Claims

1. A polynucleotide comprising an open reading frame (ORF) encoding a polypeptide, wherein

(a) (i) at least 1.03% of the codon pairs in the ORF are codon pairs shown in Table 1; or (ii) at least 1% of the codon pairs in the ORF are codon pairs shown in Table 1 and the ORF does not encode an RNA-guided DNA binding agent; or

(b) the ORF comprises a sequence with at least 95% identity to any one of SEQ ID NOs: 6-10, 29, 46, 69-73, 90-93, 96-99, 102-105, 108-111, 114-117, 120-123, 126-129, or 132-143 optionally wherein identity is determined without regard to the start and stop codons of the ORF; or

(c) at least 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% of the codons in the ORF are (i) codons listed in Table 5, or (ii) codons listed in Table 6, and wherein the polypeptide is not an RNA-guided DNA binding agent.

2. The polynucleotide of claim 1, wherein the repeat content of the ORF is less than or equal to 23.3%.

3. The polynucleotide of any one of claims 1-2, wherein the GC content of the ORF is greater than or equal to 55%.

4. A polynucleotide comprising an open reading frame (ORF) encoding a polypeptide, wherein the repeat content of the ORF is less than or equal to 23.3% and the GC content of the ORF is greater than or equal to 55%.

5. The polynucleotide of any one of claims 1-4, wherein at least 1.03% of the codon pairs in the ORF are codon pairs shown in Table 1.

6. The polynucleotide of any one of claims 1-5, wherein

(i) less than or equal to 0.9% of the codon pairs in the ORF are codon pairs shown in Table 2;

(ii) at least 60%, 65%, 70%, or 75% of the codon in the ORF are codon shown in Table 3;

(iii) less than or equal to 20% of the codons in the ORF are codons shown in Table 4;

(iv) at least 1.05% of the codon pairs in the ORF are codon pairs shown in Table 1; or

(v) less than or equal to 10% of the codon pairs in the ORF are codon pairs shown in Table 1.

7. The polynucleotide of any one of claims 1-6, wherein

(i) the GC content of the ORF is greater than or equal to 56%

(ii) the GC content of the ORF is less than or equal to 63%;

(iii) the repeat content of the ORF is less than or equal to 23.2%; or

(iv) the repeat content of the ORF is greater than or equal to 20%.

8. The polynucleotide of any one of claims 1-7, wherein

(i) less than or equal to 15% of the codons in the ORF are codons shown in Table 4;

(ii) at least 76% of the codons in the ORF are codons shown in Table 3; or

(iii) less than or equal to 87% of the codons in the ORF are codons shown in Table 3.

9. The polynucleotide of any one of claims 1-8, wherein

(i) the ORF has a uridine content ranging from its minimum uridine content to 101%, 102%, 103%, 105%, 110%, 115%, 120%, 125%, 130%, 135%, 140%, 145%, or 150% of the minimum uridine content; or

(ii) the ORF has an A+U content ranging from its minimum A+U content to 101%, 102%, 103%, 105%, 110%, 115%, 120%, 125%, 130%, 135%, 140%, 145%, or 150% of the minimum A+U content; or

(iii) the ORF has a GC content in the range of 55%-65%, such as 55%-57%, 57%-59%, 59-61%, 61-63%, or 63-65%; or

(iv) the ORF has a repeat content ranging from its minimum repeat content to 101%, 102%, 103%, 105%, 110%, 115%, 120%, 125%, 130%, 135%, 140%, 145%, or 150% of the minimum repeat content; or

(v) the ORF has a repeat content of 22%-27%.

10. The polynucleotide of any one of claims 1-9, wherein the polypeptide has a length of at least 30 amino acids, optionally wherein the polypeptide has a length of at least 50 amino acids.

11. The polynucleotide of any one of claims 1-10, wherein the polynucleotide comprises a sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identity to any one of SEQ ID NOs: 6-10, 29, 46, 69-73, 90-93, 96-99, 102-105, 108-111, 114-117, 120-123, 126-129, or 134-143.

12. The polynucleotide of any one of claims 1-11, wherein the polynucleotide comprises a sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identity to any one of SEQ ID NOs: 16-20, 78-80, 194-197, or 200-201.

