

5 **[0001]**El contenido del listado de secuencias presentado en formato electrónico en archivo de texto ASCII (Nombre: 2159_469PC02_SequenceListing_ST25.txt; tamaño: 204.138 bytes; y fecha de creación: 31 de enero de 2017) que se incorpora a la presente en su totalidad mediante esta referencia.

[0002] La vía de coagulación sanguínea, en parte, implica la formación de un complejo enzimático del factor VIIIa (FVIIIa) y factor IXa (FIXa) (complejo Xase) en la superficie de las plaquetas. La FIXa es una serín proteasa con actividad catalítica relativamente débil sin su cofactor FVIIIa. El complejo Xase escinde el factor X (FX) en el factor Xa (FXa), el cual a su vez interactúa con el factor Va (FVa) para escindir la protrombina y generar trombina. La hemofilia A es un trastorno hemorrágico provocado por las mutaciones y/o eliminaciones en el gen del FVIII (FVIII), lo cual genera deficiencia de la actividad del FVIII (Peyvandi *et ál.* 2006). En algunos casos, los pacientes tienen niveles reducidos de FVIII debido a la presencia de inhibidores de FVIII, por ejemplo, los anticuerpos anti-FVIII.

[0003] La hemofilia A se caracteriza por hemorragia espontánea y sangrado excesivo. Con el paso del tiempo, el sangrado repetido en los músculos y articulaciones, lo cual a menudo comienza en la primera infancia, provoca artropatía hemofílica y lesión articular irreversible. Esta lesión es progresiva y puede llevar a una movilidad de las articulaciones gravemente limitada, atrofia muscular y dolor crónico (Rodríguez-Merchan, E.C., *Semin. Thromb. Hemost.* 29:87-96 (2003), que se incorpora a la presente en su totalidad mediante esta referencia).

30 **[0004]** La enfermedad se puede tratar por medio de terapia sustitutiva que se dirige a la restauración de la actividad del FVIII hasta 1 a 5 % de los niveles normales para prevenir el sangrado espontáneo (véase, por ejemplo, Mannucci, P.M., et ál., *N. Engl. J. Med.* 344:1773-9 (2001), que se incorpora a la presente mediante esta referencia en su totalidad. Existen productos de FVIII recombinantes y que derivan
35 del plasma disponibles a demanda para tratar los episodios de sangrado o para evitar que se produzcan episodios de sangrado al tratarlos de forma profiláctica. De

acuerdo con la semivida de estos productos (10-12 h) (White G.C., et ál., *Thromb. Haemost.* 77:660-7 (1997); Morfini, M., *Haemophilia* 9 (suppl 1):94-99; discussion 100 (2003)), los regímenes de tratamiento exigen administración intravenosa frecuente, comúnmente dos a tres veces por semana para profilaxis y una a tres veces diarias para el tratamiento a demanda (Manco-Johnson, M.J., et ál., *N. Engl. J. Med.* 357:535-544 (2007)), cada uno de las cuales se incorpora a la presente en su totalidad mediante esta referencia. Dicha administración frecuente es inconveniente y costosa.

10 [0005] Un gran obstáculo para proporcionar una proteína de FVIII recombinante de bajo costo a pacientes es el alto costo de la producción comercial. La proteína de FVIII se expresa en forma deficiente en sistemas de expresión heterólogos, dos a tres órdenes de magnitud inferior a las proteínas con un tamaño similar. (Lynch et ál., *Hum. Gene. Ther.*; 4:259-72 (1993). La escasa expresión de FVIII se debe, en parte, a la presencia de elementos que actúan en cis en la secuencia de codificación de FVIII que inhibe la expresión de FVIII, tal como los elementos silenciadores transcripcionales (Hoeben et ál., *Blood* 85:2447-2454 (1995)), secuencias similares al acoplamiento a la matriz (MAR) (Fallux et ál., *Mol. Cell. Biol.* 16:4264-4272 (1996)) y los elementos inhibidores de elongación transcripcionales (Koeberl et ál., *Hum. Gene. Ther.*; 6:469-479 (1995)).

[0006] Los avances en nuestra comprensión de la biología de la expresión del FVIII nos condujeron al desarrollo de variantes de FVIII más potentes. Por ejemplo, los estudios bioquímicos demostraron que el dominio B del FVIII era prescindible para la actividad del cofactor de FVIII. La eliminación del dominio B provocó un aumento de 17 veces en los niveles de ARNm por sobre el FVIII de tipo salvaje y longitud completa y un 30 % de aumento en la proteína seleccionada. (Toole et ál., *Proc Natl Acad Sci EUA* 83:5939-42 (1986)). Esto llevó al desarrollo del concentrado de proteínas de FVIII con dominio B eliminado (BDD, por sus siglas en inglés), el cual es ampliamente utilizado en la clínica. Los estudios recientes, sin embargo, indican que la longitud completa y el mal plegado del hFVIII con BDD en el lumen ER provocan la activación de la respuesta a proteínas mal plegadas (UPR, por sus siglas en inglés) y apoptosis de hepatocitos murinos.

35 [0007] Por lo tanto, existe la necesidad en la técnica de secuencias de FVIII que se expresen eficazmente en los sistemas heterólogos.

COMPENDIO DE LA DESCRIPCIÓN

- [0008]** Se describen moléculas de ácido nucleico optimizadas con codones que codifican un polipéptido con actividad de FVIII. En un aspecto, la descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido del factor VIII (FVIII) y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido del FVIII; donde la primera secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-2277 y 2320-1791 de la SEQ ID NO: 3 o (ii) los nucleótidos 58-2277 y 2320-1791 de la SEQ ID NO: 4; y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII.
- [0009]** La descripción proporciona además una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido del FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido del FVIII; donde la segunda secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 o (ii) 1792-2277 y 2320-4374 de la SEQ ID NO: 6; y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica del FVIII.
- [0010]** En algunas modalidades, la descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 1. En otras modalidades, la descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 2. En una modalidad, la

descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 70. En otra modalidad, la descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 71.

[0011] En algunas modalidades, la descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 3.

[0012] En algunas modalidades, la descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 4.

[0013] En algunas modalidades, la descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 89 %, al menos 90 %, al menos

91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 5.

5 **[0014]**En algunas modalidades, la descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96
10 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 6.

[0015]En algunas modalidades, la descripción proporciona un método para aumentar la expresión de un polipéptido con actividad de FVIII en un sujeto que
15 comprende administrar una molécula de ácido nucleico aislada o un vector descrito en la presente a un sujeto que lo necesita, donde la expresión del polipéptido se aumenta con respecto a una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16 o el vector que comprende la molécula de ácido nucleico de referencia.

20

[0016]En algunas modalidades, la descripción proporciona un método para tratar un trastorno hemorrágico que comprende: administrarle a un sujeto que lo necesite una molécula de ácido nucleico, un vector o un polipéptido descrito en la presente.

25

MODALIDADES

[0017]E1. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido del factor VIII (FVIII) y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido
30 del FVIII; donde la primera secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-2277 y 2320-1791 de la SEQ ID NO: 3 o (ii) los nucleótidos 58-2277 y 2320-1791 de la SEQ ID NO: 4; y donde la parte de extremo N y la parte de extremo C juntas tienen una
35 actividad polipeptídica de FVIII.

[0018]E2. La molécula de ácido nucleico aislada de E1, donde la primera secuencia de nucleótidos comprende los nucleótidos 58-1791 de la SEQ ID NO: 3 o los nucleótidos 58-1791 de la SEQ ID NO: 4.

5 **[0019]**E3. La molécula de ácido nucleico aislada de E1 o E2, donde la segunda secuencia de nucleótidos que tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la secuencia de nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 3 o 1792-2277 y 2320-4374 de la SEQ ID
10 NO: 4.

[0020]E4. La molécula de ácido nucleico aislada de E1 o E2, donde la segunda secuencia de nucleótidos comprende los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 3 o 1792-2277 y 2320-4374 de la SEQ ID NO: 4.
15

[0021]E5. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde
20 la segunda secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 o (ii) 1792-2277 y 2320-4374 de la SEQ ID NO: 6; y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de
25 FVIII.

[0022]E6. La molécula de ácido nucleico aislada de E5, donde la segunda secuencia de ácido nucleico comprende los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 o 1792-2277 y 2320-4374 de la SEQ ID NO: 6.
30

[0023]E7. La molécula de ácido nucleico aislada de E5 o E6, donde la primera secuencia de ácido nucleico tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la secuencia de nucleótidos 58-1791 de la SEQ ID NO: 5 o los nucleótidos 58-1791 de la SEQ ID NO: 6.
35

[0024]E8. La molécula de ácido nucleico aislada de E5 o E6, donde la primera secuencia de ácido nucleico comprende los nucleótidos 58-1791 de la SEQ ID NO: 5 o los nucleótidos 58-1791 de la SEQ ID NO: 6.

5 **[0025]E9.** Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la
10 SEQ ID NO: 1.

[0026]E10. La molécula de ácido nucleico aislada de E9, donde la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 1.

15 **[0027]E11.** Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la
20 SEQ ID NO: 2.

[0028]E12. La molécula de ácido nucleico aislada de E10, donde la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 2.

25 **[0029]E13.** Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al
30 menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 70.

[0030]E14. La molécula de ácido nucleico aislada de E13, donde la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 70.

35

[0031]E15. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia

de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 71.

[0032]E16. La molécula de ácido nucleico aislada de E15, donde la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 71.

[0033]E17. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 3.

[0034]E18. La molécula de ácido nucleico aislada de E17, donde la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 3.

[0035]E19. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 4.

[0036]E20. La molécula de ácido nucleico aislada de E19, donde la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 4.

[0037]E21. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de

identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 5.

5 **[0038]**E22. La molécula de ácido nucleico aislada de E21, donde la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 5.

[0039]E23. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 89
10 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 6.

15 **[0040]**E24. La molécula de ácido nucleico aislada de E23, donde la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 6.

[0041]E25. La molécula de ácido nucleico aislada de cualquiera de E1 a E24, donde la secuencia de nucleótidos comprende además una secuencia de ácido
20 nucleico que codifica un péptido señal.

[0042]E26. La molécula de ácido nucleico aislada de E25, donde el péptido señal es un péptido señal de FVIII.

25 **[0043]**E27. La molécula de ácido nucleico aislada de E25 o E26, donde la secuencia de ácido nucleico que codifica un péptido señal está optimizada con codones.

[0044]E28. La molécula de ácido nucleico aislada de cualquier de E25 a E27,
30 donde la secuencia de ácido nucleico que codifica un péptido señal tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o al menos 100 % de identidad de secuencia con respecto a (i) la secuencia de nucleótidos 1 a 57 de la SEQ ID NO: 1; (ii) los nucleótidos 1 a 57 de la SEQ ID NO: 2; (iii) los nucleótidos 1 a 57 de la
35 SEQ ID NO: 3; (iv) los nucleótidos 1 a 57 de la SEQ ID NO: 4; (v) los nucleótidos 1 a 57 de la SEQ ID NO: 5; (vi) los nucleótidos 1 a 57 de la SEQ ID NO: 6; (vii) los

nucleótidos 1 a 57 de la SEQ ID NO: 70; (viii) los nucleótidos 1 a 57 de la SEQ ID NO: 71; o (ix) los nucleótidos 1 a 57 de la SEQ ID NO: 68.

[0045]E29. La molécula de ácido nucleico aislada de cualquiera de E1 a E28, donde la molécula de ácido nucleico comprende una o más propiedades que se seleccionan del grupo que consiste en:

- (a) el índice de adaptación de codones humanos de la molécula de ácido nucleico o una parte de esta se incrementa con respecto a la SEQ ID NO: 16;
- (b) la frecuencia de codones óptimos de la secuencia de nucleótidos o una parte de esta se incrementa con respecto a la SEQ ID NO:16;
- (c) la secuencia de nucleótidos o una parte de esta contiene un porcentaje de nucleótidos de G/C más alto en comparación con el porcentaje de nucleótidos de G/C en la SEQ ID NO: 16;
- (d) el uso de codones sinónimos relativos de la secuencia de nucleótidos o una parte de esta se incrementa con respecto a la SEQ ID NO: 16;
- (e) la cantidad eficaz de codones de la secuencia de nucleótidos o una parte de esta con se reduce respecto a la SEQ ID NO: 16;
- (f) la secuencia de nucleótidos contiene menos secuencias MARS/ARS (SEQ ID NO: 21 y 22) con respecto a la SEQ ID NO: 16;
- (g) la secuencia de nucleótidos contiene menos elementos desestabilizantes (SEQ ID NO: 23 y 24) con respecto a la SEQ ID NO: 16; y
- (h) cualquier combinación de ellos.

[0046]E30. La molécula de ácido nucleico aislada de cualquiera de E1 a E29, que comprende además una secuencia de nucleótidos heteróloga.

[0047]E31. La molécula de ácido nucleico aislada de E30, donde la secuencia de nucleótidos heteróloga codifica una secuencia de aminoácidos heteróloga que es un extensor de semivida.

[0048]E32. La molécula de ácido nucleico aislada de E30 o E31, donde la secuencia de aminoácidos heteróloga es una región constante de inmunoglobulina o una parte de esta, transferrina, albúmina o una secuencia PAS.

5 **[0049]**E33. La molécula de ácido nucleico aislada de E30 o E31, donde la secuencia de aminoácidos heteróloga es una región Fc.

[0050]E34. La molécula de ácido nucleico aislada de cualquiera de E30 a E33, donde la secuencia de aminoácidos heteróloga está unida al extremo N o extremo C
10 de la secuencia de aminoácidos codificada por la secuencia de nucleótidos o está insertada entre dos aminoácidos en la secuencia de aminoácidos codificada por la secuencia de nucleótidos.

[0051]E35. La molécula de ácido nucleico aislada de E34, donde la secuencia de
15 aminoácidos heteróloga está insertada entre dos aminoácidos en uno o más sitios de inserción que se seleccionan de la Tabla 3.

[0052]E36. La molécula de ácido nucleico aislada de cualquiera de E30 a E35, que
20 codifica una molécula híbrida de monómero-dímero que comprende FVIII.

[0053]E37. La molécula de ácido nucleico aislada de cualquiera de E1 a E36, donde el polipéptido de FVIII es un FVIII de longitud completa o un FVIII con dominio B eliminado.

25 **[0054]**E38. La molécula de ácido nucleico aislada de cualquiera de E1 a E37, unida de forma operativa a al menos una secuencia de control de transcripción.

[0055]E39. Un vector que comprende la molécula de ácido nucleico de cualquiera de E1 a E38.

30

[0056]E40. El vector de E39, donde el vector es un vector viral.

[0057]E41. Una célula hospedadora que comprende la molécula de ácido nucleico de cualquiera de E1 a E32 o el vector de E39 o E40.

35

[0058]E42. La célula hospedadora de E41, donde la célula hospedadora se selecciona del grupo que consiste en: célula CHO, una célula HEK293, a célula BHK21, una célula PER.C6, una célula NS0 y una célula CAP.

- 5 **[0059]**E43. Un polipéptido codificado por la molécula de ácido nucleico de cualquiera de E1 a E37 o el vector de E39 o E40 o producido por la célula hospedadora de E41 o E42.

- 10 **[0060]**E44. Un método para producir un polipéptido con actividad de FVIII que comprende: cultivar la célula hospedadora de E41 o E42 en condiciones en las cuales se produce un polipéptido con actividad de FVIII; y, recuperar el polipéptido con actividad de FVIII.

- 15 **[0061]**E45. El método de E44, donde se incrementa la expresión del polipéptido con actividad de FVIII con respecto a una célula hospedadora cultivada en las mismas condiciones que comprende una secuencia de nucleótidos de referencia que comprende la SEQ ID NO: 16.

- 20 **[0062]**E46. 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 cualquiera de E1 a E38 o el vector de E39 o E40 a un sujeto que lo necesita, donde se incrementa la expresión del polipéptido con respecto a una molécula de ácido nucleico aislada de referencia que comprende la SEQ ID NO: 16 o el vector que comprende la molécula de ácido nucleico de referencia.

- 25 **[0063]**E47. Un método para aumentar la expresión de un polipéptido con actividad de FVIII que comprende cultivar la célula hospedadora de E41 o E42 en condiciones en las cuales se expresa un polipéptido con actividad de FVIII mediante la molécula de ácido nucleico, donde se incrementa la expresión del polipéptido con actividad de FVIII con respecto a una célula hospedadora cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16.

- 35 **[0064]**E48. El método de cualquiera de E44 a E47, donde se incrementa la expresión del polipéptido de FVIII en al menos aproximadamente 2 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 6 veces, al menos

aproximadamente 7 veces, al menos aproximadamente 8 veces, al menos
aproximadamente 9 veces, al menos aproximadamente 10 veces, al menos
aproximadamente 11 veces, al menos aproximadamente 12 veces, al menos
aproximadamente 13 veces, al menos aproximadamente 14 veces, al menos
5 aproximadamente 15 veces, al menos aproximadamente 20 veces, al menos
aproximadamente 25 veces, al menos aproximadamente 30 veces, al menos
aproximadamente 35 veces, al menos aproximadamente 40 veces, al menos
aproximadamente 50 veces, al menos aproximadamente 60 veces, al menos
aproximadamente 70 veces, al menos aproximadamente 80 veces, al menos
10 aproximadamente 90 veces o al menos aproximadamente 100 veces.

[0065]E49. Un método para mejorar el rendimiento de un polipéptido con actividad
de FVIII que comprende cultivar la célula hospedadora de E41 o E42 en condiciones
en las cuales se produce un polipéptido con actividad de FVIII mediante la molécula
15 de ácido nucleico, donde se incrementa el rendimiento del polipéptido con actividad
de FVIII con respecto a una célula hospedadora cultivada en las mismas condiciones
que comprende una secuencia de ácido nucleico de referencia que comprende la
SEQ ID NO: 16.

20 **[0066]**E50. Un método para tratar un trastorno hemorrágico que comprende:
administrarle a un sujeto que lo necesita una molécula de ácido nucleico de
cualquiera de E1 a E38, el vector de E39 o E40 o el polipéptido de E43.

[0067]E51. El método de E50, donde el trastorno hemorrágico se caracteriza por
25 una deficiencia en FVIII.

[0068]E52. El método de E50, donde el trastorno hemorrágico es hemofilia.

[0069]E53. El método de E50, donde el trastorno hemorrágico es hemofilia A.
30

[0070]E54. El método de cualquiera de E50 a E53, donde se incrementa la
actividad de FVIII en el plasma a las 24 horas después de la administración con
respecto a un sujeto al que se le administró una molécula de ácido nucleico de
referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula
35 de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido
nucleico de referencia.

- 5 [0071]E55. El método de E54, donde se incrementa la actividad de FVIII en el plasma en al menos aproximadamente 2 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 6 veces, al menos aproximadamente 7 veces, al menos aproximadamente 8 veces, al menos aproximadamente 9 veces, al menos aproximadamente 10 veces, al menos aproximadamente 11 veces, al menos aproximadamente 12 veces, al menos aproximadamente 13 veces, al menos aproximadamente 14 veces, al menos aproximadamente 15 veces, al menos aproximadamente 20 veces, al menos aproximadamente 25 veces, al menos aproximadamente 30 veces, al menos aproximadamente 35 veces, al menos aproximadamente 40 veces, al menos aproximadamente 50 veces, al menos aproximadamente 60 veces, al menos aproximadamente 70 veces, al menos aproximadamente 80 veces, al menos aproximadamente 90 veces o al menos aproximadamente 100 veces.
- 10
- 15 [0072]E56. El vector de E39 o 40, donde vector es un vector lentiviral.
- [0073]E57. El vector de cualquiera de E39, E40 y E56 que comprende además un promotor específico del tejido, un potenciador específico del tejido o ambos, un promotor específico del tejido y un potenciador específico del tejido.
- 20
- [0074]E58. El vector de E57, donde el promotor específico del tejido y/o el potenciador específico del tejido mejora selectivamente la expresión del transgén en las células hepáticas.
- 25
- [0075]E59. El vector de E57 o E58, donde el promotor específico del tejido comprende una secuencia promotora que se selecciona del grupo que consiste en un promotor de tiretina de ratón (mTTR), un promotor del factor VIII humano endógeno (F8), promotor de alfa-1-antitripsina humano (hAAT), promotor mínimo de albúmina humana y promotor de albúmina de ratón.
- 30
- [0076]E60. El vector de cualquiera de E57 a E59, donde el promotor específico del tejido comprende el promotor mTTR.
- 35 [0077]E61. Un método para tratar un trastorno hemorrágico que comprende: administrarle a un sujeto que lo necesita el vector de cualquiera de E56 a E60.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0078] Las figuras 1A-1J proporcionan las secuencias de nucleótidos optimizadas con codones que codifican el factor VIII con dominio B eliminado (SEQ ID NO: 17). La figura 1A muestra la secuencia de nucleótidos de coFVIII-3 (SEQ ID NO: 1). La figura 1B muestra la secuencia de nucleótidos de coFVIII-4 (SEQ ID NO: 2). La figura 1C muestra la secuencia de nucleótidos de coFVIII-5 (SEQ ID NO: 70). La figura 1D muestra la secuencia de nucleótidos de coFVIII-6 (SEQ ID NO: 71). La figura 1E muestra la secuencia de nucleótidos de coFVIII-52 (SEQ ID NO: 3). La figura 1F muestra la secuencia de nucleótidos de coFVIII-62 (SEQ ID NO: 4). La figura 1G muestra la secuencia de nucleótidos de coFVIII-25 (SEQ ID NO: 5). La figura 1H muestra la secuencia de nucleótidos de coFVIII-26 (SEQ ID NO: 6). Las figuras 1I y 1J muestran las secuencias de aminoácidos y nucleótidos, respectivamente, con dominio B eliminado (BDD-FVIII) (SEQ ID NO: 16 y 17, respectivamente).

[0079] Las figuras 2A-2J muestran ajustes de tendencias de uso de codones en las secuencias de nucleótidos optimizadas con codones que codifican el BDD-FVIII. La figura 2A muestra la frecuencia relativa de los codones en la secuencia de nucleótidos de tipo salvaje (antes de la optimización con codones) que codifica el BDD-FVIII, por ejemplo, el BDD-FVIII no optimizado. El índice de adaptación de codones (CAI, por sus siglas en inglés) humano de la secuencia de BDD-FVIII no optimizado es 74 %. La figura 2B muestra la frecuencia relativa de los codones en la secuencia variante coFVIII-1, la cual tiene un CAI humano de 88 %. La figura 2C muestra la frecuencia relativa de los codones en la secuencia variante coFVIII-3, la cual tiene un CAI humano de 91 %. La figura 2D muestra la frecuencia relativa de los codones en la secuencia variante coFVIII-4, la cual tiene un CAI humano de 97 %. La figura 2E muestra la frecuencia relativa de los codones en la secuencia variante coFVIII-5, la cual tiene un CAI humano de 83 %. La figura 2F muestra la frecuencia relativa de los codones en la secuencia variante coFVIII-6, la cual tiene un CAI humano de 83 %. La figura 2G muestra la frecuencia relativa de los codones en la secuencia variante coFVIII-52, la cual tiene un CAI humano de 91 %. La figura 2H muestra la frecuencia relativa de los codones en la secuencia variante coFVIII-62, la cual tiene un CAI humano de 91 %. La figura 2I muestra la frecuencia relativa de los codones en la secuencia variante coFVIII-25, la cual tiene un CAI humano de 88 %. La figura 2J muestra la frecuencia relativa de los codones en la secuencia variante coFVIII-26, la cual tiene un CAI humano de 88 %.

[0080] La figura 3 proporciona un mapa del plásmido de FVIII-303, que comprende coFVIII-1 en una estructura principal de pcDNA3 bajo el control del promotor de transtiretina mejorado con ET, que está colocado delante del sitio de inicio de la traducción de coFVIII-1 y que comprende un potenciador sintético, un potenciador mTIR y un promotor mTIR.

[0081] La figura 4 muestra una representación gráfica de la actividad plasmática del FVIII en los ratones con HemA después de la inyección hidrodinámica de 5 µg de FVIII-303 (coFVIII-1; círculos) o 5 µg de FVIII-311 (BDD-FVIII; cuadrados). Se determinó la actividad plasmática del FVIII mediante un ensayo cromogénico específico de FVIII a las 24, 48 y 72 horas después de la inyección. Se muestran los niveles de actividad relativa a las 72 horas, normalizada con el nivel de expresión de FVIII-311.

[0082] La figura 5 proporciona un mapa del plásmido de pLV-coFVIII-52, que comprende coFVIII-52 en un plásmido lentiviral bajo el control de un promotor de ET, que se coloca delante del sitio de inicio de la traducción de coFVIII-52 y que comprende un potenciador sintético, un potenciador mTIR y un promotor mTIR.

[0083] Las figuras 6A-6C muestran representaciones gráficas de la actividad plasmática del FVIII en ratones con HemA después de la inyección hidrodinámica de diversos nucleótidos que codifican el FVIII. Se determinó la actividad plasmática del FVIII mediante un ensayo cromogénico específico para el FVIII a las 24, 48 y 72 horas después de la inyección. La figura 6A muestra la actividad plasmática del FVIII en ratones con HemA después de la inyección hidrodinámica de 5 µg de LV-coFVIII-1 (círculos rellenos), 5 µg de LV-coFVIII-3 (triángulos), 5 µg de LV-coFVIII-4 (triángulos invertidos), 5 µg de LV-coFVIII-5 (diamantes) o 5 µg de LV-coFVIII-6 (círculos abiertos). La figura 6B muestra la actividad plasmática del FVIII en ratones con HemA después de la inyección hidrodinámica de 5 µg de LV-coFVIII-1 (círculos), 5 µg de LV-coFVIII-25 (triángulos) o 5 µg de LV-coFVIII-26 (triángulos invertidos). La figura 6C muestra la actividad plasmática del FVIII en ratones con HemA después de la inyección hidrodinámica de 20 µg de LV-2116 (secuencia de nucleótidos de BDD-FVIII (WT) optimizada sin codones; círculos abiertos), 20 µg de LV-coFVIII-1 (triángulos), 20 µg de LV-coFVIII-52 (cuadrados) o 20 µg de LV-coFVIII-62 (círculos rellenos). Se muestran los niveles de actividad relativa a las 72 horas para cada plásmido, normalizados con los niveles de expresión de LV-coFVIII-1 (figuras 6A, 6B y 6C) y/o LV-2116 (figura 6C), como se indica.

[0084] La figura 7 muestra la actividad plasmática del FVIII en ratones con Hema 24 días después de la inyección con vector lentiviral de TU/ratón 1E8 que comprende coFVIII-1, coFVIII-5, coFVIII-52, coFVIII-6 o coFVIII-62 en comparación con el control LV-2116 (BDD-FVIII) y según se mide mediante un ensayo cromogénico específico para FVIII. Las barras de error indican desviaciones estándar.

[0085] Las figuras 8A-8C proporcionan las diversas secuencias de nucleótidos optimizadas con codones que codifican el BDD-FVIII fusionado con un XTEN. La figura 8A muestra la secuencia de nucleótidos de coFVIII-52-XTEN (SEQ ID NO: 19), donde una secuencia de nucleótidos que codifica un XTEN que tiene 144 aminoácidos («XTEN₁₄₄»; SEQ ID NO: 18; subrayado) se inserta dentro de la secuencia de nucleótidos de coFVIII-52. La figura 8B muestra la secuencia de nucleótidos de coFVIII-1-XTEN (SEQ ID NO: 20), donde una secuencia de nucleótidos que codifica un XTEN que tiene 144 aminoácidos («XTEN₁₄₄»; SEQ ID NO: 18; subrayado) se inserta dentro de la secuencia de nucleótidos de coFVIII-1. La figura 8C muestra la secuencia de nucleótidos de coFVIII-6-XTEN (SEQ ID NO: 72), donde una secuencia de nucleótidos que codifica un XTEN que tiene 144 aminoácidos («XTEN₁₄₄»; SEQ ID NO: 18; subrayado) se inserta dentro de la secuencia de nucleótidos coFVIII-6 (por ejemplo, residuo de aminoácidos 745 que corresponde a la secuencia FVIII madura).

[0086] La figura 9 proporciona un mapa del plásmido de pLV-coFVIII-52-XTEN, que comprende coFVIII-52-XTEN en un vector lentiviral bajo el control del promotor de ET. Los vectores lentivirales que comprenden cada una de las moléculas de ácido nucleico optimizadas con codones que codifica un polipéptido con actividad de FVIII, como se describe en la presente, se construyeron de la misma manera que pLV-coFVIII-52-XTEN, en el que se insertó la misma secuencia XTEN para remplazar el dominio B del FVIII.

[0087] Las figuras 10A y 10B muestran actividad de FVIII en ratones con Hema después de la inyección con el ADN del plásmido (figura 10A) o vector lentiviral (figura 10B) que comprende las diversas secuencias de nucleótidos optimizadas con codones que codifican el BDD-FVIII. La figura 10A muestra una representación gráfica de la actividad plasmática de FVIII en ratones con Hema después de la inyección hidrodinámica con 5 µg de FVIII-311 (optimizado sin codones, secuencia de nucleótidos que codifica el BDD-FVIII; cuadrados), 5 µg de FVIII-303 (coFVIII-1; círculos pequeños) o FVIII-306 (coFVIII-1-XTEN₁₄₄; círculos grandes). La actividad

relativa a las 72 horas, normalizada con FVIII-311 se muestra para cada plásmido. La figura 10B muestra la actividad plasmática de FVIII en ratones con HemA 21 días después de la inyección con vector lentiviral de 1E8 de TU/ratón que comprende coFVIII-52 o coFVIII-52-XTEN en comparación con el control LV-2116 (BDD-FVIII) y según se mide mediante un ensayo cromogénico específico para FVIII. Las barras de error indican desviaciones estándar.

[0088] La figura 11A muestra la secuencia de aminoácidos del factor VIII humano maduro de longitud completa. La figura 11B muestra la secuencia de aminoácidos del factor de von Willebrand humano de longitud completa (SEQ ID NO: 44). Las figuras 11C y 11D muestran las secuencias de nucleótidos y aminoácidos, respectivamente, de un polipéptido XTEN que tiene 42 aminoácidos (XTEN AE42-4; SEQ ID NO: 46 y 47, respectivamente). Se muestran las secuencias de aminoácidos de diversos polipéptidos XTEN que tienen 144 aminoácidos en las figuras 11E, 11G, 11I, 11K, 11M, 11O, 11Q, 11S, 11U y 11W (SEQ ID NO: 48, 50, 52, 54, 56, 58, 60, 62, 64 y 66, respectivamente) y las secuencias de nucleótidos correspondientes se muestran en las figuras 11F, 11H, 11J, 11L, 11N, 11P, 11R, 11T, 11V y 11X (SEQ ID NO. 49, 51, 53, 55, 57, 59, 61, 63, 65 y 67, respectivamente). La figura 11Y muestra la secuencia de nucleótidos de un promotor ET (SEQ ID NO: 69). La figura 11Z muestra la secuencia de nucleótidos para coFVIII-1 (SEQ ID NO: 68) (véase la publicación internacional n.º WO 2014/127215, SEQ ID NO: 1).

[0089] La figura 12A es una representación gráfica de la actividad plasmática del FVIII (IU/mL) en ratones con HemA de 14 días de edad después de la administración IV de aproximadamente 1.5 E10 TU/kg de LV-wtBDD-FVIII (círculos), LV-coFVIII-6 (cuadrados) o LV-coFVIII-6XTEN (triángulos). La figura 12B es una representación gráfica de la cantidad de copias del vector (VCN, por sus siglas en inglés) 150 días después del tratamiento de ratones con HemA de 14 días de edad administrado por IV aproximadamente 1.5 de E10 TU/kg de vectores lentivirales que comprende wtBDD-FVIII, coFVIII-1, coFVIII-3, coFVIII-4, coFVIII-5, coFVIII-6, coFVIII-52, coFVIII-62, coFVIII-25 o coFVIII-26. La figura 12C es una representación gráfica de la actividad plasmática de FVIII (IU/mL) 21 días después del tratamiento de ratones con HemA de 14 días de edad administrado por IV aproximadamente 1.5 de E10 TU/kg de vectores lentivirales que comprende wtBDD-FVIII, coFVIII-1, coFVIII-3, coFVIII-4, coFVIII-5, coFVIII-6, coFVIII-52, coFVIII-62, coFVIII-25 o coFVIII-26.

[0090] Las figuras 13A y 13B son representaciones gráficas que ilustran los niveles de actividad plasmática del FVIII (figuras 13A) y niveles de anticuerpos anti-FVIII (figura 13B) en cinco ratones con HemA tratados con un lentivirus que expresa la variante coFVIII-5. A las crías normales con HemA de la misma camada de catorce días de edad se les administró aproximadamente 1.5 E10 TU/kg de un lentivirus que expresa la variante coFVIII-5 mediante inyección intravenosa. Cada ratón es designado por medio de un número (es decir 1, 2, 3, 4 y 5; figuras 13A y 13B).

[0091] La figura 14 es una representación gráfica de la correlación entre el nivel de expresión de LV-FVIII, según se evidencia mediante la actividad plasmática del FVIII a los 21 días después del tratamiento lentiviral y la presencia de anticuerpos anti-FVIII. Cada punto de dato corresponde a un único ratón con HemA. Cada ratón recibió una dosis de 1,5 E10 TU/kg mediante inyección intravenosa de un lentivirus que expresa una de las variantes coFVIII descritas en la presente. Las líneas horizontales indican la actividad plasmática de FVIII promedio.

[0092] La figura 15 es una representación gráfica de la correlación entre la cantidad de copias del vector (VCN) por célula a los 150 días después del tratamiento lentiviral y la presencia de anticuerpos anti-FVIII. Cada punto de dato corresponde a un único ratón con HemA. Cada ratón recibió una dosis de 1,5 E10 TU/kg mediante inyección intravenosa de un lentivirus que expresa una de las variantes coFVIII descritas en la presente. Las líneas horizontales indican la VCN promedio.

[0093] Las figuras 16A y 16B son representaciones gráficas que ilustran los niveles de actividad plasmática de FVIII (figuras 16A) y niveles de anticuerpos anti-FVIII (figura 16B) en dos ratones con HemA (coFVIII-52-A y coFVIII-52-B) tratados con un lentivirus que expresa la variante coFVIII-52. A las crías normales con HemA de la misma camada de catorce días de edad se les administró aproximadamente 1,5 E10 TU/kg de un lentivirus que expresa la variante coFVIII-52 mediante inyección intravenosa. Las figuras 16C y 16D son imágenes que muestran la tinción por hibridación in situ del ARN para la expresión de FVIII (tinción oscura) en el tejido del hígado recolectado de coFVIII-52-A (figura 16C) y coFVIII-52-B (figura 16D) ratones de la figura 16A y 16B.

[0094] La figura 17 es una representación gráfica que muestra la expresión de FVIII a largo plazo en ratones neonatos con HemA tratados con un lentivirus que expresa un FVIII con dominio B eliminado de tipo salvaje (wtBDD-FVIII; triángulos), coFVIII-

52XTEN (círculos) o variante coFVIII-6XTEN (triángulo invertido). Se les administró a los ratones neonatos con HemA mediante inyección intravenosa aproximadamente 1,5 A10 TU/kg de un lentivirus que expresa wtBDD-FVIII, coFVIII-52XTEN o coFVIII-6XTEN. Se midió la actividad plasmática de FVIII a lo largo de aproximadamente 16 semanas.

[0095] La figura 18 es una representación gráfica del nivel de FVIII circulante en perros neonatos con HemA (S3 o K4) después de la administración de $1,3 \times 10^9$ de unidades de transducción/kg de vector lentiviral que comprenden un factor VIII que codifica nucleótidos fusionado con XTEN (SEQ ID NO: 72; LV-coFVIII-6-XTEN). Los cuadrados conectados por medio de una línea continua representan muestras de aPTT-S3 y los triángulos conectados por la línea punteada representan las muestras de aPTT-K4. El eje X muestra la actividad plasmática de FVIII como el porcentaje de normal, donde la actividad de FVIII humano es 100 %. El eje x muestra los días después del tratamiento lentiviral, donde el tratamiento lentiviral se administra el día 0.

[0096] Las figuras 19A-19C son representaciones gráficas de toda la hemostasia sanguínea según se monitorea mediante el ensayo de tromboelastometría rotacional (ROTEM, por sus siglas en inglés) para un perro con HemA sin tratamiento previo (figura 19A), perro S3 a las 2 semanas después del tratamiento lentiviral (figura 19B), y perro K4 a las 2 semanas después del tratamiento lentiviral (figura 19C). Se muestra el tiempo de coagulación (CT, por sus siglas en inglés) como segundos (s), se muestra el tiempo de formación de coágulos (CFT, por sus siglas en inglés) como segundos (s), se muestra el ángulo alfa (α) como grados ($^{\circ}$), se muestra la amplitud 5 minutos después del CT (A5) como milímetros (mm), se muestra la amplitud 20 minutos después CT (A20) como milímetros (mm) y se muestra la firmeza máxima de coágulos (MCF, por sus siglas en inglés) como milímetros (mm) para cada una de las figuras 19A-19C. La figura 19D es una tabla que resumen el intervalo normal para cada uno de los parámetros CT, CFT, α , A5, A2 y MCF que se muestran en las figuras 19A-19C.

DESCRIPCIÓN DETALLADA DE LA DESCRIPCIÓN

[0097] La presente descripción describe genes optimizaos con codones que codifican polipéptidos con actividad del factor VIII (FVIII). La presente descripción se refiere a moléculas de ácido nucleico optimizadas con codones que codifican polipéptidos con actividad de factor VIII, vectores y células hospedadoras que

comprenden moléculas de ácido nucleico optimizadas, polipéptidos codificados por moléculas de ácido nucleico optimizadas y métodos para producir dichos polipéptidos. La presente descripción se refiere también a métodos para tratar los trastornos hemorrágicos, por ejemplo, la hemofilia, que comprende administrar al

5 sujeto una secuencia de ácido nucleico del factor VIII optimizado, un vector que comprende la secuencia de ácido nucleico optimizada o el polipéptido codificado a través de esta. La presente descripción satisface una necesidad importante en la técnica al proporcionar secuencias del factor FVIII optimizadas que demuestran mayor expresión en las células hospedadoras, mejor rendimiento de la proteína del

10 factor VIII en métodos para producir el factor VIII recombinante, y potencialmente provoca mayor eficacia terapéutica cuando se utiliza en métodos de terapia génica. En determinadas modalidades, la descripción describe una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que tiene homología de secuencia con respecto a una secuencia de nucleótidos que se selecciona de las

15 SEQ ID NO: 1-14, 70 y 71.

[0098] Las construcciones de ejemplo de la descripción se ilustran en las figuras adjuntas y el listado de secuencias. Para proporcionar un mejor entendimiento de la memoria descriptiva y de las reivindicaciones, se proporcionan las siguientes

20 definiciones.

I. Definiciones

[0099] Se debe observar que el término «un» o «uno/a» se refiere a uno o más de esa entidad, por ejemplo, se entiende que «una secuencia de nucleótidos»

25 representa una o más secuencias de nucleótidos. Como tales, los términos «un» (o «uno/a»), «uno o más» y «al menos uno» pueden utilizarse de forma indistinta en la presente.

[0100] El término «aproximadamente» se usa en la presente para significar

30 aproximadamente, en líneas generales, alrededor de, o en las regiones de. Cuando el término «aproximadamente» se usa junto con un intervalo numérico, modifica ese intervalo extendiendo los límites por encima y por debajo de los valores numéricos establecidos. En general, el término «aproximadamente» se utiliza en la presente para modificar un valor numérico por encima y por debajo del valor indicado con una

35 variación del 10 %, hacia arriba y hacia abajo (más alta o más baja).

[0101] El término «aislado/a» a efectos de la presente descripción designa un material biológico (célula, polipéptido, polinucleótido o un fragmento, variante o derivado de estos) que ha sido quitado de su entorno original (el entorno en el que se encuentra naturalmente presente). Por ejemplo, un polinucleótido presente en el estado natural en una planta o un animal no está aislado, sin embargo, el mismo polinucleótido separado de los ácidos nucleicos adyacentes en los que está naturalmente presente, se considera «aislado». No se requiere ningún nivel de purificación particular. Las proteínas y los polipéptidos producidos de forma recombinante expresados en células hospedadoras se consideran aislados del objetivo de la invención, ya que son polipéptidos naturales o recombinantes que se separaron, fraccionaron, o purificaron parcial o sustancialmente mediante cualquier técnica adecuada.

[0102] «Ácidos nucleicos», «moléculas de ácido nucleico», «oligonucleótido» y «polinucleótido» se pueden usar de forma intercambiable y hacen referencia a la forma polimérica del éster de fosfato de los ribonucleótidos (adenosina, guanosina, uridina o citidina; «moléculas de ARN») o desoxirribonucleótidos (desoxiadenosina, desoxiguanosina, desoxitimidina o desoxicitidina; «molécula de ADN»), o cualquier análogo de fosfoéster de estos, tal como fosforotioatos y tioésteres en cualquier forma monocatenaria o una hélice de cadena doble. Son posibles las hélices de DNA-DNA, DNA-RNA y RNA-RNA de cadena doble. La expresión «molécula de ácido nucleico» y en particular, molécula de ADN o ARN, hace referencia a la estructura primaria y secundaria de la molécula y no la limita a ninguna forma terciaria particular. Por lo tanto, esta expresión incluye ADN de cadena doble que se encuentra, entre otros, en moléculas de ADN lineales o circulares (por ejemplo, fragmentos de restricción), plásmidos, ADN superhelicoidal y cromosomas. Al tratar la estructura de las moléculas de ADN de cadena doble particulares, en la presente se pueden describir secuencias de acuerdo con la convención normal de proporcionar solamente la secuencia en la dirección 5' a 3' a lo largo de la cadena de ADN no transcrita (es decir, la cadena que tiene una secuencia homóloga al ARNm). Una «molécula de ADN recombinante» es una molécula de ADN que experimentó manipulación biológica molecular. El ADN incluye, de modo no taxativo, ADNc, ADN genómico, ADN del plásmido, ADN sintético y ADN semisintético. Una «composición de ácido nucleico» de la descripción comprende uno o más ácidos nucleicos como se describe en la presente.

[0103] Tal como se utiliza en la presente, una «región de codificación» o «secuencia de codificación» es una parte del polinucleótido que consiste en codones que pueden traducirse en aminoácidos. Aunque un «codón de detención» (TAG, TGA o TAA) no se traduce típicamente en un aminoácido, puede considerarse parte de una

5 región de codificación, pero ninguna secuencia flanqueante, por ejemplo, promotores, sitios de unión de ribosomas, finalizadores transcripcionales, intrones y similares, no forman parte de una región de codificación. Los límites de una región de codificación se determinan típicamente mediante un codón de comienzo en el extremo 5', que codifica el extremo amino del polipéptido resultante, y un codón de

10 detención de traducción en el extremo 3', que codifica el extremo carboxilo del polipéptido resultante. Dos o más regiones de codificación pueden estar presentes en una construcción de un solo polinucleótido, por ejemplo, en un vector simple, o en construcciones de polinucleótido separadas, por ejemplo, en vectores separados (diferentes). Luego sigue que un vector simple puede contener solo una única región

15 de codificación o comprender dos o más regiones de codificación.

[0104] Determinadas proteínas secretadas por células mamíferas se asocian con un péptido señal secretor que se escinde de la proteína madura una vez iniciada la exportación de la cadena de proteínas en crecimiento a lo largo del retículo

20 endoplásmico duro. Los expertos en la técnica son conscientes de que los péptidos de señal se fusionan generalmente al extremo N del polipéptido, y se escinden del polipéptido completo o «de longitud completa» para producir una forma secretada o «madura» del polipéptido. En determinadas modalidades, un péptido de señal natural o un derivado funcional de esa secuencia retiene la capacidad de dirigir la

25 secreción del polipéptido que está asociado operativamente a esta. De forma alternativa, se puede utilizar un péptido de señal mamífero heterólogo, por ejemplo, un activador de plasminógeno de tejido humano (TPA, por sus siglas en inglés) o péptido de señal de la β -glucuronidasa de ratón, o un derivado funcional de estos.

30 **[0105]** El término «posterior» se refiere a una secuencia de nucleótidos que está ubicada 3' respecto a una secuencia de nucleótidos de referencia. En determinadas modalidades, las secuencias de nucleótidos posteriores se refieren a secuencias que siguen el punto de comienzo de transcripción. Por ejemplo, el codón de inicio de traducción de un gen está ubicado posterior al sitio de comienzo de la transcripción.

35

[0106] El término «anterior» se refiere a una secuencia de nucleótidos que está ubicada 5' respecto a una secuencia de nucleótidos de referencia. En determinadas

modalidades, las secuencias de nucleótidos anteriores se relacionan con secuencias que están ubicadas en el lado 5' de una región de codificación o punto de comienzo de la transcripción. Por ejemplo, la mayoría de los promotores están ubicados anteriores al sitio de comienzo de la transcripción.

