

VIRAL VECTORS ENCODING RECOMBINANT FVIII VARIANTS WITH INCREASED EXPRESSION FOR GENE THERAPY OF HEMOPHILIA A

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to United States Provisional Patent Application No. 62/793,058, filed January 16, 2019, the disclosure of which is hereby incorporated by reference herein in its entirety for all purposes.

BACKGROUND OF THE DISCLOSURE

[0002] Blood coagulation proceeds through a complex and dynamic biological pathway of interdependent biochemical reactions, referred to as the coagulation cascade. Coagulation Factor VIII (FVIII) is a key component in the cascade. Factor VIII is recruited to bleeding sites, and forms a Xase complex with activated Factor IX (FIXa) and Factor X (FX). The Xase complex activates FX, which in turn activates prothrombin to thrombin, which then activates other components in the coagulation cascade to generate a stable clot (reviewed in Saenko et al., Trends Cardiovasc. Med., 9:185-92 (1999); Lenting et al., Blood, 92:3983-96 (1998)).

[0003] Hemophilia A is a congenital X-linked bleeding disorder characterized by a deficiency in Factor VIII activity. Diminished Factor VIII activity inhibits a positive feedback loop in the coagulation cascade. This causes incomplete coagulation, which manifests as bleeding episodes with increased duration, extensive bruising, spontaneous oral and nasal bleeding, joint stiffness and chronic pain, and possibly internal bleeding and anemia in severe cases (Zhang et al., Clin. Rev. Allerg. Immunol., 37:114-24 (2009)).

[0004] Conventionally, hemophilia A is treated by Factor VIII replacement therapy, which consists of administering Factor VIII protein (e.g., plasma-derived or recombinantly-produced Factor VIII) to an individual with hemophilia A. Factor VIII is administered prophylactically to prevent or reduce frequency of bleeding episodes, in response to an acute bleeding episode, and/or perioperatively to manage bleeding during surgery. However, there are several undesirable features of Factor VIII replacement therapy.

[0005] First, Factor VIII replacement therapy is used to treat or manage hemophilia A, but does not cure the underlying Factor VIII deficiency. Because of this, individuals with hemophilia A require Factor VIII replacement therapy for the duration of their lives. Continuous treatment is expensive and requires the individual to maintain strict compliance, as missing only a few prophylactic doses can have serious consequences for individuals with severe hemophilia A.

[0006] Second, because Factor VIII has a relatively short half-life in vivo, conventional prophylactic Factor VIII replacement therapy requires administration every second or third day. This places a burden on the individual to maintain compliance throughout their life. While third generation "long-acting" Factor VIII drugs may reduce the frequency of administration, prophylactic Factor FVIII replacement therapy with these drugs still requires monthly, weekly, or more frequent administration in perpetuity. For example, prophylactic treatment with ELOCTATE™ [Antihemophilic Factor (Recombinant), Fc Fusion Protein] requires administration every three to five days (ELOCTATE™ Prescribing Information, Biogen Idec Inc., (2015)). Moreover, the long-term effects of chemically modified biologics (e.g., pegylated polypeptides) are not yet fully understood.

[0007] Third, between 15% and 30% of all individuals receiving Factor VIII replacement therapy form anti-Factor VIII inhibitor antibodies, rendering the therapy inefficient. Factor VIII bypass therapy (e.g., administration of plasma-derived or recombinantly-produced prothrombin complex concentrates) can be used to treat hemophilia in individuals that form inhibitor antibodies. However, Factor VIII bypass therapy is less effective than Factor VIII replacement therapy (Mannucci P.M., J Thromb Haemost., 1(7):1349-55 (2003)) and may be associated with an increased risk of cardiovascular complication (Luu and Ewenstein, Haemophilia, 10 Suppl. 2:10-16 (2004)).

[0008] Somatic gene therapy holds great promise for the treatment of hemophilia A because it would remedy the underlying under-expression functional Factor VIII activity (e.g., due to missense or nonsense mutations), rather than provide a one-time dose of Factor VIII activity to the individual. Because of this difference in the mechanism of action, as compared to Factor VIII replacement therapy, one-time administration of a Factor VIII gene therapy vector may provide an individual with Factor VIII for several years, reducing the cost of treatment and eliminating the need for continued patient compliance.

[0009] Coagulation Factor IX (FIX) gene therapy has been used effectively to treat individuals with hemophilia B, a related blood coagulation condition characterized by diminished Factor IX activity (Manno C.S., et al., Nat Med., 12(3):342-47 (2006)). However, Factor VIII gene therapy presents several unique challenges. For example, the full-length, wild-type Factor VIII polypeptide (2351 amino acids; UniProt accession number P00451) is five times larger than the full-length, wild-type Factor IX polypeptide (461 amino acids; UniProt accession number P00740). As such, the coding sequence of wild-type Factor VIII is 7053 base pairs, which is too large to be packaged in conventional AAV gene therapy vectors. Further, reported recombinant expression of B-domain deleted variants of Factor VIII (BDD-FVIII) has been poor. As such, several groups have attempted to alter the codon usage of BDD-FVIII constructs, with limited success.

SUMMARY

[0010] Accordingly, there is a need for Factor VIII variants whose coding sequences are more efficiently packaged into, and delivered via, gene therapy vectors. There is also a need for synthetic, codon-altered nucleic acids which express Factor VIII more efficiently. There is also a need for codon-altered nucleic acids encoding Factor VIII polypeptides with improved folding properties, improved secretion from expressing cells, increased activity, and/or improved circulating half-life in vivo, as compared to wild-type Factor VIII or wild-type B-domain deleted Factor VIII. Such Factor VIII variants and codon-altered nucleic acids allow for improved treatment of Factor VIII deficiencies (e.g., hemophilia A). The above deficiencies and other problems associated with the treatment of Factor VIII deficiencies (e.g., hemophilia A) are reduced or eliminated by the disclosed codon-altered Factor VIII variants.

[0011] In one aspect, the present disclosure refers to a nucleic acid composition comprising a Factor VIII polynucleotide encoding a Factor VIII protein, said Factor VIII polynucleotide having the nucleic acid sequence of CS12-FL-NA (SEQ ID NO:1). For example, nucleic acid compositions (e.g., codon-altered polynucleotides) encoding Factor VIII variants are described. In some examples, the nucleic acid compositions include polynucleotides with high sequence identity to the CS04 (SEQ ID NO:37) or CS12 (SEQ ID NO:1) sequences encoding Factor VIII variants, as described herein. In some examples, the nucleic acid compositions described herein provide increased Factor VIII expression and/or increased Factor VIII activity in the blood of an animal relative to wild-type Factor VIII coding sequences and/or other codon-optimized Factor VIII coding sequences. In some examples, the nucleic acid compositions also allow for increased production of AAV-based gene therapy virions. In some examples, the nucleic

acid compositions described herein have decreased GC content and or include fewer CpG dinucleotides, as compared to wild-type sequences encoding Factor VIII. In some examples, the Factor VIII variant encoded by the nucleic acid compositions is secreted into the blood more effectively, in vivo, and/or has an increased circulating half-life in the blood, in vivo, relative to wild-type Factor VIII and/or other Factor VIII variants.

[0012] In some examples, a nucleic acid composition includes a polynucleotide encoding a Factor VIII polypeptide having an amino acid sequence of CS12-FL-AA (SEQ ID NO:2), where the polynucleotide has a sequence with at least 95% sequence identity (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% sequence identity) to a CS12-FL-NA (SEQ ID NO:1).

[0013] In some examples, the nucleic acid composition further includes a promoter polynucleotide operatively linked to the Factor VIII polynucleotide, wherein the promoter polynucleotide has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to hTTR (SEQ ID NO:6). In some examples, the promoter is directly attached to the Factor VIII polynucleotide, e.g., as depicted in Figure 11.

[0014] In some examples, the nucleic acid composition further includes a liver-specific element operatively linked to the Factor VIII polynucleotide. In some examples, the liver-specific element is an enhancer element with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100%) to CRM8 (SEQ ID NO:5). In some examples, the nucleic acid composition includes two such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the nucleic acid composition includes three such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the one or more liver-specific elements and the promoter are directly attached, e.g., as depicted in Figure 11.

[0015] In some examples, the nucleic acid composition has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vr (SEQ ID NO:3).

[0016] In some examples, the acid composition has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vrp (SEQ ID NO:10).

[0017] In one aspect, mammalian gene therapy vectors that include a nucleic acid that encodes a Factor VIII variant are described. In some examples, the nucleic acid that encodes the Factor VIII includes a polynucleotide with high sequence identity to the CS04 (SEQ ID NO:37) or CS12 (SEQ ID NO:1) sequences encoding Factor VIII variants, as described herein. In some examples, the mammalian gene

therapy vectors described herein provide increased Factor VIII expression and/or increased Factor VIII activity in the blood of an animal relative to gene therapy vectors that include a natively encoded Factor VIII variant polynucleotide or other codon-optimized Factor VIII variant polynucleotides. In some examples, the mammalian gene therapy vectors described herein encode for a Factor VIII variant protein that is secreted into the blood more effectively, in vivo, and/or has an increased circulating half-life in the blood, in vivo, relative to wild-type Factor VIII and/or other Factor VIII variants.

[0018] In some examples, the mammalian gene therapy vector includes a polynucleotide encoding a Factor VIII polypeptide having an amino acid sequence of CS12-FL-AA (SEQ ID NO:2), where the polynucleotide has a sequence with at least 95% sequence identity (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% sequence identity) to a CS12-FL-NA (SEQ ID NO:1).

[0019] In some examples, the mammalian gene therapy vector further includes a promoter polynucleotide operatively linked to the Factor VIII polynucleotide, wherein the promoter polynucleotide has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to hTTR (SEQ ID NO:6). In some examples, the promoter is directly attached to the Factor VIII polynucleotide, e.g., as depicted in Figure 11.

[0020] In some examples, the mammalian gene therapy vector further includes a liver-specific element operatively linked to the Factor VIII polynucleotide. In some examples, the liver-specific element is an enhancer element with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100%) to CRM8 (SEQ ID NO:5). In some examples, the nucleic acid composition includes two such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the nucleic acid composition includes three such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the one or more liver-specific elements and the promoter are directly attached, e.g., as depicted in Figure 11.

[0021] In some examples, the mammalian gene therapy vector has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vr (SEQ ID NO:3).

[0022] In some examples, the mammalian gene therapy vector is included in a plasmid having a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vrp (SEQ ID NO:10).

[0023] In one aspect, the present disclosure refers to an adeno-associated virus (AAV) particle comprising capsid proteins encapsulating a nucleic acid composition as disclosed herein. In one example, adeno-associated virus (AAV) particles that include a nucleic acid that encodes a Factor VIII variant are

described. In some examples, the nucleic acid that encodes the Factor VIII includes a polynucleotide with high sequence identity to the CS04 (SEQ ID NO:37) or CS12 (SEQ ID NO:1) sequences encoding Factor VIII variants, as described herein. In some examples, the AAV particles described herein provide increased Factor VIII expression and/or increased Factor VIII activity in the blood of an animal relative to AAV particles that include a natively encoded Factor VIII variant polynucleotide or other codon-optimized Factor VIII variant polynucleotides. In some examples, the AAV particles described herein encode for a Factor VIII variant protein that is secreted into the blood more effectively, in vivo, and/or has an increased circulating half-life in the blood, in vivo, relative to wild-type Factor VIII and/or other Factor VIII variants.

[0024] In some examples, the nucleic acid contained in the AAV particles includes a polynucleotide encoding a Factor VIII polypeptide having an amino acid sequence of CS12-FL-AA (SEQ ID NO:2), where the polynucleotide has a sequence with at least 95% sequence identity (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% sequence identity) to a CS12-FL-NA (SEQ ID NO:1).

[0025] In some examples, the nucleic acid contained in the AAV particles further includes a promoter polynucleotide operatively linked to the Factor VIII polynucleotide, wherein the promoter polynucleotide has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to hTTR (SEQ ID NO:6). In some examples, the promoter is directly attached to the Factor VIII polynucleotide, e.g., as depicted in Figure 11.

[0026] In some examples, the nucleic acid contained in the AAV particles further includes a liver-specific element operatively linked to the Factor VIII polynucleotide. In some examples, the liver-specific element is an enhancer element with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100%) to CRM8 (SEQ ID NO:5). In some examples, the nucleic acid composition includes two such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the nucleic acid composition includes three such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the one or more liver-specific elements and the promoter are directly attached, e.g., as depicted in Figure 11.

[0027] In some examples, the nucleic acid contained in the AAV particles has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vr (SEQ ID NO:3).

[0028] In some examples, the nucleic acid contained in the AAV particles is produced using a plasmid having a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vrp (SEQ ID NO:10).

[0029] In one example, methods for treating hemophilia A by administering a nucleic acid composition that encodes a Factor VIII variant to a patient with hemophilia A are described. In some examples, the nucleic acid composition includes a polynucleotide with high sequence identity to the CS04 (SEQ ID NO:37) or CS12 (SEQ ID NO:1) sequences encoding Factor VIII variants, as described herein. In some examples, the methods for treating hemophilia A described herein result in increases in Factor VIII expression and/or increases in Factor VIII activity in the blood of the patient that are greater than increases in Factor VIII expression and/or increases in Factor VIII activity in the blood of a patient administered a nucleic acid composition that includes a wild type Factor VIII coding sequence and/or a different codon-optimized Factor VIII coding sequence. In some examples, the Factor VIII variant encoded by the nucleic acid composition is secreted into the blood more effectively, in vivo, and/or has an increased circulating half-life in the blood, in vivo, relative to wild-type Factor VIII and/or other Factor VIII variants.

[0030] In some examples, the administered nucleic acid composition includes a polynucleotide encoding a Factor VIII polypeptide having an amino acid sequence of CS12-FL-AA (SEQ ID NO:2), where the polynucleotide has a sequence with at least 95% sequence identity (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% sequence identity) to a CS12-FL-NA (SEQ ID NO:1).

[0031] In some examples, the administered nucleic acid composition further includes a promoter polynucleotide operatively linked to the Factor VIII polynucleotide, wherein the promoter polynucleotide has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to hTTR (SEQ ID NO:6). In some examples, the promoter is directly attached to the Factor VIII polynucleotide, e.g., as depicted in Figure 11.

[0032] In some examples, the administered nucleic acid composition further includes a liver-specific element operatively linked to the Factor VIII polynucleotide. In some examples, the liver-specific element is an enhancer element with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100%) to CRM8 (SEQ ID NO:5). In some examples, the nucleic acid composition includes two such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the nucleic acid composition includes three such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the one or more liver-specific elements and the promoter are directly attached, e.g., as depicted in Figure 11.

[0033] In some examples, the administered nucleic acid composition has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vr (SEQ ID NO:3).

[0034] In some examples, the administered nucleic acid composition has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vrp (SEQ ID NO:10).

[0035] In another aspect, the present disclosure refers to a nucleic acid composition as disclosed herein for treating hemophilia A. In another aspect, the present disclosure refers to the use of a nucleic acid composition as disclosed herein for the manufacture of a medicament for the treatment of hemophilia A. In another aspect, the present disclosure refers to a mammalian gene therapy vector as disclosed herein for treating hemophilia A. In another aspect, the present disclosure refers to the use of a mammalian gene therapy vector as disclosed herein for the manufacture of a medicament for the treatment of hemophilia A. In another aspect, the present disclosure refers to adeno-associated virus (AAV) particle as disclosed herein for treating hemophilia A. In another aspect, the present disclosure refers to the use of an adeno-associated virus (AAV) particle as disclosed herein for the manufacture of a medicament for the treatment of hemophilia A. In one example, methods for treating hemophilia A by administering a mammalian gene therapy vector that includes a nucleic acid encoding a Factor VIII variant are described. In some examples, the nucleic acid that encodes the Factor VIII includes a polynucleotide with high sequence identity to the CS04 (SEQ ID NO:37) or CS12 (SEQ ID NO:1) sequences encoding Factor VIII variants, as described herein. In some examples, the methods for treating hemophilia A described herein result in increases in Factor VIII expression and/or increases in Factor VIII activity in the blood of the patient that are greater than increases in Factor VIII expression and/or increases in Factor VIII activity in the blood of a patient administered a mammalian gene therapy vector that includes a wild type Factor VIII coding sequence and/or a different codon-optimized Factor VIII coding sequence. In some examples, the Factor VIII variant encoded by the nucleic acid within the mammalian gene therapy vector is secreted into the blood more effectively, in vivo, and/or has an increased circulating half-life in the blood, in vivo, relative to wild-type Factor VIII and/or other Factor VIII variants.

[0036] In some examples, the administered mammalian gene therapy vector includes a polynucleotide encoding a Factor VIII polypeptide having an amino acid sequence of CS12-FL-AA (SEQ ID NO:2), where the polynucleotide has a sequence with at least 95% sequence identity (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% sequence identity) to a CS12-FL-NA (SEQ ID NO:1).

[0037] In some examples, the administered mammalian gene therapy vector further includes a promoter polynucleotide operatively linked to the Factor VIII polynucleotide, wherein the promoter polynucleotide has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to hTTR (SEQ ID NO:6). In some examples, the promoter is directly attached to the Factor VIII polynucleotide, e.g., as depicted in Figure 11.

[0038] In some examples, the administered mammalian gene therapy vector further includes a liver-specific element operatively linked to the Factor VIII polynucleotide. In some examples, the liver-specific element is an enhancer element with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100%) to CRM8 (SEQ ID NO:5). In some examples, the nucleic acid composition includes two such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the nucleic acid composition includes three such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the one or more liver-specific elements and the promoter are directly attached, e.g., as depicted in Figure 11.

[0039] In some examples, the administered mammalian gene therapy vector has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vr (SEQ ID NO:3).

[0040] In some examples, the administered mammalian gene therapy vector is included in a plasmid having a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vrp (SEQ ID NO:10).

[0041] In one example, methods for treating hemophilia A by administering an adeno-associated virus (AAV) particle that includes a nucleic acid encoding a Factor VIII variant are described. In some examples, the nucleic acid that encodes the Factor VIII includes a polynucleotide with high sequence identity to the CS04 (SEQ ID NO:37) or CS12 (SEQ ID NO:1) sequences encoding Factor VIII variants, as described herein. In some examples, the methods for treating hemophilia A described herein result in increases in Factor VIII expression and/or increases in Factor VIII activity in the blood of the patient that are greater than increases in Factor VIII expression and/or increases in Factor VIII activity in the blood of a patient administered an adeno-associated virus (AAV) particle that includes a wild type Factor VIII coding sequence and/or a different codon-optimized Factor VIII coding sequence. In some examples, the Factor VIII variant encoded by the nucleic acid within the adeno-associated virus (AAV) particle is secreted into the blood more effectively, in vivo, and/or has an increased circulating half-life in the blood, in vivo, relative to wild-type Factor VIII and/or other Factor VIII variants.

