Factor IX deficiencies (e.g., hemophilia B). The above deficiencies and other problems associated with the treatment of Factor IX deficiencies (e.g., hemophilia B) are reduced or eliminated by the disclosed codon-altered nucleic acids encoding Factor IX proteins.

[0014] In one example, nucleic acid compositions (e.g., codon-altered polynucleotides) encoding Factor IX and Factor IX variants are described. The nucleic acid compositions include polynucleotides with high sequence identity to the CS02, CS03, CS04, CS05, and CS06 sequences encoding Factor IX, as described herein. The nucleic acid compositions described herein provide increased Factor IX expression relative to wild-type Factor IX coding sequences. 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 IX.

[0015] In some examples, a nucleic acid composition includes a polynucleotide encoding Factor IX that has a nucleotide sequence with at least 95% sequence identity (e.g., at least 95%, 96%, 97%, 98%, 99%, or 100% sequence identity) to a disclosed sequence selected from CS02-FL-NA (SEQ ID NO:5), CS02-MP-NA (SEQ ID NO:13), CS03-FL-NA (SEQ ID NO:6), CS03-MP-NA (SEQ ID NO:14), CS04-FL-NA (SEQ ID NO:7), CS04-MP-NA (SEQ ID NO:15), CS05-FL-NA (SEQ ID NO:8), CS05-MP-NA (SEQ ID NO:16), CS06-FL-NA (SEQ ID NO:9), and CS06-MP-NA (SEQ ID NO:17).

[0015a] In one aspect, the present disclosure refers to a nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX polynucleotide comprising a nucleic acid sequence that is least 97% identical to the nucleic acid sequence of CS06-MP-NA (SEQ ID NO:17).

[0015b] In another aspect, the present disclosure refers to a nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX polynucleotide comprising a nucleic acid sequence that is least 95% identical to the nucleic acid sequence of CS02-MP-NA (SEQ ID NO:13).

[0015c] In yet another aspect, the present disclosure refers to a nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX

polynucleotide comprising a nucleic acid sequence that is least 98% identical to the nucleic acid sequence of CS03-MP-NA (SEQ ID NO:14).

[0015d] In yet another aspect, the present disclosure refers to a nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX polynucleotide comprising a nucleic acid sequence that is least 99% identical to the nucleic acid sequence of CS04-MP-NA (SEQ ID NO:15).

[0015e] In a further aspect, the present disclosure refers to a nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX polynucleotide comprising a nucleic acid sequence that is least 98% identical to the nucleic acid sequence of CS05-MP-NA (SEQ ID NO:16).

[0016] In some examples, a nucleic acid composition includes a polynucleotide encoding Factor IX that has a nucleotide sequence with at least 95% sequence identity (e.g., at least 95%, 96%, 97%, 98%, 99%, or 100% sequence identity) to a disclosed sequence encoding a Factor IX light chain (e.g., CS02-LC-NA (SEQ ID NO:42), CS03-LC-NA (SEQ ID NO:44), CS04-LC-NA (SEQ ID NO:46), CS05-LC-NA (SEQ ID NO:48), or CS06-LC-NA (SEQ ID NO:50)) and a disclosed sequence encoding a Factor IX heavy chain (e.g., CS02-HC-NA (SEQ ID NO:41), CS03-HC-NA (SEQ ID NO:43), CS04-HC-NA (SEQ ID NO:45), CS05-HC-NA (SEQ ID NO:47), or CS06-HC-NA (SEQ ID NO:49)).

[0017] In some examples, a nucleic acid composition includes a polynucleotide that encodes a Factor IX polypeptide having a light chain, a heavy chain, and a polypeptide linker joining the C-terminus of the light chain to the N-terminus of the heavy chain (e.g., an activation

peptide). The light chain of the Factor IX polypeptide is encoded by a first nucleotide sequence with high sequence identity to one of CS02-LC-NA (SEQ ID NO:42), CS03-LC-NA (SEQ ID NO:44), CS04-LC-NA (SEQ ID NO:46), CS05-LC-NA (SEQ ID NO:48), and CS06-LC-NA (SEQ ID NO:50). The heavy chain of the Factor IX polypeptide is encoded by a second nucleotide sequence with high sequence identity to one of CS02-HC-NA (SEQ ID NO:41), CS03-HC-NA (SEQ ID NO:43), CS04-HC-NA (SEQ ID NO:45), CS05-HC-NA (SEQ ID NO:47), and CS06-HC-NA (SEQ ID NO:49). The polypeptide linker comprises a protease cleavage site (e.g., two Factor XIa cleavage sites).