5

[0107] Tal como se utiliza en la presente, la expresión «región de regulación de genes» o «región reguladora» se refiere a secuencias de nucleótidos ubicadas delante (secuencias de no codificación 5'), dentro, o posterior (secuencias de no codificación 3') a una región de codificación, y que influyen la transcripción, procesamiento de ARN, estabilidad o traducción de la región de codificación asociada. Las regiones de regulación pueden incluir promotores, secuencias líderes de traducción, intrones, secuencias de reconocimiento de poliadenilación, sitios de procesamiento de ARN, sitios de unión de efectores y estructuras bucle-tallo. Si se pretende una región de codificación para la expresión en una célula eucariota, una secuencia de finalización de transcripción y señal de poliadenilación estará ubicada normalmente 3' con respecto a la secuencia de codificación.

[0108] Un polinucleótido que codifica un producto génico, por ejemplo, un polipéptido, puede incluir un promotor y/u otros elementos de control de expresión (por ejemplo, transcripción o traducción) asociados de manera operativa con una o más regiones de codificación. En una asociación operativa de una región de codificación para un producto génico, por ejemplo, un polipéptido, se asocia con una o más regiones de regulación en forma tal de colocar la expresión del producto génico bajo la influencia o el control de la o las regiones de regulación. Por ejemplo, una región de codificación y un promotor están «asociados operativamente» en caso de que la inducción de la función de promotores ocasione la transcripción de ARNm que codifique el producto génico codificado por la región de codificación, y en caso de que la naturaleza de la unión entre el promotor y la región de codificación no interfiera con la capacidad del promotor para dirigir la expresión del producto génico o interferir con la capacidad de la plantilla de ADN que se transcribirá. Otros elementos de control de expresión, además de un promotor, por ejemplo, potenciadores, operadores, represores y señales de finalización de transcripción, también pueden asociarse operativamente con una región de codificación para dirigir la expresión del producto génico.

35

[0109] Las «secuencias de control transcripcional» hacen referencia a secuencias reguladoras de ADN, tales como promotores, potenciadores, elementos de

terminación y similares, que proporcionan la expresión de una región de codificación en una célula hospedadora. Los expertos en la técnica conocen una variedad de regiones de control de transcripción. Estos incluyen, de modo no taxativo, regiones de control de transcripción que funcionan en células vertebradas, tales como, de modo no taxativo, segmentos promotores y potenciadores de citomegalovirus (el promotor temprano inmediato, junto con el intrón A), virus del simio 40 (el promotor temprano) y los retrovirus (tal como el virus del sarcoma de Rous). Otras regiones de control de transcripción incluyen aquellas derivadas de los genes vertebrados tales como actina, proteína de choque térmico, hormona del crecimiento bovina y β -globina de conejo, así como otras secuencias capaces de controlar la expresión génica en células eucariotas. Otras regiones de control de transcripción adecuadas incluyen promotores y potenciadores específicos de tejido, así como promotores inducibles por linfocina (por ejemplo, promotores inducibles por interferones o interleucinas).

[0110] De forma similar, los expertos en la técnica conocen una variedad de elementos de control de traducción. Estos incluyen, de modo no taxativo, sitios de unión de ribosomas, codones de finalización e inicio de traducción, y elementos derivados de picornavirus (en particular, un sitio de entrada de ribosomas interno, o IRES, también conocido como una secuencia CITE).

[0111] El término «expresión», tal como se utiliza en la presente, se refiere a un proceso mediante el cual un polinucleótido produce un producto génico, por ejemplo, un ARN o un polipéptido. Incluye, de modo no taxativo, la transcripción del polinucleótido en un ARN mensajero (ARNm), ARN de transferencia (ARNt), ARN de horquilla corta (ARNhc), ARN interferente pequeño (ARNip) o cualquier otro producto de ARN, y la traducción de un ARNm en un polipéptido. La expresión produce un «producto génico». Tal como se usa en la presente, un producto del gen puede ser un ácido nucleico, por ejemplo, un ARN mensajero producido por la transcripción de un gen o un polipéptido que se traduce de una transcripción. Los productos génicos descritos en la presente incluyen además ácidos nucleicos con modificaciones transcripcionales posteriores, por ejemplo, poliadenilación o corte y empalme, o polipéptidos con modificaciones de traducción posteriores, por ejemplo, metilación, glicosilación, la adición de lípidos, asociación con otras subunidades de proteínas, o escisión proteolítica. El término «rendimiento» tal como se utiliza en la presente, hace referencia a la cantidad de un polipéptido producido por la expresión de un gen.

- [0112]** Un «vector» se refiere a cualquier vehículo para la clonación y/o transferencia de un ácido nucleico en una célula hospedadora. Un vector puede ser un replicón al cual puede unirse otro segmento de ácido nucleico con el fin de provocar la replicación del segmento unido. Un «replicón» se refiere a cualquier elemento genético (por ejemplo, plásmido, fago, cósmido, cromosoma, virus) que funciona como una unidad autónoma de replicación *in vivo*, es decir, capaz de replicación bajo su propio control. El término «vector» incluye tanto vehículos virales como no virales para introducir el ácido nucleico en una célula *in vitro*, *ex vivo* o *in vivo*. Se conoce y utiliza en la técnica una gran cantidad de vectores incluyendo, por ejemplo, plásmidos, virus eucariotas modificados o virus bacterianos modificados. La inserción de un polinucleótido en un vector adecuado puede lograrse mediante ligación de los fragmentos de polinucleótidos adecuados en un vector seleccionada que tiene extremos cohesivos complementarios.
- [0113]** Los vectores pueden modificarse genéticamente para codificar indicadores o marcadores seleccionables que proporcionan la selección o identificación de células que incorporaron el vector. La expresión de indicadores o marcadores seleccionables permite la identificación y/o selección de células hospedadoras que incorporan y expresan otras regiones de codificación 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, canamicina, higromicina, herbicida bialafos, sulfonamida y similares; y genes que se utilizan como marcadores fenotípicos, es decir, genes que regulan la antocianina, gen de transferasa isopentaniolo y similares. Los ejemplos de indicadores conocidos y utilizados en la técnica incluyen: luciferasa (Luc), proteína fluorescente verde (GFP), acetiltransferasa cloranfenicol (CAT), β -galactosidasa (LacZ), β -glucuronidasa (Gus) y similares. Los marcadores seleccionables también pueden considerarse como indicadores.
- [0114]** La expresión «marcador seleccionable» hace referencia a un factor de identificación, generalmente un antibiótico o un gen de resistencia química, que se puede seleccionar basándose en el efecto del gen marcador, es decir, la resistencia a un antibiótico, la resistencia a un herbicida, marcadores colorimétricos, enzimas, marcadores fluorescentes y similares, en donde el efecto se utiliza para rastrear 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, canamicina, higromicina, herbicida bialafos, sulfonamida y similares; y genes que se utilizan como marcadores fenotípicos, es decir, genes que regulan la antocianina, gen de transferasa isopentaniolo y similares.

5

[0115] La expresión «gen marcador» hace referencia a un ácido nucleico que codifica un factor de identificación que es capaz de ser identificado tras el efecto del gen marcador, 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 o transcripción de expresión génica. Los ejemplos de genes marcadores conocidos y utilizados en la técnica incluyen: luciferasa (Luc), proteína fluorescente verde (GFP), acetiltransferasa cloranfenicol (CAT), β -galactosidasa (LacZ), β -glucuronidasa (Gus) y similares. Los genes marcadores seleccionables se pueden considerar también genes reporteros.

15

[0116] «Promotor» y «secuencia promotora» se utilizan de forma intercambiable y hacen referencia a una secuencia de ADN capaz de controlar la expresión de una secuencia de codificación o ARN funcional. En general, una secuencia de codificación se ubica 3' respecto de una secuencia promotora. Los promotores pueden derivar en su totalidad de un gen nativo o pueden estar compuestos por diferentes elementos derivados de diferentes promotores que se encuentran en la naturaleza, o incluso comprender segmentos de ADN sintéticos. Los expertos en la técnica entienden que diferentes promotores pueden dirigir la expresión de un gen en diferentes tipos de células o tejidos, o en diferentes etapas de desarrollo, o en respuesta a diferentes condiciones ambientales o fisiológicas. Los promotores que pueden hacer que un gen se exprese en la mayoría de los tipos de células en la mayoría de los casos, se los conoce comúnmente como «promotores constitutivos». Los promotores que hacen que un gen se exprese en un tipo de célula específica se los conocen comúnmente como «promotores específicos de células» o «promotores específicos del tejido». Los promotores que hacen que un gen se exprese en una etapa específica de desarrollo o diferenciación celular se los conocen comúnmente como «promotores específicos para el desarrollo» o «promotores específicos para la diferenciación celular». Los promotores que son inducidos y hacen que un gen se exprese después de la exposición o tratamiento de la célula con un agente, molécula biológica, químico, ligando, luz o similar que induce al promotor se los conoce comúnmente como «promotores inducibles» o «promotores regulables». También se reconoce que, dado que en la mayoría de los casos no se definen completamente

35

los límites exactos de las secuencias de regulación, fragmentos de ADN de longitudes diferentes pueden tener actividad promotora idéntica.

5 [0117] La secuencia promotora está generalmente unida en su extremo 3' por el sitio de inicio de transcripción y se extiende de forma ascendente (en dirección 5') para incluir la cantidad mínima de bases o elementos necesarios para iniciar la transcripción a niveles detectables respecto al fondo. Dentro de la secuencia promotora se encontrará un sitio de inicio de la transcripción (definido de forma conveniente, por ejemplo, al mapearse con la nucleasa S1), así como también los
10 dominios de unión a proteínas (secuencias de consenso) responsables de la unión de la ARN polimerasa.

[0118] Las expresiones «endonucleasa de restricción» y «enzima de restricción» se utilizan de forma intercambiable y hacen referencia a una enzima que se une y corta
15 dentro de una secuencia de nucleótidos específica dentro del ADN de cadena doble.

[0119] 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 normalmente tiene forma de moléculas de ADN de cadena doble circular. Dichos
20 elementos pueden ser secuencias que se replican autónomamente, secuencias que integran genomas, secuencias de nucleótidos o fagos, lineales, circulares o superhelicoidales de un ADN o ARN de cadena doble o simple, derivadas de cualquier fuente, donde se unió o recombinó una cantidad de secuencias de nucleótidos en una única construcción que es capaz de introducir un fragmento de
25 promotor y secuencia de ADN para un producto génico seleccionado junto con la secuencia no traducida 3' adecuada en una célula.

[0120] Los vectores virales eucariotas que pueden utilizarse incluyen, de modo no taxativo, vectores del adenovirus, vectores del retrovirus, vectores del virus
30 adenoasociado y poxivirus, por ejemplo, vectores del vaccinia virus, vectores del baculovirus o vectores del herpesvirus. Los vectores no virales incluyen plásmidos, liposomas, lípidos cargados eléctricamente (citofectinas), complejos ADN-proteína y biopolímeros.

35 [0121] Un «vector de clonación» se refiere a un «replicón», que es una longitud unitaria de un ácido nucleico que se replica secuencialmente y que comprende un origen de replicación, tal como un plásmido, fago o cósmido, al que puede unirse

otro segmento de ácido nucleico de forma tal de provocar la replicación del segmento unido. Determinados vectores de clonación son capaces de replicación en un tipo celular, por ejemplo, bacterias y expresión en otros, por ejemplo, células eucariotas. Los vectores de clonación comprenden típicamente una o más
 5 secuencias que pueden utilizarse para la selección de células que comprenden el vector y/o uno o más sitios de clonación múltiples para la inserción de secuencias de ácido nucleico de interés.

[0122]El término «vector de expresión» se refiere a un vehículo diseñado para
 10 permitir la expresión de una secuencia de ácido nucleico insertado luego de la inserción en una célula hospedadora. La secuencia de ácido nucleico insertada se coloca en una asociación operativa con regiones de regulación tal como se describió anteriormente.

[0123]Los vectores se introducen en células hospedadoras mediante métodos muy conocidos en la técnica, por ejemplo, transfección, electroporación, microinyección, transducción, fusión celular, DEAE dextrano, precipitación de fosfato de calcio, lipofección (fusión lisosomal), uso de una pistola génica, o un transportador de vector de ADN.

[0124]Tal como se usan en la presente, «cultivo» y «cultivar» significan incubar células en condiciones *in vitro* que permiten la reproducción o división celular, o mantener las células en estado vivo. «Células cultivadas», tal como se usa en la presente, significa células que se propagan *in vitro*.

[0125]Tal como se utiliza en la presente, el término «polipéptido» pretende abarcar un «polipéptido» singular, así como «polipéptidos» plurales, y se refiere a una molécula compuesta de monómeros (aminoácidos) unidos de forma lineal mediante enlaces amida (también conocidos como enlaces peptídicos). El término
 30 «polipéptido» se refiere a cualquier cadena o cadenas de dos o más aminoácidos, y no se refiere a una longitud específica del producto. De este modo, 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, se incluyen dentro de la definición de «polipéptido», y el término
 35 «polipéptido» puede utilizarse en lugar de, o de forma indistinta con, cualquiera de estos términos. El término «polipéptido» también pretende referirse a los productos de modificaciones posteriores a la expresión del polipéptido incluyendo, de modo no

taxativo, glicosilación, acetilación, fosforilación, amidación, derivatización a través de grupos de bloqueo/protectores conocidos, escisión proteolítica, o modificación a través de aminoácidos que no tienen origen natural. Un polipéptido puede derivarse de una fuente biológica natural o tecnología recombinante producida, pero no se traduce necesariamente a partir de una secuencia de ácido nucleico designado. Puede generarse de cualquier forma, incluyendo mediante síntesis química.

[0126]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 se encuentran también dentro del alcance de la invención e incluyen norleucina, omitina, norvalina, homoserina y otros análogos de residuos de aminoácidos, tales como los descritos en Ellman *et ál.* Meth. Enzym. 202:301-336 (1991). Para generar dichos residuos de aminoácidos de origen no natural, se pueden utilizar los procedimientos de Noren *et ál.* Science 244:182 (1989) y Ellman *et ál.*, supra. En resumen, estos procedimientos implican la activación química de un ARNt supresor con un residuo aminoácido de origen no natural, seguido de transcripción y traducción del ARN *in vitro*. También se puede lograr la introducción del aminoácido no tradicional mediante químicas peptídicas conocidas en la técnica. Tal como se usa en la presente, la expresión «aminoácido polar» incluye aminoácidos que tienen carga neta cero, pero que tienen cargas parciales distintas a cero en diferentes partes de sus cadenas laterales (por ejemplo, M, F, W, S, Y, N, Q, C). Estos aminoácidos pueden participar en interacciones hidrofóbicas e interacciones electrostáticas. Tal como se usa en la presente, la expresión «aminoácido cargado» incluye aminoácidos que pueden tener carga neta distinta de cero en sus cadenas laterales (por ejemplo, R, K, H, E, D). Estos aminoácidos pueden participar en interacciones hidrofóbicas e interacciones electrostáticas.

[0127]En la presente descripción se incluyen también fragmentos o variantes de polipéptidos y cualquier combinación de estos. El término «fragmento» o «variante», cuando se refiere a moléculas de unión o dominios de unión al polipéptido de la presente invención, incluye cualquier polipéptido que retiene al menos algunas de las propiedades (por ejemplo, afinidad de unión a FcRn para una variante Fc o dominio de unión a FcRn, actividad de coagulación para una variante del FVIII, o

actividad de unión del FVIII para el fragmento del VWF) del polipéptido de referencia. Los fragmentos de polipéptidos incluyen fragmentos proteolíticos, así como fragmentos de eliminación, además de los fragmentos de anticuerpos específicos analizados en cualquier punto de la presente, pero no incluyen el polipéptido de longitud completa de origen natural (o polipéptido maduro). Las variantes de moléculas de unión o dominios de unión al polipéptido de la presente descripción incluyen fragmentos tal como se describió anteriormente, y también polipéptidos con secuencias alteradas de aminoácidos debido a inserciones, eliminaciones o sustituciones de aminoácidos. Las variantes pueden tener origen natural o no natural. Las variantes de origen no natural pueden producirse mediante el uso de técnicas de mutagénesis conocidas en la técnica. Los polipéptidos variantes pueden comprender adiciones, eliminaciones o sustituciones de aminoácidos conservadoras o no conservadoras.

[0128] Una «sustitución de aminoácidos conservadora» es una sustitución en la que el residuo de aminoácidos se reemplaza con un residuo de aminoácidos que tiene una cadena lateral similar. Las familias de residuos de aminoácidos que tienen cadenas laterales similares se han definido en la técnica, incluyendo cadenas laterales básicas (por ejemplo, lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo, ácido aspártico, ácido glutámico), cadenas laterales polares sin carga (por ejemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales ramificadas beta (por ejemplo, treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina). De este modo, si se reemplaza un aminoácido en un polipéptido con otro aminoácido de la misma familia de cadena lateral, la sustitución se considera conservadora. En otra modalidad, una hebra de aminoácidos puede reemplazarse de forma conservadora con una hebra estructuralmente similar que difiere en orden y/o composición de los miembros de la familia de cadena lateral.

[0129] Según se conoce en la técnica, el «porcentaje de identidad» es una relación entre dos o más secuencias de polipéptidos o dos o más secuencias de polinucleótidos, según se determina mediante la comparación de las secuencias. En la técnica, «identidad» también implica el grado de relación de secuencia entre las secuencias de polipéptidos o polinucleótidos, según sea el caso, determinada por la coincidencia entre cadenas de dichas secuencias. La «identidad» se puede calcular

fácilmente mediante métodos conocidos que incluyen, de modo no taxativo, aquellos 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*
5 *Data, Part 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 proporcionar la mayor coincidencia entre las
10 secuencias analizadas. Los métodos para determinar la identidad se codifican en los programas informáticos disponibles públicamente. Las alineaciones de secuencias y cálculos de identidad de porcentaje se pueden realizar usando el software de análisis de secuencias tal como el programa Megalign del paquete de programas de la bioinformática LASERGENE (DNASTAR Inc., Madison, WI), el paquete de
15 programas GCG (Wisconsin Package versión 9.0, Genetics Computer Group (GCG), Madison, WI), BLASTP, BLASTN, BLASTX (Altschul *et ál.*, *J. Mol. Biol.* 215:403 (1990)), y DNASTAR (DNASTAR, Inc. 1228 S. Park St. Madison, WI 53715 EUA). En el contexto de la presente descripción, se comprenderá que en los casos en que se usa software de análisis de secuencias para el análisis, los resultados del análisis se
20 basan en los «valores predeterminados» del programa al que se hace referencia. Según se utiliza en la presente, «valores por defecto» significará cualquier conjunto de valores o parámetros que se cargan originalmente con el software cuando se inicia por primera vez. A los efectos de determinar el porcentaje de identidad entre una secuencia de FVIII BDD optimizada de la descripción y una secuencia de
25 referencia, se usan solamente los nucleótidos en la secuencia de referencia que corresponden a los nucleótidos en la secuencia de FVIII BDD optimizada de la descripción para calcular el porcentaje de identidad. Por ejemplo, cuando se compara una secuencia de nucleótidos de FVIII de longitud completa que contiene el dominio B con una secuencia de nucleótidos de FVIII con dominio B eliminado (BDD)
30 optimizada de la descripción, la parte de la alineación que incluye el dominio A1, A2, A3, C1 y C2 se utilizará para calcular el porcentaje de identidad. Los nucleótidos en la parte de la secuencia de FVIII de longitud completa que codifica el dominio B (que dará lugar a un «espacio» grande en la alineación) no se contará como un mal apareamiento. Además, para determinar el porcentaje de identidad entre una
35 secuencia de FVIII BDD optimizada de la descripción o una parte designada de esta (por ejemplo, los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO:3), y una secuencia de referencia, el porcentaje de identidad se calculará al dividir la cantidad

de nucleótidos entre la cantidad total de los nucleótidos emparejados en la secuencia completa de la secuencia de BDD-FVIII optimizada, o una parte designada de esta, como se indica en la presente.

5 **[0130]** Tal como se utiliza en la presente, los «nucleótidos que corresponden a nucleótidos en la secuencia de FVIII con BDD optimizada de la descripción» se identifican por la alineación de la secuencia de FVIII con BDD optimizada de la descripción para maximizar la identidad con respecto a la secuencia de FVIII de referencia. La cantidad utilizada para identificar un aminoácido equivalente en una
10 secuencia de FVIII de referencia se basa en la cantidad utilizada para identificar el aminoácido correspondiente en la secuencia de FVIII con BDD optimizada.

[0131] Una proteína «quimérica» o «de fusión» comprende una primera secuencia de aminoácidos unida a la segunda secuencia de aminoácidos con la que no se une
15 naturalmente en su naturaleza. Las secuencias de aminoácidos que existen normalmente en proteínas individuales pueden juntarse en el polipéptido de fusión, o las secuencias de aminoácidos que existen normalmente en la misma proteína pueden estar ubicadas en una nueva disposición en el polipéptido de fusión, por ejemplo, fusión de un dominio del Factor VIII de la invención con un dominio Fc de
20 Ig. Una proteína de fusión se crea, por ejemplo, mediante síntesis química, o mediante creación y traducción de un polinucleótido donde las regiones peptídicas se codifican en la relación deseada. Una proteína quimérica puede comprender además una segunda secuencia de aminoácidos asociada con la primera secuencia de aminoácidos mediante un enlace no peptídico covalente o un enlace no
25 covalente.

[0132] Tal como se utiliza en la presente, el término «sitio de inserción» se refiere a una posición en un polipéptido del FVIII, o fragmento, variante o derivado de este, que se encuentra inmediatamente corriente arriba de la posición en la que se puede
30 insertar un resto heterólogo. Un «sitio de inserción» se indica como un número, donde el número es el número del aminoácido en el FVIII natural maduro (SEQ ID NO: 15; figura 11A) al que corresponde el sitio de inserción, que se encuentra inmediatamente en el extremo N con respecto a la posición de la inserción. Por ejemplo, la frase «a3 comprende un resto heterólogo en un sitio de inserción que
35 corresponde al aminoácido 1656 de la SEQ ID NO: 15» indica que el resto heterólogo está ubicado entre dos aminoácidos que corresponden al aminoácido 1656 y aminoácidos 1657 de la SEQ ID NO: 15.

[0133] La frase «inmediatamente posterior a un aminoácido», tal como se utiliza en la presente, se refiere a la posición justo al lado del grupo de extremo carboxilo del aminoácido. De forma similar, la frase «inmediatamente anterior a un aminoácido», tal como se utiliza en la presente, se refiere a la posición justo al lado del grupo de extremo amin del aminoácido.

[0134] Los términos «insertado», «está insertado», «insertado en» o términos relacionados gramaticalmente, tal como se utiliza en la presente, se refieren a la posición de un resto heterólogo en un polipéptido de FVIII recombinante con respecto a la posición análoga en el FVIII humano maduro natural. Tal como se utiliza en la presente, los términos se refieren a las características del polipéptido del FVIII recombinante con respecto al FVIII humano maduro natural y no indican, implican o infieren ningún método o proceso mediante el cual se realizó el polipéptido del FVIII recombinante.

[0135] Como se usa en la presente, el término «semivida» se refiere a una semivida biológica de un polipéptido particular *in vivo*. La semivida puede estar representada por el tiempo requerido para la mitad de la cantidad administrada a un sujeto que se eliminará de la circulación y/u otros tejidos en el animal. Cuando una curva de depuración de un polipéptido dado se construye como una función de tiempo, la curva es normalmente bifásica con una fase α rápida y fase β más duradera. La fase α representa típicamente un equilibrio del polipéptido Fc administrado entre el espacio intra y extravascular y está, en parte, determinada por el tamaño del polipéptido. La fase β representa típicamente el catabolismo del polipéptido en el espacio intravascular. En algunas modalidades, el FVIII y las proteínas quiméricas que comprenden FVIII son monofásicas y, de este modo, no tienen una fase alfa, sino solamente la simple fase beta. Por lo tanto, en determinadas modalidades, el término semivida, tal como se utiliza en la presente, se refiere a la semivida del polipéptido en la fase β .

[0136] El término «unido», tal como se utiliza en la presente, se refiere a una primera secuencia de aminoácidos o secuencia de nucleótidos unida de forma covalente o no covalente a una segunda secuencia de aminoácidos o secuencia de nucleótidos, respectivamente. La primera secuencia de aminoácidos o nucleótidos puede estar unida o yuxtapuesta directamente a la segunda secuencia de aminoácidos o nucleótidos o, de forma alternativa, una secuencia de intervención puede unir covalentemente la primera secuencia a la segunda secuencia. El término «unido»

- significa no solamente una fusión de una primera secuencia de aminoácidos a una segunda secuencia de aminoácidos en el extremo C o extremo N, pero también incluye la inserción de la primera secuencia de aminoácidos completa (o la segunda secuencia de aminoácidos) en cualesquiera dos aminoácidos en la segunda
- 5 secuencia de aminoácidos (o la primera secuencia de aminoácidos, respectivamente). En una modalidad, la primera secuencia de aminoácidos puede unirse a una segunda secuencia de aminoácidos mediante un enlace peptídico o un enlazador. La primera secuencia de nucleótidos puede unirse a una segunda secuencia de nucleótidos mediante un enlace fosfodiéster o un enlazador. El
- 10 enlazador puede ser un péptido o un polipéptido (para cadenas polipeptídicas) o un nucleótido o una cadena de nucleótidos (para cadenas de nucleótidos) o cualquier resto químico (para ambas cadenas polipeptídicas o de polinucleótidos). El término «unido» también se indica con un guión (-).
- 15 **[0137]** Tal como se utiliza en la presente, el término «asociado con» se refiere a un enlace covalente o no covalente formado entre una primera cadena de aminoácidos y una segunda cadena de aminoácidos. En una modalidad, el término «asociado con» significa un enlace no peptídico covalente o un enlace no covalente. Esta asociación puede estar indicada por dos puntos, es decir, (:). En otra modalidad,
- 20 significa un enlace covalente excepto un enlace peptídico. Por ejemplo, la cisteína del aminoácido comprende un grupo tiol que puede formar un puente o enlace disulfuro con un grupo tiol en un segundo residuo de cisteína. En la mayoría de las moléculas de IgG de origen natural, las regiones CL y CH1 están asociadas mediante un enlace disulfuro y las dos cadenas pesadas están asociadas por dos
- 25 enlaces disulfuro en posiciones que corresponden a 239 y 242 utilizando el sistema de numeración de Kabat (posición 226 o 229, sistema de numeración EU). Los ejemplos de enlaces covalentes incluyen, de modo no taxativo, un enlace peptídico, un enlace metálico, un enlace hidrógeno, un enlace disulfuro, un enlace sigma, un enlace pi, un enlace delta, un enlace glicosídico, un enlace agnóstico, un enlace
- 30 curvado, un enlace dipolar, una estructura Pi, un enlace doble, un enlace triple, un enlace cuádruple, un enlace quíntuple, un enlace séxtuple, conjugación, hiperconjugación, aromaticidad, hapticidad o antiunión. Los ejemplos no taxativos de enlaces no covalentes incluyen un enlace iónico (por ejemplo, enlace catión-pi o enlace salino), un enlace metálico, un enlace hidrógeno (por ejemplo, enlace
- 35 dihidrógeno, complejo dihidrógeno, enlace hidrógeno de barrera inferior o enlace hidrógeno simétrico), fuerza van der Waals, fuerza de dispersión de London, un

enlace mecánico, un enlace halógeno, aurofilicidad, intercalación, apilación, fuerza entrópica o polaridad química.

5 **[0138]** El término «híbrido de monómero-dímero» utilizado en la presente se refiere a una proteína quimérica que comprende una primera cadena polipeptídica y una segunda cadena polipeptídica, que están asociadas entre sí mediante un enlace disulfuro, donde la primera cadena comprende un factor de coagulación, por ejemplo, Factor VIII, y una primera región Fc y la segunda cadena comprende, consiste esencialmente en o consiste en una segunda región Fc sin el factor de
10 coagulación. De este modo, la construcción de híbrido monómero-dímero es un híbrido que comprende un aspecto de monómero que tiene únicamente un factor de coagulación y un aspecto de dímero que tiene dos regiones Fc.

15 **[0139]** La hemostasia, tal como se usa en la presente, significa la detención o ralentización del sangrado o hemorragia; o la detención o ralentización del flujo sanguíneo a través de un vaso sanguíneo o una parte del cuerpo.

20 **[0140]** El trastorno de la hemostasia, tal como se usa en la presente, significa una enfermedad heredada genéticamente o adquirida, caracterizada por una tendencia a las hemorragias, ya sea de forma espontánea o como resultado de algún traumatismo, debido a la capacidad reducida o incapacidad de formar coágulos de fibrina. Los ejemplos de dichos trastornos incluyen las hemofilias. Las tres formas principales son hemofilia A (deficiencia del factor VIII), hemofilia B (deficiencia del factor IX o «enfermedad de Navidad») y hemofilia C (deficiencia del factor XI, tendencia de sangrado leve). Otros trastornos hemostáticos incluyen, por ejemplo, enfermedad de Von Willebrand, deficiencia del factor XI (deficiencia de PTA), deficiencia del factor XII, deficiencias o anormalidades estructurales en fibrinógeno, protrombina, factor V, factor VII, factor X o factor XIII, síndrome de Bernard-Soulier, que es un defecto o deficiencia en GPIb. GPIb, el receptor del VWF, puede ser
25 defectuoso y conducir a carencia de formación de coágulos primaria (hemostasia primaria) y tendencia de sangrado aumentado, y trombostenia de Glanzman y Naegeli (trombastenia de Glanzmann). Tras insuficiencia hepática (formas agudas y crónicas), no hay producción suficiente de factores de coagulación desde el hígado; esto puede aumentar el riesgo de sangrado.

35

[0141] Las moléculas de ácido nucleico aisladas, los polipéptidos aislados o vectores que comprenden la molécula de ácido nucleico aislada de la descripción se pueden

utilizar de forma profiláctica. Tal como se utiliza en la presente, el término «tratamiento profiláctico» se refiere a la administración de una molécula antes de un episodio hemorrágico. En una modalidad, el sujeto que necesita un agente hemostático general está sometido a, o está a punto de someterse a, cirugía. Un polinucleótido, polipéptido o vector de la descripción puede administrarse antes o después de la cirugía como profiláctico. El polinucleótido, polipéptido o vector de la descripción puede administrarse durante o después de la cirugía para controlar un episodio hemorrágico agudo. La cirugía puede incluir, de modo no taxativo, trasplante de hígado, resección hepática, procedimientos dentales o trasplante de células madre.

[0142] Las moléculas de ácido nucleico aisladas, los polipéptidos aislados o vectores de la descripción se utilizan también para un tratamiento a demanda. La expresión «tratamiento a demanda» se refiere a la administración de una molécula de ácido nucleico aislada, polipéptido aislado o vector en respuesta a los síntomas de un episodio hemorrágico o antes de una actividad que pueda ocasionar sangrado. En un aspecto, el tratamiento según demanda puede proporcionarse a un sujeto cuando comienza el sangrado, tal como después de una lesión, o cuando se espera el sangrado, tal como antes de la cirugía. En otro aspecto, el tratamiento según demanda puede proporcionarse antes de las actividades que aumentan el riesgo de sangrado, tal como los deportes de contacto.

[0143] Tal como se utiliza en la presente, el término «sangrado agudo» se refiere a un episodio de sangrado independientemente de la causa subyacente. Por ejemplo, un sujeto puede tener un trauma, uremia, un trastorno de sangrado hereditario (por ejemplo, deficiencia de factor VII), un trastorno de plaquetas o resistencia debido al desarrollo de anticuerpos de los factores de coagulación.

[0144] Tratar, tratamiento, que trata, tal como se utiliza en la presente, se refiere a, por ejemplo, la reducción de la gravedad de una enfermedad o afección; reducción de la duración de un transcurso de enfermedad; mejora de uno o más síntomas asociados con una enfermedad o afección; suministro de efectos benéficos a un sujeto que padece una enfermedad o afección, sin curar necesariamente la enfermedad o afección, o la profilaxis de uno o más síntomas asociados con una enfermedad o afección. En una modalidad, el término «que trata» o «tratamiento» significa mantener un nivel de depresión del FVIII de al menos aproximadamente 1 IU/dL, 2 IU/dL, 3 IU/dL, 4 IU/dL, 5 IU/dL, 6 IU/dL, 7 IU/dL, 8 IU/dL, 9 IU/dL, 10 IU/dL,

11 IU/dL, 12 IU/dL, 13 IU/dL, 14 IU/dL, 15 IU/dL, 16 IU/dL, 17 IU/dL, 18 IU/dL, 19 IU/dL, o 20 IU/dL en un sujeto mediante la administración de una molécula de ácido nucleico aislada, polipéptido o vector aislado de la descripción. En otra modalidad, que trata o tratamiento significa mantener un nivel de depresión del FVIII entre

5 aproximadamente 1 y aproximadamente 20 IU/dL, aproximadamente 2 y aproximadamente 20 IU/dL, aproximadamente 3 y aproximadamente 20 IU/dL, aproximadamente 4 y aproximadamente 20 IU/dL, aproximadamente 5 y aproximadamente 20 IU/dL, aproximadamente 6 y aproximadamente 20 IU/dL, aproximadamente 7 y aproximadamente 20 IU/dL, aproximadamente 8 y

10 aproximadamente 20 IU/dL, aproximadamente 9 y aproximadamente 20 IU/dL, o aproximadamente 10 y aproximadamente 20 IU/dL. Tratamiento o que trata una enfermedad o afección también puede incluir mantener una actividad del FVIII en un sujeto en un nivel que se compara a al menos alrededor de 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 11 %, 12 %, 13 %, 14 %, 15 %, 16 %, 17 %, 18 %, 19 % o

15 20 % de la actividad del FVIII en un sujeto no hemofílico. El nivel de depresión mínimo requerido para el tratamiento puede medirse mediante uno o más métodos conocidos y puede ajustarse (aumentarse o disminuirse) para cada persona.

[0145] «Administrar», tal como se utiliza en la presente, significa proporcionar una

20 molécula de ácido nucleico que codifica el factor VIII farmacéuticamente aceptable, polipéptido del factor VIII o vector que comprende una molécula de ácido nucleico que codifica el factor VIII de la descripción a un sujeto mediante una vía farmacéuticamente aceptable. Las vías de administración pueden ser intravenosas, por ejemplo, inyección intravenosa e infusión intravenosa. Las vías de administración

25 adicionales incluyen, por ejemplo, administración subcutánea, intramuscular, oral, nasal y pulmonar. Las moléculas de ácido nucleico, los polipéptidos y vectores se pueden administrar como parte de una composición farmacéutica que comprende al menos un excipiente.

30 [0146] Tal como se utiliza en la presente, la expresión «sujeto que lo necesita» incluye sujetos, tales como sujetos mamíferos que se beneficiarían de la administración de una molécula de ácido nucleico, un polipéptido o vector de la descripción, por ejemplo, para mejorar la hemostasia. En una modalidad, los sujetos incluyen, de modo no taxativo, individuos con hemofilia. En otra modalidad, los

35 sujetos incluyen, de modo no taxativo, los individuos que han desarrollado un inhibidor de FVIII y, por lo tanto, necesitan de una terapia de bypass. El sujeto puede ser un adulto o un menor (por ejemplo, menor de 12 años).

- [0147]** Tal como se usa en la presente, el término «factor de coagulación» se refiere a moléculas, o análogos de estas, de origen natural o producida de forma recombinante que previenen o disminuyen la duración de un episodio de sangrado en un sujeto. En otras palabras, significa que las moléculas que tienen una actividad procoagulante, es decir, son responsables de la conversión del fibrinógeno en una red insoluble de fibrina causando que la sangre se coagule o forme coágulos. Un «factor de coagulación activable» es un factor de coagulación en forma inactiva (por ejemplo, en su forma zimógena) que es capaz de convertirse en una forma activa.
- [0148]** La actividad coagulante, tal como se usa en la presente, significa la habilidad de participar en la cascada de reacciones bioquímicas que culminan en la formación de coágulos de fibrina y/o reducen la gravedad, duración o frecuencia de un episodio de hemorragia o sangrado.
- [0149]** Tal como se utiliza en la presente, los términos «exógeno» o «heterólogo» se refieren a dichas moléculas que no se encuentran normalmente en un contexto dado, por ejemplo, en una célula o en un polipéptido. Por ejemplo, puede introducirse en una célula una molécula exógena o heteróloga y estas únicamente se encuentran presente luego de manipular una célula, por ejemplo, mediante transfección u otras formas de ingeniería genética o puede presentarse una secuencia de aminoácidos heteróloga en una proteína en la cual no se encuentra de forma natural.
- [0150]** Tal como se utiliza en la presente, la expresión «secuencia de nucleótidos heteróloga» hace referencia a una secuencia de nucleótidos que no se produce de forma natural con una secuencia de polinucleótidos dada. En una modalidad, la secuencia de nucleótidos heteróloga codifica un polipéptido capaz de extender la semivida del FVIII. En otra modalidad, la secuencia de nucleótidos heteróloga codifica un polipéptido que aumenta el radio hidrodinámico del FVIII. En otras modalidades, la secuencia de nucleótidos heteróloga codifica un polipéptido que mejora una o más propiedades farmacocinéticas del FVIII sin afectar significativamente su actividad o función biológica (por ejemplo, su actividad procoagulante). En algunas modalidades, el FVIII está unido o conectado al polipéptido codificado por la secuencia de nucleótidos heteróloga por medio de un enlazador. Los ejemplos no taxativos de restos de polipéptido codificados por las secuencias de nucleótidos heterólogas incluyen una región constante de inmunoglobulina o una parte de esta, albúmina o un fragmento de esta, un restos de

unión a la albúmina, una transferrina, los polipéptidos PAS de la solicitud de patente estadounidense n.o 20100292130, una secuencia HAP, transferrina o un fragmento de esta, el péptido de extremo C (CTP) de la subunidad β de la gonadotropina coriónica humana, molécula pequeña de unión a la albúmina, una secuencia de

5 XTEN, restos de unión a FcRn (por ejemplo, regiones Fc completas o partes de estas que se unen a FcRn), regiones Fc de cadena simple (regiones ScFc, por ejemplo, como se describe en US 2008/0260738, WO 2008/012543, o WO 2008/1439545), enlazadores de poliglicina, enlazadores de poliserina, péptidos y polipéptidos cortos de 6-40 aminoácidos de dos tipos de aminoácidos que se

10 selecciona de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) con grados distintos de estructura secundaria de menos de 50 % a más de 50 %, entre otros, o dos o más combinaciones de estos. En algunas modalidades, el polipéptido codificado por la secuencia de nucleótidos heteróloga está unido a un resto que no es un polipéptido. Los ejemplos no taxativos de restos que no son de

15 polipéptido incluyen polietilenglicol (PEG), moléculas pequeñas a la albúmina, ácido polisiálico, almidón de hidroxietilo (HES), un derivado de estos o cualesquiera combinaciones de estos.

[0151] Tal como se utiliza en la presente, la expresión «región Fc» se define como la

20 parte de un polipéptido que corresponde a la región Fc del Ig natural, es decir, según se formó mediante la asociación dimérica de los dominios Fc respectivos de sus dos cadenas pesadas. Una región Fc natural forma un homodímero con otra región Fc. Por el contrario, la expresión «región Fc fusionada genéticamente» o «región Fc de cadena simple» (región scFc), tal como se utiliza en la presente, se refiere a una

25 región Fc dimérica sintética que comprende dominios Fc unidos genéticamente dentro de una única cadena polipeptídica (es decir, codificada en una única secuencia genética contigua).

[0152] En una modalidad, la «región Fc» se refiere a la parte de una única cadena

30 pesada de Ig que comienza en la región bisagra justo anterior al sitio de escisión de papaína (es decir, residuo 216 en IgG, tomando el primer residuo de la región constante de cadena pesada como 114) y que finaliza en el extremo C del anticuerpo. Por consiguiente, un dominio Fc completo comprende al menos un dominio bisagra, un dominio CH2 y un dominio CH3.

35

[0153] La región Fc de una región constante de Ig, dependiendo del isotipo de Ig puede incluir los dominios CH2, CH3 y CH4, así como la región bisagra. Las

proteínas quiméricas que comprenden una región Fc de una Ig conceden varias propiedades deseables en una proteína quimérica incluyendo el aumento de la estabilidad, el aumento de la semivida sérica (véase Capon *et ál.*, 1989, Nature 337:525), así como unión a receptores Fc tal como el receptor Fc neonatal (FcRn) (patente estadounidense n.º 6,086,875, 6,485,726, 6,030,613; WO 03/077834; US2003-0235536A1), que se incorporan en la presente mediante esta referencia en su totalidad.

[0154] Una «secuencia de nucleótidos de referencia», tal como se utiliza en la presente en comparación con una secuencia de nucleótidos de la descripción, es una secuencia de polinucleótidos esencialmente idéntica a la secuencia de nucleótidos de la descripción excepto que las partes correspondientes a la secuencia de FVIII no están optimizadas. Por ejemplo, la secuencia de nucleótidos de referencia para una molécula de ácido nucleico consiste en el FVIII con BDD optimizado con codones de la SEQ ID NO: 1 y una secuencia de nucleótidos heteróloga que codifica una región Fc de cadena simple unida a la SEQ ID NO: 1 en su extremo 3' es una molécula de ácido nucleico que consiste en el FVIII con BDD original (o «progenitor») de la SEQ ID NO: 16 (figura 1I) y una secuencia de nucleótidos heteróloga que codifica una región Fc de cadena simple unida a la SEQ ID NO: 16 en su extremo 3'.

[0155] Un «índice de adaptación de codones», tal como se utiliza en la presente, hace referencia a una medición de la tendencia de uso del codón. Un índice de adaptación de codones (CAI) mide la desviación de una secuencia génica de codificación de proteínas dada con respecto a un conjunto de genes de referencia (Sharp PM y Li WH, *Nucleic Acids Res.* 15(3):1281–95 (1987)). El CAI se calcula determinando el promedio geométrico del peso asociado con cada codón con respecto a la longitud de la secuencia génica (medida en codones):

$$CAI = \exp(1/L \sum_{l=1}^L \ln(w_i(l))), \tag{I}$$

[0156] Para cada aminoácido, el peso de cada uno de sus codones, en el CAI, se computa como la relación entre la frecuencia observada del codón (f_i) y la frecuencia del codón sinónimo (f_j) para ese aminoácido:

$$w_i = \frac{f_i}{\max(f_j)} \quad i, j \in [\text{codones sinónimos para el aminoácido}] \tag{II}$$

[0158] Tal como se utiliza en la presente, el término «optimizado/a» con respecto a las secuencias de nucleótidos, hace referencia a una secuencia de polinucleótidos que codifica un polinucleótido, donde la secuencia de polinucleótidos ha sido mutada para mejorar una propiedad de esa secuencia de polinucleótidos. En algunas modalidades, la optimización se realiza 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 las proteínas de regulación tales como los factores de transcripción generales, aumentar o disminuir el empalme o aumentar el rendimiento del polipéptido producido mediante la secuencia de polinucleótidos. Los ejemplos de cambios que se pueden realizar a una secuencia de polinucleótidos para optimizarla incluyen la optimización con codones, la optimización del contenido de G/C, la remoción de las secuencias de repetición, remoción de elementos ricos en AT, la remoción de sitios de empalme crípticos, la remoción de elementos que actúan en cis que reprimen la transcripción o traducción, agregar o quitar las secuencias poli-T o poli-A, agregar secuencias alrededor del sitio de inicio de la transcripción que mejoran la transcripción, tal como las secuencias de consenso de Kozak, la remoción de secuencias que podrían formar estructuras de bucle-tallo, la remoción de secuencias desestabilizantes y dos o más combinaciones de estos.

Proteína de FVIII que codifica secuencias de polinucleótidos

[0159] En algunas modalidades, la presente descripción se refiere a moléculas de ácido nucleico optimizadas que codifican un polipéptido con actividad de FVIII. En algunas modalidades, el polinucleótido codifica un polipéptido de FVIII de longitud completa. En otras modalidades, la molécula de ácido nucleico codifica un polipéptido de FVIII con dominio B eliminado (BDD), donde se elimina todo el dominio B de FVIII o una parte de este. En una modalidad particular, la molécula de ácido nucleico codifica un polipéptido que comprende una secuencia de aminoácidos que tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a la SEQ ID NO: 17 (figura 1J) o un fragmento de este. En una modalidad, la molécula de ácido nucleico codifica un polipéptido que tiene la secuencia de aminoácidos de la SEQ ID NO: 17 o un fragmento de este.

[0160] En algunas modalidades, la molécula de ácido nucleico de la descripción codifica un polipéptido de FVIII que comprende un péptido señal o un fragmento de este. En otras modalidades, la molécula de ácido nucleico codifica un polipéptido de FVIII que carece de un péptido señal. En algunas modalidades, el péptido señal
5 comprende los aminoácidos 1-19 de la SEQ ID NO: 17.