[0042] In some examples, the nucleic acid contained in the administered AAV particles includes a polynucleotide encoding a Factor VIII polypeptide having an amino acid sequence of CS12-FL-AA (SEQ ID NO:2), where the polynucleotide has a sequence with at least 95% sequence identity (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% sequence identity) to a CS12-FL-NA (SEQ ID NO:1).

[0043] In some examples, the nucleic acid contained in the administered AAV particles further includes a promoter polynucleotide operatively linked to the Factor VIII polynucleotide, wherein the promoter

polynucleotide has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to hTTR (SEQ ID NO:6). In some examples, the promoter is directly attached to the Factor VIII polynucleotide, e.g., as depicted in Figure 11.

[0044] In some examples, the nucleic acid contained in the administered AAV particles further includes a liver-specific element operatively linked to the Factor VIII polynucleotide. In some examples, the liver-specific element is an enhancer element with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100%) to CRM8 (SEQ ID NO:5). In some examples, the nucleic acid composition includes two such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the nucleic acid composition includes three such liver-specific elements operatively linked to the Factor VIII polynucleotide. In some examples, the one or more liver-specific elements and the promoter are directly attached, e.g., as depicted in Figure 11.

[0045] In some examples, the nucleic acid contained in the administered AAV particles has a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vr (SEQ ID NO:3).

[0046] In some examples, the nucleic acid contained in the administered AAV particles is produced using a plasmid having a nucleic acid sequence with at least 90% sequence identity (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100%) to CS12-CRM8.2-Vrp (SEQ ID NO:10). In yet another aspect, the present disclosure refers to a method for producing an adeno-associated virus (AAV) particle comprising introducing a nucleic acid composition as disclosed herein into a eukaryotic host cell: wherein: the nucleic acid composition comprises a Factor VIII polynucleotide having the nucleic acid sequence of CS12-FL-NA (SEQ ID NO:1) flanked by a 5' inverted terminal repeat sequence (5' ITR) and a 3' inverted terminal repeat sequence (3' ITR), and the mammalian host cell comprises one or more polynucleotides encoding an AAV rep gene, an AAV cap gene, and viral replication helper genes.

BRIEF DESCRIPTION OF DRAWINGS

[0047] Figures 1A and 1B collectively show the CS12 codon-optimized nucleotide sequence (SEQ ID NO:1) encoding a Factor VIII variant in accordance with some embodiments ("CS12-FL-NA" for full-length coding sequence).

[0048] Figure 2 shows the Factor VIII variant amino acid sequence (SEQ ID NO: 2) encoded by the CS12 codon-altered nucleotide sequence in accordance with some embodiments (“CS12-FL-AA” for full-length amino acid sequence).

[0049] Figures 3A and 3B collectively show the nucleic acid sequence (SEQ ID NO:3) of the CS12-CRM8.2-Vr nucleotide-reduced gene therapy vector encoding a Factor VIII variant in accordance with some embodiments.

[0050] Figure 4 shows the nucleic acid sequence of various genetic elements useful for gene therapy vectors encoding a Factor VIII variant, including a 5'-ITR (SEQ ID NO:4), a CRM8 enhancer element (SEQ ID NO:5), a human TTR promoter (SEQ ID NO:6), a minimal Kozak sequence (SEQ ID NO:7), and synthetic poly-adenylation element (SEQ ID NO:8), and a 3'-ITR (SEQ ID NO:9), in accordance with some embodiments.

[0051] Figures 5A, 5B, and 5C collectively show the nucleic acid sequence (SEQ ID NO:10) of the CS12-CRM8.2-Vrp plasmid containing a gene therapy vector encoding a Factor VIII variant in accordance with some embodiments.

[0052] Figures 6A and 6B collectively show the nucleic acid sequence (SEQ ID NO:38) of the CS12-CRM8.2-V gene therapy vector encoding a Factor VIII variant in accordance with some embodiments.

[0053] Figures 7A and 7B collectively show the CS04 codon-optimized nucleotide sequence (SEQ ID NO:37) encoding a Factor VIII variant in accordance with some embodiments ("CS12-FL-NA" for full-length coding sequence).

[0054] Figures 8A and 8B collectively show amino acid and nucleotide sequences for exemplary glycosylation peptides that are inserted into the B-domain substituted linker of a Factor VIII variant in accordance with some embodiments. "NG1" or NG1-AA" is the code for the amino acid sequence, shown in the top line. "NG1-NA" is the code for the nucleic acid sequence, shown in the bottom line for each set. Figure 8A and 8B disclose the amino acid sequences as SEQ ID NOS 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35, and the nucleotide sequences as SEQ ID NOS 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, and 36 all, respectively, in order of appearance.

[0055] Figure 9 shows western blot analysis of Factor VIII variants, with (vNG4/CS04, vNG5/CS04, and vNG16/CS04) and without (vCS04 with SQ) glycosylation peptides engineered into the SQ linker, expressed in Huh-7 cells.

[0056] Figure 10 shows western blot analysis of Factor VIII variants, with (vNG4/CS04, vNG5/CS04, and vNG16/CS04) and without (vCS04 with SQ) glycosylation peptides engineered into the SQ linker, expressed in HepG2 cells following infection of an AAV8 gene therapy vector.

[0057] Figure 11 illustrates exemplary Factor VIII gene therapy constructs encoding a Factor VIII variant in accordance with some implementations.

[0058] Figure 12 shows Factor VIII activity levels in vivo in a “line E” hFVIII tolerant mouse model, and in vitro using HepG2 cells, post-infection with AAV8 gene therapy vectors encoding a Factor VIII variant protein in accordance with some embodiments.

[0059] Figure 13 shows agarose gel electrophoretic analysis of the vCS04, vX5/NG5/CS120, and vX5/NG5/CS12 gene therapy vectors encoding a Factor VIII variant in accordance with some embodiments.

[0060] Figure 14 shows Factor VIII activity levels in vivo in a FVIII F17 knock-in mouse model, and in vitro using HepG2 cells, a human liver cell line (ATCC #HB-8065), post-infection with AAV8 gene therapy vectors encoding a Factor VIII variant protein in accordance with some embodiments.

[0061] Figure 15 shows schematic illustrations of the wild-type and Refacto-type human Factor VIII protein constructs, as well as the Factor VIII protein encoded by the CS12 polynucleotide, which contains the X5 mutations and NG5 glycosylation peptide.

[0062] Figure 16 shows a wild-type human Factor VIII amino acid sequence (SEQ ID NO:39), in accordance with some embodiments (“FVIII-FL-AA”).

[0063] Figure 17 shows example coding sequences (SEQ ID NOS 41-53, respectively, in order of appearance) for B-domain substituted linkers in accordance with some embodiments. BDLO04 (SEQ ID NO:41) is the portion of the CS04 codon-altered nucleotide sequences that encodes the B-domain substituted linker.

[0064] Figure 18 shows an example plasmid backbone (SEQ ID NO:54) for integrating a Factor VIII gene therapy genome, in accordance with some embodiments.

[0065] Figure 19 shows an example replicon (“pMB1 Replicon” – SEQ ID NO:55) and ampicillin resistance marker (“Bla(ApR)” – SEQ ID NO:56) for a plasmid useful for integrating a Factor VIII gene therapy genome, in accordance with some embodiments.

[0066] Figures 20A, 20B, and 20C collectively show the nucleic acid sequence (SEQ ID NO:57) of the CS12-CRM8.2-Vp plasmid containing a gene therapy vector encoding a Factor VIII variant in accordance with some embodiments.

DETAILED DESCRIPTION OF DISCLOSURE

I. Introduction

[0001] AAV-based gene therapy holds great promise for the treatment of hemophiliacs. For hemophilia B, first clinical data are encouraging in that FIX levels of about 10% can be maintained in at least some patients for more than 1 year. For hemophilia A however, achieving therapeutic expression levels of 5-10% with AAV vectors remains challenging for various reasons. First, the Factor VIII coding sequence is too large for conventional AAV-based vectors. Second, engineered B-domain deleted or truncated Factor VIII constructs suffer from poor expression in vivo, even when codon-optimized. Third, these B-domain deleted or truncated Factor VIII variant constructs have short half-lives in vivo, exacerbating the effects of poor expression. Fourth, even when expressed, FVIII is not efficiently secreted from cells, as are other coagulation factors, such as Factor IX.

[0067] The present disclosure relates, in part, to the discovery of gene therapy vectors containing codon-altered Factor VIII variant coding sequences that solve these and other problems associated with Factor VIII gene therapy. For example, in some embodiments, the Factor VIII variant polynucleotides, polypeptides, and gene therapy constructs disclosed herein provide improved exogenous Factor VIII expression in mammalian cells. In some embodiments, the Factor VIII variant polynucleotides, polypeptides, and gene therapy constructs disclosed herein provide improved bioavailability (e.g., result in improved Factor VIII activity in the blood of a patient) in vivo. In some embodiments, the Factor VIII variant polynucleotides, polypeptides, and gene therapy constructs disclosed herein provide improved circulating half-life for exogenous Factor VIII in the blood of a patient. As described herein, one or more of these

advantages are realized by using any combination of one or more of the following improvements to the gene therapy system.

[0068] In some implementations, one or more of these advantages are realized by engineering the X5 mutations into a Factor VIII variant polypeptide that is encoded by a codon-altered Factor VIII polypeptide, as described herein. Advantageously, inclusion of the X5 mutations in an encoded Factor VIII polypeptide, as described herein, provides improved biopotency of exogenously expressed Factor VIII, in vivo and in vitro. For instance, as described in Example 4, inclusion of the five X5 mutations in the A1 domain of the heavy chain of a Refacto FVIII polypeptide increased in vivo exogenous Factor VIII biopotency in line E2 mice by 3-fold following administration of an AAV gene therapy vector encoding the Refacto Factor VIII-X5 variant, as compared to mice administered an otherwise identical gene therapy vector encoding a wild type Refacto Factor VIII variant (compare vX5/CS24 to vCS04 in Figure 12). Consistent with the in vivo results, inclusion of the five X5 mutations in the A1 domain of the Factor VIII heavy chain increased in vitro exogenous Factor VIII biopotency by 4-fold following infection of HepG2 cells with an AAV gene therapy vector encoding a Refacto Factor VIII-X5 variant, as compared to HepG2 cells infected with an otherwise identical gene therapy vector encoding a wild type Refacto Factor VIII variant (compare vX5/CS24 to vCS04 in Figure 12).

[0069] In some implementations, one or more of these advantages are realized by engineering an NG5 glycosylation peptide into a B-domain linker of a Factor VIII variant polypeptide that is encoded by a codon-altered Factor VIII polypeptide, as described herein. Advantageously, inclusion of the NG5 glycosylation peptide in an encoded Factor VIII polypeptide, as described herein, provides improved biopotency of exogenously expressed Factor VIII, in vivo. For instance, as described in Example 3, inclusion of the NG glycosylation peptide in the SQ-linker of a Refacto FVIII polypeptide increased in vivo exogenous Factor VIII biopotency in line E2 mice by 2-fold following administration of an AAV gene therapy vector encoding the Refacto Factor VIII-NG5 variant, as compared to mice administered an otherwise identical gene therapy vector encoding a wild type Refacto Factor VIII variant (compare vNG5/CS04 to vCS04 in Figure 12).

[0070] In some implementations, one or more of these advantages are realized by engineering both the X5 mutations and the NG5 glycosylation peptide into a Factor VIII variant polypeptide

that is encoded by a codon-altered Factor VIII polypeptide, as described herein.

Advantageously, inclusion of the X5 mutations and NG5 glycosylation peptide in an encoded Factor VIII polypeptide, as described herein, provides improved biopotency of exogenously expressed Factor VIII, in vivo and in vitro. For instance, as described in Example 5, inclusion of the five X5 mutations in the A1 domain of the heavy chain and the NG5 glycosylation peptide in the SQ-linker of a Refacto FVIII polypeptide increased in vivo exogenous Factor VIII biopotency in line E2 mice by 4.5-fold following administration of an AAV gene therapy vector encoding the Refacto Factor VIII-X5/NG5 variant, as compared to mice administered an otherwise identical gene therapy vector encoding a wild type Refacto Factor VIII variant (compare vX5/NG5/CS125 to vCS04 in Figure 12). Consistent with the in vivo results, inclusion of the five X5 mutations in the A1 domain of the heavy chain and the NG5 glycosylation peptide in the SQ-linker of a Refacto FVIII polypeptide increased in vitro exogenous Factor VIII biopotency by 3-fold following infection of HepG2 cells with an AAV gene therapy vector encoding a Refacto Factor VIII-X5/NG5 variant, as compared to HepG2 cells infected with an otherwise identical gene therapy vector encoding a wild type Refacto Factor VIII variant (compare vX5/NG5/CS125 to vCS04 in Figure 12).

[0071] In some implementations, one or more of these advantages are realized by using a human hTTR promoter and one or more liver-specific CRM8 elements upstream of the polynucleotide sequence encoding a Factor VIII variant polypeptide. Advantageously, use of the hTTR promoter and one or more liver-specific CRM8 elements provides improved biopotency of exogenously expressed Factor VIII in human cells, in vitro. For instance, as described in Example 2, use of the hTTR promoter and either one or two liver-specific CRM8 elements increased in vivo exogenous Factor VIII biopotency in HepG2 cells by about 2-fold and 4-fold, respectively, as compared to use of mouse TTR promoter and enhancer sequences (compare vCS115 and vCS116 to vCS04 in Figure 12).

[0072] In some implementations, one or more of these advantages are realized by removing extraneous nucleotides positioned between various elements of an AAV gene therapy vector encoding a Factor VIII variant protein. Advantageously, removing extraneous nucleotides between various elements provides improved biopotency of exogenously expressed Factor VIII in human cells, in vitro. For instance, as described in Example 6, removal of only 71 nucleotides from the vX5/NG5/CS120 AAV gene therapy vector encoding a Refacto Factor VIII-X5/NG5

variant improved the in vitro biopotency of the expressed Factor VIII variant by 50% (compare vX5/NG5/CS12 to vX5/NG5/CS120 in Figure 14).

[0073] In some implementations, one or more of these advantages are realized by engineering both the X5 mutations and the NG5 glycosylation peptide into a Factor VIII variant polypeptide that is encoded by a codon-altered Factor VIII polypeptide, and using a human hTTR promoter and one or more liver-specific CRM8 elements upstream of the polynucleotide sequence encoding the Factor VIII variant polypeptide. Advantageously, using this combination of improvements provides improved biopotency of exogenously expressed Factor VIII, in vivo and in vitro. For instance, as described in Example 6, use of this combination of improvements in an AAV gene therapy vector increased in vivo biopotency 14.5-fold, relative to use of an AAV gene therapy vector that includes a polynucleotide having the same codon-alteration for an encoded wild-type Refacto Factor VIII using murine TTR promoter and enhancer sequences (compare vX5/NG5/CS120 to vSC04 in Figure 14). Consistent with the in vivo results, use of this combination of improvements in an AAV gene therapy vector increased in vitro biopotency 17-fold, relative to use of an AAV gene therapy vector that includes a polynucleotide having the same codon-alteration for an encoded wild-type Refacto Factor VIII using murine TTR promoter and enhancer sequences (compare vX5/NG5/CS120 to vSC04 in Figure 14). As reported in Table 4 of WO 2017/083762 (the content of which is hereby incorporated herein by reference), the vSC04 vector provides more than 70-fold greater FVIII biopotency in vivo, relative to an equivalent gene therapy vector encoding for a Refacto Factor VIII polynucleotide using the wild-type coding sequence. Accordingly, it would be expected that the vX5/NG5/CS120 gene therapy vector would provide a 1000-fold to 1250-fold increase in FVIII biopotency, relative to use of the wild-type Refacto coding sequence.

[0074] In some implementations, one or more of these advantages are realized by engineering both the X5 mutations and the NG5 glycosylation peptide into a Factor VIII variant polypeptide that is encoded by a codon-altered Factor VIII polypeptide, using a human hTTR promoter and one or more liver-specific CRM8 elements upstream of the polynucleotide sequence encoding the Factor VIII variant polypeptide, and removing extraneous nucleotides positioned between various elements of the AAV gene therapy vector. Advantageously, using this combination of improvements provides improved biopotency of exogenously expressed Factor VIII, in vivo and in vitro. For instance, as described in Example 6, use of this combination of improvements in an

AAV gene therapy vector increased in vivo biopotency 14-fold, relative to use of an AAV gene therapy vector that includes a polynucleotide having the same codon-alteration for an encoded wild-type Refacto Factor VIII using murine TTR promoter and enhancer sequences (compare vX5/NG5/CS12 to vSC04 in Figure 14). Consistent with the in vivo results, use of this combination of improvements in an AAV gene therapy vector increased in vitro biopotency 24-fold, relative to use of an AAV gene therapy vector that includes a polynucleotide having the same codon-alteration for an encoded wild-type Refacto Factor VIII using murine TTR promoter and enhancer sequences (compare vX5/NG5/CS12 to vSC04 in Figure 14). As reported in Table 4 of WO 2017/083762 (the content of which is hereby incorporated herein by reference), the vCS04 vector provides more than 70-fold greater FVIII biopotency in vivo, relative to an equivalent gene therapy vector encoding for a Refacto Factor VIII polynucleotide using the wild-type coding sequence. Accordingly, it would be expected that the vX5/NG5/CS120 gene therapy vector would provide a 1000-fold to 1750-fold increase in FVIII biopotency, relative to use of the wild-type Refacto coding sequence.

II. Definitions

[0075] As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

[0076] As used herein, the terms "Factor VIII" and "FVIII" are used interchangeably, and refer to any protein with Factor VIII activity (e.g., active FVIII, often referred to as FVIIIa) or protein precursor (e.g., pro-protein or pre-pro-protein) of a protein with Factor IXa cofactor activity under particular conditions, e.g., as measured using the two-step chromogenic Factor X activation assay described in Chapter 2.7.4 of the European Pharmacopoeia 9.0. In an exemplary embodiment, a Factor VIII polypeptide refers to a polypeptide that has sequences with high sequence identity (e.g., at least 70%, 75%, 80%, 85%, 90%, 95%, 99%, or more) to the heavy and light chains of a wild type Factor VIII polypeptide. In some embodiments, the B-domain of a Factor VIII polypeptide is deleted, truncated, or replaced with a linker polypeptide to reduce the size of the polynucleotide encoding the Factor VIII polypeptide.

[0077] Non-limiting examples of wild type Factor VIII polypeptides include human pre-pro-Factor VIII (e.g., GenBank accession nos. AAA52485, CAA25619, AAA58466, AAA52484,

AAA52420, AAV85964, BAF82636, BAG36452, CAI41660, CAI41666, CAI41672, CAI43241, CAO03404, EAW72645, AAH22513, AAH64380, AAH98389, AAI11968, AAI11970, or AAB61261), corresponding pro-Factor VIII, and natural variants thereof; porcine pre-pro-Factor VIII (e.g., UniProt accession nos. F1RZ36 or K7GSZ5), corresponding pro-Factor VIII, and natural variants thereof; mouse pre-pro-Factor VIII (e.g., GenBank accession nos. AAA37385, CAM15581, CAM26492, or EDL29229), corresponding pro-Factor VIII, and natural variants thereof; rat pre-pro-Factor VIII (e.g., GenBank accession no. AAQ21580), corresponding pro-Factor VIII, and natural variants thereof; rat pre-pro-Factor VIII; and other mammalian Factor VIII homologues (e.g., monkey, ape, hamster, guinea pig, etc.).

