGAGCGGCATTCTGGGACCACTGCTGTACGGAGAGGTGCGGCATACCTGCTCATCTTCAAAAACAGGC CTCCCGGCTTACAACATCTACCTCACGGAATCACGACGTGCGGCCACTTACTCGCGGCGCTGCCGAAGGGCG TCAAGCACCTGAAAGACTTCCCTATCCTGCGGGGAAAATCTCAAGTATAAGTGGACCGTCACCGTGGAGGACGGG CCCACCAAGAGCGATCTAGGTGTCTGACTCGGTACTACTCCAGCTTCTGTAACATGGAACGGGACCTGGCATCGGG ACTCATTGGACCGCTGCTGATCTGCTACAAAGAGTCGGTGGATCAACGCGGCAACCAGATCATGTCCGACAAGCGCA ACGTGATCCTGTTCTCCGTGTTTATGAAAACAGATCCTGGTACCTCACTGAAAACATCCAGAGGTTCTCTCCAAAC CCCGCAGGAGTCAACTGGAGGACCTGAGTTTCAGGCCTCGAATATCATGCACTCGATTAAACGGTTACGTGTTTCA CTCGCTGCAACTGAGCGTGTGCCTCCATGAAGTCGCTTACTGGTACATTCTGTCCATCGGCGCCAGACTGACTTCC TGAGCGTGTCTTTTCCGGTTACACCTTAAAGCACAAAGATGGTGTACGAAGATACCTGACCTGTTTCCCTTTCTCC GGGAAAACGGTGTTCATGTCGATGGAGAACCCTGGGTCTGTGGATTCTGGGATGCCACAACAGCGACTTTCGGAACCG CGGAATGACTGCCCTGCTGAAGGTGCTCTATGCGACAAGAACCCGGAGACTACTACGAGGACTCTACGAGGATA TCTAGCCTACCTCTGTCCAAGAACAACCGCATCGAGCGCGCAGCTTACGCCAGAACGGCGCGCAACATCAGAG AGCGCCACCCCTGAAAGTGGTCCGGGAGCGAGCCAGCCACATCTGGGTGCGAAACGCCAGGCACAAGTGAGTCTGC AACTCCCGAGTCCGGACCTGGCTCCGAGCCTGCCACTAGCGGCTCCGAGACTCCGGGAACCTCCGAGAGCGCTACAC CAGAAAGCGGACCCGGAACAGTACCGAACCTAGCGAGGGCTCTGC TCCGGGCGAGCCAGCCGGCTCTCTACATCC ACGGAGGAGGGCACTTCCGAATCCGCCACCCCGAGTCAAGGGCCAGGATCTGAACCCGCTACCTCAGGCAGTGAGAC GCCAGGAACGAGCGAGTCCGCTACACCGGAGAGTGGGCCAGGGAGCCCTGCTGGATCTCTACGTCCACTGAGGAAG GGTACCCAGCGGGCTCGCCACCAGCACTGAAGAAGGTGCTCTGAGCCCGCTGTGCTGAAGAGGCACCGAGGAGAA ATTACCCGACCAACCTCCAATCGGATCAGGAGGAATCGACTACGACGACCACTCTCGGTGGAATGAAGAAGGA AGATTTTCGATATCTACGACGAGGACGAAAATCAGTCCCTCGCTCATTCCAAAAGAAAACAGACTACTTTATCG CCGCGGTGGAAGACTGTGGGACTATGGAATGTCATCCAGCCCTCACGTCTTTCGGAACCGGGCCAGAGCGGATCG GTGCCTCAGTTCAAGAAAGTGGTGTTCAGGAGTTCACCGACGGCAGCTTACCCAGCCGCTGTACCGGGGAGAACT GAACGAACACCTGGGCTGCTCGGTCCCTACATCCGCGCGGAAGTGGAGGATAACATCATGGTGACCTTCCGTAACC AAGCATCCAGACCTTACTCCTTCTATTCTCCCTGATCTCATACGAGGAGGACAGCGCCAAGGCGCCGAGCCCCGC AAGAAGTTCGTCGAAGCCCAACGAGACTAAGACCTACTTCTGGAAGGTCCAACACCATATGGCCCCGACCAAGGATGA