[0161] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de
10 extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3 o (ii) los
15 nucleótidos 58-1791 de la SEQ ID NO: 4; y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII. En una modalidad particular, la primera secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-1791 de la SEQ ID NO:
20 3. En otra modalidad, la primera secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-1791 de la SEQ ID NO: 4. En otras modalidades, la primera secuencia de nucleótidos comprende los nucleótidos 58-1791 de la SEQ ID NO: 3 o los nucleótidos 58-1791
25 de la SEQ ID NO: 4.

[0162] En otras modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un
30 polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1-1791 de la SEQ ID NO: 3 o (ii) los nucleótidos 1-1791
35 de la SEQ ID NO: 4; y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII. En una modalidad, la primera secuencia de nucleótidos comprende los nucleótidos 1-1791 de la SEQ ID NO: 3 o los

nucleótidos 1-1791 de la SEQ ID NO: 4. En otra modalidad, la segunda secuencia de nucleótidos tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1792-4374 de la SEQ ID NO: 3 o 1792-4374 de la SEQ ID NO: 4. En una modalidad particular, la segunda secuencia de nucleótidos comprende los nucleótidos 1792-4374 de la SEQ ID NO: 3 o 1792-4374 de la SEQ ID NO: 4. En incluso otra modalidad, la segunda secuencia de nucleótidos tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 3 o 1792-2277 y 2320-4374 de la SEQ ID NO: 4 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 3 o 1792-4374 de la SEQ ID NO: 4 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En una modalidad particular, la segunda secuencia de nucleótidos comprende los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 3 o 1792-2277 y 2320-4374 de la SEQ ID NO: 4 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 3 o 1792-4374 de la SEQ ID NO: 4 sin los nucleótidos que codifican el dominio B o fragmento de dominio B).

[0163] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-4374 de la SEQ ID NO: 5 o (ii) 1792-4374 de la SEQ ID NO: 6; y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII. En determinadas modalidades, la segunda secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1792-4374 de la SEQ ID NO: 5. En otras modalidades, la segunda secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1792-4374 de la SEQ ID NO: 6. En una modalidad particular, la segunda secuencia de ácido nucleico comprende los nucleótidos 1792-4374 de la SEQ ID NO: 5 o 1792-4374 de la SEQ ID NO: 6. En algunas modalidades, la primera secuencia de ácido nucleico

unida a la segunda secuencia de ácido nucleico enumerada anteriormente tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-1791 de la SEQ ID NO: 5 o los
 5 nucleótidos 58-1791 de la SEQ ID NO: 6. En otras modalidades, la primera secuencia de ácido nucleico unida a la segunda secuencia de ácido nucleico enumerada anteriormente tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1-1791 de la
 10 SEQ ID NO: 5 o los nucleótidos 1-1791 de la SEQ ID NO: 6.

[0164] En otras modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un
 15 polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la segunda secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 (es decir,
 20 los nucleótidos 1792-4374 de la SEQ ID NO: 5 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o (ii) 1792-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad de polipéptidos de FVIII. En
 25 determinadas modalidades, la segunda secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 5 sin los nucleótidos que codifican el dominio B o fragmento de dominio
 30 B). En otras modalidades, la segunda secuencia de ácido nucleico tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En
 35 una modalidad particular, la segunda secuencia de ácido nucleico comprende los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 o 1792-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 5 o 1792-

4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En algunas modalidades, la primera secuencia de ácido nucleico unida a la segunda secuencia de ácido nucleico enumerada anteriormente tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-1791 de la SEQ ID NO: 5 o los nucleótidos 58-1791 de la SEQ ID NO: 6. En otras modalidades, la primera secuencia de ácido nucleico unida a la segunda secuencia de ácido nucleico enumerada anteriormente tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1-1791 de la SEQ ID NO: 5 o los nucleótidos 1-1791 de la SEQ ID NO: 6.

[0165] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 1, (ii) los nucleótidos 58-1791 de la SEQ ID NO: 2, (iii) los nucleótidos 58-1791 de la SEQ ID NO: 70, o (iv) los nucleótidos 58-1791 de la SEQ ID NO: 71; y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII. En otras modalidades, la primera secuencia de nucleótidos comprende los nucleótidos 58-1791 de la SEQ ID NO: 1, los nucleótidos 58-1791 de la SEQ ID NO: 2, (iii) los nucleótidos 58-1791 de la SEQ ID NO: 70, o (iv) los nucleótidos 58-1791 de la SEQ ID NO: 71.

[0166] En otras modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1-1791 de la SEQ ID NO: 1, (ii) los nucleótidos 1-1791 de la SEQ ID NO:

2, (iii) los nucleótidos 1-1791 de la SEQ ID NO: 70, o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 71; y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII. En una modalidad, la primera secuencia de nucleótidos comprende los nucleótidos 1-1791 de la SEQ ID NO: 1, los nucleótidos 1-1791 de la SEQ ID NO: 2, (iii) los nucleótidos 1-1791 de la SEQ ID NO: 70, o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 71. En otra modalidad, la segunda secuencia de nucleótidos unida a la primera secuencia de nucleótidos tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1792-4374 de la SEQ ID NO: 1, 1792-4374 de la SEQ ID NO: 2, (iii) los nucleótidos 1792-4374 de la SEQ ID NO: 70, o (iv) los nucleótidos 1792-4374 de la SEQ ID NO: 71. En una modalidad particular, la segunda secuencia de nucleótidos unida a la primera secuencia de nucleótidos comprende (i) los nucleótidos 1792-4374 de la SEQ ID NO: 1, (ii) los nucleótidos 1792-4374 de la SEQ ID NO: 2, (iii) los nucleótidos 1792-4374 de la SEQ ID NO: 70, o (iv) los nucleótidos 1792-4374 de la SEQ ID NO: 71. En otra modalidad, la segunda secuencia de nucleótidos unida a la primera secuencia de nucleótidos tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 1, (ii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 2, (iii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 70, o (iv) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 71. En una modalidad, la segunda secuencia de nucleótidos comprende (i) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 1, (ii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 2, (iii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 70, o (iv) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 71.

[0167] En otra modalidad, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-4374 de la SEQ ID NO: 1, (ii) los nucleótidos 1792-4374 de la SEQ ID NO: 2, (iii) los nucleótidos 1792-4374 de la SEQ ID NO: 70, o (iv) los nucleótidos

1792-4374 de la SEQ ID NO: 71; y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII. En una modalidad particular, la segunda secuencia de ácido nucleico comprende (i) los nucleótidos 1792-4374 de la SEQ ID NO: 1, (ii) los nucleótidos 1792-4374 de la SEQ ID NO: 2, (iii) los nucleótidos 1792-4374 de la SEQ ID NO: 70, o (iv) los nucleótidos 1792-4374 de la SEQ ID NO: 71. En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la segunda secuencia de ácido nucleico tiene al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 1, (ii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 2, (iii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 70, o (iv) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 71 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 1, los nucleótidos 1792-4374 de la SEQ ID NO: 2, los nucleótidos 1792-4374 de la SEQ ID NO: 70 o los nucleótidos 1792-4374 de la SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la parte de extremo N y la parte de extremo C juntas tienen una actividad de polipéptidos de FVIII. En una modalidad, la segunda secuencia de nucleótidos comprende (i) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 1, (ii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 2, (iii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 70, o (iv) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 71 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 1, los nucleótidos 1792-4374 de la SEQ ID NO: 2, los nucleótidos 1792-4374 de la SEQ ID NO: 70 o los nucleótidos 1792-4374 de la SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B).

[0168] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58 a 4374 de la SEQ ID NO: 1. En otras modalidades, la secuencia de nucleótidos comprende una secuencia de nucleótidos que tiene al menos 90 %, al

menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 1 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En otras modalidades, la secuencia de ácido nucleico tiene al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la SEQ ID NO: 1. En otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 1 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 58 a 4374 de la SEQ ID NO: 1. En incluso otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 1-2277 y 2320-4374 de la SEQ ID NO: 1 (es decir, los nucleótidos 1-4374 de la SEQ ID NO: 1 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 1 a 4374 de la SEQ ID NO: 1.

[0169] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58 a 4374 de la SEQ ID NO: 2. En otras modalidades, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 2. En otras modalidades, la secuencia de ácido nucleico tiene al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la SEQ ID NO: 2. En otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 2 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 2 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 58 a 4374 de la SEQ ID NO: 2. En incluso otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 1-2277 y 2320-4374 de la SEQ ID NO: 2 (es decir, los nucleótidos 1-4374 de la SEQ ID NO: 2 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 1 a 4374 de la SEQ ID NO: 2.

[0170] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58 a 4374 de la SEQ ID NO: 70. En otras modalidades, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 70 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 70 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En otras modalidades, la secuencia de ácido nucleico tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la SEQ ID NO: 70. En otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 70 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 70 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 58 a 4374 de la SEQ ID NO: 70. En incluso otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 1-2277 y 2320-4374 de la SEQ ID NO: 70 (es decir, los nucleótidos 1-4374 de la SEQ ID NO: 70 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 1 a 4374 de la SEQ ID NO: 70.

[0171] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58 a 4374 de la SEQ ID NO: 71. En otras modalidades, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con

respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En otras modalidades, la secuencia de ácido nucleico tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la SEQ ID NO: 71. En otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 58 a 4374 de la SEQ ID NO: 71. En incluso otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 1-2277 y 2320-4374 de la SEQ ID NO: 71 (es decir, los nucleótidos 1-4374 de la SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 1 a 4374 de la SEQ ID NO: 71.

[0172] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58 a 4374 de la SEQ ID NO: 3. En otras modalidades, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 3 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 3 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En determinadas modalidades, la secuencia de ácido nucleico tiene al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la SEQ ID NO: 3. En algunas modalidades, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 3 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 3 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 58 a 4374 de la SEQ ID NO: 3. En incluso otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 3 (es decir, los nucleótidos 1-4374 de la SEQ ID NO: 3 sin los

nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 1 a 4374 de la SEQ ID NO: 3.

[0173] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58 a 4374 de la SEQ ID NO: 4. En otras modalidades, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 4 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 4 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En otras modalidades, la secuencia de ácido nucleico tiene al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la SEQ ID NO: 4. En otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 4 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 4 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 58 a 4374 de la SEQ ID NO: 4. En incluso otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 1-2277 y 2320-4374 de la SEQ ID NO: 4 (es decir, los nucleótidos 1-4374 de la SEQ ID NO: 4 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 1 a 4374 de la SEQ ID NO: 4.

[0174] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58 a 4374 de la SEQ ID NO: 5. En otras modalidades, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al

menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 5 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 5 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En determinadas modalidades, la secuencia de ácido nucleico tiene al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la SEQ ID NO: 5. En algunas modalidades, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 5 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 5 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 58 a 4374 de la SEQ ID NO: 5. En incluso otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 1-2277 y 2320-4374 de la SEQ ID NO: 5 (es decir, los nucleótidos 1-4374 de la SEQ ID NO: 5 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 1 a 4374 de la SEQ ID NO: 5.

15

[0175] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58 a 4374 de la SEQ ID NO: 6. En otras modalidades, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En determinadas modalidades, la secuencia de ácido nucleico tiene al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a la SEQ ID NO: 6. En algunas modalidades, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 58 a 4374 de la SEQ ID NO: 6. En incluso otras modalidades, la secuencia de nucleótidos comprende los nucleótidos 1-2277 y 2320-4374 de la SEQ ID NO: 6 (es

35

decir, los nucleótidos 1-4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 1 a 4374 de la SEQ ID NO: 6.

[0176] En algunas modalidades, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que codifica un péptido señal. En determinadas modalidades, el péptido señal es un péptido señal de FVIII. En algunas modalidades, la secuencia de ácido nucleico que codifica un péptido señal está optimizado con codones. En una modalidad particular, la secuencia de ácido nucleico que codifica un péptido señal tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o al menos 100 % de identidad de secuencia con respecto a (i) los nucleótidos 1 a 57 de la SEQ ID NO: 1; (ii) los nucleótidos 1 a 57 de la SEQ ID NO: 2; (iii) los nucleótidos 1 a 57 de la SEQ ID NO: 3; (iv) los nucleótidos 1 a 57 de la SEQ ID NO: 4; (v) los nucleótidos 1 a 57 de la SEQ ID NO: 5; (vi) los nucleótidos 1 a 57 de la SEQ ID NO: 6; (vii) los nucleótidos 1 a 57 de la SEQ ID NO: 70; (viii) los nucleótidos 1 a 57 de la SEQ ID NO: 71; o (ix) los nucleótidos 1 a 57 de la SEQ ID NO: 68.

[0177] SEQ ID NO: 1-6, 70 y 71 son versiones optimizadas de la SEQ ID NO: 16, la secuencia de nucleótidos de FVIII de inicio o parental o «de tipo salvaje». SEQ ID NO: 16 codifica un FVIII humano con dominio B eliminado. Mientras que las SEQ ID NO: 1-6, 70 y 71 derivan de una forma de FVIII con el dominio B eliminado específico (SEQ ID NO: 16), se debe entender que la presente descripción se refiere también a versiones optimizadas de ácidos nucleicos que codifican otras versiones del FVIII. Por ejemplo, otras versiones del FVIII pueden incluir FVIII de longitud completa, otras eliminaciones del dominio B del FVIII (que se describen más adelante) u otros fragmentos del FVIII que conservan la actividad del FVIII.

[0178] Tal como se utiliza en la presente, «un polipéptido con actividad de FVIII» significa un polipéptido del FVIII funcional en su papel normal de coagulación, a menos que se especifique lo contrario. La expresión un polipéptido con actividad de FVIII incluye un fragmento, variante, análogo o derivado funcional de este que conserva la función del factor VIII de tipo salvaje y longitud completa en la vía de coagulación. «Un polipéptido con actividad de FVIII» se utiliza de forma indistinta con una proteína del FVIII, polipéptido del FVIII o FVIII. Los ejemplos de las funciones del FVIII incluyen, de modo no taxativo, una capacidad para activar la coagulación, una capacidad para actuar como un cofactor para el factor IX o una capacidad para 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. En una modalidad, un polipéptido que tiene actividad de FVIII comprende dos cadenas de polipéptidos, la primera cadena tiene la cadena pesada del FVIII y la segunda cadena tiene la cadena ligera del FVIII. En otra modalidad, el polipéptido que tiene actividad de FVIII es un FVIII de cadena simple. El FVIII de cadena simple puede contener una o más mutaciones o sustituciones en el residuo de aminoácidos 1645 y/o 1648 que corresponde a la secuencia del FVIII maduro. Véase la solicitud internacional n.º PCT/US2012/045784, que se incorpora en su totalidad mediante esta referencia. La proteína del FVIII puede ser la proteína del FVIII humano, porcino, canino, de rata o murino. Además, las comparaciones entre el FVIII de humanos y otras especies identificaron residuos conservados que probablemente se requieren para funcionar (Cameron et ál., *Thromb. Haemost.* 79:317-22 (1998); US 6,251,632).

[0179] El «dominio B» del FVIII, tal como se utiliza en la presente, es el mismo que el dominio B que se conoce en la técnica que se define mediante una identidad de secuencia de aminoácidos interna y sitios de escisión proteolítica, por ejemplo, los residuos Ser741-Arg1648 del FVIII humano de longitud completa. Los otros dominios del FVIII humano se definen mediante los siguientes residuos de aminoácidos: A1, residuos Ala1-Arg372; A2, residuos Ser373-Arg740; A3, residuos Ser1690-Ile2032; C1, residuos Arg2033-Asn2172; C2, residuos Ser2173-Tyr2332. La secuencia A3-C1-C2 incluye residuos Ser1690-Tyr2332. En la secuencia remanente, los residuos Glu1649-Arg1689, se conocen normalmente como el péptido de activación de cadena ligera del FVIII. Las ubicaciones de los límites para todos los dominios, incluyendo los dominios B, para el FVIII porcino, de ratón y canino también se conocen en la técnica. Un ejemplo de un FVIII con BDD es el FVIII de BDD recombinante REFACTO® (Wyeth Pharmaceuticals, Inc.).

[0180] Un «FVIII con dominio B eliminado» puede tener eliminaciones parciales o completas descritas en la patente estadounidense n.º 6,316,226, 6,346,513, 7,041,635, 5,789,203, 6,060,447, 5,595,886, 6,228,620, 5,972,885, 6,048,720, 5,543,502, 5,610,278, 5,171,844, 5,112,950, 4,868,112 y 6,458,563, cada una de las cuales se incorpora en su totalidad a la presente mediante esta referencia. En algunas modalidades, una secuencia del FVIII con dominio B eliminado de la presente descripción comprende cualquiera de las eliminaciones descritas en col. 4, línea 4 a col. 5, línea 28 y ejemplos 1-5 de la patente estadounidense n.º 6,316,226 (también en US 6,346,513). En algunas modalidades, un FVIII con dominio B

eliminado de la presente descripción tiene una eliminación descrita en col. 2, líneas 26-51 y ejemplos 5-8 de la patente estadounidense n.º 5,789,203 (también US 6,060,447, US 5,595,886 y US 6,228,620). En algunas modalidades, un FVIII con dominio B eliminado tiene una eliminación descrita en col. 1, líneas 25 a col. 2, línea 40 de la patente estadounidense n.º 5,972,885; col. 6, líneas 1-22 y el ejemplo 1 de la patente estadounidense n.º 6,048,720; col. 2, líneas 17-46 de la patente estadounidense n.º 5,543,502; col. 4, línea 22 a col. 5, línea 36 de la patente estadounidense n.º 5,171,844; col. 2, líneas 55-68, figura 2, y ejemplo 1 de la patente estadounidense n.º 5,112,950; col. 2, línea 2 a col. 19, línea 21 y tabla 2 de la patente estadounidense n.º 4,868,112; col. 2, línea 1 a col. 3, línea 19, col. 3, línea 40 a col. 4, línea 67, col. 7, línea 43 a col. 8, línea 26, y col. 11, línea 5 a col. 13, línea 39 de la patente estadounidense no. 7,041,635; o col. 4, líneas 25-53, de la patente estadounidense n.º 6,458,563. En algunas modalidades, el FVIII con dominio B eliminado tiene una eliminación de la mayoría de los dominios B, pero todavía contiene secuencias terminales amino del dominio B que son esenciales para el procesamiento proteolítico *in vivo* del producto primario de traducción en dos cadenas de polipéptidos, como se describe en WO 91/09122, que se incorpora en su totalidad a la presente mediante esta referencia. En algunas modalidades, un FVIII con dominio eliminado B se construye con una eliminación de los aminoácidos 747-1638, es decir, prácticamente una eliminación completa del dominio B. Hoebe R.C., *et ál. J. Biol. Chem.* 265 (13): 7318-7323 (1990), que se incorpora a la presente en su totalidad mediante esta referencia. Un FVIII con dominio B eliminado también puede contener una eliminación de los aminoácidos 771-1666 o aminoácidos 868-1562 del FVIII. Meulien P., *et ál. Protein Eng.* 2(4): 301-6 (1988), que se incorpora a la presente en su totalidad mediante esta referencia. Las eliminaciones adicionales del dominio B que forman parte de la descripción incluyen, por ejemplo, eliminación de los aminoácidos 982 hasta 1562 o 760 hasta 1639 (Toole *et ál., Proc. Natl. Acad. Sci. EUA* (1986) 83, 5939-5942)), 797 a 1562 (Eaton, *et ál. Biochemistry* (1986) 25:8343-8347)), 741 a través de 1646 (Kaufman (solicitud PCT publicada n.º WO 87/04187)), 747-1560 (Sarver, *et ál., DNA* (1987) 6:553-564)), 741 a través de 1648 (Pasek (solicitud PCT n.º 88/00831)), o 816 a través de 1598 o 741 a través de 1689 (Lagner (Behring Inst. Mitt. (1988) n.º 82:16-25, EP 295597)), cada una de las cuales se incorpora a la presente en su totalidad mediante esta referencia. Cada una de las eliminaciones anteriores pueden realizarse en cualquier secuencia del FVIII.

[0181] Se describe una cantidad de moléculas del FVIII funcionales que incluyen eliminaciones del dominio B, en las siguientes patentes US 6,316,226 y US

6,346,513, ambas asignadas a Baxter; US 7,041,635 asignada a In2Gen; US 5,789,203, US 6,060,447, US 5,595,886, y US 6,228,620 asignada a Chiron; US 5,972,885 y US 6,048,720 asignada a Biovitrum, US 5,543,502 y US 5,610,278 asignada a Novo Nordisk; US 5,171,844 asignada a Immuno Ag; US 5,112,950
5 asignada a Transgene S.A.; US 4,868,112 asignada a Genetics Institute, cada una de los cuales se incorpora a la presente mediante esta referencia en su totalidad.

Optimización con codones

[0182]En una modalidad, la presente descripción proporciona una molécula de ácido
10 nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de ácido nucleico ha sido optimizada con codones. En otra modalidad, la secuencia de ácido nucleico de partida que codifica un polipéptido con actividad de FVIII y que está sujeta a la optimización con codones es la SEQ ID NO: 16. En algunas modalidades, la
15 secuencia que codifica un polipéptido con actividad de FVIII está optimizada con codones para la expresión humana. En otras modalidades, la secuencia que codifica un polipéptido con actividad de FVIII está optimizada con codones para la expresión murina. SEQ ID NO: 1-6, 70 y 71 son versiones optimizadas con codones de la SEQ ID NO: 16, optimizadas para la expresión humana.

20 **[0183]**La expresión «optimizado/a con codones» cuando habla de regiones de codificación o génicas de las moléculas de ácido nucleico para la transformación de diversos hospedadores, hace referencia a la alteración de codones en las regiones de codificación o génicas de las moléculas de ácido nucleico para reflejar el uso del
25 codón típico del organismo hospedador sin alterar el polipéptido codificado por el ADN. Dicha optimización incluye remplazar al menos uno, o más de uno, o una cantidad considerable de codones con uno o más codones que se utilizan con más frecuencia en los genes de ese organismo.

30 **[0184]**Las desviaciones en la secuencia de nucleótidos que comprende los codones que codifican los aminoácidos de cualquier cadena de polipéptidos permite las variaciones en la secuencia que codifica el gen. Dado que dicho codón consiste en tres nucleótidos, y los nucleótidos que comprenden ADN se limitan a cuatro bases específicas, hay 64 combinaciones posibles de nucleótidos, 61 de las cuales
35 codifican aminoácidos (los tres codones restantes codifican señales que terminan la traducción). El «código genético» que muestra qué codones codifican qué aminoácidos se reproduce en la presente como Tabla 1. Como resultado, se

designan muchos aminoácidos por más de un codón. Por ejemplo, los aminoácidos alanina y prolina están codificados por cuatro tripletes, la serina y arginina en seis, mientras que el triptófano y la metionina están codificados por apenas un triplete. Esta degeneración permite que la composición de base de ADN varíe a lo largo de un amplio intervalo sin alterar la secuencia de aminoácidos de las proteínas codificadas por el ADN

Tabla 1: El código genético estándar

	T	C	A	G
T	TTT Phe (F) TTC " TTA Leu (L) TTG "	TCT Ser (S) TCC " TCA " TCG "	TAT Tyr (Y) TAC " TAA Detención TAG Detención	TGT Cys (C) TGC TGA Detención TGG Trp (W)
C	CTT Leu (L) CTC " CTA " CTG "	CCT Pro (P) CCC " CCA " CCG "	CAT His (H) CAC " CAA Gln (Q) CAG "	CGT Arg (R) CGC " CGA " CGG "
A	ATT Ile (I) ATC " ATA " ATG Met (M)	ACT Thr (T) ACC " ACA " ACG "	AAT Asn (N) AAC " AAA Lys (K) AAG "	AGT Ser (S) AGC " AGA Arg (R) AGG "
G	GTT Val (V) GTC " GTA " GTG "	GCT Ala (A) GCC " GCA " GCG "	GAT Asp (D) GAC " GAA Glu (E) GAG "	GGT Gly (G) GGC " GGA " GGG "

[0185] Muchos organismos muestran una tendencia de uso de codones particulares para codificar la inserción de un aminoácido particular en una cadena de péptidos

creciente. La preferencia de codones, o tendencia a los codones, las diferencias en el uso del codón entre organismos, es otorgada por la degeneración del código genético y está bien documentado entre otros organismos. La tendencia a los codones se correlaciona a menudo con la eficiencia de traducción del ARN mensajero (ARNm), que se cree a su vez que depende, entre otros, de las propiedades de los codones que se están traduciendo y la disponibilidad de las moléculas de ARN de transferencia (ARNt) particulares. La predominancia de los ARNt seleccionados en una célula es generalmente un reflejo de los codones más frecuentemente utilizados en la síntesis de péptidos. Por consiguiente, los genes se pueden confeccionar para obtener una expresión génica óptima en un organismo dado de acuerdo con la optimización con codones.

[0186] Dada la gran cantidad de secuencias génicas disponibles para una gran variedad de especies animales, vegetal y microbianas, se han calculado las frecuencias relativas del uso del codón. Las tablas con el uso del codón se encuentran disponibles, por ejemplo, en la «base de datos de uso de codones» disponible en www.kazusa.or.jp/codon/ (visitado el 18 de junio de 2012). Véase Nakamura, Y., et ál. Nucl. Acids Res. 28:292 (2000).

[0187] La asignación aleatoria de codones a una frecuencia optimizada para codificar una secuencia de polipéptidos dada se puede llevar a cabo calculando las frecuencias de los codones para cada aminoácido y luego asignando los codones a la secuencia de polipéptidos de forma aleatoria. Adicionalmente, se pueden utilizar diversos algoritmos y programas de software informático para calcular una secuencia óptima.

[0188] En algunas modalidades, la molécula de ácido nucleico comprende una o más propiedades: (a) la molécula de ácido nucleico o una parte de esta tiene un mayor índice de adaptación de codones humanos con respecto a la SEQ ID NO: 16; (b) la secuencia de nucleótidos o una parte de esta tiene una mayor frecuencia de codones óptimos con respecto a la EQ ID NO:16; (c) la secuencia de nucleótidos o una parte de esta contiene un porcentaje mayor de nucleótidos de G/C en comparación con el porcentaje de nucleótido de G/C en la SEQ ID NO: 16; (d) la secuencia de nucleótidos o una parte de esta tiene un mayor uso de codones sinónimos relativos con respecto a la SEQ ID NO: 16; (e) la secuencia de nucleótidos o una parte de esta tiene una menor cantidad eficaz de codones con respecto a la SEQ ID NO: 16; (f) la secuencia de nucleótidos contiene menos

secuencias MARS/ARS (SEQ ID NO: 21 y 22) con respecto a la SEQ ID NO: 16; (g) la secuencia de nucleótidos contiene menos elementos desestabilizantes (SEQ ID NO: 23 y 24) con respecto a la SEQ ID NO: 16; (i) la secuencia de nucleótidos no contiene una secuencia poli-T, (j) la secuencia de nucleótidos no contiene una secuencia poli-A; o (k) cualquier combinación de estos. En algunas modalidades, las moléculas de ácido nucleico contienen al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o diez características de (a) a (j).

10 Índice de adaptación de codones

[0189] En una modalidad, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en la presente que codifica un polipéptido con actividad de FVIII, donde el índice de adaptación de codones humanos se incrementa con respecto a la SEQ ID NO: 16.

15 Por ejemplo, la secuencia de nucleótidos puede tener un índice de adaptación de codones humanos que es al menos aproximadamente 0,75 (75 %), al menos aproximadamente 0,76 (76 %), al menos aproximadamente 0,77 (77 %), al menos aproximadamente 0,78 (78 %), al menos aproximadamente 0,79 (79 %), al menos aproximadamente 0,80 (80 %), al menos aproximadamente 0,81 (81 %), al menos aproximadamente 0,82 (82 %), al menos aproximadamente 0,83 (83 %), al menos aproximadamente 0,84 (84 %), al menos aproximadamente 0,85 (85 %), al menos aproximadamente 0,86 (86 %), al menos aproximadamente 0,87 (87 %), al menos aproximadamente 0,88 (88 %), al menos aproximadamente 0,89 (89 %), al menos aproximadamente 0,90 (90 %), al menos aproximadamente 0,91 (91 %), al menos aproximadamente 0,92 (92 %), al menos aproximadamente 0,93 (93 %), al menos aproximadamente 0,94 (94 %), al menos aproximadamente 0,95 (95 %), al menos aproximadamente 0,96 (96 %), al menos aproximadamente 0,97 (97 %), al menos aproximadamente 0,98 (98 %) o al menos aproximadamente 0,99 (99 %). En algunas modalidades, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .88 (88 %). En otras modalidades, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .91 (91 %). En otras modalidades, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .91 (97 %).

35

[0190] En una modalidad particular, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos

que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos

5 aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos

10 aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID

15 NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde se incrementa el índice de adaptación de codones humanos de la secuencia de nucleótidos con respecto a la SEQ ID NO: 16. En algunas modalidades, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente 0,75 (75 %), al menos

20 aproximadamente 0,76 (76 %), al menos aproximadamente 0,77 (77 %), al menos aproximadamente 0,78 (78 %), al menos aproximadamente 0,79 (79 %), al menos aproximadamente 0,80 (80 %), al menos aproximadamente 0,81 (81 %), al menos aproximadamente 0,82 (82 %), al menos aproximadamente 0,83 (83 %), al menos aproximadamente 0,84 (84 %), al menos aproximadamente 0,85 (85 %), al menos

25 aproximadamente 0,86 (86 %), al menos aproximadamente 0,87 (87 %), al menos aproximadamente 0,88 (88 %), al menos aproximadamente 0,89 (89 %), al menos aproximadamente 0,90 (90 %) o al menos aproximadamente ,91 (91 %). En una particular, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .88 (88 %). En otra modalidad, la

30 secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .91 (91 %).

[0191] En otra modalidad particular, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos

35 que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la segunda

secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos
 aproximadamente 85 %, al menos aproximadamente 86 %, al menos
 aproximadamente 87 %, al menos aproximadamente 88 %, al menos
 aproximadamente 89 %, al menos aproximadamente 90 %, al menos
 5 aproximadamente 91 %, al menos aproximadamente 92 %, al menos
 aproximadamente 93 %, al menos aproximadamente 94 %, al menos
 aproximadamente 95 %, al menos aproximadamente 96 %, al menos
 aproximadamente 97 %, al menos aproximadamente 98 %, o al menos
 10 aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos
 1792-2277 y 2320-4374 de la SEQ ID NO: 5 o (ii) 1792-2277 y 2320-4374 de la SEQ
 ID NO: 6; donde la parte de extremo N y la parte de extremo C juntas tienen una
 actividad polipeptídica de FVIII; y donde se incrementa el índice de adaptación de
 codones humanos de la secuencia de nucleótidos con respecto a la SEQ ID NO: 16.
 En algunas modalidades, la secuencia de nucleótidos tiene un índice de adaptación
 15 de codones humanos que es al menos 0.75 (75 %), al menos aproximadamente 0,76
 (76 %), al menos aproximadamente 0,77 (77 %), al menos aproximadamente 0,78
 (78 %), al menos aproximadamente 0,79 (79 %), al menos aproximadamente 0,80
 (80 %), al menos aproximadamente 0,81 (81 %), al menos aproximadamente 0,82
 (82 %), al menos aproximadamente 0,83 (83 %), al menos aproximadamente 0,84
 20 (84 %), al menos aproximadamente 0,85 (85 %), al menos aproximadamente 0,86
 (86 %), al menos aproximadamente 0,87 (87 %), o al menos aproximadamente 0,88
 (88 %). En una particular, la secuencia de nucleótidos tiene un índice de adaptación
 de codones humanos que es al menos aproximadamente ,83 (83 %). En otra
 modalidad, la secuencia de nucleótidos tiene un índice de adaptación de codones
 25 humanos que es al menos aproximadamente ,88 (88 %).

[0192] En otras modalidades, la presente descripción proporciona una molécula de
 ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un
 polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende
 30 una secuencia de ácido nucleico que tiene al menos 80 %, al menos 85 %, al menos
 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94
 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 %
 de identidad de secuencia con respecto a los nucleótidos 58-2277 y 2320-4374 de
 una secuencia de aminoácidos que se selecciona de la SEQ ID NO: 1, 2, 3, 4, 5, 6,
 35 70 y 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71
 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde se
 incrementa el índice de adaptación de codones humanos de la secuencia de

- nucleótidos con respecto a la SEQ ID NO: 16. En algunas modalidades, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos 0,75 (75 %), al menos aproximadamente 0,76 (76 %), al menos aproximadamente 0,77 (77 %), al menos aproximadamente 0,78 (78 %), al menos aproximadamente 0,79 (79 %), al menos aproximadamente 0,80 (80 %), al menos aproximadamente 0,81 (81 %), al menos aproximadamente 0,82 (82 %), al menos aproximadamente 0,83 (83 %), al menos aproximadamente 0,84 (84 %), al menos aproximadamente 0,85 (85 %), al menos aproximadamente 0,86 (86 %), al menos aproximadamente 0,87 (87 %), o al menos aproximadamente 0,88 (88 %). En una particular, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .75 (75 %). En otra modalidad, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .83 (83 %). En otra modalidad, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .88 (88 %). En otra modalidad, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .91 (91 %). En otra modalidad, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente .97 (97 %).
- 20 **[0193]** En algunas modalidades, la molécula de ácido nucleico aislada de la presente descripción tiene una mayor frecuencia de codones óptimos (FOP) con respecto a la SEQ ID NO: 16. En determinadas modalidades, el FOP de la molécula de ácido nucleico aislada es al menos aproximadamente 40, al menos aproximadamente 45, al menos aproximadamente 50, al menos aproximadamente 55, al menos aproximadamente 60, al menos aproximadamente 64, al menos aproximadamente 65, al menos aproximadamente 70, al menos aproximadamente 75, al menos aproximadamente 79, al menos aproximadamente 80, al menos aproximadamente 85 o al menos aproximadamente 90.
- 30 **[0194]** En otras modalidades, la molécula de ácido nucleico aislada de la presente descripción tiene un mayor uso de codones sinónimos relativos (RCSU) con respecto a la SEQ ID NO: 16. En algunas modalidades, el RCSU de la molécula de ácido nucleico aislada es mayor que 1,5. En otras modalidades, el RCSU de la molécula de ácido nucleico aislada es mayor que 2,0. En determinadas modalidades, el RCSU de la molécula de ácido nucleico aislada es al menos aproximadamente 1,5, al menos aproximadamente 1,6, al menos aproximadamente 1,7, al menos aproximadamente 1,8, al menos aproximadamente 1,9, al menos

aproximadamente 2,0, al menos aproximadamente 2,1, al menos aproximadamente 2,2, al menos aproximadamente 2,3, al menos aproximadamente 2,4, al menos aproximadamente 2,5, al menos aproximadamente 2,6 o al menos aproximadamente 2,7.

5

[0195] En incluso otras modalidades, la molécula de ácido nucleico aislada de la presente descripción tiene una menor cantidad efectiva de codones relativa con respecto a la SEQ ID NO: 16. En algunas modalidades, la molécula de ácido nucleico aislada tiene una cantidad eficaz de codones inferior a aproximadamente 50, inferior a aproximadamente 45, inferior a aproximadamente 40, inferior a aproximadamente 35, inferior a aproximadamente 30 o inferior a aproximadamente 25. En una modalidad particular, la molécula de ácido nucleico aislada tiene una cantidad eficaz de codones de aproximadamente 40, aproximadamente 35, aproximadamente 30, aproximadamente 25 o aproximadamente 20.

15

Optimización del contenido de G/C

[0196] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en la presente que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos contiene un porcentaje de nucleótidos de G/C más alto en comparación con el porcentaje de nucleótidos de G/C en la SEQ ID NO: 16. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 45 %, al menos aproximadamente 46 %, al menos aproximadamente 47 %, al menos aproximadamente 48 %, al menos aproximadamente 49 %, al menos aproximadamente 50 %, al menos aproximadamente 51 %, al menos aproximadamente 52 %, al menos aproximadamente 53 %, al menos aproximadamente 54 %, al menos aproximadamente 55 %, al menos aproximadamente 56 %, al menos aproximadamente 57 %, al menos aproximadamente 58 %, al menos aproximadamente 59 % o al menos aproximadamente 60 %.

[0197] En una modalidad particular, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera

35

secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos
 aproximadamente 85 %, al menos aproximadamente 86 %, al menos
 aproximadamente 87 %, al menos aproximadamente 88 %, al menos
 aproximadamente 89 %, al menos aproximadamente 90 %, al menos
 5 aproximadamente 91 %, al menos aproximadamente 92 %, al menos
 aproximadamente 93 %, al menos aproximadamente 94 %, al menos
 aproximadamente 95 %, al menos aproximadamente 96 %, al menos
 aproximadamente 97 %, al menos aproximadamente 98 %, o al menos
 10 aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos
 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los
 nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID
 NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una
 actividad polipeptídica del FVIII; y donde la secuencia de nucleótidos contiene un
 mayor porcentaje de nucleótidos de G/C en comparación con el porcentaje de
 15 nucleótidos de G/C en la SEQ ID NO: 16. En algunas modalidades, la secuencia de
 nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de
 G/C que es al menos aproximadamente 45 %, al menos aproximadamente 46 %, al
 menos aproximadamente 47 %, al menos aproximadamente 48 %, al menos
 aproximadamente 49 %, al menos aproximadamente 50 %, al menos
 20 aproximadamente 51 %, al menos aproximadamente 52 %, al menos
 aproximadamente 53 %, al menos aproximadamente 54 %, al menos
 aproximadamente 55 %, al menos aproximadamente 56 %, al menos
 aproximadamente 57 %, o al menos aproximadamente 58 % En una modalidad
 particular, la secuencia de nucleótidos que codifica un polipéptido con actividad de
 25 FVIII tiene un contenido de G/C que es al menos aproximadamente 58 %.

[0198] En otra modalidad, la presente descripción proporciona una molécula de ácido
 nucleico aislada que comprende una secuencia de nucleótidos que comprende una
 primera secuencia de ácido nucleico que codifica una parte de extremo N de un
 30 polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una
 parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido
 nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %,
 al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos
 aproximadamente 88 %, al menos aproximadamente 89 %, al menos
 35 aproximadamente 90 %, al menos aproximadamente 91 %, al menos
 aproximadamente 92 %, al menos aproximadamente 93 %, al menos
 aproximadamente 94 %, al menos aproximadamente 95 %, al menos

aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-4374 de la SEQ ID NO: 5; (ii) los nucleótidos 1792-4374 de la SEQ ID NO: 6, (iii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 5 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o (iv) 1792-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica del FVIII; y donde la secuencia de nucleótidos contiene un mayor porcentaje de nucleótidos de G/C en comparación con el porcentaje de nucleótidos de G/C en la SEQ ID NO: 16. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 45 %, al menos aproximadamente 46 %, al menos aproximadamente 47 %, al menos aproximadamente 48 %, al menos aproximadamente 49 %, al menos aproximadamente 50 %, al menos aproximadamente 51 %, al menos aproximadamente 52 %, al menos aproximadamente 53 %, al menos aproximadamente 54 %, al menos aproximadamente 55 %, al menos aproximadamente 56 %, o al menos aproximadamente 57 %. En una modalidad particular, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 52 %. En otra modalidad, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 55 %. En otra modalidad, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 57 %.

[0199] En otras modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 80 %, al menos 85 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 o (ii) los nucleótidos 58-2277 y 2320-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 sin los nucleótidos que codifican el

dominio B o fragmento de dominio B); y donde la secuencia de nucleótidos contiene un mayor porcentaje de nucleótidos de G/C en comparación con el porcentaje de nucleótidos de G/C en la SEQ ID NO: 16. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 45 %. En una modalidad particular, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 52 %. En otra modalidad, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 55 %. En otra modalidad, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 57 %. En otra modalidad, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 58 %. En incluso otra modalidad, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 60 %.

[0200] «Contenido de G/C» (o contenido de guanina-citosina) o «porcentaje de nucleótido de G/C» hace referencia al porcentaje de bases nitrogenadas en una molécula de ADN que son guanina o citosina. Es posible calcular el contenido de G/C mediante el uso de la siguiente fórmula:

$$\frac{G + C}{A + T + G + C} \times 100 \tag{III}$$

[0201] Los genes humanos son altamente heterogéneos en su contenido de G/C, es decir que algunos genes tienen un contenido de G/C que no supera el 20 % y otros genes tienen un contenido de G/C que puede alcanzar el 95 %. En general, los genes ricos en G/C se expresan considerablemente más. De hecho, se ha demostrado que aumentar el contenido de G/C de un gen puede conducir a una mayor expresión del gen, debido principalmente a un aumento en la transcripción y niveles de ARNm de estado estacionario. Véase Kudla *et ál.*, PLoS Biol., 4(6): e180 (2006).

Secuencias similares a la región de acoplamiento a la matriz

[0202] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en la presente que codifica un polipéptido con actividad de FVIII, donde la secuencia de

nucleótidos contiene menos secuencias MARS/ARS con respecto a la SEQ ID NO: 16. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 6, como máximo 5, como máximo 4, como máximo 3 o como máximo 2 secuencias MARS/ARS. En otras modalidades, la

5 secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 1 secuencia MARS/ARS. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene una secuencia MARS/ARS.

10 **[0203]**En una modalidad particular, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera

15 secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos

20 aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los

25 nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos contiene menos secuencias MARS/ARS con respecto a la SEQ ID NO: 16. En otras modalidades, las secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene

30 como máximo 6, como máximo 5, como máximo 4, como máximo 3 o como máximo 2 secuencias MARS/ARS. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 1 secuencia MARS/ARS. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene una secuencia MARS/ARS.

35

[0204]En otra modalidad, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una

primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-4374 de la SEQ ID NO: 5; (ii) los nucleótidos 1792-4374 de la SEQ ID NO: 6, (iii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 5 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o (iv) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos contiene menos secuencias MARS/ARS con respecto a la SEQ ID NO: 16. En otras modalidades, las secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 6, como máximo 5, como máximo 4, como máximo 3 o como máximo 2 secuencias MARS/ARS. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 1 secuencia MARS/ARS. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene una secuencia MARS/ARS.

[0205] En otras modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 80 %, al menos 85 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2,

3, 4, 5, 6, 70 o 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la secuencia de nucleótidos contiene menos secuencias MARS/ARS con respecto a la SEQ ID NO: 16. En otras modalidades, las secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como

5 máximo 6, como máximo 5, como máximo 4, como máximo 3 o como máximo 2 secuencias MARS/ARS. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 1 secuencia MARS/ARS. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene una secuencia MARS/ARS.

10

[0206] Se han identificado los elementos ricos en AT en la secuencia de nucleótidos del FVIII humano que comparten la similitud de secuencia con las secuencias de replicación de forma autónoma (ARS, por sus siglas en inglés) de *Saccharomyces cerevisiae* y las regiones de acoplamiento de matrices nucleares (MAR, por sus

15 siglas en inglés). (Fallux et ál., *Mol. Cell. Biol.* 16:4264-4272 (1996). Se ha demostrado que uno de estos elementos se une a los factores nucleares *in vitro* y reprime la expresión de un gen marcador de acetiltransferasa cloranfenicol (CAT). *Id.* Se ha planteado la hipótesis de que estas secuencias pueden contribuir a la represión transcripcional del gen del FVIII humano. Por lo tanto, en una modalidad,

20 se han suprimido todas las secuencias MAR/ARS en el gen del FVIII de la presente descripción. Hay cuatro secuencias MAR/ARS ATATTT (SEQ ID NO: 21) y tres secuencias MAR/ARS AAATAT (SEQ ID NO: 22) en la secuencia de FVIII parental (SEQ ID NO: 16). Todos estos sitios se mutaron para destruir las secuencias MAR/ARS en las secuencias del FVIII optimizadas (SEQ ID NO: 1-6). La ubicación

25 de cada uno de estos elementos y la secuencia de los nucleótidos correspondientes en las secuencias optimizadas se muestran en la Tabla 2 que sigue a continuación.