[0078] As used herein, a Factor VIII polypeptide includes natural variants and artificial constructs with Factor IX cofactor activity. As used in the present disclosure, Factor VIII encompasses any natural variants, alternative sequences, isoforms, or mutant proteins that retain some basal Factor IX cofactor activity (e.g., at least 5%, 10%, 25%, 50%, 75%, or more of the corresponding wild type activity).

[0079] Specifically included within the definition of “Factor VIII” are Factor VIII variants, sometimes also referred to as “variant FVIII”. Variant FVIII proteins have at least one amino acid modification as compared to human wild type FVIII. Examples of Factor VIII amino acid variations (relative to FVIII-FL-AA (SEQ ID NO: 19)) found in the human population include, without limitation, S19R, R22T, Y24C, Y25C, L26P/R, E30V, W33G, Y35C/H, G41C, R48C/K, K67E/N, L69P, E72K, D75E/V/Y, P83R, G89D/V, G92A/V, A97P, E98K, V99D, D101G/H/V, V104D, K108T, M110V, A111T/V, H113R/Y, L117F/R, G121S, E129V, G130R, E132D, Y133C, D135G/Y, T137A/I, S138R, E141K, D145H, V147D, Y155H, V159A, N163K, G164D/V, P165S, C172W, S176P, S179P, V181E/M, K185T, D186G/N/Y, S189L, L191F, G193R, L195P, C198G, S202N/R, F214V, L217H, A219D/T, V220G, D222V, E223K, G224W, T252I, V253F, N254I, G255V, L261P, P262L, G263S, G266F, C267Y, W274C, H275L, G278R, G280D, E284K, V285G, E291G/K, T294I, F295L, V297A, N299I, R301C/H/L, A303E/P, I307S, S308L, F312S, T314A/I, A315V, G323E, L326P, L327P/V, C329F, I331V, M339T, E340K, V345A/L, C348R/S/Y, Y365C, R391C/H/P, S392L/P, A394S, W401G, I405F/S, E409G, W412G/R, K427I, L431F/S, R437P/W, I438F, G439D/S/V, Y442C, K444R, Y450D/N, T454I, F455C, G466E, P470L/R/T, G474E/R/V, E475K, G477V, D478N, T479R, F484C, A488G, R490G, Y492C/H, Y492H, I494T, P496R, G498R, R503H, G513S/V, I522Y,

K529E, W532G, P540T, T541S, D544N, R546W, R550C/G/H, S553P, S554C/G, V556D, R560T, D561G/H/Y, I567T, P569R, S577F, V578A, D579A/H, N583S, Q584H/K/R, I585R/T, M586V, D588G/Y, L594Q, S596P, N601D/K, R602G, S603I/R, W604C, Y605H/S, N609I, R612C, N631K/S, M633I, S635N, N637D/I/S, Y639C, L644V, L650F, V653A/M, L659P, A663V, Q664P, F677L, M681I, V682F, Y683C/N, T686R, F698L, M699T/V, M701I, G705V, G710W, N713I, R717L/W, G720D/S, M721I/L, A723T, L725Q, V727F, E739K, Y742C, R795G, P947R, V1012L, E1057K, H1066Y, D1260E, K1289Q, Q1336K, N1460K, L1481P, A1610S, I1698T, Y1699C/F, E1701K, Q1705H, R1708C/H, T1714S, R1715G, A1720V, E1723K, D1727V, Y1728C, R1740G, K1751Q, F1762L, R1768H, G1769R, L1771P, L1775F/V, L1777P, G1779E/R, P1780L, I1782R, D1788H, M1791T, A1798P, S1799H, R1800C/G/H, P1801A, Y1802C, S1803Y, F1804S, L1808F, M1842I, P1844S, T1845P, E1848G, A1853T/V, S1858C, K1864E, D1865N/Y, H1867P/R, G1869D/V, G1872E, P1873R, L1875P, V1876L, C1877R/Y, L1882P, R1888I, E1894G, I1901F, E1904D/K, S1907C/R, W1908L, Y1909C, A1939T/V, N1941D/S, G1942A, M1945V, L1951F, R1960L/Q, L1963P, S1965I, M1966I/V, G1967D, S1968R, N1971T, H1973L, G1979V, H1980P/Y, F1982I, R1985Q, L1994P, Y1998C, G2000A, T2004R, M2007I, G2013R, W2015C, R2016P/W, E2018G, G2022D, G2028R, S2030N, V2035A, Y2036C, N2038S, 2040Y, G2045E/V, I2051S, I2056N, A2058P, W2065R, P2067L, A2070V, S2082N, S2088F, D2093G/Y, H2101D, T2105N, Q2106E/P/R, G2107S, R2109C, I2117F/S, Q2119R, F2120C/L, Y2124C, R2135P, S2138Y, T2141N, M2143V, F2145C, N2148S, N2157D, P2162L, R2169C/H, P2172L/Q/R, T2173A/I, H2174D, R2178C/H/L, R2182C/H/P, M2183R/V, L2185S/W, S2192I, C2193G, P2196R, G2198V, E2200D, I2204T, I2209N, A2211P, A2220P, P2224L, R2228G/L/P/Q, L2229F, V2242M, W2248C/S, V2251A/E, M2257V, T2264A, Q2265R, F2279C/I, I2281T, D2286G, W2290L, G2304V, D2307A, P2319L/S, R2323C/G/H/L, R2326G/L/P/Q, Q2330P, W2332R, I2336F, R2339T, G2344C/D/S, and C2345S/Y. Factor VIII proteins also include polypeptides containing post-translational modifications.

[0080] Generally, polynucleotides encoding Factor VIII encode for an inactive single-chain polypeptide (e.g., a pre-pro-protein) that undergoes post-translational processing to form an active Factor VIII protein (e.g., FVIIIa). For example, referring to Figure 15, the wild type human Factor VIII pre-pro-protein is first cleaved to release the encoded signal peptide (not shown), forming a first single-chain pro-protein (shown as "human wild-type FVIII). The pro-

protein is then cleaved between the B and A3 domains to form a first polypeptide that includes the Factor VIII heavy chain (e.g., the A1 and A2 domains) and B-domain, and a second polypeptide that includes the Factor VIII light chain (e.g., including the A3, C1, and C3 domains). The first polypeptide is further cleaved to remove the B-domain, and also to separate the A1 and A2 domains, which remain associated with the Factor VIII light chain in the mature Factor VIIIa protein. For review of the Factor VIII maturation process, see Graw et al., *Nat Rev Genet.*, 6(6):488-501 (2005), the content of which is incorporated herein by reference in its entirety for all purposes.

[0081] As used herein, the terms "Factor VIII heavy chain," or simply "heavy chain," refers to the aggregate of the A1 and A2 domains of a Factor VIII polypeptide. In an exemplary embodiment, amino acids 20-759 of hFVIII-FL-AA (SEQ ID NO:39) constitute a Factor VIII heavy chain.

[0082] As used herein, the term "Factor VIII light chain," or simply "light chain," refers to the aggregate of the A3, C1, and C2 domains of a Factor VIII polypeptide. In an exemplary embodiment, amino acids 1668-2351 of hFVIII-FL-AA (SEQ ID NO:39) constitute a Factor VIII light chain. In some embodiments, a Factor VIII light chain excludes the acidic a3 peptide, which is released during maturation in vivo.

[0083] Generally, Factor VIII heavy and light chains are expressed as a single polypeptide chain, e.g., along with an optional B-domain or B-domain substituted linker. However, in some embodiments, a Factor VIII heavy chain and Factor VIII light chain are expressed as separate polypeptide chains (e.g., co-expressed), and reconstituted to form a Factor VIII protein (e.g., in vivo or in vitro).

[0084] As used herein, the terms "B-domain substituted linker" and "Factor VIII linker" are used interchangeably, and refer to truncated versions of a wild type Factor VIII B-domain (e.g., amino acids 760-1667 of hFVIII-FL-AA (SEQ ID NO:39)) or peptides engineered to replace the B-domain of a Factor VIII polypeptide. As used herein, a Factor VIII linker is positioned between the C-terminus of a Factor VIII heavy chain and the N-terminus of a Factor VIII light chain in a Factor VIII variant polypeptide in accordance with some embodiments. Non-limiting examples of B-domain substituted linkers are disclosed in U.S. Patent Nos. 4,868,112, 5,112,950, 5,171,844, 5,543,502, 5,595,886, 5,610,278, 5,789,203, 5,972,885, 6,048,720, 6,060,447,

6,114,148, 6,228,620, 6,316,226, 6,346,513, 6,458,563, 6,924,365, 7,041,635, and 7,943,374; U.S. Patent Application Publication Nos. 2013/024960, 2015/0071883, and 2015/0158930; and PCT Publication Nos. WO 2014/064277 and WO 2014/127215, the disclosures of which are hereby incorporated by reference, in their entireties, for all purposes.

[0085] Unless otherwise specified herein, the numbering of Factor VIII amino acids refers to the corresponding amino acid in the full-length, wild-type human Factor VIII sequence (hFVIII-FL-AA), presented as SEQ ID NO:39 in Figure 16. As such, when referring to an amino acid substitution in a Factor VIII variant protein disclosed herein, the recited amino acid number refers to the analogous (e.g., structurally or functionally equivalent) and/or homologous (e.g., evolutionarily conserved in the primary amino acid sequence) amino acid in the full-length, wild-type Factor VIII sequence. For example, a T2105N amino acid substitution refers to a T to N substitution at position 2105 of the full-length, wild-type human Factor VIII sequence (hFVIII-FL-AA; SEQ ID NO:39), a T to N substitution at position 1218 of the Factor VIII variant protein encoded by CS12 (CS12-FL-AA; SEQ ID NO: 2).

[0086] As described herein, the Factor VIII amino acid numbering system is dependent on whether the Factor VIII signal peptide (e.g., amino acids 1-19 of the full-length, wild-type human Factor VIII sequence) is included. Where the signal peptide is included, the numbering is referred to as "signal peptide inclusive" or "SPI". Where the signal peptide is not included, the numbering is referred to as "signal peptide exclusive" or "SPE." For example, F328S is SPI numbering for the same amino acid as F309S, in SPE numbering. Unless otherwise indicated, all amino acid numbering refers to the corresponding amino acid in the full-length, wild-type human Factor VIII sequence (hFVIII-FL-AA), presented as SEQ ID NO:39 in Figure 16.

[0087] As described herein, the codon-altered polynucleotides provide increased expression of transgenic Factor VIII in vivo (e.g., when administered as part of a gene therapy vector), as compared to the level of Factor VIII expression provided by a natively-coded Factor VIII construct (e.g., a polynucleotide encoding the same Factor VIII construct using the wild-type human codons). As used herein, the term "increased expression" refers to an increased level of transgenic Factor VIII activity in the blood of an animal administered the codon-altered polynucleotide encoding Factor VIII, as compared to the level of transgenic Factor VIII activity in the blood of an animal administered a natively-coded Factor VIII construct. The activity

levels can be measured using any Factor VIII activity known in the art, e.g., the two-step chromogenic Factor X activation assay described in Chapter 2.7.4 of the European Pharmacopoeia 9.0. An exemplary assay for determining Factor VIII activity is the Technochrome FVIII assay (Technoclone, Vienna, Austria).

[0088] In some embodiments, increased bioavailability refers to at least 25% greater transgenic Factor VIII polypeptide in the blood of an animal administered the codon-altered Factor VIII polynucleotide, as compared to the level of transgenic Factor VIII polypeptide in the blood of an animal administered a natively coded Factor VIII polynucleotide. In some embodiments, increased bioavailability refers to at least 25% greater transgenic Factor VIII polypeptide in the blood of an animal administered an improved gene therapy vector that includes a codon-altered Factor VIII polynucleotide, as compared to the level of a transgenic Factor VIII polypeptide in the blood of an animal administered a different gene therapy vector that includes the same or a different codon-altered Factor VIII polynucleotide. In some embodiments, increased expression refers to at least 50% greater, at least 75% greater, at least 100% greater, at least 3-fold greater, at least 4-fold greater, at least 5-fold greater, at least 6-fold greater, at least 7-fold greater, at least 8-fold greater, at least 9-fold greater, at least 10-fold greater, at least 15-fold greater, at least 20-fold greater, at least 25-fold greater, at least 30-fold greater, at least 40-fold greater, at least 50-fold greater, at least 60-fold greater, at least 70-fold greater, at least 80-fold greater, at least 90-fold greater, at least 100-fold greater, at least 125-fold greater, at least 150-fold greater, at least 175-fold greater, at least 200-fold greater, at least 225-fold greater, or at least 250-fold greater transgenic Factor VIII activity in the blood of an animal administered the codon-altered Factor VIII polynucleotide.

[0089] By “Factor VIII activity” herein is meant the ability to promote cleavage of a Factor X polypeptide by Factor IXa, e.g., Factor IXa co-factor activity, via hydrolysis of the Arg194-Ile195 peptide bond in wild-type Factor IX, thus activating Factor X to Factor Xa. The activity levels can be measured using any Factor VIII activity known in the art; suitable assays are outlined herein. An exemplary assay for determining Factor VIII activity is the two-step chromogenic Factor X activation assay described in Chapter 2.7.4 of the European Pharmacopoeia 9.0.

[0090] As used herein, the term “biopotency” refers to the amount of Factor VIII activity in the blood a subject, in vivo, or in a cell culture supernatant, in vitro. In some embodiments, biopotency will refer to an amount of activity per unit volume, such as units of Factor XIa co-factor activity per mL of blood, in vivo, or per mL cell culture supernatant, in vitro. In some embodiments, biopotency will be expressed as a fold increase with respect to a first level, e.g., a natively coded Factor VIII protein or a codon-optimized native Factor VIII (e.g., a ‘wild-type’ Refecto FVIII protein). In some embodiments, as used herein, the biopotency of exogenously expressed Factor VIII refers to the amount of Factor VIII activity provided by a recombinant Factor VIII protein expressed from a gene therapy vector. That is, the amount of Factor VIII activity in the blood or cell culture supernatant after accounting for any baseline amount of native Factor VIII activity. Thus, increases in biopotency can be achieved by either, or both, increasing the expression level of an exogenous Factor VIII protein and/or increasing the specific activity of an exogenous Factor VIII protein, e.g., by including amino acid substitutions (such as the X5 mutation) that confer greater specific activity.

[0091] In some embodiments, the therapeutic potential of a Factor VIII polynucleotide composition is evaluated by the increase in Factor VIII activity in the blood of an animal administered a Factor VIII polynucleotide, e.g., instead of, or in addition to, increased Factor VIII expression and/or bioavailability. In some embodiments, as used herein, increased Factor VIII activity refers to a greater increase in Factor VIII activity in the blood of an animal administered a codon-altered Factor VIII polynucleotide, relative to a baseline Factor VIII activity in the blood of the animal prior to administration of the codon-altered Factor VIII polynucleotide, as compared to the increase in Factor VIII activity in the blood of an animal administered a natively-coded Factor VIII polynucleotide, relative to a baseline Factor VIII activity in the blood of the animal prior to administration of the natively-coded Factor VIII polynucleotide. In some embodiments, increased Factor VIII activity refers to at least a 25% greater increase in Factor VIII activity in the blood of an animal administered the codon-altered Factor VIII polynucleotide, relative to a baseline level of Factor VIII activity in the blood of the animal prior to administration of the codon-altered Factor VIII polynucleotide, as compared to the increase in the level Factor VIII activity in the blood of an animal administered a natively-coded Factor VIII polynucleotide, relative to the baseline level of Factor VIII activity in the animal prior to administration of the natively-coded Factor VIII polynucleotide. In some

embodiments, increased Factor VIII activity refers to at least a 25% greater increase in Factor VIII activity in the blood of an animal administered an improved gene therapy vector that includes a codon-altered Factor VIII polynucleotide, relative to a baseline level of Factor VIII activity in the blood of the animal prior to administration of the improved gene therapy vector, as compared to the increase in the level of Factor VIII activity in the blood of an animal administered a different gene therapy vector that includes the same or a different codon-altered Factor VIII polynucleotide, relative to the, relative to the baseline level of Factor VIII activity in the animal prior to administration of the different gene therapy vector. In some embodiments, increased Factor VIII activity refers to at least 50% greater, at least 75% greater, at least 100% greater, at least 3-fold greater, at least 4-fold greater, at least 5-fold greater, at least 6-fold greater, at least 7-fold greater, at least 8-fold greater, at least 9-fold greater, at least 10-fold greater, at least 15-fold greater, at least 20-fold greater, at least 25-fold greater, at least 30-fold greater, at least 40-fold greater, at least 50-fold greater, at least 60-fold greater, at least 70-fold greater, at least 80-fold greater, at least 90-fold greater, at least 100-fold greater, at least 125-fold greater, at least 150-fold greater, at least 175-fold greater, at least 200-fold greater, at least 225-fold greater, or at least 250-fold greater increase in Factor VIII activity in the blood of an animal administered the codon-altered Factor VIII polynucleotide, relative to a baseline level of Factor VIII activity in the blood of the animal prior to administration of the codon-altered Factor VIII polynucleotide, as compared to the increase in the level Factor VIII activity in the blood of an animal administered a natively-coded Factor VIII polynucleotide, or a different gene therapy vector that includes the same or a different codon-altered Factor VIII polynucleotide, relative to the baseline level of Factor VIII activity in the animal prior to administration of the natively-coded Factor VIII polynucleotide or different gene therapy vector that includes the same or a different codon-altered Factor VIII polynucleotide. Activity is measured using the two-step chromogenic Factor X activation assay described in Chapter 2.7.4 of the European Pharmacopoeia 9.0, as described herein.

[0092] As described herein, the codon-altered polynucleotides provide increased vector production, as compared to the level of vector production provided by a natively-coded Factor VIII construct (e.g., a polynucleotide encoding the same Factor VIII construct using the wild-type human codons). As used herein, the term "increased virus production" refers to an increased vector yield in cell culture (e.g., titer per liter culture) inoculated with the codon-

altered polynucleotide encoding Factor VIII, as compared to the vector yield in cell culture inoculated with a natively-coded Factor VIII construct. The vector yields can be measured using any vector titer assay known in the art. An exemplary assay for determining vector yield (e.g., of an AAV vector) is qPCR targeting the AAV2 inverted terminal repeats (Aurnhammer, Human Gene Therapy Methods: Part B 23:18-28 (2012)).