GTTTGACTGCAAGGCTGGGCTACTTCTCCGACGTGGACCTTGAGAAGGATGTCCATTCCGGCTGATCGGGCCG TGCTCGTGTGTACACCAACACCTGAACCCAGCCGATGGAATGAGAGCCAGGTCACCGTCCAGGAGTTTCTGTCTTCTT ACCATTTTGTACGAAACTAAGTCTGGTACTTCAACCGAGGATGGAATATGGAGCGAACTGTAGAGCGCCCTGC AATATCCA GATGGAAGATCCGACTTTCAGGAGAACTATAGATTCCACGCCATCAACGGGTACATCATGGATACTCTGCCGGGGC TGGTCATGGCCAGGATCAGAGGATTCGGTGGTACTTGTGTCAATGGGATCGAACGAAAACATTCCTCCATTAC TTCTCCGGTCACGTGTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCGCTGTACAATCTGTACCCCGGGGTGTT CGAAACTGTGGAGATGCTGCCGTCCAAGGCCGGCATCTGGAGAGTGGAGTGCCTGATCGGAGAGCACTCCACGCGG GGATGTCCACCCTCTTCTGGTGTACTCGAATAAGTGCCAGACCCGCTGGGCATGGCTCGGGCCACATCAGAGAC TTCCAGATCACAGCAAGCGGACAATACGGCCAATGGGCGCCGAAGCTGGCCCCGCTTGCACTACTCCGGATCGATCAA CGCATGGTCCACCAAGGAACCGTTCTCGTGGATTAAAGTGGACCTCTTGCCCCATGATTATCCACGGAATTAAGA CCAGCGCGCCAGGACAGAAGTTCTCTCCCTGATCATCTCGCAATTATCATCATGTACAGCTTGGACGGGAAGAG TGGCAGACTTACAGGGGAACTCCACCGGCACCTGATGGTCTTTTTCGGAACGTTGGATTCTCCGGCATTAAAGCA CAACATCTTCAACCCACCGATCATAGCCAGATATATTAGGCTCCACCCACTCACTACTCAATCCGCTCAACTCTTC GGATGGAACCTCATGGGGTGCACCTGAACTCCTGCTCCATGCCGTTGGGGATGGAATCAAAGGCTATTAGCGACGCC CAGATCACCGCGAGCTCCTACTTCACTAACATGTTCCGCCACCTGGAGCCCCCTCAAGGCCAGGCTGCATTGACGGG ACGGTCAAATGCCTGGCGGCGCAAGTGAACAATCCGAAGGAATGGCTTCAAGTGGATTCCAAAAGACCATGAAAG TGACCGGAGTCAACCCAGGGAGTGAAGTCCCTTCTGACCTCGATGTATGTGAAGGAGTTCCTGATTAGCAGCAGC CAGGACGGGCACCACTGGACCTGTTCTTCAAAACGGAAGGTCAAGGTGTTCCAGGGGAACCAAGGACTCGTTAC ACCCGTGGTGAACCTCCTGGACCCCACTGCTGACGCGGTACTTGAGGATTCTATCTCAGTCTGGGTCCATCAGA TTGCATTGCGAATGGAAGTCTGGGCTGCGAGGCCAGGACCTGTACTGA
SEQ ID NO: 35 V3.0 Casete de expresión A1AT-Intrón- BDDFVIIIITE N-WPRE-	ATCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGTCCAAGTGGCCCTTGGCAGCATTTACTCTCTCTGTT TGCTCTGGTTAATAATCTCAGGAGCACAAACATTCTGGAGGCAGGAGAAGAAATCAACATCTGGACTTATCCTC TGGGCTCTCCCCACCTTCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGTCCAAGTGGCCCTTGGCAGC ATTTACTCTCTCTGTTTGTCTGTTTAAATATCTCAGGAGCACAAACATTCTGGAGGCAGGAGAAGAAATCAACA TCCTGGACTTATCTCTGGGCTCTCCCCACCTTCAAGTACGCACTAGCCTGAGGCTGGTCAAATTTGAACCTCC TCCTGCTCTGAGCAGCTGGGGGCGAGCTAAGCAGAGGGCTGTGCGAGCCACATAAAGAGCCTACTGTGTGCCA GGCACTTCAACCGAGGCACTTCAAGCATGCTTGGGAATGAACTTCAACTCTTGGGATGCAGGTGAAACAGT TCCTGGTTCAAGAGGTTGAAGCGGCTGCTGAGGCAGCACAGCTCTTCTTACAGATGTGCTTCCCCACCTCTAC CCTGTCTACGGCCCCCATGCCAGCCTGACGGTTGTGTCTGCCTCAGTCATGCTCCATTTTTCCATCGGGACCAT

bGHPoliA optimizado en codones	CAAGAGGGTGTGTTGTGTCTAAGGCTGACTGGGTAACCTTTGGATGAGCGGTCTCTCCGCTCTGAGCCTGTTTCTCA TCTGTCAAATGGGCTCTAACCCTCTGATCTCCAGGGCGGCAGTAAGTCTTCAGCATCAGGCATTTTGGGGTGA CTCAGTAAATGGTAGATCTTGCTACCACTGGAACAGCCACTAAGGATTCTGCACTGAGAGCAGAGGGCCAGCTAAG TGGTACTCTCCAGAGACTGTCTGACTCACGCCACCCCTCCACCTTTGGACACAGGACGCTGTGGTTTCTGAGCCA GGTACAATGACTCCTTTTCGGTAAGTGCAGTGGAAGCTGTACACTGCCAGGCAAAGCGTCCGGGCAGCGTAGGCGG GCGACTCAGATCCAGCCAGTGGAAGTCTAGCCCTGTTTGTCTCCGATAACTGGGGTACCTTGGTTAATATTCA CCAGCAGCTCCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAGGACAGGGCCCTGTCTCCTCAGCTT CAGGCAACCACCTGACCTGGGACAGGAATTCTCAGGAGCACAAACATTCTGGAGGACAGGAGAAGAAATCAACAT CCTGGACTTATCTCTGGGCTCTCCCAACCATATCTACCTGCTGATCGCCCGGCCCTGTTCAAACATGTCCTA ATACTCTGTGCGGGCAAAGGTGCGCAGTAGTTTTCCATCTTACTCAACATCCTCCAGTGTACGTAGGATCCTGTC TGTCTGCACATTTCTGAGAGCAGTGTTCGATACTCTAATCTCCGGGGCAAAGGTCTGATTGACTTAGGTTACT TATTCTCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCTGAGTGGCAGGGATCAGCAGCCTG GGTTGGAAGGAGGGGGTATAAAAGCCCTTACCAGGAGAAGCCGTACACAGATCCACAAGCTCTGTCTAGAGTC GCTGCGCGTGCCTTCGCCCCGTGCCCGCTCCGCCGCCCTCGCGCCGCCCGCCCGGCTCTGACTGACCGCT TACTCCACAGGTGAGCGGGCGGGACGGCCTTCTCTCCGGCTGTAATTAGCGCTTGGTTTATTGACGGCTTGT TTCTTTTCTGTGGCTGCGTGAAAGCCTTGAGGGGCTCCGGAAAGGCCCTTGTGCGGGGGAGCGGCTCGGGGGT GCGTGCGTGTGTGTGCGTGGGAGCGCCGCTGCGGCTCCGCGCTGCCGGCGGCTGTGAGCGCTCGGGGCGG GCGCGGGGCTTTGTGCGCTCCGAGTGTGCGGAGGGGAGCGCGGCCGGGGCGGTGCCCGCGGTGCGGGGGGG CTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCTCGGTGCGG CTGCAACCCCCCTGCACCCCCCTCCCGAGTTGCTGAGCACGGCCCGGCTTCCGGTGCGGGGCTCCGTACGGGGC GTGGCGCGGGGCTCGCGTGCGGGGCGGGGGTGGCGGAGGTGGGGGTGCCGGGCGGGGCGGGGGCGCCTCGGGC CGGGAGGGGCTCGGGGGAGGGGCGGGCGGCCCGGAGCGCCGGCGGTGTCGAGGCGCGGCGAGCCGACGCT TGCCTTTTATGGTAATCGTGCAGAGGGCGCAGGGACTTCTTTGTCCAAATCTGTGCGGAGCCGAAATCTGGGA GGCGCGCGCACCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCGCGAGGAAGGAAATGGGCGGGGAGGG CCTTCGTGCGTGCCTCGCGGCCCGCTCCCTTCTCCCTCTCCAGCTCGGGGTGTCCGCGGGGGGACGGCTGCCTT CGGGGGGACGGGACAGGGCGGGGTTCCGCTTCTGCGTGTGACCGCGGCTCTAGAGCTCTGCTAACCTTGTTC TTGCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATATTTTGGCAAAGAATT ACTCAGAGGCCACCATGCAGATTGAAGTGTCCACTTGCTTCTTCTGTGCTCCTGCGGTTTTGCTTCTCGGCCACC CGCCGGTATTACTTAGGTGCTGTGGAAGTGTGAGTGGGACTACATGCAGTCCGACCTGGGAGAACTGCCGGTGGACG CGAGATTCCCACCTAGAGTCCCGAAGTCTTCCACTCAACACCTCCGTGGTCTACAAAAAGACCTGTTTCGTGGA GTTCACTGACCACCTTTTCAATATTGCAAGCGCGCCCGCTGGATGGGCTCTTGTGCTTACGATCCAGCA GAGGTCTACGACACCGTGGTCTACACTGAAGAACATGGCTCACACCCGTGTGCTGCTGCTGTGGAGTGT CCTACTGGAAGGCTCAGAGGGTGCCTAATATGATGACAGACAGCCAGAGGGGAAAGGAGGATGACAAAGTGT CCCGGTGGCAGCCACACTTACGTGTGCAAGTGTGAAGGAAACGGGCTATGGCGTGGACCCCTATGCTG ACCTACTCCTACCTGTCCATGTGGACCTTGTGAAGATCTCAACTCGGGACTGATCGGCGCCCTTGTGTGTGCA GAGAAGGCAGCTGGCGAAGGAAAAGACTCAGACCTGCACAAGTTCATTCTGTTGTTTGTGTGTGCTGATGAAGG AAAGTCTGGCACTCAGAAACCAAGAACTCGCTGATGCAGGATAGAGATGCGGCCTCGGCCAGAGCCTGGCCTAAA ATGCACACCGTCAACGGATATGTGAACAGGTGCTCCCTGGCCCTCATCGGCTGCCACAGAAAGCTCGTATTGGC ATGTGATCGGCATGGTACTACTCCGGAAGTGCATAGTATCTTTCTGGAGGGCCATCTTCTGGTGCACACCA CAGACAGGCTCGCTGGAATCTCGCTATCACTTTCTGACTGCGCAGACCTCTTATGGACCTTGGACCTTCTC CTGTGTTCTGTACATCAGCTCCATCAGCATGATGGGATGGAGGCTATGTCAAAGTGGACTCTGCCCTGAGG AGCCACAGCTCCGGATGAAGAACAATGAGGAAGCGGAGGATTACGACGACGACCTGACTGACAGCGAAATGGACGT CGTGCGATTGATGACGACAACAGCCGCTCTTATCCAAATTAGATCAGTGGCGAAGAAGCACCCCAAGACCTGG GTGCACTACATTGCCGCGGAGGAAGAGGACTGGGACTACGCGCGCTGGTGTGCGGCCAGACGACAGGAGCTACA AGTCCAGTACCTCAACAACGGGCGCAGCGCAATTGGCAGGAAGTACAAGAAAGTCCGCTTATGGCTACACTGA