Tabla 2: Resumen de cambios a los elementos represivos

Ubicación del elemento	Secuenc ia del FVIII BDD de partida (SEQ ID NO: 16)	Secuencia de FVIII con BDD optimizada							
		SEQ ID NO: 1	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 4	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 70	SEQ ID NO: 71
Secuencias desestabilizantes									
639	ATTTA	GTT TA	GTT CA	GTT CA	GTT CA	GTT CA	GTT CA	GTT CA	GTTC A
1338	ATTTA	GTT TA	GTT CA	CTT CA	GTT CA	GTT CA	GTT CA	CTT CA	GTTC A
1449	ATTTA	CTTT A	CTT CA	CTT CA	CTT CA	CTT CA	CTT CA	CTT CA	CTTC A
1590	TAAAT	TAA AT	CAA GT	CAA GT	TAA GT	CAA GT	CAA GT	CAA GT	TAAG T
1623	TAAAT	CAA AA	GAA GA	CTA AG	CAA GA	CAA GA	CAA GA	TAA GT	CAAGA
2410	ATTTA	ATC TA	ATC TA	ATC TA	ATC TA	ATC TA	ATC TA	ATC TA	ATCT A
2586	ATTTA	GTT TA	GTT CA	GTT CA	GTT CA	GTT CA	GTT CA	GTT CA	GTTC A
2630	TAAAT	TGA AT	TGA AC	TGA AC	TGA AC	TCA AT	TGA AC	TCA AT	TGAA C
3884	ATTTA	ATC TG	ACC TG	ACC TG	ACC TG	ATC TG	ACC TG	ATC TG	ACCT G
3887	TAAAT	TGA AC	TGA AC	TGA AC	TGA AC	TGA AC	TGA AC	TGA AC	TGAA C
Sitios de unión al promotor potenciales									
641	TTATA	TTAT C	TCA TC	TCA TT	TCA TC	TCA TC	TCA TC	TCA TT	TCATC
1275	TATAA	CTA TA	TTA CA	CTA CA	GTA CA	CTA CA	CTA CA	CTA CA	GTAC A
1276	TTATA	TATA A	TAC AA	TAC AA	TAC AA	TAC AA	TAC AA	TAC AA	TACAA
1445	TTATA	TCA	TCA	TTA	TCA	TCA	TCA	TTA	TCATC

		TC	TC	TC	TC	TC	TC	TC	
1474	TATAA	TATA A	TAC AA	TAC AA	TAC AA	TAC AA	TAC AA	TAC AA	TACAA
1588	TATAA	TATA A	TAC AA	TAC AA	TATA A	TAC AA	TAC AA	TAC AA	TATAA
2614	TTATA	CTG TA	CTG TA	CTG TA	CTG TA	TTG TA	CTG TA	TTG TA	CTGT A
2661	TATAA	CAT CA	CAT CA	CAT CA	CAT CA	CAT CA	CAT CC	CAT CA	CATC C
3286	TATAA	TAT AA	TAC AA	TAC AA	TAC AA	TAC AA	TAC AA	TAC AA	TACAA
3840	TTATA	TTAT A	TTA CT	CTA CA	CTA CA	CTA CA	CTA CT	CTA CA	CTACT
Secuencias similares al acoplamiento de matriz (MARS/ARS)									
1287	ATATTT	GTA TCT	GTA CCT	GTA CCT	GTA TCT	GTA CCT	GTA CCT	GTA CCT	GTAT CT
1447	ATATTT	ATC TTT	ATC TTC	ATC TTC	ATC TTC	ATC TTC	ATC TTC	ATC TTC	ATCTT C
1577	AAATAT	AAA TCT	AGA TCT	AAA TCT	AAA TCT	AGA TCT	AGA TCT	AAA TCT	AAATC T
1585	AAATAT	AAG TAT	AAG TAC	AAG TAC	AAG TAT	AAG TAC	AAG TAC	AAG TAC	AAGTA T
2231	ATATTT	ACA TCA	ATA TCA	ACA TCA	ACA TCA	ACA TCT	ATA TCT	ACA TCT	ATATC T
3054	AAATAT	AAA CAT	GAA TAT	GAA CAT	GAA CAT	GAA CAT	GAA TAT	GAA CAT	GAATA T
3788	ATATTT	ATA TCT	ATA TCT	ACA TCT	ACA TCT	ACA TCT	ACA TCT	ACA TCT	ACATC T
Elementos de la secuencia rica en AU (ARE)									
2468	1. ATTT TATT	ACTT CATC	ACTTC ATC	ACTT CATT	ACTTC ATT	ACTTT ATT	ACTTT ATC	ACT TTA TT	ACTTT ATC
3790	2. ATTT TTAA	ATCTT TAA	ATCTT CAA	ATCT TCAA	ATCTT CAA	ATCTT CAA	ATCTT CAA	ATCT TCAA	ATCTT CAA
Secuencias poli A/poli T									
3273	AAAAAA	GAAA	GAA	GAA	GAA	GAA	CAA	GAA	CAAGA

	A	AAA	GAA	GAA	GAA	GAA	GAA	GAA	AG
			G	G	G	G	G	G	
4195	TTTTTT	TTCT	TTCT	TTCT	TTCT	TTCT	TTCT	TTCT	TTCTT
		TT	TC	TC	TC	TC	TC	TCC	CC
Sitios de empalme									
2203	GGTGAT	GGGG	GGCG	GGGG	GGGG	GGAG	GGAG	GGA	GGAGA
		AC	AC	AC	AC	AC	AC	GAC	C

Secuencias desestabilizantes

[0207]En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en la presente que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos contiene menos elementos desestabilizantes con respecto a la SEQ ID NO: 16. En otras modalidades, las secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 9, como máximo 8, como máximo 7, como máximo 6 o como máximo 5 elementos desestabilizantes. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 4, como máximo 3, como máximo 2 o como máximo 1 elemento desestabilizante. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un elemento desestabilizante.

[0208]En una modalidad particular, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los

nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos contiene menos elementos desestabilizantes con respecto a la SEQ ID NO: 16. En otras

5 modalidades, las secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 9, como máximo 8, como máximo 7, como máximo 6 o como máximo 5 elementos desestabilizantes. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como

10 máximo 4, como máximo 3, como máximo 2 o como máximo 1 elemento desestabilizante. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un elemento desestabilizante.

[0209] En otra modalidad, la presente descripción proporciona una molécula de ácido

15 nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos

20 aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos

25 aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-4374 de la SEQ ID NO: 5; (ii) los nucleótidos 1792-4374 de la SEQ ID NO: 6, (iii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 5

30 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o (iv) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos

35 contiene menos elementos desestabilizantes con respecto a la SEQ ID NO: 16. En otras modalidades, las secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 9, como máximo 8, como máximo 7, como

máximo 6 o como máximo 5 elementos desestabilizantes. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 4, como máximo 3, como máximo 2 o como máximo 1 elemento desestabilizante. En incluso otras modalidades, la secuencia de nucleótidos que
 5 codifica un polipéptido con actividad de FVIII no contiene un elemento desestabilizante.

[0210] En otras modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un
 10 polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos
 15 aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 o (ii) los
 20 nucleótidos 58-2277 y 2320-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la secuencia de nucleótidos contiene menos elementos desestabilizantes con respecto a la SEQ ID NO: 16. En otras
 25 modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 9, como máximo 8, como máximo 7, como máximo 6 o como máximo 5 elementos desestabilizantes. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 4, como máximo 3, como máximo 2 o como máximo 1 elemento
 30 desestabilizante. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un elemento desestabilizante.

[0211] Existen diez elementos desestabilizantes en la secuencia FVIII parental (SEQ
 35 ID NO: 16): seis secuencias ATTTA (SEQ ID NO: 23) y cuatro secuencias TAAAT (SEQ ID NO: 24). En una modalidad, las secuencias de estos sitios se mutaron para destruir los elementos desestabilizantes en el FVIII optimizado de las SEQ ID NO: 1-

6, 70 y 71. La ubicación de cada uno de estos elementos y la secuencia de los nucleótidos correspondientes en las secuencias optimizadas se muestran en la Tabla 2.

5 Sitios de unión al promotor potenciales

[0212] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en la presente que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos contiene menos sitios de unión al promotor con respecto a la SEQ ID NO: 16. En otras modalidades, las secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 9, como máximo 8, como máximo 7, como máximo 6 o como máximo 5 elementos sitios de unión al promotor. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 4, como máximo 3, como máximo 2 o como máximo 1 sitio de unión al promotor. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un sitio de unión al promotor.

[0213] En una modalidad particular, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos contiene menos sitios de unión al promotor potenciales con respecto a la SEQ ID NO: 16. En otras

modalidades, las secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 9, como máximo 8, como máximo 7, como máximo 6 o como máximo 5 elementos sitios de unión al promotor. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 4, como máximo 3, como máximo 2 o como máximo 1 sitio de unión al promotor. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un sitio de unión al promotor.

[0214] En otra modalidad, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-4374 de la SEQ ID NO: 5; (ii) los nucleótidos 1792-4374 de la SEQ ID NO: 6, (iii) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 5 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 5 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o (iv) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 6 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos contiene menos sitios de unión al promotor con respecto a la SEQ ID NO: 16. En otras modalidades, las secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 9, como máximo 8, como máximo 7, como máximo 6 o como máximo 5 elementos sitios de unión al promotor. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 4, como máximo 3, como máximo 2 o como máximo 1 sitio de unión al promotor. En incluso otras modalidades, la secuencia de

nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un sitio de unión al promotor.

[0215] En otras modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la secuencia de nucleótidos contiene menos sitios de unión al promotor con respecto a la SEQ ID NO: 16. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 9, como máximo 8, como máximo 7, como máximo 6 o como máximo 5 elementos sitios de unión al promotor. En otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene como máximo 4, como máximo 3, como máximo 2 o como máximo 1 sitio de unión al promotor. En incluso otras modalidades, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un sitio de unión al promotor.

[0216] Las cajas TATA son secuencias reguladoras que se encuentran a menudo en las regiones promotoras de las eucariotas. Estas sirven como sitio de unión de la proteína de unión a TATA (TBP), un factor de transcripción general. Las cajas TATA comprende comúnmente la secuencia TATAA (SEQ ID NO: 28) o una variante cercana. Las cajas TATA dentro de una secuencia de codificación, sin embargo, pueden inhibir la traducción de la proteína de longitud completa. Existen diez secuencias de unión al promotor en la secuencia del FVIII BDD de tipo salvaje (SEQ ID NO: 16): cinco secuencias TATAA (SEQ ID NO: 28) y cinco secuencias TTATA (SEQ ID NO: 29). En algunas modalidades, se suprime al menos 1, al menos 2, al

menos 3 o al menos 4 de los sitios de unión al promotor en los genes del FVIII de la presente descripción. En algunas modalidades, se suprimen al menos 5 de los sitios de unión al promotor en los genes del FVIII de la presente descripción. En otras modalidades, se suprime al menos 6, al menos 7 o al menos 8 de los sitios de unión al promotor en los genes del FVIII de la presente descripción. En una modalidad, se suprimen al menos 9 de los sitios de unión al promotor en los genes del FVIII de la presente descripción. En una modalidad particular, se suprimen todos los sitios de unión al promotor en los genes del FVIII de la presente descripción. La ubicación de cada sitio de unión al promotor posible y la secuencia de los nucleótidos correspondientes en las secuencias optimizadas se muestran en la Tabla 2.

Otros elementos reguladores negativos que actúan en cis

[0217] Además de las secuencias MAR/ARS, los elementos desestabilizantes y los posibles sitios del promotor descritos anteriormente, se pueden identificar varias secuencias potencialmente inhibidoras adicionales en la secuencia del FVIII BDD de tipo salvaje (SEQ ID NO: 16). Se pueden identificar dos elementos de secuencia ricos en AU (ARE) (ATTTTATT (SEQ ID NO: 30); y ATTTTTAA (SEQ ID NO: 31), junto con un sitio de poli-A (AAAAAAA; SEQ ID NO: 26), un sitio de poli-T (TTTTTT; SEQ ID NO: 25), y un sitio de empalme (GGTGAT; SEQ ID NO: 27) en la secuencia del FVIII BDD no optimizada. Uno o más de estos elementos se puede quitar de las secuencias del FVIII optimizadas. La ubicación de cada uno de estos sitios y la secuencia de los nucleótidos correspondientes en las secuencias optimizadas se muestran en la Tabla 2.

[0218] En determinadas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos

aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos no contiene uno o más elementos reguladores negativos que actúan en cis, por ejemplo, un sitio de empalme, una secuencia poli-T, una secuencia poli-A, una secuencia ARE y cualquier combinación de estos.

10 **[0219]** En otra modalidad, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido
15 nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos
20 aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 1792-4374 de la SEQ ID NO: 5; (ii) los nucleótidos 1792-4374 de la SEQ ID NO: 6, (iii) los nucleótidos 1792-2277 y 2320-
25 4374 de la SEQ ID NO: 5 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 5 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o (iv) los nucleótidos 1792-2277 y 2320-4374 de la SEQ ID NO: 6 (es decir, los nucleótidos 1792-4374 de la SEQ ID NO: 6; sin los nucleótidos que codifican el dominio B o fragmento de dominio B); donde la parte de extremo N y la parte de extremo C juntas
30 tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos no contiene uno o más elementos reguladores negativos que actúan en cis, por ejemplo, un sitio de empalme, una secuencia poli-T, una secuencia poli-A, una secuencia ARE y cualquier combinación de estos.

35 **[0220]** En otras modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende

una secuencia de ácido nucleico que tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la secuencia de nucleótidos no contiene uno o más elementos reguladores negativos que actúan en cis, por ejemplo, un sitio de empalme, una secuencia poli-T, una secuencia poli-A, una secuencia ARE y cualquier combinación de estos.

[0221] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos no contiene el sitio de empalme GGTGAT (SEQ ID NO: 27). En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una

secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos

5 aproximadamente 80 %, al menos aproximadamente 85 %, al menos
 aproximadamente 86 %, al menos aproximadamente 87 %, al menos
 aproximadamente 88 %, al menos aproximadamente 89 %, al menos
 aproximadamente 90 %, al menos aproximadamente 91 %, al menos
 aproximadamente 92 %, al menos aproximadamente 93 %, al menos

10 aproximadamente 94 %, al menos aproximadamente 95 %, al menos
 aproximadamente 96 %, al menos aproximadamente 97 %, al menos
 aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de

15 secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos no contiene una secuencia poli-T (SEQ ID NO: 25). En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una

20 primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos

25 aproximadamente 88 %, al menos aproximadamente 89 %, al menos
 aproximadamente 90 %, al menos aproximadamente 91 %, al menos
 aproximadamente 92 %, al menos aproximadamente 93 %, al menos
 aproximadamente 94 %, al menos aproximadamente 95 %, al menos
 aproximadamente 96 %, al menos aproximadamente 97 %, al menos

30 aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de

35 secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos no contiene una secuencia poli-A (SEQ ID NO: 26). En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una

primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido de FVIII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C de un polipéptido de FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-1791 de la SEQ ID NO: 3; (ii) los nucleótidos 1-1791 de la SEQ ID NO: 3; (iii) los nucleótidos 58-1791 de la SEQ ID NO: 4; o (iv) los nucleótidos 1-1791 de la SEQ ID NO: 4; donde la parte de extremo N y la parte de extremo C juntas tienen una actividad polipeptídica de FVIII; y donde la secuencia de nucleótidos no contiene un elemento ARE (SEQ ID NO: 30 o la SEQ ID NO: 31).

[0222] En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la secuencia de nucleótidos no contiene el sitio de empalme GGTGAT (SEQ ID NO: 27). En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad

de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la secuencia de nucleótidos no contiene una secuencia poli-T (SEQ ID NO: 25). En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la secuencia de nucleótidos no contiene una secuencia poli-A (SEQ ID NO: 26). En algunas modalidades, la presente descripción proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos

aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de una secuencia de aminoácidos que se selecciona de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 y 71 (es decir, los nucleótidos 58-4374 de la SEQ ID NO: 1, 2, 3, 4, 5, 6, 70 o 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y donde la secuencia de nucleótidos no contiene un elemento ARE (SEQ ID NO: 30 o la SEQ ID NO: 31).

[0223] En otras modalidades, una secuencia del FVIII optimizada de la descripción no comprende uno o más motivos antivirales, estructuras de bucle-tallo y secuencias de repetición.

15

[0224] En incluso otras modalidades, los nucleótidos que rodean el sitio de inicio de la transcripción se cambian a una secuencia de consenso de kozak (GCCGCCACCATGC (SEQ ID NO: 32), donde los nucleótidos subrayados son el codón de inicio). En otras modalidades, los sitios de restricción se pueden agregar o quitar para facilitar el proceso de clonación.

20

Secuencias de nucleótidos heterólogos

[0225] En algunas modalidades, las moléculas de ácido nucleico aisladas de la descripción comprenden además una secuencia de nucleótidos heterólogo. En algunas modalidades, las moléculas de ácido nucleico aisladas de la descripción comprenden además al menos una secuencia de nucleótidos heterólogo. La secuencia de nucleótidos heterólogo puede unirse con las secuencias de nucleótidos de FVIII BDD optimizadas de la descripción en el extremo 5', en el extremo 3' o insertarse en la mitad de la secuencia de nucleótidos del BDD-FVIII optimizada. Por lo tanto, en algunas modalidades, la secuencia de aminoácidos heteróloga codificada por la secuencia de nucleótidos heterólogo está unida al extremo N o el extremo C de la secuencia de aminoácidos del FVIII codificada por la secuencia de nucleótidos o insertada entre dos aminoácidos en la secuencia de aminoácidos del FVIII. En algunas modalidades, la secuencia de aminoácidos heteróloga puede estar insertada entre dos aminoácidos en uno o más sitios de inserción que se seleccionan de la Tabla 3. En algunas modalidades, la secuencia de aminoácidos heteróloga se puede insertar dentro del polipéptido del FVIII codificado por la

25

30

35

molécula de ácido nucleico de la descripción en cualquier sitio descrito en la publicación internacional n.º WO 2013/123457 A1, WO 2015/106052 A1 o la publicación estadounidense n.º 2015/0158929 A1, las cuales se incorporan a la presente mediante referencia en su totalidad.

5

[0226] En algunas modalidades, la secuencia de aminoácidos heteróloga codificada por la secuencia de nucleótido heteróloga se inserta dentro del dominio B o un fragmento de este. En algunas modalidades, la secuencia de aminoácidos heteróloga se inserta dentro del FVIII inmediatamente después de un aminoácido correspondiente al aminoácido 745 del FVIII humano maduro (SEQ ID NO:15). En una modalidad particular, el FVIII comprende una eliminación de aminoácidos 746-1646 que corresponde al FVIII humano maduro (SEQ ID NO:15), y la secuencia de aminoácidos heteróloga codificada por la secuencia de nucleótidos heteróloga se inserta inmediatamente después del aminoácido 745, que corresponde al FVIII humano maduro (SEQ ID NO:15).

10

15

TABLA 3: Sitios de inserción de restos heterólogos

Sitio de inserción	Dominio	Sitio de inserción	Dominio	Sitio de inserción	Dominio
3	A1	375	A2	1749	A3
18	A1	378	A2	1796	A3
22	A1	399	A2	1802	A3
26	A1	403	A2	1827	A3
40	A1	409	A2	1861	A3
60	A1	416	A2	1896	A3
65	A1	442	A2	1900	A3
81	A1	487	A2	1904	A3
116	A1	490	A2	1905	A3
119	A1	494	A2	1910	A3
130	A1	500	A2	1937	A3
188	A1	518	A2	2019	A3
211	A1	599	A2	2068	C1
216	A1	603	A2	2111	C1
220	A1	713	A2	2120	C1
224	A1	745	B	2171	C2
230	A1	1656	Región	2188	C2

			a3		
333	A1	1711	A3	2227	C2
336	A1	1720	A3	2332	CT
339	A1	1725	A3		

Nota: Los sitios de inserción indican la posición del aminoácido que corresponde a una posición del aminoácido del FVIII humano maduro (SEQ ID NO: 15).

5 **[0227]**En otras modalidades, las moléculas de ácido nucleico aisladas de la descripción comprenden además dos, tres, cuatro, cinco, seis, siete u ocho secuencias de nucleótidos heterólogas. En algunas modalidades, todas las secuencias de nucleótidos heterólogas son idénticas. En algunas modalidades, al menos una secuencia de nucleótidos heteróloga es diferente a las otras secuencias
10 de nucleótidos heterólogas. En algunas modalidades, la descripción puede comprender dos, tres, cuatro, cinco, seis o más de siete secuencias de nucleótidos heterólogas en tándem.

15 **[0228]**En algunas modalidades, la secuencia de nucleótidos heteróloga codifica una secuencia de aminoácidos. En algunas modalidades, la secuencia de aminoácidos codificada por la secuencia de nucleótidos heteróloga es un resto heterólogo que puede comprender la semivida (un «extensor de semivida») de una molécula del FVIII.

20 **[0229]**En algunas modalidades, el resto heterólogo es un péptido o un polipéptido con características estructuradas o no estructuradas que se asocian con la prolongación de la semivida *in vivo* cuando se incorporan en una proteína de la descripción. Los ejemplos no taxativos incluyen albúmina, fragmentos de albúmina, fragmentos Fc de inmunoglobulinas, el péptido de extremo C (CTP) de la subunidad
25 β de la gonadotropina coriónica humana, una secuencia HAP, una secuencia XTEN, una transferrina o un fragmento de esta, un polipéptido PAS, enlazadores de poliglicina, enlazadores de poliserina, restos de unión a la albúmina, o cualquier fragmento, derivados, variantes o combinaciones de estos polipéptidos. En una modalidad particular, la secuencia de aminoácidos heteróloga es una región
30 constante de inmunoglobulina o una parte de esta, transferrina, albúmina o una secuencia PAS. En algunos aspectos, un resto heterólogo incluye el factor de von Willebrand o un fragmento de este. En otros aspectos relacionados, un resto heterólogo puede incluir un sitio de acoplamiento (por ejemplo, un aminoácido de

cisteína) para un resto no polipeptídico tal como polietilenglicol (PEG), almidón hidroxietílico (HES), ácido polisiálico o cualquier derivado, variante o combinación de estos elementos. En algunos aspectos, un resto heterólogo comprende un aminoácido de cisteína que funciona como un sitio de acoplamiento para un resto no polipeptídico tal como polietilenglicol (PEG), almidón hidroxietílico (HES), ácido polisiálico o cualquier derivado, variante o combinación de estos elementos.

[0230] En una modalidad específica, una primera secuencia de nucleótidos heteróloga codifica un primer resto heterólogo que es una molécula que extiende la semivida y una segunda secuencia de nucleótidos heteróloga codifica un segundo resto heterólogo que puede también ser una molécula que extiende la semivida la cual es conocida en la técnica. En determinadas modalidades, el primer resto heterólogo (por ejemplo, un primer resto Fc) y el segundo resto heterólogo (por ejemplo, un segundo resto Fc) se asocian entre sí para formar un dímero. En una modalidad, el segundo resto heterólogo es un segundo resto Fc, donde el segundo resto Fc está unido o asociado con el primer resto heterólogo, por ejemplo, el primer resto Fc. Por ejemplo, el segundo resto heterólogo (por ejemplo, el segundo resto Fc) puede unirse al primer resto heterólogo (por ejemplo, el primer resto Fc) mediante un enlazador o puede asociarse con el primer resto heterólogo mediante un enlace covalente o no covalente.

[0231] En algunas modalidades, el resto heterólogo es un polipéptido que comprende, consiste esencialmente o consiste en al menos aproximadamente 10, al menos aproximadamente 100, al menos aproximadamente 200, al menos aproximadamente 300, al menos aproximadamente 400, al menos aproximadamente 500, al menos aproximadamente 600, al menos aproximadamente 700, al menos aproximadamente 800, al menos aproximadamente 900, al menos aproximadamente 1000, al menos aproximadamente 1100, al menos aproximadamente 1200, al menos aproximadamente 1300, al menos aproximadamente 1400, al menos aproximadamente 1500, al menos aproximadamente 1600, al menos aproximadamente 1700, al menos aproximadamente 1800, al menos aproximadamente 1900, al menos aproximadamente 2000, al menos aproximadamente 2500, al menos aproximadamente 3000 o al menos aproximadamente 4000 aminoácidos. En otras modalidades, el resto heterólogo es un polipéptido que comprende, consiste esencialmente o consiste en aproximadamente 100 a aproximadamente 200 aminoácidos, aproximadamente 200 a aproximadamente 300 aminoácidos, aproximadamente 300 to aproximadamente

400 aminoácidos, aproximadamente 400 a aproximadamente 500 aminoácidos, aproximadamente 500 to aproximadamente 600 aminoácidos, aproximadamente 600 a aproximadamente 700 aminoácidos, aproximadamente 700 to aproximadamente 800 aminoácidos, aproximadamente 800 a aproximadamente 900 aminoácidos o
5 aproximadamente 900 to aproximadamente 1000 aminoácidos.

[0232] En determinadas modalidades, un resto heterólogo mejora una o más propiedades farmacocinéticas de la proteína de FVIII sin afectar significativamente su actividad o función biológica.

10

[0233] En determinadas modalidades, un resto heterólogo aumenta la semivida *in vivo* y/o *in vitro* de la proteína de FVIII de la descripción. En otras modalidades, un resto heterólogo facilita la visualización o localización de la proteína de FVIII de la descripción o un fragmento de esta (por ejemplo, un fragmento que comprende un
15 resto heterólogo después de la escisión proteolítica de la proteína de FVIII). La visualización y/o ubicación de la proteína de FVIII de la descripción o un fragmento de esta puede ser *in vivo*, *in vitro*, *ex vivo*, o combinaciones de estos.

[0234] En otras modalidades, un resto heterólogo aumenta la estabilidad de la
20 proteína de FVIII de la descripción o un fragmento de esta (por ejemplo, un fragmento que comprende un resto heterólogo después de la escisión proteolítica de la proteína de FVIII). Tal como se utiliza en la presente, el término «estabilidad» se refiere a una medida reconocida en la técnica del mantenimiento de una o más propiedades físicas de una proteína del FVIII en respuesta a una condición
25 ambiental (por ejemplo, temperatura elevada o disminuida). En determinados aspectos, la propiedad física puede ser el mantenimiento de la estructura covalente de la proteína de FVIII (por ejemplo, la ausencia de escisión proteolítica, oxidación o desamidación indeseada). En otro aspecto, la propiedad física puede ser también la presencia de la proteína de FVIII en un estado adecuadamente plegado (por
30 ejemplo, la ausencia de agregados o precipitados solubles o insolubles). En un aspecto, la estabilidad de la proteína de FVIII se mide sometiendo a ensayo una propiedad biofísica de la proteína de FVIII, por ejemplo, estabilidad térmica, perfil de pH de despliegue, eliminación estable de la glicosilación, solubilidad, función bioquímica (por ejemplo, capacidad para unirse a una proteína, receptor o ligando),
35 etc., y/o combinaciones de las mismas. En otro aspecto, la función bioquímica se demuestra por medio de la afinidad de unión de la interacción. En un aspecto, una medida de estabilidad de proteínas es la estabilidad térmica, es decir, resistencia a

la prueba térmica. La estabilidad puede medirse utilizando métodos conocidos en la técnica, tales como, HPLC (cromatografía líquida de alto rendimiento), SEC (cromatografía de exclusión por tamaño), DLS (dispersión de luz dinámica), etc. Los métodos para medir la estabilidad térmica incluyen, a modo no taxativo, calorimetría diferencial barrido (DSC, por sus siglas en inglés), fluorimetría diferencial de barrido (DSF, por sus siglas en inglés), dicroísmo circular (CD, por sus siglas en inglés) y ensayo de desafío térmico.

[0235] En determinados aspectos, una proteína de FVIII codificada por la molécula de ácido nucleico de la descripción comprende al menos un extensor de semivida, es decir, un resto heterólogo que aumenta la semivida *in vivo* de la proteína de FVIII con respecto a la semivida *in vivo* de la proteína de FVIII correspondiente que carece de dicho resto heterólogo. La semivida *in vivo* de una proteína de FVIII puede determinarse mediante cualquier método conocido para los expertos en la técnica, por ejemplo, ensayos de actividad (ensayo cromogénico o ensayo aPTT de coagulación en una etapa), ELISA, ROTEMTM, etc.

[0236] En algunas modalidades, la presencia de uno o más extensores de semivida resulta en la semivida de la proteína de FVIII que se aumentará en comparación con la semivida de la proteína correspondiente que carece de dicho uno o más extensores de semivida. La semivida de la proteína de FVIII que comprende un extensor de semivida es al menos aproximadamente 1,5 veces, al menos aproximadamente 2 veces, al menos aproximadamente 2,5 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 6 veces, al menos aproximadamente 7 veces, al menos aproximadamente 8 veces, al menos aproximadamente 9 veces, al menos aproximadamente 10 veces, al menos aproximadamente 11 veces o al menos aproximadamente 12 veces mayor que la semivida *in vivo* de la proteína de FVIII correspondiente que carece de dicho extensor de semivida.

[0237] En una modalidad, la semivida de la proteína de FVIII que comprende un extensor de semivida es aproximadamente 1,5 veces a aproximadamente 20 veces, aproximadamente 1,5 veces a aproximadamente 15 veces, o aproximadamente 1,5 veces a aproximadamente 10 veces mayor que la semivida *in vivo* de la proteína correspondiente que carece de dicha extensión de semivida. En otra modalidad, la semivida de la proteína de FVIII que comprende un extensor de semivida se

extiende aproximadamente 2 veces a aproximadamente 10 veces, aproximadamente
 2 veces a aproximadamente 9 veces, aproximadamente 2 veces a aproximadamente
 8 veces, aproximadamente 2 veces a aproximadamente 7 veces, aproximadamente
 2 veces a aproximadamente 6 veces, aproximadamente 2 veces a aproximadamente
 5 veces, aproximadamente 2 veces a aproximadamente 4 veces, aproximadamente
 2 veces a aproximadamente 3 veces, aproximadamente 2,5 veces a
 aproximadamente 10 veces, aproximadamente 2,5 veces a aproximadamente 9
 veces, aproximadamente 2,5 veces a aproximadamente 8 veces, aproximadamente
 2,5 veces a aproximadamente 7 veces, aproximadamente 2,5 veces a
 aproximadamente 6 veces, aproximadamente 2,5 veces a aproximadamente 5
 veces, aproximadamente 2,5 veces a aproximadamente 4 veces, aproximadamente
 2,5 veces a aproximadamente 3 veces, aproximadamente 3 veces a
 aproximadamente 10 veces, aproximadamente 3 veces a aproximadamente 9 veces,
 aproximadamente 3 veces a aproximadamente 8 veces, aproximadamente 3 veces a
 aproximadamente 7 veces, aproximadamente 3 veces a aproximadamente 6 veces,
 aproximadamente 3 veces a aproximadamente 5 veces, aproximadamente 3 veces a
 aproximadamente 4 veces, aproximadamente 4 veces a aproximadamente 6 veces,
 aproximadamente 5 veces a aproximadamente 7 veces o aproximadamente 6 veces
 a aproximadamente 8 veces en comparación con la semivida *in vivo* de la proteína
 correspondiente que carece de dicho extensor de semivida.

[0238] En otras modalidades, la semivida de la proteína de FVIII que comprende un
 extensor de semivida es de al menos aproximadamente 17 horas, al menos
 aproximadamente 18 horas, al menos aproximadamente 19 horas, al menos
 aproximadamente 20 horas, al menos aproximadamente 21 horas, al menos
 aproximadamente 22 horas, al menos aproximadamente 23 horas, al menos
 aproximadamente 24 horas, al menos aproximadamente 25 horas, al menos
 aproximadamente 26 horas, al menos aproximadamente 27 horas, al menos
 aproximadamente 28 horas, al menos aproximadamente 29 horas, al menos
 aproximadamente 30 horas, al menos aproximadamente 31 horas, al menos
 aproximadamente 32 horas, al menos aproximadamente 33 horas, al menos
 aproximadamente 34 horas, al menos aproximadamente 35 horas, al menos
 aproximadamente 36 horas, al menos aproximadamente 48 horas, al menos
 aproximadamente 60 horas, al menos aproximadamente 72 horas, al menos
 aproximadamente 84 horas, al menos aproximadamente 96 horas, o al menos
 aproximadamente 108 horas.

[0239] En incluso otras modalidades, la semivida de la proteína de FVIII que comprende un extensor de semivida es aproximadamente 15 horas a aproximadamente dos semanas, aproximadamente 16 horas a aproximadamente una semana, aproximadamente 17 horas a aproximadamente una semana, aproximadamente 18 horas a aproximadamente una semana, aproximadamente 19 horas a aproximadamente una semana, aproximadamente 20 horas a aproximadamente una semana, aproximadamente 21 horas a aproximadamente una semana, aproximadamente 22 horas a aproximadamente una semana, aproximadamente 23 horas a aproximadamente una semana, aproximadamente 24 horas a aproximadamente una semana, aproximadamente 36 horas a aproximadamente una semana, aproximadamente 48 horas a aproximadamente una semana, aproximadamente 60 horas a aproximadamente una semana, aproximadamente 24 horas a aproximadamente seis días, aproximadamente 24 horas a aproximadamente cinco días, aproximadamente 24 horas a aproximadamente cuatro días, aproximadamente 24 horas a aproximadamente tres días, o aproximadamente 24 horas a aproximadamente dos días.

[0240] En algunas modalidades, la semivida promedio de la proteína de FVIII por sujeto es de aproximadamente 15 horas, aproximadamente 16 horas, aproximadamente 17 horas, aproximadamente 18 horas, aproximadamente 19 horas, aproximadamente 20 horas, aproximadamente 21 horas, aproximadamente 22 horas, aproximadamente 23 horas, aproximadamente 24 horas (1 día), aproximadamente 25 horas, aproximadamente 26 horas, aproximadamente 27 horas, aproximadamente 28 horas, aproximadamente 29 horas, aproximadamente 30 horas, aproximadamente 31 horas, aproximadamente 32 horas, aproximadamente 33 horas, aproximadamente 34 horas, aproximadamente 35 horas, aproximadamente 36 horas, aproximadamente 40 horas, aproximadamente 44 horas, aproximadamente 48 horas (2 días), aproximadamente 54 horas, aproximadamente 60 horas, aproximadamente 72 horas (3 días), aproximadamente 84 horas, aproximadamente 96 horas (4 días), aproximadamente 108 horas, aproximadamente 120 horas (5 días), aproximadamente seis días, aproximadamente siete días (una semana), aproximadamente ocho días, aproximadamente nueve días, aproximadamente 10 días, aproximadamente 11 días, aproximadamente 12 días, aproximadamente 13 días o aproximadamente 14 días.

35

[0241] Se pueden fusionar uno o más extensores de semivida a los extremos N o extremos C del FVIII o insertarse dentro del FVIII.

1. Una región constante de inmunoglobulina o una parte de esta

[0242] En otro aspecto, un resto heterólogo comprende una o más regiones constantes de inmunoglobulina o partes de estas (por ejemplo, una región Fc). En una modalidad, una molécula de ácido nucleico aislada de la descripción comprende
5 además una secuencia de ácido nucleico heteróloga que codifica una región constante de inmunoglobulina o una parte de esta. En algunas modalidades, la región constante de inmunoglobulina o una parte de esta es una región Fc.

[0243] Una región constante de inmunoglobulina está compuesta por dominios denotados dominios CH (pesados constantes) (CH1, CH2, etc.). Dependiendo del
10 isotipo, (es decir, IgG, IgM, IgA, IgD o IgE), la región constante puede comprender tres o cuatro dominios CH. *Algunas regiones constantes de isotipos (por ejemplo, IgG)* contienen también una región bisagra. Véase Janeway et ál. 2001, *Immunobiology*, Garland Publishing, N.Y., N.Y.

[0244] Una región constante de inmunoglobulina o una parte de esta para producir la proteína de FVIII de la presente invención puede obtenerse de una cantidad de fuentes diferentes. En una modalidad, una región constante de inmunoglobulina o una parte de esta deriva de una inmunoglobulina humana. No obstante, se entiende
15 que la región constante de inmunoglobulina o una parte de esta puede derivar de una inmunoglobulina de otra especie de mamífero que incluye, por ejemplo, una especie de roedor (por ejemplo, ratón, rata, conejo, cobayo) o primate no humano (por ejemplo, chimpancé, macaco). Asimismo, la región constante de inmunoglobulina o una parte de esta puede derivar de cualquier clase de
20 inmunoglobulina, lo que incluye IgM, IgG, IgD, IgA e IgE, y cualquier isotipo de inmunoglobulina, lo que incluye IgG1, IgG2, IgG3 e IgG4. En una modalidad, se utiliza la IgG1 de isotipo humano.

[0245] Una variedad de secuencias génicas de la región constante de
30 inmunoglobulina (por ejemplo, secuencias génicas de la región constante humana) se encuentran disponibles en la forma de depósitos accesibles públicamente. Se puede seleccionar una secuencia de dominios de la región constante que tenga una función efectora particular (o que carezca de una función efectora particular) o con una modificación particular para reducir la inmunogenicidad. Se han publicado varias
35 secuencias de anticuerpos y genes que codifican anticuerpos y secuencias adecuadas de la región constante de Ig (por ejemplo, secuencias CH2 y/o CH3, bisagra, o partes de estas) pueden obtenerse de estas secuencias que utilizan

técnicas reconocidas en la materia. El material genético obtenido utilizando cualquiera de los métodos anteriores puede alterarse o sintetizarse entonces para obtener los polipéptidos de la presente descripción. Se apreciará además que el alcance de la presente descripción abarca alelos, variantes y mutaciones de las

5 secuencias de ADN de la región constante.

[0246] Las secuencias de la región constante de inmunoglobulina o una parte de esta pueden clonarse, por ejemplo, utilizando la reacción en cadena de la polimerasa y cebadores que se seleccionan para amplificar el dominio de interés.

10 Para clonar una secuencia de la región constante de Ig o una parte de esta de un anticuerpo, el ARNm puede aislarse del hibridoma, bazo o células linfáticas, ADN de transcripción inversa y genes de anticuerpos amplificados por PCR. Los métodos de amplificación por PCR se describen en detalle en las Patentes estadounidenses n.º 4,683,195; 4,683,202; 4,800,159; 4,965,188 y en, por ejemplo, "PCR Protocols: A

15 Guide to Methods and Applications" Innis et ál. eds., Academic Press, San Diego, CA (1990); Ho et ál. 1989. *Gene* 77:51; Horton et ál. 1993. *Methods Enzymol.* 217:270). La PCR puede iniciarse mediante cebadores de región constante de consenso o mediante cebadores más específicos basados en el ADN de cadena pesada y ligera y en las secuencias de aminoácidos publicados. También se puede

20 utilizar la PCR para aislar clones de ADN que codifican las cadenas ligeras y pesadas del anticuerpo. En este caso, las bibliotecas se pueden analizar mediante cebadores de consenso o sondas homólogas más grandes, tales como sondas de regiones constantes de ratón. En la técnica se conocen numerosos grupos de cebadores adecuados para la amplificación de los genes de anticuerpos (por

25 ejemplo, cebadores 5' basados en la secuencia de extremo N de anticuerpos purificados (Benhar y Pastan. 1994. *Protein Engineering* 7:1509); rapid amplification of cDNA ends (Ruberti, F. et ál. 1994. *J. Immunol. Methods* 173:33); secuencias líderes de anticuerpos (Larrick et ál. 1989 *Biochem. Biophys. Res. Commun.* 160:1250). La clonación de secuencias de anticuerpos se describe adicionalmente

30 en Newman et ál., patente estadounidense n.º 5,658,570, presentada el 25 de enero de 1995, que se incorpora en la presente mediante esta referencia.

[0247] Una región constante de inmunoglobulina utilizada en la presente puede incluir todos los dominios y la región bisagra o partes de estos. En una modalidad, la

35 región constante de inmunoglobulina o una parte de esta comprende un dominio CH2, dominio CH3 y una región bisagra, es decir, una región Fc o un compañero de unión de FcRn.

[0248] Tal como se utiliza en la presente, la expresión «región Fc» se define como la parte de un polipéptido que corresponde a la región Fc del Ig natural, es decir, según se formó mediante la asociación dimérica de los dominios Fc respectivos de sus dos cadenas pesadas. Una región Fc natural forma un homodímero con otra región Fc.

5 Por el contrario, la expresión «región Fc fusionada genéticamente» o «región Fc de cadena simple» (región scFc), tal como se utiliza en la presente, se refiere a una región Fc dimérica sintética que comprende dominios Fc unidos genéticamente dentro de una única cadena polipeptídica (es decir, codificada en una única secuencia genética contigua). Véase la publicación internacional n.º WO
10 2012/006635, que se incorpora en su totalidad mediante esta referencia.

[0249] En una modalidad, la «región Fc» se refiere a la parte de una única cadena pesada de Ig que comienza en la región bisagra justo antes del sitio de escisión de papaína (es decir, residuo 216 en IgG, tomando el primer residuo de la región
15 constante de cadena pesada como 114) y que finaliza en el extremo C del anticuerpo. Por consiguiente, una región Fc completa comprende al menos un dominio bisagra, un dominio CH2 y un dominio CH3.

[0250] Una región constante de inmunoglobulina o una parte de esta puede ser un
20 compañero de unión de FcRn. El FcRn es activo en tejidos epiteliales adultos y se expresa en el lumen de los intestinos, vías respiratorias pulmonares, superficies nasales, superficies vaginales, superficies rectales y del colon (patente estadounidense n.º 6,485,726). Un compañero de unión de FcRn es una parte de una inmunoglobulina que se une a FcRn.

25

[0251] El receptor de FcRn se aisló de varias especies mamíferas incluyendo humanos. Las secuencias del FcRn humano, FcRn de mono, FcRn de rata y FcRn de ratón se conocen (Story et ál. 1994, J. Exp. Med. 180:2377). El receptor FcRn se une a IgG (pero no a otras clases de inmunoglobulina tales como IgA, IgM, IgD e
30 IgE) con un pH relativamente bajo, transporta activamente el IgG de forma transcelular en una dirección luminal a serosal, y luego libera el IgG con un pH relativamente alto en los fluidos intersticiales. Se expresa en tejido epitelial adulto (patentes estadounidenses n.º 6,485,726, 6,030,613, 6,086,875; WO 03/077834; US2003-0235536A1) incluyendo el epitelio intestinal y pulmonar (Israel et ál. 1997,
35 Immunology 92:69) epitelio tubular proximal renal (Kobayashi et ál. 2002, Am. J. Physiol. Renal Physiol. 282:F358), así como epitelio nasal, superficies vaginales y superficies del árbol biliar.

[0252] Los compañeros de unión de FcRn útiles en la presente invención abarcan moléculas que pueden unirse específicamente mediante el receptor de FcRn incluyendo la IgG completa, el fragmento Fc de la IgG y otros fragmentos que incluyen la región de unión completa del receptor de FcRn. La región de la parte Fc de la IgG que se une al receptor de FcRn se describió con base en cristalografía de rayos X (Burmeister et ál. 1994, Nature 372:379). El área de contacto principal del Fc con el FcRn se encuentra cerca de la unión de los dominios CH2 y CH3. Los contactos Fc-FcRn se encuentran todos dentro de una única cadena pesada de Ig. Los compañeros de unión de FcRn incluyen la IgG completa, el fragmento Fc de la IgG y otros fragmentos de la IgG que incluyen la región de unión completa de FcRn. Los sitios de contacto principales incluyen los residuos de aminoácidos 248, 250-257, 272, 285, 288, 290-291, 308-311 y 314 del dominio CH2 y los residuos de aminoácidos 385-387, 428 y 433-436 del dominio CH3. Las referencias a la numeración de aminoácidos de inmunoglobulinas o fragmentos o regiones de inmunoglobulina, se basan en Kabat *et ál.* 1991, Sequences of Proteins of Immunological Interest, Departamento de Salud Pública de EUA, Bethesda, Md.

[0253] Las regiones Fc o compañeros de unión de FcRn unidos a FcRn pueden transportarse eficazmente a lo largo de las barreras epiteliales mediante el FcRn, proporcionando así un medio no invasivo para administrar sistemáticamente una molécula terapéutica deseada. De forma adicional, las proteínas de fusión que comprenden una región Fc o un compañero de unión de FcRn son endocitosadas mediante células que expresan el FcRn. Pero, en vez de marcarse para su degradación, estas proteínas de fusión se reciclan para una nueva circulación, aumentando así la semivida *in vivo* de estas proteínas. En determinadas modalidades, las partes de regiones constantes de Ig son una región Fc o un compañero de unión de FcRn que se asocia típicamente, a través de enlaces disulfuro y otras interacciones no específicas, con otra región Fc u otro compañero de unión de FcRn para formar dímeros y multímeros de mayor orden.

30

[0254] Dos receptores de FcRn pueden unir una única molécula Fc. Los datos cristalográficos sugieren que cada molécula de FcRn une un único polipéptido del homodímero Fc. En una modalidad, la unión del auxiliar de unión de FcRn, por ejemplo, un fragmento Fc de una IgG a una molécula biológicamente activa proporciona un medio para proporcionar la molécula biológicamente activa de forma oral, bucal, sublingual, rectal, vaginal, como un aerosol administrado por vía nasal o a través de una vía pulmonar, o mediante una vía ocular. En otra modalidad, la

35

proteína de FVIII puede administrarse invasivamente, por ejemplo, por vía subcutánea, intravenosa.