[0093] In some embodiments, increased virus production refers to at least 25% greater codon-altered vector yield, as compared to the yield of a natively-coded Factor VIII construct in the same type of culture. In some embodiments, increased virus production refers to at least 25% greater yield of an improved vector that includes a codon-altered Factor VIII polynucleotide, as compared to the yield of a different vector that includes the same or a different codon-altered Factor VIII polynucleotide. In some embodiments, increased vector production refers to at least 50% greater, at least 75% greater, at least 100% greater, at least 3-fold greater, at least 4-fold greater, at least 5-fold greater, at least 6-fold greater, at least 7-fold greater, at least 8-fold greater, at least 9-fold greater, at least 10-fold greater, at least 15-fold greater, or at least 20-fold greater codon-altered vector yield.

[0094] As used herein, the term "hemophilia" refers to a group of disease states broadly characterized by reduced blood clotting or coagulation. Hemophilia may refer to Type A, Type B, or Type C hemophilia, or to the composite of all three diseases types. Type A hemophilia (hemophilia A) is caused by a reduction or loss of factor VIII (FVIII) activity and is the most prominent of the hemophilia subtypes. Type B hemophilia (hemophilia B) results from the loss or reduction of factor IX (FIX) clotting function. Type C hemophilia (hemophilia C) is a consequence of the loss or reduction in factor XI (FXI) clotting activity. Hemophilia A and B are X-linked diseases, while hemophilia C is autosomal. Conventional treatments for hemophilia include both prophylactic and on-demand administration of clotting factors, such as FVIII, FIX, including Bebulin®-VH, and FXI, as well as FEIBA-VH, desmopressin, and plasma infusions.

[0095] As used herein, the term "FVIII gene therapy" includes any therapeutic approach of providing a nucleic acid encoding Factor VIII to a patient to relieve, diminish, or prevent the reoccurrence of one or more symptoms (e.g., clinical factors) associated with hemophilia. The term encompasses administering any compound, drug, procedure, or regimen comprising a nucleic acid encoding a Factor VIII molecule, including any modified form of Factor VIII (e.g.,

Factor VIII variant), for maintaining or improving the health of an individual with hemophilia. One skilled in the art will appreciate that either the course of FVIII therapy or the dose of a FVIII therapeutic agent can be changed, e.g., based upon the results obtained in accordance with the present disclosure.

[0096] As used herein, the term “Factor VIII gene therapy,” or “FVIII gene therapy,” includes any therapeutic approach of providing a nucleic acid encoding a Factor VIII polypeptide to a patient to relieve, diminish, or prevent the reoccurrence of one or more symptoms (e.g., clinical factors) associated with a Factor VIII deficiency (e.g., hemophilia A). The term encompasses administering any compound, drug, procedure, or regimen comprising a nucleic acid encoding a Factor VIII molecule, including any modified form of Factor VIII (e.g., a Factor VIII variant having the X5 mutations, a B-domain deletion, and/or a glycosylation peptide inserted into a B-domain linker polypeptide), for maintaining or improving the health of an individual with a Factor VIII deficiency (e.g., hemophilia A). One skilled in the art will appreciate that either the course of FVIII gene therapy or the dose of a FVIII gene therapy therapeutic agent can be changed, e.g., based upon the results obtained in accordance with the present disclosure.

[0097] As used herein, the term “bypass therapy” includes any therapeutic approach of providing non-Factor VIII hemostatic agents, compounds or coagulation factors to a patient to relieve, diminish, or prevent the reoccurrence of one or more symptoms (e.g., clinical factors) associated with hemophilia. Non-Factor VIII compounds and coagulation factors include, but are not limited to, Factor VIII Inhibitor Bypass Activity (FEIBA), recombinant activated factor VII (FVIIa), prothrombin complex concentrates, and activated prothrombin complex concentrates. These non-Factor VIII compounds and coagulation factors may be recombinant or plasma-derived. One skilled in the art will appreciate that either the course of bypass therapy or the dose of bypass therapy can be changed, e.g., based upon the results obtained in accordance with the present disclosure.

[0098] As used herein, a “combination therapy” including administration of a nucleic acid encoding a Factor VIII molecule and a conventional hemophilia A therapeutic agent includes any therapeutic approach of providing both a nucleic acid encoding a Factor VIII molecule and a Factor VIII molecule and/or non-Factor VIII hemostatic agent (e.g., bypass therapeutic agent) to a patient to relieve, diminish, or prevent the reoccurrence of one or more symptoms (e.g., clinical

factors) associated with hemophilia. The term encompasses administering any compound, drug, procedure, or regimen including a nucleic acid encoding a Factor VIII molecule, including any modified form of factor VIII, which is useful for maintaining or improving the health of an individual with hemophilia and includes any of the therapeutic agents described herein.

[0099] The terms "therapeutically effective amount or dose" or "therapeutically sufficient amount or dose" or "effective or sufficient amount or dose" refer to a dose that produces therapeutic effects for which it is administered. For example, a therapeutically effective amount of a drug useful for treating hemophilia can be the amount that is capable of preventing or relieving one or more symptoms associated with hemophilia.

[00100] **[0083]** In some embodiments, a therapeutically effective treatment results in a decrease in the frequency and/or severity of bleeding incidents in a subject.

[00101] In some embodiments, a therapeutically effective treatment results in increased Factor VIII activity in the blood stream of a patient, as compared to the activity prior to the treatment.

[00102] As used herein, the term "gene" refers to the segment of a DNA molecule that codes for a polypeptide chain (e.g., the coding region). In some embodiments, a gene is positioned by regions immediately preceding, following, and/or intervening the coding region that are involved in producing the polypeptide chain (e.g., regulatory elements such as a promoter, enhancer, polyadenylation sequence, 5' untranslated region, 3' untranslated region, or intron).

[00103] As used herein, the term "regulatory elements" refers to nucleotide sequences, such as promoters, enhancers, terminators, polyadenylation sequences, introns, etc, that provide for the expression of a coding sequence in a cell.

[00104] As used herein, the term "promoter element" refers to a nucleotide sequence that assists with controlling expression of a coding sequence. Generally, promoter elements are located 5' of the translation start site of a gene. However, in certain embodiments, a promoter element may be located within an intron sequence, or 3' of the coding sequence. In some embodiments, a promoter useful for a gene therapy vector is derived from the native gene of the target protein (e.g., a Factor VIII promoter). In some embodiments, a promoter useful for a gene

therapy vector is specific for expression in a particular cell or tissue of the target organism (e.g., a liver-specific promoter). In yet other embodiments, one of a plurality of well characterized promoter elements is used in a gene therapy vector described herein. Non-limiting examples of well-characterized promoter elements include the CMV early promoter, the β -actin promoter, and the methyl CpG binding protein 2 (MeCP2) promoter. In some embodiments, the promoter is a constitutive promoter, which drives substantially constant expression of the target protein. In other embodiments, the promoter is an inducible promoter, which drives expression of the target protein in response to a particular stimulus (e.g., exposure to a particular treatment or agent). For a review of designing promoters for AAV-mediated gene therapy, see Gray et al. (Human Gene Therapy 22:1143-53 (2011)), the contents of which are expressly incorporated by reference in their entirety for all purposes.

[00105] [0087] As used herein, a “CRM8” element refers to cis-acting regulatory module derived from the SERPINA1 gene (NCBI accession number NM_000295.4) that enhances expression of an operatively linked gene, e.g., a sequence encoding a Factor VIII polypeptide, in a liver-specific fashion having high sequence identity to SEQ ID NO:5. In some embodiments, the CRM8 element is identical to SEQ ID NO:5. As used herein, a CRM8 element refers to a single copy of the regulatory element which, in some embodiments, is included in one or more copies within a Factor VIII polynucleotide, e.g., 1, 2, 3, or more copies. For further information on CRM elements, such as CRM8, see Chuah MK et al., Mol Ther., 22(9):1605-13 (2014), which is hereby incorporated by reference.

[00106] [0089] As used herein, the term “operably linked” refers to the relationship between a first reference nucleotide sequence (e.g., a gene) and a second nucleotide sequence (e.g., a regulatory control element) that allows the second nucleotide sequence to affect one or more properties associated with the first reference nucleotide sequence (e.g., a transcription rate). In the context of the present disclosure, a regulatory control element is operably linked to a Factor VIII transgene when the regulatory element is positioned within a gene therapy vector such that it exerts an affect (e.g., a promotive or tissue-selective affect) on transcription of the Factor VIII transgene.

[00107] As used herein, the term “vector” refers to any vehicle used to transfer a nucleic acid (e.g., encoding a Factor VIII gene therapy construct) into a host cell. In some embodiments,

a vector includes a replicon, which functions to replicate the vehicle, along with the target nucleic acid. Non-limiting examples of vectors useful for gene therapy include plasmids, phages, cosmids, artificial chromosomes, and viruses, which function as autonomous units of replication in vivo. In some embodiments, a vector is a viral vehicle for introducing a target nucleic acid (e.g., a codon-altered polynucleotide encoding a Factor VIII variant). Many modified eukaryotic viruses useful for gene therapy are known in the art. For example, adeno-associated viruses (AAVs) are particularly well suited for use in human gene therapy because humans are a natural host for the virus, the native viruses are not known to contribute to any diseases, and the viruses illicit a mild immune response.

[00108] As used herein, the term “Factor VIII viral vector” refers to a recombinant virus comprising a Factor VIII polynucleotide, encoding a Factor VIII polypeptide, which is sufficient for expression of the Factor VIII polypeptide when introduced into a suitable animal host (e.g., a human). Specifically included within the definition of Factor VIII viral vector are recombinant viruses in which a codon-altered Factor VIII polynucleotide, which encodes a Factor VIII polypeptide, has been inserted into the genome of the virus. Also specifically included within the definition of Factor VIII viral vectors are recombinant viruses in which the native genome of the virus has been replaced with a Factor VIII polynucleotide, which encodes a Factor VIII polypeptide. Included within the definition of Factor VIII viral vectors are recombinant viruses comprising a Factor VIII polynucleotide, which encodes a variant of Factor VIII having the X5 mutations, a B-domain deletion, and/or a glycosylation peptide inserted into a B-domain linker polypeptide.

[00109] As used herein, the term “Factor VIII viral particle” refers to a viral particle encapsulating a Factor VIII polynucleotide, encoding a Factor VIII polypeptide, which is specific for expression of the Factor VIII polypeptide when introduced into a suitable animal host (e.g., a human). Specifically included within the definition of Factor VIII viral particles are recombinant viral particles encapsulating a genome in which a codon-altered Factor VIII polynucleotide, which encodes a Factor VIII polypeptide, has been inserted. Also specifically included within the definition of Factor VIII viral particles are recombinant viral particles encapsulating a Factor VIII polynucleotide, which encodes a Factor VIII polypeptide, which replaces the native genome of the virus. Included within the definition of Factor VIII viral particles are recombinant viral particles encapsulating a Factor VIII polynucleotide, which encodes a variant of Factor VIII

having the X5 mutations, a B-domain deletion, and/or a glycosylation peptide inserted into a B-domain linker polypeptide.

[00110] By “AAV” or “adeno-associated virus” herein is meant a Dependoparvovirus within the Parvoviridae genus of viruses. As used herein, AAV can refer to a virus derived from a naturally occurring “wild-type” AAV genome into which a Factor VIII polynucleotide has been inserted, a recombinant virus derived from a recombinant Factor VIII polynucleotide packaged into a capsid using capsid proteins encoded by a naturally occurring AAV cap gene, or a recombinant virus derived from a recombinant Factor VIII polynucleotide packaged into a capsid using capsid proteins encoded by a non-natural capsid cap gene. Included within the definition of AAV are AAV type 1 (AAV1), AAV type 2 (AAV2), AAV type 3 (AAV3), AAV type 4 (AAV4), AAV type 5 (AAV5), AAV type 6 (AAV6), AAV type 7 (AAV7), AAV type 8 (AAV8), and AAV type 9 (AAV9) viruses encapsulating a Factor VIII polynucleotide and viruses formed by one or more variant AAV capsid proteins encapsulating a Factor VIII polynucleotide.

[00111] By “AAV8,” “AAV-8,” or “AAV serotype 8” herein is meant a virus formed by an AAV8 capsid viral protein that encapsulates a Factor VIII polynucleotide.

[00112] As used herein, the term “CpG” refers to a cytosine-guanine dinucleotide along a single strand of DNA, with the “p” representing the phosphate linkage between the two.

[00113] As used herein, the term “CpG island” refers to a region within a polynucleotide having a statistically elevated density of CpG dinucleotides. As used herein, a region of a polynucleotide (e.g., a polynucleotide encoding a codon-altered Factor VIII protein) is a CpG island if, over a 200-base pair window: (i) the region has GC content of greater than 50%, and (ii) the ratio of observed CpG dinucleotides per expected CpG dinucleotides is at least 0.6, as defined by the relationship:

$$\frac{N[CpG]*N[length\ of\ window]}{N[C]*N[G]} \geq 0.6.$$

For additional information on methods for identifying CpG islands, see Gardiner-Garden M. et al., J Mol Biol., 196(2):261-82 (1987), the content of which is expressly incorporated herein by reference, in its entirety, for all purposes.

[00114] As used herein, the term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form, and complements thereof. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, and peptide-nucleic acids (PNAs).

[00115] The term "amino acid" refers to naturally occurring amino acids, including those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Naturally occurring amino acids can include, e.g., D- and L-amino acids. Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

[00116] By "modification" herein is meant an amino acid substitution, insertion, and/or deletion in a polypeptide sequence or an alteration to a moiety chemically linked to a protein. For example, a modification may be an altered carbohydrate or PEG structure attached to a protein. By "amino acid modification" herein is meant an amino acid substitution, insertion, and/or deletion in a polypeptide sequence. For clarity, unless otherwise noted, the amino acid modification is always to an amino acid coded for by DNA, e.g. the 20 amino acids that have codons in DNA and RNA.

[00117] By "amino acid substitution" or "substitution" herein is meant the replacement of an amino acid at a particular position in a parent polypeptide sequence with a different amino acid. In particular, in some embodiments, the substitution is to an amino acid that is not naturally occurring at the particular position, either not naturally occurring within the organism or in any organism. For example, the substitution G151K refers to a variant polypeptide, in which the glycine at position 151 is replaced with lysine. For clarity, a protein which has been engineered to change the nucleic acid coding sequence but not change the starting amino acid (for example exchanging CGG (encoding arginine) to CGA (still encoding arginine) to increase

host organism expression levels) is not an "amino acid substitution"; that is, despite the creation of a new gene encoding the same protein, if the protein has the same amino acid at the particular position that it started with, it is not an amino acid substitution. Accordingly, each variation of a nucleic acid which encodes a same polypeptide is implicit in each described sequence with respect to the expression product, but not with respect to actual gene therapy constructs.

[00118] By "amino acid insertion" or "insertion" as used herein is meant the addition of an amino acid sequence at a particular position in a parent polypeptide sequence. For example, -233E or 233E designates an insertion of glutamic acid after position 233 and before position 234. Additionally, -233ADE or A233ADE designates an insertion of AlaAspGlu after position 233 and before position 234.

[00119] By "amino acid deletion" or "deletion" as used herein is meant the removal of an amino acid sequence at a particular position in a parent polypeptide sequence. For example, E233- or E233#, E233() or E233del designates a deletion of glutamic acid at position 233. Additionally, EDA233- or EDA233# designates a deletion of the sequence GluAspAla that begins at position 233. The terms "identical" or percent "identity," in the context of two or more nucleic acids or peptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 60% identity, preferably 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection.

[00120] As is known in the art, a number of different programs may be used to identify whether a protein (or nucleic acid as discussed below) has sequence identity or similarity to a known sequence. Sequence identity and/or similarity is determined using standard techniques known in the art, including, but not limited to, the local sequence identity algorithm of Smith & Waterman, Adv. Appl. Math., 2:482 (1981), by the sequence identity alignment algorithm of Needleman & Wunsch, J. Mol. Biol., 48:443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Natl. Acad. Sci. U.S.A., 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin

Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, WI), the Best Fit sequence program described by Devereux et al., Nucl. Acid Res., 12:387-395 (1984), preferably using the default settings, or by inspection. Preferably, percent identity is calculated by FastDB based upon the following parameters: mismatch penalty of 1; gap penalty of 1; gap size penalty of 0.33; and joining penalty of 30, "Current Methods in Sequence Comparison and Analysis," Macromolecule Sequencing and Synthesis, Selected Methods and Applications, pp 127-149 (1988), Alan R. Liss, Inc, all of which are incorporated by reference.

[00121] An example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pair wise alignments. It may also plot a tree showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, J. Mol. Evol. 35:351-360 (1987); the method is similar to that described by Higgins & Sharp CABIOS 5:151-153 (1989), both incorporated by reference. Useful PILEUP parameters including a default gap weight of 3.00, a default gap length weight of 0.10, and weighted end gaps.

[00122] Another example of a useful algorithm is the BLAST algorithm, described in: Altschul et al., J. Mol. Biol. 215, 403-410, (1990); Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997); and Karlin et al., Proc. Natl. Acad. Sci. U.S.A. 90:5873-5787 (1993), both incorporated by reference. A particularly useful BLAST program is the WU-BLAST-2 program which was obtained from Altschul et al., Methods in Enzymology, 266:460-480 (1996); [http://blast.wustl.edu/blast/ README.html](http://blast.wustl.edu/blast/README.html)]. WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable parameters are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched; however, the values may be adjusted to increase sensitivity.

[00123] An additional useful algorithm is gapped BLAST, as reported by Altschul et al., Nucl. Acids Res., 25:3389-3402, incorporated by reference. Gapped BLAST uses BLOSUM-62 substitution scores; threshold T parameter set to 9; the two-hit method to trigger ungapped extensions; charges gap lengths of k a cost of 10+k; Xu set to 16, and Xg set to 40 for database

search stage and to 67 for the output stage of the algorithms. Gapped alignments are triggered by a score corresponding to ~22 bits.

[00124] A % amino acid sequence identity value is determined by the number of matching identical residues divided by the total number of residues of the "longer" sequence in the aligned region. The "longer" sequence is the one having the most actual residues in the aligned region (gaps introduced by WU-Blast-2 to maximize the alignment score are ignored). In a similar manner, "percent (%) nucleic acid sequence identity" with respect to the coding sequence of the polypeptides identified is defined as the percentage of nucleotide residues in a candidate sequence that are identical with the nucleotide residues in the coding sequence of the cell cycle protein. A preferred method utilizes the BLASTN module of WU-BLAST-2 set to the default parameters, with overlap span and overlap fraction set to 1 and 0.125, respectively.

[00125] The alignment may include the introduction of gaps in the sequences to be aligned. In addition, for sequences which contain either more or fewer amino acids than the protein encoded by the sequence of Figure 1 (SEQ ID NO:1), it is understood that in one embodiment, the percentage of sequence identity will be determined based on the number of identical amino acids or nucleotides in relation to the total number of amino acids or nucleotides. Thus, for example, sequence identity of sequences shorter than that shown in Figure 1 (SEQ ID NO:1), as discussed below, will be determined using the number of nucleotides in the shorter sequence, in one embodiment. In percent identity calculations relative weight is not assigned to various manifestations of sequence variation, such as, insertions, deletions, substitutions, etc.