[0018] In some examples of the polynucleotides described above, the polypeptide linker has an amino acid sequence with high sequence identity to the wild-type Factor IX activation peptide FIX-AP-AA (SEQ ID NO:56; amino acids 192-226 of FIX-FL-AA (SEQ ID NO:2)). In some examples, the polypeptide linker is encoded by a third nucleic acid sequence having high sequence identity to one of CS02-AP-NA (SEQ ID NO:57), CS03-AP-NA (SEQ ID NO:58), CS04-AP-NA (SEQ ID NO:59), CS05-AP-NA (SEQ ID NO:60), and CS06-AP-NA (SEQ ID NO:61).

[0019] In some examples, the codon-altered polynucleotides described herein encode a pre-pro-Factor IX polypeptide, e.g., where the encoded Factor IX protein includes a signal peptide and a pro-peptide. In some examples, the signal peptide, the pro-peptide, or both the signal peptide and the pro-peptide are encoded by a codon-altered sequence. In some examples, the signal peptide, the pro-peptide, or both the signal peptide and the pro-peptide are encoded by a wild-type sequence, while the portion of the nucleic acid encoding the mature Factor IX single-chain polypeptide (e.g., FIX-MP-AA (SEQ ID NO:10); amino acids 47-461 of FIX-FL-AA (SEQ ID NO:2)) is codon altered.

[0020] In some examples, the codon-altered polynucleotides described herein encode a Factor IX variant polypeptide, e.g., having one or more amino acid substitution with respect to the wild-type Factor IX amino acid sequence (e.g., FIX-FL-AA (SEQ ID NO:2) or FIX-MP-AA (SEQ ID NO:10)). In some examples, the Factor IX variant is a hyperactive Factor IX variant with increased Factor IX activity, as compared to wild-type Factor IX. In a particular example, the encoded Factor IX polypeptide has a ‘Padua’ R384L amino acid substitution (relative to the Factor

IX pre-pro-polypeptide sequence FIX-FL-AA (SEQ ID NO:2), R338L amino acid substitution relative to the mature Factor IX single-chain sequence FIX-MP-AA (SEQ ID NO:10)).

[0021] In one example, methods for treating hemophilia B are described. The methods include administering to a patient in need thereof a nucleic acid composition (e.g., a codon-altered Factor IX polynucleotide construct) described herein (e.g., a polynucleotide having high sequence identity to a CS02, CS03, CS04, CS05, or CS06 Factor IX coding sequence). In some examples, the Factor IX polynucleotide construct is a mammalian gene therapy vector, as described herein. In a particular example, the Factor IX polynucleotide construct is an adeno-associated virus (AAV) vector. In some examples, the gene therapy vector includes one or more copies of a liver-specific regulatory control element (e.g., 1 to 3 copies of a CRM8 regulatory control element).

[0021a] In one aspect, the present disclosure refers to a nucleic acid composition disclosed herein for treating hemophilia B.

[0021b] In another aspect, the present disclosure refers to use of a nucleic acid composition disclosed herein for the manufacture of a medicament for the treatment of hemophilia B.

[0022] In one example, methods for producing an adeno-associated virus (AAV) particle are described. The method includes introducing a nucleic acid composition (e.g., a codon-altered Factor IX polynucleotide construct) described herein (e.g., a polynucleotide having high sequence identity to a CS02, CS03, CS04, CS05, or CS06 Factor IX coding sequence) into a mammalian host cell, wherein the polynucleotide is competent for replication in the mammalian host cell.