[0255] Una región de compañero de unión de FcRn es una molécula o una parte de esta que puede unirse específicamente mediante el receptor de FcRn con un transporte activo consecuente mediante el receptor de FcRn de la región Fc. Unirse específicamente se refiere a dos moléculas que forman un complejo que es relativamente estable en condiciones fisiológicas. La unión específica se caracteriza mediante una afinidad alta y una capacidad de baja a moderada tal como se distinguió a partir de una unión no específica que normalmente tiene una afinidad baja con una capacidad de moderada a alta. Comúnmente, la unión se considera específica cuando la constante de afinidad K_A es mayor a 10^6 M^{-1} , o mayor a 10^8 M^{-1} . En caso de ser necesario, la unión no específica se puede reducir sin afectar sustancialmente la unión específica variando las condiciones de unión. Las condiciones de unión adecuadas, tales como la concentración de las moléculas, la fuerza iónica de la solución, la temperatura, el tiempo permitido para la unión, la concentración de un agente de bloqueo (por ejemplo, albúmina de suero, caseína de la leche), etc. pueden ser optimizados por un experto en la técnica mediante técnicas de rutina.

20

[0256] En determinadas modalidades, una proteína de FVIII codificada por la molécula de ácido nucleico de la descripción comprende una o más regiones Fc truncadas que son, no obstante, suficientes para conferir propiedades de unión del receptor Fc (FcR) a la región Fc. Por ejemplo, la parte de una región Fc que se une a FcRn (es decir, la parte de unión de FcRn) comprende aproximadamente aminoácidos 282-438 de IgG1, numeración EU (donde los sitios de contacto principales son los aminoácidos 248, 250-257, 272, 285, 288, 290-291, 308-311 y 314 del dominio CH2 y residuos de aminoácidos 385-387, 428 y 433-436 del dominio CH3. De este modo, una región Fc de la descripción puede comprender o consistir en una parte de unión de FcRn. Las partes de unión de FcRn pueden derivar de cadenas pesadas de cualquier isotipo, lo que incluye IgG1, IgG2, IgG3 e IgG4. En una modalidad, se utiliza una parte de la unión de FcRn de un anticuerpo de la IgG1 de isotipo humano. En otra modalidad, se utiliza una parte de la unión de FcRn de un anticuerpo de la IgG4 de isotipo humano.

35

[0257] Las regiones Fc se pueden obtener a partir de una cantidad de fuentes diferentes. En una modalidad, una región Fc del polipéptido deriva de una

inmunoglobulina humana. Sin embargo, se entiende que el resto Fc puede derivar de una inmunoglobulina de otra especie de mamífero, lo que incluye, por ejemplo, especies de roedores (por ejemplo, un ratón, una rata, un conejo, un cobayo) o de primates no humano (por ejemplo, chimpancé, macaco). Asimismo, el polipéptido de los dominios Fc o partes de estos pueden derivar de cualquier clase de inmunoglobulina, que incluye IgM, IgG, IgD, IgA e IgE, y cualquier isotipo de inmunoglobulina, que incluye IgG1, IgG2, IgG3 e IgG4. En otra modalidad, se utiliza la IgG1 de isotipo humano.

[0258] En determinadas modalidades, la variante Fc confiere un cambio en al menos una función efectora impartida por un resto Fc que comprende dicho dominio Fc de tipo salvaje (por ejemplo, una mejora o reducción de la capacidad de la región Fc de unirse a los receptores Fc (por ejemplo, FcγRI, FcγRII o FcγRIII) o proteínas de complemento (por ejemplo, C1q), o de desencadenar citotoxicidad dependiente de anticuerpos (ADCC), fagocitosis o citotoxicidad dependiente de complemento (CDCC)). En otras modalidades, la variante Fc proporciona un residuo de cisteína modificado genéticamente.

[0259] La región Fc de la descripción pueden emplear variantes Fc reconocidos en la técnica que se sabe imparten un cambio (por ejemplo, una mejora o reducción) en la función efectora y/o unión de FcR o FcRn. De manera específica, una región Fc de la descripción puede incluir, por ejemplo, un cambio (por ejemplo, una sustitución) en una o más de las posiciones de aminoácidos descritas en las publicaciones internacionales PCT WO88/07089A1, WO96/14339A1, WO98/05787A1, WO98/23289A1, WO99/51642A1, WO99/58572A1, WO00/09560A2, WO00/32767A1, WO00/42072A2, WO02/44215A2, WO02/060919A2, WO03/074569A2, WO04/016750A2, WO04/029207A2, WO04/035752A2, WO04/063351A2, WO04/074455A2, WO04/099249A2, WO05/040217A2, WO04/044859, WO05/070963A1, WO05/077981A2, WO05/092925A2, WO05/123780A2, WO06/019447A1, WO06/047350A2 y WO06/085967A2; publicaciones de patentes estadounidenses n.º US2007/0231329, US2007/0231329, US2007/0237765, US2007/0237766, US2007/0237767, US2007/0243188, US20070248603, US20070286859, US20080057056; o patentes estadounidenses 5,648,260; 5,739,277; 5,834,250; 5,869,046; 6,096,871; 6,121,022; 6,194,551; 6,242,195; 6,277,375; 6,528,624; 6,538,124; 6,737,056; 6,821,505; 6,998,253; 7,083,784; 7,404,956 y 7,317,091, cada una de las cuales se incorporan en la presente mediante esta referencia. En una modalidad, el cambio específico (por

ejemplo, la sustitución específica de uno o más aminoácidos descritos en la técnica) puede realizarse en una o más de las posiciones de aminoácidos descritas. En otra modalidad, se puede hacer un cambio diferente en una o más de las posiciones de aminoácidos descritas (por ejemplo, la sustitución diferente de una o más posiciones de aminoácidos descritas en la técnica).

[0260] La región Fc o compañero de unión de FcRn de la IgG se puede modificar de acuerdo con procedimientos bien reconocidos tal como mutagénesis dirigida al sitio y similares para proporcionar IgG modificadas o fragmentos Fc o partes de estos que se unirán mediante el FcRn. Dichas modificaciones incluyen modificaciones remotas desde los sitios de contacto de FcRn así como modificaciones dentro de los sitios de contacto que conservan o aun potencian la unión al FcRn. Por ejemplo, los siguientes residuos de aminoácidos simples en la IgG1 Fc (Fc γ 1) humana pueden ser sustituidos sin pérdida significativa de afinidad de unión de Fc por FcRn: P238A, S239A, K246A, K248A, D249A, M252A, T256A, E258A, T260A, D265A, S267A, H268A, E269A, D270A, E272A, L274A, N276A, Y278A, D280A, V282A, E283A, H285A, N286A, T289A, K290A, R292A, E293A, E294A, Q295A, Y296F, N297A, S298A, Y300F, R301A, V303A, V305A, T307A, L309A, Q311A, D312A, N315A, K317A, E318A, K320A, K322A, S324A, K326A, A327Q, P329A, A330Q, P331A, E333A, K334A, T335A, S337A, K338A, K340A, Q342A, R344A, E345A, Q347A, R355A, E356A, M358A, T359A, K360A, N361A, Q362A, Y373A, S375A, D376A, A378Q, E380A, E382A, S383A, N384A, Q386A, E388A, N389A, N390A, Y391F, K392A, L398A, S400A, D401A, D413A, K414A, R416A, Q418A, Q419A, N421A, V422A, S424A, E430A, N434A, T437A, Q438A, K439A, S440A, S444A y K447A, donde, por ejemplo, P238A representa la prolina de tipo salvaje sustituida por alanina en la posición número 238. A modo de ejemplo, una modalidad específica incorpora la mutación N297A, eliminando un sitio de glicosilación N altamente conservado. Además de la alanina, se pueden sustituir otros aminoácidos por los aminoácidos de tipo salvaje en las posiciones especificadas anteriormente. Las mutaciones se pueden introducir únicamente en Fc dando lugar a más de cien regiones Fc distintas del Fc natural. De forma adicional, se pueden introducir juntas combinaciones de dos, tres o más de estas mutaciones individuales, dando lugar a más de cientos de regiones Fc.

[0261] Determinadas de las mutaciones anteriores pueden conferir nueva funcionalidad tras la región Fc o el compañero de unión de FcRn. Por ejemplo, una modalidad específica incorpora N297A, eliminando un sitio de glicosilación N

- altamente conservado. El efecto de esta mutación es reducir la inmunogenicidad, potenciando así la semivida circulante de la región Fc, y hacer que la región Fc no pueda unirse a FcγRI, FcγRIIA, FcγRIIB y FcγRIIIA, sin comprometer la afinidad para FcRn (Routledge et ál. 1995, Transplantation 60:847; Friend et ál. 1999, Transplantation 68:1632; Shields et ál. 1995, J. Biol. Chem. 276:6591). Como ejemplo adicional de una nueva funcionalidad que surge de mutaciones descritas anteriormente, la afinidad para FcRn se puede aumentar más allá de las de tipo salvaje en algunos casos. Esta afinidad aumentada puede reflejar un índice de «asociación» aumentado, un índice de «asociación» disminuido o tanto un índice de «asociación» aumentado como un índice de «disociación» disminuido. Los ejemplos de mutaciones que se cree imparten una afinidad aumentada para el FcRn incluyen, de modo no taxativo, T256A, T307A, E380A y N434A (Shields et ál. 2001, J. Biol. Chem. 276:6591).
- 15 **[0262]** De forma adicional, aparecen al menos tres receptores gamma Fc humanos para reconocer un sitio de unión sobre la IgG dentro de la región bisagra inferior, generalmente los aminoácidos 234-237. Por lo tanto, otro ejemplo de nueva funcionalidad e inmunogenicidad disminuida potencial puede surgir de mutaciones de esta región, como por ejemplo reemplazando los aminoácidos 233-236 de la IgG1 humana «ELLG» (SEQ ID NO: 45) con la secuencia correspondiente de «PVA» de la IgG2 (con una eliminación de aminoácidos). Se ha demostrado que FcγRI, FcγRII y FcγRIII, los cuales median varias funciones efectoras no se unirán a la IgG1 cuando se introduzcan tales mutaciones. Ward y Ghetie 1995, Therapeutic Immunology 2:77 y Armour et ál. 1999, Eur. J. Immunol. 29:2613.
- 25 **[0263]** En otra modalidad, la región constante de inmunoglobulina o una parte de esta comprende una secuencia de aminoácidos en la región bisagra o una parte de esta que forma uno o más enlaces disulfuro con una segunda región constante de inmunoglobulina o una parte de esta. La segunda región constante de inmunoglobulina o una parte de esta se puede unir a un segundo polipéptido, uniendo la proteína de FVIII y el segundo polipéptido. En algunas modalidades, el segundo polipéptido es un resto de potenciador. Tal como se utiliza en la presente, el término «resto de potenciador» se refiere a una molécula, fragmento de esta o un componente de un polipéptido que es capaz de mejorar la actividad procoagulante del FVIII. El resto de potenciador puede ser un cofactor, tal como el factor tisular soluble (sTF, por sus siglas en inglés) o un péptido procoagulante. Por lo tanto, tras

la activación del FVIII, el resto de potenciador se encuentra disponible para mejorar la actividad de FVIII.

[0264] En determinadas modalidades, una proteína de FVIII codificada por una molécula de ácido nucleico de la descripción comprende una sustitución de aminoácidos a una región constante de inmunoglobulina o una parte de esta (por ejemplo, variantes Fc), que altera las funciones efectoras independientes de antígenos de la región constante de inmunoglobulina, en particular la semivida en circulación de la proteína.

2. Regiones scFc

[0265] En otro aspecto, un resto heterólogo comprende una región scFc (cadena Fc simple). En una modalidad, una molécula de ácido nucleico aislada de la descripción comprende además una secuencia de ácido nucleico heteróloga que codifica una región scFc. La región scFc comprende al menos dos regiones constantes de inmunoglobulina o porciones de estas (por ejemplo, restos o dominios Fc (por ejemplo, 2, 3, 4, 5, 6 o más restos o dominios Fc)) dentro de la misma cadena polipeptídica lineal que son capaces de plegarse (por ejemplo, plegarse intra o intermolecularmente) para formar una región scFc funcional que se conecta mediante un conector peptídico Fc. Por ejemplo, en una modalidad, un polipéptido de la descripción es capaz de unirse, a través de su región scFc, a al menos un receptor Fc (por ejemplo, un FcRn, un receptor FcγR (por ejemplo, FcγRIII) o una proteína de complemento (por ejemplo, C1q)) a efectos de mejorar la semivida o desencadenar una función efectora inmunitaria (por ejemplo, citotoxicidad dependiente de anticuerpo (ADCC), fagocitosis o citotoxicidad dependiente de complemento (CDCC) y/o para mejorar la factibilidad de fabricación).

3. CTP

[0266] En otro aspecto, un resto heterólogo comprende un péptido de extremo C (CTP) de la subunidad β de la gonadotropina coriónica humana o fragmento, variante o derivado de esta. Se sabe que uno o más péptidos CTP insertados en una proteína recombinante aumenta la semivida *in vivo* de esa proteína. Véase, por ejemplo, la patente estadounidense n.º 5,712,122, incorporada mediante esta referencia en su totalidad.

[0267] Los ejemplos de péptidos CTP incluyen DPRFQDSSSSKAPPPSLPSRLPGPSDTPIL (SEQ ID NO: 33) o

SSSSKAPPSLPSPSRLPGPSDTPILPQ (SEQ ID NO: 34). Véase, por ejemplo, la publicación de solicitud de patente estadounidense n.º US 2009/0087411 A1, incorporada mediante esta referencia.

5 **4. Secuencia XTEN**

[0268] En algunas modalidades, un resto heterólogo comprende una o más secuencias XTEN, fragmentos, variantes o derivados de este. Tal como se utiliza en la presente, «secuencia XTEN» se refiere a polipéptidos de longitud prolongada con secuencias sustancialmente no repetitivas y de origen natural que están compuestas principalmente de aminoácidos hidrofílicos pequeños, con la secuencia que tiene una estructura de bajo grado o no secundaria o terciaria en condiciones fisiológicas. Como un resto heterólogo, las XTEN pueden servir como un resto de extensión de semivida. Además, XTEN puede proporcionar propiedades deseables que incluyen, de modo no taxativo, características de solubilidad y parámetros farmacocinéticos
10 potenciados.
15

[0269] La incorporación de un resto heterólogo que comprende una secuencia XTEN en una proteína de la descripción puede conferir a la proteína una o más de las siguientes propiedades ventajosas: flexibilidad conformacional, solubilidad acuosa potenciada, alto grado de resistencia a proteasa, baja inmunogenicidad, baja unión a
20 receptores mamíferos o radio hidrodinámico aumentado (o de Stokes).

[0270] En determinados aspectos, una secuencia XTEN puede aumentar las propiedades farmacocinéticas tales como semivida *in vivo* más prolongada o mayor
25 área bajo la curva (AUC, por sus siglas en inglés), de manera que una proteína de la descripción permanece *in vivo* y tiene actividad procoagulante durante un mayor período de tiempo en comparación con una proteína con el mismo el resto heterólogo de XTEN pero sin este.

[0271] En algunas modalidades, la secuencia XTEN útil de la descripción es un péptido o un polipéptido que tiene más de aproximadamente 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1200, 1400, 1600, 1800 o 2000 residuos de aminoácidos. En determinadas modalidades, XTEN es un péptido o un polipéptido que tiene más de
30 aproximadamente 20 a aproximadamente 3000 residuos de aminoácidos, más de 30 a aproximadamente 2500 residuos, más de 40 a aproximadamente 2000 residuos, más de 50 a aproximadamente 1500 residuos, más de 60 a aproximadamente 1000

residuos, más de 70 a aproximadamente 900 residuos, más de 80 a aproximadamente 800 residuos, más de 90 a aproximadamente 700 residuos, más de 100 a aproximadamente 600 residuos, más de 110 a aproximadamente 500 residuos o más de 120 a aproximadamente 400 residuos. En una modalidad particular, el XTEN comprende una secuencia de aminoácidos de más de 42 aminoácidos y menos de 144 aminoácidos de longitud.

[0272] La secuencia XTEN de la descripción puede comprender uno o más motivos de secuencias de 5 a 14 (por ejemplo, 9 a 14) residuos de aminoácidos o una secuencia de aminoácidos al menos 80 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntica al motivo de secuencias, donde el motivo comprende, consiste esencialmente o consiste en 4 a 6 tipos de aminoácidos (por ejemplo, 5 aminoácidos) que se seleccionan del grupo que consiste en glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P). Véase US 2010-0239554 A1.

15

[0273] En algunas modalidades, el XTEN comprende motivos de secuencias que no se superponen donde aproximadamente 80 %, o al menos aproximadamente 85 %, o al menos aproximadamente 90 %, o aproximadamente 91 %, o aproximadamente 92 %, o aproximadamente 93 %, o aproximadamente 94 %, o aproximadamente 95 %, o aproximadamente 96 %, o aproximadamente 97 %, o aproximadamente 98 %, o aproximadamente 99 % o aproximadamente 100 % de la secuencia consiste en múltiples unidades de secuencias que no se superponen que se seleccionan de una única familia de motivos seleccionada de la Tabla 4, lo que produce una secuencia de familias. Tal como se usa en la presente, «familia» significa que el XTEN tiene motivos seleccionados solo de una única categoría de motivos de la TABLA 4; es decir., AD, AE, AF, AG, AM, AQ, BC o BD XTEN, y que cualquier otro aminoácido en el XTEN que no pertenece a la familia de motivos se selecciona para lograr una propiedad necesaria, tal como permitir la incorporación de un sitio de restricción por medio de nucleótidos codificantes, incorporación de una secuencia de escisión o para lograr una mejor conexión con FVIII. En algunas modalidades de familias XTEN, una secuencia XTEN comprende varias unidades de motivos de secuencias que no se superponen de la familia de motivos AD, o de la familia de motivos AE, o de la familia de motivos AF, o de la familia de motivos AG, o de la familia de motivos AM, o de la familia de motivos AQ, o de la familia BC, o de la familia BD, con el XTEN resultante que presenta el intervalo de homología descrito en la presente. En otras modalidades, el XTEN comprende varias unidades de las secuencias de motivos de dos o más de las familias de motivos de la Tabla 2A. Estas secuencias

30

35

pueden seleccionarse para lograr las características físicas/químicas deseadas, incluyendo dichas propiedades como carga neta, hidrofili­cidad, falta de estructura secundaria o falta de repetitividad que se confieren mediante la composición de aminoácidos de los motivos, descritos en más detalle a continuación. En las

5 modalidades descritas anteriormente en este párrafo, los motivos incorporados en el XTEN pueden seleccionarse y ensamblarse utilizando los métodos descritos en la presente para lograr un XTEN de aproximadamente 36 a aproximadamente 3000 de residuos de aminoácidos.

10 **Tabla 4. Motivos de secuencias de XTEN de 12 aminoácidos y familias de motivos**

Familia de motivos*	SECUENCIA DE MOTIVOS	SEQ ID NO:
AD	GESPGGSSGSES	73
AD	GSEGSSGPGESS	74
AD	GSSEGSSEGGP	75
AD	GSGGEPSESGSS	76
AE, AM	GSPAGSPTSTEE	77
AE, AM, AQ	GSEPATSGSETP	78
AE, AM, AQ	GTSESATPESGP	79
AE, AM, AQ	GTSTEPSEGSAP	80
AF, AM	GSTSESPSGTAP	81
AF, AM	GTSTPESGSASP	82
AF, AM	GTSPSGESSTAP	83
AF, AM	GSTSSTAESPGP	84
AG, AM	GTPGSGTASSSP	85
AG, AM	GSSTPSGATGSP	86
AG, AM	GSSPSASTGTGP	87
AG, AM	GASPGTSSTGSP	88
AQ	GEPAGSPTSTSE	89
AQ	GTGEPSSTPASE	90
AQ	GSGPSTESAPTE	91
AQ	GSETPSGPSETA	92
AQ	GPSETSTSEPGA	93
AQ	GSPSEPTEGTSA	94
AQ	GSPSEPTEGTSA	94

Familia de motivos*	SECUENCIA DE MOTIVOS	SEQ ID NO:
BC	GSGASEPTSTEP	95
BC	GSEPATSGTEPS	96
BC	GTSEPSTSEPGA	97
BC	GTSTEPSEPGSA	98
BD	GSTAGSETSTEА	99
BD	GSETATSGSETA	100
BD	GTSESATSESGA	101
BD	GTSTEASEGSAS	102

* Denota secuencias de motivos individuales que, cuando se utilizan juntos en varias permutaciones, ocasionan una «secuencia de familias».

5 **[0274]** Los ejemplos de secuencias XTEN que se pueden utilizar como restos heterólogos en proteínas quiméricas de la descripción se describen en la publicación de patente estadounidense n.º 2010/0239554 A1, 2010 /0323956 A1, 2011/0046060 A1, 2011/0046061 A1, 2011/0077199 A1, o 2011/0172146 A1, o la publicación de patente Internacional n.º WO 2010091122 A1, WO 2010 144502 A2, WO
10 2010144508 A1, WO 2011028228 A1, WO 2011028229 A1 o WO 2011028344 A2, cada una de las cuales se incorpora mediante referencia en su totalidad.

[0275] XTEN puede tener diversas longitudes para su inserción en el FVIII o acoplamiento a este. En una modalidad, la longitud de la o las secuencias XTEN se
15 escoge con base en la propiedad o función que se alcanzará en la proteína de fusión. Dependiendo de la función o propiedad pretendida, el XTEN puede tener una secuencia de longitud corta o intermedia o una secuencia más larga que puede ser útil como portador. En determinadas modalidades, el XTEN incluye segmentos cortos de aproximadamente 6 a aproximadamente 99 residuos de aminoácidos,
20 longitudes intermedias de aproximadamente 100 a aproximadamente 399 residuos de aminoácidos, y longitudes más largas de aproximadamente 400 a aproximadamente 1000 y hasta aproximadamente 3000 residuos de aminoácidos. Por lo tanto, el XTEN insertado en el FVIII o unido a este puede tener longitudes de
25 40, aproximadamente 42, aproximadamente 72, aproximadamente 96, aproximadamente 144, aproximadamente 288, aproximadamente 400, aproximadamente 500, aproximadamente 576, aproximadamente 600,

aproximadamente 700, aproximadamente 800, aproximadamente 864, aproximadamente 900, aproximadamente 1000, aproximadamente 1500, aproximadamente 2000, aproximadamente 2500, o hasta aproximadamente 3000 residuos de aminoácidos de longitud. En otras modalidades, las secuencias XTEN

5 tienen de aproximadamente 6 to aproximadamente 50, aproximadamente 50 a aproximadamente 100, aproximadamente 100 a 150, aproximadamente 150 a 250, aproximadamente 250 a 400, aproximadamente 400 a aproximadamente 500, aproximadamente 500 a aproximadamente 900, aproximadamente 900 a 1500, aproximadamente 1500 a 2000, o aproximadamente 2000 a aproximadamente 3000

10 residuos de aminoácidos de longitud. La longitud exacta de un XTEN insertado en el FVIII o unido a este puede variar sin afectar negativamente la actividad del FVIII. En una modalidad, uno o más de los XTEN utilizados en la presente tienen 42 aminoácidos, 72 aminoácidos, 144 aminoácidos, 288 aminoácidos, 576 aminoácidos o 864 aminoácidos de longitud y se pueden seleccionar de una o más de las

15 secuencias de familias XTEN; es decir, AD, AE, AF, AG, AM, AQ, BC o BD.

[0276] En algunas modalidades, la secuencia XTEN utilizada en la descripción es al menos 60 %, 70 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % idéntica a una secuencia que se selecciona del grupo que

20 consiste en AE42, AG42, AE48, AM48, AE72, AG72, AE108, AG108, AE144, AF144, AG144, AE180, AG180, AE216, AG216, AE252, AG252, AE288, AG288, AE324, AG324, AE360, AG360, AE396, AG396, AE432, AG432, AE468, AG468, AE504, AG504, AF504, AE540, AG540, AF540, AD576, AE576, AF576, AG576, AE612, AG612, AE624, AE648, AG648, AG684, AE720, AG720, AE756, AG756, AE792,

25 AG792, AE828, AG828, AD836, AE864, AF864, AG864, AM875, AE912, AM923, AM1318, BC864, BD864, AE948, AE1044, AE1140, AE1236, AE1332, AE1428, AE1524, AE1620, AE1716, AE1812, AE1908, AE2004A, AG948, AG1044, AG1140, AG1236, AG1332, AG1428, AG1524, AG1620, AG1716, AG1812, AG1908, AG2004 y cualquier combinación de estos. Véase US 2010-0239554 A1. En una modalidad

30 particular, el XTEN comprende AE42, AE72, AE144, AE288, AE576, AE864, AG 42, AG72, AG144, AG288, AG576, AG864 o cualquier combinación de estos.

[0277] Las secuencias XTEN de ejemplo se pueden utilizar como restos heterólogos en la proteína quimérica de la descripción incluye XTEN AE42-4 (SEQ ID NO: 46, codificado por la SEQ ID NO: 47; figuras 11C y 11D, respectivamente), XTEN 144-2A (SEQ ID NO: 48, codificado por la SEQ ID NO: 49; figuras 11E y 11F, respectivamente), XTEN A144-3B (SEQ ID NO: 50, codificado por la SEQ ID NO: 51;

35

figuras 11G y 11H, respectivamente), XTEN AE144-4A (SEQ ID NO: 52, codificado por la SEQ ID NO: 53; figuras 11I y 11J, respectivamente), XTEN AE144-5A (SEQ ID NO: 54, codificado por la SEQ ID NO: 55; figuras 11K y 11L, respectivamente), XTEN AE144-6B (SEQ ID NO: 56, codificado por la SEQ ID NO: 57; figuras 11M y 11N, respectivamente), XTEN AG144-1 (SEQ ID NO: 58, codificado por la SEQ ID NO: 59; figuras 11O y 11P, respectivamente), XTEN AG144-A (SEQ ID NO: 60, codificado por la SEQ ID NO: 61; figuras 11Q y 11R, respectivamente), XTEN AG144-B (SEQ ID NO: 62, codificado por la SEQ ID NO: 63; figuras 11S y 11T, respectivamente), XTEN AG144-C (SEQ ID NO: 64, codificado por la SEQ ID NO: 65; figuras 11U y 11V, respectivamente), y XTEN AG144-F (SEQ ID NO: 66, codificado por la SEQ ID NO: 67; figuras 11W y 11X, respectivamente). En una modalidad particular, el XTEN es codificado por la SEQ ID NO: 18.

[0278] En algunas modalidades, menos del 100 % de aminoácidos de un XTEN se seleccionan de la glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) o menos del 100 % de la secuencia consiste en los motivos de secuencia de la Tabla 2A o una secuencia XTEN proporcionada en la presente. En dichas modalidades, los residuos de aminoácidos restantes del XTEN se seleccionan de cualquiera de los otros 14 L-aminoácidos naturales, pero se pueden seleccionar preferentemente de aminoácidos hidrofílicos de manera que la secuencia XTEN contiene al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o al menos aproximadamente 99 % de aminoácidos hidrofílicos. El contenido de aminoácidos hidrofóbicos en el XTEN utilizado en las construcciones de conjugación puede ser inferior al 5 %, o inferior al 2 % o inferior al 1 % del contenido de aminoácidos hidrofóbicos. Los residuos hidrofóbicos menos favorecidos en la construcción de XTEN incluyen triptófano, fenilalanina, tirosina, leucina, isoleucina, valina y metionina. Además, las secuencias XTEN pueden contener menos de 5 % o menos de 4 % o menos de 3 % o menos de 2 % o menos de 1 % o ninguno de los siguientes aminoácidos: metionina (por ejemplo, para evitar la oxidación), o asparagina y glutamina (para evitar la desamidación).

[0279] Las una o más secuencias de XTEN se pueden insertar en el extremo N o en el extremo C de la secuencia de aminoácidos codificada por la secuencia de nucleótidos o está insertada entre dos aminoácidos en la secuencia de aminoácidos codificada por la secuencia de nucleótidos. Por ejemplo, el XTEN puede estar insertada entre dos aminoácidos en uno o más sitios de inserción que se seleccionan de la Tabla 3. Los ejemplos de sitios dentro del FVIII que son

permisibles para la inserción de XTEN se pueden encontrar en, por ejemplo, la publicación internacional n.º 2013/123457 A1 o la publicación estadounidense n.º 2015/0158929 A1, las cuales se incorporan mediante referencia en su totalidad.

5. Albúmina o fragmento, derivado o variante de esta

[0280] En algunas modalidades, el resto heterólogo comprende albúmina o un fragmento funcional de esta. La seroalbúmina humana (HSA o HA), una proteína de 609 aminoácidos en su forma de longitud completa, es responsable de una proporción considerable de la presión osmótica del suero y también funciona como portadora de ligandos endógenos y exógenos. El término «albúmina», tal como se usa en la presente, incluyen albúmina de longitud completa o un fragmento, variante, derivado o análogo funcional de esta. Los ejemplos de albúmina o los fragmentos o variantes de esta se describen en la publicación de patente estadounidense n.º 2008/0194481A1, 2008/0004206 A1, 2008/0161243 A1, 2008/0261877 A1, o 2008/0153751 A1 o la publicación de solicitud PCT n.º 2008/033413 A2, 2009/058322 A1, o 2007/021494 A2, que se incorporan en la presente mediante esta referencia en su totalidad.

[0281] En una modalidad, la proteína de FVIII codificada por una molécula de ácido nucleico de la descripción comprende albúmina, un fragmento o una variante de esta que se une adicionalmente a un segundo resto heterólogo que se selecciona del grupo que consiste en una región constante de inmunoglobulina o parte de esta (por ejemplo, una región Fc), una secuencia PAS, HES y PEG.

6. Restos de unión a la albúmina

[0282] En determinadas modalidades, el resto heterólogo es un resto de unión a la albúmina, que comprende un péptido de unión a la albúmina, un dominio de unión a la albúmina bacteriana, un fragmento de anticuerpo de unión a la albúmina o cualquier combinación de estos.

[0283] Por ejemplo, la proteína de unión a la albúmina puede ser una proteína de unión a la albúmina bacteriana, un anticuerpo o un fragmento de anticuerpo que incluye anticuerpos de dominio (véase la patente estadounidense n.º 6,696,245). Una proteína de unión a la albúmina, por ejemplo, puede ser un dominio de unión a la albúmina bacteriana, tal como la proteína G estreptocócica (Konig, T. y Skerra, A. (1998) *J. Immunol. Methods* 218, 73-83). Otros ejemplos de péptidos de unión a la albúmina que se pueden utilizar como compañero de conjugación son, por ejemplo,

aquellos que tiene un secuencia de consenso Cys-Xaa₁-Xaa₂-Xaa₃-Xaa₄-Cys, donde Xaa₁ es Asp, Asn, Ser, Thr, o Trp; Xaa₂ es Asn, Gln, H es, Ile, Leu, o Lys; Xaa₃ es Ala, Asp, Phe, Trp, o Tyr; y Xaa₄ es Asp, Gly, Leu, Phe, Ser, o Thr como se describe en la solicitud de patente estadounidense 2003/0069395 o Dennis et ál.

5 (Dennis et ál. 2002) J. Biol. Chem. 277, 35035-35043).

[0284]El dominio 3 de la proteína G de estreptococo, como se describe en Kraulis et ál., FEBS Lett. 378:190-194 (1996) y Linhult et ál., Protein Sci. 11:206-213 (2002) es un ejemplo de un dominio de unión a albúmina bacteriana. Ejemplos de péptidos de

10 unión a albúmina incluyen una serie de péptidos que tienen una secuencia núcleo DICLPRWGCLW (SEQ ID NO: 35). Véase, por ejemplo, Dennis et ál., J. Biol. Chem. 2002, 277: 35035-35043 (2002)). Ejemplos de fragmentos de anticuerpo de unión a albúmina se describen en Muller y Kontermann, Curr. Opin. Mol. Ther. 9:319-326 (2007); Roovers et ál., Cancer Immunol. Immunother. 56:303-317 (2007) y Holt et ál.,

15 Prot. Eng. Design Sci., 21:283-288 (2008), que se incorporan a la presente en su totalidad mediante esta referencia. Un ejemplo de dicho resto de unión a la albúmina es hexanoato de 2-(3-maleimidopropanamido)-6-(4-(4-yodofenil)butanamido) (etiqueta «Albu») tal como se divulga en Trussel et ál., Bioconjugate Chem. 20:2286-2292 (2009).

20 **[0285]**Los ácidos grasos, en particular, los ácidos grasos de cadena larga (LCFA, por sus siglas en inglés) particular y compuestos de unión a la albúmina similares a ácidos grasos de cadena larga pueden utilizarse para extender la semivida *in vivo* de proteínas de FVIII de la descripción. Un ejemplo de un compuesto de unión a la

25 albúmina similar a LCFA ácido 16-(1-(3-(9-(((2,5-dioxopirrolidin-1-iloxi) carboniloxi)-metil)-7-sulfo-9H-fluoren-2-ilamino)-3-oxopropil)-2,5-dioxopirrolidin-3-iltio) hexadecanoico (véase, por ejemplo, WO 2010/140148).

7. Secuencia PAS

30 **[0286]**En otras modalidades, el resto heterólogo es una secuencia PAS. Una secuencia PAS, tal como se usa en la presente, significa una secuencia de aminoácidos que comprende principalmente residuos de alanina y serina o que comprende principalmente residuos de alanina, serina y prolina, formando la secuencia de aminoácidos una conformación de bobina aleatoria en condiciones

35 fisiológicas. Por lo tanto, la secuencia PAS es una unidad estructural, un polímero de aminoácido, o un casete de secuencia que comprende, consiste esencialmente en, o consiste en alanina, serina y prolina que puede usarse como parte del resto

heterólogo en la proteína quimérica. Sin embargo, el experto sabrá que un polímero de aminoácidos también puede formar una conformación de bobina aleatoria cuando se agregan residuos distintos de alanina, serina y prolina como constituyentes minoritarios en la secuencia PAS. La expresión «constituyente minoritario» tal como se usa en la presente, significa que se pueden agregar aminoácidos distintos de alanina, serina y prolina en la secuencia PAS hasta un cierto grado, por ejemplo, hasta aproximadamente 12 %, es decir, aproximadamente 12 a 100 aminoácidos de la secuencia PAS, hasta aproximadamente 10 %, es decir, aproximadamente 10 a 100 aminoácidos de la secuencia PAS, hasta aproximadamente 9 %, es decir, aproximadamente 9 a 100 aminoácidos, hasta aproximadamente 8 %, es decir, aproximadamente 8 de 100 aminoácidos, aproximadamente 6 %, es decir, aproximadamente 6 de 100 aminoácidos, aproximadamente 5 %, es decir, aproximadamente 5 de 100 aminoácidos, aproximadamente 4 %, es decir, aproximadamente 4 de 100 aminoácidos, aproximadamente 3 %, es decir, aproximadamente 3 de 100 aminoácidos, aproximadamente 2 %, es decir, aproximadamente 2 de 100 aminoácidos, aproximadamente 1 %, es decir, aproximadamente 1 de 100 de los aminoácidos. Los aminoácidos diferentes de alanina, serina y prolina pueden seleccionarse del grupo que consiste en Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Thr, Trp, Tyr, y Val.

[0287] En condiciones fisiológicas, la secuencia PAS al estirarse forma una conformación de bobina aleatoria y puede mediar así una estabilidad *in vivo* y/o *in vitro* aumentada para la proteína de FVIII. Dado que el dominio de bobina aleatorio no adopta una estructura o función estable por sí mismo, la actividad biológica mediada por la proteína de FVIII mediada se conserva esencialmente. En otras realizaciones, las secuencias PAS que forman un dominio de bobina aleatorio son biológicamente inertes, especialmente con respecto a la proteólisis en el plasma sanguíneo, inmunogenicidad, punto isoeléctrico/comportamiento electrostático, unión a receptores de superficie celular o internalización, pero siendo aún biodegradables, lo cual proporciona ventajas claras con respecto a polímeros sintéticos tales como PEG.

[0288] Los ejemplos no taxativos de las secuencias PAS que forman la conformación de bobina aleatoria comprenden una secuencia de aminoácidos que se selecciona del grupo que consiste en ASPAAPAPASPAAPAPSAPA (SEQ ID NO: 36), AAPASPAPAAPSAPAPAAPS (SEQ ID NO: 37), APSSPSPSAPSSPSPASPSS (SEQ ID NO: 38), APSSPSPSAPSSPSPASPS (SEQ ID NO: 39),

SSPSAPSPSSPASPSPPSPA (SEQ ID NO: 40), AASPAAPSAPPAAASPAAPSAPPA (SEQ ID NO: 41) y ASAAAPAAASAAASAPSAAA (SEQ ID NO: 42) o cualquier combinación de estos. Los ejemplos adicionales de las secuencias PAS se conocen, por ejemplo, de publicación de patente PCT n.º 2010/0292130 A1 y solicitud PCT n.º WO 2008/155134 A1.

8. Secuencia HAP

[0289] En determinadas modalidades, el resto heterólogo es un polímero de homoaminoácidos rico en glicina (HAP, por sus siglas en inglés). La secuencia HAP puede comprender una secuencia repetitiva de glicina, que tiene al menos 50 aminoácidos, al menos 100 aminoácidos, 120 aminoácidos, 140 aminoácidos, 160 aminoácidos, 180 aminoácidos, 200 aminoácidos, 250 aminoácidos, 300 aminoácidos, 350 aminoácidos, 400 aminoácidos, 450 aminoácidos o 500 aminoácidos de longitud. En una modalidad, la secuencia HAP es capaz de extender la semivida de un resto fusionado o conectado con la secuencia HAP. Los ejemplos no taxativos de la secuencia HAP incluyen, a modo no taxativo $(\text{Gly})_n$, $(\text{Gly}_4\text{Ser})_n$ o $\text{S}(\text{Gly}_4\text{Ser})_n$, donde n es 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20. En una modalidad, n es 20, 21, 22, 23, 24, 25, 26, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40. En otra modalidad, n es 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190 o 200.

9. Transferrina o fragmento de esta

[0290] En determinadas modalidades, el resto heterólogo es transferrina o un fragmento de esta. Se puede utilizar cualquier transferrina para producir las proteínas de FVIII de la descripción. A modo de ejemplo, un TF humano natural (TF) es una proteína de 679 aminoácidos, de aproximadamente 75 KDa (sin dar cuenta de la glicosilación), con dos dominios principales, N (aproximadamente 330 aminoácidos) y C (aproximadamente 340 aminoácidos), que parecen originarse de una duplicación de genes. Véanse los números de acceso GenBank NM001063, XM002793, M12530, XM039845, XM 039847 y S95936 (www.ncbi.nlm.nih.gov), todos los cuales se incorporan a la presente en su totalidad mediante esta referencia. La transferrina comprende dos dominios, el dominio N y el dominio C. El dominio N comprende dos subdominios, el dominio N1 y el dominio N2, y el dominio C comprende dos subdominios, el dominio C1 y el dominio C2.

[0291] En una modalidad, el resto heterólogo de transferrina incluye una variante de empalme de transferrina. En un ejemplo, una variante de empalme de transferrina

puede ser una variante de empalme de transferrina humana, por ejemplo, acceso de GenBank AAA61140. En otra modalidad, la parte de transferrina de la proteína quimérica incluye uno o más dominios de la secuencia de transferrina, por ejemplo, dominio N, dominio C, dominio N1, dominio N2, dominio C1, dominio C2 o cualquier
5 combinación de estos.

10. Receptores de depuración

[0292] En determinadas modalidades, el resto heterólogo es un receptor de depuración, fragmento, variante o derivado de este. El LRP1 es una proteína de
10 membrana integral de 600 kDa que participa en la depuración mediada por receptores de una variedad de proteínas, tales como factor X. Véase, por ejemplo, Narita et ál., Blood 91:555-560 (1998).

11. Factor de von Willebrand o fragmentos de este

15 [0293] En determinadas modalidades, el resto heterólogo es el factor de von Willebrand (VWF) o uno o más fragmentos de este.

[0294] VWF (también conocido como F8VWF) es una gran glucoproteína multimérica presente en el plasma sanguíneo y que se produce constitutivamente en el endotelio
20 (en los cuerpos de Weibel-Palade), megacariocitos (gránulos α de plaquetas) y tejido conectivo subendotelial. El monómero del VWF básico es una proteína del aminoácido 2813. Cada monómero contiene una cantidad de dominios específicos con una función específica, los dominios D' y D3 (que se une al factor VIII), el dominio A1 (que se une al receptor de GPIb de plaquetas, heparina y/o
25 posiblemente colágeno), el dominio A3 (que se une al colágeno), el dominio C1 (donde el dominio RGD se une a la integrina de plaquetas α IIb β 3 cuando esta está activada), y el dominio «nudo de cisteína» en el extremo C de la proteína (donde VWF comparte el factor de crecimiento derivado de plaquetas (PDGF, por sus siglas en inglés), el factor de crecimiento de transformante β (TGF β) y la gonadotropina
30 coriónica humana β (β HCG)).

[0295] La secuencia de aminoácidos del monómero 2813 para el VWF humano se informa como el número de acceso NP000543.2 en Genbank. La secuencia de nucleótidos que codifica el VWF humano se informa como el número de acceso
35 NM000552.3 en Genbank. SEQ ID NO: 44 (figura 11B) es la secuencia de aminoácidos codificada por la SEQ ID NO: 43. El dominio D' incluye aminoácidos

764 a 866 de la SEQ ID NO: 44. El dominio D3 incluye aminoácidos 867 a 1240 de la SEQ ID NO: 44.

5 [0296] En el plasma, 95-98 % del FVIII circula en un complejo no covalente ajustado con VWF de longitud completa. La información de este complejo es importante para el mantenimiento de niveles plasmáticos adecuados de FVIII *in vivo*. Lenting *et ál.*, *Blood*. 92(11): 3983-96 (1998); Lenting *et ál.*, *J. Thromb. Haemost.* 5(7): 1353-60 (2007)). Cuando el FVIII se activa debido a proteólisis en las posiciones 372 y 740 en la cadena pesada y en la posición 1689 en la cadena ligera, se elimina el VWF
10 unido a FVIII del FVIII activado.

[0297] En determinadas modalidades, el resto heterólogo es factor de von Willebrand de longitud completa. En otras modalidades, el resto heterólogo es un fragmento del factor de von Willebrand. La expresión «proteína del VWF» o «fragmentos del
15 VWF», tal como se utiliza en la presente, significa cualquier fragmento del VWF que interactúa con el FVIII y conserva al menos una o más propiedades que se proporcionan normalmente al FVIII mediante un VWF de longitud completa, por ejemplo, evitando una activación prematura al FVIIIa, evitando una proteólisis prematura, evitando una asociación con membranas de fosfolípidos que podrían
20 conducir a una depuración prematura, evitando una unión a receptores de depuración del FVIII que pueden unirse a FVIII puros pero no a FVIII unidos a VWF, y/o estabilizando las interacciones de cadena ligera y cadena pesada del FVIII. En una modalidad específica, el resto heterólogo es un fragmento (VWF) que comprende un dominio D' y un dominio D3 del VWF. El fragmento del VWF que
25 comprende el dominio D' y el dominio D3 puede comprender además un dominio del VWF que se selecciona del grupo que consiste en un dominio A1, un dominio A2, un dominio A3, un dominio D1, un dominio D2, un dominio D4, un dominio B1, un dominio B2, un dominio B3, un dominio C1, un dominio C2, un dominio CK, uno o más fragmentos de estos, y cualquier combinación de estos. Los ejemplos
30 adicionales del polipéptido que tiene actividad de FVIII fusionado con el fragmento del VWF se describen en la solicitud de patente provisional estadounidense n.º 61/667,901, presentada el 3 de julio de 2012, y la publicación estadounidense n.º 2015/0023959 A1, las cuales se incorporan a la presente mediante referencia en su totalidad.

35

12. Restos de enlazador

[0298] En determinadas modalidades, el resto heterólogo es un péptido enlazador.

[0299] Tal como se utiliza en la presente, la expresión «enlazadores polipeptídicos» o «restos de enlazador» se refiere a una secuencia de péptidos o polipéptidos (por ejemplo, una secuencia sintética de péptidos o polipéptidos) que conecta dos dominios en una secuencia de aminoácidos lineal de una cadena de polipéptidos.

5

[0300] En algunas modalidades, las secuencias de nucleótidos heterólogas que codifican los enlazadores peptídicos se pueden insertar entre las secuencias de polinucleótidos del FVIII optimizadas de la descripción y una secuencia de nucleótidos heteróloga que codifica, por ejemplo, uno de los restos heterólogos descritos anteriormente, tal como la albúmina. Los enlazadores peptídicos pueden proporcionar flexibilidad a la molécula de polipéptidos quiméricos. Los enlazadores no son escindibles normalmente, sin embargo, puede que sea deseable tal escisión. En una modalidad, estos enlazadores no se quitan durante el procesamiento.

[0301] Un tipo de enlazador que puede estar presente en una proteína quimérica de la descripción es un enlazador escindible de proteasa, que comprende un sitio de escisión (es decir, un sustrato del sitio de escisión de proteasa, un factor Xla, Xa, o un sitio de escisión de trombina) y este puede incluir enlazadores adicionales en cualquiera de los extremos N o C o en ambos lados del sitio de escisión. Estos enlazadores escindibles, cuando se incorporan en una construcción de la descripción, dan como resultado una molécula quimérica que tiene un sitio de escisión heterólogo.