[00126] In one embodiment, only identities are scored positively (+1) and all forms of sequence variation including gaps are assigned a value of "0", which obviates the need for a weighted scale or parameters as described below for sequence similarity calculations. Percent sequence identity may be calculated, for example, by dividing the number of matching identical residues by the total number of residues of the "shorter" sequence in the aligned region and multiplying by 100. The "longer" sequence is the one having the most actual residues in the aligned region.

[00127] The term "allelic variants" refers to polymorphic forms of a gene at a particular genetic locus, as well as cDNAs derived from mRNA transcripts of the genes, and the polypeptides encoded by them. The term "preferred mammalian codon" refers a subset of

codons from among the set of codons encoding an amino acid that are most frequently used in proteins expressed in mammalian cells as chosen from the following list: Gly (GGC, GGG); Glu (GAG); Asp (GAC); Val (GTG, GTC); Ala (GCC, GCT); Ser (AGC, TCC); Lys (AAG); Asn (AAC); Met (ATG); Ile (ATC); Thr (ACC); Trp (TGG); Cys (TGC); Tyr (TAT, TAC); Leu (CTG); Phe (TTC); Arg (CGC, AGG, AGA); Gln (CAG); His (CAC); and Pro (CCC).

[00128] As used herein, the term codon-altered refers to a polynucleotide sequence encoding a polypeptide (e.g., a Factor VIII variant protein), where at least one codon of the native polynucleotide encoding the polypeptide has been changed to improve a property of the polynucleotide sequence. In some embodiments, the improved property promotes increased transcription of mRNA coding for the polypeptide, increased stability of the mRNA (e.g., improved mRNA half-life), increased translation of the polypeptide, and/or increased packaging of the polynucleotide within the vector. Non-limiting examples of alterations that can be used to achieve the improved properties include changing the usage and/or distribution of codons for particular amino acids, adjusting global and/or local GC content, removing AT-rich sequences, removing repeated sequence elements, adjusting global and/or local CpG dinucleotide content, removing cryptic regulatory elements (e.g., TATA box and CCAAT box elements), removing of intron/exon splice sites, improving regulatory sequences (e.g., introduction of a Kozak consensus sequence), and removing sequence elements capable of forming secondary structure (e.g., stem-loops) in the transcribed mRNA.

[00129] As discussed herein, there are various nomenclatures to refer to components of the disclosure herein. "CS-number" (e.g. "CS12", "CS04", etc.) refer to codon altered polynucleotides encoding FVIII polypeptides and/or the encoded polypeptides, including variants. For example, CS12-FL refers to the Full Length codon altered CS12 polynucleotide sequence or amino acid sequence (sometimes referred to herein as "CS12-FL-AA" for the Amino Acid sequence and "CS12-FL-NA" for the Nucleic Acid sequence) encoded by the CS12 polynucleotide sequence. Similarly, "CS12-LC" refers to either the codon altered nucleic acid sequence ("CS12-LC-NA") encoding the light chain of a FVIII polypeptide or the amino acid sequence (also sometimes referred to herein as "CS12-LC-AA") of the FVIII light chain encoded by the CS12 polynucleotide sequence. Likewise, CS12-HC, CS12-HC-AA and CS12-HC-NA are the same for the FVIII heavy chain. As will be appreciated by those in the art, for constructs such as CS04, that are only codon-altered (e.g. they do not contain additional amino acid

substitutions as compared to Refacto), the amino acid sequences will be identical, as the amino acid sequences are not altered by the codon optimization. Thus, sequence constructs of the disclosure include, but are not limited to, CS12-FL-NA, CS12-FL-AA, CS12-LC-NA, CS12-LC-AA, CS12-HC-AA, and CS12-HC-NA.

[00130] This nomenclature also applies to glycosylation peptides as shown in Figure 8, such that "NG1-AA" refers to the amino acid sequence and NG1-NA refers to the nucleic acid sequence.

[00131] The disclosure also includes additional new Factor VIII variants, as described below, with the appropriate nomenclature.

[00132] As used herein, the term "liver-specific expression" refers to the preferential or predominant in vivo expression of a particular gene (e.g., a codon-altered, transgenic Factor VIII gene) in hepatic tissue, as compared to in other tissues. In some embodiments, liver-specific expression means that at least 50% of all expression of the particular gene occurs within hepatic tissues of a subject. In other embodiments, liver-specific expression means that at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99%, or 100% of all expression of the particular gene occurs within hepatic tissues of a subject. Accordingly, a liver-specific regulatory element is a regulatory element that drives liver-specific expression of a gene in hepatic tissue.

[00133] As used herein, the terms "less than" X and "less than" X% refer to a range of from 0 to X, exclusive of the value X, e.g., from 0% to X%, exclusive of X%. As used herein, the terms are used interchangeably with a range starting at 0 or 0% through, but not including, X or X%.

[00134] As used herein, the terms "no more than" X or "no more than" X% refer to a range of from 0 to X, inclusive of the value X, e.g., from 0% to X%, inclusive of X%. As used herein, the terms are used interchangeably with a range starting at 0 or 0% through, and including, X or X%.

[00135] As used herein, the terms "greater than" X or "greater than" X% refer to a range of from X to an upper limit, exclusive of the value X, e.g., from X% to 100%, exclusive of X%. As used herein, the terms are used interchangeably with a range starting at, but not including, X or X% through an upper limit which is 100% in the context of a percentage.

[00136] As used herein, the terms “at least” X or “at least” X% refer to a range of from X to an upper limit, inclusive of the value X, e.g., from X% to 100%, inclusive of X%. As used herein, the terms are used interchangeably with a range starting at, and including, X or X% through an upper limit which is 100% in the context of a percentage.

[00137] As used herein, the terms “between ‘X’ and ‘Y’,” “between ‘X’% and ‘Y’%,” “from ‘X’ to ‘Y’,” and “from ‘X’% to ‘Y’%” refer to a range of from X to Y, inclusive of the values X and Y, e.g., from X% to Y%, inclusive of X% and Y%. As used herein, the terms are used interchangeably with a range starting at X or X% through, and including, Y or Y%.

III. Codon-Altered Factor VIII Polynucleotides

[00138] In some embodiments, the present disclosure provides codon-altered polynucleotides encoding Factor VIII variants. These codon-altered polynucleotides provide markedly improved Factor VIII biopotency (e.g., activity) when administered in an AAV-based gene therapy construct. The codon-altered polynucleotides also demonstrate improved AAV-virion packaging, as compared to conventionally codon-optimized constructs.

[00139] Wild-type Factor VIII is encoded with a 19 amino acid signal peptide (amino acids 1-19 of SEQ ID NO:39), which is cleaved from the encoded polypeptide prior to activation of Factor VIII. As appreciated by those in the art the Factor VIII signal peptide may be mutated, replaced by signal peptides from other genes or Factor VIII genes from other organisms, or completely removed, without affecting the sequence of the mature polypeptide left after the signal peptide is removed by cellular processing.

[00140] Accordingly, in some embodiments, a codon-altered polynucleotide (e.g., a nucleic acid composition) provided herein has a nucleotide sequence with high sequence identity to the portions of CS12-FL-NA (SEQ ID NO:1) encoding a Factor VIII heavy and light chains, and a short, 14 amino acid, B-domain substituted linker (e.g., the “SQ” linker containing a furin cleavage site to facilitate maturation of an active FVIIIa protein in vivo), that further includes one or more of the five “X5 mutations” (e.g., one, two, three, four, or all five of the I105V/A127S/G151K/M166T/L171P mutations (SPI numbering; (SPE numbering is I86V/A108S/G132K/M147T/L152P, respectively)), relative to the full-length human wild type

Factor VIII sequence), and/or a short glycosylation peptide (e.g., NG5; SEQ ID NO:15) inserted into the B-domain substituted linker (e.g., an SQ linker).

[00141] Specifically, the X5 mutation set is based on the fact that substitution of porcine amino acids 82-176 for the corresponding human amino acids in a B-domain deleted gene therapy construct increased Factor VIII activity when expressed in HEK293 cells (W. Xiao, communication). Back-mutation of single porcine amino acids into the human BDD-FVIII construct identified five amino acids within the A1 domain that contribute to this phenomenon: I105V, A127S, G151K, M166T, and L171P (SPI). Introduction of the combination of these mutations into the human construct recapitulated the improved activity of the larger porcine substitution. (W. Xiao, communication). Accordingly, in some embodiments, the encoded Factor VIII polypeptides include one or more amino acid substitutions selected from I105V, A127S, G151K, M166T, and L171P, with the entire 5 amino acid set finding particular use in many embodiments.

CS12 Codon Altered Polynucleotides

[00142] In some embodiments, a nucleic acid composition provided herein includes a codon-altered Factor VIII polynucleotide that has a nucleotide sequence with high sequence identity to CS12-FL-NA (SEQ ID NO:1) and encodes a Factor VIII polypeptide with human Factor VIII heavy and light chains, and a short, 14 amino acid, B-domain substituted linker (the “SQ” linker) containing a furin cleavage site to facilitate maturation of an active FVIIIa protein in vivo, where the heavy chain of the Factor VIII polypeptide includes the five X5 mutations (I105V, A127S, G151K, M166T, and L171P (SPI), relative to the full-length human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker.

[00143] In some embodiments, the Factor VIII polynucleotide has a nucleotide sequence has at least 95% identity to CS12-FL-NA (SEQ ID NO:1). In some embodiments, the nucleotide sequence has at least 96% identity to CS12-FL-NA (SEQ ID NO:1). In some embodiments, the nucleotide sequence has at least 97% identity to CS12-FL-NA (SEQ ID NO:1). In some embodiments, the nucleotide sequence has at least 98% identity to CS12-FL-NA (SEQ ID NO:1). In some embodiments, the nucleotide sequence has at least 99% identity to CS12-FL-NA (SEQ ID NO:1). In some embodiments, the nucleotide sequence has at least 99.5% identity to

CS12-FL-NA (SEQ ID NO:1). In some embodiments, the nucleotide sequence has at least 99.9% identity to CS12-FL-NA (SEQ ID NO:1). In some embodiments, the nucleotide sequence is identical to CS12-FL-NA (SEQ ID NO:1).

[00144] In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS12-FL-AA (SEQ ID NO: 2), including the five X5 mutations (I105V/A127S/G151K/M166T/L171P (SPI), relative to the full-length human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker.

[00145] In some embodiments, the amino acid sequence of the encoded Factor VIII variant has at least 97% identity to CS12-FL-AA (SEQ ID NO: 2), including the five X5 mutations (I105V/A127S/G151K/M166T/L171P (SPI), relative to the full-length human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker.

[00146] In some embodiments, the amino acid sequence of the encoded Factor VIII variant has at least 98% identity to CS12-FL-AA (SEQ ID NO: 2), including the five X5 mutations (I105V/A127S/G151K/M166T/L171P (SPI), relative to the full-length human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker.

[00147] In some embodiments, the amino acid sequence of the encoded Factor VIII variant has at least 99% identity to CS12-FL-AA (SEQ ID NO: 2), including the five X5 mutations (I105V/A127S/G151K/M166T/L171P (SPI), relative to the full-length human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker.

[00148] In some embodiments, the amino acid sequence of the encoded Factor VIII variant has at least 99.5% identity to CS12-FL-AA (SEQ ID NO: 2), including the five X5 mutations (I105V/A127S/G151K/M166T/L171P (SPI), relative to the full-length human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker.

[00149] In some embodiments, the amino acid sequence of the encoded Factor VIII variant has at least 99.9% identity to CS12-FL-AA (SEQ ID NO: 2), including the five X5 mutations (I105V/A127S/G151K/M166T/L171P (SPI), relative to the full-length human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker.

[00150] In some embodiments, the amino acid sequence of the encoded Factor VIII variant is to CS12-FL-AA (SEQ ID NO: 2).

Factor VIII B-domain Substituted Linkers

[00151] As described above, the linkage between the FVIII heavy chain and the light chain (e.g., the B-domain in wild-type Factor VIII) is altered in the Factor VIII variants described herein. Removal of the B-domain from wild type Factor VIII constructs does not appear to affect the activity of the activated enzyme (e.g., FVIIIa), presumably because the B-domain is removed during activation. Due to size constraints of AAV packaging capacity, B-domain deleted, truncated, and or linker substituted variants should improve the efficacy of the FVIII gene therapy construct. The most conventionally used B-domain substituted linker is that of SQ FVIII, which retains only 14 amino acids of the B domain as linker sequence. Another variant of porcine VIII ("OBI-1," described in U.S. Patent No. 6,458,563) is well expressed in CHO cells, and has a slightly longer linker of 24 amino acids. In some embodiments, the Factor VIII constructs encoded by the codon-altered polynucleotides described herein include an SQ-type B-domain linker sequence. In other embodiments, the Factor VIII constructs encoded by the codon-altered polynucleotides described herein include an OBI-1-type B-domain linker sequence.

[00152] Although the Factor VIII B-domain is dispensable for activity, the B-domain contains several residues that are post-translationally modified, e.g., by N- or O-linked glycosylation. In silico analysis (Prediction of N-glycosylation sites in human proteins, R. Gupta, E. Jung and S. Brunak, in preparation (2004)) of the wild-type Factor VIII B-domain predicts that at least four of these sites are glycosylated in vivo. It is thought that these modifications within the B-domain contribute to the post-translational regulation and/or half-life of Factor VIII in vivo. Thus, in some embodiments, the polypeptide linker of the encoded Factor VIII constructs described herein includes one or more glycosylation sequences, to allow for

glycosylation in vivo. In some embodiments, the polypeptide linker includes at least one consensus glycosylation sequence (e.g., an N- or O-linked glycosylation consensus sequence). In some embodiments, the polypeptide linker includes at least two consensus glycosylation sequences. In some embodiments, the polypeptide linker includes at least three consensus glycosylation sequences. In some embodiments, the polypeptide linker includes at least four consensus glycosylation sequences. In some embodiments, the polypeptide linker includes at least five consensus glycosylation sequences. In some embodiments, the polypeptide linker includes at least 6, 7, 8, 9, 10, or more consensus glycosylation sequences.

[00153] In some embodiments, the polypeptide linker contains at least one N-linked glycosylation sequence N-X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least two N-linked glycosylation sequences N-X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least three N-linked glycosylation sequences N-X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least four N-linked glycosylation sequences N-X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least five N-linked glycosylation sequences N-X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least 6, 7, 8, 9, 10, or more N-linked glycosylation sequences N-X-S/T, where X is any amino acid other than P, S, or T.

[00154] In some embodiments, the encoded Factor VIII polypeptides described herein include an SQ-type B-domain linker, including amino acids 760-762/1657-1667 of the wild-type human Factor VIII B-domain (FVIII-FL-AA; SEQ ID NO:39) (Sandberg et al. *Thromb. Haemost.* 85:93 (2001), the content of which is hereby incorporated herein by reference). In some embodiments, the SQ-type B-domain linker has one amino acid substitution relative to the corresponding wild-type sequence. In some embodiments, the SQ-type B-domain linker has two amino acid substitutions relative to the corresponding wild-type sequence.

[00155] In some embodiments, a glycosylation peptide is inserted into the SQ-type B-domain linker. In some embodiments, the glycosylation peptide is selected from those shown in Figure 8 (SEQ ID NOS: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35, respectively, in order of appearance). In some embodiments, the glycosylation peptide is encoded by a

respective polynucleotide shown in Figure 8 (SEQ ID NOS: 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, and 36, respectively, in order of appearance. In a particular embodiment, the NG5 glycosylation peptide (SEQ ID NO:15) is inserted into the SQ-type B-domain linker of the Factor VIII polypeptides described herein. In a particular embodiment, the NG5 glycosylation peptide is encoded by a polynucleotide having a nucleic acid sequence of SEQ ID NO:16.

[00156] In some embodiments, the SQ-type B-domain linker containing the glycosylation peptide encoded by a respective polynucleotide shown in Figure 17 (SEQ ID NOS: 40-53, respectively, in order of appearance. In a particular embodiment, the SQ-type B-domain linker containing the glycosylation peptide is encoded by a polynucleotide having a nucleic acid sequence of SEQ ID NO:43.

[00157] In some embodiments, the polypeptide linker (e.g., the SQ-type B-domain linker) includes a glycosylation peptide with high sequence identity to any one of SEQ ID NOS: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35, respectively, in order of appearance, as shown in Figures 8A-8B. In some embodiments, the glycosylation peptide has no more than two amino acid substitutions relative to any one of SEQ ID NOS: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35, respectively, in order of appearance, as shown in Figures 8A-8B. In some embodiments, the glycosylation peptide has no more than one amino acid substitution relative to any of SEQ ID NOS: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35, respectively, in order of appearance, as shown in Figures 8A-8B. In some embodiments, the glycosylation peptide has an amino acid sequence selected from any of SEQ ID NOS: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35, respectively, in order of appearance, as shown in Figures 8A-8B. In some embodiments, the glycosylation peptide is encoded by a polynucleotide sequence having high sequence identity to a corresponding nucleotide sequence selected from SEQ ID NOS: 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, and 36, respectively, in order of appearance, as shown in Figures 8A-8B.

[00158] In some embodiments, the glycosylation peptide has no more than two amino acid substitutions relative to SEQ ID NO:15 and is encoded by a polynucleotide sequence having at least 90% identity to SEQ ID NO:16. In some embodiments, the glycosylation peptide has no more than two amino acid substitutions relative to SEQ ID NO:15 and is encoded by a polynucleotide sequence having at least 95% identity to SEQ ID NO:16. In some embodiments,

the glycosylation peptide has no more than two amino acid substitutions relative to SEQ ID NO:15 and is encoded by a polynucleotide sequence having at least 98% identity to SEQ ID NO:16.

[00159] In some embodiments, the glycosylation peptide has no more than one amino acid substitution relative to SEQ ID NO:15 and is encoded by a polynucleotide sequence having at least 90% identity to SEQ ID NO:16. In some embodiments, the glycosylation peptide has no more than one amino acid substitution relative to SEQ ID NO:15 and is encoded by a polynucleotide sequence having at least 95% identity to SEQ ID NO:16. In some embodiments, the glycosylation peptide has no more than one amino acid substitution relative to SEQ ID NO:15 and is encoded by a polynucleotide sequence having at least 98% identity to SEQ ID NO:16.

[00160] In some embodiments, the glycosylation peptide has an amino acid sequence of SEQ ID NO:15 and is encoded by a polynucleotide sequence having at least 90% identity to SEQ ID NO:16. In some embodiments, the glycosylation peptide has an amino acid sequence of SEQ ID NO:15 and is encoded by a polynucleotide sequence having at least 95% identity to SEQ ID NO:16. In some embodiments, the glycosylation peptide has an amino acid sequence of SEQ ID NO:15 and is encoded by a polynucleotide sequence having at least 98% identity to SEQ ID NO:16.

