FACTOR IX ENCODING NUCLEOTIDES

Field of the Invention

The present invention relates to polynucleotides comprising a nucleotide sequence encoding Factor IX, viral particles comprising the polynucleotides and treatments utilising the polynucleotides.

Background to the Invention

Haemophilia B, an X-linked life threatening bleeding disorder affects 1:30,000 males. Current treatment involves frequent intravenous injections (2-3 times per week) of Factor IX (FIX) protein. This treatment is highly effective at arresting bleeding but it is not curative and is extremely expensive (£150,000/patient/year), thus making it unaffordable by the majority of haemophilia B patients in the world. Gene therapy for haemophilia B offers the potential for a cure through persistent, endogenous production of Factor IX following the transfer of a functioning copy of the Factor IX gene to an affected patient.

The present application relates to a gene therapy approach for treating haemophilia B, involving administering a vector comprising a polynucleotide encoding Factor IX. Such a gene therapy approach would avoid the need for frequent intravenous injections of Factor IX. However, it is difficult to provide an effective gene therapy vector, *i.e.* one that allows for a high level of Factor IX expression and of the expression of Factor IX which is highly active.

Summary

The present application demonstrates that various modifications to a polynucleotide comprising a Factor IX nucleotide sequence can help to improve the expression level and the activity of the expressed Factor IX polypeptide. For example, the present application demonstrates that the following can improve the efficacy of a polynucleotide comprising a Factor IX nucleotide sequence for treatment of haemophilia B:

- using a codon optimised sequence;
- maintaining a portion of the Factor IX polypeptide that is not codon optimised;
- including an intron or a fragment of an intron;
- providing sequences flanking the intron or fragment of an intron that are not codon optimised;
- using a gain of function mutation;
- using a specific promoter; and/or
- maintaining an AAV genome, comprising the nucleotide, in single stranded form.

These modifications provide a Factor IX sequence which is expressed highly, and which encodes a highly active Factor IX polypeptide or fragment thereof. As demonstrated in the Examples, the polynucleotide of the disclosure expresses, and provides overall Factor IX activity, at higher levels than other Factor IX encoding polynucleotides, for example those disclosed in WO16/075473.

In one aspect, the present disclosure provides a polynucleotide comprising a Factor IX nucleotide sequence corresponding to a reference wild-type Factor IX sequence of SEQ ID NO: 9, wherein: (i) the Factor IX nucleotide sequence comprises a coding sequence that encodes a Factor IX protein or fragment thereof; (ii) a portion of the coding sequence is codon optimized compared to the reference wild-type Factor IX sequence of SEQ ID NO: 9; (iii) the portion of the coding sequence that is codon optimized is at least 1100 nucleotides in length; (iv) in the portion that is codon optimized at least 73% of the codons are selected from the group consisting of: TTC encoding phenylalanine, CTG encoding leucine, ATC encoding isoleucine, GTG or GTC encoding valine, AGC encoding serine, CCC encoding proline, ACC encoding threonine, GCC encoding alanine, TAC encoding tyrosine, CAC encoding histidine, CAG encoding glutamine, AAC encoding asparagine, AAA or AAG encoding lysine, GAC encoding aspartate, TGC encoding cysteine, AGG encoding arginine, GGC encoding glycine, and GAG encoding glutamate; (v) the portion of the coding sequence that is codon optimized is CpG free; (vi) the polynucleotide further comprises a transcription regulatory element comprising: (a) an A1 AT promoter or a fragment of an A1 AT promoter between 150 and 300 nucleotides in length; and/or (b) an HCR enhancer or a fragment of an HCR enhancer; between 100 and 250 nucleotides in

length and (vii) the Factor IX nucleotide sequence comprises a codon that encodes a leucine at a position corresponding to position 384 of the reference wild type Factor IX sequence of SEQ ID NO: 9.

Accordingly, in a first example of the disclosure, there is provided a polynucleotide comprising a Factor IX nucleotide sequence, wherein the Factor IX nucleotide sequence comprises a coding sequence that encodes a Factor IX protein or fragment thereof and wherein a portion of the coding sequence is not wild type.