[0302] En una modalidad, un polipéptido del FVIII codificado por una molécula de ácido nucleico de la presente descripción comprende dos o más dominios Fc o restos unidos por medio de un enlazador cscFc para formar una región Fc comprendida en una cadena de polipéptidos simple. El enlazador cscFc se encuentra flanqueado por al menos un sitio de procesamiento intracelular, es decir, un sitio escindido por una enzima intracelular. La escisión del polipéptido en al menos un sitio de procesamiento intracelular da como resultado un polipéptido que comprende al menos dos cadenas de polipéptidos.

[0303] Otros enlazadores peptídicos se pueden usar opcionalmente en una construcción de la descripción, por ejemplo, para conectar una proteína de FVIII a una región Fc. Algunos enlazadores de ejemplo que pueden utilizarse en conexión con la descripción incluyen, por ejemplo, polipéptidos que comprenden los aminoácidos GlySer descritos más detalladamente a continuación.

35

[0304] En una modalidad, el enlazador peptídico es sintético, es decir, de origen no natural. En una modalidad, un enlazador peptídico incluye péptidos (o polipéptidos) (que pueden ser de origen natural o no) que comprenden una secuencia de aminoácidos que une o fusiona genéticamente una primera secuencia lineal de aminoácidos a una segunda secuencia lineal de aminoácidos a la cual no se encuentra unida de manera natural o fusionada genéticamente en la naturaleza. Por ejemplo, en una modalidad el enlazador peptídico puede comprender polipéptidos de origen no natural que son formas modificadas de polipéptidos de origen natural (por ejemplo, que comprenden una mutación, tal como una adición, sustitución o eliminación). En otra modalidad, el enlazador peptídico puede comprender aminoácidos de origen no natural. En otra modalidad, el enlazador peptídico puede comprender aminoácidos de origen natural que se producen en una secuencia lineal que no existe en la naturaleza. En incluso otra modalidad, el enlazador peptídico puede comprender una secuencia polipeptídica de origen natural.

[0305] Por ejemplo, en determinadas modalidades, se puede usar un enlazador peptídico para fusionar restos Fc idénticos, lo que de esta manera forma una región scFc homomérica. En otras modalidades, puede utilizarse un enlazador peptídico para fusionar diferentes restos Fc (por ejemplo, un resto Fc de tipo salvaje y una variante del resto Fc), lo que forma una región scFc heteromérica.

[0306] En otra modalidad, un enlazador peptídico comprende o consiste en un enlazador gly/ser. En una modalidad, un enlace scFc o cscFc comprende al menos una parte de una bisagra de inmunoglobulina y un enlace gly-ser. Tal como se usa en la presente, el término «enlace gly-ser» se refiere a un péptido que consiste en residuos de glicina y serina. En determinadas modalidades, dicho enlazador gly-ser se puede insertar entre dos secuencias adicionales del enlazador peptídico. En otras modalidades, un enlazador gly-ser está unido en uno o ambos extremos de otra secuencia del enlazador peptídico. En incluso otras modalidades, dos o más enlaces gly-ser se incorporan en serie en un enlazador peptídico. En una modalidad, un enlazador peptídico de la descripción comprende al menos una parte de una región bisagra superior (por ejemplo, derivado de una molécula IgG1, IgG2, IgG3, o IgG4), al menos una parte de una región bisagra media (por ejemplo, derivada de una molécula de IgG1, IgG2, IgG3, o IgG4) y una serie de residuos de aminoácidos gly/ser.

[0307] Los enlazadores peptídicos de la descripción tienen al menos un aminoácido de longitud y pueden ser de distintas longitudes. En una modalidad, un enlazador peptídico de la descripción tiene entre aproximadamente 1 y aproximadamente 50 aminoácidos de longitud. Como se utiliza en este contexto, el término «aproximadamente» indica +/- dos residuos de aminoácidos. Dado que la longitud del enlace debe ser un entero positivo, la longitud de aproximadamente 1 a aproximadamente 50 aminoácidos de longitud significa una longitud de aproximadamente 1-3 a 48-52 aminoácidos de longitud. En otra modalidad, un enlazador peptídico de la descripción tiene entre aproximadamente 10 y aproximadamente 20 aminoácidos de longitud. En otra modalidad, un enlazador peptídico de la descripción tiene entre aproximadamente 15 y aproximadamente 50 aminoácidos de longitud. En otra modalidad, un enlazador peptídico de la descripción tiene entre aproximadamente 20 y aproximadamente 45 aminoácidos de longitud. En otra modalidad, un enlazador peptídico de la descripción tiene entre aproximadamente 15 y aproximadamente 35 o aproximadamente 20 y aproximadamente 30 aminoácidos de longitud. En otra modalidad, un enlazador peptídico de la descripción tiene de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 500, 1000, o 2000 aminoácidos de longitud. En una modalidad, un enlazador peptídico de la descripción tiene 20 o 30 aminoácidos de longitud.

[0308] En algunas modalidades, el enlazador peptídico puede comprender al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos 10, al menos 20, al menos 30, al menos 40, al menos 50, al menos 60, al menos 70, al menos 80, al menos 90 o al menos 100 aminoácidos. En otras realizaciones, el conector peptídico comprende al menos 200, al menos 300, al menos 400, al menos 500, al menos 600, al menos 700, al menos 800, al menos 900 o al menos 1000 aminoácidos. En algunas realizaciones, el conector peptídico puede comprender al menos aproximadamente 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900 o 2000 aminoácidos. El enlazador puede comprender 1 a 5 aminoácidos, 1 a 10 aminoácidos, 1 a 20 aminoácidos, 10 a 50 aminoácidos, 50 a 100 aminoácidos, 100 a 200 aminoácidos, 200 a 300 aminoácidos, 300 a 400 aminoácidos, 400 a 500 aminoácidos, 500 a 600 aminoácidos, 600 a 700 aminoácidos, 700 a 800 aminoácidos, 800 a 900 aminoácidos o 900 a 1000 aminoácidos.

[0309] Los enlazadores peptídicos pueden ser introducidos en secuencias polipeptídicas usando técnicas conocidas en la técnica. Las modificaciones pueden ser confirmadas por análisis de secuencias de ADN. El ADN plásmido puede usarse a transformar células hospedadoras para la producción estable de los polipéptidos producidos.

Híbridos de monómero-dímero

[0310] En algunas modalidades, las moléculas de ácido nucleico aisladas de la descripción comprenden además una secuencia de nucleótidos heteróloga codifica una molécula híbrida de monómero-dímero que comprende el FVIII.

[0311] La expresión «híbrido de monómero-dímero» que se utiliza en la presente se refiere a una proteína quimérica que comprende una primera cadena polipeptídica y una segunda cadena polipeptídica, que están asociadas entre sí mediante un enlace disulfuro, donde la primera cadena comprende un factor VIII, y una primera región Fc y la segunda cadena comprende, consiste esencialmente o consiste en una segunda región Fc sin el FVIII. De este modo, la construcción de híbrido monómero-dímero es un híbrido que comprende un aspecto de monómero que tiene únicamente un factor de coagulación y un aspecto de dímero que tiene dos regiones Fc.

Elemento de control de expresión

[0312] En algunas modalidades, la molécula de ácido nucleico o vector de la descripción comprende además al menos una secuencia de control de expresión. Una secuencia de control de expresión, tal como se utiliza en la presente, es cualquier secuencia de nucleótidos regulatoria, tal como una secuencia promotora o combinación de potenciador-promotor, que facilita la traducción y transcripción eficaces del ácido nucleico de codificación al cual está unido operativamente. Por ejemplo, la molécula de ácido nucleico aislada de la descripción se puede unir operativamente a al menos una secuencia de control de transcripción. La secuencia de control de expresión génica puede, por ejemplo, ser un promotor viral o de mamífero, tal como un promotor inducible o constitutivo. Los promotores mamíferos constitutivos incluyen, de modo no taxativo, los promotores para los siguientes genes: hipoxantina fosforribosil transferasa (HPRT), adenosina desaminasa, piruvato cinasa, promotor beta-actina y otros promotores constitutivos. Los ejemplos de promotores virales que funcionan constitutivamente en células eucariotas incluyen, por ejemplo, promotores del citomegalovirus (CMV), virus de simio (por ejemplo, SV40), virus del papiloma, adenovirus, virus de inmunodeficiencia humana (VIH),

virus del sarcoma de Rous, citomegalovirus, las 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. Los expertos en la técnica conocen otros promotores constitutivos. Los promotores útiles como secuencias de expresión

5 génica de la descripción también incluyen promotores inducibles. Los promotores inducibles se expresan en presencia de un agente de inducción. Por ejemplo, se induce el promotor de metalotioneína para promover la transcripción y traducción en presencia de determinados iones metálicos. Los expertos en la técnica conocen otros promotores inducibles.

10

[0313] En una modalidad, la descripción incluye la expresión de un transgén bajo el control de un promotor y/o potenciador específico del tejido. En otra modalidad, el promotor u otra secuencia de control de expresión mejora selectivamente la expresión del transgén en las células hepáticas. Los ejemplos de promotores

15 específico del hígado incluyen, de modo no taxativo, un promotor de tiretina de ratón (mTTR), un promotor del factor VIII humano endógeno (F8), promotor de alfa-1-antitripsina humano (hAAT), promotor mínimo de albúmina humana y promotor de albúmina de ratón. En una modalidad particular, el promotor comprende un promotor mTTR. El promotor mTTR se describe en R. H. Costa et ál., 1986, Mol. Cell. Biol.

20 6:4697. El promotor F8 se describe en Figueiredo y Brownlee, 1995, J. Biol. Chem. 270:11828-11838.

25

[0314] Los niveles de expresión pueden mejorarse también para lograr la eficacia terapéutica usando uno o más potenciadores. Se pueden proporcionar uno o más

25 potenciadores ya sea solos o juntos con uno o más elementos promotores. Comúnmente, la secuencia de control de expresión comprende múltiples elementos potenciadores y un promotor específico del tejido. En una modalidad, un potenciador comprende una o más copias del potenciador de la α -1-microglobulina/bikunina (Rouet et ál., 1992, J. Biol. Chem. 267:20765-20773; Rouet et ál., 1995, Nucleic

30 Acids Res. 23:395-404; Rouet et ál., 1998, Biochem. J. 334:577-584; Ill et ál., 1997, Blood Coagulation Fibrinolysis 8:S23-S30). En otra modalidad, un potenciador se deriva de los sitios de unión al factor de transcripción específico del hígado, tales como EBP, DBP, HNF1, HNF3, HNF4, HNF6, con Enh1, que comprende HNF1, (sentido)-HNF3, (sentido)-HNF4, (antisentido)-HNF1, (antisentido)-HNF6, (sentido)-

35 EBP, (antisentido)-HNF4 (antisentido).

[0315] En un ejemplo particular, un promotor útil para la descripción comprende la SEQ ID NO: 69 (es decir, el promotor ET; figura 11Y), que también es conocido como n.º de GenBank AY661265. Véase también Vigna et ál., *Molecular Therapy* 11(5):763 (2005). Los ejemplos de otros vectores adecuados y elementos reguladores génicos se describen en WO 02/092134, EP1395293, o las patentes estadounidenses n.º 6,808,905, 7,745,179, o 7,179,903, las cuales se incorporan a la cual se incorpora a la presente en su totalidad mediante esta referencia.

[0316] En general, las secuencias de control de expresión deberían incluir, según sea necesario, secuencias de no transcripción 5' y no traducción 5' involucradas en el inicio de la transcripción y traducción, respectivamente, tal como caja TATA, secuencia de limitación, secuencia CAAT y similares. Específicamente, dichas secuencias de no transcripción 5' incluirán una región promotora que incluye una secuencia promotora para un control transcripcional del ácido nucleico de codificación unido operativamente. Las secuencias de expresión génica incluyen opcionalmente secuencias potenciadoras o secuencias activadoras anteriores según se desea.

Sistemas de vectores

[0317] Algunas modalidades de la presente descripción se refieren a vectores que comprenden una o más moléculas de ácido nucleico optimizadas con codones que codifican un polipéptido con actividad de FVIII descrita en la presente, células hospedadoras que comprenden los vectores y métodos para tratar un trastorno hemorrágico mediante el uso de vectores. La presente descripción satisface una necesidad importante en la técnica al proporcionar un vector que comprende una secuencia de FVIII optimizada que demuestra mayor expresión en un sujeto y potencialmente provoca mayor eficacia terapéutica cuando se utiliza en métodos de terapia génica.

[0318] En algunas modalidades, la presente descripción proporciona un vector que comprende una secuencia de ácido nucleico aislada que comprende una secuencia de nucleótidos que comprende una primera secuencia de ácido nucleico que codifica una parte de extremo N de un polipéptido del FVII y una segunda secuencia de ácido nucleico que codifica una parte de extremo C del polipéptido del FVIII; donde la primera secuencia de ácido nucleico tiene al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 86 %, al menos aproximadamente 87 %, al menos aproximadamente 88 %, al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos

aproximadamente 91 %, al menos aproximadamente 92 %, al menos
aproximadamente 93 %, al menos aproximadamente 94 %, al menos
aproximadamente 95 %, al menos aproximadamente 96 %, al menos
aproximadamente 97 %, al menos aproximadamente 98 %, o al menos
5 aproximadamente 99 % de identidad de secuencia con respecto a (i) los nucleótidos
58-1791 de la SEQ ID NO: 3 o (ii) los nucleótidos 58-1791 de la SEQ ID NO: 4; y
donde la parte de extremo N y la parte de extremo C juntas tienen una actividad
polipeptídica de FVIII. En otras modalidades, la molécula de ácido nucleico
comprende una secuencia de nucleótidos que comprende una primera secuencia de
10 ácido nucleico que codifica una parte de extremo N de un polipéptido del FVIII y una
segunda secuencia de ácido nucleico que codifica una parte de extremo C de un
polipéptido del FVIII; donde la segunda secuencia de ácido nucleico tiene al menos
aproximadamente 80 %, al menos aproximadamente 85 %, al menos
aproximadamente 86 %, al menos aproximadamente 87 %, al menos
15 aproximadamente 88 %, al menos aproximadamente 89 %, al menos
aproximadamente 90 %, al menos aproximadamente 91 %, al menos
aproximadamente 92 %, al menos aproximadamente 93 %, al menos
aproximadamente 94 %, al menos aproximadamente 95 %, al menos
aproximadamente 96 %, al menos aproximadamente 97 %, al menos
20 aproximadamente 98 %, o al menos aproximadamente 99 % de identidad de
secuencia con respecto a (i) los nucleótidos 1792-4374 de la SEQ ID NO: 5; (ii) los
nucleótidos 1792-4374 de la SEQ ID NO: 6, (iii) los nucleótidos 1792-2277 y 2320-
4374 de la SEQ ID NO: 5, o (iv) los nucleótidos 1792-2277 y 2320-4374 de la SEQ
ID NO: 6; y donde la parte de extremo N y la parte de extremo C juntas tienen una
25 actividad polipeptídica de FVIII.

[0319] En algunas modalidades, la presente descripción proporciona un vector que
comprende una secuencia de ácido nucleico aislada que comprende una secuencia
de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia
30 de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos
aproximadamente 91 %, al menos aproximadamente 92 %, al menos
aproximadamente 93 %, al menos aproximadamente 94 %, al menos
aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos
99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de la
35 SEQ ID NO: 1 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 1 y está
operativamente unida a un promotor, un secuencia diana o ambos. En otras

modalidades, la secuencia de ácido nucleico comprende (i) los nucleótidos 58-4374 de la SEQ ID NO: 1 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 1.

5 **[0320]** En algunas modalidades, la presente descripción proporciona un vector que comprende una secuencia de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) 10 los nucleótidos 58-4374 de la SEQ ID NO: 2 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 2 y está operativamente unida a un promotor, un secuencia diana o ambos. En otras modalidades, la secuencia de ácido nucleico comprende (i) los nucleótidos 58-4374 de la SEQ ID NO: 2 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 2.

15

[0321] En algunas modalidades, la presente descripción proporciona un vector que comprende una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 20 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de la SEQ ID NO: 70 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 70 y está operativamente unida a un promotor, un secuencia diana o ambos. En otras 25 modalidades, la secuencia de ácido nucleico comprende (i) los nucleótidos 58-4374 de la SEQ ID NO: 70 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 70.

[0322] En algunas modalidades, la presente descripción proporciona un vector que 30 comprende una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al 35 menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de la SEQ ID NO: 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 71 y está operativamente unida a un promotor, un secuencia diana o ambos. En otras

modalidades, la secuencia de ácido nucleico comprende (i) los nucleótidos 58-4374 de la SEQ ID NO: 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 71.

- 5 **[0323]** En algunas modalidades, la presente descripción proporciona un vector que comprende una secuencia de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos
- 10 aproximadamente 94 %, al menos aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de la SEQ ID NO: 3 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 3 y está operativamente unida a un promotor, un secuencia diana o ambos. En otras modalidades, la secuencia de ácido nucleico comprende (i)
- 15 los nucleótidos 58-4374 de la SEQ ID NO: 3 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 3.

- [0324]** En algunas modalidades, la presente descripción proporciona un vector que comprende una secuencia de ácido nucleico aislada que comprende una secuencia
- 20 de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos
- 25 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de la SEQ ID NO: 4 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 4 y está operativamente unida a un promotor, un secuencia diana o ambos. En otras modalidades, la secuencia de ácido nucleico comprende (i) los nucleótidos 58-4374 de la SEQ ID NO: 4 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 4.

30

- [0325]** En algunas modalidades, la presente descripción proporciona un vector que comprende una secuencia de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos
- 35 aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos

aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de la SEQ ID NO: 5 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 5 y está operativamente unida a un promotor, un secuencia diana o ambos. En otras
5 modalidades, la secuencia de ácido nucleico comprende (i) los nucleótidos 58-4374 de la SEQ ID NO: 5 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 5.

[0326] En algunas modalidades, la presente descripción proporciona un vector que comprende una secuencia de ácido nucleico aislada que comprende una secuencia
10 de nucleótidos que codifica un polipéptido con actividad de FVIII, donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos aproximadamente 89 %, al menos aproximadamente 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos
15 aproximadamente 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con respecto a (i) los nucleótidos 58-4374 de la SEQ ID NO: 6 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 6 y está operativamente unida a un promotor, un secuencia diana o ambos. En otras modalidades, la secuencia de ácido nucleico comprende (i) los nucleótidos 58-4374
20 de la SEQ ID NO: 6 o (ii) los nucleótidos 58-2277 y 2320-4374 de la SEQ ID NO: 6.

[0327] Los vectores adecuados para la descripción incluyen vectores de expresión, vectores virales y vectores plasmídicos. En una modalidad, el vector es un vector viral.

25

[0328] Tal como se utiliza en la presente, un vector de expresión se refiere a cualquier construcción de ácido nucleico que contiene los elementos necesarios para la transcripción y traducción de una secuencia de codificación insertada o, en el caso de un vector viral de ARN, los elementos necesarios para replicación y
30 traducción, cuando se introduce en una célula hospedadora adecuada. Los vectores de expresión pueden incluir plásmidos, fagémidos, virus y derivados de estos.

[0329] Los vectores de expresión de la descripción incluirán polinucleótidos optimizados que codifican la proteína del FVIII BDD descrita en la presente. En una
35 modalidad, las secuencias de codificación optimizadas para la proteína de FVIII BDD están operativamente unidas a una secuencia de codificación de expresión. Tal como se utiliza en la presente, dos secuencias de ácido nucleico están unidas

operativamente cuando están unidas covalentemente de modo tal que permita que cada secuencia de ácido nucleico componente retenga su funcionalidad. Se dice que una secuencia de codificación y una secuencia de control de expresión génica están unidas operativamente cuando están unidas covalentemente de modo tal de colocar la expresión o transcripción y/o traducción de la secuencia de codificación bajo la influencia o el control de la secuencia de control de expresión génica. Se dice que dos secuencias de ADN están unidas operativamente si la inducción de un promotor en la secuencia de expresión génica 5' resulta en la transcripción de la secuencia de codificación y si la naturaleza de la unión entre las dos secuencias de ADN no (1) resultan en la introducción de una mutación marco-cambio, (2) interfieren con la capacidad de la región promotora de dirigir la transcripción de la secuencia de codificación o (3) interfieren con la capacidad de la transcripción de ARN correspondiente de traducirse en una proteína. De este modo, una secuencia de expresión génica se uniría operativamente a una secuencia de ácido nucleico de codificación si la secuencia de expresión génica fuera capaz de efectuar la transcripción de esa secuencia de ácido nucleico de codificación de forma tal que la transcripción resultante se traduzca en la proteína o el polipéptido deseado.

[0330] Los vectores virales incluyen, de modo no taxativo, secuencias de ácido nucleico de los siguientes virus: retrovirus, tal como virus de la leucemia murina de Moloney, virus del sarcoma murino de Harvey, virus del tumor mamario murino y virus del sarcoma de Rous; adenovirus, virus adeno-asociado; virus tipo SV40; poliomavirus; virus de Epstein-Barr; virus del papiloma; herpesvirus; vaccinia virus; polio virus y virus del ARN tal como un retrovirus. Se pueden emplear fácilmente otros vectores conocidos en la técnica. Determinados vectores virales se basan en virus eucariotas no citopáticos donde los genes no esenciales se reemplazaron con el gen de interés. Los virus no citopáticos incluyen retrovirus, el ciclo vital que involucra una transcripción inversa de ARN viral genómico en ADN con una posterior integración proviral en un ADN celular hospedador. Los retrovirus se aprobaron para análisis de terapias génicas humanas. Son más útiles aquellos retrovirus que carecen de replicación (es decir, capaz de dirigir la síntesis de las proteínas deseadas, pero incapaz de fabricar una partícula infecciosa). Dichos vectores de expresión retrovirales alterados genéticamente tienen utilidad general para la transducción de alta eficacia de genes *in vivo*. Los protocolos estándar para producir retrovirus que carecen de replicación (incluyendo los pasos de incorporación de material genético exógeno en un plásmido, transfección de una línea celular de empaquetamiento con plásmidos, producción de retrovirus recombinante mediante la

línea celular de empaquetamiento, recolección de partículas virales del medio de cultivo celular e infección de células diana con partículas virales) se proporcionan en Kriegler, M., Gene Transfer and Expression, A Laboratory Manual, W.H. Freeman Co., Nueva York (1990) y Murry, E. J., Methods in Molecular Biology, Vol. 7, Humana Press, Inc., Clifton, N.J. (1991).

[0331] En una modalidad, el virus es un virus adeno-asociado, un virus de ADN de cadena doble. El virus adeno-asociado puede modificarse genéticamente para carecer de replicación y es capaz de infectar una gran variedad de especies y tipos celulares. Además, tiene ventajas tales como estabilidad de solvente de lípidos y calor; altas frecuencias de transducción en células de diversos linajes, incluyendo células hematopoyéticas y falta de inhibición de superinfección permitiendo así varias series de transducciones. Aparentemente, el virus adeno-asociado puede integrar en el ADN celular humano de forma específica al sitio, minimizando así la posibilidad de mutagénesis por inserción y variabilidad de una característica de expresión génica insertada de infección retroviral. Además, las infecciones del virus adeno-asociado de tipo salvaje se han seguido en cultivo tisular en más de 100 pasajes en ausencia de presión selectiva, lo que implica que la integración genómica del virus adeno-asociado es un evento relativamente estable. El virus adeno-asociado también puede funcionar de forma extracromosómica.

[0332] En otra modalidad, el vector viral es un virus adeno-asociado (AAV) que ha sido manipulado para portar un polinucleótido que codifica una proteína de FVIII como se describe en la presente. Se han descrito los métodos generales para obtener AAV recombinantes (rAAV). Ver, por ejemplo, USP 8,734,809, 2013/0195801, así como también las referencias citadas en la presente. En algunas modalidades, un vector rAAV comprende una o más repeticiones terminales invertidas (ITR) de AAV y un transgén de interés (por ejemplo, una secuencia de polinucleótidos del FVIII optimizada). En determinadas modalidades, los métodos para elaborar rAAV implican cultivar una célula hospedadora deseada que contiene una secuencia de ácido nucleico que codifica una proteína de cápside de AAV o un fragmento de esta; un gen rep funcional; un vector de rAAV compuesto por repeticiones terminales invertidas (ITR) de AAV y un transgén de interés; y funciones auxiliares suficientes para permitir el empaquetamiento del vector de AAV recombinante en las proteínas de cápside de AAV. Los materiales y métodos para realizar estos procedimientos y otros relacionados se han descrito, por ejemplo, en USP 8,734,809, 2013/0195801, PCT/US1997/015692, PCT/US2002/033692,

PCT/US2002/033630, WO2007/148971, WO00/20561, WO03/042361 y WO2007/04670.

[0333] Es posible utilizar uno o más secuencias de vectores AAV diferentes derivadas de casi cualquier serotipo de acuerdo con la presente descripción. La elección de una secuencia de vector AAV particular estará guiada por parámetros conocidos tales como tropismo de interés, rendimientos del vector requeridos, etc. Generalmente, los serotipos AAV tienen secuencias genómicas de homología considerable en los niveles de aminoácido y ácido nucleico, proporcionan un conjunto relacionado de funciones genéticas, producen viriones que están relacionados y se replican y ensamblan de forma similar. Para la secuencia genómica de los diversos serotipos AAV y un resumen general de las similitudes genómicas véase, por ejemplo, número de acceso a GenBank U89790; número de acceso a GenBank J01901; número de acceso a GenBank AF043303; número de acceso a GenBank AF085716; Chlorini et ál. (1997, J. Vir. 71: 6823-33); Srivastava et ál. (1983, J. Vir. 45:555-64); Chlorini et ál. (1999, J. Vir. 73:1309-1319); Rutledge et ál. (1998, J. Vir. 72:309-319); y Wu et ál. (2000, J. Vir. 74: 8635-47). Los serotipos 1, 2, 3, 4 y 5 de AAV son una fuente ilustrativa de secuencias de nucleótidos de AAV para usar en el contexto de la presente descripción. Los AAV6, AAV7, AAV8 o AAV9 o las partículas similares al AAV recientemente desarrolladas obtenidas, por ejemplo, mediante técnicas de trasposición de cápside y bibliotecas de cápsides de AAV o a partir de ITR recientemente diseñado, desarrollado o evolucionado son también adecuados para determinadas aplicaciones de la descripción. Véase Dalkara, D et ál. (2013), Sci. Transl. Med. 5(189): 189ra76; Kotterman, MA Nat. Rev. Genet. (2014) 15(7):455.

[0334] En determinadas modalidades, sin embargo, los vectores de AAV con tropismo considerable al hígado y tejidos relacionados serán de interés para expresar las proteínas del FVIII descritas en la presente. Los ejemplos no taxativos incluyen serotipos 1, 2, 6 y 8 del AAV. Véase, por ejemplo. Torres-Torranteras et ál. (2014) 22: 901 y las referencias citadas en la presente.

[0335] En otras modalidades, el vector deriva del lentivirus. En determinadas modalidades, el vector es un vector de un lentivirus recombinante capaz de infectar células que no se dividen.

[0336] El genoma lentiviral y el ADN proviral tienen comúnmente los tres genes que se encuentran en los retrovirus: gag, pol y env, los cuales están flanqueados por dos secuencias de repetición terminal larga (LTR, por sus siglas en inglés). El gen gag codifica las proteínas estructurales internas (matriz, cápside y nucleocápside); el gen pol codifica la ADN polimerasa dirigida al ARN (transcriptasa inversa), una proteasa y una integrasa; y el gen env codifica las glicoproteínas de membrana viral. Las LTR 3' y 5' sirven para promover la transcripción y poliadenilación del virión del ARN. La LTR contiene todas las secuencias que actúan en cis necesarias para la replicación viral. Los lentivirus tiene genes adicionales que incluyen vif, vpr, tat, rev, vpu, nef y vpx (en VIH-I, VIH-2 y/o VIS).

[0337] Adyacentes a la LTR 5' se encuentran secuencias necesarias para la transcripción inversa del genoma (el sitio de unión al cebador de ARNt) y para la envoltura eficaz del ARN viral en partículas (el sitio Psi). Si faltan las secuencias necesarias para la encapsulación (o empaquetamiento del ARN retroviral en viriones infecciosos) en el genoma viral, el defecto cis evita la encapsulación del ARN genómico.

[0338] Sin embargo, el mutante resultante sigue siendo capaz de direccionar la síntesis de todas las proteínas del virión. La descripción proporciona un método para producir un lentivirus recombinante capaz de infectar una célula que no se divide que comprende transfectar una célula hospedadora adecuada con dos o más vectores que portan las funciones de empaquetamiento, a saber, gag, pol y env, así como también rev y tat. Como se describirá en la presente más adelante, los vectores que carecen de un gen tat funcional son deseable para determinadas aplicaciones. Por lo tanto, por ejemplo, un primer vector puede proporcionar un ácido nucleico que codifica un gag viral y un pol viral y otro vector puede proporcionar un ácido nucleico que codifica un env viral para producir una célula de empaquetamiento. Introducir un vector que proporciona un gen heterólogo, identificado en la presente como un vector de transferencia, en esa célula de empaquetamiento genera una célula productora que libera las partículas virales infecciosas que portan el gen de interés extraño.

[0339] De acuerdo con la configuración de vectores que se indicó anteriormente y los genes extraños, el segundo vector puede proporcionar un ácido nucleico que codifica un gen de envoltura (env) vírica. El gen env puede derivar de casi cualquier virus adecuado, lo que incluye los retrovirus. En algunas modalidades, la proteína

env es una proteína de envoltura amfotrópica que permite la transducción de células de humano y otras especies.

5 **[0340]** Los ejemplos de genes env derivados retrovirales incluyen, de modo no taxativo: virus de la leucemia murina de Moloney (MoMuLV o MMLV), virus del sarcoma murino de Harvey (HaMuSV o HSV), virus del tumor mamario murino (MuMTV o MMTV), virus de la leucemia del mono gibón (GaLV o GALV), virus de inmunodeficiencia humana (VIH) y virus del sarcoma de Rous (RSV). Se pueden utilizar también otros genes env tal como la proteína G del virus de estomatitis
10 vesicular (VSV) (VSV G), el de los virus de la hepatitis y de la gripe.

[0341] El vector que proporciona la secuencia de ácido nucleico de env viral se asocia operativamente con las secuencias reguladoras descritas en otra parte en la presente.

15

[0342] En determinadas modalidades, el vector incluye un vector lentiviral en el que los genes de virulencia del VIH env, vif, vpr, vpu y nef se eliminan sin comprometer la capacidad del vector de transducir las células que no se dividen.

20 **[0343]** En algunas modalidades, el vector incluye un vector lentiviral que comprende una eliminación de la región U3 de la LTR 3'. La eliminación de la región U3 puede ser una eliminación completa o una eliminación parcial.

[0344] En algunas modalidades, el vector lentiviral de la descripción que comprende
25 la secuencia de nucleótidos del FVIII descrita en la presente se puede transfectar en una célula con (a) una primera secuencia de nucleótidos que comprende un gag, a pol o gag o genes pol y (b) una segunda secuencia de nucleótidos que comprende un gen env heterólogo; donde el vector lentiviral carece un gen tat funcional. En otras modalidades, la célula se transfecta además con una cuarta secuencia de
30 nucleótidos que comprende un gen rev. En determinadas modalidades, el vector lentiviral carece de genes funcionales que se seleccionan de if, vpr, vpu, vpx y nef, o una combinación de estos.

[0345] En determinadas modalidades, un vector lentiviral comprende una o más
35 secuencias de nucleótidos que codifican una proteína gag, un elemento de respuesta Rev, un tracto de polipurina central (cPPT), o cualquier combinación de estos.

[0346] Los ejemplos de vectores lentivirales se describen en WO9931251, WO9712622, WO9817815, WO9817816 y WO9818934, las cuales se incorporan a la presente mediante esta referencia en su totalidad.

- 5 **[0347]** Otros vectores incluyen vectores de plásmidos. Los vectores de plásmidos se describieron en detalle en la técnica y se conocen bien por los expertos en la técnica. Véase, por ejemplo, Sambrook *et ál.*, Molecular Cloning: A Laboratory Manual, Segunda Edición, Cold Spring Harbor Laboratory Press, 1989. En los últimos años, se descubrió que los vectores de plásmidos son particularmente
- 10 ventajosos para proporcionar genes a células *in vivo* debido a su incapacidad de replicarse dentro de un genoma hospedador o integrarse en este. Estos plásmidos, no obstante, que tienen un promotor compatible con la célula hospedadora, pueden expresar un péptido de un gen codificado operativamente dentro del plásmido. Algunos plásmidos utilizados comúnmente que se encuentran disponibles de
- 15 proveedores comerciales incluyen pBR322, pUC18, pUC19, varios plásmidos de ADNpc, pRC/CMV, varios plásmidos de pCMV, pSV40 y pBlueScript. Ejemplos adicionales de plásmidos específicos incluyen pcDNA3.1, número de catálogo V79020; pcDNA3.1/higro, número de catálogo V87020; pcDNA4/myc-His, número de catálogo V86320 y pBudCE4.1, número de catálogo V53220, todos en Invitrogen
- 20 (Carlsbad, CA.). Los expertos en la técnica conocen otros plásmidos. De forma adicional, los plásmidos se pueden diseñar a medida utilizando técnicas de biología molecular estándar para quitar y/o agregar fragmentos específicos de ADN.

Expresión específica del tejido

- 25 **[0348]** En determinadas modalidades, será útil incluir dentro del vector una o más secuencias diana del miARN que, por ejemplo, están operativamente unidas al transgén del FVIII optimizado. Por lo tanto, la descripción proporciona también al menos una secuencia de miARN diana unida operativamente a la secuencia de nucleótidos del FVIII optimizada o de otro modo insertada dentro de un vector. Más
- 30 de una copia de una secuencia diana de miARN incluida en el vector puede aumentar la eficacia del sistema. Se incluyen también diferentes secuencias diana del miARN. Por ejemplo, los vectores que expresan más de un transgén pueden tener el transgén bajo control de más de una secuencia diana de miARN, que puede ser la misma o diferente. Las secuencias diana del miARN pueden estar en tándem,
- 35 pero otras configuraciones se incluyen también. El casete de expresión de transgenes que contiene las secuencias diana del miARN, se puede insertar también dentro del vector en la orientación antisentido. La orientación antisentido puede ser

útil en la producción de partículas virales para evitar la expresión de productos génicos que de otro modo pueden ser tóxicos para las células productoras. En otras modalidades, el vector comprende 1, 2, 3, 4, 5, 6, 7 u 8 copias de la misma secuencia diana del miARN o una diferente. Sin embargo, en otras determinadas

5 modalidades, el vector no incluirá ninguna secuencia diana del miARN. La elección de si incluir una secuencia diana de miARN o no (y cuántas) se guiará por parámetros conocidos tales como el objetivo tisular previsto, el nivel de expresión requerido, etc.

10 **[0349]** En una modalidad, la secuencia diana es una diana miR-223 que se ha informado que bloquea la expresión de manera más eficaz en progenitores comprometidos mieloides y al menos parcialmente en el HSPC más primitivo. La diana miR-223 puede bloquear la expresión en diferentes células mieloides diferenciadas que incluyen, granulocitos, monocitos, macrófagos, células dendríticas

15 mieloides. La diana miR-223 también puede ser adecuada para las aplicaciones de terapia génica que se basan en la expresión potente de transgenes en el linaje eritroide o linfoide. La diana miR-223 también puede bloquear la expresión de forma muy eficaz en el HSC humano.

20 **[0350]** En otra modalidad, la secuencia diana es una diana miR142 (tccataaagt aggaacact aca (SEQ ID NO: 43)). En una modalidad, el vector comprende 4 copias de las secuencias diana miR-142. En determinadas modalidades, la secuencia complementaria de los micro ARN específicos hematopoyéticos tales como miR-142 (142T), se incorpora en la región no traducida 3' de un vector, por ejemplo, vectores

25 lentivirales (LV), elaborando la transcripción que codifica genes susceptibles a la regulación por disminución mediada por miARN. Mediante este método, la expresión de transgenes se puede evitar en células presentadoras de antígenos (APC, por sus siglas en inglés) de linaje hematopoyético, mientras que se mantienen en células no hematopoyéticas (Brown et ál., Nat Med 2006). Esta estrategia puede imponer un

30 control postrancricional riguroso en la expresión de transgenes y, por lo tanto, permite el suministro estable y la expresión a largo plazo de transgenes. En algunas modalidades, la regulación del miR-142 evita la depuración mediada por el sistema inmunológico de las células transducidas y/o induce los linfocitos T reguladores (T regs) específicos para el antígeno y median la potente tolerancia inmunológica con

35 respecto al antígeno codificado por transgenes.

[0351] En algunas modalidades, la secuencia diana es una diana miR181. Chen C-Z y Lodish H, *Seminars in Immunology* (2005) 17(2):155-165 describe el miR-181, un miARN expresado específicamente en linfocitos B dentro de la médula ósea del ratón (Chen y Lodish, 2005). Además, se describe que algunos miARN humano están unidos a leucemias.

[0352] La secuencia diana puede ser parcial o completamente complementaria al miARN. La expresión «completamente complementario/a» significa que la secuencia diana tiene una secuencia de ácido nucleico que es 100 % complementaria a la secuencia del miARN que la reconoce. La expresión «parcialmente complementario/a» significa que la secuencia diana es solo en parte complementaria a la secuencia del miARN que la reconoce, a través de la cual la secuencia parcialmente complementaria es aún reconocida por el miARN. En otras palabras, una secuencia diana parcialmente complementaria en el contexto de la presente descripción es eficaz para reconocer el miARN correspondiente y efectuar la prevención o reducción de la expresión de transgenes en las células que expresan ese miARN. Los ejemplos de secuencias diana de miARN se describen en WO2007/000668, WO2004/094642, WO2010/055413, o WO2010/125471, las cuales se incorporan a la presente mediante esta referencia en su totalidad.

Células hospedadoras

[0353] La descripción proporciona también una célula hospedadora que comprende una molécula de ácido nucleico o vector de la descripción. Tal como se usa en la presente, el término «transformación» deberá usarse en un amplio sentido para referirse a la introducción de ADN en una célula hospedadora receptora que cambia el genotipo y consecuentemente da como resultado un cambio en la célula receptora.

[0354] Las «células hospedadoras» hacen referencia a células que se han transformado con vectores contruidos utilizando técnicas de ADN recombinante y codificando al menos un gen heterólogo. Las células hospedadoras de la presente descripción son preferentemente de origen mamífero, más preferentemente de origen humano o de ratón. Los expertos en la técnica tienen la capacidad de determinar de forma preferencial las líneas de células hospedadoras particulares que son más adecuadas para su propósito. Los ejemplos de líneas de células hospedadoras incluyen, pero de modo no taxativo, CHO, DG44 y DUXB11 (líneas de ovario de hámster chino, menos DHFR), HELA (carcinoma de cuello uterino

humano), CVI (línea de riñón de mono), COS (un derivado de CVI con antígeno SV40 T), R1610 (fibroblasto de hámster chino), BALBC/3T3 (fibroblasto de ratón), HAK (línea de riñón de hámster), SP2/O (mieloma de ratón), P3.veces.63-Ag3.653 (mieloma de ratón), BFA-1c1BPT (células endoteliales bovinas), RAJI (linfocito humano), PER.C6[®], NS0, CAP, BHK21, y HEK 293 (riñón humano). En una modalidad particular, la célula hospedadora se selecciona del grupo que consiste en: célula CHO, una célula HEK293, a célula BHK21, una célula PER.C6[®], una célula NS0 y una célula CAP. Las líneas de células hospedadoras normalmente se encuentran disponibles en servicios comerciales, American Tissue Culture Collection o en la bibliografía publicada.

[0355] La introducción de moléculas de ácido nucleico aisladas o vectores en la célula hospedadora se puede lograr mediante diversas técnicas que los expertos en la técnica conocen bien. Estas incluyen, de modo no taxativo, transfección incluyendo electroforesis y electroporación), fusión protoplástica, precipitación de fosfato de calcio, fusión celular con ADN cubierto, microinyección e infección con virus intacto. Véase, Ridgway, A. A. G "*Mammalian Expression Vectors*" Chapter 24.2, págs. 470-472 Vectors, Rodriguez y Denhardt, Eds. (Butterworths, Boston, Mass. 1988). Más preferentemente, la introducción de plásmidos en el hospedador se hace por medio de electroporación. Las células transformadas que alojan la construcción de expresión se cultivan en condiciones adecuadas para la producción de cadenas ligeras y cadenas pesadas y se analizan para determinar la síntesis de proteína de cadena pesada y/o ligera. Los ejemplos de técnicas de ensayo incluyen el ensayo por inmunoabsorción ligado a enzimas (ELISA), radioinmunoensayo (RIA), o análisis distribuido celular activado por fluorescencia (FACS), inmunohistoquímica y similares.

[0356] Las células hospedadoras que comprenden las moléculas de ácido nucleico aisladas o vectores de la descripción se cultivan en un medio de cultivo adecuado. Tal como se utiliza en la presente, el término «medio de crecimiento adecuado» significa un medio que contiene nutrientes que se necesitan para el crecimiento de células. Los nutrientes que se necesitan para el crecimiento celular 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 ternero bovino o suero de ternero fetal (FCS). En una modalidad, el medio no contiene sustancialmente ninguna IgG. El medio de cultivo seleccionará

generalmente células que contienen la construcción de ADN mediante, por ejemplo, selección de fármacos o carencia de un nutriente esencial que se complementa mediante el marcador seleccionable sobre la construcción de ADN o cotransfectada con la construcción de ADN. Las células de mamífero cultivadas se cultivan generalmente en un medio libre de suero o que contiene suero comercialmente disponible (por ejemplo, MEM, DMEM, DMEM/F12). En una modalidad, el medio es CDoptiCHO (Invitrogen, Carlsbad, CA.). En otra modalidad, el medio es CD17 (Invitrogen, Carlsbad, CA.). La selección de un medio adecuado para la línea celular particular utilizada se encuentra dentro del nivel de los expertos en la técnica.

10

Preparación de polipéptidos

[0357] La descripción proporciona también un polipéptido codificado por una molécula de ácido nucleico de la descripción. En otras modalidades, el polipéptido de la descripción está codificado por un vector que comprende las moléculas de ácido nucleico aisladas de la descripción. En incluso otras modalidades, el polipéptido de la descripción es producido por una célula hospedadora que comprende las moléculas de ácido nucleico aisladas de la descripción.

[0358] En otras modalidades, la descripción proporciona también un método para producir un polipéptido con actividad de FVIII que comprende cultivar una célula hospedadora de la descripción en condiciones en las cuales se produce un polipéptido con actividad de FVIII y, recuperar el polipéptido con actividad de FVIII. En algunas modalidades, se incrementa la expresión del polipéptido con actividad de FVIII con respecto a una célula hospedadora cultivada en las mismas condiciones pero que comprende una secuencia de nucleótidos de referencia que comprende la SEQ ID NO: 16, la secuencia génica del FVIII parental.

[0359] En otras modalidades, la descripción proporciona un método para aumentar la expresión de un polipéptido con actividad de FVIII que comprende cultivar una célula hospedadora de la descripción en condiciones en las cuales se expresa un polipéptido con actividad de FVIII mediante la molécula de ácido nucleico, donde se incrementa la expresión del polipéptido con actividad de FVIII con respecto a una célula hospedadora cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16.

35

[0360] En otras modalidades, la descripción proporciona un método para mejorar el rendimiento de un polipéptido con actividad de FVIII que comprende cultivar una

célula hospedadora de la descripción en condiciones en las cuales se produce un polipéptido con actividad de FVIII mediante la molécula de ácido nucleico, donde se incrementa el rendimiento del polipéptido con actividad de FVIII con respecto a una célula hospedadora cultivada en las mismas condiciones que comprende una
5 secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16.

[0361] En otras modalidades, la descripción proporciona un método para aumentar el rendimiento de un polipéptido con actividad de FVIII que comprende cultivar una célula hospedadora que comprende una secuencia de nucleótidos que codifica el
10 polipéptido, donde se incrementa el índice de adaptación de codones de una parte 3' de la secuencia de nucleótidos con respecto a una parte 5' de la secuencia de nucleótidos; donde se incrementa el rendimiento del polipéptido con actividad de FVIII con respecto a una célula hospedadora cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la
15 SEQ ID NO: 16. En algunas modalidades, el índice de adaptación de codones de la parte 5' de la secuencia de nucleótidos se incrementa, disminuye o no cambia con respecto al índice de adaptación de codones de la SEQ ID NO: 16.