Cis-Regulatory Elements

[00161] In some embodiments, a nucleic acid composition encoding a Factor VIII variant, as provided herein, also includes one or more cis-acting regulatory elements, e.g. promoter and/or enhancer elements, that drives gene expression in vivo,, which is operably linked to the polynucleotide encoding the Factor VIII variant. IN this context, “cis acting” means that the regulatory element are present on the same molecule of DNA as the gene they regulate. As will be appreciated by those in the art and discussed below, suitable regulatory elements for use in the invention include, but are not limited to, promoters, enhancer elements, polyadenylation signal elements, and/or inverted terminal repeat elements.

Promoters

[00162] The promoters of use in the invention are operably linked to the coding region, generally directly linked (e.g. no additional nucleotides are included between the promoter and the coding region) although in some embodiments, indirect linkages can be used, for example through the use of non-functional linkers, or in cases where additional regulatory elements that function “downstream” of the promoter but “upstream” of the coding region can be used. However, due to the size constraints of the vectors of the invention, this is generally not preferred.

Enhancer Elements

[00163] In some embodiments, one or more enhancer element is used in the Factor VIII variant construct. As is known in the art, enhancer elements generally drive expression in a tissue dependent fashion, e.g., predominantly in a specific tissue. In general, as described below, enhancer elements are generally positioned “upstream” of the promoter elements. Because Factor VIII is synthesized primarily in the liver, in some embodiments, the nucleic acid compositions described herein include a liver-specific regulatory element, which substantially limits expression of the encoded Factor VIII variant to hepatic cells.

[00164] Generally, liver-specific regulatory elements can be derived from any gene known to be exclusively expressed in the liver. WO 2009/130208 identifies several genes expressed in a liver-specific fashion, including, serpin peptidase inhibitor, clade A member 1, also known as α -antitrypsin (SERPINA1; GeneID 5265), apolipoprotein C-I (APOC1; GeneID 341), apolipoprotein C-IV (APOC4; GeneID 346), apolipoprotein H (APOH; GeneID 350); transthyretin (TTR; GeneID 7276), albumin (ALB; GeneID 213), aldolase B (ALDOB; GeneID 229), cytochrome P450, family 2, subfamily E, polypeptide 1 (CYP2E1; GeneID 1571), fibrinogen alpha chain (FGA; GeneID 2243), transferrin (TF; GeneID 7018), haptoglobin related protein (HPR; GeneID 3250). In some embodiments, the nucleic acid compositions described herein include a liver-specific regulatory element derived from the genomic loci of one or more of these proteins. Several examples of such elements are described in WO 2009/130208, the content of which is expressly incorporated herein by reference, in its entirety, for all purposes.

[00165] One example of a liver-specific regulatory element is from the transthyretin (TTR) gene, commonly referred to as "TTR_{enh}" or "TTR_{enh}." Hsieh J.L., et al., Cancer Sci., 100(3):537-45 (2009), the content of which is expressly incorporated herein by reference, in its

entirety, for all purposes. In some embodiments, nucleic acid composition encoding a Factor VIII variant, as described herein, includes a human TTR promoter element. In one embodiment, the human TTR promoter has a nucleic acid sequence with high sequence identity to the hTTR promoter shown in Figure 4 (SEQ ID NO:6). In some embodiments, the human TTR promoter has a nucleic acid sequence that is at least 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identical to SEQ ID NO:6.

[00166] Another example of a liver-specific regulatory element is from the SERPINA1 gene, as described in PCT Publication No. WO 2016/146757, the content of which is expressly incorporated herein by reference, in its entirety, for all purposes. An example of such an element is the CRM8 regulatory control element shown in Figure 4 (SEQ ID NO:5). In some embodiments, a SERPINA1-derived regulatory control element has a nucleic acid sequence that is at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to CRM8 (SEQ ID NO:5).

[00167] In some embodiments, a nucleic acid composition encoding a Factor VIII variant, as provided herein, includes one or more SERPINA1-derived regulatory control element, as exemplified by the constructs illustrated in Figures 11. In one embodiment, a Factor VIII polynucleotide includes one SERPINA1-derived regulatory control element (e.g., CRM8). In another embodiment, a Factor VIII polynucleotide includes two SERPINA1-derived regulatory control elements (e.g., CRM8). In yet other embodiments, a Factor VIII polynucleotide includes 3, 4, 5, 6, or more SERPINA1-derived regulatory control elements (e.g., CRM8).

[00168] In one embodiment, a nucleic acid composition encoding a Factor VIII variant, as provided herein, includes one or more SERPINA1-derived regulatory control element (e.g., CRM8) and a human TTR promoter element, as exemplified in Figure 11. In one embodiment, a nucleic acid composition encoding a Factor VIII variant, as provided herein, includes two CRM8 elements and a human TTR promoter element operably linked to the polynucleotide encoding the Factor VIII variant.

[00169] In some embodiments, a nucleic acid composition encoding a Factor VIII variant, as provided herein, includes one or more CRM8 element positioned upstream of a human TTR promoter. E.g., the one or more CRM8 element is positioned 5' of the TTR promoter in a double stranded construct, relative to the transcriptional orientation of the molecule. This means that in

a (+) single-stranded construct (e.g., where the single strand encodes for the Factor VIII variant), the one or more CRM8 element is positioned 5' of the TTR promoter.

[00170] As reported in Example 6, because of the large size of the Factor VIII variant constructs described herein, a small reduction in the number of nucleotides in a Factor VIII polynucleotide that is part of an AAV gene therapy vector can significantly increase the Factor VIII biopotency of the vector. Accordingly, in some embodiments, the one or more CRM8 element is directly attached to the 5' end of the TTR promoter, e.g., there are no extraneous nucleotides positioned between the CRM8 element and the TTR promoter. Likewise, in some embodiments, the TTR promoter is directly attached to the 5' end of the coding sequence, or to a translational initiation sequence (e.g., a Kozak sequence), for the Factor VIII variant polypeptide, e.g., there are no extraneous nucleotides positioned between the TTR promoter and the Factor VIII variant gene.

Polyadenylation Signals

[00171] In some embodiments, the regulatory element for the constructs described herein (e.g. a nucleic acid composition encoding a Factor VIII variant) is a regulatory element that is a polyadenylation signal, e.g., as illustrated in the examples constructs in Figure 11. The polyadenylation signal directs synthesis of a poly-A tail on the 3' end of the mRNA transcript generated from the Factor VIII polynucleotide. Accordingly, the polyadenylation signal is positioned 3' to the Factor VIII variant coding sequence. Non-limiting examples of polyadenylation signals that can be used in the Factor IX gene therapy constructs described herein include synthetic polyadenylation signals, poly-adenylation signals derived from a Simian virus 40 (SV40) late gene, a bovine growth hormone (BGH) polyadenylation signal, and a minimal rabbit β -globin (mRBG) gene polyadenylation signal.

[00172] In some embodiments, a nucleic acid composition encoding a Factor VIII variant, as provided herein, includes a synthetic polyadenylation signal, as exemplified by the constructs illustrated in Figures 11. In one embodiment, the synthetic polyadenylation signal has a nucleic acid sequence that is at least 90%, 95%, 97%, or 100% identical to the synthetic Poly-A signal shown in Figure 4 (SEQ ID NO:8).

[00173] As reported in Example 6, because of the large size of the Factor VIII variant constructs described herein, a small reduction in the number of nucleotides in a Factor VIII

polynucleotide that is part of an AAV gene therapy vector can significantly increase the Factor VIII biopotency of the vector. Accordingly, in some embodiments, the polyadenylation signal is directly attached to the 3' end of the coding sequence of the Factor VIII variant polypeptide, including one or more stop codons that are positioned at the end of the coding sequence. E.g., there are no extraneous nucleotides positioned between the Factor VIII variant gene and the polyadenylation sequence.

Inverted Terminal Repeats

[00174] In some embodiments, a nucleic acid composition encoding a Factor VIII variant, as provided herein, also includes adeno-associated virus (AAV) internal terminal repeats (ITRs) flanking the Factor VIII variant coding sequence and associated regulatory elements (e.g., promoters, enhancers, polyadenylation signals, etc.). The inverted terminal repeats each form a hairpin, which facilitates self-priming for primase-independent synthesis of the second DNA strand. The ITRs also help facilitate encapsulation within an AAV virion and integration of the AAV genome into the host cell genome.

[00175] In some embodiments, a nucleic acid composition encoding a Factor VIII variant, as provided herein, includes a 5' ITR having high sequence identity (e.g., at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical) to the AAV2 5' ITR shown in Figure 4 (SEQ ID NO:4). In some embodiments, a nucleic acid composition encoding a Factor VIII variant, as provided herein, includes a 3' ITR having high sequence identity (e.g., at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical) to the AAV2 3' ITR shown in Figure 4 (SEQ ID NO:9).

[00176] As reported in Example 6, because of the large size of the Factor VIII variant constructs described herein, a small reduction in the number of nucleotides in a Factor VIII polynucleotide that is part of an AAV gene therapy vector can significantly increase the Factor VIII biopotency of the vector. Accordingly, in some embodiments, the 5' ITR is directly attached to the 5' end of the liver-specific element (e.g., one or more CRM8 elements), such that no extraneous nucleotides are positioned between the 5' ITR sequence and the liver-specific element. Similarly, in some embodiments, the 3' ITR is directly attached to the 3' end of the polyadenylation signal, such that no extraneous nucleotides are positioned between the polyadenylation signal and the 3' ITR sequence.

IV. Factor VIII Expression Vectors

[00177] In some embodiments, the codon-altered polynucleotides described herein are integrated into expression vectors. Non-limiting examples of expression vectors include viral vectors (e.g., vectors suitable for gene therapy), plasmid vectors, bacteriophage vectors, cosmids, phagemids, artificial chromosomes, and the like. In general, there are two basic types of expression vectors of use in the invention: those that are used in cell culture to produce the Factor VIII polypeptides, and those that are used as gene therapy vectors to administer to patients such that endogeneous levels of Factor VIII (whether protein or activity) are increased.

[00178] Non-limiting examples of viral vectors include: retrovirus, e.g., Moloney murine leukemia virus (MMLV), Harvey murine sarcoma virus, murine mammary tumor virus, and Rous sarcoma virus; adenoviruses, adeno-associated viruses; SV40-type viruses; polyomaviruses; Epstein-Barr viruses; papilloma viruses; herpes viruses; vaccinia viruses; and polio viruses.

[00179] In many embodiments, the codon-optimized polynucleotides of the invention are used in gene therapy applications, such that the administration to a patient results in the production of Factor VIII as generally described herein. In general, gene therapy viral vectors are preferably replication deficient, such that the introduction of the gene therapy vector into a patient does not result in viral propagation.

[00180] Accordingly, in some embodiments, the codon-altered polynucleotides described herein are integrated into a gene therapy vector. In some embodiments, the gene therapy vector is a retrovirus, and particularly a replication-deficient retrovirus. In some embodiments, the codon-altered polynucleotides described herein are integrated into a retroviral expression vector. These systems have been described previously, and are generally well known in the art (Mann *et al.*, *Cell*, 33:153-159, 1983; Nicolas and Rubinstein, In: *Vectors: A survey of molecular cloning vectors and their uses*, Rodriguez and Denhardt, eds., Stoneham: Butterworth, pp. 494-513, 1988; Temin, In: *Gene Transfer*, Kucherlapati (ed.), New York: Plenum Press, pp. 149-188, 1986). In a specific embodiment, the retroviral vector is a lentiviral vector (*see*, for example, Naldini *et al.*, *Science*, 272(5259):263-267, 1996; Zufferey *et al.*, *Nat Biotechnol*, 15(9):871-875, 1997; Blomer *et al.*, *J Virol*, 71(9):6641-6649, 1997; U.S. Pat. Nos. 6,013,516 and 5,994,136). Protocols for the production of replication-deficient retroviruses are known in the art. For

review, see Kriegler, M., *Gene Transfer and Expression, A Laboratory Manual*, W.H. Freeman Co., New York (1990) and Murry, E. J., *Methods in Molecular Biology*, Vol. 7, Humana Press, Inc., Clifton, N.J. (1991).

[00181] A wide variety of vectors can be used for the expression of a Factor VIII polypeptide from a codon-altered polypeptide in cell culture, including eukaryotic and prokaryotic expression vectors. In certain embodiments, a plasmid vector is contemplated for use in expressing a Factor VIII polypeptide in cell culture. In general, plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell are used in connection with these hosts. The vector can carry a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. The plasmid will include the codon-altered polynucleotide encoding the Factor VIII polypeptide, operably linked to one or more control sequences, for example, a promoter.

[00182] Non-limiting examples of vectors for prokaryotic expression include plasmids such as pRSET, pET, pBAD, etc., wherein the promoters used in prokaryotic expression vectors include lac, trc, trp, recA, araBAD, etc. Examples of vectors for eukaryotic expression include: (i) for expression in yeast, vectors such as pAO, pPIC, pYES, pMET, using promoters such as AOX1, GAP, GAL1, AUG1, etc; (ii) for expression in insect cells, vectors such as pMT, pAc5, pIB, pMIB, pBAC, etc., using promoters such as PH, p10, MT, Ac5, OpIE2, gp64, polh, etc., and (iii) for expression in mammalian cells, vectors such as pSVL, pCMV, pRc/RSV, pcDNA3, pBPV, etc., and vectors derived from viral systems such as vaccinia virus, adeno-associated viruses, herpes viruses, retroviruses, etc., using promoters such as CMV, SV40, EF-1, UbC, RSV, ADV, BPV, and β -actin.

Adeno-Associated Virus (AAV) Vectors

[00183] In one embodiment, a codon-altered polynucleotide, as described herein, is integrated into an adeno-associated virus (AAV)-based gene therapy vector. AAV systems have been described previously and are generally well known in the art (Kelleher and Vos, *Biotechniques*, 17(6):1110-17 (1994); Cotten et al., *Proc Natl Acad Sci USA*, 89(13):6094-98 (1992); Curiel, *Nat Immun*, 13(2-3):141-64 (1994); Muzyczka, *Curr Top Microbiol Immunol*, 158:97-129 (1992); and Asokan A, et al., *Mol. Ther.*, 20(4):699-708 (2012), each incorporated herein by reference in their entireties for all purposes). Details concerning the generation and

use of rAAV vectors are described, for example, in U.S. Patent Nos. 5,139,941 and 4,797,368, each incorporated herein by reference in their entireties for all purposes.

[00184] Accordingly, in some embodiments, an AAV gene therapy vector is provided that includes a codon-altered polynucleotide (e.g., a nucleic acid composition), as described herein, that includes a nucleotide sequence with high sequence identity to the portions of CS12-FL-NA (SEQ ID NO:1) encoding a Factor VIII heavy and light chains, and a short, 14 amino acid, B-domain substituted linker (e.g., the “SQ” linker containing a furin cleavage site to facilitate maturation of an active FVIIIa protein in vivo), and that further includes one or more of the five X5 mutations (e.g., one, two, three, four, or all five of the I105V, A127S, G151K, M166T, L171P mutations (SPI), relative to the full-length human wild type Factor VIII sequence), and/or a short glycosylation peptide (e.g., NG5; SEQ ID NO:15) inserted into the B-domain substituted linker (e.g., an SQ linker).

[00185] In some embodiments, the AAV gene therapy vector includes a codon-altered Factor VIII polynucleotide that has a nucleotide sequence with high sequence identity to CS12-FL-NA (SEQ ID NO:1) and encodes a Factor VIII polypeptide with human Factor VIII heavy and light chains, and a short, 14 amino acid, B-domain substituted linker (the “SQ” linker) containing a furin cleavage site to facilitate maturation of an active FVIIIa protein in vivo, where the heavy chain of the Factor VIII polypeptide includes the five X5 mutations (I105V, A127S, G151K, M166T, and L171P (SPI), relative to the full-length human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker. In some embodiments, the Factor VIII polynucleotide has a nucleotide sequence has at least 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or 100% identity to CS12-FL-NA (SEQ ID NO:1).

[00186] In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide of the AAV gene therapy vector has an amino acid sequence with high sequence identity to CS12-FL-AA (SEQ ID NO: 2), including the five X5 mutations (I105V/A127S/G151K/M166T/L171P (SPI), relative to the full-length human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker. In some embodiments, the amino acid sequence of the encoded Factor VIII variant has at least 97%, 98%, 99%, 99.5%, 99.9%, or 100% identity to CS12-FL-AA (SEQ ID NO: 2), including the five X5 mutations (I105V/A127S/G151K/M166T/L171P (SPI), relative to the full-length

human wild type Factor VIII sequence), and an NG5 glycosylation peptide (SEQ ID NO:15) inserted into the SQ linker.

[00187] In some embodiments, the polynucleotide encoding the Factor VIII variant within the AAV gene therapy vector is operatively linked to a promoter element having a nucleic acid sequence with high sequence identity to the hTTR promoter shown in Figure 4 (SEQ ID NO:6). In some embodiments, the promoter element has a nucleic acid sequence that is at least 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identical to SEQ ID NO:6.

[00188] In some embodiments, the polynucleotide encoding the Factor VIII variant within the AAV gene therapy vector is operatively linked to one or more liver-specific regulatory elements having a nucleic acid sequence with high sequence identity to the CRM8 element shown in Figure 4 (SEQ ID NO:5). In some embodiments, the liver-specific regulatory elements have a nucleic acid sequence that is at least 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identical to SEQ ID NO:5. In some embodiments, as illustrated in Figure 11, the polynucleotide includes one CRM8 element. In some embodiments, as illustrated in Figure 11, the polynucleotide includes two CRM elements.

[00189] In some embodiments, the polynucleotide encoding the Factor VIII variant within the AAV gene therapy vector is operatively linked to one CRM8 element and a human TTR promoter element, as exemplified in Figure 11. In some embodiments, the polynucleotide encoding the Factor VIII variant within the AAV gene therapy vector is operatively linked to two CRM8 element and a human TTR promoter element, as exemplified in Figure 11. As described in Example 2, use of the hTTR promoter and either one or two liver-specific CRM8 elements increased in vivo exogenous Factor VIII biopotency in HepG2 cells by about 2-fold and 4-fold, respectively, as compared to use of mouse TTR promoter and enhancer sequences (compare vCS115 and vCS116 to vCS04 in Figure 12).

[00190] As reported in Example 6, because of the large size of the Factor VIII variant constructs described herein, a small reduction in the number of nucleotides in a Factor VIII polynucleotide that is part of an AAV gene therapy vector can significantly increase the Factor VIII biopotency of the vector. Accordingly, in some embodiments, the one or more CRM8 element is directly attached to the 5' end of the TTR promoter, e.g., there are no extraneous nucleotides positioned between the CRM8 element and the TTR promoter. Likewise, in some

embodiments, the TTR promoter is directly attached to the 5' end of the coding sequence, or to a translational initiation sequence (e.g., a Kozak sequence), for the Factor VIII variant polypeptide, e.g., there are no extraneous nucleotides positioned between the TTR promoter and the Factor VIII variant gene.