In a second example of the disclosure, there is provided a polynucleotide comprising a Factor IX nucleotide sequence, wherein the Factor IX nucleotide sequence comprises a coding sequence that encodes a Factor IX protein or a fragment thereof and the coding sequence comprises: (i) a sequence that is at least 95%, at least 98%, at least 99%, at least 99.5%, at least 99.8%, or 100% identical to SEQ ID NO.1; and (ii) a sequence that is at least 98%, at least 99%, at least 99.5%, at least 99.8%, or 100% identical to SEQ ID NO. 15.

In a third example of the disclosure, there is provided a polynucleotide comprising a Factor IX nucleotide sequence, wherein the Factor IX nucleotide sequence encodes a Factor IX protein or fragment thereof and has at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.8%, or 100% identity to SEQ ID NO. 5.

In a fourth example of the disclosure, there is provided a viral particle comprising a recombinant genome comprising the polynucleotide of the disclosure.

In a fifth example of the disclosure, there is provided a composition comprising the polynucleotide or viral particle of the disclosure and a pharmaceutically acceptable excipient.

In a sixth example of the disclosure, there is provided a method of treatment comprising administering an effective amount of the polynucleotide or viral particle of the disclosure to a patient.

In a seventh example of the disclosure, there is provided a use of the polynucleotide, viral particle or composition of the disclosure in the manufacture of a medicament for use in a method of treatment.

In an eighth example, the polynucleotide, viral particle or composition of the disclosure is for use in expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 12 and a transcription regulatory element of SEQ ID NO. 7, thereby treating a disease.

In a ninth example, the polynucleotide, viral particle or composition of the disclosure is for use in treating a disease by expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 12 and a transcription regulatory element of SEQ ID NO. 7.

In a tenth example, there is provided a method of expressing a polypeptide encoded by a Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 12 and a transcription regulatory element of SEQ ID NO. 7 comprising

administering the polynucleotide, viral particle or composition of the disclosure to a patient.

In an eleventh example, there is provided a method of treating a disease by expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 12 and a transcription regulatory element of SEQ ID NO. 7 comprising administering the polynucleotide, viral particle or composition of the disclosure to a patient.

In a twelfth example, there is provided a use of the polynucleotide, viral particle or composition of the disclosure in expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 12 and a transcription regulatory element of SEQ ID NO. 7, thereby treating a disease.

In a thirteenth example, there is provided a use of the polynucleotide, viral particle or composition of the disclosure in treating a disease by expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 12 and a transcription regulatory element of SEQ ID NO. 7.

In a fourteenth example, the polynucleotide, viral particle or composition of the disclosure is for use in expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 18 and a transcription regulatory element of SEQ ID NO. 6, thereby treating a disease.

In a fifteenth example, the polynucleotide, viral particle or composition of the disclosure is for use in treating a disease by expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 18 and a transcription regulatory element of SEQ ID NO. 6.

In a sixteenth example, there is provided a method of expressing a polypeptide encoded by a Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 18 and a transcription regulatory element of SEQ ID NO. 6 comprising administering the polynucleotide, viral particle or composition of the disclosure to a patient.

In a seventeenth example, there is provided a method of treating a disease by expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 18 and a transcription regulatory element of SEQ ID NO. 6 comprising administering the polynucleotide, viral particle or composition of the disclosure to a patient.

In an eighteenth example, there is provided a use of the polynucleotide, viral particle or composition of the disclosure in expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 18 and a transcription regulatory element of SEQ ID NO. 6, thereby treating a disease.

In a nineteenth example, there is provided a use of the polynucleotide, viral particle or composition of the disclosure in treating a disease by expressing a polypeptide encoded by the Factor IX nucleotide sequence at a level at least 2, or at least 3 times greater than a polypeptide encoded by a nucleotide sequence comprising a Factor IX nucleotide sequence of SEQ ID NO. 18 and a transcription regulatory element of SEQ ID NO. 6.

In a twentieth example, the polynucleotide, viral particle or composition of the disclosure is for use in treating a disease by expressing Factor IX protein or a fragment thereof having a Factor IX activity greater than the activity of Factor IX expressed from a viral particle comprising a Factor IX nucleotide sequence of SEQ ID NO: 12 and a transcription regulatory element of SEQ ID NO. 7 and/or a viral particle comprising a Factor IX

nucleotide sequence of SEQ ID NO. 18 and a transcription regulatory element of SEQ ID NO. 6.