[0362] En otras modalidades, la descripción proporciona un método para aumentar el
20 rendimiento de un polipéptido con actividad de FVIII que comprende cultivar una célula hospedadora que comprende una secuencia de nucleótidos que codifica el polipéptido, donde se incrementa el índice de adaptación de codones de una parte 5' de la secuencia de nucleótidos con respecto a una parte 3' de la secuencia de nucleótidos; donde se incrementa el rendimiento del polipéptido con actividad de
25 FVIII con respecto a una célula hospedadora cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16. En algunas modalidades, el índice de adaptación de codones de la parte 3' de la secuencia de nucleótidos se incrementa, disminuye o no cambia con respecto al índice de adaptación de codones de la SEQ ID NO: 16.

30

[0363] En otras modalidades, la descripción proporciona un método para mejorar el rendimiento de un polipéptido con actividad de FVIII que comprende cultivar una célula hospedadora que comprende una secuencia de nucleótidos que codifica el polipéptido, donde se incrementa el índice de adaptación de codones de una parte
35 del nucleótido que codifica una parte de extremo C del polipéptido con respecto a una parte del nucleótido que codifica una parte de extremo N del polipéptido; donde se incrementa el rendimiento del polipéptido con actividad de FVIII con respecto a

una célula hospedadora cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16. En algunas modalidades, el índice de adaptación de codones de la parte del nucleótido que codifica la parte de extremo N del polipéptido se incrementa, disminuye o no cambia con respecto al índice de adaptación de codones de la SEQ ID NO: 16.

[0364] En otras modalidades, la descripción proporciona un método para mejorar el rendimiento de un polipéptido con actividad de FVIII que comprende cultivar una célula hospedadora que comprende una secuencia de nucleótidos que codifica el polipéptido, donde se incrementa el índice de adaptación de codones de una parte del nucleótido que codifica una parte de extremo N del polipéptido con respecto a una parte del nucleótido que codifica una parte de extremo C del polipéptido; donde se incrementa el rendimiento del polipéptido con actividad de FVIII con respecto a una célula hospedadora cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16. En algunas modalidades, el índice de adaptación de codones de la parte del nucleótido que codifica la parte de extremo C del polipéptido se incrementa, disminuye o no cambia con respecto al índice de adaptación de codones de la SEQ ID NO: 16.

[0365] En determinadas modalidades para mejorar el rendimiento de un polipéptido con actividad de FVIII, la parte 5' de la secuencia de nucleótidos, cuando se alinea de forma adecuada, se corresponde con aproximadamente los nucleótidos 1-1791 de la SEQ ID NO: 1, los nucleótidos 58-1791 de la SEQ ID NO: 1 o un fragmento de estos. En otras modalidades, un polipéptido codificado por la parte 5' del nucleótido, cuando se alinea de forma adecuada, se corresponde con aproximadamente los nucleótidos 1-497 de la SEQ ID NO: 17, aminoácidos 20-497 de la SEQ ID NO: 17 o un fragmento de estos. En determinadas modalidades, la parte de la secuencia de nucleótidos que codifica la parte de extremo N del polipéptido, cuando se alinea de forma adecuada, se corresponde con aproximadamente los nucleótidos 1-1791 de la SEQ ID NO: 1, los nucleótidos 58-1791 de la SEQ ID NO: 1 o un fragmento de estos. En algunas modalidades, la parte 3' de la secuencia de nucleótidos, cuando se alinea de forma adecuada, se corresponde aproximadamente con los nucleótidos 1792-4374 de la SEQ ID NO: 1 o un fragmento de este. En otras modalidades, un polipéptido codificado por la parte 3' del nucleótido, cuando se alinea de forma adecuada, se corresponde con aproximadamente los nucleótidos 498-1458 de la SEQ ID NO: 17 o un fragmento de estos. En determinadas modalidades, la parte de la secuencia de nucleótidos que codifica la parte de extremo C del polipéptido,

cuando se alinea de forma adecuada, se corresponde aproximadamente con los nucleótidos 1792-4374 de la SEQ ID NO: 1 o un fragmento de este

[0366] En algunas modalidades, se incrementa la expresión del polipéptido del FVIII en al menos aproximadamente 1,5 veces, al menos aproximadamente 2 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 6 veces, al menos aproximadamente 7 veces, al menos aproximadamente 8 veces, al menos aproximadamente 9 veces, al menos aproximadamente 10 veces, al menos aproximadamente 11 veces, al menos aproximadamente 12 veces, al menos aproximadamente 13 veces, al menos aproximadamente 14 veces, al menos aproximadamente 15 veces, al menos aproximadamente 20 veces, al menos aproximadamente 25 veces, al menos aproximadamente 30 veces, al menos aproximadamente 35 veces, al menos aproximadamente 40 veces, al menos aproximadamente 50 veces, al menos aproximadamente 60 veces, al menos aproximadamente 70 veces, al menos aproximadamente 80 veces, al menos aproximadamente 90 veces, al menos aproximadamente 100 veces, al menos aproximadamente 150 veces, o al menos aproximadamente 200 veces con respecto a una célula hospedadora cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16.

[0367] Se encuentra disponible una cantidad de métodos para producir de forma recombinante una proteína de FVIII a partir de la molécula de ácido nucleico optimizada de la descripción. Un polinucleótido de la secuencia deseada puede ser producido mediante síntesis de ADN de fase sólida *de novo* o mediante mutagénesis por PCR de un polinucleótido preparado previamente. La mutagénesis mediada por oligonucleótidos es un método para preparar una sustitución, inserción, eliminación o alteración (por ejemplo, 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 a una plantilla de ADN de cadena simple. Luego de la hibridación, se utiliza una ADN polimerasa para sintetizar una segunda cadena complementaria completa de la plantilla que incorpora el cebador de oligonucleótidos. En una modalidad, la modificación genética, por ejemplo, mutagénesis por PCR basada en cebadores, es suficiente como para incorporar una alteración, tal como se define en la presente, con el fin de producir un polinucleótido de la descripción.

[0368] Para la producción de proteínas recombinantes, se inserta una secuencia de polinucleótidos optimizada de la descripción que codifica la proteína de FVIII en un vehículo de expresión adecuado, es decir, un vector que contiene los elementos necesarios para la transcripción y traducción de la secuencia de codificación insertada, o en el caso de un vector viral de ARN, los elementos necesarios para la replicación y traducción.

[0369] La secuencia de polinucleótidos de la descripción se inserta en el vector en un marco de lectura adecuado. Posteriormente, el vector de expresión se transfecta a una célula diana adecuada que expresará el polipéptido. Las técnicas de transfección conocidas en la técnica incluyen, de modo no taxativo, la precipitación con fosfato de calcio (Wigler *et ál.* 1978, *Cell* 14: 725) y la electroporación (Neumann *et ál.* 1982, *EMBO, J.* 1: 841). Se puede utilizar una variedad de sistemas de vectores de expresión hospedadores para expresar las proteínas de FVIII descritas en la presente en células eucariotas. En una modalidad, la célula eucariótica es una célula animal, que incluye células de mamíferos (por ejemplo, células HEK293, células PER.C6[®], CHO, BHK, Cos, HeLa). Una secuencia de polinucleótidos de la descripción puede codificar también una secuencia de señal que permitirá que se secrete la proteína de FVIII. Un experto en la técnica entenderá que, al traducirse la proteína de FVIII, la secuencia de señal es escindida por la célula para formar la proteína madura. En la técnica se conocen varias secuencias de señal, por ejemplo, secuencia de señal del factor VII natural, la secuencia de señal IX del factor natural y la secuencia de señal de cadena ligera IgK de ratón. De manera alternativa, cuando no se incluye una secuencia de señal, la proteína de FVIII puede ser recuperada por medio del lisado de las células.

[0370] La proteína de FVIII de la descripción puede ser sintetizada en un animal transgénico, tal como un roedor, cabra, oveja, cerdo o vaca. La expresión «animales transgénicos» hace referencia a animales no humanos que han incorporado un gen extraño en su genoma. Debido a que este gen está presente en tejidos germinales, se pasa de los progenitores a su descendencia. Los genes exógenos se introducen en embriones de una célula (Brinster *et ál.* 1985, *Proc. Natl. Acad. Sci. EUA* 82:4438). En la técnica se conocen métodos de producción de animales transgénicos que incluyen transgénicos que producen moléculas de inmunoglobulina (Wagner *et ál.* 1981, *Proc. Natl. Acad. Sci. EUA* 78: 6376; McKnight *et ál.* 1983, *Cell* 34: 335; Brinster *et ál.* 1983, *Nature* 306: 332; Ritchie *et ál.* 1984, *Nature* 312: 517;

Baldassarre et ál. 2003, Theriogenology 59: 831; Robl et ál. 2003, Theriogenology 59: 107; Malassagne et ál. 2003, Xenotransplantation 10 (3): 267).

[0371] Los vectores de expresión pueden codificar las etiquetas que permiten la fácil
 5 purificación o identificación de la proteína producida de forma recombinante. Los
 ejemplos incluyen, de modo no taxativo, el vector pUR278 (Ruther et ál. 1983,
 EMBO J. 2: 1791) en el cual la proteína de FVIII descrita en la presente que codifica
 la secuencia puede estar ligada en el vector en marco con la región de codificación
 lac Z de manera que se produzca una proteína híbrida; los vectores pGEX pueden
 10 utilizarse para expresar proteínas con una etiqueta de glutatión S-transferasa
 (GST). Estas proteínas son normalmente solubles y se pueden purificar fácilmente a
 partir de células mediante la adsorción de cuentas glutatión-agarosa seguido de
 elución en presencia de glutatión libre. Los vectores incluyen sitios de escisión, por
 ejemplo, para PreCission Protease (Pharmacia, Peapack, N. J.) para una
 15 eliminación fácil de la etiqueta después de la purificación.

[0372] A los efectos de esta descripción se pueden emplear numerosos sistemas de
 vector de expresión. Estos vectores de expresión normalmente se replican en los
 organismos hospedadores ya sea como episomas o como una parte integral del
 20 ADN cromosómico huésped. Los vectores de expresión pueden incluir secuencias
 de control de expresión que incluyen, de modo no taxativo, promotores (por ejemplo,
 promotores heterólogos o asociados naturalmente), potenciadores, secuencias de
 señal, señales de empalme, elementos potenciadores y secuencias de terminación
 de la transcripción. Preferentemente, las secuencias de control de expresión son
 25 sistemas de promotores eucariotas en vectores capaces de transformar o transfectar
 células hospedadoras eucariotas. Los vectores de expresión también pueden utilizar
 elementos de ADN que derivan de virus de animales, tales como el virus del
 papiloma bovino, virus del polio, adenovirus, virus vaccinia, baculovirus, retrovirus
 (RSV, MMTV o MOMLV), citomegalovirus (CMV) o virus SV40. Otros implican el uso
 30 de sistemas policistrónicos con sitios de unión a ribosomas internos.

[0373] Comúnmente, los vectores de expresión contienen marcadores de selección
 (por ejemplo, resistencia a la ampicilina, resistencia a la higromicina, resistencia a la
 tetraciclina o resistencia a la neomicina) para permitir la detección de esas células
 35 transformadas con las secuencias de ADN deseado (véase, por ejemplo, Itakura et
 ál., patente estadounidense 4,704,362). Las células que tienen el ADN integrado a
 sus cromosomas se pueden seleccionar mediante la introducción de uno o más

marcadores que permiten la selección de células hospedadoras transfectadas. El marcador puede proporcionar prototrofia a un hospedador auxotrófico, resistencia a biocidas (por ejemplo, antibióticos) o resistencia a metales pesados como cobre. El gen marcador seleccionable puede estar unido directamente a las secuencias de ADN para expresarse o puede introducirse en la misma célula mediante cotransformación.

[0374] Un ejemplo de un vector útil para expresar una secuencia de FVIII optimizada es NEOSPLA (patente estadounidense n.º 6,159,730). Este vector contiene el promotor/potenciador de citomegalovirus, el promotor principal de beta globina de ratón, el origen SV40 de replicación, la secuencia de poliadenilación de la hormona de crecimiento bovina, exón 1 y exón 2 de fosfotransferasa de neomicina, el gen dihidrofolato reductasa y la secuencia líder. Se ha encontrado que este vector da como resultado un nivel muy alto de expresión de anticuerpos tras la incorporación de genes de región variable y constante, la transfección de células, seguido por la selección en G418 que contiene medio y amplificación de metotrexato. Los sistemas de vector también se enseñan en las patentes estadounidenses n.º 5,736,137 y 5,658,570, cada una de las cuales se incorpora en su totalidad en la presente mediante esta referencia. El sistema proporciona altos niveles de expresión, por ejemplo, > 30 pg/célula/día. Otros ejemplos de sistemas de vectores se describen, por ejemplo, en la patente estadounidense n.º 6,413,777.

[0375] En otras modalidades, los polipéptidos de la descripción de la presente invención pueden ser expresados usando construcciones policistrónicas. En estos sistemas de expresión, se pueden producir múltiples productos génicos de interés, tales como polipéptidos múltiples de proteína que se une a multímero a partir de una construcción policistrónica simple. Estos sistemas utilizan de manera ventajosa un sitio interno de entrada al ribosoma (IRES) para proporcionar niveles relativamente altos de polipéptidos de unión en células hospedadoras eucariotas. Las secuencias IRES compatibles se describen en la patente estadounidense n.º 6,193,980, la cual también se incorpora a la presente.

[0376] Más generalmente, una vez que se ha preparado el vector o secuencia de ADN que codifica un polipéptido, el vector de expresión se puede introducir en una célula hospedadora adecuada. Esto es, las células hospedadoras pueden ser transformadas. La introducción del plásmido en la célula hospedadora se puede lograr mediante diversas técnicas que los expertos en la técnica conocen bien. Las

células transformadas se cultivan en condiciones adecuadas para la producción del polipéptido del FVIII y se analizan para determinar la síntesis del polipéptido de FVIII. Los ejemplos de técnicas de ensayo incluyen el ensayo por inmunoabsorción ligado a enzimas (ELISA), radioinmunoensayo (RIA), o análisis distribuido celular activado por fluorescencia (FACS), inmunohistoquímica y similares.

[0377] En las descripciones de procesos para el aislamiento de polipéptidos de hospedadores recombinantes, los términos «célula» y «cultivo celular» se usan de manera intercambiable para indicar la fuente de polipéptido, a menos que se especifique claramente lo contrario. En otras palabras, la recuperación de polipéptidos desde las «células» puede significar desde células enteras centrifugadas o desde el cultivo celular que contiene tanto el medio como las células suspendidas.

[0378] La línea de células hospedadoras utilizada para la expresión de proteínas tiene preferentemente origen mamífero, más preferentemente origen humano o de ratón, ya que los ácidos nucleicos aislados de la descripción se han optimizado para la expresión en células humanas. Las líneas de células hospedadoras se describieron anteriormente. En una modalidad del método para producir un polipéptido con actividad de FVIII, la célula hospedadora es una célula HEK293. En otra modalidad del método para producir un polipéptido con actividad de FVIII, la célula hospedadora es una célula CHO.

[0379] Los genes que codifican los polipéptidos de la descripción pueden expresarse también en células no mamíferas, tales como bacterias o levadura o células de plantas. En este sentido, se apreciará que varios microorganismos no mamíferos unicelulares tales como bacterias también pueden transformarse; es decir, es posible cultivarlos en cultivos o fermentación. Las bacterias, que son susceptibles a la transformación, incluyen miembros de enterobacteriaceae, tales como cepas de *Escherichia coli* o *Salmonella*; *Bacillaceae*, tales como *Bacillus subtilis*; *Pneumococcus*; *Streptococcus* y *Haemophilus influenzae*. También se observará que, cuando se expresan en bacterias, los polipéptidos se vuelven, normalmente, parte de los cuerpos de inclusión. Los polipéptidos deben aislarse, purificarse y luego ensamblarse en moléculas funcionales.

[0380] De manera alternativa, las secuencias de nucleótidos optimizadas se pueden incorporar en transgenes para la introducción en el genoma de un animal

transgénico y la expresión posterior en la leche del animal transgénico (véase, por ejemplo, Deboer *et ál.*, US 5,741,957, Rosen, US 5,304,489 y Meade *et ál.*, US 5,849,992). Los transgenes adecuados incluyen secuencias de codificación en unión operativa con un promotor y potenciador de un gen específico de glándula mamaria, tales como la caseína o la beta lactoglobulina.

[0381] La producción *in vitro* permite que el aumento proporcione grandes cantidades de los polipéptidos deseados. En la técnica se conocen técnicas de cultivo de células de mamíferos en condiciones de cultivo de tejidos e incluyen el cultivo de suspensiones homogéneas, por ejemplo, en un reactor aerotransportado o en un reactor de agitación continuo o el cultivo celular inmovilizado o atrapado, por ejemplo, en fibras huecas, microcápsulas, en microperlas de agarosa o cartuchos de cerámica. Si se necesita y/o desea, las soluciones de los polipéptidos se pueden purificar mediante los métodos de cromatografía habituales, por ejemplo, filtración de gel, cromatografía de intercambio iónico, cromatografía sobre celulosa DEAE o cromatografía por (inmuno)afinidad, por ejemplo, antes de la biosíntesis preferencial de un polipéptido de región bisagra sintética o antes o después del paso de la cromatografía HIC descritos en la presente. Opcionalmente, se puede acoplar o incluir una secuencia de etiqueta de afinidad (*por ejemplo*, una etiqueta His(6)) en la secuencia polipeptídica para facilitar la purificación posterior.

[0382] Una vez expresada, la proteína de FVIII puede purificarse según los métodos estándar de la técnica, lo que incluye una precipitación de sulfato de amonio, cromatografía en columna por afinidad, purificación HPLC, electroforesis en gel y similares (véase en general Scopes, Protein Purification (Springer-Verlag, Nuevas York, (1982)). Para uso farmacéutico, se prefieren las proteínas sustancialmente puras de aproximadamente 90 al 95 % de homogeneidad, y se prefieren aún más del 98 al 99 % o más de homogeneidad.

Composición farmacéutica

[0383] 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 hospedadora de la presente descripción pueden contener un portador farmacéuticamente aceptable adecuado. Por ejemplo, pueden contener excipientes y/o auxiliares que facilitan el procesamiento de los compuestos activos en preparaciones diseñadas para el suministro al sitio de acción.

[0384] La composición farmacéutica se puede formular para administración parenteral (es decir, intravenosa, subcutánea o intramuscular) mediante inyección en bolo. Las formulaciones para inyección se pueden presentar en forma de dosificación unitaria, por ejemplo, en ampollas o en recipientes de varias dosis con un conservante agregado. Las composiciones pueden tomar dichas formas como suspensiones, soluciones o emulsiones en vehículos oleosos o acuosos, y contienen agentes de formulación tales como agentes de suspensión, estabilización y/o dispersión. De forma alternativa, el ingrediente activo puede estar en forma de polvo para su constitución con un vehículo adecuado, por ejemplo, agua libre de pirógeno.

10

[0385] Las formulaciones adecuadas para administración parenteral también incluyen soluciones acuosas de los compuestos activos en una forma soluble en agua, por ejemplo, sales solubles en agua. Además, se pueden administrar suspensiones de los compuestos activos como suspensiones de inyección oleosas adecuadas. Los vehículos o solventes lipofílicos adecuados incluyen aceites grasos, por ejemplo, aceite de sésamo o ésteres de ácidos grasos sintéticos, por ejemplo, etil oleato o triglicéridos. Las suspensiones de inyección acuosas pueden contener sustancias, que aumentan la viscosidad de la suspensión lo que incluye, por ejemplo, carboximetilcelulosa de sodio, sorbitol y dextrano. Opcionalmente, la suspensión también puede contener estabilizantes. Los liposomas también se pueden utilizar para encapsular las moléculas de la descripción para su suministro en células o espacios intersticiales. Los ejemplos de portadores farmacéuticamente aceptables son solventes fisiológicamente compatibles, medios de dispersión, recubrimientos, agentes antibacterial y antifúngica, agentes de retraso de absorción e isotónicos, agua, solución salina, solución salina tamponada con fosfato, dextrosa, glicerol, etanol y similares. En algunas modalidades, la composición comprende agentes isotónicos, por ejemplo, azúcares, polialcoholes tal como manitol, sorbitol o cloruro de sodio. En otras modalidades, las composiciones comprenden sustancias farmacéuticamente aceptables tales como, agentes humectantes o cantidades menores de sustancias auxiliares tales como agentes humectantes o emulsionantes, conservantes o amortiguadores, que potencian la semivida o efectividad de los ingredientes activos.

25

30

35

[0386] Las composiciones de la descripción pueden estar en diversas formas, que incluyen, por ejemplo, soluciones líquidas (por ejemplo, inyectables e infusibles), dispersiones, suspensiones, formas de dosificación sólidas y semisólidas. La forma preferida depende del modo de administración y aplicación terapéutica.

- [0387] La composición puede formularse como una solución, microemulsión, dispersión, liposoma u otra estructura ordenada adecuada para una concentración alta de fármaco. Las soluciones inyectables estériles se pueden preparar incorporando el ingrediente activo en la cantidad requerida en un solvente adecuado con una o más combinaciones de ingredientes enumerados anteriormente, según se requiere, seguido de esterilización filtrada. Generalmente, las dispersiones se preparan incorporando el ingrediente activo en un vehículo estéril que contiene un medio de dispersión básico y los otros ingredientes requeridos de los enumerados anteriormente. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los métodos preferidos de preparación son secado al vacío y liofilización que proporcionan un polvo del ingrediente activo más cualquier ingrediente deseado adicional de una solución filtrada previamente estérilmente. La fluidez adecuada de una solución se puede mantener, por ejemplo, mediante el uso de un recubrimiento, tal como lecitina, por el mantenimiento del tamaño de partícula requerido en el caso de la dispersión y por el uso de tensioactivos. La absorción prolongada de composiciones inyectables se puede lograr incluyendo en la composición un agente que retrasa la absorción, por ejemplo, sales de monostearato y gelatina.
- [0388] El ingrediente activo se puede formular con un dispositivo o formulación de liberación controlada. Los ejemplos de dichas formulaciones y dispositivos incluyen implantes, parches transdérmicos y sistemas de administración microencapsulados. Se pueden utilizar polímeros biocompatibles y biodegradables, por ejemplo, etilvinilacetato, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres y ácido poliláctico. Los métodos para la preparación de dichas formulaciones y dispositivos se conocen en la técnica. Véase, por ejemplo, Sustained and Controlled Release Drug Delivery Systems, J. R. Robinson, ed., Marcel Dekker, Inc., Nueva York, 1978.
- [0389] Las formulaciones de depósito inyectables se pueden realizar mediante la formación de matrices microencapsuladas del fármaco en polímeros biodegradables tal como, poliláctido-poliglicólido. Dependiendo de la relación del fármaco con respecto al polímero, y de la naturaleza del polímero empleado, se puede controlar la relación de liberación del fármaco. Otros ejemplos de polímeros biodegradables son poliortoésteres y polianhídridos. Las formulaciones inyectables de depósito pueden también prepararse reteniendo el fármaco en liposomas o microemulsiones.

[0390] Los compuestos activos complementarios se pueden incorporar en las composiciones. En una modalidad, la proteína quimérica de la descripción se formula con otro factor de coagulación, o una variante, fragmento, análogo o derivado de estos. Por ejemplo, el factor de coagulación incluye, de modo no

5 taxativo, factor V, factor VII, factor IX, factor X, factor XI, factor XII, factor XIII, protrombina, fibrinógeno, factor von Willebrand o factor tisular soluble recombinante (rsTF) o formas activadas de cualquiera de los anteriores. El factor de coagulación del agente hemostático también puede incluir fármacos anti-fibrinolíticos, por ejemplo, ácido épsilon-amino-caproico, ácido tranexámico.

10

[0391] Los regímenes de dosificación se pueden ajustar para proporcionar la respuesta deseada óptima. Por ejemplo, se puede administrar un único bolo, se pueden administrar varias dosis divididas durante el tiempo o la dosis se puede reducir o aumentar proporcionalmente según se indica en las exigencias de la

15 situación terapéutica. Es ventajoso formular composiciones parenterales en forma unitaria de dosificación para facilitar la administración y uniformidad de la dosificación. Véase, por ejemplo, Remington's Pharmaceutical Sciences (Mack Pub. Co., Easton, Pa. 1980).

20 **[0392]** Además del compuesto activo, la forma de dosificación líquida puede contener ingredientes inertes tales como agua, alcohol etílico, carbonato de etilo, acetato de etilo, alcohol bencílico, benzoato de bencilo, propilenglicol, 1,3-butilenglicol, dimetilformamida, aceites, glicerol, alcohol de tetrahidrofurfurilo, polietilenglicoles y ésteres de ácidos grasos de sorbitán.

25

[0393] Los ejemplos no taxativos de portadores farmacéuticos adecuados también se describen en Remington's Pharmaceutical Sciences por E. W. Martin. Algunos ejemplos de excipientes incluyen almidón, glucosa, lactosa, sacarosa, gelatina, malta, arroz, harina, creta, gel de sílice, estearato de sodio, monoestearato de

30 glicerol, talco, cloruro de sodio, leche descremada en polvo, glicerol, propileno, glicol, agua, etanol y similares. Las composiciones también pueden contener reactivos de amortiguación del pH y agentes humectantes o emulsionantes.

[0394] Para la administración oral, la composición farmacéutica puede tener la forma

35 de comprimidos o cápsulas preparados mediante medios convencionales. La composición también se puede preparar como un líquido, por ejemplo, un jarabe o una suspensión. El líquido puede incluir agentes de suspensión (por ejemplo, jarabe

de sorbitol, derivados de celulosa o grasas comestibles hidrogenadas), agentes emulsionantes (lecitina o acacia), vehículos no acuosos (por ejemplo, aceite de almendras, ésteres oleosos, alcohol etílico o aceites vegetales fraccionados) y conservantes (por ejemplo, metil o propil-p-hidroxibenzoatos o ácido sórbico). Las preparaciones también pueden incluir agentes saborizantes, colorantes y endulzantes. De forma alternativa, la composición se puede presentar como un producto seco para su constitución con agua u otro vehículo adecuado.

[0395] Para la administración bucal, la composición puede tener la forma de comprimidos o grageas de acuerdo con los protocolos convencionales.

[0396] Para administración mediante inhalación, los compuestos para su uso de acuerdo con la presente invención se administran de forma conveniente en la forma de un aerosol nebulizado con o sin excipientes o en la forma de un pulverizador en aerosol a partir de un nebulizador o paquete presurizado, opcionalmente con un propelante, por ejemplo, diclorodifluorometano, triclorofluorometano, diclorotetrafluorometano, dióxido de carbono u otro gas adecuado. En el caso de un aerosol presurizado la dosificación unitaria se puede determinar proporcionando una válvula para administrar una cantidad medida. Las cápsulas y cartuchos de, por ejemplo, gelatina para su uso en un inhalador o insuflador se pueden formular con una mezcla de polvo del compuesto y una base de polvo adecuada, tal como lactosa o almidón.

[0397] La composición farmacéutica también puede formularse para administración rectal como supositorio o enema de retención, por ejemplo, que contiene bases de supositorio convencionales tales como manteca de cacao u otros glicéridos.

[0398] En una modalidad, una composición farmacéutica comprende un polipéptido que tiene actividad de FVIII, una molécula de ácido nucleico optimizada que codifica el polipéptido que tiene actividad del factor VIII, el vector que comprende la molécula de ácido nucleico o la célula hospedadora que comprende el vector y un portador farmacéuticamente aceptable. En algunas modalidades, la composición se administra mediante una vía que se selecciona 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.

[0399] En otras modalidades, la composición se utiliza para tratar una enfermedad o afección hemorrágica en un sujeto que lo necesita. La enfermedad o afección hemorrágica se selecciona del grupo que consiste en un trastorno de coagulación hemorrágica, hemartrosis, sangrado muscular, sangrado oral, hemorragia, hemorragia en los músculos, hemorragia oral, trauma, trauma capitis, sangrado gastrointestinal, hemorragia intracraneal, hemorragia intra-abdominal, hemorragia intratorácica, fractura ósea, sangrado del sistema nervioso central, sangrado en el espacio retrofaríngeo, sangrado en el espacio retroperitoneal, sangrado en la vaina de iliopsoas y cualquier combinación de estos. En aun otras modalidades, se coordina una cirugía para el sujeto. En aun otras modalidades, el tratamiento es profiláctico o según demanda.

Métodos de tratamiento

[0400] La descripción proporciona un método para tratar un trastorno hemorrágico que comprende administrarle a un sujeto que lo necesite una molécula de ácido nucleico, un vector o un polipéptido de la presente. En algunas modalidades, el trastorno hemorrágico se caracteriza por una deficiencia en FVIII. En algunas modalidades, el trastorno hemorrágico es hemofilia. En algunas modalidades, el trastorno hemorrágico es hemofilia A. En algunas modalidades del método para tratar un trastorno hemorrágico, se incrementa la actividad de FVIII en el plasma a las 24 horas después de la administración con respecto a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia.

[0401] En algunas modalidades, la actividad de FVIII en el plasma se incrementa en aproximadamente 6 horas, en aproximadamente 12 horas, en aproximadamente 18 horas, en aproximadamente 24 horas, en aproximadamente 36 horas, en aproximadamente 48 horas, en aproximadamente 3 días, en aproximadamente 4 días, en aproximadamente 5 días, en aproximadamente 6 días, en aproximadamente 7 días, en aproximadamente 8 días, en aproximadamente 9 días, en aproximadamente 10 días, en aproximadamente 11 días, en aproximadamente 12 días, en aproximadamente 13 días, en aproximadamente 14 días, en aproximadamente 15 días, en aproximadamente 16 días, en aproximadamente 17 días, en aproximadamente 18 días, en aproximadamente 19 días, en aproximadamente 20 días, en aproximadamente 21 días, en aproximadamente 22 días, en aproximadamente 23 días, en aproximadamente 24 días, en

aproximadamente 25 días, en aproximadamente 26 días, en aproximadamente 27 días, o en aproximadamente 28 días después de la administración relativa a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En determinadas modalidades, la actividad de FVIII en el plasma se incrementa en aproximadamente 24 horas después de la administración relativa a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En otra modalidad, la actividad de FVIII en el plasma se incrementa en aproximadamente 21 horas después de la administración relativa a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia.

[0402] En algunas modalidades, la actividad de FVIII en el plasma después de la administración se incrementa en al menos aproximadamente 1,5 veces, al menos aproximadamente 2 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 6 veces, al menos aproximadamente 7 veces, al menos aproximadamente 8 veces, al menos aproximadamente 9 veces, al menos aproximadamente 10 veces, al menos aproximadamente 11 veces, al menos aproximadamente 12 veces, al menos aproximadamente 13 veces, al menos aproximadamente 14 veces, al menos aproximadamente 15 veces, al menos aproximadamente 20 veces, al menos aproximadamente 25 veces, al menos aproximadamente 30 veces, al menos aproximadamente 35 veces, al menos aproximadamente 40 veces, al menos aproximadamente 50 veces, al menos aproximadamente 60 veces, al menos aproximadamente 70 veces, al menos aproximadamente 80 veces, al menos aproximadamente 90 veces, al menos aproximadamente 100 veces, al menos aproximadamente 150 veces, al menos aproximadamente 200 veces, al menos aproximadamente 250 veces, al menos aproximadamente 300 veces, al menos aproximadamente 350 veces, al menos aproximadamente 400 veces, al menos aproximadamente 450 veces, o al menos aproximadamente 500 veces con respecto a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un

vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas modalidades, se incrementa la actividad de FVIII en el plasma después de la administración en al menos aproximadamente 150 %, al menos aproximadamente 200 %, al menos aproximadamente 250 %, al menos aproximadamente 300 %, al menos aproximadamente 350 %, al menos aproximadamente 400 %, al menos aproximadamente 450 %, al menos aproximadamente 500 %, al menos aproximadamente 550 %, al menos aproximadamente 600 %, al menos aproximadamente 650 %, al menos aproximadamente 700 %, al menos aproximadamente 750 %, al menos aproximadamente 800 %, al menos aproximadamente 850 %, al menos aproximadamente 900 %, al menos aproximadamente 950 %, al menos aproximadamente 1000 %, al menos aproximadamente 1500 %, al menos aproximadamente 2000 %, al menos aproximadamente 2500 %, al menos aproximadamente 3000 %, al menos aproximadamente 3500 %, al menos aproximadamente 4000 %, al menos aproximadamente 4500 %, al menos aproximadamente 5000 %, al menos aproximadamente 5500 %, al menos aproximadamente 6000 %, al menos aproximadamente 7000 %, al menos aproximadamente 8000 %, al menos aproximadamente 9000 %, al menos aproximadamente 10.000 % respecto a los niveles de FVIII circulantes fisiológicamente normales. En una modalidad, se incrementa la actividad de FVIII en el plasma después de la administración en al menos aproximadamente 3000 y aproximadamente 5000 % con respecto a los niveles de FVIII circulantes fisiológicamente normales. En algunas modalidades, 24 horas después de la administración de un plásmido que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descrito en la presente, o 21 días después de la administración de un vector lentiviral o AAV que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descritos en la presente, se incrementa la actividad del FVIII en el plasma en al menos aproximadamente 6 veces con respecto a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas modalidades, 24 horas después de la administración de un plásmido que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descrito en la presente, o 21 días después de la administración de un vector lentiviral o AAV que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descritos en la

presente, se incrementa la actividad del FVIII en el plasma en al menos aproximadamente 10 veces con respecto a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas modalidades, 24 horas después de la administración de un plásmido que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descrito en la presente, o 21 días después de la administración de un vector lentiviral o AAV que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descritos en la presente, se incrementa la actividad del FVIII en el plasma en al menos aproximadamente 23 veces con respecto a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas modalidades, 24 horas después de la administración de un plásmido que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descrito en la presente, o 21 días después de la administración de un vector lentiviral o AAV que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descritos en la presente, se incrementa la actividad del FVIII en el plasma en al menos aproximadamente 18 veces con respecto a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas modalidades, 24 horas después de la administración de un plásmido que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descrito en la presente, o 21 días después de la administración de un vector lentiviral o AAV que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descritos en la presente, se incrementa la actividad del FVIII en el plasma en al menos aproximadamente 30 veces con respecto a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas modalidades, 24 horas después de la administración de un plásmido que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descrito en la presente, o 21 días después de la administración de un vector lentiviral o AAV que comprende un gen optimizado con

codones que codifica polipéptidos con actividad del factor VIII (FVIII) descritos en la presente, se incrementa la actividad del FVIII en el plasma en al menos aproximadamente 50 veces con respecto a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un

5 vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas modalidades, 24 horas después de la administración de un plásmido que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descrito en la presente, o 21 días después de la administración de un vector lentiviral

10 o AAV que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII) descritos en la presente, se incrementa la actividad del FVIII en el plasma en al menos aproximadamente 100 veces con respecto a un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16, un vector que comprende la molécula de ácido

15 nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia.

[0403] La descripción se refiere también a un método para tratar, mejorar o evitar un trastorno hemostático en un sujeto que comprende administrar una cantidad

20 terapéuticamente eficaz de una molécula de ácido nucleico aislada de la descripción o un polipéptido que tiene actividad de FVIII codificado por la molécula de ácido nucleico de la descripción. El tratamiento, la mejora y la prevención mediante la molécula de ácido nucleico aislada o el polipéptido codificado puede ser una terapia con bypass. En sujeto que recibe la terapia con bypass ya puede haber desarrollado

25 un inhibidor para un factor de coagulación, por ejemplo, FVIII, o está sujeto a desarrollar un inhibidor de factor de coagulación.

[0404] Las moléculas de ácido nucleico, vectores o polipéptidos de la descripción tratan o previenen un trastorno hemostático promoviendo la formación de un coágulo

30 de fibrina. El polipéptido que tiene actividad de FVIII codificado por la molécula de ácido nucleico de la descripción puede activar un miembro de una cascada de coagulación. El factor de coagulación puede participar de la vía extrínseca, la vía intrínseca o ambas.

[0405] Las moléculas de ácido nucleico, vectores o polipéptidos de la descripción se pueden utilizar para tratar los trastornos hemostáticos conocidos por ser tratables con el FVIII. Los trastornos hemostáticos que se pueden tratar usando métodos de la

descripción incluyen, de modo no taxativo, hemofilia A, hemofilia B, enfermedad de von Willebrand, deficiencia del factor XI (deficiencia en PTA), deficiencia del factor XII, así como también deficiencias o anomalías estructurales en fibrinógeno, protrombina, factor V, factor VII, factor X o factor XIII, hemartrosis, hemorragia muscular, sangrado oral, hemorragia, hemorragia en los músculos, hemorragia oral, trauma capitis, sangrado gastrointestinal, hemorragia intracraneal, hemorragia intra-abdominal, hemorragia intratorácica, fractura ósea, sangrado del sistema nervioso central, sangrado en el espacio retrofaríngeo, sangrado en el espacio retroperitoneal y sangrado en la vaina de iliopsoas. Las composiciones para su administración a un sujeto incluyen moléculas de ácido nucleico que comprenden una secuencia de nucleótidos optimizada de la descripción que codifica un factor de coagulación del FVIII (para las aplicaciones de tratamientos genéticos) así como también moléculas de polipéptidos de FVIII.

[0406] En algunas modalidades, el trastorno hemostático es un trastorno heredado. En una modalidad, el sujeto tiene hemofilia A. En otras modalidades, el trastorno hemostático es el resultado de una deficiencia en el FVIII. En otra modalidad, el trastorno hemostático puede ser el resultado de un factor de coagulación del FVIII deficiente.

[0407] En otra modalidad, el trastorno hemostático puede ser un trastorno adquirido. El trastorno adquirido puede ser el resultado de una enfermedad o afección secundaria subyacente. La afección no relacionada puede ser, como ejemplo de modo no taxativo, cáncer, una enfermedad autoinmunitaria o embarazo. El trastorno adquirido puede ser el resultado de edad avanzada o de medicación para tratar un trastorno secundario subyacente (por ejemplo, quimioterapia contra el cáncer).

[0408] La descripción también se refiere a métodos para tratar a un sujeto que no tiene un trastorno hemostático o una enfermedad o afección secundaria que da como resultado la adquisición de un trastorno hemostático. La descripción se refiere, por lo tanto, a un método para tratar a un sujeto que necesita un agente hemostático general que comprende administrar una cantidad terapéuticamente eficaz de al menos una molécula de ácido nucleico aislada, vector o polipéptido del FVIII de la descripción. Por ejemplo, en una modalidad, el sujeto que necesita un agente hemostático general se ha sometido o está por someterse a cirugía. La molécula de ácido nucleico aislada, vector o polipéptido de FVIII de la descripción se puede administrar antes o luego de una cirugía como profiláctico. La molécula de ácido

nucleico aislada, vector o polipéptido de FVIII de la descripción puede administrarse durante o después de la cirugía para controlar un episodio hemorrágico agudo. La cirugía puede incluir, de modo no taxativo, trasplante de hígado, resección de hígado o trasplante de células madre.

5

[0409] En otra modalidad, la molécula de ácido nucleico aislada, vector o polipéptido de FVIII de la descripción puede usarse para tratar a un sujeto que tiene un episodio de sangrado agudo y que no tiene un trastorno hemostático. El episodio hemorrágico agudo puede ser el resultado de un traumatismo grave, por ejemplo, cirugía, un accidente automovilístico, heridas, laceración por disparo o cualquier otro evento traumático que ocasione un sangrado incontrolado.

10

[0410] La molécula de ácido nucleico aislada, vector o proteína de FVIII puede usarse para tratar de manera profiláctica a un sujeto con un trastorno hemostático.

15

La molécula de ácido nucleico aislada, vector o proteína de FVIII puede usarse para tratar un episodio de sangrado agudo en un sujeto con un trastorno hemostático.

[0411] En otra modalidad, la expresión de la proteína de FVIII mediante la administración de la molécula de ácido nucleico aislada o vector de la descripción no induce una respuesta inmunitaria en un sujeto. En algunas modalidades, la respuesta inmunitaria comprende el desarrollo de anticuerpos frente al FVIII. En algunas modalidades, la respuesta inmunitaria es una secreción de citocinas. En algunas modalidades, la respuesta inmunitaria comprende la activación de linfocitos B, linfocitos T o ambos linfocitos B y linfocitos T. En algunas modalidades, la respuesta inmunitaria es una respuesta inmunitaria inhibidora, donde la respuesta inmunitaria en el sujeto reduce la actividad de la proteína de FVIII con respecto a la actividad del FVIII en un sujeto que no ha desarrollado una respuesta inmunitaria. En determinadas modalidades, la expresión de la proteína de FVIII mediante la administración de la molécula de ácido nucleico aislada o vector de la descripción evita una respuesta inmunitaria inhibidora frente a la proteína de FVIII o la proteína de FVIII expresada a partir de la molécula de ácido nucleico aislada o el vector.

20

25

30

[0412] En algunas modalidades, una composición de molécula de ácido nucleico aislada, vector o proteína de FVIII de la descripción se administra en combinación con al menos un agente adicional que promueve la hemostasia. Dicho agente adicional que promueve la hemostasia es un producto terapéutico con actividad coagulante demostrada. Como ejemplo, pero de modo no taxativo, el agente

35

hemostático puede incluir el factor V, factor VII, factor IX, factor X, factor XI, factor XII, factor XIII, protrombina o fibrinógeno o formas activadas de cualquiera de los anteriores. El factor de coagulación o agente hemostático también puede incluir fármacos antifibrinolíticos, por ejemplo, ácido épsilon aminocaproico, ácido tranexámico.

[0413] En una modalidad de la descripción, la composición (por ejemplo, la molécula de ácido nucleico aislada, vector o polipéptido del FVIII) es una en la que el FVIII está presente en forma activable cuando se administra a un sujeto. Dicha molécula activable puede activarse *in vivo* en el sitio de la coagulación después de la administración a un sujeto.

[0414] La molécula de ácido nucleico aislada, vector, o polipéptido del FVIII puede administrarse de forma intravenosa, subcutánea, intramuscular o a través de cualquier superficie mucosa, por ejemplo, oral, sublingual, bucal, sublingual, nasal, rectal, vaginal o por medio de una vía pulmonar. La proteína de FVIII puede implantarse dentro de un soporte sólido de biopolímero o unirse a este que permite la liberación lenta de la proteína quimérica en el sitio deseado.

[0415] Para la administración oral, la composición farmacéutica puede tener la forma de comprimidos o cápsulas preparados mediante medios convencionales. La composición también se puede preparar como un líquido, por ejemplo, un jarabe o una suspensión. El líquido puede incluir agentes de suspensión (por ejemplo, jarabe de sorbitol, derivados de celulosa o grasas comestibles hidrogenadas); agentes emulsionantes (lecitina o acacia); vehículos no acuosos (por ejemplo, aceite de almendra, ésteres oleosos, alcohol etílico o aceites vegetales fraccionados) y conservantes (por ejemplo, p-hidroxibenzoatos de metilo o propilo o ácido sórbico). Las preparaciones también pueden incluir agentes saborizantes, colorantes y endulzantes. De forma alternativa, la composición se puede presentar como un producto seco para su constitución con agua u otro vehículo adecuado.

[0416] Para la administración bucal o sublingual, la composición puede adoptar la forma de comprimidos, grageas o películas de rápida disolución, de acuerdo con los protocolos convencionales.

[0417] Para la administración mediante inhalación, el polipéptido que tiene actividad de FVIII para usar de acuerdo con la presente descripción se administran

convenientemente en forma de un aerosol en spray a partir de un paquete o nebulizador presurizados (por ejemplo, en PBS), con un propulsor adecuado, por ejemplo, diclorodifluorometano, triclorofluorometano, diclorotetrafluorometano, dióxido de carbono u otro gas adecuado. En el caso de un aerosol presurizado la dosificación unitaria se puede determinar proporcionando una válvula para administrar una cantidad medida. Las cápsulas y cartuchos de, por ejemplo, gelatina para usarse en un inhalador o insuflador se pueden formular de forma que contengan una mezcla del compuesto y una base de polvo adecuada tal como lactosa o almidón.

10

[0418] En una modalidad, la vía de administración de la molécula de ácido nucleico aislada, vector, o polipéptido de FVIII es parenteral. El término parenteral, según se emplea en la presente, incluye la administración intravenosa, intraarterial, intraperitoneal, intramuscular, subcutánea, rectal o vaginal. Se prefiere la forma intravenosa de la administración parenteral. Mientras que todas estas formas de administración están claramente comprendidas dentro del alcance de la invención, una forma de administración sería una solución para inyección, en particular para inyección o goteo intravenoso o intraarterial. Frecuentemente, una composición farmacéutica adecuada para inyección puede comprender un amortiguador (por ejemplo, amortiguador de acetato, fosfato o citrato), un tensioactivo (por ejemplo, polisorbato), opcionalmente un agente estabilizante (por ejemplo, albúmina humana), etc. Sin embargo, en otros métodos compatibles con las enseñanzas de la presente, la molécula de ácido nucleico aislada, vector o polipéptido de FVIII pueden administrarse directamente en el sitio de la población celular adversa, aumentando de esta forma la exposición del tejido enfermo al agente terapéutico.