13. The polynucleotide of any one of claims 1-12, wherein the ORF encodes an RNA-guided DNA binding agent.

14. The polynucleotide of claim 13, wherein the RNA-guided DNA-binding agent has double-stranded endonuclease activity or nickase activity or comprises a dCas DNA binding domain.

15. The polynucleotide of any one of claims 1-14, wherein the ORF encodes an endonuclease.

16. The polynucleotide of any one of claims 1-15, wherein the polynucleotide further comprises a 5' UTR with at least 90% identity to any one of SEQ ID NOs: 177-181 or 190-192; and/or a 3' UTR with at least 90% identity to any one of SEQ ID NOs: 182-186 or 202-204.

17. The polynucleotide of any one of claims 1-16, wherein the encoded polypeptide comprises a nuclear localization signal (NLS).

18. The polynucleotide of claim 17, wherein the NLS comprises a sequence having at least 80%, 85%, 90%, or 95% identity to any one of SEQ ID NOs: 163-176.

19. The polynucleotide of any of claims 1-18, wherein at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% of the uridine is substituted with a modified uridine, optionally, wherein the modified uridine is one or more of N1-methyl-pseudouridine, pseudouridine, 5-methoxyuridine, or 5-iodouridine.

20. The polynucleotide of any one of claims 1-19, wherein the polynucleotide is an mRNA.

21. The polynucleotide of any one of claims 1-20, wherein the polynucleotide is an expression construct comprising a promoter operably linked to the ORF.

22. A plasmid comprising the polynucleotide of claim 21.

23. A host cell comprising the polynucleotide of claim 21 or the plasmid of claim 22.

24. A method of preparing an mRNA comprising contacting the polynucleotide of claim 21 or the plasmid of claim 22 with an RNA polymerase under conditions permissive for transcription of the mRNA, optionally wherein the contacting step is performed *in vitro*.

25. A method of expressing a polypeptide, comprising contacting a cell with the polynucleotide of any one of claims 1-21.

26. A composition comprising a polynucleotide according to any one of claims 1-21 and at least one guide RNA, wherein the polynucleotide encodes an RNA-guided DNA binding agent.

27. A lipid nanoparticle comprising a polynucleotide according to any one of claims 1-21.

28. A pharmaceutical composition comprising a polynucleotide according to any one of claims 1-21 and a pharmaceutically acceptable carrier.

29. A method of genome editing or modifying a target gene comprising contacting a cell with the polynucleotide according to any one of claims 1-21, the composition according to claim 26, the lipid nanoparticle according to claim 27, or the pharmaceutical composition according to claim 28, wherein the polynucleotide encodes an RNA-guided DNA binding agent.

30. Use of the polynucleotide according to any one of claims 1-21 for genome editing or modifying a target gene, wherein the polynucleotide encodes an RNA-guided DNA binding agent.

31. The polynucleotide according to any one of claims 1-21 for use in therapy.

32. The method of claim 29, wherein the genome editing or modification of the target gene occurs in a liver cell, optionally wherein the liver cell is a hepatocyte.

33. The use of claim 30, wherein the genome editing or modification of the target gene occurs in a liver cell, optionally wherein the liver cell is a hepatocyte.

34. A method of generating an open reading frame (ORF) sequence encoding a polypeptide, the method comprising:

a) providing a polypeptide sequence of interest;

b) assigning a codon for each amino acid position of the polypeptide sequence, wherein if the amino acid position is a member of a dipeptide shown in Table 1, then the codon pair for that dipeptide is used, but if the amino acid position is a member of more than one dipeptide shown in Table 1 and the codon pairs for those dipeptides provide different codons for the position or the amino acid position is not a member of a dipeptide shown in Table 1, then one or more of the following is performed:

   i. selecting a codon from a wild-type sequence encoding the polypeptide if a naturally occurring polypeptide is encoded;

   ii. if the amino acid is a member of more than one dipeptide shown in Table 1 and the codon pairs for those dipeptides provide different codons for the position, eliminating codons that appear in Table 4 and/or that would result in the presence of a codon pair shown in Table 2, and/or selecting a codon that appears in Table 3;

   iii. using a codon set of Table 5, 6, or 7 to supply the codon for the amino acid position, optionally wherein if steps (i) and/or (ii) are performed then step (iii) is performed if a unique codon for the amino acid position has not been provided; and/or

   iv. selecting a codon that (1) minimizes uridine content, (2) minimizes repeat content, and/or (3) maximizes GC content.