In a twenty-first example, there is provided a method of treating a disease by expressing Factor IX protein or a fragment thereof having a Factor IX activity greater than the activity of Factor IX expressed from a viral particle comprising a Factor IX nucleotide sequence of SEQ ID NO: 12 and a transcription regulatory element of SEQ ID NO. 7 and/or a viral particle comprising a Factor IX nucleotide sequence of SEQ ID NO. 18 and a transcription regulatory element of SEQ ID NO. 6, comprising administering an effective amount of the polynucleotide, viral particle or composition of the disclosure to a patient.

In a twenty-second example, there is provided a use of the polynucleotide, viral particle or composition of the disclosure in treating a disease by expressing Factor IX protein or a fragment thereof having a Factor IX activity greater than the activity of Factor IX expressed from a viral particle comprising a Factor IX nucleotide sequence of SEQ ID NO: 12 and a transcription regulatory element of SEQ ID NO. 7 and/or a viral particle comprising a Factor IX nucleotide sequence of SEQ ID NO. 18 and a transcription regulatory element of SEQ ID NO. 6.

In a twenty-third example, the polynucleotide, viral particle or composition of the disclosure is for use in achieving a stable Factor IX activity level of at least 20% in a patient having haemophilia B.

In a twenty-fourth example, there is provided a method of achieving a stable Factor IX activity level of at least 20% in a patient having haemophilia B, comprising administering an effective amount of the polynucleotide, viral particle or composition of the disclosure to a patient.

In a twenty-fifth example there is provided a use of the polynucleotide, viral particle or composition of the disclosure in achieving a stable Factor IX activity level of at least 20% in a patient having haemophilia B.

In a twenty-sixth example, there is provided the polynucleotide, viral particle or composition of the disclosure for use in treating haemophilia B comprising administering the polynucleotide, viral particle or composition to a patient at a dose effective to achieve a stable Factor IX activity level of at least 20% in a patient.

In a twenty-seventh example, there is provided a method of treating haemophilia B comprising administering the polynucleotide, viral particle or composition of the disclosure to a patient at a dose effective to achieve a stable Factor IX activity level of at least 20% in a patient.

In a twenty-eighth example, there is provided a use of the polynucleotide, viral particle or composition of the disclosure in treating haemophilia B comprising administering the polynucleotide, viral particle or composition to a patient at a dose effective to achieve a stable Factor IX activity level of at least 20% in a patient.

Description of the Figures

Figure 1 – Schematic of FIX transgene cassettes ssLP1.FIXco, ssHLP2.TI-codop-FIX-GoF (HTFG) and ssHLP2.TI-ACNP-FIX-GoF (HTAG). ITR = Inverted Terminal Repeat; HLP2 and LP1 are transcription regulatory elements of SEQ ID NOs: 6 and 7 respectively; E = Exon; T1 = Truncated Intron 1A; WT = Wild Type; CO = Codon Optimised; ACNP = a codon optimised sequence of the invention.

Figure 2 – Results from HUH7 transduction with AAV2/Mut C vectors – Experiment 1; (A) shows the level of FIX antigen in supernatant; (B) shows the level of FIX antigen after normalisation using the number of vector genomes present in cell lysate; Error bars represent mean ± SD of n=2. 1e3 = 1 x 10^3; 5e3 = 5 x 10^3; 1e4 = 1 x 10^4; MOI = multiplicity of infection.

Figure 3 – Results from HUH7 transduction with AAV2/Mut C vectors – Experiment 2, showing the level of FIX antigen in supernatant. Error bars represent mean ± SD of n=3. 1e3 = 1 x 10^3; 5e3 = 5 x 10^3; 1e4 = 1 x 10^4, MOI = multiplicity of infection.