TGAAACCTTCAAGACGAGGGAAGCCATCCAGCAGAGTCAAGCATCTGGGACCGCTCTTTACGGCGAAGTCCGG GATACCTGCTCATCATTTTCAAGAACCAGGCATCGCGCCCTACAACATCTACCTCACGGGATCACAGACCTGC GCCCCCTCACTCCCGCGGCTGCCAAGGGAGTGAAGCACTGAAGGATTTTCCATCCTCGCGGAGAAATCTT CAAGTACAAGTGGACCGTACTGTGGAAGATGGCCCTACCAAGTGGACCTCGCTGTGACCCGGTACTATTCC TCGTTTGTGAACATGGAGCGCGACCTGGCTCGGGGCTGATTGGTCCGCTGCTGATCTGCTACAAGGAGTCCGTGG ACCAGCGCGGGAACAGATCATGTCCGACAAGCGCAACGTGATCTGTTCTGTCTTTGATGAAAACAGATCGTG GTACTTGACTGAGAATATCCAGCGTTCTGCCCCAACCCAGCGGGAGTGAACTGGAGGACCCGGAGTTCCAGGCC TCAAACATTATGCACTCTATCAACGGCTATGTGTTGACTCGCTCCAAGTGAAGCGTGTGCTGATGAAGTGGCAT ACTGGTACATTCTGTCCATCGGAGCCAGACCGACTTCTGTCCGTGTTCTTCTCCGGATACACCTTCAAGCATAA GATGGTGTACGAGGACACTCTGACCTCTTCCATTTTCCGGAGAACTGTGTTTCAATGGAAGAACCCGGGC TTGTGATTCTGGGTTGCCATAACTCGGACTTCCGGAATAGAGGGATGACCCCTGTGAAAGTGTCCAGCTGTG ACAAGAATACCGGCGATTACTACGAGGACAGCTATGAGGACATCTCCGCTTATCTGCTGTCCAAGAACAACGCCAT TGAACCCAGGTCTTCTCCAAAACGGTGCACCGACCTCCGAAAGCGCCACCCAGAGTCAAGACCTGGCTCGGAA CCGGCTACCTCGGGCTCAGAGACACCGGGGACTTCCGAGTCCGCAACCCCGAGAGTGGACCCGGATCCGAACAG CAACCTCAGGATCAGAAACCCCGGGAATTCGGAATCCGCCACTCCCGAGTCCGGGACCGGACCTCCACTGAGCC TTCCGAGGGAAGCGCCCCGGATCCCTGCTGGATCCCTACCAGCAC TGAAGAAGGACCTCAGAATCCGCGACC CCTGAGTCCGGCCCTGGAAGCGAACCCGCCACTCCGGTCCGAAACCCCTGGGACTAGCGAGAGCGCCACTCCGG AATCGGGCCCAGGAAGCCCTGCCGATCCCGACAGCAGGAGGGAAGCCCGGGGTCAACGACTTCCAG TGAGGAGGAGCCTCATCCCCCGTGTGAAGCGGACATCAAAGAGAGATCACCAGGACACTCTCCAGTCCGAT CAGGAAGAAATTGACTACGACGATACTATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGATGAGGATG AGAACCAGTCCCCCGGAGCTTTCAGAAGAAAACCCGCCACTACTTATCGCTGCCGTGGAGCGGCTGTGGGATTA CGGGATGTCCAGTCAACCGCATGTGCTGCGGAATAGAGCGCAGTCAAGATCGGTGCCCGAGTTCAAGAAGGTCTGT
--------------------------------------	--