[00191] In some embodiments, the polynucleotide encoding the Factor VIII variant within the AAV gene therapy vector is operatively linked to a polyadenylation signal, e.g., as illustrated in the examples constructs in Figure 11. In one embodiment, the synthetic polyadenylation signal has a nucleic acid sequence that is at least 90%, 95%, 97%, or 100% identical to the synthetic Poly-A signal shown in Figure 4 (SEQ ID NO:8).

[00192] As reported in Example 6, because of the large size of the Factor VIII variant constructs described herein, a small reduction in the number of nucleotides in a Factor VIII polynucleotide that is part of an AAV gene therapy vector can significantly increase the Factor VIII biopotency of the vector. Accordingly, in some embodiments, the polyadenylation signal is directly attached to the 3' end of the coding sequence of the Factor VIII variant polypeptide, including one or more stop codons that are positioned at the end of the coding sequence. E.g., there are no extraneous nucleotides positioned between the Factor VIII variant gene and the polyadenylation sequence.

[00193] Internal terminal repeats are required cis-regulatory elements for AAV-based recombinant vectors. Accordingly, the Factor VIII variant encoding polynucleotides used in the AAV gene therapy vectors described herein include 5' and 3' ITR sequences. In some embodiments, the 5' ITR has high sequence identity (e.g., at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical) to the AAV2 5' ITR shown in Figure 4 (SEQ ID NO:4). In some embodiments, the 3' ITR has high sequence identity (e.g., at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical) to the AAV2 3' ITR shown in Figure 4 (SEQ ID NO:9).

[00194] As reported in Example 6, because of the large size of the Factor VIII variant constructs described herein, a small reduction in the number of nucleotides in a Factor VIII polynucleotide that is part of an AAV gene therapy vector can significantly increase the Factor VIII biopotency of the vector. Accordingly, in some embodiments, the 5' ITR is directly attached to the 5' end of the liver-specific element (e.g., one or more CRM8 elements), such that no extraneous nucleotides are positioned between the 5' ITR sequence and the liver-specific

element. Similarly, in some embodiments, the 3' ITR is directly attached to the 3' end of the polyadenylation signal, such that no extraneous nucleotides are positioned between the polyadenylation signal and the 3' ITR sequence.

[00195] In a specific embodiment, as exemplified in Figure 11, a polynucleotide included within an AAV gene therapy vector provided herein includes, in a 5' to 3' orientation, a 5' ITR sequence (e.g., having high sequence identity to SEQ ID NO:4), one or two CRM elements having high sequence identity to SEQ ID NO:5, an hTTR promoter element having high sequence identity to SEQ ID NO:6, a minimal Kozak consensus sequence having high sequence identity to SEQ ID NO:7, a Factor VIII variant polynucleotide having high sequence identity to SEQ ID NO:1, a polyadenylation sequence (e.g., having high sequence identity to SEQ ID NO:8), and a 3' AAV ITR sequence (e.g., having high sequence identity to SEQ ID NO:9). In some embodiments, the described polynucleotide has high sequence identity (e.g., at least 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or 100% identity) to the CS12-CRM8.2-Vr vector shown in Figure 3 (SEQ ID NO:3).

[00196] The AAV gene therapy vectors described herein are used with AAV capsid proteins that encapsulate the polynucleotide encoding the Factor VIII variant polypeptide, as described herein. That is, the delivery of the viral vector that will produce the Factor VIII is done by using a viral particle including the “shell” of capsid proteins that encapsulate the viral vector.

[00197] The serotype of an AAV vector is typically defined by the capsid proteins used. Several AAV serotypes have been characterized, including AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, and AAV9. Generally, any AAV serotype may be used for the Factor VIII gene therapy constructs described herein. However, the serotypes have different tropisms, e.g., they preferentially infect different tissues. In one embodiment, because Factor VIII is produced primarily in the liver, an AAV serotype for the disclosed gene therapy constructs is selected based on a liver tropism, found in at least serotypes AAV7, AAV8, and AAV9. Accordingly, in one embodiment, a Factor VIII gene therapy construct is an AAV7 serotype vector. In another embodiment, a Factor VIII gene therapy construct is an AAV8 serotype vector. In yet another embodiment, a Factor VIII gene therapy construct is an AAV9 serotype vector.

[00198] In some embodiments, plasmid polynucleotides that incorporate a codon-altered Factor VIII gene therapy genomes are also provided. The plasmids are useful for the production of the final AAV particles (e.g., AAV virions carrying the polynucleotide encoding the variant Factor VIII polypeptide), e.g., when introduced into a mammalian cell competent for recombinant AAV production (e.g., a cell harboring nucleic acids encoding AAV *rep* and *cap* genes, as well as helper genes (e.g., adenovirus genes) for AAV production. In some embodiments, the plasmids include regulatory elements that allow for replication of the plasmid (e.g., to scale-up the plasmid) in a host cell (e.g., a prokaryotic host cell, such as a bacterium, or a eukaryotic host cell, such as a yeast).

[00199] For instance, the sequence of an example plasmid carrying a codon-altered Factor VIII gene therapy genome is shown in Figure 5 (CS12-CRM8.2-Vrp; SEQ ID NO:10), in accordance with an embodiment of the disclosure. The CS12-CRM8.2-Vrp plasmid includes the CS12-CRM8.2-Vr Factor VIII gene therapy genome (shown as SEQ ID NO:3, in Figure 3), and plasmid backbone (shown as SEQ ID NO:54, in Figure 18). The genetic elements of the CS12-CRM8.2-Vr Factor VIII gene therapy genome are shown in Figure 4, as described above. The plasmid backbone of the CS12-CRM8.2-Vrp plasmid includes a pMB1 replicon (shown as SEQ ID NO:55 in Figure 19; Bolívar F., Life Sci., 25(10):807-17 (1979)) that facilitates replication of the plasmid in a bacterial host cell, and a Bla(ApR) ampicillin resistance gene (shown as SEQ ID NO:56 in Figure 19; Sutcliffe, P.N.A.S. U.S.A., 75(8):3737-41 (1978)) that facilitates selection of bacterial host cells transformed by the plasmid. The location of each element in the CS12-CRM8.2-Vrp plasmid is shown in Table 1, below.

Table 1. Elements present in the CS12-CRM8.2-Vrp plasmid.

Name of element	Nucleotide position	SEQ ID NO:
AAV2 5'-ITR	1 - 145	SEQ ID NO:4
CRM8	146 - 217; 219 - 290	SEQ ID NO:5
Human TTR promoter	291 - 523	SEQ ID NO:6
Kozak sequence	524 - 528	SEQ ID NO:7
FVIII-BDD coding sequence with X5 and NG5	529 - 4923	SEQ ID NO:1
ATG start codon	529 - 531	
TGA stop codons	4921 - 4923; 4924 - 4926	

X5 variant (I86V; A108S; G132K; M147T; L152P)	841 – 843; 907 – 909; 979 – 981; 1024 – 1026; 1039 – 1041	
NG5 sequence	2821 – 2841	SEQ ID NO:16
Synthetic polyA	4927 – 4975	SEQ ID NO:8
AAV2 3'-ITR	4976 – 5120	SEQ ID NO:9
Plasmid backbone	5121 – 7794	SEQ ID NO:54
• Rep (pMB1)	• 5541 – 6155	SEQ ID NO:55
• Bla(ApR)	• 6315 – 7175	SEQ ID NO:56

[00200] In one embodiment, a plasmid incorporating a Factor VIII variant gene therapy genome has high sequence identity to the CS12-CRM8.2-Vrp plasmid shown in Figure 5 (SEQ ID NO:10). In some embodiments, the described polynucleotide has high sequence identity (e.g., at least 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or 100% identity) to the CS12-CRM8.2-Vrp plasmid shown in Figure 5 (SEQ ID NO:10). In some embodiments, the described polypeptide includes a Factor VIII-BDD coding sequence, e.g., a sequence having at least 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or 100% identity to a Factor VIII-BDD coding sequence disclosed herein, and high sequence identity to the remaining portion of the CS12-CRM8.2-Vrp plasmid shown in Figure 5, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or 100% identity to nucleotides 1-528 and 4924-7794 of SEQ ID NO:10. In some embodiments, the described polynucleotide is a plasmid that comprises some or all of the elements shown in Table 1. In some embodiments, one or more of the elements shown in Table 1 are replaced by a comparable element.

Production of AAV Vectors

[00201] The codon-altered Factor VIII polynucleotides and viral vectors described herein are produced according to conventional methods for nucleic acid amplification and vector production. Several platforms have been developed for large-scale production of recombinant AAV vectors. A first platform is based on introduction of a plasmid containing the sequence for the desired viral genome into a mammalian cell containing polynucleotides encoding AAV *rep* and *cap* genes, as well as viral replication helper genes. For review, see, Kotin R.M., Hum. Mol. Genet., 20(R1):R2-6 (2011); Penaud-Budloo, M. et al., Mol Ther Methods Clin Dev., 8(8):166-80 (2018); and Aponte-Ubillus JJ et al., Appl Microbiol Biotechnol., 102(3):1045-54 (2018), the contents of which are expressly incorporated herein by reference, in its entirety, for all purposes.

A second platform is based on construction of a stable mammalian cell line having the desired viral genome integrated in the mammalian cell genome, e.g., by co-infection of a mammalian cell with wild type adenovirus and recombinant adenovirus harboring the sequence for the desired viral genome. For review, see, Penaud-Budloo, M. et al., *Mol Ther Methods Clin Dev.*, 8(8):166-80 (2018), the content of which is expressly incorporated herein by reference, in its entirety, for all purposes. A third platform is based on co-infection of a mammalian cell with a first recombinant HSV harboring the sequence for the desired viral genome and a second recombinant HSV encoding AAV *rep* and *cap* genes. For review, see, Penaud-Budloo, M. et al., *Mol Ther Methods Clin Dev.*, 8(8):166-80 (2018); and Aponte-Ubillus JJ et al., *Appl Microbiol Biotechnol.*, 102(3):1045-54 (2018), the contents of which are expressly incorporated herein by reference, in its entirety, for all purposes. A fourth platform is based on co-infection of insect cells with a first recombinant baculovirus harboring the sequence for the desired viral genome and a second recombinant baculovirus encoding AAV *rep* and *cap* genes. For review, see, Penaud-Budloo, M. et al., *Mol Ther Methods Clin Dev.*, 8(8):166-80 (2018); and Aponte-Ubillus JJ et al., *Appl Microbiol Biotechnol.*, 102(3):1045-54 (2018), the contents of which are expressly incorporated herein by reference, in its entirety, for all purposes. A fifth platform is based on based on introduction of a plasmid containing the sequence for the desired viral genome into a yeast cell containing polynucleotides encoding AAV *rep* and *cap* genes, as well as viral replication helper genes. For review, see, Aponte-Ubillus JJ et al., *Appl Microbiol Biotechnol.*, 102(3):1045-54 (2018), the content of which is expressly incorporated herein by reference, in its entirety, for all purposes.

V. Methods of Treating Hemophilia A

[00202] In some embodiments, the nucleic acid compositions (e.g., codon-altered polynucleotides encoding a Factor VIII variant) and gene therapy vectors (e.g., AAV particles containing a codon-altered polynucleotide encoding a Factor VIII variant) described herein are administered to a patient with hemophilia A for the treatment of hemophilia A, in accordance with known administrative methods. Methods for administering gene therapy vectors are well known in the art. These include, without limitation, intravenous administration, intramuscular injection, interstitial injection, and intra-hepatic administration (e.g., intra-hepatic artery or vein). For example, see Chuah MK et al., *Hum Gene Ther.*, 23(6):557-65 (2012); Chuah MK et al., J

Thromb Haemost., 10(8):1566-69 (2012); Chuah MK et al., J Thromb Haemost. 11 Suppl 1:99-110 (2013); VandenDriessche et al., Hum Gene Ther. 23(1):4-6 (2012); High KA, Blood, 120(23):4482-87 (2012); Matrai et al., Mol Ther., 18(3):477-90 (2010); and Matrai et al., Curr Opin Hematol., 17(5):387-92 (2010), the content of each of which is hereby incorporated by reference herein, for review.

[00203] Accordingly, the disclosure provides methods for treating a Factor VIII deficiency (e.g., hemophilia A). In some embodiments, the methods include administering to a patient in need thereof a nucleic acid composition (e.g., a codon-altered Factor VIII polynucleotide construct and/or recombinant AAV vector), as described herein. In some embodiments, the nucleic acid composition includes a codon-altered polynucleotide encoding a Factor VIII variant polypeptide, e.g., having high nucleic acid sequence identity (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or 100%) to CS12-FL-NA (SEQ ID NO:1). As described herein, in some embodiments, the codon-altered polynucleotide encoding the Factor VIII variant polypeptide is operably linked to a promoter (e.g., a human TTR promoter, as described herein) and one or more liver-specific regulatory elements (e.g., one or two CRM8 elements, as described herein).

[00204] In some embodiments, the nucleic acid composition is part of a mammalian gene therapy vector. In a specific embodiment, the mammalian gene therapy vector is a viral vector, e.g., a lentivirus, retrovirus, adeno virus, or adeno-associated virus vector.

[00205] In one embodiment, the gene therapy vector is an adeno-associated virus (AAV) particle harboring a viral vector encoding the codon-altered Factor VIII variant coding sequence. Generally, the viral vector includes inverted terminal repeats (ITR) at each termini, one or more expression regulatory elements (e.g., a promoter (e.g., a human TTR promoter, as described herein) and one or more liver-specific regulatory elements (e.g., one or two CRM8 elements, as described herein)), a codon-altered Factor VIII coding sequence, and a poly-A signal sequence.

Assessing Therapeutic Efficacy

[00206] The therapeutic efficacy of a hemophilia A treatment can be evaluated, for example, by measuring the Factor VIII-dependent coagulation potential of blood from a subject being treated. Metrics for assessing coagulation potential include, without limitation, in vitro activated partial thromboplastin time assay (APPT), Factor IX chromogenic activity assays, blood clotting times, and Factor VIII antigen levels (e.g., using a Factor VIII-specific ELISA). It

should be noted that a therapeutic dose need not result in wild-type levels of Factor VIII in a patient; rather, sufficient expression to decrease symptoms in a meaningful or measurable way is considered therapeutic for the purposes of the disclosure.

[00207] According to the National Hemophilia Foundation, a subject is classified as having mild hemophilia A when their blood plasma contains between 6% and 49% of the Factor VIII activity of normal human blood plasma. Subjects with mild hemophilia A typically experience bleeding only after serious injury, trauma or surgery. In many cases, mild hemophilia is not diagnosed until an injury, surgery or tooth extraction results in prolonged bleeding. The first episode may not occur until adulthood. Women with mild hemophilia often experience menorrhagia, heavy menstrual periods, and can hemorrhage after childbirth.

[00208] According to the National Hemophilia Foundation, a subject is classified as having moderate hemophilia A when their blood plasma contains between 1% and 5% of the Factor VIII activity of normal human blood plasma. Subjects with moderate hemophilia A tend to have bleeding episodes after injuries. Bleeds that occur without obvious cause are called spontaneous bleeding episodes.

[00209] According to the National Hemophilia Foundation, a subject is classified as having severe hemophilia A when their blood plasma contains less than 1% of the Factor VIII activity of normal human blood plasma. Subjects with severe hemophilia A experience bleeding following an injury and may have frequent spontaneous bleeding episodes, often into their joints and muscles.

[00210] In some embodiments, normal human blood plasma is defined as containing 1 IU of Factor VIII activity per mL. Thus, in some embodiments, blood plasma from a subject classified with mild hemophilia A contains between 0.06 and 0.49 IU of Factor VIII activity per mL. In some embodiments, blood plasma from a subject classified with moderate hemophilia A contains between 0.01 and 0.05 IU of Factor VIII activity per mL. In some embodiments, blood plasma from a subject classified with severe hemophilia A contains between 0.01 and 0.05 IU of Factor VIII activity per mL.

[00211] In some embodiments, a therapy is therapeutically effective when it lessens the severity of a symptom of hemophilia A, e.g., by raising the average level of Factor VIII activity in the subject's blood. Accordingly, in some embodiments, hemophilia A therapy is

therapeutically effective when it raises the average level of Factor VIII activity in the subject's blood/plasma. In some embodiments, a therapeutically effective treatment raises the average level of Factor VIII activity in the subject's blood/plasma by at least 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, or more.

[00212] In some embodiments, a therapeutically effective treatment raises the average level of Factor VIII activity in the subject's blood such that the subject is classified as having a less severe form of hemophilia A. For example, in one embodiment, a subject originally classified with severe hemophilia A is reclassified with moderate hemophilia A or mild hemophilia A after undergoing a therapeutically effective treatment. In another embodiment, a subject originally classified with moderate hemophilia A is reclassified with mild hemophilia A after undergoing a therapeutically effective treatment. In another embodiment, a subject originally classified with mild hemophilia A is reclassified as not having hemophilia A after undergoing a therapeutically effective treatment.

Formulations

[00213] Compositions for use in treatment of hemophilia A are provided herein. Such compositions contain a therapeutically effective amount of a nucleic acid composition, e.g., an AAV gene therapy vector including a codon-altered polynucleotide encoding for Factor VIII, as described herein. Therapeutically effective amounts of the codon-altered VIII polynucleotide (e.g., an AAV gene therapy vector including the codon-altered Factor VIII coding sequence) are mixed with a suitable pharmaceutical carrier or vehicle for, e.g., systemic administration. Final formulation of the codon-altered Factor VIII polynucleotides disclosed herein will be within the abilities of those skilled in the art.

Dosages

[00214] The nucleic acid compositions of the invention are administered to patients in need thereof. The amount or dose of the therapeutic gene therapy agent administered depends on factors such as the particular codon-altered VIII polynucleotide construct, the delivery vector used, the severity of the disease, and the general characteristics of the subject. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (see, e.g., Lieberman, *Pharmaceutical Dosage Forms* (vols. 1-3, 1992); Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999); Pickar,

Dosage Calculations (1999); and Remington: The Science and Practice of Pharmacy, 20th Edition, 2003, Gennaro, Ed., Lippincott, Williams & Wilkins). It is within the abilities of the skilled physician to determine a particular dosage and dosing regimen for treatment of a particular subject.

VI. Examples

Example 1 - Improvement of a single-stranded AAV8 vector construct expressing FVIII by addition of N-glycosylation linker sequences

[00215] To address whether additional N-glycosylation sites within the SQ sequence of BDD-FVIII increases FVIII protein expression, a set of short peptide sequences containing putative N-linked glycosylation sites were designed. Previously, McIntosh et al. (Blood 121(17):3335-44 (2013)) showed that the concept of N-glycosylation with 6 potential sites (“V3 peptide”) resulted in enhanced FVIII expression levels in the plasma of mice. Interestingly, in silico prediction of the “V3 peptide” identified two out of 6 sites to be potentially N-glycosylated in vivo.