CLAIMS

1. A polynucleotide comprising a Factor IX nucleotide sequence corresponding to a reference wild-type Factor IX sequence of SEQ ID NO: 9, wherein:

(i) the Factor IX nucleotide sequence comprises a coding sequence that encodes a Factor IX protein or fragment thereof;

(ii) a portion of the coding sequence is codon optimized compared to the reference wild-type Factor IX sequence of SEQ ID NO: 9;

(iii) the portion of the coding sequence that is codon optimized is at least 1100 nucleotides in length;

(iv) in the portion that is codon optimized at least 73% of the codons are selected from the group consisting of:

TTC encoding phenylalanine,

CTG encoding leucine,

ATC encoding isoleucine,

GTG or GTC encoding valine,

AGC encoding serine,

CCC encoding proline,

ACC encoding threonine,

GCC encoding alanine,

TAC encoding tyrosine,

CAC encoding histidine,

CAG encoding glutamine,

AAC encoding asparagine,

AAA or AAG encoding lysine,

GAC encoding aspartate,

TGC encoding cysteine,

AGG encoding arginine,

GGC encoding glycine, and

GAG encoding glutamate;

(v) the portion of the coding sequence that is codon optimized is CpG free;

(vi) the polynucleotide further comprises a transcription regulatory element comprising:

(a) an A1 AT promoter or a fragment of an A1 AT promoter between 150 and 300 nucleotides in length; and/or

(b) an HCR enhancer or a fragment of an HCR enhancer; between 100 and 250 nucleotides in length and

(vii) the Factor IX nucleotide sequence comprises a codon that encodes a leucine at a position corresponding to position 384 of the reference wild type Factor IX sequence of SEQ ID NO: 9.

2. The polynucleotide of claim 1, wherein the polynucleotide comprises a promoter selected from the group consisting of:

(i) a promoter comprising at least 80% identity to SEQ ID NO. 14;

(ii) a promoter comprising at least 98% identity to SEQ ID NO. 14;

(iii) a promoter comprising the sequence of SEQ ID NO. 14; and

(iv) a liver-specific promoter.

3. The polynucleotide of claim 1, wherein:

the portion of the coding sequence that is codon optimized comprises at least a portion of:

exon 3,

at least a portion of exon 4,

at least a portion of exon 5,

at least a portion of exon 6,

at least a portion of exon 7,

and at least a portion of exon 8.

4. The polynucleotide of claim 1, wherein, in the portion of the coding sequence that is codon optimized, the codon optimization is selected from the group consisting of:

(i) a) at least 1, at least 2, at least 4, or at least 5 codons that encode phenylalanine is/are replaced with TTC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 60%, at least 65%, or at least 70% of the codons that encode phenylalanine are TTC;

c) at least 60%, at least 65%, or at least 70% of the codons that encode phenylalanine are TTC and the remainder are TTT; or

d) the codons that encode phenylalanine are TTC, except where the following codon starts with a G;

(ii) a) at least 5, at least 10, at least 15, or at least 16 codons that encode leucine is/are replaced with CTG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 90%, or at least 94% of the codons that encode leucine are CTG; or

c) at least 90%, or at least 94% of the codons that encode leucine are CTG and the remainder are CTC;

(iii) a) at least 5, at least 10, at least 11, or at least 12 codons that encode isoleucine is/are replaced with ATC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1 of codon ATC is/are replaced with ATT compared to a reference wild type Factor IX sequence, where the following codon starts with a G;

c) at least 60%, at least 70%, or at least 75% of the codons that encode isoleucine are ATC;

d) at least 60%, at least 70%, or at least 75% of the codons that encode isoleucine are ATC and the remainder are ATT; or

e) the codons that encode isoleucine are ATC, except where the following codon starts with a G;

(iv) a) at least 10, at least 15, at least 20, or at least 25 codons that encode valine is/are replaced with GTG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1 codon that encodes valine is/are replaced with GTC compared to a reference wild type Factor IX sequence of SEQ ID NO: 9;

c) at least 80%, at least 90%, or at least 95% of the codons that encode valine are GTG; or

d) at least 80%, at least 90%, or at least 95% of the codons that encode valine are GTG and the remainder are GTC;

(v) a) at least 5, at least 10, at least 12, or at least 13 codons that encode serine is/are replaced with AGC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1, at least 2, or at least 4 codons that encode serine is/are replaced with TCT compared to a reference wild type Factor IX sequence of SEQ ID NO: 9, where the following codon starts with a G;

c) at least 60%, at least 65%, or at least 70% of the codons that encode serine are AGC; or

d) at least 60%, at least 65%, or at least 70% of the codons that encode serine are AGC and the remainder are TCT or TCC;