[0022a] In one aspect, the present disclosure refers to an *in vitro* method for producing an adeno-associated virus (AAV) particle comprising introducing a nucleic acid composition disclosed herein into a mammalian host cell, wherein the nucleic acid composition is competent for replication in the mammalian host cell.

BRIEF DESCRIPTION OF DRAWINGS

[0023] Figure 1 illustrates exemplary Factor IX gene therapy constructs, in accordance with some implementations. The sequence elements for self-complementary (A, B) and single-stranded (C, D) vectors are shown without (A, C) and with (B, D) liver-specific cis-regulatory modules (CRM8).

[0024] Figure 2 shows the wild-type Factor IX coding sequence (SEQ ID NO:1) for accession number CCDS14666.1 (“FIX-FL-NA”).

[0025] Figure 3 shows the amino acid sequences of two wild-type Factor IX pre-pro-polypeptide isoforms expressed in humans. Figure 3A shows the wild-type amino acid sequence for the first, longer Factor IX pre-pro-polypeptide isoform (SEQ ID NO:2) corresponding to UniProt accession number P00740 and NCBI accession number NP_000124.1 (“FIX-FL-AA”). Figure 3B shows the wild-type amino acid sequence for the second, shorter Factor IX pre-pro-

WHAT IS CLAIMED IS:

1. A nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX polynucleotide comprising a nucleic acid sequence that is least 97% identical to the nucleic acid sequence of CS06-MP-NA (SEQ ID NO:17).

2. The nucleic acid composition of claim 1, wherein the Factor IX polynucleotide comprises a nucleic acid sequence that is at least 99% identical to the nucleic acid sequence of CS06-MP-NA (SEQ ID NO:17).

3. The nucleic acid composition of claim 1 or 2, wherein the Factor IX polynucleotide has no more than 10 CpG dinucleotides.

4. The nucleic acid composition of claim 1 or 2, wherein the Factor IX polynucleotide has no more than 3 CpG dinucleotides.

5. The nucleic acid composition of claim 1, wherein the Factor IX polynucleotide comprises the nucleic acid sequence of CS06-MP-NA (SEQ ID NO:17).

6. The nucleic acid composition according to any one of claims 1-5, wherein the Factor IX polynucleotide encodes for leucine at nucleotide positions 1150-1152, relative to FIX-FL-NA (SEQ ID NO:1).

7. The nucleic acid composition according to any one of claims 1-6, wherein the Factor IX protein encoded by the Factor IX polynucleotide has from 1 to 10 amino acid substitutions as compared to FIXp-MP-AA (SEQ ID NO:12).

8. The nucleic acid composition according to any one of claims 1-6, wherein the Factor IX protein encoded by the Factor IX polynucleotide has the amino acid sequence of FIXp-MP-AA (SEQ ID NO:12).

9. The nucleic acid composition according to any one of claims 1-8, wherein the Factor IX polynucleotide further comprises a pre-pro-leader polynucleotide encoding a pre-pro-leader peptide, said pre-pro-leader peptide comprising the amino acid sequence of FIX-PPP-AA (SEQ ID NO:36).

10. The nucleic acid composition of claim 9, wherein the pre-pro-leader polynucleotide has the nucleic acid sequence of CS06-PPP-NA (SEQ ID NO:23).

11. The nucleic acid composition according to any one of claims 1-10, wherein the Factor IX polynucleotide has a nucleic acid sequence that is at least 99% identical to the nucleic acid sequence of CS06-FL-NA (SEQ ID NO:9).

12. The nucleic acid composition according to any one of claims 1-10, wherein the Factor IX polynucleotide has the nucleic acid sequence of CS06-FL-NA (SEQ ID NO:9).

13. A nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX polynucleotide comprising a nucleic acid sequence that is least 95% identical to the nucleic acid sequence of CS02-MP-NA (SEQ ID NO:13).

14. A nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX polynucleotide comprising a nucleic acid sequence that is least 98% identical to the nucleic acid sequence of CS03-MP-NA (SEQ ID NO:14).

15. A nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX polynucleotide comprising a nucleic acid sequence that is least 99% identical to the nucleic acid sequence of CS04-MP-NA (SEQ ID NO:15).