[00216] 12 different linker sequences, shown as NG1-NG21 in Figures 8A-8B, were designed within the context of a codon-optimized BDD-FVIII “CS01” sequence (see, WO 2017/083762, the contents of which are hereby incorporated herein by reference) by applying the program NetNGlyc-Database (Steentoft et al., H. EMBO J, 32(10):1478-88, 2013), with the goal of generating short sequences containing multiple N-glycosylation sites. The NetNGly platform combines nine neural networks analyzing human protein sequences for their N-glycosylation pattern. Based on the NetNGly database, the designed peptides were analyzed for the likelihood to transfer N-glycosylation as post-translational modification, as described in Steentoft et al., H. EMBO J, 32(10):1478-88, (2013), the contents of which are hereby incorporated by reference herein. Of the 12 novel alternative peptides, four promising NG linkers with predicted low immunogenicity (as explained below) were inserted into the 14 amino acid-sized SQ sequence (SFSQN – novel peptide – PPVLKRHQR) of a codon-optimized BDD-FVIII sequence termed “CS04” (SEQ ID NO:10).

[00217] Post-translational modification of the predicted N-glycosylation sites was confirmed experimentally for three vectors including vNG4/CS04, vNG5/CS04, and

vNG16/CS04, by transfecting human hepatic Huh-7 cells and detecting modified BDD-FVIII by anti-FVIII Western blot analyses (Figure 9). In comparison to BDD-FVIII containing the SQ sequence only, larger-sized heavy chains of the modified versions of FVIII were detected by gel electrophoresis, indicating the presence of novel N-glycosylation sites. In addition, AAV8 infection of the human liver cell line HepG2 confirmed the N-glycosylation as shown exemplary for the expression vector vNG5/CS04 and the highly secreting X5 variant vX5/NG5/CS125 (Figure 10; vX5/NG5/CS125 will be described in Examples 3 and 5).

[00218] The focus for the linker sequence design was to generate short peptide sequences (7-17 amino acid residues) accompanied by low risk of immunogenicity. For this, in silico immunogenicity profiling Epibase™ (Lonza) was applied (HLA binding). Epibase™'s epitope prediction method comprises a structural and a statistical layer for the prediction of immunogenicity. The structural part estimates the binding affinity based on the Pepscope technology, as described in Desmet et al., *Proteins*. 58(1):53-69 (2005), the contents of which are hereby incorporated herein by reference. The statistical part extracts information from peptide sequences and their experimental binding affinities. Based on the critical epitope counts, the affected HLA class II allotypes (Krischmann et al., *J Immunol*. 15;155(12):5655-62 (1995); Verreck et al., *Int Immunol*. 8(3):397-404 (1996), the contents of which are hereby incorporated herein by reference) and DRB1 (Laupeze et al., *Hum Immunol*. 60(7):591-7 (1999), the content of which is hereby incorporated herein by reference) risk score were used to estimate the immunogenic risk of the proteins. Based on this scoring system, the immunogenic risk of peptide NG16 is moderate, that of NG4 and NG10 is low and that of NG5 is even lower than the unmodified SQ sequence of BDD-FVIII (Table 2).

Table 2. FVIII activity two weeks after AAV treatment *in vivo*.

AAV construct	Linker			Epibase	FVIII knock-out mice		“line E” mice	
	ID	Number of amino acid residues	Number of predicted N-glycosylation sites		FVIII activity [IU/mL]	Fold increase vs Orth04	FVIII activity [IU/mL]	Fold increase vs Orth04
vCS04	-	0	0	low	1.7 ± 0.5	1.0	1.4 ± 0.6	1.0
vNG4/CS04	NG4	11	3	low	3.1 ± 0.6	1.8	n.d.	n.d.
vNG5/CS04	NG5	7	2	very low	3.8 ± 0.8	2.3	4.0 ± 2.8	3.0
vNG10/CS04	NG10	17	2	low	4.1 ± 0.5	2.4	n.d.	n.d.
vNG16/CS04	NG16	9	2	moderate	2.5 ± 0.8	1.4	n.d.	n.d.

[00219] Further, vectors vNG4/CS04, vNG5/CS04, vNG10/CS04 and vNG16/CS04 were analyzed in exon 16 FVIII knock-out mouse model generated in the laboratory of Haig Kazazian at the University of Pennsylvania (Table 2). Two weeks post AAV8 infection, FVIII expression levels were determined by chromogenic activity of mouse plasma samples. Two weeks post AAV8 infection at a dose of 4.0×10^{12} , FVIII expression levels were 1.4 to 2.4-fold higher than the peptide-free vCS04 construct. Overall, construct vNG5/CS04 displayed the most favorable features, including the high FVIII expression levels (3.8 IU/mL), the lowest immunogenic risk, and the shortest peptide in size (7 amino acid residues). Construct vNG5/CS04 was further tested in another FVIII mouse model, the “line E” model, which is immunologically tolerant to human FVIII (described in Reipert et al., Haemophilia. 16 (Suppl 5):47-53 (2010) and van Helden et al., Blood 118(13):3698-707 (2011), the contents of which are hereby expressly incorporated herein by reference). In this independent mouse model, increased FVIII expression of vNG5/CS04 was confirmed and assessed to be 3-fold higher than vCS04.

[00220] Taken together, an improved vector termed vNG5/CS04 containing a short novel N-glycosylation peptide sequence with a very low immunogenic risk and enhanced levels of FVIII expression *in vivo* was developed.

Example 2 - Improvement of a single-stranded AAV8 vector construct expressing FVIII by promoter/enhancer replacement

[00221] The promoter cassette of the initial FVIII-expressing AAV8-based single-stranded vector construct termed vCS04 contains a core promoter and enhancer sequence derived from the liver-specific murine transthyretin (TTR) gene (Yan et al., EMBO J, 9:869-78 (1990), the content of which is hereby incorporated by reference herein), see Figure 11. This liver-specific mTTR promoter/enhancer cassette was modified two-fold, firstly the murine core (basal) promoter sequence was fully replaced by the corresponding human sequence. Secondly, the murine enhancer sequence was replaced by the recently described liver-specific cis-regulatory module CRM8 (Nair et al., Blood, 123:3195-99 (2014) and Chuah et al., Mol Ther, 22:1605-13 (2014), the contents of which are hereby incorporated herein by reference). One or two CRM8 elements were inserted upstream of the human TTR core (basal) promoter, resulting in vector constructs vCS115 and vCS116, respectively (Figure 11).

[00222] The strength of the novel promoter cassette was assessed by in vitro and in vivo analyses, in a human liver-derived HepG2 cell line and in “line E” mice at a dose of 4.0E+12 vg/kg. In vitro, the constructs with one CRM8 element (vCS115) and two CRM8 elements (vCS116) resulted in 2.2- and 3.7-fold higher biopotency units, respectively, in comparison to the reference (vCS04) (Figure 12). In vivo, the CRM8/hTTR promoter cassette performed comparable to the mTTR promoter/enhancer (Figure 12).

[00223] The CRM8-related augmenting effect on FVIII expression was further evaluated in triple sets of constructs, in which the described promoter cassettes including mTTR promoter/enhancer, 1xCRM8/hTTR, and 2xCRM8/hTTR were combined with novel modifications of BDD-FVIII (see below, Examples 3 -5). Overall, the in vitro data reveal an increase in biopotency units based on the presence of one and two CRM8 elements by directly comparing (1) vNG5/CS04 with vNG5/CS117 (2.5 fold higher) and vNG5/CS118 (4.0-fold higher), (2) vX5/CS24 with vX5/CS101 (1.8-fold higher) and vX5/CS105 (5.3-fold higher) and (3) vX5/NG5/CS125 with vX5/NG5/CS119 (4.4-fold higher) and vX5/NG5/CS120 (6.2-fold higher). No CRM8-dependent effect on FVIII levels was observed in vivo. Nevertheless, the in vivo data clearly show that both the novel 1xCRM8/hTTR and the 2xCRM8/hTTR promoter cassette perform equally well in the mouse model compared to the mTTR promoter/enhancer.

Example 3 - Improvement of a single-stranded AAV8 vector construct expressing FVIII by introduction of the “NG5” variant

[00224] Based on the data shown in Example 1, the N-glycosylation linker NG5 in construct vNG5/CS04 was selected to be combined with the novel human liver-specific promoters 1xCRM8/hTTR and 2xCRM8/hTTR, resulting in constructs vNG5/CS117 and vNG5/CS118 (Figure 11). The vNG5/CS04 construct contains the mTTR promoter/enhancer cassette (Figure 11). The in vivo biopotency levels of vNG5/CS04 were 1.9-fold higher than vCS04, attributing the positive effect on the NG5 linker sequence (Figure 12). In vitro, an increase in biopotency was observed for the CRM8 element(s)-containing vectors vNG5/CS117 and vNG5/CS118. In vivo, expression levels of vNG5/CS117 and vNG5/CS118 were similar to vNG5/CS04) (Figure 12).

Example 4 - Improvement of a single-stranded AAV8 vector construct expressing FVIII by introduction of the “X5” variant

[00225] In order to further improve the FVIII-expressing vectors, the X5 variant (described as mutation ‘m2’ in WO 2017/083762, the content of which is hereby incorporated herein by reference) was introduced into BDD-FVIII. The X5 variant contains five porcine FVIII amino acid residues within the A1 domain of the heavy chain that confer efficient secretion to human FVIII (Cao et al., 2014, ASGCT abstract #460; details of variants disclosed in oral presentation). Specifically, the X5 variant of BDD-FVIII was combined with the mTTR promoter/enhancer, resulting in vX5/CS24, and with one or two CRM8 element(s) plus hTTR promoter, resulting in constructs vX5/CS101 and vX5/CS105 (Figure 11). The three novel constructs were analyzed in vitro in a liver-derived HepG2 cell line and in vivo in “line E2” mice, a transgenic FVIII knock-out mouse line which expresses minimal amounts of human FVIII cDNA leading to immunological tolerance to human FVIII (van Helden PM, et al., Blood 118(13):3698-707 (2011), the content of which is incorporated by reference herein, in its entirety, for all purposes), and compared to the reference construct vCS04. In vivo, vX5/CS24, vX5/CS101, and vX5/CS105 resulted in enhanced levels of FVIII activity by a factor of 2.5 – 3.1 (Figure 12). In vitro, the increase in biopotency ranged between 4.0 and 21.2. The in vitro test system revealed a 4.0-fold increase stemming from the X5 variant (construct vX5/CS24), and

was further increased by a factor of 7.3 and 21.2 by the constructs containing X5 and one or two copies of the CRM8 element (constructs vX5/CS101 and vX5/CS105, respectively).

Example 5 - Improvement of a single-stranded AAV8 vector construct expressing FVIII by introducing both the “X5” and the “NG5” variant

[00226] In a further set of vectors, the N-glycosylation linker NG5 (Example 3) and the X5 variant (Example 4) were introduced into BDD-FVIII in parallel, and additionally combined with three different promoters, including the mTTR promoter/enhancer, 1x CRM8/hTTR, and 2xCRM8/hTTR, resulting in constructs vX5/NG5/CS125, vX5/NG5/CS119 and vX5/NG5/CS120, respectively (Figure 11). In comparison to vCS04, in vivo and in vitro biopotencies of vX5/NG5/CS125, vX5/NG5/CS119, and vX5/NG5/CS120 were elevated by a factor of 3.6 – 5.5 and 3.2 – 19.8, respectively. The introduction of the two novel modifications X5 and NG5 reveal further increased expression, in vivo, as shown for the three series of constructs each having a common promoter. For example, for the series with the 2xCRM8/hTTR promoter (constructs vCS116, vNG5/CS118, vX5/CS105, and vX5/NG5/CS120), the presence of NG5 alone raises the FVIII expression level from 1.9 to 2.9 IU/mL, the presence of X5 alone raises the FVIII expression from 1.9 to 5.1 IU/mL, and the presence of both X5 and NG5 raises the FVIII expression from 1.9 to 11.4 IU/mL.

[00227] Taken together, construct vX5/NG5/CS120 carrying the novel 2xCRM8/hTTR promoter cassette and additionally a codon-optimized BDD-FVIII “CS04” nucleotide sequence with two modifications, the X5 variant and the NG5 sequence, was determined to be the construct with highest in vitro and in vivo biopotencies.

Example 6 - Nucleotide reduction of a single-stranded AAV8 vector construct

[00228] In order to reduce the vector size of the slightly oversized vector construct vX5/NG5/CS120 which should result in more efficient packaging, all non-function DNA sequences within the flanking ITR's were deleted. This resulted in the reduction of a total of 71 nucleotides, from a 5191 to a 5120 nucleotide-sized expression cassette. The deletions included 19 nucleotides between the 5'-ITR and the CRM8 sequence, 9 nucleotides between the human TTR promoter and the Kozak sequence, 27 nucleotides between the BDD-FVIII coding sequence and the synthetic polyadenylation site, and 16 nucleotides between the synthetic polyadenylation

site and the 3'-ITR sequence. The size-reduced expression cassette, termed vX5/NG5/CS12, consists of a promoter with two CRM8 elements and the core human TTR promoter sequence, the BDD-FVIII sequence including the X5 variant as well as the NG5 sequence and a synthetic polyadenylation site and is flanked by two AAV2-based inverted terminal repeats. Construct vX5/NG5/CS12 is schematically shown in Figure 11.

[00229] Vector genome integrity of the AAV vector genome preparations vCS04, vX5/NG5/CS120 and vX5/NG5/CS12 was addressed by agarose gel electrophoresis. The results shown in Figure 13 demonstrate that the vCS04, vX5/NG5/CS120 and vX5/NG5/CS12 viral vectors have a similar-sized genome, indicated by a distinct band of approximately 5kb. Despite a calculated vector size of approx. 5.2 kb, the genome is a homogenous band confirming correct packaging of the marginally oversized genome (relative to an AAV wild-type genome of 4.7 kb). The shorter vX5/NG5/CS12 variant is preferred.

[00230] A set of AAV8-based viral vectors including vCS04, vX5/NG5/CS120, and vX5/NG5/CS12 were administered to FVIII F17 knock-in mice at vector doses of 5×10^{11} vg/kg, 1×10^{12} vg/kg, and 4×10^{12} vg/kg and FVIII activity levels were determined at day 14. As shown in Figure 14, both vX5/NG5/CS12 and vX5/NG5/CS120 resulted in comparable expression levels of approximately 4 IU/mL at the vector dose of 1×10^{12} vg/kg. In contrast to the reference construct vCS04 with very low FVIII expression levels of 0.3 IU/mL at a dose of 1×10^{12} vg/kg, the improved vectors vX5/NG5/CS12 and vX5/NG5/CS120 showed elevated FVIII expression levels by a factor of approximately 14 (Figure 14). Even at the dose of 5×10^{11} vg/kg expression levels of 1.5 IU/mL of FVIII could be obtained for the vector vX5/NG5/CS12. In accordance to the in vivo data, the in vitro biopotencies of vX5/NG5/CS120 and vX5/NG5/CS12 are strongly elevated approximately by a factor of 17 and 24, respectively, infecting HepG2 cells at equal multiplicity of infection (MOI) (Figure 14). Thus, because all vectors were administered at the same concentration (multiplicity of infection), the differences in biopotency are due to the variants used and are not size-dependent.

WHAT IS CLAIMED IS:

1. A nucleic acid composition comprising a Factor VIII polynucleotide encoding a Factor VIII protein, said Factor VIII polynucleotide having the nucleic acid sequence of CS12-FL-NA (SEQ ID NO:1).
2. The nucleic acid composition of claim 1, further comprising a promoter polynucleotide operatively linked to the Factor VIII polynucleotide, wherein the promoter polynucleotide has the nucleic acid sequence of hTTR (SEQ ID NO:6)
3. The nucleic acid composition of claim 2, wherein the promoter is directly attached to the Factor VIII polynucleotide.
4. The nucleic acid composition according to any one of claims 1 to 3, further comprising a liver-specific element operatively linked to the Factor VIII polynucleotide, wherein the liver-specific element has a sequence of CRM8 (SEQ ID NO:5).
5. The nucleic acid composition of claim 4 wherein a second liver-specific element is operatively linked to said Factor VIII polypeptide.
6. The nucleic acid composition of claim 4 or 5, wherein the liver-specific element and the promoter are directly attached.
7. The nucleic acid composition of claim 1, having a nucleic acid sequence comprising CS12-CRM8.2-Vr (SEQ ID NO: 3).
8. A mammalian gene therapy vector comprising a nucleic acid composition according to any one of claims 1 to 7.
9. The mammalian gene therapy vector of claim 8, wherein the mammalian gene therapy vector is an adeno-associated virus (AAV) vector.

10. The mammalian gene therapy vector of claim 9, wherein the AAV vector is a serotype 8 adeno-associated virus (AAV-8) vector.
11. The mammalian gene therapy vector according to any one of claims 8 to 10, wherein the nucleic acid composition is a single-stranded polynucleotide encoding the Factor VIII protein.
12. An adeno-associated virus (AAV) particle comprising capsid proteins encapsulating a nucleic acid composition according to any one of claims 1 to 7.
13. The AAV particles of claim 12, wherein the capsid proteins comprise serotype 8 adeno-associated virus (AAV-8) capsid proteins.
14. The AAV particles of claim 12 or 13, wherein the nucleic acid composition is a single-stranded polynucleotide encoding the Factor VIII protein.
15. A nucleic acid composition according to any one of claims 1 to 7 for treating hemophilia A.
16. Use of a nucleic acid composition according to any one of claims 1 to 7 for the manufacture of a medicament for the treatment of hemophilia A.
17. A mammalian gene therapy vector according to any one of claims 8 to 11 for treating hemophilia A.
18. Use of a mammalian gene therapy vector according to any one of claims 8 to 11 for the manufacture of a medicament for the treatment of hemophilia A.
19. An adeno-associated virus (AAV) particle according to any one of claims 12 to 14 for treating hemophilia A.

20. Use of an adeno-associated virus (AAV) particle according to any one of claims 12 to 14 for the manufacture of a medicament for the treatment of hemophilia A.

21. A method for producing an adeno-associated virus (AAV) particle comprising introducing a nucleic acid composition according to any one of claims 1 to 7 into a eukaryotic host cell: wherein:

the nucleic acid composition comprises a Factor VIII polynucleotide having the nucleic acid sequence of CS12-FL-NA (SEQ ID NO:1) flanked by a 5' inverted terminal repeat sequence (5' ITR) and a 3' inverted terminal repeat sequence (3' ITR), and

the mammalian host cell comprises one or more polynucleotides encoding an AAV *rep* gene, an AAV *cap* gene, and viral replication helper genes.

22. The method of claim 21, wherein the nucleic acid composition is a plasmid having the nucleic acid sequence of CS12-CRM8.2-Vrp (SEQ ID NO:10).

23. The method of claim 21 or 22, wherein the AAV *cap* gene is a serotype 8 adeno-associated virus (AAV-8) *cap* gene.