(vi) a) at least 1, at least 2, or at least 5 codons that encode proline is/are replaced with CCC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1 codon that encode proline is/are replaced with CCT compared to the reference wild type Factor IX sequence of SEQ ID NO: 9, where the following codon starts with a G;

c) at least 50%, at least 55%, or at least 60% of the codons that encode proline are CCC;

d) at least 50%, at least 55%, or at least 60% of the codons that encode proline are CCC and the remainder are CCA or CCT; or

e) the codons that encode proline are CCC, except where the following codon starts with a G;

(vii) a) at least 6, at least 7, at least 8, or at least 10 codons that encode threonine is/are replaced with ACC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1, or at least 2, codons that encode threonine is/are replaced with ACT compared to the reference wild type Factor IX sequence of SEQ ID NO: 9, where the following codon starts with a G;

c) at least 45%, at least 50%, or at least 55% of the codons that encode threonine are ACC;

d) at least 45%, at least 50%, or at least 55% of the codons that encode threonine are ACC and the remainder are ACT; or

the codons that encode threonine are ACC, except where the following codon starts with a G;

(viii) a) at least 1, at least 2, or at least 3codons that encode alanine is/are replaced with GCC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1, at least 2, at least 3, or at least 4 codons that encode alanine is/are replaced with GCT compared to the reference wild type Factor IX sequence of SEQ ID NO: 9, where the following codon starts with a G;

c) at least 35%, at least 40%, or at least 43% of the codons that encode alanine are GCC;

d) at least 35%, at least 40%, or at least 43% of the codons that encode alanine are GCC and the remainder are GCT; or

the codons that encode alanine are GCC, except where the following codon starts with a G;

(ix) a) at least 1, or at least 2 codons that encode tyrosine is/are replaced with TAC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1 of codon TAC is/are replaced with TAT compared to the reference wild type Factor IX sequence of SEQ ID NO: 9, where the following codon starts with a G;

c) at least 40%, at least 45%, or at least 48% of the codons that encode tyrosine are TAC;

d) at least 40%, at least 45%, or at least 48% of the codons that encode tyrosine are TAC and the remainder are TAT; or

e) the codons that encode tyrosine are TAC, except where the following codon starts with a G;

(x) a) at least 1 codon that encode histidine is/are replaced with CAC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 50%, at least 60%, or at least 65% of codons that encode histidine are CAC;

c) at least 50%, at least 60%, or at least 65% of the codons that encode histidine are CAC and the remainder are CAT; or

d) the codons that encode histidine are CAC, except where the following codon starts with a G;

(xi) a) at least 1, at least 2, at least 4, or at least 5 codons that encode glutamine is/are replaced with CAG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1 of codon CAG is/are replaced with CAA compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

c) at least 80%, at least 85%, or at least 90% of the codons that encode glutamine are CAG; or

d) at least 80%, at least 85%, or at least 90% of the codons that encode glutamine are CAG and the remainder are CAA;

(xii) a) at least 1, at least 2, at least 4, or at least 5 codons that encode asparagine is/are replaced with AAC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 60%, at least 65%, or at least 70% of the codons that encode asparagine are AAC;

c) at least 60%, at least 65%, or at least 70% of the codons that encode asparagine are AAC and the remainder are AAT; or

d) the codons that encode asparagine are AAC, except where the following codon starts with a G;

(xiii) a) at least 5, at least 7, at least 8, or at least 9 codons that encode lysine is/are replaced with AAG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1 of codon AAG is/are replaced with AAA compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

c) at least 80%, at least 90%, or at least 95% of the codons that encode lysine are AAG; or

d) at least 80%, at least 90%, or at least 95% of the codons that encode lysine are AAG and the remainder are AAA;

(xiv) a) at least 1, at least 2, at least 3, or at least 4 codons that encode aspartate is/are replaced with GAC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1 of codon GAC is/are replaced with GAT compared to the reference wild type Factor IX sequence of SEQ ID NO: 9, where the following codon starts with a G;

c) at least 45%, at least 50%, or at least 60% of the codons that encode aspartate are GAC;

d) at least 45%, at least 50%, or at least 60% of the codons that encode aspartate are GAC and the remainder are GAT; or

e) the codons that encode aspartate are GAC, except where the following codon starts with a G;