[0419] Las preparaciones para la administración parenteral incluyen soluciones, suspensiones y emulsiones acuosas o no acuosas estériles. Algunos ejemplos de disolventes no acuosos son el propilenglicol, polietilenglicol, aceites vegetales como aceite de oliva y ésteres orgánicos inyectables como el oleato de etilo. Los portadores acuosos incluyen agua, soluciones alcohólicas/acuosas, emulsiones o suspensiones, incluido medios salinos y amortiguados. En la descripción en cuestión, los portadores farmacéuticamente aceptables incluyen, de modo no taxativo, amortiguador de fosfato 0,01-0,1M y preferentemente 0,05M o salino al 0,8 %. Otros vehículos parenterales comunes incluyen las soluciones de fosfato de sodio, dextrosa de Ringer, dextrosa y cloruro de sodio, Ringer lactado o aceites fijos. Los vehículos intravenosos incluyen regeneradores de fluidos y nutrientes,

30

35

regeneradores de electrolitos, tales como aquellos basados en dextrosa de Ringer y similares. Los conservantes y otros aditivos también pueden estar presentes como, por ejemplo, antimicrobianos, antioxidantes, agentes quelantes y gases inertes y similares.

5

[0420] Más particularmente, las composiciones farmacéuticas adecuadas para el uso inyectable incluyen dispersiones o soluciones acuosas estériles (en casos en que sean solubles en agua) y polvos estériles para la preparación extemporánea de dispersiones o soluciones inyectables estériles. En dichos casos, la composición debe ser estéril y fluida en la medida en que exista la posibilidad de colocarla en una jeringa fácilmente. Debe ser estable en las condiciones de fabricación y almacenamiento y preferentemente, debe conservarse contra la acción contaminante de microorganismos, tales como bacterias y hongos. El portador puede ser un solvente o un medio de dispersión que contenga, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido, y similares) y mezclas adecuadas de estos. Puede mantenerse la fluidez adecuada, por ejemplo, mediante el uso de un recubrimiento tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de la dispersión y mediante el uso de tensioactivos.

[0421] La prevención de la acción de microorganismos se puede lograr mediante diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido ascórbico, timerosal y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares, polialcoholes tales como manitol, sorbitol o cloruro de sodio en la composición. La absorción prolongada de las composiciones inyectables puede provocarse mediante la inclusión en la composición de un agente que retarde la absorción, por ejemplo, monoestearato de aluminio y gelatina.

[0422] En cualquier caso, las soluciones inyectables estériles se pueden preparar mediante la incorporación de un compuesto activo (es decir, un polipéptido solo o en combinación con otros agentes activos) en la cantidad necesaria en un solvente apropiado con un ingrediente de los mencionados anteriormente o una combinación de estos, según sea necesario, seguido por una esterilización filtrada. Generalmente, las dispersiones se preparan mediante la incorporación del compuesto activo en un vehículo esterilizado que contiene un medio de dispersión básico y los otros ingredientes necesarios de los mencionados 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 liofilización, que proporcionan un polvo de un ingrediente activo más cualquier ingrediente adicional deseado a partir de una solución de estos previamente filtrada en condiciones esterilizadas. Las preparaciones para inyecciones se procesan, se colocan en
5 recipientes como ampollas, bolsas, frascos, jeringas o viales y se sellan en condiciones asépticas de acuerdo con los métodos conocidos en la técnica. Adicionalmente, las preparaciones pueden empaquetarse y venderse en forma de un kit. Dichos artículos de fabricación tendrán preferentemente etiquetas o prospectos que indiquen que las composiciones asociadas son útiles para tratar un sujeto que
10 padece o está predispuesto a padecer trastornos de coagulación.

[0423] La composición farmacéutica puede también formularse para administración rectal como supositorio o enema de retención, por ejemplo, que contiene bases de supositorio convencionales tales como manteca de cacao u otros glicéridos.

15

[0424] Las dosis eficaces de las composiciones de la presente invención para el tratamiento de las afecciones varían dependiendo de muchos factores diferentes, incluyendo medios de administración, sitio objetivo, estado fisiológico del sujeto, si el sujeto es humano o un animal, otros medicamentos administrados y si el tratamiento
20 es profiláctico o terapéutico. Generalmente, el paciente es un ser humano, pero también se pueden tratar mamíferos no humanos incluido mamíferos transgénicos. Las dosificaciones de tratamiento se pueden titular utilizando métodos de rutina conocidos por los expertos en la técnica para optimizar la seguridad y la eficacia.

[0425] Las dosificaciones para administrar un plásmido que comprende un gen optimizado con codones que codifica polipéptidos con actividad del factor VIII (FVIII), como se describe en la presente, o un polipéptido codificado por el gen optimizado con codones puede variar de 1000 ug/kg a 0,1 ng/kg de peso corporal. En una modalidad, el rango de dosificación es de 1 ug/kg a 100 ug/kg. Las dosificaciones
25 para administrar un vector lentiviral que comprende un gen optimizado con codones que codifica polipéptidos con actividad de factor VIII (FVIII), como se describe en la presente, pueden variar entre 10^3 y 10^{15} TU/kg. Las dosificaciones para administrar un vector AAV que comprende un gen optimizado con codones que codifica polipéptidos con actividad de factor VIII (FVIII), como se describe en la presente,
30 pueden variar entre 10^5 y 10^{18} TU/kg.

[0426] La molécula de ácido nucleico aislada, plásmido, vector o polipéptido de FVIII se puede administrar como una dosis única o como múltiples dosis, donde las múltiples dosis se pueden administrar de forma continua o a intervalos de tiempo específicos. Se pueden emplear ensayos *in vitro* para determinar rangos y/o regímenes de dosificación óptimos para la administración. En la técnica se conocen ensayos *in vitro* para medir la actividad de factores de coagulación. Adicionalmente, las dosis efectivas pueden extrapolarse de curvas de dosis-respuesta obtenidas de modelos animales, por ejemplo, un perro hemofílico (Mount et ál. 2002, Blood 99 (8): 2670).

10

[0427] Se pretende que los intermedios de dosis dentro de los rangos antes mencionados también estén dentro del rango de la descripción. Los sujetos pueden recibir dichas dosis diariamente, en días alternativos, semanalmente o de acuerdo con cualquier otro plan determinado a través de análisis empíricos. Cualquier tratamiento ejemplar comprende la administración en múltiples dosificaciones durante un período prolongado, por ejemplo, de al menos seis meses. En algunos métodos, dos o más polipéptidos se pueden administrar de manera simultánea, en cuyo caso la dosificación de cada polipéptido administrado se encuentra dentro de los rangos indicados.

20

[0428] La molécula de ácido nucleico aislada, vector o polipéptido de FVIII de la descripción se puede administrar en múltiples ocasiones. Los intervalos entre dosificaciones simples pueden ser semanales, mensuales o anuales. Los intervalos también pueden ser irregulares, como se indica al medir los niveles de sangre del polipéptido o antígeno modificado en el paciente. De manera alternativa, los polipéptidos se pueden administrar como una formulación de liberación sostenida, en cuyo caso se necesita una administración menos frecuente. La dosificación y frecuencia varía dependiendo de la semivida del polipéptido o polinucleótido en el paciente.

30

[0429] La dosificación y la frecuencia de administración puede variar según si el tratamiento es profiláctico o terapéutico. En aplicaciones profilácticas, las composiciones que contienen la molécula de ácido nucleico aislada, vector o polipéptido de FVIII, o un cóctel de estos se administran a un paciente que ya no se encuentra en estado de enfermedad para potenciar la resistencia del paciente o minimizar los efectos de la enfermedad. Dicha cantidad se define como una «dosis eficaz profiláctica». Se administra una dosificación relativamente baja en intervalos

35

relativamente poco frecuentes durante un largo período. Algunos pacientes continúan recibiendo tratamiento por el resto de sus vidas.

5 [0430] La molécula de ácido nucleico aislada, vector o polipéptidos de FVIII de la descripción pueden opcionalmente administrarse en combinación con otros agentes que son eficaces para tratar el trastorno o afección que necesita tratamiento (por ejemplo, profiláctico o terapéutico).

10 [0431] Tal como se usa en la presente, la administración de moléculas de ácido nucleico aisladas, vectores o polipéptidos de FVIII de la descripción junto o en combinación con una terapia adjunta significa la administración o aplicación secuencial, simultánea, coextensiva, concurrente, concomitante o contemporánea de la terapia y los polipéptidos descritos. Los expertos en la técnica apreciarán que la administración o aplicación de los varios componentes del régimen terapéutico
15 combinado se puede programar para mejorar la eficacia general del tratamiento. Un experto en la técnica (por ejemplo, un médico) será capaz de discernir fácilmente regímenes terapéuticos combinados eficaces sin experimentación indebida en función de la terapia adjunta seleccionada y las enseñanzas de la presente memoria descriptiva.

20

[0432] También se observará que la molécula de ácido nucleico aislada, vector o polipéptido de FVIII de la presente descripción pueden usarse junto o en combinación con un agente o agentes (por ejemplo, para proporcionar una pauta posológica combinada). Los agentes de ejemplo con los que se puede combinar un
25 polipéptido de la descripción incluyen agentes que representan el estándar actual de atención médica para un trastorno particular que se esté tratando. Dichos agentes pueden ser de naturaleza química o biológica. El término «biológico/a» o «agente biológico» se refiere a cualquier agente farmacéuticamente activo preparado a partir de organismos vivos y/o sus productos que se indica para su uso como producto
30 terapéutico.

[0433] La cantidad de agente a ser usado en combinación con los polinucleótidos o polipéptidos de la presente invención puede variar según el sujeto o puede administrarse de acuerdo con lo que se conoce en la técnica. Véase, por ejemplo,
35 Bruce A Chabner *et ál.*, *Antineoplastic Agents*, en GOODMAN & GILMAN'S THE PHARMACOLOGICAL BASIS OF THERAPEUTICS 1233-1287 ((Joel G. Hardman *et ál.*, eds.,

9.^a ed. 1996). En otra modalidad, se administra la cantidad de dicho agente consistente con el estándar de atención médica.

5 **[0434]** Como se describió anteriormente, los polinucleótidos y polipéptidos de la presente descripción, pueden administrarse en una cantidad farmacéuticamente eficaz para el tratamiento *in vivo* de los trastornos de coagulación. En este aspecto, se observará que los polipéptidos o polinucleótidos de la descripción pueden formularse para facilitar la administración y promover la estabilidad del agente activo. Preferentemente, las composiciones farmacéuticas de acuerdo con la presente
10 descripción comprenden un portador estéril, no tóxico, farmacéuticamente aceptable tal como una solución salina fisiológica, amortiguadores no tóxicos, conservantes y similares. Por supuesto, las composiciones farmacéuticas de la presente descripción se pueden administrar en dosificaciones únicas o múltiples para proporcionar una cantidad farmacéuticamente eficaz del polipéptido.

15

[0435] Una cantidad de análisis se encuentran disponibles para evaluar la función del sistema de coagulación: análisis de tiempo de tromboplastina parcial activada (aPTT), ensayo cromogénico, ensayo ROTEM[®], análisis de tiempo de protrombina (PT) (también utilizado para determinar el INR), análisis de fibrinógenos (a menudo
20 mediante el método de Clauss), conteo de plaquetas, análisis de función de plaquetas (a menudo mediante PFA-100), TCT, tiempo de sangrado, análisis de mezclado (si una anormalidad corrige si el plasma del paciente se mezcla con plasma normal), ensayos del factor de coagulación, anticuerpos de antifosfolípidos, dímero D, análisis genético (por ejemplo, factor V Leiden, mutación de protrombina
25 G20210A), ensayo del veneno de la víbora de Russell diluido (dRVVT), análisis de función plaquetaria miscelánea, tromboelastografía (TEG o Sonoclot), tromboelastometría (TEM[®], por ejemplo, ROTEM[®]) o tiempo de lisis de euglobulina (ELT).

30 **[0436]** El análisis de aPTT es un indicador de rendimiento que mide la eficacia de ambas vías de coagulación «intrínsecas» (también referidas a la vía de activación de contacto) y comunes. Este análisis se utiliza comúnmente para medir la actividad de coagulación de factores de coagulación recombinantes comercialmente disponibles, por ejemplo, FVIII o FIX. Se utiliza junto con el tiempo de protrombina (PT), que mide
35 la vía extrínseca.

[0437] El análisis ROTEM[®] proporciona información sobre la cinética completa de la hemostasia: tiempo de coagulación, formación de coágulos, estabilidad de coágulos y lisis. Los diversos parámetros en la tromboelastometría dependen de la actividad del sistema de coagulación plasmático, función de plaquetas, fibrinólisis o varios factores que influyen en estas interacciones. Este ensayo puede proporcionar una vista completa de hemostasia secundaria.

Terapia génica

[0438] La descripción 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 descripción a un sujeto que lo necesita, donde se incrementa la expresión del polipéptido con respecto a una molécula de ácido nucleico de referencia que comprende la SEQ ID NO: 16. La descripción proporciona también un método para aumentar la expresión de un polipéptido con actividad de FVIII en un sujeto que comprende administrar un vector de la descripción a un sujeto que lo necesita, donde se incrementa la expresión del polipéptido con respecto a un vector que comprende una molécula de ácido nucleico de referencia.

[0439] Se ha explorado la terapia génica somática como un posible tratamiento para la hemofilia A. La terapia génica es un tratamiento particularmente interesante para la hemofilia debido a su potencial para curar la enfermedad a través de la producción endógena continua del FVIII después de la administración única del vector. La hemofilia A es adecuada para un enfoque de reemplazo de genes debido a que sus manifestaciones clínicas son atribuibles por completo a la falta de un producto génico único (FVIII) que circula en pequeñas cantidades (200ng/ml) en el plasma.

[0440] Una proteína de FVIII de la descripción se puede producir *in vivo* en un mamífero, por ejemplo, un paciente humano, utilizando una terapia génica para tratar una enfermedad o trastorno hemorrágico que se selecciona del grupo que consiste en un trastorno de coagulación hemorrágico, hemartrosis, sangrado muscular, sangrado oral, hemorragia, hemorragia en músculos, hemorragia oral, trauma, trauma capitis, sangrado gastrointestinal, hemorragia intracraneal, hemorragia intra-abdominal, hemorragia intratorácica, fractura ósea, sangrado del sistema nervioso central, sangrado en el espacio retrofaríngeo, sangrado en el espacio retroperitoneal y sangrado en la vaina de iliopsoas sería terapéuticamente beneficioso. En una modalidad, la enfermedad o trastorno de sangrado es hemofilia.

En otra modalidad, la enfermedad o trastorno hemorrágico es hemofilia A. Esto implica la administración de un ácido nucleico que codifica un FVIII optimizado unido operativamente a secuencias de control de expresión adecuadas. En determinada modalidad, estas secuencias se incorporan a un vector viral. Dichos vectores virales para dicha terapia génica incluyen vectores adenovirales, vectores lentivirales, vectores baculovirales, vectores del virus Epstein Barr, vectores papovavirales, vectores vaccinia virus, vectores del virus herpes simple y vectores del virus adeno-asociado (AAV). El vector viral puede ser un vector viral que carece de replicación. En otras modalidades, un vector adenoviral tiene una eliminación en su gen E1 o gen E3. En otras modalidades, las secuencias se incorporan en un vector no viral conocido por los expertos en la técnica.

[0441] Todos los distintos aspectos, modalidades y opciones descritas en la presente se pueden combinar en cualquier y todas las variaciones.

[0442] Todas las publicaciones, patentes y solicitudes de patente mencionadas en la presente memoria descriptiva se incorporan a la presente mediante esta referencia en la misma medida que si se indicara que cada publicación, patente o solicitud de patente individual se incorpora de forma específica e individual mediante referencia.

[0443] Tras haber descrito en general la presente descripción, se puede obtener una mejor comprensión mediante esta referencia con respecto a los ejemplos proporcionados en la presente. Estos ejemplos se proporcionan solamente a efectos ilustrativos y no se pretende que sean limitativos.

EJEMPLOS

Ejemplo 1. Estrategia de optimización con codones

[0444] Se crearon ocho variantes de FVIII BDD optimizadas con codones al controlar las tendencias de uso de codones, lo que incluye coFVIII-3 (SEQ ID NO: 1; figura 1A), coFVIII-4 (SEQ ID NO: 2; figura 1B), coFVIII-5 (SEQ ID NO: 70; figura 1C), coFVIII-6 (SEQ ID NO: 71; figura 1D), coFVIII-52 (SEQ ID NO: 3; figura 1E), coFVIII-62 (SEQ ID NO: 4; figura 1F), coFVIII-25 (SEQ ID NO: 5; figura 1G) y coFVIII-26 (SEQ ID NO: 6; figura 1H). Se utilizó Eugene la herramienta en línea para facilitar la optimización con codones como se describió previamente (véase Gaspar et ál.,

"EuGene: maximizing synthetic gene design for heterologous expression," *Bioinformatics* 28:2683-84 (2012)), y se monitorearon varios parámetros de uso de codones, tales como el índice de adaptación de codones (CAI) y el uso de codones sinónimos relativos (RSCU) (Tabla 5). Todas las variantes se ajustaron a un CAI $\geq$ 83 % y RSCU $\geq$ 1,63, mientras que la secuencia del FVIII con dominio B eliminado parenteral, antes de la optimización, tiene un CAI de 74 % y un RSCU de 1,12 (Tabla 5).

Tabla 5: Parámetros de optimización con codones

	FVIII BDD parent al	coF VIII- 3	coF VIII- 4	coFVI II-5	coFV III-6	coFVI II-52	coFV III-62	coFV III-25	coFV III-26
Índice de adaptación de codones (CAI; %)	74	91	97	83	83	91	91	88	88
Frecuencia de codones óptimos (FOP)	39	65	92	64	64	79	79	74	75
Contenido de GC (%)	44,10	52,10	60,80	55,7	55,9 %	58,30	58,30	57,30	57,60
Uso de codones sinónimos relativos (RSCU)	1,12	2,32	2,72	1,63	1,63	2,22	2,19	2,04	2,58
Tendencias de par de codones	0,19	0,43	0,04	0,11	0,11	0,27	0,27	0,23	0,48
Cantidad efectiva de codones	54,2	25,6	22,8	39,7	39,1	30,9	31,4	34,1	26,7

[0445] Además del aumento general en el CAI, se diseñaron ocho variantes en tres clases según la distribución del CAI a lo largo de la región de codificación, como se ilustra en la figura 2, con respecto a la secuencia del FVIII BDD no optimizada (figura 2A). La primera clase comprende variantes de FVIII BDD con una distribución

5 uniforme del CAI alto a lo largo de toda la región de codificación (véanse las figuras 2C-2F). La primera clase incluye coFVIII-3 (figura 2C), coFVIII-4 (figura 2D), coFVIII-5 (figura 2E), coFVIII-6 (figura 2F), así como también el coFVIII-1 descrito previamente (véase la publicación internacional n.º WO 2014/127215 (SEQ ID NO: 1)) (figura 2B). La segunda clase comprende variantes de BDD-FVIII con un CAI más

10 bajo en la mitad de extremo N de la secuencia de codificación y un CAI más alto en la mitad de extremo C de la secuencia de codificación (véanse las figuras 2G y 2H). La segunda clase incluye coFVIII-52 (figura 2G) y coFVIII-62 (figura 2H). La tercera clase comprende variantes de BDD-FVIII con un CAI más alto en la mitad de extremo N de la secuencia de codificación y un CAI más bajo en la mitad de extremo

15 C de la secuencia de codificación (véanse las figuras 2I y 2J). La tercera clase incluye coFVIII-25 (figura 2I) y coFVIII-26 (figura 2J).

[0446] Sin limitarnos a ninguna teoría, se especuló que un CAI más alto puede correlacionarse con la traducción de proteínas más rápida y estas tres clases

20 pueden representar diferentes tasas de síntesis de proteínas desde el comienzo hasta el final. Por ejemplo, la traducción de una región que tiene un CAI más bajo puede proseguir lentamente con respecto a la traducción de una región que tiene un CAI más alto. Si es así, la traducción de, por ejemplo, la mitad de extremo N de coFVIII-52 de coFVIII-62, que tiene un CAI más bajo, puede proceder inicialmente de

25 forma lenta seguido de la traducción más rápida de la mitad de extremo C que tiene un CAI más alto. Esto podría preferirse para el plegamiento de proteínas y la modificación postraducciona durante la traducción sin enlentecer la síntesis de proteínas general. El efecto opuesto puede observarse para las variantes coFVIII-25 y coFVIII-26, las cuales tienen un CAI más alto en la mitad de extremo N y un CAI

30 más bajo en la mitad del extremo C.

[0447] Para asegurar la estabilidad del ARNm, todas las variantes optimizadas con codones del FVIII se ajustaron para evitar una cantidad de sitios, que incluyen sitios de empalme crípticos, sitios poliA prematuros, motivos de inestabilidad del ARN

35 (ARE), y secuencias de repetición y para ajustar el contenido de GC (véase la Tabla 2).

Ejemplo 2. Clonación y expresión de variantes coFVIII a partir de un plásmido pcDNA3

[0448] Se diseñaron plásmidos de expresión que contienen las diversas variantes de FVIII para la expresión *in vivo*. El FVIII BDD no optimizado (figura 1I; SEQ ID NO: 16) y coFVIII-1 (figura 11Z; SEQ ID NO: 68) los polinucleótidos se clonaron en una estructura principal del pcDNA3 (Invitrogen), donde el promotor CMV se remplazó por un promotor ET (véase la figura 3). Los plásmidos resultantes, FVIII-311 (BDD FVIII) y FVIII-303 (coFVIII-1), conducen la expresión de FVIII BDD no optimizada y coFVIII-1, respectivamente.

[0449] La expresión *in vivo* de FVIII-311 y FVIII-303 se evaluó en ratones con Hema mediante inyección hidrodinámica de 5 µg de ADN/ratón del FVIII-303 o FVIII-311. Las muestras de plasma se recolectaron a las 24, 48 y 72 horas después de la inyección, y se determinó la actividad de FVIII mediante un ensayo cromogénico específico para FVIII.

[0450] Como se muestra en la figura 4, la actividad de FVIII en el plasma de los ratones tratados con FVIII-311 (BDD FVIII; cuadrados) fue de 74 ± 43 mU/mL a las 72 horas después de la inyección, mientras que la actividad de FVIII en plasma de ratones tratados con FVIII-303 (coFVIII-1; círculos) fue de 452 ± 170 mU/mL a las 72 horas después de la inyección (figura 4). Esto representa un aumento de aproximadamente seis veces en la expresión de coFVIII-1 con respecto al FVIII BDD no optimizado.

Ejemplo 3. Clonación y expresión de variantes de coFVIII mediante el uso de un sistema de vectores lentivirales

[0451] Para evaluar además el nivel de expresión de las variantes del FVIII BDD optimizadas con codones, las secuencias de codificación se clonaron en plásmidos lentivirales bajo el control de un promotor ET (véase Amendola *et ál.*, "Coordinate dual-gene transgenesis by lentiviral vectors carrying synthetic bidirectional promoters," *Nature Biol.* 23:108-16 (2005); publicación internacional n.º WO 2000/066759 A1). Un mapa de plásmidos de pLV-coFVIII-52 se muestra en la figura 5; y los plásmidos que contienen FVIII BDD no optimizado (LV-2116), coFVIII-1 (LV-coFVIII-1), coFVIII-3 (LV-coFVIII-3), coFVIII-4 (LV-coFVIII-4), coFVIII-5 (LV-coFVIII-5), y coFVIII-6 (LV-coFVIII-6), coFVIII-62 (LV-coFVIII-62), coFVIII-25 (LV-coFVIII-25),

y coFVIII-26 (LV-coFVIII-26) se construyeron de la misma manera, excepto que el fragmento de coFVIII-52 se reemplazó por cada secuencia de codificación indicada usando los sitios NheI y Sall (Tabla 6).

Tabla 6: Plásmidos de expresión que codifican variantes de FVIII

ID del plásmido	Descripción
FVIII-303 (coFVIII-1)	coFVIII-1 bajo promotor ET en pcDNA3
FVIII-311 (BDD FVIII)	BDD-FVIII parental bajo promotor ET en pcDNA3
LV-2116 (BDD FVIII)	BDD-FVIII parental bajo promotor ET en el plásmido lentiviral
LV-coFVIII-1	coFVIII-1 bajo promotor ET en el plásmido lentiviral
LV-coFVIII-3	coFVIII-3 bajo promotor ET en el plásmido lentiviral
LV-coFVIII-4	coFVIII-4 bajo promotor ET en el plásmido lentiviral
LV-coFVIII-5	coFVIII-5 bajo promotor ET en el plásmido lentiviral
LV-coFVIII-6	coFVIII-6 bajo promotor ET en el plásmido lentiviral
LV-coFVIII-52	coFVIII-52 bajo promotor ET en el plásmido lentiviral
LV-coFVIII-62	coFVIII-62 bajo promotor ET en el plásmido lentiviral
LV-coFVIII-25	coFVIII-25 bajo promotor ET en el plásmido lentiviral
LV-coFVIII-26	coFVIII-26 bajo promotor ET en el plásmido lentiviral

5

[0452] Se evaluaron las variantes de FVIII optimizadas con codones lentivirales en ratones con Hema mediante inyección hidrodinámica a una dosis de 5 µg de ADN/ratón (figuras 6A, 6B) o 20 µg de ADN/ratón (figura 6C). Como se muestra en la figura 6, cada uno de coFVIII-3 (figura 6A; triángulos), coFVIII-4 (figura 6A; triángulos invertidos), coFVIII-5 (figura 6A; diamantes), coFVIII-6 (figura 6A; círculos abiertos), coFVIII-25 (figura 6B; triángulos), coFVIII-26 (figura 6B; triángulos invertidos), coFVIII-52 (figura 6C; cuadrados) y coFVIII-62 (figura 6C; círculos rellenos) actividad de FVIII mayor mostrada que coFVIII-1 (figura 6A, círculos; figura 6B, círculos; y figura 6C, triángulos). En particular, coFVIII-25 y coFVIII-26 mostraron un nivel de expresión similar a las 72 horas después de la inyección, alcanzando aproximadamente una actividad 3 veces mayor que la del coFVIII-1 (figura 6B), lo que se traduce en actividad de FVIII 24 veces mayor en comparación con el FVIII BDD parental, no optimizado (véase la figura 4). Tanto coFVIII-52 (cuadrados) como coFVIII-62 (círculos rellenos) lograron una expresión incluso mayor a las 72 horas después de la inyección, mostrando una expresión 4 veces y 6 veces mayor, respectivamente, que coFVIII-1 (triángulos) y una expresión 30 veces y 50 veces mayor, respectivamente, que el FVIII BDD parental, no optimizado (círculos abiertos) (figura 6C). Estos datos indican que la combinación de un CAI más bajo en la mitad

10

15

20

de extremo N de la secuencia de codificación y un CAI más alto en la mitad de extremo C de la secuencia de codificación puede ser más beneficioso para la expresión de FVIII en comparación con la distribución inversa del CAI.

5 Ejemplo 4: Expresión lentiviral a largo plazo de variantes de FVIII optimizadas con codones en ratones con HemA

[0453] Se evaluaron las variantes identificadas para conducir la expresión alta del FVIII en ratones con HemA a las 72 horas después de la inyección hidrodinámica para la expresión de FVIII a largo plazo mediante transferencia genética mediada por vectores lentivirales. Los vectores lentivirales se produjeron en células 293T mediante transfección transitoria y se concentraron mediante ultracentrifugación a aproximadamente 5E9 TU/ml. Luego los vectores lentivirales se administraron en ratones con HemA de 12-14 días de edad mediante inyección retro-orbital a una dosis de 1E8 TU/ratón. A los 21 días después de la inyección lentiviral, la actividad de FVIII en el plasma promedio fue de aproximadamente 0,04 IU/ml para ratones a los que se les inyectó LV-2116 (BDD FVIII; figura 7). Cada uno de coFVIII-1, coFVIII-5, coFVIII-52, coFVIII-6 y coFVIII-62 provocó un nivel de FVIII circulante mayor a los 21 días después de la inyección con respecto al control LV-2116 (FVIII con dominio B eliminado no optimizado). En particular, la inyección de coFVIII-1 y coFVIII-5 produjo niveles de actividad de FVIII en el plasma de aproximadamente 1,8 IU/mL, coFVIII-52 produjo un nivel de actividad del FVIII en el plasma de aproximadamente 4,9 IU/mL, coFVIII-6 produjo niveles de actividad de FVIII en el plasma de aproximadamente 4,6 IU/mL, y coFVIII-62 produjo un nivel de actividad de FVIII en el plasma de aproximadamente 2,5 IU/mL a los 21 días después de la inyección (figura 7). Los niveles de FVIII en el plasma observados en ratones a los que se les inyectó LV-coFVIII-6 y LV-coFVIII-52, 4,6 IU/ml y 4,9 IU/ml, respectivamente, son 100 veces más altos que los niveles plasmáticos observados en ratones a los que se les inyectó el control LV-2116 (BDD-FVIII no optimizado).

30

Ejemplo 5. Construcciones de fusión de coFVIII-XTEN

[0454] Se probó la capacidad del XTEN de mejorar la expresión de FVIII de estado estacionario. En primer lugar, se insertó la secuencia de codificación de un XTEN de 144 aminoácidos («XTEN₁₄₄»; SEQ ID NO: 18) en el nucleótido 1193 (o después de los primeros 764 aminoácidos del polipéptido codificado) de coFVIII-52 y coFVIII-1 para generar coFVIII-52-XTEN (figura 8A; SEQ ID NO: 19) y coFVIII-1-XTEN (figura

35

8B; SEQ ID NO: 20), respectivamente. Luego la secuencia de coFVIII-1-XTEN se clonó en la estructura principal de pcDNA3 (Invitrogen) bajo el control de un promotor ET, como se describió anteriormente, para crear el plásmido de expresión de FVIII-306; y la secuencia de coFVIII-52-XTEN se clonó en un plásmido lentiviral bajo el control de un promotor ET, como se describió anteriormente, para crear el pLV-coFVIII-52-XTEN (figura 9). Se administró FVIII-306 (coFVIII-1-XTEN) a los ratones con Hema a 5 µg de ADN/ratón mediante inyección hidrodinámica. En comparación con FVIII-303 (coFVIII-1; figura 10A, círculos pequeños) y FVIII-311 (BDD FVIII; figura 10A, cuadrado), la fusión de XTEN₁₄₄ a coFVIII-1 (FVIII-306; figura 10A, círculos grandes) provocó una expresión de FVIII 5 veces y 33 veces mayor, respectivamente, en ratones Hema a las 72 horas después de la inyección. El efecto de la inserción de XTEN en la expresión de FVIII se evaluó también usando el vector lentiviral en ratones con Hema (figura 10B). Se administró LV-coFVIII-52-XTEN a los ratones con Hema de 12-14 días de edad a 1E8 TU/ratón mediante inyección retro-orbital. En comparación con LV-coFVIII52 y LV-2116 (BDD-FVIII), fusión de XTEN₁₄₄ a coFVIII-52 (figura 10B) provocó una expresión de FVIII 4 veces y 450 veces mayor, respectivamente, en ratones con Hema a 21 días después de la inyección.

[0455] Los vectores lentivirales que comprenden cada uno de coFVIII-3, co-FVIII-4, coFVIII-5, coFVIII-6, coFVIII-62, coFVIII-25, y coFVIII-26 fusionados a XTEN₁₄₄ y fusionados a un promotor ET se elaborarán como se describió anteriormente. Los vectores se probarán para determinar su expresión de proteínas de FVIII.

Ejemplo 6. Expresión de construcciones coFVIII

[0456] Se clonaron las variantes de FVIII optimizadas con codones en plásmidos lentivirales, como se ilustra en la figura 9, mediante técnicas de clonación molecular estándar. Los vectores lentivirales se produjeron luego en células HEK293 mediante transfección transitoria y se aislaron mediante ultracentrifugación.

[0457] Se administraron vectores lentivirales de FVIII a crías de ratones con Hema de 14 días de edad mediante inyección intravenosa a una dosis de 1,5E10 TU/kg de variante de FVIII. Se midió la actividad de FVIII en el plasma a los 21 días después del tratamiento con LV-FVIII y se midió la cantidad de copias del vector (VCN) por célula en las muestras de necropsia del hígado recolectadas de animales tratados con LV-FVIII en el día 150 después del tratamiento con LV-FVIII. Mientras que los valores VCN fueron similares en todos los animales independientemente de las

variantes de LV-FVIII administradas (figura 12B), los niveles de actividad de FVIII en animales tratados con variantes de coFVIII resultaron 30 a 100 veces más altos que en los animales tratados con wtBDD-FVIII (figuras 12A y 12C; Tabla 7). Estos datos indican que la optimización con codones de FVIII mejora la expresión de FVIII en una configuración de vector lentiviral.

Tabla 7: Expresión relativa de construcciones de FVIII optimizadas con codones

Variantes LV-FVIII	CoFVIII-1	CoFVIII-3	CoFVIII-4	CoFVI II-5	CoFVIII-6
Mejora en veces de la expresión de FVIII con respecto al LV-wtBDD-FVIII	57	34	33	74	107
Variantes LV-FVIII	CoFVIII-52	CoFVIII-62	CoFVIII-25	CoFVI II-26	
Mejora en veces de la expresión de FVIII con respecto al LV-wtBDD-FVIII	96	59	87	83	

Ejemplo 7. Respuesta inmunitaria mediada por expresión transgénica de FVIII en ratones con HemA después del tratamiento lentiviral

[0458] Los ratones tratados con LV-FVIII del Ejemplo 6 se evaluaron para determinar la expresión de FVIII a largo plazo y la formación de anticuerpos anti-FVIII. La expresión de FVIII, como se evidencia mediante actividad de FVIII en el plasma, varió entre animales dentro del mismo grupo de tratamiento (figura 13A). Por ejemplo, tres ratones (designados 1, 2 y 3) tratados con un vector lentiviral que expresa la variante coFVIII-5 mostraron expresión de FVIII consistente a lo largo de aproximadamente 16 semanas, mientras que tres crías normales de la misma camada (designados 4, 5 y 6), las cuales se trataron con el mismo vector lentiviral, mostraron fuertes caídas en los niveles de actividad de FVIII en el plasma en aproximadamente 10 semanas después del tratamiento (figura 13A). La actividad de FVIII en el plasma consistente observado en ratones 1, 2 y 3 se correlacionó con niveles no detectables o muy bajos de anticuerpos anti-FVIII (figura 13B; ratones 1, 2 y 3). Por el contrario, los ratones que mostraron caídas fuertes en la actividad de

FVIII en el plasma mostraron también mayores niveles de anticuerpos anti-FVIII (figura 13B; ratones 4, 5 y 6). Estos datos sugieren que la expresión del transgén de FVIII induce la formación de anticuerpos anti-FVIII en un subconjunto de animales, y los anticuerpos anti-FVIII resultantes eliminaron de circulación la proteína de FVIII transgénica.

[0459] Se evaluó la relación entre la expresión de FVIII y la formación de anticuerpos anti-FVIII. Los ratones tratados con LV-FVIII del Ejemplo 6 se dividieron en dos grupos: ratones que eran negativos para el anticuerpo anti-FVIII y ratones que eran positivos para el anticuerpo anti-FVIII. Como se muestra en la figura 14, la expresión del FVIII transgénico a niveles fisiológicos no induce una respuesta inmunitaria al FVIII transgénico (figura 14, círculos). Sin embargo, los niveles suprafisiológicos de la expresión de FVIII parecen inducir la formación del anticuerpo anti-FVIII de manera que cuanto mayor sea el nivel de expresión de FVIII, mayor será la posibilidad de inducción del anticuerpo anti-FVIII. Estos datos sugieren que puede ser beneficioso mantener los niveles fisiológicos de la expresión de FVIII en pacientes sometidos a un tratamiento de terapia génica con FVIII.

[0460] Para determinar si la respuesta inmunitaria inducida por expresión de FVIII provoca la pérdida de células hepáticas que expresan el transgén, se evaluaron la cantidad de copias del vector (figura 15) y el nivel de transcripción de ARN de FVIII (figura 16) en las muestras de necropsia del hígado de ratones negativos y positivos para el anticuerpo anti-FVIII. Como se muestra en la figura 15, la distribución de la cantidad de copias del vector fue la misma en los ratones negativos y positivos para el anticuerpo anti-FVIII, lo que indica que se conservaron las células con integración de LV-FVIII pese al desarrollo del anticuerpo anti-FVIII. Esto sugiere que la expresión del transgén de FVIII mediada con LV-FVIII no induce una respuesta de linfocitos T citotóxicos (CTL) frente a las células hepáticas que expresan FVIII. Para confirmar además estos resultados, se evaluó la transcripción del ARN de FVIII mediante hibridación in situ del ARN (figuras 16C y 16D). En el momento de la recolección del hígado, el coFVIII-52-B de ratón no tuvo ningún FVIII circulante y un nivel alto de anticuerpos anti-FVIII (figuras 16A y 16B). Sin embargo, la señal de transcripción del ARN y la cantidad de células positivas para FVIII-RNA en el tejido del hígado del ratón coFVIII-52-B se pudieron comparar con el ratón FVIII-52-A, que tuvo aproximadamente 4 IU/ml del FVIII circulante en el momento de la necropsia. Por lo tanto, la expresión de FVIII no indujo la respuesta CTL en ratones con Hema experimentales.

Ejemplo 8. Expresión a largo plazo de FVIII en neonatos de ratones con HemA tratados con LV-FVIII

[0461] Para evaluar la eficacia de usar un sistema lentiviral para el tratamiento de
 5 pacientes pediátricos con HemA a través del direccionamiento del hígado, se les
 administró a los ratones con HemA de 2 días de edad mediante inyección en la vena
 temporal aproximadamente 1,5 E10 TU/kg de LV-coFVIII-52XTEN, LV-coFVIII6-
 XTEN, o un vector lentiviral que expresa wtBDD-FVIII. Se observó expresión de FVIII
 10 a largo plazo consistente para ambas variantes y el control, lo que demuestra que la
 cascada de expresión del FVIII integrada se mantuvo en las células hepáticas que
 se dividen de los ratones tratados (figura 17). Estos datos sugieren que el LV-FVIII
 podría utilizarse potencialmente para tratar ambos pacientes pediátricos y adultos
 con HemA.

15 Ejemplo 9. Evaluación del LV-FVIII en caninos neonatos HemA

[0462] Para evaluar adicionalmente la eficacia del LV-FVIII en modelos de animales
 más grandes, se les administró a dos caninos neonatos con HemA de una semana
 de edad (designados S3 y K4) mediante inyección intravenosa $1,3 \times 10^9$ TU/kg de
 20 LV-coFVIII-6-XTEN. Esta dosis fue 10 veces más baja que la dosis utilizada
 previamente en modelos de ratones HemA. Después de la administración de los
 vectores lentivirales, se monitoreó la actividad de FVIII en el plasma mediante un
 ensayo de coagulación de una etapa (aPTT) (figura 18) y se monitoreó la
 hemostasia de sangre entera mediante ensayo de tromboelastometría rotacional
 25 (ROTEM) (figuras 19A-19D). Antes del tratamiento con LV-FVIII, el nivel de FVIII
 para S3 fue 0,7 % del normal (figura 18). Después del tratamiento con vector
 lentiviral, el nivel de FVIII de S3 aumentó a 79 % y 103 % del normal en el día 7 y día
 14, respectivamente (figura 18). Para el K4, el nivel de FVIII de administración previa
 fue 1,4 % del normal (figura 18). Después del tratamiento con vector lentiviral, el
 30 nivel de FVIII aumentó a 22 % y 25 % del normal en el día 6 y día 14,
 respectivamente (figura 18).

[0463] Correlacionado con el nivel de FVIII, se observó el ROTEM normalizado para
 ambos animales a las 2 semanas después del tratamiento (figuras 19A-19C), lo que
 35 demuestra el beneficio terapéutico mediado por LV-FVIII. El nivel de expresión de
 FVIII terapéuticamente beneficioso logrado por LV-FVIII en el perro HemA, confirma
 el uso potencial del LV-FVIII para el tratamiento de la hemofilia A.

Ejemplo 10. Evaluación del LV-FVIII en ratones neonatos con HemA

- [0464]** La terapia génica *ex vivo* con vectores lentivirales (LV) para el reemplazo génico demostró eficacia clínica para múltiples indicaciones y con seguimiento plurianual en pacientes tratados que no muestran ninguna evidencia de tumorigénesis. El suministro sistémico de LV-FIX media la expresión de FIX persistente y es bien tolerado en modelos de animales con hemofilia. La gran capacidad de empaquetamiento, capacidad de sostener la expresión del transgén a largo plazo mediante integración génica, falta de anticuerpos (ab) anti-LV preexistentes en poblaciones humanas y perfiles *in vivo* alentadores se demostraron en entornos preclínicos y clínicos, lo que hace de LV un vehículo prometedor para el suministro de genes *in vivo*, especialmente para candidatos génicos con gran tamaño de ADNc tal como FVIII.
- [0465]** Para evaluar el potencial uso del LV-FVIII para el tratamiento de la hemofilia A (HemA), las variantes de FVIII humano (hFVIII) optimizadas con codones colocadas bajo un promotor específico para hepatocitos se construyeron en un sistema de LV que contiene múltiples copias de secuencias diana de microARN-142 para minimizar la expresión de FVIII en las células presentadoras de antígenos y reducir la probabilidad de inducir anticuerpos anti-FVIII. Los vectores de LV-hFVIII se produjeron mediante transfección transitoria de células 293T, seguido de una concentración de 1000 veces mediante ultracentrifugación y se evaluaron en modelos de ratones con HemA. Después de la administración intravenosa del LV-hFVIII, se monitoreó el nivel de hFVIII circulante mediante actividad de FVIII y ensayos de antígenos, se evaluó la eficiencia de transducción en el hígado al medir las copias de ADN de LV mediante PCR cuantitativa y ARN del transgén mediante hibridación *in situ*, se midieron los anticuerpos anti-hFVIII mediante ELISA con anticuerpos anti-hFVIII total.
- [0466]** Se observó la expresión de FVIII persistente para determinar las variantes de LV-hFVIII en ratones con HemA que se trataron en la etapa neonatal. En las unidades de transducción de $1,5 \times 10^6$ /kg dosis, el LV que codifica el hFVIII optimizado con codones (LV-cohFVIII) dio como resultado un FVIII circulante 30 a 100 veces más alto que el LV que codifica el hFVIII de tipo salvaje (figura 12C), mientras que se pudieron comparar la cantidad de copias del vector en células hepáticas y el porcentaje de células positivas para el ARN de FVIII en todos los grupos probados (figura 12B). La combinación de optimización con codones con XTEN (LV-cohFVIII-

XTEN), un polipéptido hidrofílico no estructurado que mejora presumiblemente la semivida circulante al aumentar el tamaño hidrodinámico de la carga útil, provocó una actividad de FVIII de 30-50 IU/mL en el plasma, lo que representa 3.000 a 5.000 % del nivel de FVIII circulante normal (figura 12A, figura 17). Asimismo, se detectaron solamente ab anti-hFVIII en ratones con nivel suprafisiológico de hFVIII (figura 14), pero no se observó ninguna respuesta de linfocitos T citotóxicos frente a células transducidas con LV en ratones positivos para el anticuerpo anti-hFVIII (figuras 15 y 16A-16D). Nuestro resultado apoya el desarrollo adicional del LV-FVIII para la terapia génica in vivo de la hemofilia A.

10

[0467] La descripción anterior de las modalidades específicas revelará por completo la naturaleza general de la descripción de forma que otros puedan, aplicando su conocimiento dentro de la técnica, modificar y/o adaptar fácilmente para diversas aplicaciones de dichas modalidades específicas, sin experimentación indebida, sin apartarse del concepto general de la presente descripción. Por lo tanto, dichas adaptaciones y modificaciones pretenden estar dentro del significado e intervalo de equivalentes de las modalidades descritas, con base en las enseñanzas y guías presentadas en la presente. Se debe entender que la fraseología o terminología en la presente tiene como objetivo la descripción y no la limitación, de forma tal que la terminología o fraseología de la presente memoria descriptiva se interprete en vista de las enseñanzas y guía por los expertos en la técnica.

[0468] Otras modalidades de la descripción serán evidentes para los expertos en la técnica al considerar la memoria descriptiva y la puesta en práctica de la descripción descrita en la presente. Se pretende que la memoria descriptiva y los ejemplos sean considerados únicamente como ejemplos y que las siguientes reivindicaciones indiquen el alcance y el espíritu de la descripción.

[0469] Todas las patentes y publicaciones mencionadas en la presente se incorporan la presente mediante referencia en su totalidad.

[0470] La presente solicitud reivindica el beneficio de la solicitud de patente provisional estadounidenses n.º 62/289,696, presentada el 1 de febrero de 2016 y 62/409,739, presentada el 18 de octubre de 2016, cuyas solicitudes se incorporan a la presente en su totalidad mediante esta referencia.