	TTCCAAGAGTTACCCGACGGGTCTTCACTCAACCCCTGTACCGGGGCGAACTCAACGAACACCTGGGACTGCTTG GGCCGTATATCAGGGCAGAAGTGGAAGATAACATCATGGTCACCTTCCGCAACCAGGCCCTCCGGCCGTACAGCTT CTACTCTTCACTGATCTCCTACGAGGAAGATCAGCGGCAGGGAGCCGAGCCCCGGAAGAACTTCGTCAAGCCTAAC GAAACTAAGACCTACTTTTGGAAAGGTCCAGCATCACATGGCCCCGACCAAAGACGAGTTCGACTGTAAAGCCTGGG CCTACTTCTCCGATGTGGACCTGGAGAAGGACGTGCACTCGGGACTCATTGGCCCCGCTCCTTGTGTGCCATACTAA TACCCTGAACCTGCTCACGGTCGCCAAGTCACAGTGCAGGAGTTCGCCCTCTTCTTACCATCTTCGATGAAACA AAGTCCTGGTACTTTACTGAGAACATGGAACGCAATTGCAGGGCACCTGCAACATCCAGATGGAAGATCCCACCT TCAAGGAAAACCTACCGGTTTCATGCCATTAACGGCTACATAATGGACACGTTCGACAGGACTGGTCATGGCCCAGGA CCAGAGAATCCGGTGGTATCTGCTCTCCATGGGCTCCAACGAAAACATTCACAGCATTCAATTTTCCGGCCATGTG TTCACCGTCCGGAAGAAGGAAGAGTACAAGATGGCTCTGTACAACCTCTACCCTGGAGTGTTCGAGACTGTGGA TGCTGCCTAGCAAGGCCGGCATTGGAGAGTGGAATGCCTGATCGGAGAGCATTGACGCCGGAATGTCCACCCCT GTTTCTTGTGTAATCCAACAAGTGCCAGACCCCGCTGGGAATGGCCTCAGGTCAATATTAGGGATTTCCAGATCACT GCTTCGGGGCAGTACGGGCAGTGGGCACCTAAGTTGGCCCCGGCTGCACTACTCTGGCTCCATCAATGCCTGGTCCA CCAAGGAACCTTCTCCTGGATTAAAGTGGACCTCCGGCCCCAATGATTATTCACGGTATTAAGACCCAGGGTGC CCGACAGAAGTTCTCCTCACTCTACATCTCGCAATTATCATATAATGTACAGCCTGGATGGGAAGAAGTGGCAGACC TACCGGGGAACTCCACTGGAACGCTCATGGTGTCTTTCGGCAACGTGGACTCCTCCGGCATTAAAGCACAACATCT TCAACCTCCGATCATTGCTCGGTACATCCGGCTGCACCAACTACTACAGCATCCGGTCCACCTGCGGATGGA ACTGATGGGTTGTGACCTGAACTCTGCTCCATGCCCTTGGGATGGAATCCAAGGCCATTAGCGATGCACAGATC ACCGCTCTTTCATACTTACCAACATGTTGCGGACCTGGTCCCGCTCGAAGGCCCGCTGCACCTCAAGGTGCT CCAATGCGTGGCGGCCTCAAGTGAACAACCCCAAGGAGTGGCTCCAGGTCGACTTCCAAAGACCATGAAGGTAC CGGAGTGACCAACCCAGGGCGTGAAGTCCCTGCTGACCTCTATGTACGTTAAGGAGTTCTCATCTCTCAAGCCAA GACGGACATCAGTGGACCTGTTCTTCCAAAACGAAAAAGTCAAAGTATTCAGGGCAACCAGGACTCCTTACCC CTGTGGTCAACAGCCTGGACCCCCATTGCTGACCCGCTACCTCCGCATCCACCCCAAGCTGGGTCCACAGAT CGCACTGCGCATGGAGGTCTTGGATGCGAAGCCCAAGATCTGTACTAAGCGGCCGCTCATAATCAACCTCTGGAT TACAAAATTTGTGAAAGATTGACTGGTATTCTTAATATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAA TGCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCTCCTGTATAAATCCTGGTTGCTGTCTCT TTATGAGGAGTTGTGGCCCGTGTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGTGTCAGCAACCCCCAC TGGTGGGGCATTGGCAACACCTGTGAGCTCCTTTCGGGACTTTCGCTTTCCTTCCCTATTGCCACGGCGGAAC TCAATCATGTCCTTTCCTTGGCTGCTCGCTGTGTTGCCACCTGGATTCTGCGCGGGAGCTCCTTCTGCTACGCTCCT TCCGCCCTCAATCCAGCGGACCTTCTTCCCGCGGCTGCTGCCGCTCTGCGGCTCTTCCGCTTCTCGCTTTC GCCCTCAGACGAGTCCGATCTCCTTGGGCGGCTCCCGCTGCCTAGGCGACTGTGCCTTCTAGTTGCCAGCCA TCTGTTGTTTGGCCCTCCCGCTGCTTCTTACCCTGGAAGGTGCCACTCCACTGTCTTCTAATAAAATG AGGAAATTGCATGCATTGTCTGAGTAGGTGTCTATTCTATTCTGGGGGTGGGTGGGGCAGGACAGCAAGGGGGA GGATTGGGAAGACAATAGCAGGCATGTGGGGAAGACCATGGGCGCGCCAGGCTGTGACGCCCGGGCGGTACCG CGATCGCTCGCGACGCATAAAG
SEQ ID NO: 36 promotor de la antitripsina alfa-1 humana (A1AT)	ATCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGTCCAGTGGCCCTTGGCAGCATTACTCTCTGT TGCTCTGGTTAATAATCTCAGGAGCACAAACATTCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTATCCTC TGGGCCTCTCCCCACCTTCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGTCCAGTGGCCCTTGGCAGC ATTTACTCTCTCTGTTTGTCTGGTTAATAATCTCAGGAGCACAAACATTCTGGAGGCAGGAGAAGAAATCAACA TCCTGGACTTATCCTCTGGGCTCTCCCCACCTTCGAACTAGCCACTAGCCTGAGGCTGGTCAAAATTGAACCTCC TCCTGCTCTGAGCAGCTGGGGGCGAGCTAAGCAGAGGGCTGTGACAGCCACATAAAGAGCCTACTGTGTGCCA GGCACTTACCCGAGGCACTTCAAGCATGCTTGGGAATGAACTTCCAACCTTTGGGATGCAGGTGAAACAGT TCCTGGTTTCAAGAGAGGTGAAGCGGCTGCCTGAGGCAGCAGCTCTTCTTACAGATGTGCTTCCCCACCTTAC CCTGTCTCACGGCCCCCATGCCAGCCTGACGTTGTGTCTGCCCTCAGTCATGCTCATTTTTCCATCGGGACCAT CAAGAGGGTGTGTGCTAAGGCTGACTGGGTAACCTTGGATGAGCGGTCTCTCCGCTCTGAGCCTGTTTCTCA TCTGTCAAATGGGCTCTAACCCTCTGATCTCCAGGGCGGCAGTAAGTCTTCAAGCATCAGGCATTTTGGGGTGA CTCAGTAAATGGTAGATCTTGCTACAGTGGAACAGCCACTAAGGATTCTGAGTGAGAGCAGAGGGCCAGCTAAG TGGTACTCTCCAGAGACTGTCTGACTCACGCCACCCCTCCACCTTGGACACAGGACGCTGTGGTTTCTGAGCCA GGTACAATGACTCCTTTCGGTAAGTGCAGTGGAAGCTGTACACTGCCAGGCAAGCGTCCGGGACGCTAGGCGG GCGACTCAGATCCAGCCAGTGACTTAGCCCTGTTTGTCTCCGATAACTGGGGTGACCTTGGTTAATATTCA CCAGCAGCCTCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAGGACAGGGCCCTGTCTCCTCAGCTT CAGGCACCACCACTGACCTGGGACAG