(xv) a) at least 15, at least 20, at least 25, or at least 26 codons that encode glutamate is/are replaced with GAG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 80%, at least 90%, or at least 95% of the codons that encode glutamate are GAG; or

c) at least 80%, at least 90%, or at least 95% of the codons that encode glutamate are GAG and the remainder are GAA;

(xvi) a) at least 5, at least 6, at least 7, or at least 8 codons that encode cysteine is/are replaced with TGC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1 of codon TGC is/are replaced with TGT compared to the reference wild type Factor IX sequence of SEQ ID NO: 9, where the following codon starts with a G;

c) at least 40%, at least 50%, or at least 55% of the codons that encode cysteine are TGC;

d) at least 40%, at least 50%, or at least 55% of the codons that encode cysteine are TGC and the remainder are TGT; or

e) the codons that encode cysteine are TGC, except where the following codon starts with a G;

(xvii) wherein, the codons that encode tryptophan are TGG;

(xviii) a) at least 5, at least 8, at least 10, or at least 11 codons that encode arginine is/are replaced with AGG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 1 codon that encodes arginine is/are replaced with AGA compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

c) at least 60%, at least 70%, or at least 75% of the codons that encode arginine are AGG; or

d) at least 60%, at least 70%, or at least 75% of the codons that encode arginine are AGG and the remainder are AGA; and

(xix) a) at least 5, at least 10, at least 12, or at least 13 codons that encode glycine is/are replaced with GGC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 5, at least 6, at least 7, or at least 8 codons that encode glycine is/are replaced with GGG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9, where the following codon starts with a G;

c) at least 50%, at least 55%, or at least 60% of the codons that encode glycine are GGC;

d) at least 50%, at least 55%, or at least 60% of the codons that encode glycine are GGC and the remainder are GGG; or

e) the codons that encode glycine are GGC, except where the following codon starts with a G.

5. The polynucleotide of claim 1 wherein the portion of the coding sequence that is codon optimized comprises codons encoding phenylalanine, leucine, isoleucine, valine, serine, proline, threonine, alanine, tyrosine, histidine, glutamine, asparagine, lysine, aspartate, glutamate, cysteine, tryptophan, arginine, and glycine, and in the portion of the coding sequence that is codon optimized:

a) at least 5 codons that encode phenylalanine is/are replaced with TTC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

b) at least 16 codons that encode leucine is/are replaced with CTG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

c) at least 12 codons that encode isoleucine is/are replaced with ATC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

d) at least 25 codons that encode valine is/are replaced with GTG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

e) at least 13 codons that encode serine is/are replaced with AGC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

f) at least 5 codons that encode proline is/are replaced with CCC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

g) at least 10 codons that encode threonine is/are replaced with ACC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

h) at least 4 codons that encode alanine is/are replaced with GCC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

i) at least 2 codons that encode tyrosine is/are replaced with TAC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

j) at least 1 codon that encode histidine is/are replaced with CAC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

k) at least 5 codons that encode glutamine is/are replaced with CAG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

l) at least 5 codons that encode asparagine is/are replaced with AAC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

m) at least 9 codons that encode lysine is/are replaced with AAG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

n) at least 4 codons that encode aspartate is/are replaced with GAC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

o) at least 26 codons that encode glutamate is/are replaced with GAG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

p) at least 8 codons that encode cysteine is/are replaced with TGC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9;

q) the codons that encode tryptophan are TGG;

r) at least 11 codons that encode arginine is/are replaced with AGG compared to the reference wild type Factor IX sequence of SEQ ID NO: 9; and

s) at least 13 codons that encode glycine is/are replaced with GGC compared to the reference wild type Factor IX sequence of SEQ ID NO: 9.

6. The polynucleotide of claim 1, wherein the portion of the coding sequence that is codon optimized comprises codons encoding phenylalanine, leucine, isoleucine, valine, serine, proline, threonine, alanine, tyrosine, histidine, glutamine, asparagine, lysine, aspartate, glutamate, cysteine, tryptophan, arginine, and glycine, and in the portion of the coding sequence that is codon optimised:

a) at least 70% of codons that encode phenylalanine are TTC;

b) at least 94% of codons that encode leucine are CTG;

c) at least 75% of codons that encode isoleucine are ATC;

d) at least 95% of codons that encode valine are GTG;

e) at least 70% of codons that encode serine are AGC;