16. A nucleic acid composition, comprising a Factor IX polynucleotide encoding a Factor IX protein, said Factor IX polynucleotide comprising a nucleic acid sequence that is least 98% identical to the nucleic acid sequence of CS05-MP-NA (SEQ ID NO:16).

17. The nucleic acid composition according to any one of claims 13-16, wherein the Factor IX polynucleotide has no more than 10 CpG dinucleotides.

18. The nucleic acid composition according to any one of claims 13-16, wherein the Factor IX polynucleotide has no more than 3 CpG dinucleotides.

19. The nucleic acid composition according to any one of claims 13-16, wherein the Factor IX polynucleotide encodes for leucine at nucleotide positions 1150-1152, relative to FIX-FL-NA (SEQ ID NO:1).

20. The nucleic acid composition according to any one of claims 13-19, wherein the Factor IX protein encoded by the Factor IX polynucleotide has from 1 to 10 amino acid substitutions as compared to FIXp-MP-AA (SEQ ID NO:12).

21. The nucleic acid composition according to any one of claims 13-19, wherein the Factor IX protein encoded by the Factor IX polynucleotide has the amino acid sequence of FIXp-MP-AA (SEQ ID NO:12).

22. The nucleic acid composition according to any one of claims 13-21, wherein the Factor IX polynucleotide further comprises a pre-pro-leader polynucleotide encoding a pre-pro-leader peptide, said pre-pro-leader peptide comprising the amino acid sequence of FIX-PPP-AA (SEQ ID NO:36).

23. The nucleic acid composition according to any one of claims 1-22, further comprising a liver-specific promoter element operably linked to the Factor IX polynucleotide.

24. The nucleic acid composition of claim 23, wherein the liver-specific promoter element comprises one copy of a promoter polynucleotide, said promoter polynucleotide comprising a nucleic acid sequence that is least 95% identical to CRM8 (SEQ ID NO:39).

25. The nucleic acid composition of claim 23, wherein the liver-specific promoter element comprises three copies of a promoter polynucleotide, said promoter polynucleotide comprising a nucleic acid sequence that is least 95% identical to CRM8 (SEQ ID NO:39).

26. The nucleic acid composition of claim 24 or 25, wherein said promoter polynucleotide comprises the nucleic acid sequence of CRM8 (SEQ ID NO:39).

27. The nucleic acid composition according to any one of claims 1-22, further comprising an intron operatively linked to the Factor IX polynucleotide.

28. The nucleic acid composition of claim 27, wherein the intron comprises an MVM intron polynucleotide comprising a nucleic acid sequence that is at least 95% identical to MVMI (SEQ ID NO:53).

29. The nucleic acid composition of claim 28, wherein said MVM intron polynucleotide comprises the nucleic acid sequence of MVMI (SEQ ID NO:53).

30. The nucleic acid composition according to any one of claims 27-29, wherein said intron is positioned between a promoter element and the translation initiation site of the nucleotide sequence encoding a Factor IX polypeptide.

31. The nucleic acid composition of claim 1, comprising the nucleic acid sequence of CS06-CRM8.3-ssV (SEQ ID NO:40).

32. The nucleic acid composition according to any one of claims 1-31, comprising a mammalian gene therapy vector.

33. The nucleic acid composition of claim 32, wherein the mammalian gene therapy vector is an adeno-associated virus (AAV) vector.

34. The nucleic acid composition of claim 33, wherein said adeno-associated virus vector is a serotype 8 adeno-associated virus (AAV-8) vector.

35. The nucleic acid composition according to any one of claims 32-34, wherein the mammalian gene therapy vector comprises a single-stranded polynucleotide encoding the Factor IX protein.

36. A nucleic acid composition according to any one of claims 1-35 for treating hemophilia B.

37. Use of a nucleic acid composition according to any one of claims 1-35 for the manufacture of a medicament for the treatment of hemophilia B.

38. An *in vitro* method for producing an adeno-associated virus (AAV) particle comprising introducing a nucleic acid composition according to any one of claims 1-35 into a mammalian host cell, wherein the nucleic acid composition is competent for replication in the mammalian host cell.