f) at least 60% of codons that encode proline are CCC;

g) at least 55% of codons that encode threonine are ACC;

h) at least 43% of codons that encode alanine are GCC;

i) at least 48% of codons that encode tyrosine are TAC;

j) at least 65% of codons that encode histidine are CAC;

k) at least 90% of codons that encode glutamine are CAG;

l) at least 70% of codons that encode asparagine are AAC;

m) at least 95% of codons that encode lysine are AAG;

n) at least 60% of codons that encode aspartate are GAC;

o) at least 95% of codons that encode glutamate are GAG;

p) at least 55% of codons that encode cysteine are TGC;

q) codons that encode tryptophan are TGG;

r) at least 75% of codons that encode arginine are AGG; and

s) at least 60% of codons that encode glycine are GGC.

7. The polynucleotide of claim 1, wherein the portion of the coding sequence that is codon optimized comprises codons encoding phenylalanine, leucine, isoleucine, valine, serine, proline, threonine, alanine, tyrosine, histidine, glutamine, asparagine, lysine, aspartate, glutamate, cysteine, tryptophan, arginine, and glycine, and in the portion of the coding sequence that is codon optimized:

a) at least 70% of the codons that encode phenylalanine are TTC and the remainder are TTT;

b) at least 94% of codons that encode leucine are CTG and the remainder are CTC;

c) at least 75% of codons that encode isoleucine are ATC and the remainder are ATT;

d) at least 95% of codons that encode valine are GTG;

e) at least 70% of codons that encode serine are AGC;

f) at least 60% of codons that encode proline are CCC and the remainder are CCA or CCT;

g) at least 55% of codons that encode threonine are ACC and the remainder are ACT;

h) at least 43% of codons that encode alanine are GCC and the remainder are GCT;

i) at least 48% of codons that encode tyrosine are TAC and the remainder are TAT;

j) at least 65% of codons that encode histidine are CAC and the remainder are CAT;

k) at least 90% of codons that encode glutamine are CAG and the remainder are CAA;

l) at least 70% of codons that encode asparagine are AAC and the remainder are AAT;

m) at least 95% of codons that encode lysine are AAG and the remainder are AAA;

n) at least 60% of codons that encode aspartate are GAC and the remainder are GAT;

o) at least 95% of codons that encode glutamate are GAG and the remainder are GAA;

p) at least 55% of codons that encode cysteine are TGC and the remainder are TGT;

q) codons that encode tryptophan are TGG;

r) at least 75% of codons that encode arginine are AGG and the remainder are AGA; and

s) at least 60% of codons that encode glycine are GGC and the remainder are GGG.

8. The polynucleotide of claim 1, wherein the polynucleotide comprises an enhancer selected from the group consisting of:
(i) an enhancer comprising at least 80% identity to SEQ ID NO. 13;
(ii) an enhancer comprising at least 98% identity to SEQ ID NO. 13; and
(iii) the polynucleotide comprises an enhancer comprising the sequence of SEQ ID NO. 13.

9. The polynucleotide of claim 2, wherein a polypeptide encoded by the Factor IX nucleotide sequence is expressed in human liver cells at higher levels compared to the reference wild type Factor IX sequence of SEQ ID NO: 9.

10. The polynucleotide of claim 1, wherein the codon that encodes an amino acid at the position corresponding to position 384 of the reference wild type Factor IX sequence of SEQ ID NO: 9 is CTC.

11. The polynucleotide of claim 1, wherein the codon that encodes an amino acid at the position corresponding to position 384 of the reference wild type Factor IX sequence of SEQ ID NO: 9 is CTG.

12. An AAV viral particle comprising a recombinant genome comprising the polynucleotide of claim 1.

13. The AAV viral particle of claim 12, wherein the recombinant genome comprises two resolvable ITRs.

14. The AAV viral particle of claim 12, wherein after transduction into a population of Huh7 cells, the AAV viral particle expresses the Factor IX protein or a fragment thereof, wherein the Factor IX protein or the fragment thereof has an activity greater than an activity of a Factor IX protein or a fragment thereof expressed from a comparable AAV viral particle comprising a Factor IX nucleotide sequence of SEQ ID NO: 12 and a transcription regulatory element of SEQ ID NO. 7 transduced into a comparable population of Huh7 cells.